Final Year Project Report
Full Unit - Project Plan

———————

Resourceful Robots

Cougar Tasker

———————

A report submitted in part fulfilment of the degree of

MSci (Hons) in Computer Science

Supervisor: Dr. Anand Subramoney

Department of Computer Science
Royal Holloway, University of London

October 4, 2023

Declaration

This report has been prepared on the basis of my own work. Where other published and unpublished source materials have been used, these have been acknowledged.

Word Count:

Student Name: Cougar Tasker

Date of Submission:

Signature:

Table of Contents

Abstract

Autonomous robots such as Boston-dynamic's Spot are increasing in prevalence across a wide range of fields[2]. Furthermore, these robots are increasingly integral in modern society, taking on ever more advanced tasks[4]. As autonomous robots take on increasingly complex and resource-intensive tasks, optimising their resource consumption while meeting objectives becomes an increasingly significant challenge. These autonomous robots operate in diverse environments with varying objectives, so a singular algorithm will not perform optimally for all cases.

This project aims to explore how reinforcement learning can provide a solution for effectively prioritising these objectives and managing resources. Reinforcement Learning, a form of machine learning, involves an agent learning to make decisions through interactions with its environment. Its ability to handle delayed rewards sets it apart from other machine-learning approaches, making it particularly well-suited for addressing this problem. In this project, we will simulate the environment, initially using a grid world and later incorporating more advanced environments from OpenAI's gym library[3].

//todo add more detail about reinforment learning add more refrences -> refrence the sutton book! and the other one

Motivations

My inspiration for studying this degree in artificial intelligence comes in large part from my belief that AI is becoming increasingly pivotal in shaping the future of technology and industry. For this purpose, this project presents an invaluable opportunity for my personal and professional growth. It is a fantastic platform to improve my comprehension of reinforcement learning while offering hands-on experience.

What is unique about this resource-gathering robot project is its structured progression of complexity, Starting from fundamental concepts and culminating in advanced techniques. This gradient makes the complex nature of reinforcement learning more approachable than it may be in industry.

This project interests me because of its generality and applicability to many different scenarios. Resource-gathering has the potential to incorporate many real-world constraints like energy, visibility and obstacles. I would like to see how this impacts different exploration strategies.

Last year, I completed my year-long internship at Zing, a digital communications company that is progressively incorporating AI systems for its customers. This experience has demonstrated to me the value of understanding the internals of these AI systems. «insert reference to the increased value of ai in business». Through this project, I aim to improve my understanding of autonomous agents' benefits, biases, and limitations. This knowledge will be desirable for many companies like Zing working with artificial agents.

Objectives

The primary objective is to develop and understand reinforcement learning agents. The project will include two parts: a report and a graphical program. In the report, I will describe the reinforcement learning concepts that are implemented by the program, such as Markov decision processes and finding optimal policies by the Bellman equations. The report will also cover the software engineering process of the program and the results of different parameter choices and exploration strategies. The program will be developed in Python using the libraries Kivy for graphics and PyTorch for Deep Neural networks. The program should achieve the following goals:

- Energy-Efficient Navigation: Create a robot that can autonomously navigate a grid world while making optimal decisions to conserve energy. The robot should learn to prioritise energy-efficient paths.

- Resource Gathering Strategies: Design intelligent resource-gathering strategies for the robot, ensuring it collects essential resources while balancing energy expenditure. The robot should adapt its behaviour based on the availability of resources and its current energy levels.

- Dynamic Energy Management: The robot should monitor its energy reserves and adjust its actions accordingly. This includes learning when to engage in resource gathering and when to return to a charging station.

- Effective Exploration: Develop exploration strategies that enable the robot to learn and adapt to different grid world scenarios.

- Deep Reinforcement Learning: As an extension, the program should integrate deep neural networks (DNN) with Q-learning. The advantage of using a DNN instead of a table is the DNN's ability to generalise knowledge[1]. This unlocks working in environments with far more extensive observations and action spaces, allowing for added complexity to our existing environment or an extension of integrating a different environment, such as one from OpenAI's gym library[3].

Chapter 1: **Timeline**

1.1   Term One

Week	Goals	Explanation/Motivation

1.2   Term Two

Week	Goals	Explanation/Motivation

Chapter 2: **Risks And Mitigations**

2.1   Hardware Issues

Likelihood:	Low
Importance:	High

Hardware failures, such as a lost or malfunctioning laptop, could disrupt the project's progress. Project data should be regularly backed up externally to reduce this risk. The project will be stored under Git with a GitHub remote repository. Under this arrangement, frequent code pushes upstream can minimise potential data loss. The report will be written on the Google Docs platform, providing cloud backup and version control.

2.2   Time Management Issues

Likelihood:	Moderate
Importance:	Moderate

Underestimating the time required for tasks could result in missing milestones or goals. Perfect time estimates are impossible; however, tasks can subdivided appropriately to avoid significant surprises. Spreading out tasks evenly and leaving some buffer time can assist in avoiding work being cut short.

2.3   Machine Learning Risks

Likelihood:	Moderate
Importance:	Low

Machine learning can be a slow and computationally expensive task; this risks slowing the development or failing to train an effective model. The Q-learning table's complexity and size can be gradually increased and dynamically controlled to avoid computational bottlenecks. Manual fine-tuning of the system can be timely. Instead, the hyperparameters will be optimised with grid search.

2.4   Software Development Challenges

Likelihood:	High
Importance:	Moderate

Developing a complete graphical application can be complex, and unexpected software bugs may arise. Modern software engineering principles can minimise bugs and improve software

quality. That is why this project will have version control, test-driven development, documentation, and static code analysis.

2.5   GUI Development Challenges

Likelihood:	Moderate
Importance:	Low

Designing and implementing the Graphical User Interface may be more time-consuming or challenging than anticipated. This project will use GUI development libraries or frameworks that streamline the process. Start GUI development early in the project to allow for iterative improvements.

2.6   Understanding of Reinforcement Concepts

Likelihood:	Moderate
Importance:	High

Reinforcement learning has several abstract concepts, such as MDPs, Bellman equations, dynamic programming, or Q-learning, which may lead to incorrect implementations or interpretations. Referencing textbooks and research papers to check results can validate findings. Thoroughly studying reinforcement learning concepts can avoid misinterpretations.

Bibliography

[1] Vincent François-Lavet, Peter Henderson, Riashat Islam, Marc G Bellemare, Joelle Pineau, et al. An introduction to deep reinforcement learning. *Foundations and Trends® in Machine Learning*, 11(3-4), 2018.

[2] Martin Hagele. Robots conquer the world [turning point]. *IEEE Robotics & Automation Magazine*, 23(1), 2016.

[3] OpenAI. Gym documentation.

[4] Cédrick Zaouter, Alexandre Joosten, Joseph Rinehart, Michel MRF Struys, and Thomas M Hemmerling. Autonomous systems in anesthesia: where do we stand in 2020? a narrative review. *Anesthesia & Analgesia*, 130(5), 2020.