
RL Project

David S. Hippocampus
Department of Computer Science
University of Bath
Bath, BA2 7AY
hippo@bath.ac.uk

1 Problem Definition

A clear, precise and concise description of your chosen problem, including the states, actions, transition dynamics, and the reward function. You will lose marks for an unclear, incorrect, or incomplete problem definition.

2 Background

A discussion of reinforcement learning methods that may be effective at solving your chosen problem, their strengths and weaknesses for your chosen problem, and any existing results in the scientific literature (or publicly available online) on your chosen problem or similar problems.

2.1 DQN (Mirco)

2.2 A2C (Chris)

- Intro to Policy Gradient Methods

A set of alternative methods to value-based reinforcement methods that can be applied to this problem is policy gradient methods. Whilst value-based methods such as DQN learn values of actions and use these estimates to select actions, policy gradient methods do not need to consult a value function and, instead, learn a parameterised policy to select actions (Sutton and Barto, 2018). In selecting actions using a parameterised policy, policy gradient methods use *action preferences* rather than action values.

- Intro to Actor Critic
- Intro to A2C including A2C vs A3C results

2.3 Async Q-Learning (Peter)

3 Method

A description of the method(s) used to solve your chosen problem, an explanation of how these methods work (in your own words), and an explanation of why you chose these specific methods.

3.1 DQN (Mirco)

3.2 A2C (Chris)

3.3 Async Q-Learning (Peter)

4 Results

result comparison between the 3 approaches. how quickly each agent learns how well do they perform in absolute terms how do they compare with respect to a human player

4.1 DQN (Mirco)

4.2 A2C (Chris)

4.3 Async Q-Learning (Peter)

5 Discussion

An evaluation of how well you solved your chosen problem.

5.1 DQN (Mirco)

5.2 A2C (Chris)

5.3 Async Q-Learning (Peter)

6 Future Work

What other techniques can be used to improve further the performances of the 3 agents.

7 Personal Experience

A discussion of your personal experience with the project, such as difficulties or pleasant surprises you encountered while completing it.

References

Appendices

If you have additional content that you would like to include in the appendices, please do so here. There is no limit to the length of your appendices, but we are not obliged to read them in their entirety while marking. The main body of your report should contain all essential information, and content in the appendices should be clearly referenced where it's needed elsewhere.

Appendix A: Example Appendix 1

Appendix B: Example Appendix 2