Cardiff University
School of Computer Science and Informatics
CM3203 One Semester Individual Project
40 Credits

Visualizing Deep Reinforcement Learning

Initial Plan

Author
Tian Sou Zen Tan

Supervisor
Dr Jing Wu

Contents

Project Description....2
Project Aims and Objectives....3
Primary....3
Secondary....3
Ethics....3
Work Plan....4
Week 1....4
Week 2....4
Week 3....4
Week 4 & 5....4
Week 6....4
Week 7 & 8....4
Week 9....4
Easter & Week 10....4
Week 11 & 12....5
References....6

Project Description

Deep reinforcement learning (DRL) is the combination of both reinforcement learning and deep learning to allow for high dimensional problems using function approximation. There has been much work applied to improving RL algorithms but understanding the learning process is still a problem. Finding an explanation can help recognize the capabilities of the agents and identify problems in the decision-making process for better optimisation. Some attempts have made by experts to develop visual analytics to help better understand these behaviours such as DQNViz [1] and DRLViz [2]. These experts are interested to understand the agent's memory, statistics from the learning process, and the relations between the agent's observation and action.

For this project I will be focusing on investigating several visualization techniques to help experts interpret the learning process. I will be using Gym [3], a toolkit that provides a range of simulated environments in which RL algorithms can be applied for training agents. The environment that I will be using is called Breakout by Atari where the objective is to obtain the maximum score by breaking bricks using a ball and a paddle while avoiding life loss. Deep Q-Network (DQN) is a RL method that I will be using to collect data during the learning process. It is a type of deep reinforcement learning model that uses a neural network, and it is popular for achieving superhuman level performance in Atari games.

Data collected during the learning process may vary according to different RL algorithms therefore I will need to decide on a visualization technique that is most suitable for the type of data collected. Some of the potential visualizations could be parallel coordinates or matrix visualisation etc. D3.js could be a possible library for producing visualizations on a web browser hosted on GitHub. There are also visualization libraries in python that are in consideration for a smoother learning curve, however D3.js provides easier visualization implementations and better features.

In conclusion, I will need to create a visualization system that can help experts gain better understanding in how the agent behaves and learns by looking at certain aspects. For example, looking at the statistics about the actions taken by the agent through the entire training process, and how the actions change over time through the Epoch view. There may also be a need for certain interactions with the system to analyse the data better.

Project Aims and Objectives

The project aims and objectives are separated into two sections, the primary being the main goal and the secondary being desirables. Once the primary objectives are achieved in accordance with the work plan, the secondary objectives should be attempted.

The primary aim of this project is to produce visualization techniques on the agent's behaviour so that experts can gain a better understanding on deep reinforcement learning. The secondary aim is to evaluate the differences on the agent's behaviour when applying different RL methods.

Primary

- Research on reinforcement learning algorithms and its environments.
- Train and collect data for visualization purposes.
- Design a suitable visualization technique depending on the specific type of data collected.
- Implement a visualization tool for the specific RL algorithm and task.

Secondary

- Collect data using a different RL method on the same task.
- Apply the same visualization tool to compare the agent's behaviour between the two RL algorithms.

Ethics

In one of the meetings with my supervisor, we have concluded that there will be a need for ethical approval. This project requires the feedback from RL experts for the visualization tools to be useful and effective. The feedback obtained will not include any personal information.

Work Plan

The work plan contains a list of tasks and milestones over the span of 12 weeks. The plan may vary and be adjusted depending on certain features and the outcome of the meeting with my supervisor.

There is a scheduled meeting with my supervisor every week on Thursday from 11:30am to 12:00pm. The meeting frequency should be flexible and subject to change.

Week 1

- Background research on reinforcement learning.
- Meeting with supervisor.
- **Milestone:** Decide on a RL algorithm and the task where the agent interacts with an environment.
- **Deliverables:** Submit initial plan.

Week 2

- Apply RL algorithm and start collecting data.
- Meeting with supervisor.
- **Milestone:** Initial data collected to be used for visualization.

Week 3

- Continue to run algorithm and collect data.
- Research and design different visualization techniques suitable for the data collected.
- **Milestone:** Decide on the most appropriate visualization technique.
- Meeting with supervisor.

Week 4 & 5

- Implementation of specific visualization system.
- **Milestone:** Apply visualization system on collected data to produce charts.
- Display charts and data on a website hosted on GitHub.
- **Deliverables:** First prototype created.
- Meeting with supervisor.

Week 6

- Evaluate on prototype with RL experts and obtain feedback.
- Meeting with supervisor.

Week 7 & 8

- Improve system according to feedback and add the required interactions needed.
- **Deliverables:** Second prototype created.
- Meeting with supervisor.

Week 9

- Evaluate prototype with RL experts again.
- Start drafting final report.
- Meeting with supervisor.

Easter & Week 10

- Continue to work on tasks that may be incomplete.
- Work on secondary aim and objectives (desirables).

Week 11 & 12

- Focus on completing final report.
- Meeting with supervisor.
- **Deliverables:** Submit final report, code, and charts.

References

[1] J. Wang, L. Gou, H.-W. Shen and H. Yang, “DQNViz: A Visual Analytics Approach to Understand,” IEEE TVCG. 25(1): 288-298, 2019.

[2] T. Jaunet, R. Vuillemot and C. Wolf, “DRLViz: Understanding Decisions and Memory in Deep,” Computer Graphics Forum, 2020.

[3] OpenAI, “Gym,” [Online]. Available: https://gym.openai.com/. [Accessed 2022].