Section B – Interim report

Manwel Borg – IT-MSD-6.3B

Paper 1

Paper name: Simultaneous use of imitation learning and reinforcement learning in artificial intelligence development for video games

Authors: Karavaev Vadim, Kiseleva Tatyana and Orlinskaya Oksana

Methodology used: Quantitative and qualitative; they tested two approaches and compared them in terms of training time and agent performance. Creswell (2014) states that "in a quantitative project, the problem is best addressed by understanding what factors or variables influence an outcome." In their study, the authors have analysed the training graphs to determine which approach took the least time to train. Creswell (2014) also states that "in a qualitative project, the author will describe a research problem that can best be understood by exploring a concept or phenomenon", and that "a mixed methods study can employ either the qualitative or the quantitative approach (or some combination) to writing an introduction." The authors observed the agents they had created to determine how they behaved.

Please note that the statements by Creswell were not rewritten for each paper in this report because they apply to all the papers.

For both approaches, the agent was trained by having it play a series of matches against five tanks on a flat terrain. In their first approach, they trained their agent using solely reinforcement learning; in their other approach, the agent was trained via imitation learning and reinforcement learning simultaneously. In order to do this, they first recorded a set of human demonstrations for the agent to imitate. Then, they provided these demonstrations as inputs to the neural network during training.

Objectives: In their study, Vadim, Tatyana and Oksana (2018) explored the simultaneous use of imitation learning and reinforcement learning to create an intelligent agent that can play realistically and effectively without making use of unfair advantage.

Implications: In their paper, the authors state that in most cases, video game NPCs are controlled by traditional deterministic algorithms, and that as a result, their behaviour may look unnatural and unpleasant. They imply that their study was an attempt at addressing this.

Findings: When the agent was trained via reinforcement learning exclusively, it was able to fight a maximum of two opponents after 10 hours of training. After combining the two, the agent was able to fight 3 – 4 opponents after only 40 minutes of training.

Conclusion: Vadim, Tatyana and Oksana (2018) concluded by stating that the agent they have created and trained was able to use the available game mechanics effectively while behaving similarly to a human.

Paper 2

Paper name: Building your kingdom Imitation Learning for a Custom Gameplay Using Unity ML-agents

Authors: Amira E.Youssef, Sohaila El Missiry, Islam Nabil El-gaafary, Jailan S. ElMosalami, Khaled M. Awad, Khaled Yasser

Methodology used: They made use of imitation learning to clone human behaviour and used its learning output as a dataset for reinforcement learning, reducing training time and increase learning performance. The approaches taken by Youssef et al. (2019) are comparable to those taken by Vadim, Tatyana and Oksana (2018) but for this study, they have chosen to create two separate environments; one was created specifically for recording demonstrations while a harder one was used to train the agents via imitation learning and reinforcement learning simultaneously. They've worked with both quantitative and qualitative data for their study; their results were training statistics presented in the form of graphs along with agent inference observations.

Objectives: The authors' research objectives are the same as those in the paper by Vadim, Tatyana, and Oksana (2018). Youssef et al. (2019) explored the simultaneous use of imitation learning and reinforcement learning to create an intelligent agent that can play realistically and effectively without making use of unfair advantage.

Implications: It was implied by the authors that imitating human-like behaviour in action games is a challenging task in machine learning research. In their study, they attempt to address this problem by training an agent using human demonstrations in conjunction with reinforcement learning.

Findings: They first trained the agent via imitation learning for 3 hours and 14 minutes. They then applied the recorded demonstrations by using them in conjunction with reinforcement learning. Training was stopped after 1 hour and 36 minutes. It was found by the authors that the optimum solution for creating agents with human-like behaviour is to use imitation learning and reinforcement learning simultaneously.

Conclusion: The authors concluded by stating that, using Unity and ML-Agents, they have been able to create and implement several AI agents which behave similarly to humans in a 3D game they have developed.

Paper 3

Paper name: Training an Agent for Third-person Shooter Game Using Unity ML-Agents

Authors: Jun LAI, Xi-liang CHEN and Xue-zhen ZHANG

Methodology used: In their study, Lai, Chen and Zhang (2019) trained several agents to play in a third-person shooter environment. They used the Unity game engine along with the ML-Agents toolkit to create the environment and train the agents. Their research was both a quantitative and qualitative one. They presented graphs to indicate a successful training run and described how the agent performed post-training.

Objectives: As stated by the authors, the purpose of their paper was to study the behaviour of a reinforcement learning agent in a third-person shooter environment on the Unity platform. They also highlight specific tasks, being target search, collection of objects, and elimination of enemies.

Implications: In the introduction, the authors highlight three AI research platforms: ViZDoom, Roboschool and ML-Agents. However, they only mention the limitations of the first two. Additionally, they state that experiments have shown that in video games, deep reinforcement learning can provide a higher level of intelligence for an agent than finite-state machines. Notwithstanding this, it is believed that they are implying the following:

- There is a lack of existing information about the use of Unity and ML-Agents for a third-person shooter game;
- The system proposed in their study can be used in place of finite-state machines.

Findings: During training, a steady increase in the cumulative reward was observed. The agents, after being trained, were able to do the following: search for enemies in unfamiliar areas; search for health pick-ups after being injured; collect ammunition; fight multiple enemies.

Conclusion: In their conclusion, Lai, Chen and Zhang (2019) state that they have successfully trained several agents to play in a complex 3D battlefield environment. They continued by stating that Unity3D and ML-Agents have shown to be an excellent reinforcement learning platform. Finally, they stated that when compared to finite-state machines, the agents in their proposed method implement a higher level of intelligence while requiring less coding.

Paper 4

Paper name: Real Time Strategy Games: A Reinforcement Learning Approach

Authors: Harshit Sethy, Amit Patel and Vineet Padmanabhan

Methodology used: In their study, the authors implemented the Q-learning and SARSA algorithms in a game called BattleCity. They trained the agent using these algorithms in conjunction with a generalised reward function. Their research method was quantitative; in addition to their implementations, they tested Darmok2's performance in-game. Darmok2 is a case-based planning system designed for real-time domains such as RTS games. Tests for the three algorithms were conducted by having the agent play a series of matches in different maps and against two different opponent types; AI-Random and AI-Follower. Then, they compared the algorithms in terms of win ratio and training time. They presented their results in the form of graphs. Unity and ML-Agents were not used in this study.

Objectives: They attempted to prove that their system could provide good results without having to learn from human demonstrations or requiring hardcoding.

Implications: The authors stated that their proposed method has two advantages over other works in the RTS genre. Firstly, they could ignore the concept of a simulator which is often game specific and usually hardcoded in RTS games, and secondly, their system could learn from interaction with opponents and quickly change the strategy accordingly without needing to observe expert demonstrations. They implied that their system is an improvement over other works on the genre.

Findings: The authors compared SARSA, Q-Learning and Darmok2. Their results were as follows:

- The agent won more than 90% of the matches when it played against both opponents in simple maps;
- The agent won 80% - 90% of the matches when it played against AI-Random in complex maps;
- The agent won 60% - 80% of the matches when it played against AI-Follower in complex maps.

They observed that SARSA provided the best results out of all three methods because it required the least time to train.

Conclusion: The authors stated that in terms of learning time and win ratio, their approach has shown to be better than previous approaches, albeit without using human demonstrations or requiring a hardcoded approach. They also mention the fact that as opposed to Darmok2, the two reinforcement learning algorithms could benefit from the reward function for better performance.

Bibliography:

Vadim, K., Tatyana, K. and Oksana, O. (2018). Simultaneous use of imitation learning and reinforcement learning in artificial intelligence development for video games, [online] Available at: http://ceur-ws.org/Vol-2254/10000154.pdf [Accessed 16 Nov. 2019].

Youssef, A., Missiry, S., Nabil El-gaafary, I., ElMosalami, J., Awad, K. and Yasser, K., 2019. Building your kingdom Imitation Learning for a Custom Gameplay Using Unity ML-agents. 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON),.

LAI, Jun & Xiliang, Chen & ZHANG, Xue-zhen. (2019). Training an Agent for Third-person Shooter Game Using Unity ML-Agents. DEStech Transactions on Computer Science and Engineering. 10.12783/dtcse/icaic2019/29442.

Sethy, H., Patel, A. and Padmanabhan, V. (2015). Real Time Strategy Games: A Reinforcement Learning Approach. Procedia Computer Science, [online] 54, pp.257-264. Available at: https://www.sciencedirect.com/science/article/pii/S187705091501354X#! [Accessed 18 Nov. 2019].

Creswell, J., 2013. *Research Design: Qualitative, Quantitative, And Mixed Methods Approaches*. SAGE Publications Ltd.