1- Objective

The objective of this project is to implement and test a state-of-the-art Deep RL algorithm and note the differences between it and vanilla RL. More concretely, this project wishes to compare the performance of Q-learning (implemented in Project 3) and a chosen Distributional RL algorithm, a suitable algorithm would be Quantile-Regression Deep Q-Network (QR-DQN).

The result of this project is significant because if QR-DQN outperforms Q-learning in most situations, this opens paths to apply QR-DQN to agents in fields other than game agents and results in more rational agents that are able to learn more efficiently in the real world

2- Related Work

Classical DQN aims to approximate the expected discounted rewards (Q-values). In other words, it tries to find the average reward of each action in a particular state.However, Bellemare et. al [1] have shown that estimating the mean of the reward does not always yield a desirable result. This is why they proposed a novel method by which they find the probability distribution of the discounted rewards, rather than only the mean.

This probability distribution is estimated by minimizing the loss function of Quantile Regression. This regression differs from Ordinary Least Squares in that it finds the best weights for the specified quantiles (a.k.a percentiles). Thus, we can successfully estimate the true probability distribution by computing the weights for each quantile [2].

3- Technical Outline

Our principal aim is to reimplement the given algorithm for different types of action spaces. For instance, the current implementations for QR-DQN do not have support for Box, MultiDiscrete, and MultiBinary. We will try to implement the algorithm for at least one of them. Furthermore, we will test the performance of the algorithm with benchmarks other than the authors of the original paper used.

References

[1] Bellemare, M. G., Dabney, W., & Munos, R. (2017, July). A distributional perspective on reinforcement learning. In International Conference on Machine Learning (pp. 449-458). PMLR.

[2] Dabney, W., Rowland, M., Bellemare, M., & Munos, R. (2018, April). Distributional reinforcement learning with quantile regression. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 32, No. 1).