Abstract

The field of Reinforcement Learning, and the use of Reinforcement Learning agents has been growing increasingly in popularity in the most recent years. Despite this, there is still an insufficient number of established testing techniques which are designed for Reinforcement Learning problems. There are currently a number of proposals to address this gap, but many are still insufficiently tested, and not widely used.

In the thesis, we aim to adapt an existing proposed technique - Mutation Testing - which was originally adapted from its general use in traditional Supervised Learning, for application on Reinforcement Learning agents. We will be further adapting this approach to work on a wider range of algorithms and environments, in this field. The existing work, can currently only facilitate discrete action space-based environments, and three algorithms: Deep Q-Network, Proximal Policy Optimization, and Advantage Actor Critic. Therefore, we will be expanding on this, with particular focus on the implementing the Deep Deterministic Policy Gradient algorithm, and the CartPoleContinuous environment - a continuous action space.

This implementation will then be assessed using two evaluation metrics: AVG and R, which were designed specifically for Mutation Testing on Reinforcement Learning problems. The results will be analysed under the evaluation objectives: the effectiveness of Mutation Testing overall, the effectiveness of Mutation Testing on a continuous environment in comparison to a discrete environment, and finally the effectiveness of the evaluation metrics being utilised.

We find that the proposed technique - Mutation Testing - is effective, dependent on the mutation introduced to system (agent, environment, policy). We found that some mutants had more of a significant impact on the system, than others, which was evidenced by certain mutants which were consistently detected and killed, and others that remained undetected. We find that based on the results, the effectiveness of Mutation Testing is virtually the same on both continuous, and discrete action space-based environments. Finally, we conclude that both evaluation metrics proved to be useful and effective, and when used in conjunction with one another, produce convincing and reliable conclusions.