A COMPARISON BETWEEN DEEP Q-LEARNING AND ACTOR-CRITIC IN THE CARTPOLE ENVIRONMENT

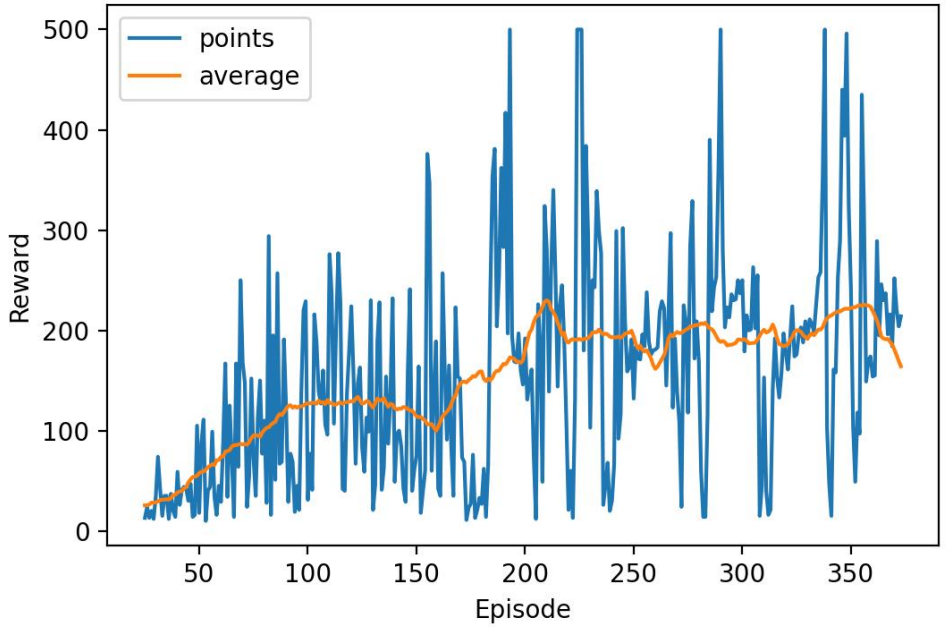
Giulio Vaccari

Autonomous and Adaptive Systems 2020/2021

06/2021

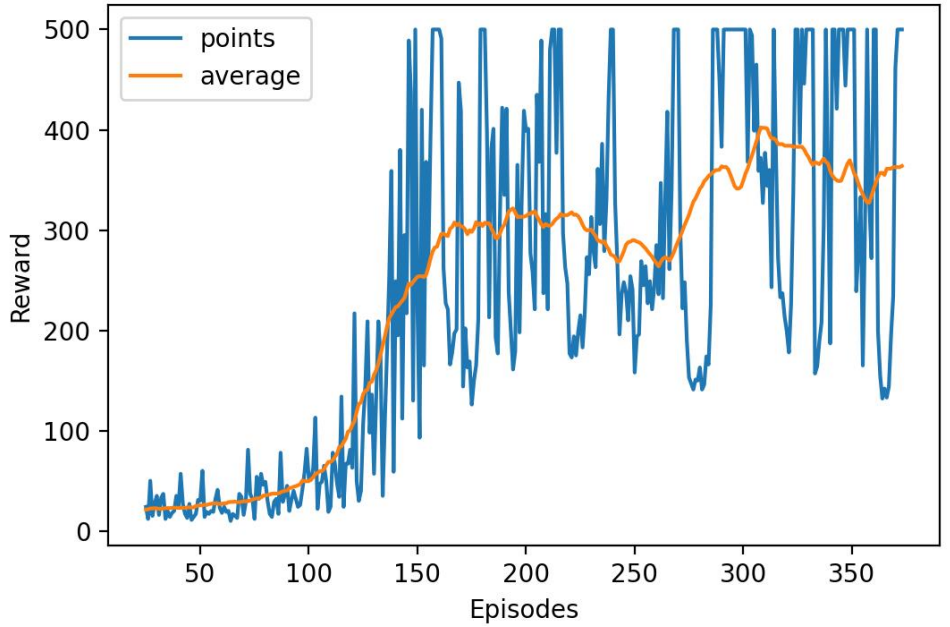
DQN Agent

- Value network represented with a Deep Neural Network consisting of 3 dense layers.
- Time-difference approach.
- ϵ-policy with ϵ value linearly decreasing over training episodes.
- Experience replay buffer of size 2000
- Learning rate equal to $1e-3$



Actor-Critic Agent

- Critic and Actor components represented with two distinct Deep Neural Network consisting of 3 dense layers.
- Time-difference approach.
- Small replay buffer of size 500 to train only on recent experiences.
- Learning rate of $1e-3$ for the Critic and of $1e-5$ for the Actor.



Results

- Actor-Critic shown higher and more stable results during training.
- In the final evaluation Actor-Critic performed **considerably better**, with an average score in 20 simulations of $\mathbf{488}$ compared to the score of $\mathbf{182}$ for DQN.