Assignment 3

Jacqueline Wu, Sam Zhang

March 2024

1 Value-based method with linear function approximation

We implement Q-Learning and Expected SARSA and experiment them with two environments from the Gymnasium, an open-source fork of the OpenAI Gym. [1] During the experiment, the environment's states are discretized through tile-coding into a binary vector.

1.1 Results

We experiment with 3 ϵ per algorithm-environment pair, and use a fixed set of 3 different learning rates $\alpha \in \{\frac{1}{4}, \frac{1}{8}, \frac{1}{16}\}$. The results are shown in Figure 1.

1.2 Discussion

2 Policy Gradient Theorem

3 Policy-based methods with linear function approximation
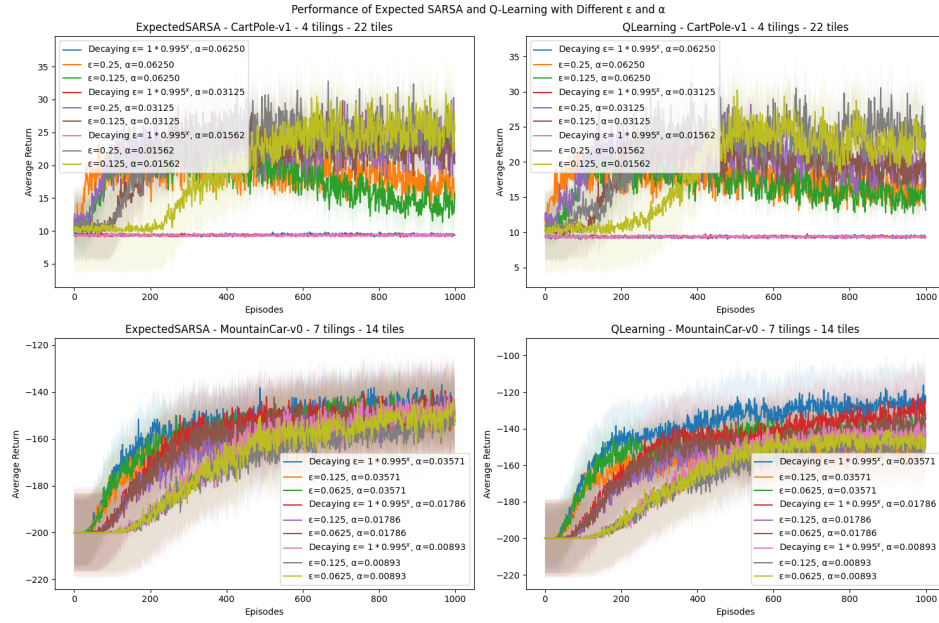
3.1 Results

3.2 Discussion

Figure 1: ExpectedSARSA and Q-Learning's Performances on CartPole-v1 and MountainCar-v0

References

[1] Mark Towers, Jordan K. Terry, Ariel Kwiatkowski, John U. Balis, Gianluca de Cola, Tristan Deleu, Manuel Goulão, Andreas Kallinteris, Arjun KG, Markus Krimmel, Rodrigo Perez-Vicente, Andrea Pierré, Sander Schulhoff, Jun Jet Tai, Andrew Tan Jin Shen, and Omar G. Younis. Gymnasium, March 2023.