

Reinforcement Learning

Coding Exam

18/11/2022

Sanjit K. Kaul

Instructions: You have an hour and twenty minutes to work on the questions and submit your code (python files) and PDF with plots and an interpretation of your results on Classroom. Delayed uploads will not be graded. Exam is open book and class notes. You can also use your own HW submissions, if applicable. No other resources, human or otherwise are allowed. Violation of the policy will be treated as use of unfair means and plagiarism. All academic penalties will apply.

Question 1. 100 marks You will implement two algorithms including Double-Q Learning (DQL) and Double-Q Learning With a Twist (DQLT). You will stick to tabular methods. Recall that in DQL, we maintain two copies of Q-values. When updating either, we pick the greedy action. In DQLT, we will change this a bit. For one of the copies, we will instead choose the ϵ -greedy action.

The goal is for the agent to accumulate rewards as fast as possible. You will compare the learning performance of DQL and DQLT for the following environment.

Both environments have the six states 0, 1, 2, 3, 4, 5. In states $i \in \{1, 2, 3, 4\}$ your agent can choose either the action Left or the action Right. The action Left changes state i to $i - 1$ and the action Right changes the state i to $i + 1$. In state 0 the agent can only choose Right and in state 5 the agent can only choose Left. The action Left in any state results in a reward of 0. The action Right in states $\{0, 1, 2, 3\}$ results in a reward of 1. The action Right in 4 results in a reward that is Gaussian with mean -5 and standard deviation 3. Assume $\gamma = 0.98$.

Compare the performance of DQL and DQLT. Make any assumptions that are necessary. For example, make suitable choices of α and ϵ . Preferably, demonstrate the learning performance of each algorithm over more than one choice of the parameters. You must create a PDF with plots that compare the performance of the algorithms (for any number of choices of parameters as you choose). All plots must be properly labeled and each plot must be accompanied by your interpretation of the same. There must be at least one plot that compares DQL and DQLT. In addition to the PDF, you will upload your python code as part of your submission. The code must include your implementation of the algorithms and also the code you used to plot your results. When grading your submission, we will run the provided code as is.