Prof. Joschka Bödecker, Gabriel Kalweit

REINFORCEMENT LEARNING
Exercise 4

This week, we use the methods from last week for control. You find tests for the following tasks in exercise-04_test.py. Additionally, there is a visualization script. You again need matplotlib. Run the tests and the visualization by the usual commands.

0 Lecture

Watch *Lecture 05: Value Function Approximation*[1] before the upcoming session on Friday, November 23.

1 Off-Policy MC Control with Importance Sampling

This task is again based on the Blackjack environment from the lecture and the last exercise. Implement First Visit Off-Policy MC Control with Importance Sampling,

mc_control_importance_sampling(env, num_episodes, behavior_policy, discount_factor=1.0),

in off_policy_mc.py.

2 Q-learning

The tests for this task are based on the Cliff Walking example from the lecture. An implementation can be found in cliff_walking.py. Implement the Q-learning algorithm,

q_learning(env, num_episodes, discount_factor=1.0, alpha=0.5, epsilon=0.1),

in q_learning.py.

3 Experiences

Make a post in thread *Week 04: Model-free Control* in the forum[2], where you provide a brief summary of your experience with this exercise, the corresponding lecture and the last meeting.

---

[1] https://ilias.uni-freiburg.de/goto.php?target=xvid_1121350&client_id=unifreiburg
[2] https://ilias.uni-freiburg.de/goto.php?target=frm_1121060&client_id=unifreiburg