Reinforcement Learning
Prof. Milica Gašić

Release: Thursday, 21 November 2024
Discussion: Wednesday, 28 November 2024

Exercise set #7

You do not have to hand in your solutions to the exercises and they will **not** be graded. However, there will be four short tests during the semester. You need to reach ≥50% of the total points in order to be admitted to the final exam (Klausur). The tests are held at the start of a lecture (room 2522.U1.74) at the following dates:

Test 1: Thursday, 31 October 2024, 10:30-10:45
Test 2: Thursday, 21 November 2024, 10:30-10:45
Test 3: Thursday, 5 December 2024, 10:30-10:45
Test 4: Thursday, 9 January 2025, 10:30-10:45

Please ask questions in the RocketChat
The exercises are discussed every Wednesday, 14:30-16:00 in room 2512.02.33.

1. **Expected Sarsa**

   Continue implementing our TD Agent from last week. This time, implement expected Sarsa and apply them to the cliff walking environment from the lecture. We will finish recreating Figure 6.3 from the textbook[1], which compares different learning rates. Follow the instructions in the Jupyter notebook `td-control-2.ipynb`.

2. **Double Q-learning**

   Implement double Q-learning to recreate Figure 6.5 from lecture 6. Follow the instructions in the Jupyter notebook `double-q-learning.ipynb`.

3. **n-step return**

   Implement the n-step return . Follow the instructions in the Jupyter notebook `td-returns.ipynb`.

Exercises by Stefan Harmeling, used with permission
[1] `http://incompleteideas.net/book/the-book-2nd.html`