Prof. Frank Hutter, Prof. Joschka Bödecker

Dr. Marius Lindauer
Gabriel Kalweit

Reinforcement Learning
Exercise 7
Submit until **Thursday, January 18 at 2:00pm**

1 Regret of UCB1 and Decaying ϵ-greedy (20p)

(a) Implement the different exploration strategies – UCB1, ϵ-greedy, random exploration and decaying ϵ-greedy exploration for some schedule of your choice – in

`YOUR_REPO/exercise-07/scripts/exploration.py`.

Explain your schedule. Evaluate the exploration strategies on the given toy problem by plotting the regret of UCB1, decaying ϵ-greedy, ϵ-greedy and random exploration and compare the results. (15p)

(b) Imagine you want to apply the algorithms from this lecture on a real physical system and some actions in some states may break your robot, so you have avoid them (but you do not know those states beforehand). However, the presented algorithms **need** to explore in order to find a good solution. What problems occur with some of the presented exploration strategies and how would you approach the exploration? (5p)

2 Bonus: Experiences (1p)

Submit an `experiences.txt`, where you provide a brief summary of your experience with this exercise, the corresponding lecture and the last meeting. As a minimum, say how much time you invested and if you had major problems – and if yes, where.

Please push your solutions to subdirectory `exercise-07` in your assigned git-repository by **Thursday, January 18 at 2:00pm**. **Solutions after that or via email will not be accepted.**