

## REINFORCEMENT LEARNING Exercise 1



### 1 Exploration

Implement the different exploration strategies – UCB, Softmax Action Selection,  $\epsilon$ -greedy, random exploration and decaying  $\epsilon$ -greedy exploration for some schedule of your choice – in `exploration.py`. Play around with the hyperparameters. Evaluate the exploration strategies on the given toy problem by plotting the regret and compare the results.

### 2 Experiences

Make a post in thread *Week 01: Bandits and MDPs* in the forum<sup>1</sup>, where you provide a brief summary of your experience with this exercise and the corresponding lecture.

---

<sup>1</sup>[https://ilias.uni-freiburg.de/goto.php?target=frm\\_1837317&client\\_id=unifreiburg](https://ilias.uni-freiburg.de/goto.php?target=frm_1837317&client_id=unifreiburg)