Prof. Frank Hutter, Prof. Joschka Bödecker

Dr. Marius Lindauer
Gabriel Kalweit

<div align="center">

REINFORCEMENT LEARNING
Exercise 5
Submit until **Thursday, December 14 at 2:00pm**

</div>

Preliminaries

This exercise is based on Lecture 6[1] from David Silver's RL course[2]. Watch before the upcoming meeting on Friday, December 8.

1 TensorFlow (6p)

Read through the TensorFlow tutorials[3]. Implement the `NeuralNetwork` class based on TensorFlow in

<div align="center">

`YOUR_REPO/exercise-05/scripts/q_learning_fa.py`.

</div>

2 Q-learning with Function Approximator (14p)

(a) Implement Q-learning with function approximator in

<div align="center">

`YOUR_REPO/exercise-05/scripts/q_learning_fa.py`,

</div>

based on your implementation such that it uses your neural function approximator, then apply it on the modified Gym Mountain Car environment which you find in `lib.envs.mountain_car`. It includes a shaped reward and a deterministic initial state. Run the control loop. You can play around with the parameters and an exemplary setting is already given. Provide learning curves of your experiments. A plotting function is given in the script. (7p)

(b) Now add experience replay and a target network. You now do not update on the transition you collect in a time step, but sample a minibatch and train on that. The targets are now calculated using the target network. Again, submit learning curves and a short comparison about your experiences with Q-learning – with and without experience replay and target networks. (7p)

---

[1] https://youtu.be/UoPei5o4fps
[2] http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching.html
[3] https://www.tensorflow.org/get_started/

3 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-05` in your assigned git-repository by **Thursday, December 14 at 2:00pm. Solutions after that or via email will not be accepted.**