Reinforcement Learning
Exercise 6

Jim Mainprice, Philipp Kratzer
Machine Learning & Robotics lab, U Stuttgart
Universitätsstraße 38, 70569 Stuttgart, Germany

June 5, 2020

Submission Instructions:

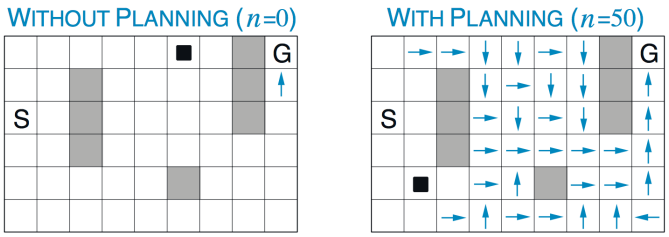
The submission deadline for this exercise sheet is 16.06., 23:55.

Put your answers into a single pdf. Your python code should be a single python script. Upload both files to ilias. Make sure that the code runs with *python3 yourscript.py* without any errors.

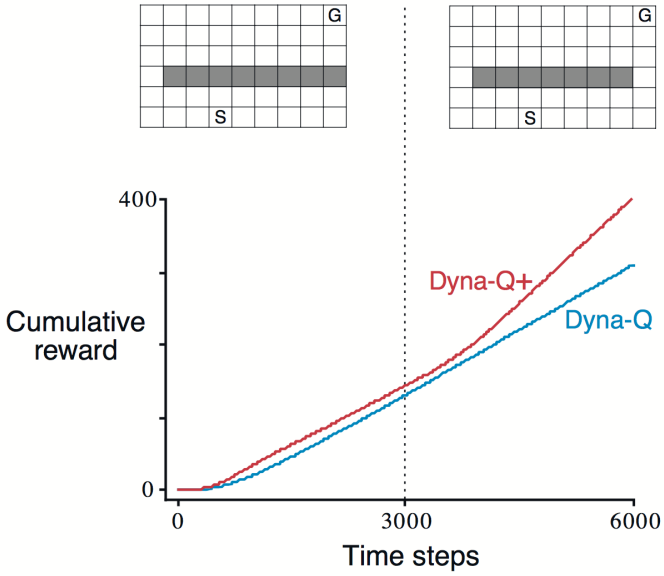
Group submissions of up to three students are allowed.

1 Planning and Learning (4P)

a) The nonplanning method looks particularly poor in the figure below; a method using n-step bootstrapping would do better. Do you think one of the n-step bootstrapping methods could do as well as the Dyna-Q method? Explain why or why not. (2P)



b) Why did the Dyna-Q+ (i.e., with exploration bonus) perform better in the first phase as well as in the second phase of the blocking and shortcut experiments (see figure below)? (2P)



2 n-step sarsa on the FrozenLake (4P)

The code template can be found on github (`https://github.com/humans-to-robots-motion/rl-course`) in *ex06-nstep/ex06-nstep.py*.

Implement n-step Sarsa and evaluate it on the 8×8 *Frozen Lake* environment. Evaluate the performance for different choices of n and α. Visualize your results (plot the performance over α for different choices of n, similar to lecture 6 slide 9).