

Homework

- Use the MountainCar-v0 domain from OpenAI gym
 - See Example 10.1, pg 244 in Sutton and Barto [2018]
- Implement the SARSA semi-gradient algorithm with tile coding (8 tiles)
 - Use the value function code provided on Moodle. The `__call__` function computes the value of a given state-action pair, the `update` function updates the Q-function, and the `act` function derives an epsilon-greedy policy
 - Use ϵ -greedy policies with $\epsilon = 0.1$ and a learning rate of $\alpha = 0.1$

By next week's lecture, submit on Moodle:

1. Run the algorithm for 500 episodes, and plot the number of steps per episode (on a log scale) as a function of number of episodes. Average your results over 100 runs and submit the graph
2. Use the rendering ability of OpenAI gym to render the learned Q-function acting in the environment (do not use exploration for this). Create a video of it solving the task and upload the video
3. Your code