Homework 02 Review for Group 09

General Feedback

- Overall a nice notebook, everything works out of the box :)
- Some markdown formatting to structure the notebook and separate the different tasks would have been nice
- Nice commenting and use of docstrings, though the latter are incomplete (i.e. missing arguments and return description) or are missing completely (i.e., Gridworld class and init of Tab_SARASAN class)

Task 01

- Nice way of initializing a board randomly, especially checking whether a path exists to the terminal state via search
- Good visualization with a nice legend for overview
- Intermediate negative reward for all other states seems relatively high compared to the terminal state reward

Task 02

- Minor remark: you could have cleared the output after each epoch to not clutter the screen
- Maybe also include the cumulative reward in the visualize=false setting to at least have that information as well, i.e.print it together with the current episode
- Minor remark: would have been nice to be able to provide the gridworld as an argument for the Tab_SARASAN class in order to train on the same grid world that one configures beforehand
- Apart from that, everything works nicely. Well done! :)

Homework 02 Review for Group 16

General Feedback

- In general, this notebook seems well coded, formatted and everything runs smoothly

Task 01

- Missing Docstrings, although you've provided extensive comments - still nice to have them
- Cool idea to use the recursive backtracking algorithm to create random mazes
- Why exactly has the size to be odd in order for the agent to find a goal?

Task 02

- Legend of the different colors used in the gridworld would be helpful, otherwise really nice visualization - almost looks like a real board game (really cool!)
- General description and returns within the docstrings are missing
- Maybe use the clear_output function at the beginning of each episode to not clutter the notebook output
- Only one-step Sarsa is implemented
- It seems as if the agent cannot move through the white tiles. Therefore, in some cases, the agent is trapped and can't reach the goal
- Apart from that, everything works nicely. Well done! :)

Homework 02 Review for Group 46

General Feedback

- Overall, your work looks really good and like you put a lot of effort into this. I also love how clean and smooth your code looks!

Task 01

- Your grid world has all it needs and it even checks for exceptions!
- You allow for arbitrary grid sizes, which is neat.
- The “random state transition” works, but could have been a little more creative.

Task 02

- Looks great overall. Your implementation of SARSA should be correct.
- The idea of using a distance measure to measure training progress looks really useful. However, defining it on the policies can lead to misleading results. For example, if action a is suboptimal in state s, but its Q-Value is highly overestimated, it may take a few iterations to change the policy in s. During this time, your distance measure between both policies will be zero (for this state), even though the Q-values may be changing a lot on every iteration.
- I also like that you compared the runtimes for different epsilon-values. Plotting both against each other also would have been interesting too.