Dear Group 8,

Your code is very well documented and has a nice structure.
Sometimes it's still a bit hard to understand the code though, maybe in some cases the variables and function names could be more meaningful, or it's just us being too tired ^^.

We really like your visualization. You could reduce the number of times it gets plotted to every x-th episode or something, as the console gets pretty clutted this way and we're not sure if it is really necessary to print every step of every episode.
Additionally, it seems like the agent isn't learning so fast, even in the last episode it still often happens that he goes opposite to the "right" direction or hops back and forth between states and the Policy you print doesn't seem to be very optimal to the environment.

Best,
Group 20