Homework #4

In programming assignments 1, 2, 3, use the supplied gridworld_lec4.py. For problem 1, use the optimal policy for gridworld 5x5. You should also print the number of visited states.

1) Implement MC exploring starts to compute value function of GridWorld. Compare against exact algorithm using the model dynamics.

2) Implement MC exploring starts control of GridWorld.

3) Implement on-policy first visit MC control of GridWorld without exploring starts.

4) Implement off-policy MC control of GridWorld.

5) Prove that the incremental version of weighted importance sampling from:

$$V_n \doteq \frac{\sum_{k=1}^{n-1} W_k G_k}{\sum_{k=1}^{n-1} W_k},$$

Can be given by the following set of equations.

$$V_{n+1} \doteq V_n + \frac{W_n}{C_n}\Big[G_n - V_n\Big], \qquad n \geq 1,$$

and

$$C_{n+1} \doteq C_n + W_{n+1},$$

6) Install Gymnasium and test it with Cartpole. You should create an agent that plays Cartpole using MC.

7) Implement a python program based on keyboard input to play CarRacing-v0.

8) Implement a MC to play tic-tac-toe.