Lab 1: Policy and Value Iteration

Ibrahim Sammour

Consider a "Grid World" with dimensions (X,Y), a player should begin moving from the start position till reaching the goal while collecting the maximum amount of rewards.

Rules:

- The player can move (Up, Down, Left, Right) when possible.
- The probability of taking an action and getting it executed properly is 100%. (i.e. if I took a decision to move up I will certainly go up).
- The player is not allowed to pass through the same state twice during an episode.

Tasks:

1- Understand the code of the grid world and run it (do some adjustments to make sure you understand the code).
2- Now in a separate code file:
    a. Add and initialize the proper reinforcement learning based variables (Value function i.e.).
    b. Add the variables needed to run the code in episodic manner (take into consideration the rules given above).
    c. Implement policy iteration and test it.
    d. Vary the discount factor gamma and compare the results.

e. Implement value iteration and test it.

3- Modify the Grid World code to show the best path (colored).

4- Use the Grid World code to create a separate Grid showing the value of each state.