For both versions of Magneto, your job will be to compute the wolverine's optimal policy, i.e. the action that should be taken at each state to maximize the wolverine's expected discounted reward, where we fix the discount factor (gamma) to be 0.85.

Task 1: Implement value iteration for both versions of Magneto on MDP (10 points)

Task 2: Implement policy iteration for both versions of Magneto (10 points).

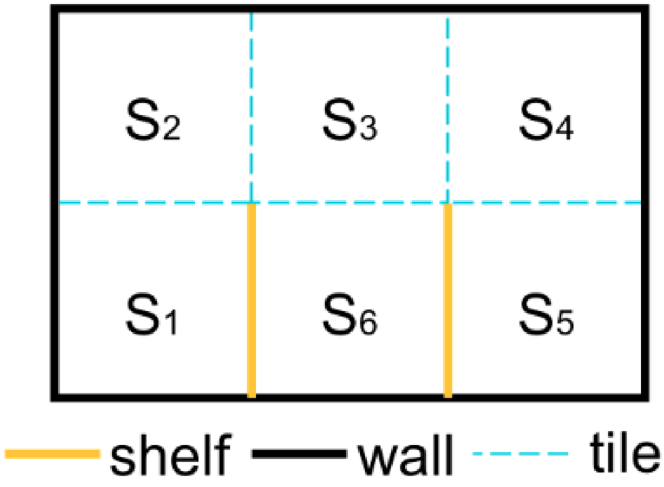
Task 3: Implement and visualize the MDP board and strategy (policy) graphically. (10 points)

Task 4: Compare the results for all the four approaches - Value iteration for lazy Magneto, Value iteration for active Magneto, Policy iteration for lazy Magneto, Policy iteration for active Magneto with proper tables/graphs/statistics. Comment which one is best among all the four mentioned approaches. (10 points)

PART IV
Robot localization using Hidden Markov Model (HMM) [20 Points]

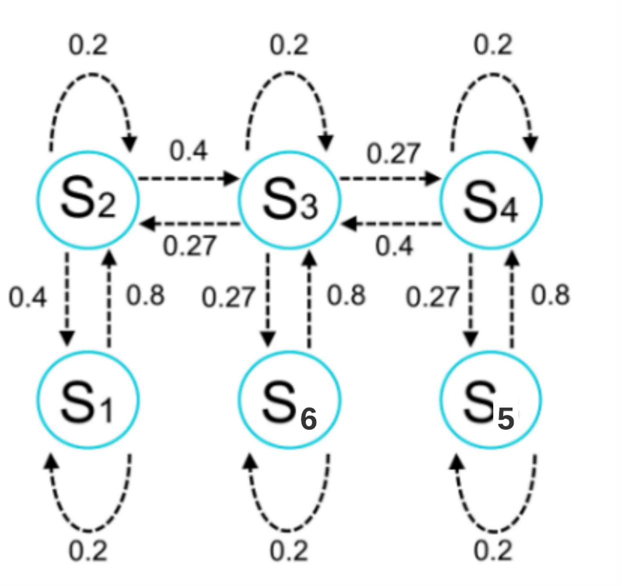
Problem of mobile robot in a warehouse. The agent is randomly placed in an environment and we, its supervisors, cannot observe what happens in the room. The only information we receive are the sensor readings from the robot.

Environment



The agent can move within an area of 6 square tiles. In the mini-warehouse there is one shelf located between tiles S_1 and S_6 and a second shelf between S_6 and S_5

The environment of the agent consists of six discrete states. The time is also discrete. At each subsequent time step the robot is programmed to change its position with a probability of 80% and moves randomly to a different neighboring tile. As soon as the robot makes a move, we receive four readings from the sensing system.



Sensors

Just as we humans can localize ourselves using senses, robots use sensors. Our agent is equipped with the sensing system composed of a compass and a proximity sensor, which detects obstacles in four directions: north, south, east and west. The sensor values are conditionally independent given the position of the robot. Moreover, the device is not perfect, the sensor has an error rate of e = 25%

You need to implement the above as an HMM to track the robot and report its location after every 10 iterations (total 100 iterations) alongwith the estimated probabilities.

Best of Luck!!!