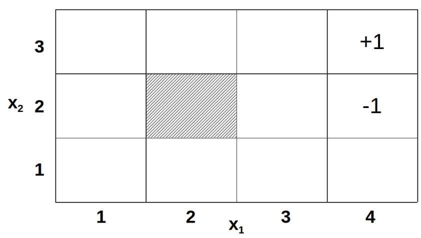
[Pre-Lab] Reinforcement Learning

Jae Yun JUN KIM*

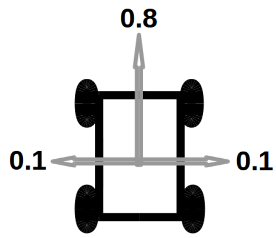
Due: Before the start of the coming ab session.
Remark:

- Only groups of two or three people accepted (preferably three). Forbidden groups of fewer or larger number of people.
- Do the below tasks in order so that you can work efficiently during the coming two-hour lab session.
- No plagiarism. If plagiarism happens, both the "lender" and the "borrower" will have a zero.
- Code yourself from scratch. No lab work will be considered if you solve the problem using any ML library.
- Do thoroughly all the demanded tasks.
- Study the theory for the questions.

For this lab session, you are asked to plan the motion of a 2D mobile robot using the Markov Decision Process formalism. Consider the following 2D map for the autonomous navigation of a mobile robot



(a) 2D map



(b) 2D mobile robot

Figure 1: 2D map and 2D mobile robot

This map consists of 12 cells. The dashed cell at (x_1, x_2)=$(2, 2)$ represents an obstacle to be avoided. The cell with reward "+1" at (x_1, x_2)=$(4, 3)$ is a desired absorbing cell (the goal), while the cell with reward "-1" at (x_1, x_2)=$(4, 2)$ is an undesired absorbing cell (e.g., a pit). On the other hand, the mobile robot can take four actions: A=$\{N, S, E, W\}$, where N, S, E, W represent north, south, east and west, respectively. If A=N, then the mobile robot behaves following transition probability distribution indicated in Figure 1(b). This is also true for the rest of actions. Further, the reward function is defined as follows

$$R = \begin{cases} +1 & (x_1, x_2)=(4,3) \\ -1 & (x_1, x_2)=(4,2) \\ -0.02 & \text{otherwise} \end{cases} \tag{1}$$

Finally, assign a value for the discount factor (γ) at your convenience.

1 Tasks

1) For any state in the map illustrated in Figure 1(a), implement a function in Python that returns a next state, following the state-transition distribution expressed by the robot model indicated in Figure 1(b), while satisfying the constraints imposed by the map given in Figure 1(a).

*ECE Paris Graduate School of Engineering, 37 quai de Grenelle 75015 Paris, France; jae-yun.jun-kim@ece.fr