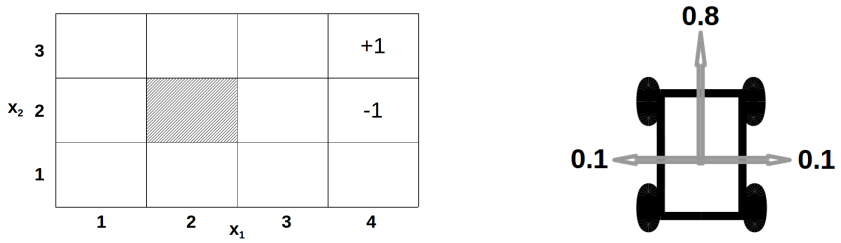
[Lab] Reinforcement Learning

Jae Yun JUN KIM*

Lab due: Before the end of today lab session
Evaluation: By showing your code and results, and explaining them to the Professor.
Remark:

- Only groups of two or three people accepted (preferably three).
- No late evaluation is accepted.
- No plagiarism. If plagiarism happens, both the "lender" and the "borrower" will have a zero.
- Code yourself from scratch **strictly following the theory given during lecture**.
- No lab work will be considered if any ML library is used.
- 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

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

1 Tasks

1) Download the Python scripts from the course website to your working folder.

2) For all states, find the optimal value function $Q^*(s, a)$ and the optimal policy function $\pi^*(s)$ using the *Q-learning* algorithm. Choose a value for the discount factor (γ) at your will.

3) **You must do the lab task using these Python scripts and following the indications given in these files. No code other than these will be considered during the evaluation time.**

*ECE Paris, 10 Rue Sextius Michel 75015 Paris, France; jae-yun.jun-kim@ece.fr