Reinforcement Learning

Jae Yun JUN KIM*

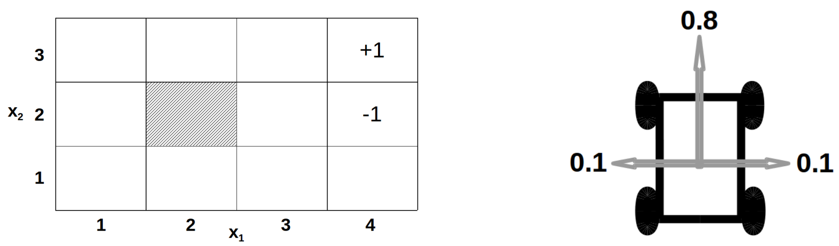
Due: Upload your code to the course website before the deadline indicated on campus.ece.fr.
Evaluation:

- code (in group of two or three people)

Remark:

- Only groups of two or three people accepted. Forbidden groups one or larger than three people.
- No late homework will be accepted.
- No plagiarism. If plagiarism happens, both the "lender" and the "borrower" will have a zero.
- Code yourself from scratch. No homework will be considered if you solve the problem using any ML library.
- Do thoroughly all the demanded tasks.
- Study the theory for the interrogation.

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 the discount factor (γ) to be 0.99.

1 Tasks

1) For all states, find the optimal value function $V^*(s)$ and the optimal policy function $\pi^*(s)$ using the *value iteration* algorithm.

2) For all states, find the optimal value function $V^*(s)$ and the optimal policy function $\pi^*(s)$ using the *policy iteration* algorithm.

3) Compare the results obtained in 1) to those of 2).

---

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