Machine, Data and Learning

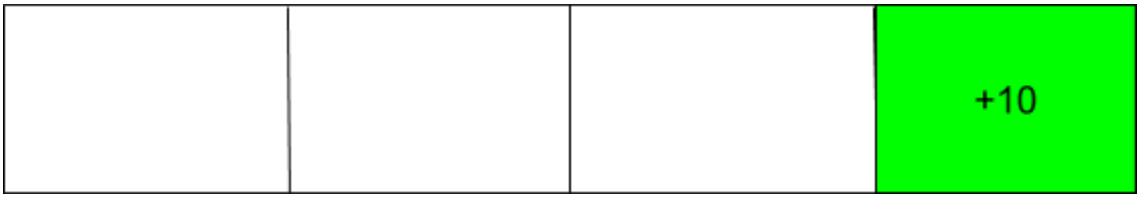
Assignment-2 Part-1

Task:

Dry-run/apply the value iteration algorithm on the following scenario to obtain the optimal policy and the state reward values corresponding to it:

Scenario:

- **State-space**: 1D array having 4 states



  - The rightmost state is a **terminal state** with a reward **+10.**

- **Actions** and **Transition probabilities: (stochastic)**
  - **Move right -** moves right with a probability of **0.8** and moves left with a probability of **0.2.**
  - **Move left -** moves left with a probability of **0.8** and moves right with a probability of **0.2.**

- **Reward:**
  - For each step taken, a **negative reward = -1** is assigned to the agent.

- **Parameters:**
  - **Gamma = 0.25** (The discount factor).
  - **Delta = 0.01** (Convergence or Bellman error).

Note:

- This is a handwritten assignment(part-1).
- In each iteration, calculations corresponding to each state and its transitions should be clearly calculated and shown. Nothing will be assumed during evaluation unless it is specified in the solution/dry-run.
- Round off the calculated values to **3 decimal places**(in every iteration).

- The VI algorithm has to be applied **until convergence.**
- This will be an **individual assignment.**
- Part-1 weightage = **20%** of assignment-2.
- Part-2 will be released soon.

Submission Details:

- **Deadline:** 6:00 pm, 17th Feb 2020(Monday).
- Hand over the assignments to Sai Kamal between 5:00 pm and 6:00 pm on Monday in Machine Learning Lab(MLL).