

# HW 1- Policy Iteration for Taxi Cab Problem

Due date: 2/22- before the class

Let's design a policy for a taxi cab using policy iteration. The Smartcab's job is to pick up the passenger at one location and drop them off in another.

**State Space** In reinforcement learning, the agent encounters a state, and then takes action according to the state it is in. The state space is the set of all possible situations our taxi could inhabit. The state should contain useful information the agent needs to make the right action.

**Action Space** The agent encounters one of the states and it takes an action. The action in our case can be to move in a direction or decide to pickup/dropoff a passenger. In other words, we have six possible actions as follows: south, north, east, west, pickup, and dropoff. The taxi cannot perform certain actions in certain states due to walls. In the environment's code, we will simply provide a -1 penalty for every wall hit and the taxi won't move anywhere. This will just rack up penalties causing the taxi to consider going around the wall.

**Reward** The agent should receive a high positive reward for a successful dropoff because this behavior is highly desired. The agent should be penalized if it tries to drop off a passenger in the wrong location. Also, The agent should get a slight negative reward for not making it to the destination after every time-step. Slight negative because we would prefer our agent to reach late instead of making wrong moves trying to reach to the destination as fast as possible.

**Tasks** All of these components have already been implemented by OpenAI Gym environment [1]. Your task is to implement the policy iteration algorithm discussed in class and find the optimal value function and optimal policy. *It is your job to show the optimal value function inside the Jupyter notebook.* Find the best

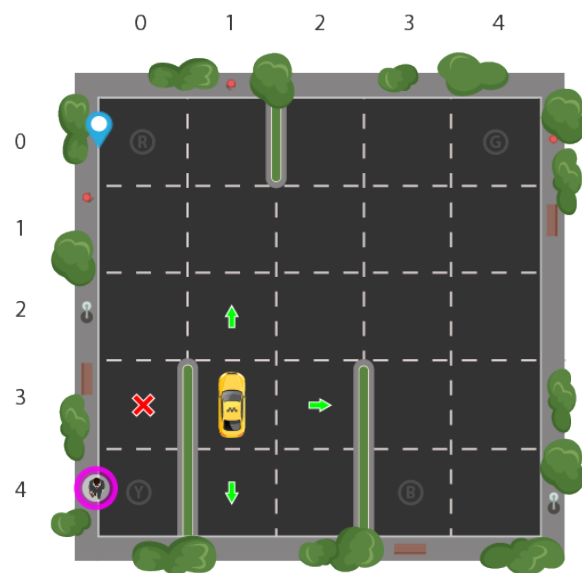


Figure 1: Taxi Cap OpenAI Gym Problem

way to do that! Upload a Jupyter notebook in your GitHub account and send the link via email with the following subject line: your name – RLC2021 – HW1.

### **More Reading**

- Reinforcement Learning: let's teach a taxi-cab how to drive
- Reinforcement Learning for Taxi-v2

### **References**

[1] [www.learndatasci.com/tutorials](http://www.learndatasci.com/tutorials).