Waterpark World

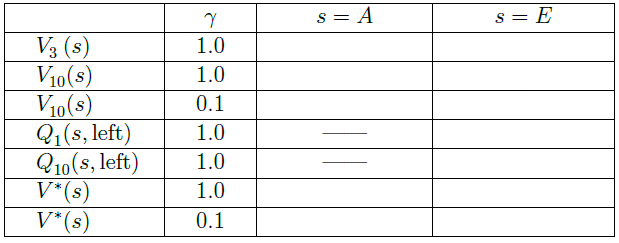
Part I: MDP Values

In the “Waterpark World” environment, an agent can move either clockwise or counterclockwise, except in 3 states where the agent must go clockwise due to the presence of a water slide. There are two states that represent being on a ladder leading up to the slide. Rewards for moving are 0, except that going down either portion of the slide gives a +2 reward (fun!) and climbing either part of the ladder has a reward of -1 (feels like work!). The agent has to move on every turn. Assume noise = 0.0 (i.e., actions are deterministic).



1. How many distinct policies are possible for this MDP? 2^11

(b) Fill in the blank cells of this table with values that are correct for the corresponding function, discount, and state. Hint: You should not need to do substantial calculation here.



|  |  |
| --- | --- |
| 0 | 4 |
| 2 | 4 |
| 0 | 2.2 |
| 0 | 0 |
| 2 | 3 |
| inf | inf |
| 0 | 2.2 |

(V-ST-19-11-01. This exercise is based on an example used at U.C. Berkeley)