COMP 560: Assignment 2

Reinforcement Learning

Group Members: Arturo Schmidt (only 1)

**Environment Setup**

The environment is parsed from stdin as is represented as a dictionary mapping a state and action to a state with certain probability. The agents interact with the environment by taking an action in a state, which returns a reward and a new state based on the probability model described by the dictionary. The environment takes as parameters the reward of the goal state and intermediary states. The agents have no knowledge of the probability table maintained by the environment.

**Model-Free Learning**

The technique used for model-free learning was using q-learning. The agent receives a starting state by the environment, it gets the maxQ action possible from that state and executes the action. The environment return a new state and its reward and the q-table[previous state, action] gets updated. This is repeated until the agent reaches a terminal state or the number of actions taken in the current epoch is greater than accepted. The game is repeated n times, updating the same q-table.