CSE 571 Fall 2020

Homework 1 Solution

Exercise 1.1

Any sequential environment in which rewards may take time to arrive will work, because then we can arrange for the reward to be "over the horizon." Suppose that in any state there are two action choices, a and b, and consider two cases: the agent is in state s at time T or at time T − 1. In state s, action a reaches state s′ with reward 0, while action b reaches state s again with reward 1; in s′ either action gains reward 10. At time T − 1, it's rational to do a in s, with expected total reward 10 before time is up; but at time T , it's rational to do b with total expected reward 1 because the reward of 10 cannot be obtained before time is up.

Students may also provide common-sense examples from real life: investments whose payoff occurs after the end of life, exams where it doesn't make sense to start the high-value question with too little time left to get the answer, and so on. The environment state can include a clock, of course; this doesn't change the gist of the answer—now the action will depend on the clock as well as on the non-clock part of the state—but it does mean that the agent can never be in the same state twice.
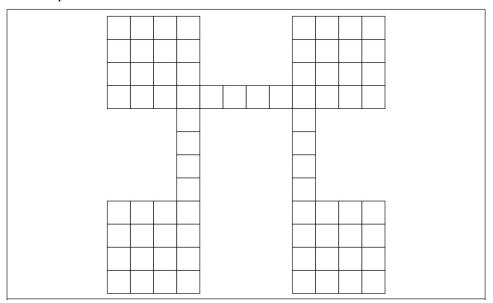
Exercise 1.2

a. False. Perfect rationality refers to the ability to make good decisions given the sensor information received.
b. True. A pure reflex agent ignores previous percepts, so cannot obtain an optimal state estimate in a partially observable environment.
c. True. For example, in an environment with a single state, such that all actions have the same reward, it doesn't matter which action is taken. More generally, any environment that is reward-invariant under permutation of the actions will satisfy this property.
d. True. This is a special case of (c); if it doesn't matter which action you take, selecting randomly is rational.
e. False. Some actions are stupid—and the agent may know this if it has a model of the environment—even if one cannot perceive the environment state.

Exercise 1.3

a. Because the agent does not know the geography and perceives only location and local dirt, and cannot remember what just happened, it will get stuck forever against a wall when it tries to move in a direction that is blocked—that is, unless it randomizes.

b. A randomized simple reflex agent might outperform a deterministic simple reflex agent. This is fairly close to what the Roomba vacuum cleaner does (although the Roomba has a bump sensor and randomizes only when it hits an obstacle). It works reasonably well in nice, compact environments. In maze-like environments or environments with small connecting passages, it can take a very long time to cover all the squares.

c. An example is shown below. It is an environment in which random motion will take a long time to cover all the squares.

Exercise 1.4

a. Simple reflex agent. Any reasonable form of the agent function will be accepted. An example of the agent function: if not possible to head towards goal at current location then follow the contour of the obstacle; if possible to head towards goal at current location then head towards the goal.

b. No, it might be stuck. Any reasonable environment will be accepted.