Homework 3. Partially observable Markov decision problems

Consider once again the problem of the knight that must save the captive princess.[1] After arriving at the castle where the princess is held captive, the knight realizes that the princess may be in any one of two towers (towers A and B). The knight must thus decide which tower to invade. However, he does not know which tower the princess is at. He can try to peer at the towers with a scope, in which case he is able to detect the princess with a probability of 0.9. However, with a 0.1 probability, the knight will see the princess in the wrong tower.

When the knight invades a tower, he either saves the princess or is expelled from those lands. For the purposes of this homework, whichever the outcome of the knight's invasion, we assume that the world resets—i.e., the princess is captured again, placed randomly in one of the two towers, and the knight must again rescue her.

In this homework, you will model the decision of the knight as a partially observable MDP (POMDP).

Exercise 1.

(a) Identify the state space, $\mathcal{X}$, the action space $\mathcal{A}$, and the observation space, $\mathcal{Z}$. You should explicitly model the fact that, when the knight does not peer into the towers with his scope, he sees *nothing*.

(b) Write down the transition probabilities, the observation probabilities and the cost function for this problem. Make sure that the values in your cost function all lie in the interval $[0, 1]$.

(c) Suppose that, at some time step t, the knight believes that the princess is in Tower A with a probability 0.7, decides to peer at the tower, and observes the princess in Tower B. Compute the resulting belief.

[1] Aside from the common theme, there is no other relation between this homework and the previous one, where the knight must navigate the grid.