

Homework 4: MDPs and POMDPs (due before class Nov. 25)

Please read the rules for assignments on the course web page. Contact Lirong (lxia@cs.duke.edu) or Vince (conitzer@cs.duke.edu) with any questions.

1 (40 points). Create your own MDP for something that you think is interesting! You should clearly specify the states, actions, transition probabilities, and rewards. There should be at least 3 states and at least 2 actions. Try to make it a relatively “interesting” MDP, *i.e.*, make it so that there are not too many deterministic transitions, and that it is not immediately obvious what the optimal policy is. The most interesting/entertaining MDP will receive some special recognition and bonus points. Note: every person should submit a different MDP (unless you want to create something *significantly* more interesting than a small example, in which case you may collaborate). You may also want to look at question 3 before doing this question.

2 (40 points). Solve for the optimal policy for your MDP, in at least two different ways; at least one of the solutions must be by computer. For example, you could (1) code up value iteration to solve it, and then (2) also solve it by hand by guessing the optimal policy (or reading it off from the value iteration solution...), computing the values of the states, and verifying that the solution is indeed optimal. (Other options include policy iteration and linear programming.) In both of your solutions, you should derive the optimal policy, as well as the optimal values at each state.

3 (20 points). Now let us assume that the states of your MDP are not directly observable. Turn your MDP into a POMDP by specifying observations and probabilities for these observations (these probabilities can be conditional on only the state, or on the state and the action).