COMPSCI 570	Instructor: Ronald Parr
Homework 7	Due: Thursday, December 1, 2016

1 Value of Information

Prove that the value of perfect information, as defined on the decision theory slides, is invariant under shifts in the value function, i.e., if you add a constant c to the utility of all events, VPI does not change. (This should be intuitively obvious, but you should still show it mathematically and not give a hand-waving argument.)

2 Value of Information

This question is Adapted from Russell & Norvig, problem 16.17:

A used car buyer can decide to carry out various tests with various costs (e.g., kick the tires, take the car to a qualified mechanic) and then, depending on the outcome of the tests, decide which car to buy. We will assume that the buyer is deciding whether to buy car and that there is time to carry out at most one test which can help to figure out the quality of the car. A car can be in good shape (of good quality $Q = q$) or in bad shape (of bad quality $Q = \bar{q}$), and the test might help to indicate what shape the car is in. There are only two outcomes for the test T: pass ($T = pass$) or fail ($T = fail$). The car costs \$1,500, and its market value is \$2,000 if it is in good shape; if not, \$700 in repairs will be needed to make it in good shape. The buyer's initial estimate is that the car has 70% chance of being in good shape.

a) Calculate the expected net gain in utility for the buyer given no test.

b) Suppose $P(T = pass|Q = q) = 0.8$ and $P(T = pass|Q = \bar{q}) = 0.35$. Calculate the optimal decisions given either pass or fall, and the utility of each.

c) Calculate the maximum the buyer should be willing to pay for the test.

d) Suppose the test were perfect so that $P(T = pass|Q = q) = 1.0$ and $P(T = pass|Q = \bar{q}) = 0.0$. How much should the buyer be willing to pay then?

3 MDPs I

Prove that multiplying all rewards in an MDP by a positive constant c does not change the optimal policy for the MDP.

4 MDPs II

Suppose there are two coins. Coin A has probability 0.25 of heads and coin B has probability 0.1 of heads. You are given a chance to play a game with the following rules: At the start of the game, you pick a coin. The selected coin is then flipped until it yields heads. The game then stops. When coin A comes up heads, you get a payoff of 100. When coin B comes up heads you get a payoff of 300. There can be a significant delay between flips, so we'll assume that there is a discount factor of γ applied per time step to future payoffs.

(a) Formulate this problem as an MDP. Indicate the states, actions, and rewards for this MDP. Hints: In some states for your MDP there will be no action choices. You can also view the end of the game as a state with 0 reward and a deterministic transition to itself, i.e., a state which must have value 0. (This is called an *absorbing state*.)

(b) State the optimal policy for this MDP and justify your answer mathematically. Note that the policy may depend upon the value of γ.