EECS 649: PROBLEM SET #9

Reading:

- R&N, 4e: Sections 16.1-5; 17.1 through end of 17.1.2; 17.2
  [you can do PS#9 with this reading and the lectures Making Simple Decisions and Making Complex Decisions]

Total Points: 100

Notes:

- Submitted electronically (via Gradescope)
- This includes some programming.

Problem 9.1 [10 points] *Lottery Tickets*
Tickets to a lottery cost $1. There are two possible prizes: a $10 payoff with probability 1/50, and a $1,000,000 payoff with probability 1/2,000,000. What is the expected monetary value of a lottery ticket? When (if ever) is it rational to buy a ticket? Be precise—show an equation involving utilities. You may assume current wealth of $k and that U(Sk) = 0. You may also assume that U(Sk+10) = 10 × U(Sk+1), but you may not make any assumptions about U(Sk+1,000,000).

Carefully compute the EMV (5 points); say something reasonable, backed up by precision as asked, for all the rest (5 points).

Note: *For this problem, you are asked to compute the EMV and then compare that to the cost of $1, so there is no need to include the cost in the EMV calculation.*

Problem 9.2 [20 points] *Buying a Used Car*
(Adapted from Pearl (1988).) 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 c1, that there is time to carry out at most one test, and that t1 is the test of c1 and costs $50.

A car can be in good shape (quality q+) or bad shape (quality q−), and the tests might help indicate what shape the car is in. Car c1 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 estimate is that c1 has a 70% chance of being in good shape.

a. Draw the decision network that represents this problem **[USE THE ONE GIVEN BELOW]**
b. Calculate the expected net gain from buying c1, given no test.

c. Tests can be described by the probability that the car will pass or fail the test given that the car is in good or bad shape. We have the following information:

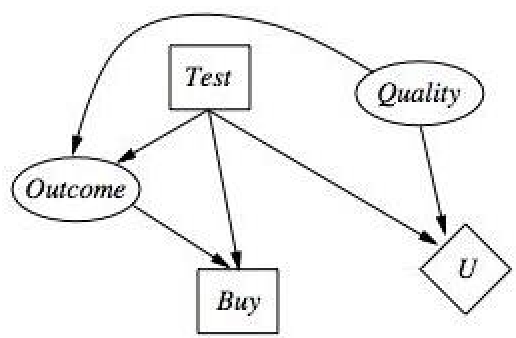
   $P(pass(c1, t1)|q+(c1)) = 0.8$

   $P(pass(c1, t1)|q-(c1)) = 0.35$

   Use Bayes' theorem to calculate the probability that the car will pass (or fail) its test and hence the probability that it is in good (or bad) shape given each possible test outcome.

d. Calculate the optimal decisions given either a pass or a fail, and their expected utilities.

Notes: *Use the decision network below. In part (c), they are asking you to compute P(Pass), P(not Pass), P(q+|Pass), P(q-|Pass), P(q+|not Pass), and P(q-|not Pass).*



Problem 9.3 [20 points] *Sequential Movements*

For the 4x3 world shown in Figure 17.1 of R&N, 4e, compute the probabilities for [*Up, Up, Right*], with points assigned as follows and **starting from the (1, 1) square**:

a. (10 points) For correct probabilities after [*Up, Up*]
b. (10 points) For correct probabilities after [*Up, Up, Right*]

To obtain full credit, enter your computed probabilities from parts (a) and (b) in the top and bottom grids on the linked sheet, respectively. Enter 0 for states not reachable. Make sure your probabilities add to one!

Hints:

- Read the paragraph before Figure 17.1 (bottom of page 562) of R&N, 4e.
- Consider the following, also from R&N: This question helps to bring home the difference between deterministic and stochastic environments. Here, even a short sequence spreads the agent all over the place. The easiest way to answer the question systematically is to draw a tree showing the state reached after each step and the transition probabilities. Then the probability of reaching each leaf is the product of the probabilities along the path, because the transition probabilities are Markovian. If the same state appears at more than one leaf, the probabilities of the leaves are summed because the events corresponding to the two paths are disjoint.

- While the above hints from R&N are correct, they can lead to a **lot** of calculations. You may find it **much, much** quicker to "collapse" the tree after each expansion, so that there is only **one copy** of each node (whose probability is computed using their method) at each level.

The following problem requires some programming.

Problem 9.4 [50 points] *Implement Value Iteration*
Implement the Value-Iteration Algorithm (Figure 17.6 of R&N, 4e) and use it to find the optimal utility values and policy for the 4x3 world shown in Figure 17.1 of R&N, 4e in the case where R=−1 (and gamma=1).

Turn in your **code** and your program's **output**.

Note: For ease of grading, enter your utility values and policy in the top and bottom grids on the linked sheet, respectively. Caption the drawings appropriately and comment on your results.

Notes:

- I have posted Python, C++, and Java programs with code that computes the transition probabilities for the 4x3 gridworld in this linked Python notebook mdp.ipynb
- If you use this code for transition probabilities, do **not** include it in the code you turn in (but do note you used my routine).
- Also, feel free to look at the code in the book's code directory, e.g., its Python MDP code
- As always, please check over whatever code you choose to incorporate in order to make sure it does what is advertised.
- The Value Iteration algorithm also appears in the lecture material.

Note: You will have to use a different ending condition than that in the algorithm in Figure 17.6, because gamma=1 would lead to an infinite loop.

Further, you should also note that the values of the terminal states are given as +1 and −1, and they should retain these values throughout both algorithms. Therefore, you need only loop over the other states in both algorithms.