CSci5512, Fall-2021
ASSIGNMENT 4:
Assigned: 11/6/21 Due: 11/18/21 at 11:55 PM (submit via Canvas, you may scan or take a picture of your paper answers) Please organize your work before submitting.
On all problems you must show work to receive full credit; all answers done individually

Problem 1. (15 points)
There were six “reasonable” assumptions made about preferences. One of them was “monotonicity”, which says:
If A›B, then random variables x and y:
{ Probabilities p>q if and only if: x=[(p,A), (1-p,B)] › y=[(q,A), (1-q,B)] }

Come up with an example (not a proof) to show how you can “exploit” someone (i.e. bad things can happen) if this property is not valid. Do this for both ways of the “if and only if”.

Problem 2. (15 points)
Pretend you were transported back in time before the advent of money and instead everything is done with bartering. After talking to the locals, and proposing a number of trades/exchanges, these are the ones they are willing to accept with you.

The first item listed is what you give, second is what you receive. So if you give someone 10 (or 11, 12, 13...) bushels of wheat to someone, they are willing to will give you one horse in exchange (but they are unwilling to exchange 9 bushels for a horse).

1. 10 bushels of wheat for 1 horse
2. 1 horse for 10 bushels of wheat
3. 1 cow for 1 pig
4. 3 cows for 1 horse
5. 1 horse for 3 goats
6. 1 pig for 1 goat
7. 2 goats for 1 cow
8. 1 goat for 2 bushels of wheat

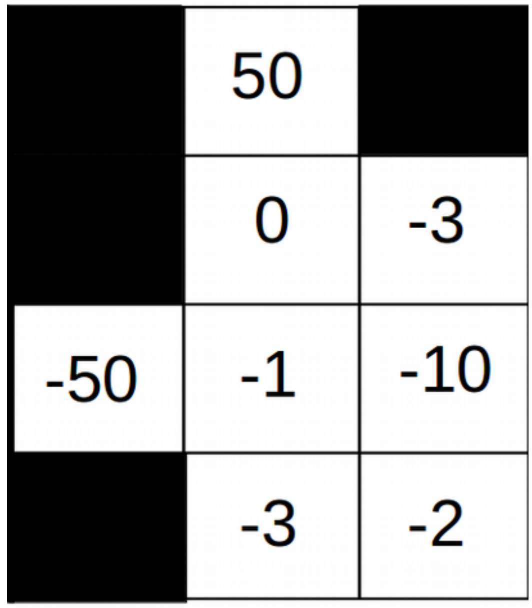
(1) Write these “willing trades” as preferences.

(2) Convert these preferences into a mathematical representation for the utility of the items.

(3) Find valid utility values for each of the items.

(4) Can you do any affine transformation on these utilities? Why or why not?

Problem 3 & 4 use the following Markov Decision Process (MDP) with rewards as shown:



Assume that when moving there is a 70% chance to end up where you want to go and a 15% chance to end up 90 degrees left/right of where you want to go.

So for example, if you intend to go "up": there is a 70% chance you go up, 15% chance you go right and 15% chance you go left

You may assume that when you hit the 50 or -50, that you cannot move anymore and just get that reward then stop the "game".

Problem 3. (25 points)
For all parts of this problem, use the MDP given above and assume $\gamma=0.8$.

(1) Run value iteration until convergence and report the utilities for every state.

(2) If your initial guesses for the utilities are all zero, what is the least amount of iterations to find best policy?

(3) On the iteration you found in part (2), what is largest difference the estimated utility vs. actual utility (i.e. between parts (1) and (2)). How does this compare to the theoretical bound for the utility?

Problem 4. (20 points)
Use policy iteration to solve the MDP shown above. Start by assuming all actions are "Up" and $\gamma=0.8$.

Problem 5. (25 points)
Assume you have the following POMDP with rewards as shown below. Assume your initial guess of where you are in this POMDP is 20% in the top-left, 30% in the top-right and 50% in the bottom-right (also shown below). Assume the movement is the same as in Problems 3 & 4 (70%, 15%, 15%) , but the only actions are moving "left" or "down". There is a boolean evidence variable, e, and $P(e|s)$ is shown for all possible states in the third picture. Use $\gamma=1$.

(1) Find all possible belief states that result from taking two actions and their associated likelihoods of occurring (in the "online algorithm" fashion).

(2) Which sequence of actions is best if you only take two actions?

Rewards:

-1	-4
2	-2

Initial guesses for states:

20%	30%
0%	50%

P(e|s):

0.3	0
0.9	0.2