CS188 Fall 2018 Section 4: Games and MDPs

1 Utilities

1. Consider a utility function of $U(x) = 2x$. What is the utility for each of the following outcomes?
   (a) 3

   (b) L($\frac{2}{3}$, 3; $\frac{1}{3}$, 6)

   (c) -2

   (d) L(0.5, 2; 0.5, L(0.5, 4; 0.5, 6))

2. Consider a utility function of $U(x) = x^2$. What is the utility for each of the following outcomes?
   (a) 3

   (b) L($\frac{2}{3}$, 3; $\frac{1}{3}$, 6)

   (c) -2

   (d) L(0.5, 2; 0.5, L(0.5, 4; 0.5, 6))

3. What is the expected monetary value (EMV) of the lottery $L(\frac{2}{3}, \$3; \frac{1}{3}, \$6)$?

4. For each of the following types of utility function, state how the utility of the lottery $U(L)$ compares to the utility of the amount of money equal to the EMV of the lottery, $U(EMV(L))$. Write $<, >, =$, or ? for can't tell.
   (a) U is an arbitrary function.
   $U(L)$ ___ $U(EMV(L))$

   (b) U is monotonically increasing and its rate of increase is increasing (its second derivative is positive).
   $U(L)$ ___ $U(EMV(L))$

   (c) U is monotonically increasing and linear (its second derivative is zero).
   $U(L)$ ___ $U(EMV(L))$

   (d) U is monotonically increasing and its rate of increase is decreasing (its second derivative is negative).
   $U(L)$ ___ $U(EMV(L))$

2 MDPs: Micro-Blackjack

In micro-blackjack, you repeatedly draw a card (with replacement) that is equally likely to be a 2, 3, or 4. You can either Draw or Stop if the total score of the cards you have drawn is less than 6. If your total score is 6 or higher, the game ends, and you receive a utility of 0. When you Stop, your utility is equal to your total score (up to 5), and the game ends. When you Draw, you receive no utility. There is no discount ($\gamma = 1$). Let's formulate this problem as an MDP with the following states: $0, 2, 3, 4, 5$ and a *Done* state, for when the game ends.

1. What is the transition function and the reward function for this MDP?

2. Fill in the following table of value iteration values for the first 4 iterations.

States	0	2	3	4	5
V_0					
V_1					
V_2					
V_3					
V_4					

3. You should have noticed that value iteration converged above. What is the optimal policy for the MDP?

States	0	2	3	4	5
π^*					

3 (Optional) Minimax and Expectimax

In this problem, you will investigate the relationship between expectimax trees and minimax trees for zero-sum two player games. Imagine you have a game which alternates between player 1 (max) and player 2. The game begins in state s_0, with player 1 to move. Player 1 can either choose a move using minimax search, or expectimax search, where player 2's nodes are chance rather than min nodes.

1. Draw a (small) game tree in which the root node has a larger value if expectimax search is used than if minimax is used, or argue why it is not possible.

2. Draw a (small) game tree in which the root node has a larger value if minimax search is used than if expectimax is used, or argue why it is not possible.

3. Under what assumptions about player 2 should player 1 use minimax search rather than expectimax search to select a move?

4. Under what assumptions about player 2 should player 1 use expectimax search rather than minimax search?

5. Imagine that player 1 wishes to act optimally (rationally), and player 1 knows that player 2 also intends to act optimally. However, player 1 also knows that player 2 (mistakenly) believes that player 1 is moving uniformly at random rather than optimally. Explain how player 1 should use this knowledge to select a move. Your answer should be a precise algorithm involving a game tree search, and should include a sketch of an appropriate game tree with player 1's move at the root. Be clear what type of nodes are at each ply and whose turn each ply represents.