CS188 Fall 2018 Section 4: Games and MDPs

1 Game Trees

1. **Expectiminimax**: Consider the zero-sum game tree below. We have a minimizing player (triangle pointing down), a maximizing player (triangle pointing up), and a chance node (circle) that will select one of the values uniformly at random. Fill in the expectiminimax value of each node.

![Game Tree Diagram]

2. Which nodes can be pruned from the game tree above through alpha-beta pruning? If no nodes can be pruned, explain why not.

3. **Non Zero-sum Games**: Let’s look at a non-zero-sum game. In this formulation, player A’s utility will be represented as the first of the two leaf numbers, and player B’s utility will be represented as the second of the two leaf numbers. Player A plays first (the triangle pointing left), and Player B plays second (the triangles pointing right). Fill in this non-zero game tree assuming each player is acting optimally.

![Non Zero-sum Game Tree Diagram]

4. Which nodes can be pruned from the game tree above through alpha-beta pruning? If no nodes can be pruned, explain why not.
Consider the MDP above, with states represented as nodes and transitions as edges between nodes. The rewards for the transitions are indicated by the numbers on the edges. For example, going from state $B$ to state $A$ gives a reward of 10, but going from state $A$ to itself gives a reward of 0. Some transitions are not allowed, such as from state $A$ to state $B$. Transitions are deterministic (if there is an edge between two states, the agent can choose to go from one to the other and will reach the other state with probability 1).

1. What is the value of each node if $\gamma = 1$?

2. Write down the optimal action at each step if the discount factor is $\gamma = 1$.

3. Let us consider this problem with $\gamma = 0.5$. What is the new value of each node? What are the optimal actions?

4. For each state, does the optimal action depend on $\gamma$? If so, for each state, write an equation that would let you determine the value for $\gamma$ at which the optimal action changes.