CS182 Fall 2021
Final Exam

Name: ______________________________

Harvard ID: ___________________________

Question 1: Comprehension [30 pts]

1. [**5 pts**] The graph-coloring problem is the following: *Given a set C of colors and a graph G with vertices V and edges E, how can we color every vertex such that no two vertices joined by an edge have the same color, if possible?* Formulate this problem as an integer program.

Name: ______________________________

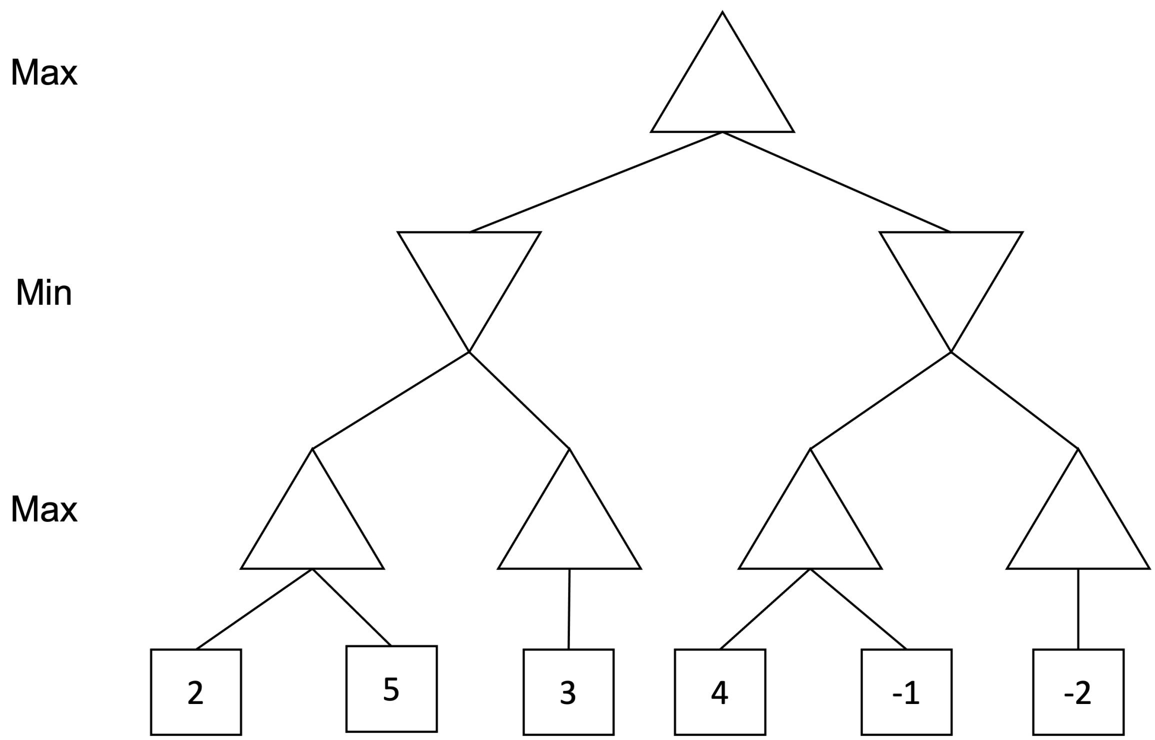
2. **[5 pts]** Given the task of searching a graph for a goal state from a start state, what is one reason why we might prefer to use a BFS over a DFS?

3. **[5 pts]** In a normal game of rock-paper-scissors between Alice and Bob, there is a payoff matrix in which each player wins +1 if they win the round and -1 if they lose, with both player getting 0 if they tie. Now consider a modified unfair game of rock-paper-scissors where Alice earns -2 utility if she plays rock or paper, no matter what Bob does. Find all the Nash equilibria (including mixed) in this modified game.

Name: ____________________________________

4. **[5 pts]** Recall the softmax layer from lecture: $\text{softmax}(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{d} e^{z_j}}$. A common technique is to add a temperature parameter $T \in (0, \infty)$ to the softmax to alter the confidence of the model's predictions: $\text{softmax}(\mathbf{z}, T)_i = \frac{e^{z_i/T}}{\sum_{j=1}^{d} e^{z_j/T}}$. A less confident model will predict closer to the uniform distribution and a more confident model will make more spiked predictions (i.e. higher probability of a 1 or a few classes). How does the parameter T affect model confidence?

5. **[5 pts]** Fill the nodes of the game tree below with the values assigned by the minimax algorithm.



Name: ____________________________________

6. **[5 pts]** The following table gives the group membership G, the true label Y, and the predicted label $\hat{Y}$ of 4 individuals.

G	Y	$\hat{Y}$
0	0	0
0	1	1
1	0	1
1	1	0

Does this classifier satisfy demographic parity? Does it satisfy equalized odds?

Name: ______________________________

Question 2: Decision Trees [20 pts]

1. **[5 pts]** Can a feature ever appear twice in a decision tree (i.e. the decision tree splits on the same feature in two different nodes of the tree) under the Learn-DT algorithm with entropy as the importance function? Explain why or why not. (Specific examples are not needed.)

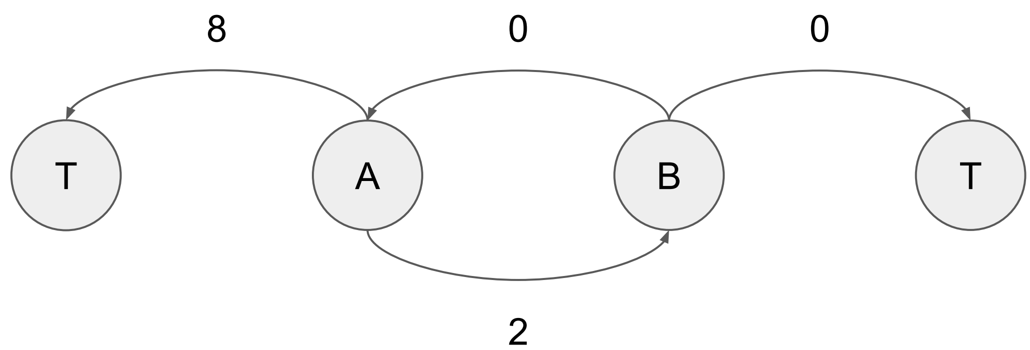
2. **[10 pts]** Suppose that the training data is completely separable on the features, i.e. no two examples have the same feature values but different labels. Prove that the Learn-DT algorithm will return a decision tree with 0% training error, i.e., one that labels the entire training set correctly.

Name: ____________________________

3. **[5 pts]** If the data is completely separable on the features, what are the potential pitfalls of a decision tree with 0% training error?

Name: ________________________________

Question 3: MDP [20 pts]



Consider the above MDP with states {A, B, T} and actions {left, right}. The two T states are terminal and the agent gets no more actions (and no more reward). In states A and B, the agent can take either action, and with probability p the action gets executed and probability *1-p* the other action gets executed. The rewards associated with moving from state to state are denoted by the numbers on the arrows (they are independent of the actions taken). The discount factor γ is 1.

Let π_p^* be the optimal policy as a function of the probability p, let $U^{\pi_p^*}(s)$ be the utility (sum of "discounted" rewards) of π_p^* at state s, and let $Q^{\pi_p^*}(s, a)$ be the utility when taking action a at state s and then following the policy π_p^*.

1. **[5 pts]** If $p = 1$, then what is π_p^*? Select one by circling your answer. Add a brief (1 sentence) explanation.

   (a) $\pi_p^*(A) = \text{left}$, $\pi_p^*(B) = \text{left}$

   (b) $\pi_p^*(A) = \text{left}$, $\pi_p^*(B) = \text{right}$

   (c) $\pi_p^*(A) = \text{right}$, $\pi_p^*(B) = \text{left}$

   (d) $\pi_p^*(A) = \text{right}$, $\pi_p^*(B) = \text{right}$

Name: ______________________________

2. **[10 pts]** If $\pi_p^*(A)$ = left, which of the following statements must be true? Select all that apply by circling your answers. Add a brief (1-3 sentences) explanation for each group of 3.

   (a) $Q^{\pi_p^*}(A, \text{left}) \leq Q^{\pi_p^*}(A, \text{right})$

   (b) $Q^{\pi_p^*}(A, \text{left}) \geq Q^{\pi_p^*}(A, \text{right})$

   (c) $Q^{\pi_p^*}(A, \text{left}) = Q^{\pi_p^*}(A, \text{right})$

   (d) $U^{\pi_p^*}(A) \leq U^{\pi_p^*}(B)$

   (e) $U^{\pi_p^*}(A) \geq U^{\pi_p^*}(B)$

   (f) $U^{\pi_p^*}(A) = U^{\pi_p^*}(B)$

   (g) $U^{\pi_p^*}(A) \leq Q^{\pi_p^*}(A, \text{left})$

   (h) $U^{\pi_p^*}(A) \geq Q^{\pi_p^*}(A, \text{left})$

   (i) $U^{\pi_p^*}(A) = Q^{\pi_p^*}(A, \text{left})$

   (j) $U^{\pi_p^*}(A) \leq Q^{\pi_p^*}(A, \text{right})$

   (k) $U^{\pi_p^*}(A) \geq Q^{\pi_p^*}(A, \text{right})$

   (l) $U^{\pi_p^*}(A) = Q^{\pi_p^*}(A, \text{right})$

Name: ____________________________________

3. **[5 pts]** Assume $p \geq 0.5$: $U^{\pi_p^*}(B) = \alpha * U^{\pi_p^*}(A) + \beta$. Find α and β in terms of p.

   Hint: Think of the optimal policy under this p and of the Bellman equations for A and B.

Name: ______________________________

Question 4: Social Choice [15 pts]

We say that a voting rule f is *unanimous* if, whenever it is given a preference profile (consisting of n rankings, one for each voter) where all voters rank the same alternative x at the top, the voting rule selects x as the winner.

1. [**5 pts**] Show that unanimity does not imply Condorcet consistency by giving an example of a voting rule that is unanimous but not Condorcet consistent.

2. [**10 pts**] Prove that any Condorcet consistent voting rule is unanimous.

   Hint: This can/should be done in 2-3 lines.

Name: ____________________________________

Question 5: Convex Optimization [15 pts]

Recall that in a minimization problem, a point $\mathbf{x} \in \mathbb{R}^n$ is globally optimal if $\mathbf{x} \in \mathcal{F}$ and for all $\mathbf{y} \in \mathcal{F}$, $f(\mathbf{x}) \leq f(\mathbf{y})$; and it is locally optimal if $\mathbf{x} \in \mathcal{F}$ and there exists $R > 0$ such that for all $\mathbf{y} \in \mathcal{F}$ with $\|\mathbf{x} - \mathbf{y}\|_2 \leq R$, $f(\mathbf{x}) \leq f(\mathbf{y})$.

Prove that in a *convex* optimization problem, all locally optimal points are globally optimal.

Hint: Even if you do not remember the proof from class, you should be able to establish it directly using the following guidance and an understanding of the definitions. Suppose $\mathbf{x}$ is locally optimal within radius R, but not globally optimal, then there is $\mathbf{y} \in \mathcal{F}$ such that $f(\mathbf{y}) < f(\mathbf{x})$. Define $\mathbf{z} = (1 - \theta)\mathbf{x} + \theta\mathbf{y}$ for $\theta = \frac{R}{2\|\mathbf{x}-\mathbf{y}\|_2}$. Derive a contradiction to the local optimality of $\mathbf{x}$.

Name: ______________________________

Name: ______________________________