CS221 Final Exam Solutions

Spring 2022

Please read all of the following information before starting the exam:

- This test has 7 questions, each with multiple subparts.
- **You will have 180 minutes to complete and submit the exam.**
- Note that different questions are worth different amounts of points. Budget your time accordingly!
- Keep your answers precise and concise. We may award partial credit so show all your work clearly and in order.
- Don't spend too much time on one problem. Read through all the problems carefully and do the easier ones first.
- If you are unsure about a problem statement when taking the exam, state your assumptions in your answer. We will take all reasonable assumptions into account when grading.
- You are only allowed a 1-page cheatsheet (front and back) to refer to during this exam.
- Being subject to the provisions of the Honor Code means in part that you must observe the rules established for this exam, which are: you may consult only inanimate sources. You may not consult or collaborate with anyone about the questions. Such collaboration is a violation of the Honor Code.
- Good luck!

Problem	Part	Max Score	Score
	a	6	
	b	9	
	c	2	
1	**Total**	17	
	a	6	
	b	11	
2	**Total**	17	
	a	8	
	b	10	
3	**Total**	18	
	a	4	
	b	6	
	c	6	
	e	2	
4	**Total**	18	
	a	10	
5	**Total**	10	
	a	8	
	b	13	
	c	2	
6	**Total**	23	
	a	8	
	b	8	
	c	2	
7	**Total**	18	

0. Honor Code (*0 points*) Please **sign your name** underneath the honor code below. By signing, you agree to abide by the honor code statement as well as the rules stated on Page 1.

Your exam will **not** be graded if this question is not completed.

"I will not consult or collaborate with anyone about the questions. Such collaboration is a violation of the Honor Code."

1. Machine Learning (*17 points*)

a. (*6 points*) **K-Means Clustering** You are given the following six points: $x_1 = [-1, -2], x_2 = [-2, -2], x_3 = [4, -1], x_4 = [3, -1], x_5 = [0, -2], x_6 = [2, -1]$. Answer the following questions.

(i) [4 points] In class we saw the use of the L-2 distance for the computation of the K-means objective. Now, we will instead be using the L-1 distance. Recall that the L-1 distance between two points $x = [x_1, x_2]$ and $y = [y_1, y_2]$ is $|x_1 - y_1| + |x_2 - y_2|$. Furthermore, recall that for a group of points, the point that minimizes the sum of the L-1 distance to all of the points is the element-wise median of the points.

If we select x_1 and x_6 as the initial centroids, write down the simulation of the 2-means clustering process until convergence. Include the new positions of the centroids μ_1 and μ_2, the new cluster assignments z and the cost (i.e. K-means objective) at every step of the simulation. Note: you may not need all the rows.

Iteration	μ_1	μ_2	Cluster Assignments z	Cost
0	[-1,-2]	[2,-1]		
1				
2				
3				
4				

Solution

(a) Iteration 0: $\mu_1 = [-1, -2]$, $\mu_2 = [2, -1]$, $z = [1, 1, 2, 2, 1, 2]$, Cost $= 5$

(b) Iteration 1: $\mu_1 = [-1, -2]$, $\mu_2 = [3, -1]$, $z = [1, 1, 2, 2, 1, 2]$, Cost $= 4$

(c) Converged

(ii) [2 points] If we can arbitrarily set the value of K (i.e. the number of clusters) and randomly select K points (without replacement) from the set as initial centroids for each of the clusters, is it possible to achieve the global minimum of the K-means objective cost function? If yes, what is the minimum value of K that achieves the global minimum **for this problem**? If no, what is the minimum cost achieved?

Solution Yes, we can achieve the global minimum (i.e. cost of 0) using $K = 6$ clusters (e.g., one cluster for each point)..

b. (*9 points*) **Neural Networks and Backpropagation** You are trying to classify images to determine whether or not a dog is present (1) or absent (0) using a 2-layer neural network with sigmoid activations.

(i) [2 points] You think that your model is overfitting because your test error is much higher than your train error. Which of the following methods can you use to reduce overfitting? Select all that apply.

(a) Collect more data and increase the size of your dataset m

(b) Use some form of regularization

(c) Reduce the size (number of parameters) in the network

Solution (a), (b), (c)

(ii) [2 points] Suppose you have access to a large training and test set. You do not know what values to use for the hyperparameters of your network (such as the learning rate). What would be an acceptable strategy regarding the dataset splits that you can use to go about choosing the suitable hyperparameters for your network?

Solution Split the training set into a training set and a validation set. The hyperparameters can be optimized for over the validation set. Other answers that make use of the validation set, such as k-fold cross-validation, are also accepted.

You somehow reduce our image down to three *scalar* features x_1, x_2, x_3, and design a neural network (defined by the following equations) to perform the binary classification:

The architecture of your network (in the form of the feed-forward, loss, and cost equations) is given below:

$$a_1 = w_1 x_1 + w_2 x_2$$

$$a_2 = w_1 x_2 + w_2 x_3$$

$$z = max(a_1, a_2)$$

$$a_3 = w_3 z$$

$$\hat{y} = \sigma(a_3)$$

$$L = y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})$$

w_1, w_2, w_3, w_4 are the trainable parameters of the network, and L denotes the binary cross-entropy loss function applied on the model's prediction $\hat{y}$ and the true label y.

We will now determine some of the equations required to update the w_is via backpropagation. Suppose that we have already computed the following derivatives:

$$\delta_0 = \partial L / \partial \hat{y}$$

$$\delta_1 = \partial \hat{y} / \partial a_3$$

$$\delta_2 = \partial a_3 / \partial z$$

$$\delta_3 = \partial z / \partial a_1$$

$$\delta_3 = \partial z / \partial a_2$$

(iii) [2 points] What is $\partial L / \partial w_3$? You may use the δ_is. in your answer.

Solution

$$\partial L / \partial w_3 = z \delta_1 \delta_0$$

(iv) [3 points] What is $\partial L / \partial w_2$? You may use the δ_is in your answer.

Solution

$$\partial L / \partial w_2 = (x_2 \delta_3 + x_3 \delta_4) \delta_2 \delta_1 \delta_0$$

c. (*2 points*) **Identifying ethical concerns** Consider a deep learning model trained on public GitHub repositories that can convert natural language instructions into code. On testing, it was discovered that it was possible to make the model output sensitive user information like SSH keys, database passwords, etc. Recall the NeurIPS Ethical Guidelines. Which of the following is the model likely in violation of?
Select all that apply. No explanation required.

1. Directly facilitate injury to living beings
2. Contain any personally identifiable information or sensitive personally identifiable information
3. The dataset used has been discredited by the creators
4. Contain information that could be deduced about individuals that they have not consented to share

Solution Answer: 2 and 4.
Rubric: 2 points if both correct options are chosen. 1 point if one of them is chosen. 0 otherwise.

2. Search (*17 points*)

a. (*6 points*) **Warm-up** Consider the following search algorithms as they were presented in this course:

A: DFS

B: BFS

C: DFS with iterative deepening

D: Dynamic Programming

E: UCS

For each of the following problems, you may assume that the state graph is acyclic.

(i) [2 points] Suppose you would like to find the minimum cost path between two states s_1, s_2. You know that the cost function is a positive constant α for every edge in the state-graph. Which algorithms can you use to find the optimal path? Select all that apply or **None** from **A, B, C, D, E**. **No explanation is necessary.**

Solution B, C, D, E

(ii) [2 points] Now you know that the cost is a **non-negative function**. Which algorithms can you use to find the optimal path? Select all that apply or **None** from **A, B, C, D, E**. **No explanation is necessary.**

Solution D, E

(iii) [2 points] You now know nothing about the cost, e.g. the cost is an arbitrary function that is bounded. Which algorithms can you use to find the optimal path? Select all that apply or **None** from **A, B, C, D, E**. **No explanation is necessary.**

Solution D

b. (*11 points*) **Encroaching the enemy territory** Samar is at the start position on a 5×5 grid and wants to reach the goal position by avoiding the mountains along the way (see figure below). To make things harder for himself, **he decides that he will never take the same action more than 3 times in a row.** Samar decides to model this as a search problem where the actions are {left, right, down, up}. The cost for every legal step he takes is $+1$. If he tries to run into a wall or a mountain or repeats an action for a fourth time in a row, he incurs a cost of $+\infty$ and the episode ends. Help Samar model the rest of the search problem and reach the enemy camp with the least cost path!

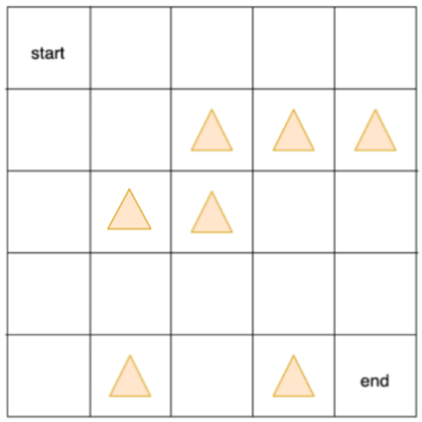


Figure 1: Samar's mission

(i) [3 points] Clearly define a minimal necessary state representation for this problem **as a tuple**.

Solution The state space is
$\{x, y$, previous move, # of times previous move has been taken consecutively$\}$

(ii) [3 points] Clearly define a successor function $T(s'|s, a)$ (where s' is the next state, s is the current state, and a is the action taken) in the form of a piece-wise function. To simplify things, you can ignore the illegal actions (cases where Samar would run into a wall or a mountain or take an action for the fourth consecutive time).
Hint: To simplify notation, assume the actions left, right, up and down can be encoded as tuples $a = (a_1, a_2)$ as $(-1, 0), (1, 0), (0, -1), (0, 1)$ respectively.

Solution The successor function $T(s' \mid s, a)$ is deterministic and is given as follows. To simplify notation, assume the actions left, right, up and down can be encoded as tuples (a_1, a_2) as $(-1, 0), (1, 0), (0, -1), (0, 1)$ respectively. Say the current state is (x, y, A, k)

$$\begin{cases} (x + a_1, y + a_2, a, 1) & a \neq A \\ (x + a_1, y + a_2, A, k + 1) & a = A \end{cases}$$

(iii) [4 points] We wish to run A^* search to find an optimal solution to this problem. To find a heuristic, we define a relaxed search problem in which the mountains are removed and the infinite cost on repeating actions more than 3 times is removed.

In this relaxed search problem, the L1 distance between two points is the cost of the optimal path and is a potential heuristic function for the original problem. **Is this heuristic consistent? If yes, prove why (your proof should be a clear written or mathematical explanation explaining why you think this heuristic satisfies the consistency definition). If no, provide a counter-example. Your answer should fit in the space provided below.**

Solution Yes, this heuristic is consistent.
To show this consider some state S_1 and a neighboring state S_2. Let the L1-distance between S_1 and the goal be X_1 and from S_2 to the goal be X_2. If S_2 is along one of the optimal paths from S_1 to the goal in the relaxed problem, then $X_1 = 1 + X_2$. If not, then $X_1 < 1 + X_2$ because S_2 is in an opposite direction to the goal from S_1. Combining, we have $X_1 \leq X_2 + 1$ which is what we need for a consistent heuristic. Note that the heuristic also evaluates to 0 at the goal state.

(iv) [1 point] Suppose you change the cost such that the cost for every legal step taken by Samar is c, where $0 < c < 1$. **Is the L1 heuristic consistent in this setting? Provide a brief explanation as to why or why not.**

Solution No, it is no longer consistent. The heuristic needs to be $c*$ L1-distance to make it consistent.

3. MDPs and Reinforcement Learning (*18 points*)

a. (*8 points*) **Debugging Q-learning** Sharon wants to train a Q-learning model for her CS221 project but doesn't quite remember the algorithm. Assume an MDP with finite number of states and actions. **Sharon initializes her Q-value table with all 0s.** She remembers that the update rule is as given below but does not remember how to vary ϵ and η. Remember ϵ is from ϵ-greedy and η is the learning rate. The exploratory policy/ data-generating policy is ϵ-greedy as we have seen in the lectures.

$$\hat{Q}_{opt}(s,a) \leftarrow (1-\eta)\hat{Q}_{opt}(s,a) + \eta(r + \gamma \max_{a'} \hat{Q}_{opt}(s',a'))$$

She tries the following variations of the algorithm:

A: initialize $\epsilon = 1$ and never decrease it during training.

B: initialize $\epsilon = 0$ and never increase it during training.

C: ϵ is initialized to 1 is decayed to 0 during training.

For each of parts i-iv, provide a brief (one-sentence) explanation.

(i) [2 points] Sharon finds that **A** converges *very* slowly to Q_{opt}. What could be the cause of this slow convergence?

Solution If we don't decay ϵ, we may not visit good states that appear later in our trajectory frequently enough, which will result in slow convergence.

(ii) [2 points] Sharon finds that **B** does not converge to Q_{opt}. Why could this be?

Solution If we don't have a non-zero ϵ, we will not explore enough states to reach the optimal Q-functions.

(iii) [2 points] In which variations is the data-generating policy changing during the course of training? Select all that apply or **None** from **A, B, C**.

Solution **B, C**: the data-generating policy changes in these three settings. When ϵ changes, clearly the ϵ-greedy policy changes and so the data-generating policy changes. In B as well, even though the ϵ doesn't change, the Q-values are being updated and hence the data-generating policy will change. However, in A, we are running a policy with $\epsilon = 1$ which is the completely random policy which assigns equal weighting to all actions. This doesn't change during the course of training.

(iv) [2 points] **For some MDP**, say we know that $Q^{\pi}(s, a_1) > V^{\pi}(s)$. Which of the following is true? Select all that apply. **Note: this part is unrelated to Sharon's experiments.**

(a) action a_1 is the best action that can be taken in state A

(b) π may be an optimal policy

(c) π is not an optimal policy

(d) None of the above

Solution (c). (a) is wrong because we know that a_1 is better than the "average" action at s but need not be the best. (b) is wrong because if (b) were true then $V^{\pi}(s) = \max_a Q^{\pi}(s, a)$. But the question says that $Q^{\pi}(s, a_1)$ is strictly greater than $V^{\pi}(s)$

b. (*10 points*) **Constantly changing reward functions** Amrita lives in an MDP and wants to find the optimal policy. She uses value-iteration and finds an optimal policy π_{old} with a corresponding value function V_{old}. She wants to determine what happens if she adds a constant scalar C to all rewards in the MDP; to do this, she constructs a new MDP with the new reward function, and computes the optimal policy and value function π_{new} and V_{new}.

(i) [5 points] **Suppose that her MDP is *infinite horizon*, with a discount factor of γ.** Find $\max_{s \in \mathcal{S}}(V_{new}(s) - V_{old}(s))$ where $\mathcal{S}$ is the set of all the states in the MDP. **Show your work. We expect a mathematical derivation or an explanation.** **Hint**: The utility of a infinite horizon trajectory $\tau = (s_1, a_1, r_1), (s_2, a_2, r_2), (s_3, a_3, r_3)...$ is $r_1 + \gamma r_2 + \gamma^2 r_3 + ...$

Solution We can show that for all states,
$V_{old}(s) = r_1 + \gamma r_2 + \dots$
$V_{new}(s) = r_1 + C + \gamma(r_2 + C) + \dots$
$V_{new}(s) = V_{old}(s) + \sum_{i=0}^{\infty} \gamma^i C$
$V_{new}(s) = V_{old}(s) + \frac{C}{1-\gamma}$
$V_{new}(s) - V_{old}(s) = \frac{C}{1-\gamma}$
$\max\{V_{new}(s) - V_{old}(s)\} = \frac{C}{1-\gamma}$

(ii) [2 points] Is π_{new} the same as π_{old}? Explain why or why not. Assume both policies are deterministic and we break ties in the same way in the new and old MDPs.

Solution Yes. The value function in all the states are scaled equally. So the policy will not change.

(iii) [3 points] Would your answers to the previous parts change **if the MDP was of *finite horizon***? Explain why or why not.

Hint: Recall that a finite horizon MDP is one in which trajectories are no longer than some length $T \in \mathbb{Z}^+$. You also do not need to re-derive your answer from part (i).

Solution Yes. We no longer have the infinite sum over $\gamma^i C$. This sum is now finite and depends on how many timesteps are left in the trajectory so the value function at each state may change by a different amount. Therefore, the policies may changes too.

4. Games (*18 points*)

a. (*4 points*) **Zero-sum or not?** For each of the following scenarios, state whether it can be represented as a zero-sum game. Answer "yes" or "no", and **give a brief (1-2 lines) explanation.**

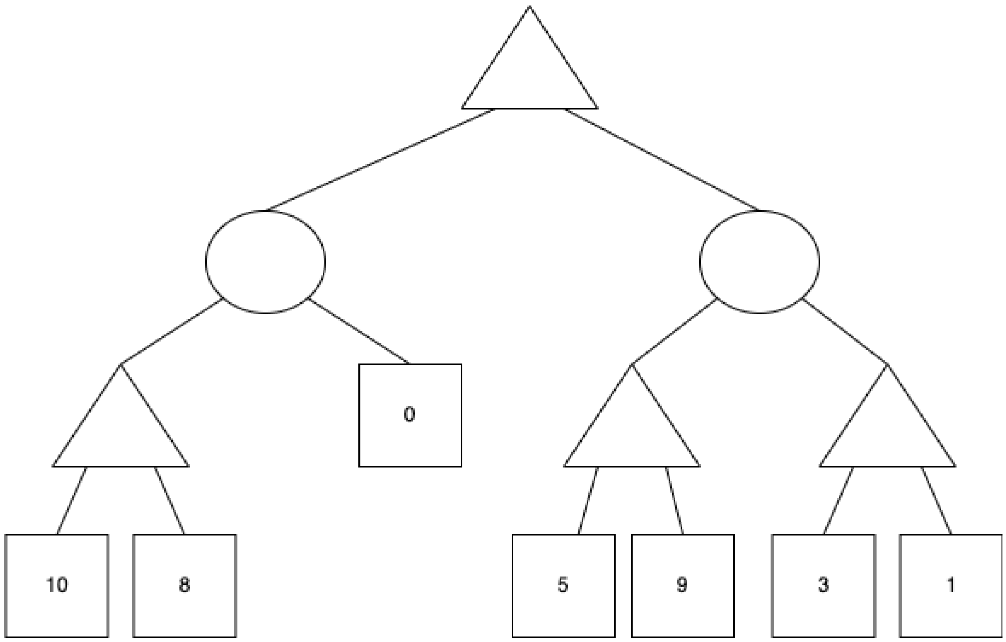
(i) [2 points] **Heads up Holdem**: You are playing a game of poker against an opponent. If you win the current round, you win the money in the pot, but if you lose, your opponent gets the money in the pot. Both players want to maximize their earnings.

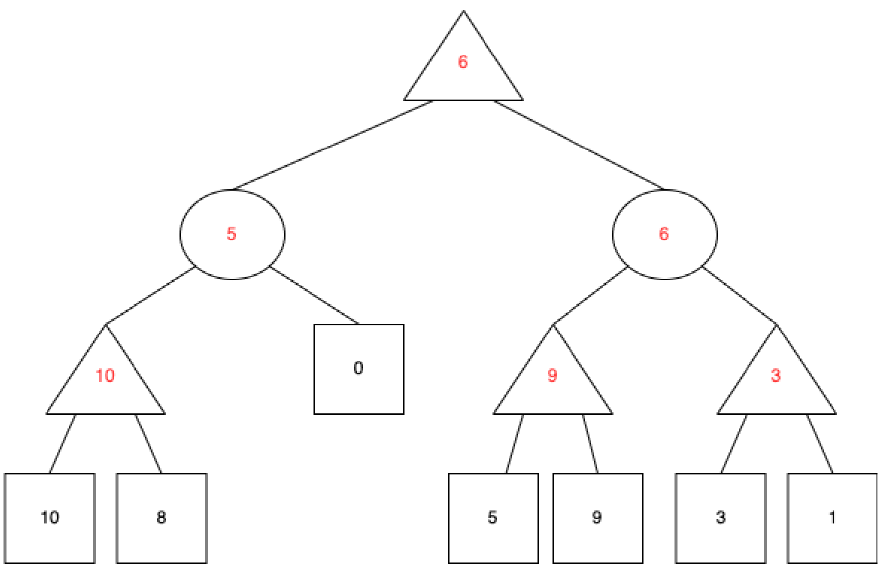
Solution **Yes.** Your earnings are your opponent's losses, and your opponent's earnings are your losses. Hence, the sum of the utilities between you and your opponent will always be 0.

(ii) [2 points] **Autonomous Driving**: Two cars are driving next to one another, and are each trying to merge into the lane ahead of the other (without crashing into one another). Each car has two available actions: merge, or wait. If both cars merge, a crash occurs (resulting in a large negative utility). If one car merges and the other waits, the merged car gets a utility of 10 and the waiting car gets a utility of -10. If both cars wait, they both get a utility of 0.

Solution **No.** Again, this scenario has an aspect of cooperation since both cars must avoid crashing into one another. Specifically, if both cars merge, then the sum of the utilities is not 0, violating the conditions for a 0-sum game.

b. (*6 points*) **Expectimax** Consider the following Expectimax game. **Suppose that the chance nodes (circle) play the left and right actions with 0.5 probability each**. Fill in the values of the maximizer (triangle) and chance (circle) nodes in the figure below.

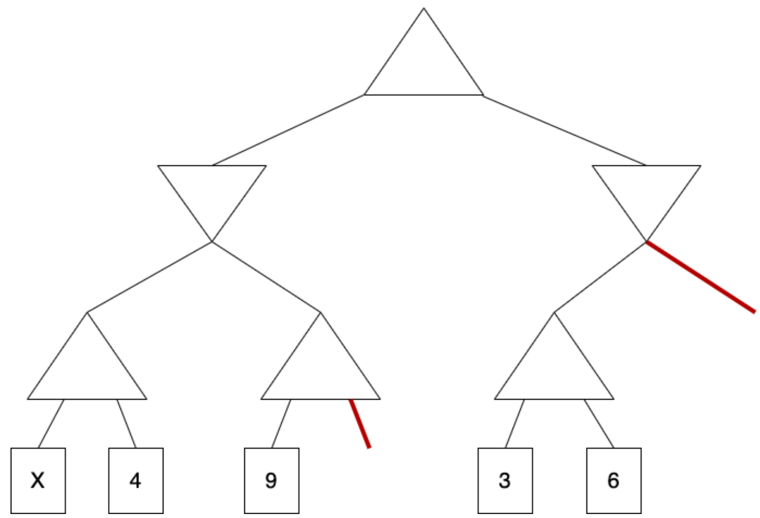




Solution

c. (*6 points*) **Alpha-beta pruning**

Consider the following minimax game tree, in which the top node is a maximizer. Determine the range of values for X which could result in the red edges being pruned by left-to-right alpha beta pruning. Fill your answer in the provided blanks (hint: both the upper and lower bounds are finite values).



______ $\leq X \leq$ ______

Solution $6 \leq X \leq 9$. In order for the left most edge to be pruned, it must be true that $X \leq 9$. In order for the right most edge to be pruned, it must be true that $X \geq 6$.

d. (*2 points*) **Wicked Problems** Recall that problems that have multiple, potentially conflicting objectives, a high degree of uncertainty and risk, and stakeholder disagreement about what would count as a solution to the problem are sometimes called "wicked problems." Another characteristic of wicked problems is that each attempt to solve them has consequences for people at the time it is made, even if another choice is made later. Which of the following are wicked problems?

Select all that apply. No explanation required.

1. Mitigating homelessness in California
2. Developing an agent to play Pac-Man
3. Assigning vaccine priority for Covid-19
4. Creating an optimal chip layout for a circuit board using a simulator

Solution Answer: 1 and 3.
Rubric: 2 points if both correct options are chosen. 1 point if one of them is chosen. 0 otherwise.

5. Constraint Satisfaction Problems (*10 points*)

a. (*10 points*) **Sam and his Sorting CSP** Sam is the new headmaster at Hogwarts and is in charge of splitting the students into houses. However, Hogwarts has expanded and now there are more than just the 4 traditional houses – there are now H houses $X_1, X_2, \ldots X_H$ and N students at Hogwarts. Alice (A), Bob (B), Candice (C), David (D) and Ed (E) are a part of the new cohort of N students at Hogwarts and must be sorted into houses. To add to the problem, Sam needs to adhere to some constraints specified by the school and some students as given below:

1. Every house must have at-least one student.
2. No house can have more than K students.
3. Every student must be assigned exactly one house.
4. Bob and David must be in the same house.
5. Candice must be in the house right after Alice i.e. if Alice is in house h, Candice must be in house $h + 1$. If Alice is in the last house H, then Candice must be in house 1.

This problem seems too complex for the Sorting Hat, so Sam decides to model the problem as a CSP and assign the students to different houses.
Model the problem as a CSP where the variables are the houses and help Sam find the optimal assignment in the CSP.

(i) [2 points] What is the domain of each variable in the CSP? Write it in English or as a set.
Hint: Don't worry about enforcing any constraints here, we will do that when we define factors.

Solution Each variable has a domain that is the power set of all N students.
Alternate solution (not really if you pay attention to the hint): Each variable has a domain that is the set of all subsets of the N students where the size of each subset S is such that $1 \leq S \leq K$.

(ii) [8 points] Express the constraints as factors on the variables. You are allowed to use only one n-ary factor. All your other factors must be unary/ binary.
Hint: There may be several correct answers.

Solution This is one possible solution. One point each for the unary factors and 2 points each for the other factors
One set of unary factors for each variable that counts the number of students in that house and ensures its more than 0 and not greater than K: $f_i(X) = \mathbb{1}[1 \leq \mid X \mid \leq K]$. i runs over all houses.
One set of unary factors over all houses that ensures that B and D are in the same house: $g_i(X) = \mathbb{1}[BD \in X \vee ((B \notin X) \wedge (D \notin X))]$. i runs over all houses.
One set of binary factors over consecutive houses to ensure that C is in the house after A: $h_{i,i+1}(X, Y) = \mathbb{1}[(A \in X \implies C \in Y)]$. i run over all houses. If i is the last house, $i+1$ is the first house. This is a corner case but the factor is still a binary factor.

One n-ary factor over all the houses to make sure that every student is assigned exactly one house: $c(X_1, X_2, \ldots X_H) = \mathbb{1}[(X_1 + X_2 + \ldots X_H).count(S) == 1, \forall S \in \{A, B, C, D, E\}]$

6. Bayes Nets (*23 points*)

a. (*8 points*) **Bayes' Casino**

Suppose you are in a casino and have been observing the behavior a slot machine, which at timestep t has a true state of S_t, who's distribution only depends on S_{t-1}. Your observations have resulted in a set of evidence variables $E_1, ..., E_t$. You know the distribution $P(E_i|S_i)$.

For each of the following questions, **No explanation is required**.

(i) [2 points] **True or false?** This situation can be modeled as a Hidden Markov Model (HMM).

Solution True

(ii) [2 points] Suppose you wish to answer the query $P(S_t|E_1, ..., E_t)$. What type of query does this correspond to?

(A) Smoothing

(B) Filtering

(C) None of the above

Solution (B). This follows by definition.

(iii) [2 points] Suppose you wish to answer the query from the previous question. Which of the following algorithms will guarantee you to answer the query **exactly**?

(A) Gibbs sampling with finite number of iterations.

(B) Forward-backward algorithm

(C) Particle filtering with finite number of iterations.

(D) None of the above

Solution (B). Gibbs sampling and particle filtering are approximate inference algorithms when using finite iterations. The forward-backward algorithm will give an exact solution.

(iv) [2 points] Your friend has a crystal ball and tells you what E_{t+5} will be. Suppose you wish to again find the distribution of S_t given this new piece of information, and your observations up till time t. What type of query does this correspond to?

(A) Smoothing

(B) Filtering

(C) None of the above

Solution (A). This also follows by definition of smoothing.

b. (*13 points*) **Bayesian Sports** You and your friend are fans of the Golden State Warriors and wish to predict the outcome of playoff games. In order to achieve this, you decide to model a Warriors game's outcome with a Bayesian Network with the following variables:

- $H \in \{0, 1\}$: whether the game is being played at home (H = 1) or away (H = 0).
- $C \in \{0, 1\}$: whether Steph Curry scores more than 30 points.
- $R \in \{0, 1\}$: whether refs are helping the Warriors (R = 1) or their opponents (R = 0).
- $O \in \{0, 1\}$: whether the Warrior's offense plays well or not.
- $D \in \{0, 1\}$: whether the Warrior's defense plays well or not.
- $W \in \{0, 1\}$: whether the Warriors win the game or not.

With these variables, you model this situation with the following Bayesian Network:

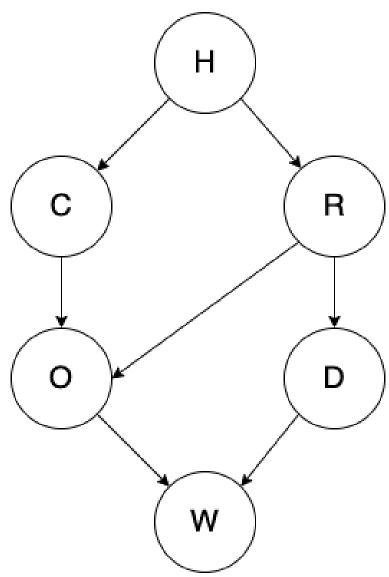


Figure 2: Bayes Net used to model the outcome of a GSW Playoff Game

Assume that offense and defense both positively impact winning
(e.g., $P(W = 1|O = 1) \geq P(W = 1|O = 0)$ and $P(W = 1|D = 1) \geq P(W = 1|D = 0)$).

(i) [2 points] **True or false?** $P(O = 1|W = 1) \geq P(O = 1|W = 1, D = 1)$. Explain your answer briefly (one sentence).

Solution True. This follows from the property of explaining away.

(ii) [3 points] You now want to examine the impact the ref can have by siding with the Warriors. Suppose you know that both the offense and defense play well (e.g., you know $O = 1$ and $D = 1$). Given this evidence, will knowledge of R (whether the ref will help the Warriors) affect the distribution over whether the Warriors win the game? **State yes or no, and explain your answer briefly**.
Hint: this question is asking whether $(W \perp\!\!\!\perp R)|O, D$.

Solution No. W is conditionally independent of R given O and D. This is because in a Bayesian Network, a variable is conditionally independent of non-descendants given its parents.

Suppose now you do not care about whether the Warriors win and only want to analyze the performance of the Warrior's offense and defense. You remove the W variable and define a new Bayes Net:

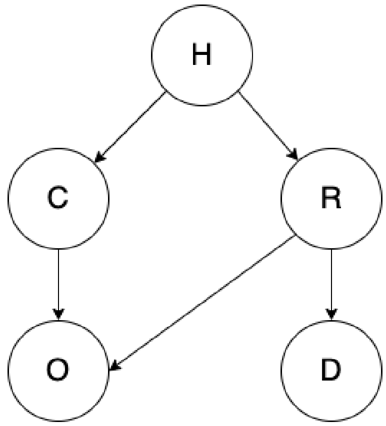


Figure 3: New Bayes Net to model the performance of the offense and defense.

The defense is playing well and you want to determine if the refs have been favoring the Warriors and if the Warriors are playing at home, e.g. you wish to compute $P(R, H|D = 1)$. To do this, you define the following Markov Network. The tables denotes the values of the factors in terms of α, β, θ which are known to you.

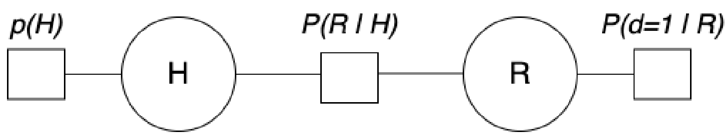


Figure 4: Markov Net for parts (ii) and (iii).

$P(H)$

h	$p(h)$
0	$1-\alpha$
1	α

$P(R \mid H)$

h \ r	0	1
0	β	$1-\beta$
1	$1-\beta$	β

$P(D \mid R)$

r \ d	0	1
0	θ	$1-\theta$
1	$1-\theta$	θ

Table 1: The parameters of the Markov Network in Fig. 4.

You realize that you have no idea what values for α, β, θ are, and decide to collect data to estimate them. However, your data collection assistant is careless and forgets to record values for whether the games were played at home or not. However, you have learned the EM algorithm from CS221 and decide to use that the EM algorithm to estimate the parameters. The table below shows your dataset.

i	h_i	r_i	d_i
1	?	1	1
2	?	1	0
...	...	...	...
100	?	0	1

Table 2: The dataset $\mathcal{D}$ containing 100 sample games.

(ii) [4 points] **E-Step:** In our situation, the E-step will compute the posterior probability $q_i(h) = \mathbb{P}(H = h \mid R = r_i, D = d_i)$ for each $i \in \{1, 2, ..., 100\}, s \in \{0, 1\}$.

Compute $q_1(1)$, expressing your answer in terms of α, β, θ (i.e., your current estimates). Show your work.

Solution

$$
\begin{aligned}
q_1(0) &\propto (1-\alpha)(1-\beta)\theta \\
q_1(1) &\propto \alpha\beta\theta \\
\therefore q_1(1) &= \frac{\alpha\beta}{\alpha\beta + (1-\alpha)(1-\beta)}
\end{aligned}
$$

(iii) [4 points] **M-Step:** Having computed the values of $q_i(h), i \in \{1, 2, ..., 100\}, s \in \{0, 1\}$ from the E-step, what is the value of α after the M-step, expressing your answer in terms of $q_i(h)$? Show your work.

Solution Count and normalize with weights from the E-step:

$$\frac{\alpha}{1-\alpha} = \frac{\sum_{i=1}^{100} q_i(1)}{\sum_{i=1}^{100} q_i(0)}$$
$$\therefore \alpha = \frac{\sum_{i=1}^{100} q_i(1)}{\sum_{i=1}^{100} q_i(1) + \sum_{i=1}^{100} q_i(0)} = \frac{\sum_{i=1}^{100} q_i(1)}{100}$$

c. (*2 points*) **Household Robotics** A robotics company has created a new household robot that can help people with chores at home like washing dishes and cleaning, cooking, etc. However, the company doesn't get approval to test their robot in California. To test their robot, they decide to test it in another country where fire safety laws are not as strict. What ethics concept that you learned in this course would this choice be in violation of, and why?

What we expect: Your answer should be 1-2 sentences and should clearly name the relevant concept, explain why it is in violation in a way that makes it clear that you understand what the concept is.

Solution Answer: This would be considered "ethics dumping" because the company is purposefully choosing to test a product in another country with looser safety regulations than their home country, thereby potentially endangering human subjects involved in tests in the other country.
Rubric: one point for identifying ethics dumping and one point for explanation of why that makes it clear that the student understands the concept.

7. Logic (*18 points*)

a. (*8 points*) **Knowledge Bases** Consider a setting with four propositional symbols: `Healthy`, `Fatigue`, `Flu`, and `Fever`. Let your current knowledge base be:

$$\mathrm{KB} = \{\texttt{Fatigue} \to \neg\texttt{Flu}, \texttt{Fever} \land \texttt{Flu}\}$$

You are asked by the CDC to update the knowledge base with some new formulas, but the CDC would like to know how the new formulas will change the set of models representing the knowledge base. For each of the following questions, **no explanation is required.**

(i) [3 points] First the CDC would like to know if the knowledge base is satisfiable. True or false: The knowledge base is satisfiable.

Solution True

(ii) [2 points] CDC researcher Sharan gives you the following new formula to compare with your knowledge base: $f_1 = \texttt{Fatigue} \lor \texttt{Healthy}$. Which of the following are true? Circle all that apply.

(a) KB entails f_1

(b) KB contradicts f_1

(c) f_1 is consistent

(d) None of the above

Solution (c)

(iii) [3 points] The CDC now gives you a new knowledge base:
KB = {`Headache`,
`Headache` $\land$ `Fatigue`,
`Headache` $\land$ `Fatigue` $\to$ `Fever`,
`Fever` $\land$ `Headache` $\land$ `Fatigue` $\to$ `Flu`
`Flu` $\to$ *False*}. Which of the following statements are true? Circle the correct answer.

(a) Modus ponens is sound on the given KB, but not necessarily complete.

(b) Modus ponens is complete on the given KB, but not necessarily sound.

(c) Modus ponens is sound and complete on the given KB.

(d) Modus ponens is neither sound nor complete on the given KB.

Solution (a)

b. (*8 points*) **Conjunctive Normal Form (CNF)** Modus ponens asserts that if we have two formulas, $A \rightarrow B$ and A in our knowledge base, then we can derive B. Resolution asserts that if we have two formulas, $A \vee B$ and $\neg B \vee C$ in our knowledge base, then we can derive $A \vee C$. If $A \wedge B$ is in the knowledge base, then we can derive both A and B.

Hint: For some of the questions, consider expressing $A \rightarrow B$ in other ways.

(i) [2 points] Your friend Skanda can only read first order logic and needs your help to convert an English sentence into first order logic. Let us consider the following formulas: $Student(x)$ means x is a student, $Classroom(x)$ means x is a classroom, $Safe(x)$ means x is safe, and $Wears(x, mask)$ means x wears a mask. Select the correct conversion(s) of the following English sentence into a first order logic formula. Circle all that apply.

"None of the classrooms are safe if not all students wear masks."

(a) $(\forall s\; Student(s) \wedge \rightarrow Wears(s, mask)) \rightarrow (\forall c\; Classroom(c) \rightarrow \neg Safe(c))$

(b) $(\exists s\; Student(s) \wedge \neg Wears(s, mask)) \rightarrow (\exists c\; Classroom(c) \wedge \neg Safe(c))$

(c) $(\exists s\; Student(s) \rightarrow \neg Wears(s, mask)) \rightarrow (\forall c\; Classroom(c) \wedge \neg Safe(c))$

(d) $(\exists s\; Student(s) \wedge \neg Wears(s, mask)) \rightarrow (\forall c\; Classroom(c) \rightarrow \neg Safe(c))$

Solution (d)

(ii) [2 points] You are given the symbols A, B, C, D. Convert the following formula to conjunctive normal form (CNF): $(A \wedge B) \rightarrow (C \wedge D)$.

Solution $(\neg A \vee \neg B \vee C) \wedge (\neg A \vee B \vee D)$

(iii) [4 points] Suppose the knowledge base contains the following two formulas: KB $= \{(A \vee B) \rightarrow C, A\}$. Convert the KB into CNF and then apply modus ponens to derive C. Show how the KB changes as you apply the derivation rules.

Solution

$$KB = \{(A \lor B) \rightarrow C, A\}$$
$$\Rightarrow \{\neg(A \lor B) \lor C, A\}$$
$$\{(\neg A \land \neg B) \lor C, A\}$$
$$\{(\neg A \lor C) \land (\neg B \lor C), A\}$$

Split the clauses in the KB to get:

$$\Rightarrow \{\neg A \lor C, \neg B \lor C, A\}$$

$$\Rightarrow \{A \rightarrow C, B \rightarrow C, A\}$$

Apply modus ponens on $A \rightarrow C$ and A to derive C.

c. (*2 points*) **Dual-use technologies** Recall that dual-use technologies are technologies that serve two purposes, typically a military and a civilian purpose. In the car-tracking assignment, we saw how a Bayes Net-based tracking system could be used for lethal secondary uses. Give an example of another dual-use AI-based technology related to what we learned in this course (e.g. a technology that includes AI or machine learning). Clearly specify how this technology could be used for both civilian and military purposes.
What we expect: Your answer should be 2-3 sentences and should (a) clearly describe a specific algorithm or AI-based technology related to what we learned in this course and (b) mention one civilian use and one (actual or plausible) military use of this technology.

Solution Rubric: 1 point for naming a plausible and concrete dual-use AI-based technology and half point each for identifying one one civilian and one actual or plausible military purpose.