CS440/ECE448 Exam 2

Jinxiong You

TOTAL POINTS

43 / 50

QUESTION 1

1 q1 8 / 8

✓ **+ 8 pts** *Correct choice with correct reasoning*

+ 5 pts Correct choice but incorrect reasoning

+ 3 pts Attempt to answer the question but incorrect answer and incorrect reasoning

+ 0 pts No attempt

QUESTION 2

2 q2 3 / 7

+ 7 pts Each node in the game tree is filled correctly and a correct minimax move sequence has been provided.

+ 5 pts Either of the following

1) 1 or 2 of the nodes are incorrect and correct move sequence provided

2) All nodes are correct but move sequence not provided/incorrect

+ 4 pts Either of the following

1) 1 or 2 incorrect nodes and move sequence not provided/incorrect

2) 3 or 4 incorrect nodes but correct move sequence provided

✓ **+ 3 pts** *Either of the following*

1) 3 or 4 incorrect nodes and move sequence not provided/incorrect

2) More than 4 incorrect nodes or node values not written but correct move sequence provided

+ 1 pts More than 4 incorrect nodes and move sequence is not provided/incorrect

+ 0 pts No attempt

QUESTION 3

3 q3 7 / 7

✓ **+ 7 pts** *Fully correct answer*

+ 2 pts Correct proposition (R) unified with antecedent of T

+ 1 pts Correct resulting unified proposition

+ 0.5 pts Human(George) used as unified proposition instead of Uses(George, cellphone)

+ 1 pts Correct resulting substitution dictionary

+ 0.5 pts Substitution dictionary is partially correct

+ 1 pts Correct new fact added to the database

+ 0.5 pts Uses(George, cellphone) added as fact instead of Human(George)

+ 2 pts Any reasonable attempt at answering question

+ 1 pts Any attempt at answering question

+ 0 pts No answer provided

QUESTION 4

4 q4 7 / 7

✓ **+ 7 pts** *Reached correct answer with supporting work*

+ 6 pts Applied reasonable methods but

reached wrong answer

+ 4 pts Correct formula but left incomplete

+ 3 pts Correct answer but no supporting work

+ 2 pts Applied irrelevant methods and reached wrong answer

+ 0 pts Reached wrong answer with no supporting work

+ 0 pts Incomplete / did not attempt

QUESTION 5

5 q5 7 / 7

✓ **+ 7 pts** *Reached Correct Answer.*

+ 6 pts Applied Reasonable Methods but Reached Wrong Answer

+ 4 pts Applied Some Reasonable Methods or Concepts but Wrong Answer

+ 2 pts Applied Irrelevant or Incorrect or Limited Methods and Reached Wrong Answer

+ 0 pts Reached Wrong Answer with No Supporting Work

+ 0 pts Reached Correct Answer with No or Wrong Supporting Work

QUESTION 6

6 q6 6 / 7

+ 0 pts No attempt.

+ 2 pts No/incomplete probability sequence shown. Final answer incorrect.

+ 4 pts Incorrect probability sequence shown. Final answer incorrect.

✓ **+ 6 pts** *Correct probability sequence shown. Final answer incorrect.*

+ 5 pts No/incorrect probability sequence shown. Final answer correct.

+ 7 pts Correct probability sequence shown. Final answer correct.

QUESTION 7

7 q7 5 / 7

+ 0 pts No attempt

+ 1 pts No use of expression for context $$c_i = \text{softmax}(q_i@k_t)@v$$ (beyond writing it down), or incorrect/no softmax calculated

Final answer incorrect

+ 3 pts Incorrect use of expression for context $$c_i = \text{softmax}(q_i@k_t)@v$$, but correct softmax

Final answer incorrect

✓ **+ 5 pts** *Correct use of expression for context $$c_i = \text{softmax}(q_i@k_t)@v$$, expression is a single value and shows some form of dot product with $$v$$*

Softmax may be incorrect

Final answer incorrect

+ 7 pts Correct use of expression for context $$c_i = \text{softmax}(q_i@k_t)@v$$

Final answer correct

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

CS440/ECE448 Artificial Intelligence

Exam 2

Spring 2023

April 3, 2023

Your Name: Jinxiong You

Your NetID: you24

Instructions

- Please write your name on the top of every page.
- Have your ID ready; you will need to show it when you turn in your exam.
- This will be a CLOSED BOOK, CLOSED NOTES exam. You are permitted to bring and use only one 8.5x11 page of notes, front and back, handwritten or typed in a font size comparable to handwriting.
- No electronic devices (phones, tablets, calculators, computers etc.) are allowed.
- SHOW YOUR WORK. Correct answers derivation may not receive full credit if you don't show your work.
- Make sure that your answer includes only the variables that it should include. Solve integrals and summations. After that is done, do not further simplify explicit numerical expressions. For example, the answer $x = \frac{1}{1+\exp(-0.1)}$ is MUCH preferred (much easier for us to grade) than the answer $x = 0.524979$.

Possibly Useful Formulas

Consistent Heuristic: $h(p) \leq d(p,r) + h(r)$

Alpha-Beta Max Node: $v = \max(v, \text{child}); \quad \alpha = \max(\alpha, \text{child})$

Alpha-Beta Min Node: $v = \min(v, \text{child}); \quad \beta = \min(\beta, \text{child})$

Variance Network: $\mathscr{L} = \frac{1}{n-1} \sum_{i=1}^{n} \left(f_2(x_i) - (f_1(x_i) - x_i)^2 \right)^2$

Unification: $U = S(P) = S(Q); \; U \Rightarrow \exists x : Q; \; U \Rightarrow \exists x : P$

Bayes Rule: $P(Y = y | X = x) = \frac{P(X = x | Y = y) P(Y = y)}{\sum_{y'} P(X = x | Y = y') P(Y = y')}$

Unnormalized Relevance: $\tilde{R}(f_c, x_d) = \frac{\partial f_c}{\partial x_d} x_d f_c$

Normalized Relevance: $R(f_c, x_d) = \frac{\frac{\partial f_c}{\partial x_d} x_d}{\sum_{d'} \frac{\partial f_c}{\partial x_{d'}} x_{d'}} f_c$

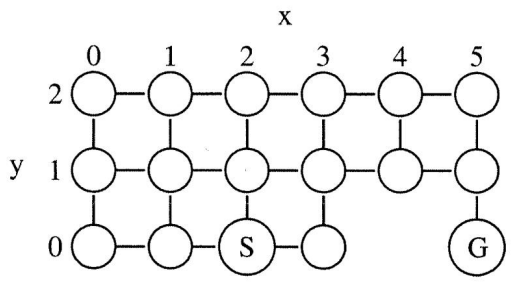
Softmax: $\underset{j}{\text{softmax}}(e) = \frac{\exp(e_j)}{\sum_k \exp(e_k)}$

Softmax Deriv: $\frac{\partial \, \text{softmax}_m(e)}{\partial e_n} = \underset{m}{\text{softmax}}(e) \delta[m-n] - \underset{m}{\text{softmax}}(e) \underset{n}{\text{softmax}}(e), \quad \delta[m-n] = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$

Viterbi: $v_t(j) = \max_i v_{t-1}(i) a_{i,j} b_j(x_t)$

Transformer: $c_i = \text{softmax}(q_i @ k^T) @ v$

Question 1 ***(8 points)***



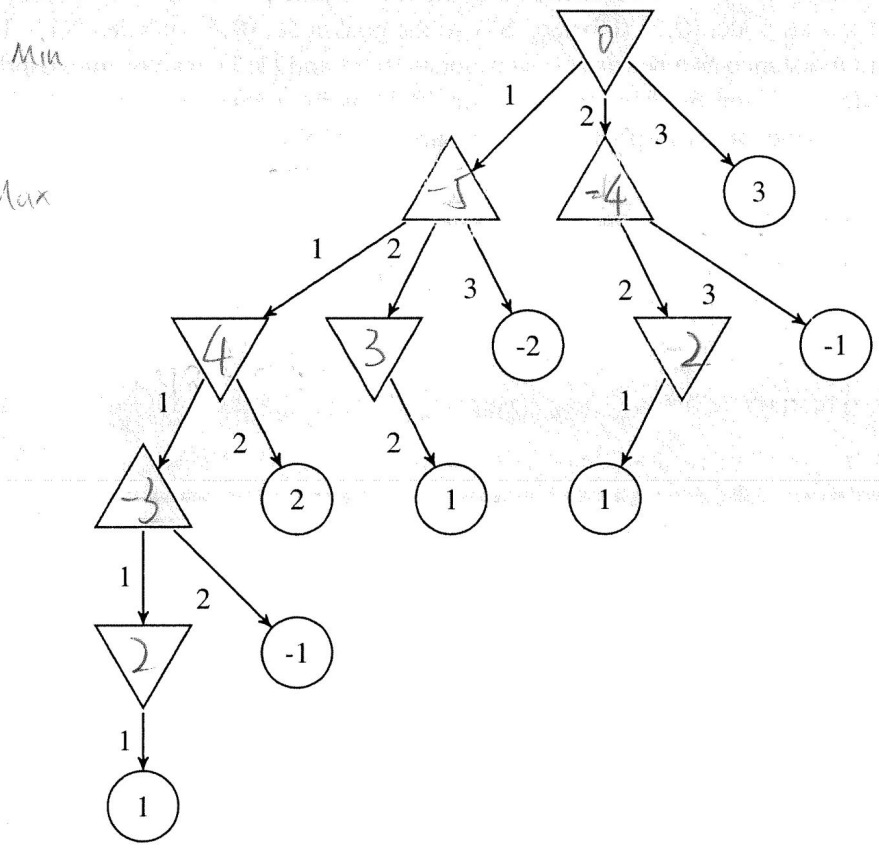
In the maze shown above, nodes are named by their (y,x) coordinates, where y is the row number (starting from the bottom), and x is the column number (starting from the left). A robot is trying to find a path from the start node, $(0,2)$ (labeled "S"), to the goal node, $(0,5)$ (labeled "G"). It uses A* search, with Manhattan distance as a heuristic. After nodes $(0,3)$ and $(1,2)$ have been expanded, there are two copies of nodes $(1,3)$ on the frontier, one with $(0,3)$ as its parent, and one with $(1,2)$ as its parent. Which of these two copies was placed on the frontier first? Why?

$h(0,3)=2$, $h(1,2)=4$

The copy with $(0,3)$ as its parent will be placed on the frontier first, because A* expands nodes in order of increasing cost.

Question 2 ***(7 points)***

Consider a game in which Max and Min each start with three stones. Min plays first. On their turn, each player must discard one, two, or three stones. The number of stones a player discards must be greater than or equal to the number of stones their opponent discarded on the immediate preceding turn (if a player does not have enough stones to satisfy this rule, they must discard all of their remaining stones). When one player loses their last stone, the other player wins a number of points equal to the number of stones they have not yet discarded. The figure below shows the game tree for this game. In each triangle (each Min or Max node), enter a number specifying the value of that node. Specify the minimax value of the game. Indicate a minimax move sequence (there are several different sequences that a pair of optimal players might play; specify one of those sequences).



Min discards 1, Max discards 2, Min discards 2, Max wins 1 point

Question 3 ***(7 points)***

u, v, w, x, y, and z are variables. You are trying to determine whether or not it's possible to perform a step of forward-chaining using the rule $T = \text{Uses}(u, \text{cellphone}) \Rightarrow \text{Human}(u)$. The facts currently available to you in the database $\mathscr{D}$ are:

$$\mathscr{D} = \left\{ \begin{array}{c} P = \text{Eats}(\text{tiger}, \text{cellphone}) \\ Q = \forall v : \text{Uses}(v, \text{landline}) \\ R = \forall w : \text{Uses}(\text{george}, w) \\ S = \exists x : \text{Zambonis}(x, \text{icerink}) \end{array} \right\}$$

Which proposition (P, Q, R, or S) can be unified with the antecedent of T? What is the resulting unified proposition, what is the resulting substitution dictionary, and what new fact is added to the database?

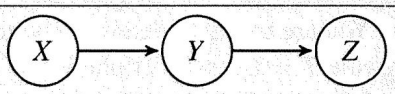
R can be unified with antecedent of T.

The result is Uses(george, cellphone).

Substitution dictionary is {u: george, w: cellphone}

Human(george) is added to the database.

Question 4 ***(7 points)***



X, Y, and Z are random variables, each of which can take the values -1, 0, or 1. Their causal dependencies are shown in the Bayes network above. The parameters of this model are:

	x = -1	x = 0	x = 1
$P(X=x)$	a	b	c
$P(Y=-1\|X=x)$	d	e	f
$P(Y=0\|X=x)$	g	h	i
$P(Y=1\|X=x)$	j	k	l

	y = -1	y = 0	y = 1
$P(Z=-1\|Y=y)$	m	n	o
$P(Z=0\|Y=y)$	p	q	r
$P(Z=1\|Y=y)$	s	t	u

In terms of the parameters a through u, what is $P(X=-1|Z=0)$?

$$P(X=-1|Z=0) = \frac{P(X=-1,Z=0)}{P(Z=0)} = \frac{\sum_y P(X=-1,Y=y,Z=0)}{\sum_y \sum_x P(X=x,Y=y,Z=0)}$$

$$= \frac{adp+agq+ajr}{adp+agq+ajr+bep+bhq+bkr+cfp+ciq+clr}$$

Question 5 ***(7 points)***

Suppose $f_j = \text{softmax}_j(w@h)$, where w is a 2×2 matrix, and h is a 2-vector. In terms of h_m, $w_{n,m}$, and/or f_k for appropriate values of k, m, and n, what is the unnormalized relevance of h_0 to the output f_1?

$$\tilde{R}(f_1, h_0) = \frac{\partial f_1}{\partial h_0} h_0 f_1$$

$$\frac{\partial f_1}{\partial h_0} = \sum_j \frac{\partial \text{softmax}_1(e)}{\partial e_j} \frac{\partial e_j}{\partial h_0}$$

$$= \sum_j \left(\text{softmax}_1(e)\,\delta[1-j] - \text{softmax}_1(e)\,\text{softmax}_j(e) \right) w_{j,0}$$

$$= \sum_j \left(f_1 \delta[1-j] - f_1 f_j \right) w_{j,0}$$

$$\tilde{R}(f_1, h_0) = \sum_j \left(f_1 \delta[1-j] - f_1 f_j \right) w_{j,0} \cdot h_0 f_1$$

where $\delta[m-n] = \begin{cases} 1, & m=n \\ 0, & m \neq n \end{cases}$

Question 6 ***(7 points)***

Slarti is in Paris, attempting to walk home from a pizza restaurant. He is on a sidewalk whose west edge is a steep drop into the river; he wants to make sure he does not fall into the river. Let Y_t be the true distance between Slarti and the cliff edge at time t, measured in meters (m). Suppose you know that $Y_0 = 3$m, for sure. Since Slarti is too full to walk straight, he wobbles as he walks. Since it's foggy, he does not always see clearly: X_t is how far away the cliff edge looks, at time t, which may or may not be equal to the true distance Y_t. The transition probabilities and observation probabilities are

$$P(Y_t = k|Y_{t-1} = j) = \begin{cases} \frac{1}{4} & k = j-1 \\ \frac{1}{2} & k = j \\ \frac{1}{4} & k = j+1 \end{cases}, \quad P(X_t = k|Y_t = j) = \begin{cases} \frac{1}{3} & k = j-1 \\ \frac{1}{3} & k = j \\ \frac{1}{3} & k = j+1 \end{cases}$$

What is $P(Y_2 = 2, X_2 = 2)$?

$$P(Y_2=2, X_2=2) = \sum_{X_1, Y_1} P(Y_1) P(X_1|Y_1) P(Y_2=2|Y_1) P(X_2=2|Y_2=2)$$

$$P(Y_1=\quad) = \sum_{X_0 Y_0 X_1} P(Y_0) P(X_0|Y_0) P(Y_1|Y_0) \cdot P(X_1|Y_1=\quad)$$

$$P(Y_1=2) = \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{3}$$

Question 7 ***(7 points)***

Suppose the the input to a transformer is the sequence of scalar values $v_t = \left(\frac{t}{1000}\right)$, where $0 \le t \le 999$. You are trying to find the context, c_i, for a query, q_i, whose inner product with the keys is

$$q_i @ k_t = \begin{cases} 0 & t \in \{500, 501, 502\} \\ -\infty & \text{otherwise} \end{cases}$$

Find the numerical value of c_i.

$c_i = \text{softmax}\, q_i @ k^T @ v$.

$$\text{softmax}_i(q_i @ k) = \frac{\exp(q_i @ k)}{\sum_k \exp(q_i @ k_k)}$$

$$= \begin{cases} \frac{\exp(0)}{3\exp(0) + 997\exp(-\infty)}, & t \in \{500, 501, 502\} \\ \frac{\exp(-\infty)}{3\exp(0) + 997\exp(-\infty)}, & \text{otherwise.} \end{cases}$$

$$= \begin{cases} \frac{1}{3}, & t \in \{500, 501, 502\} \\ 0, & \text{otherwise.} \end{cases}$$

$$c_i = \frac{1}{3}\cos\left(\frac{500}{1000}\right) + \frac{1}{3}\cos\left(\frac{501}{1000}\right) + \frac{1}{3}\cos\left(\frac{502}{1000}\right)$$