UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
CS440/ECE448 Artificial Intelligence
Practice Exam 2
Spring 2023

April 3, 2023

Your Name: ______________________________

Your NetID: ______________________________

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)***

A robot crane is trying to build a building. Suppose that the state variable is $s = \{r_1(s), \ldots, r_n(s)\}$, where $r_i(s) = (x_i(s), y_i(s), z_i(s))$ are the lattitudinal, longitudinal, and vertical positions of the i^{th} brick during the s^{th} state of partial construction. The goal of the search is to find a way to put the n available bricks into n desired positions. Order doesn't matter: it doesn't matter which particular brick ends up in each of the n desired target positions. Suppose that the cost of moving the i^{th} brick into the j^{th} target position, $r_j(g) = (x_j(g), y_j(g), z_j(g))$, is

$$\|r_i(s) - r_j(g)\| = \sqrt{(x_i(s) - x_j(g))^2 + (y_i(s) - y_j(g))^2 + (z_i(s) - z_j(g))^2}$$

Prove that the following heuristic is admissible for this problem:

$$h(s) = \sum_{i=1}^{n} \min_{j=1}^{n} \|r_i(s) - r_j(g)\|$$

Solution: Each brick needs to move into one of the goal positions. Let $j(i)$ be the position the i^{th} brick needs to move to. The total cost of moving all bricks from node s to their goal positions is

$$d(s,g) = \sum_{i=1}^{n} \|r_i(s) - r_{j(i)}(g)\|.$$

For each brick,

$$\min_{j=1}^{n} \|r_i(s) - r_j(g)\| \le \|r_i(s) - r_{j(i)}(g)\|,$$

therefore

$$h(s) \le d(s,g)$$

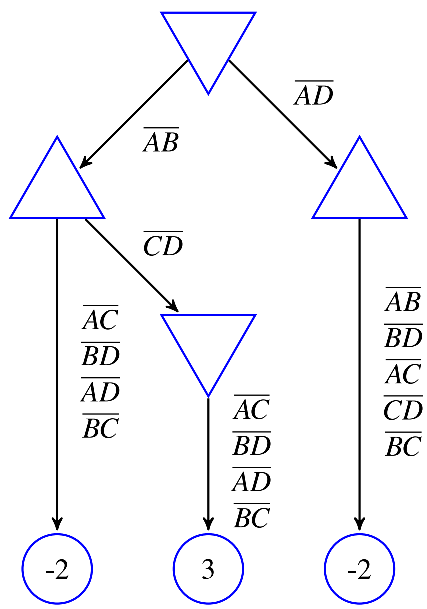
Question 2 ***(7 points)***

C D

A B

Consider a game in which Min plays first, and on the first move, Min must draw either the line segment $\overline{AB}$ or the line segment $\overline{AD}$. Thereafter players take turns; on each turn after the first one, the person playing must draw one line segment connecting any two of the four corners A, B, C, or D. The game ends when one of the players draws a line segment that touches or crosses any previously-drawn line segment; at that point, the other player wins a number of points equal to the total number of line segments that have been drawn. Draw a complete game tree, indicating each move by the letters of the endpoints of the line segment drawn by that player. If many different moves result in the same subtree, draw just one edge that lists all of those moves. Indicate a minimax sequence of moves, and specify the minimax value of the game.

Solution:



The minimax value is -2. The minimax sequence must begin with Min playing $\overline{AD}$, then any move that Max plays will result in the minimax game value.

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

Two propositions, P and Q, can be unified to create the proposition $U = \text{Flies}(\text{mary}, \text{hotairballoons})$. The substitution dictionary creating this unification is $S = \{u : \text{mary}, v : \text{hotairballoons}\}$. Is this information sufficient to specify the two propositions P and Q? Prove your answer.

Solution: No. The propositions $P = \text{Flies}(\text{mary}, \text{hotairballoons})$ and $Q = \text{Flies}(u, v)$ would produce this result. The propositions $P = \text{Flies}(u, \text{hotairballoons})$ and $Q = \text{Flies}(\text{mary}, v)$ would also produce this result.

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

Suppose your burglar alarm has four modes: quiet ($A=0$), warning ($A=1$), alarm ($A=2$), and klaxon ($A=3$). Suppose the probability of an earthquake is $P(E=T)=0.01$, and the probability of a burglary is $P(B=T)=0.001$. Your neighbor, Jack, sends you a text message (event $J=T$) with a probability that depends only on the status of your alarm. Conditional probabilities for the variables A and J are given as the variables a through u in the following table:

	$a=0$	$a=1$	$a=2$	$a=3$
$P(A=a\|B=F,E=F)$	b	c	d	e
$P(A=a\|B=F,E=T)$	f	g	h	i
$P(A=a\|B=T,E=F)$	j	k	l	m
$P(A=a\|B=T,E=T)$	n	o	p	q
$P(J=T\|A=a)$	r	s	t	u

In terms of the parameters a through u, what is $P(B=T|A=3)$?

Solution:

$$\begin{aligned}
P(B=T|A=3) &= \frac{P(B=T,A=3)}{P(A=3)} \\
&= \frac{\sum_{e'=F}^{T} P(B=T,E=e',A=3)}{\sum_{b'=F}^{T} P(B=b',E=e',A=3)} \\
&= \frac{(0.001)(0.01q+0.99m)}{(0.001)(0.01q+0.99m)+(0.999)(0.01i+0.99e)}
\end{aligned}$$

Question 5 *(7 points)*

Consider a neural network with three input nodes, $x = [x_0, x_1, x_2]$, three hidden nodes, $h = [h_0, h_1, h_2]$, and three output nodes, $f = [f_0, f_1, f_2]$, related by

$$h = \text{ReLU}(w_0 @ x)$$
$$f = \text{softmax}(w_1 @ h)$$

where

$$w_0 = \begin{bmatrix} 0.8 & 0.3 & 0.5 \\ -0.4 & -0.9 & -0.5 \\ -0.2 & -0.8 & 0.7 \end{bmatrix}, \quad w_1 = \begin{bmatrix} 0.4 & -0.6 & 0.6 \\ -0.7 & 0.2 & -0.3 \\ 0.9 & -0.1 & 0.1 \end{bmatrix}$$

Suppose $x = [1, 0, -1]$. It can be computed that, in this case, $f = [0.33, 0.26, 0.41]$. What is the unnormalized relevance of the hidden node h_0 to the output f_2?

Solution: Unnormalized relevance is

$$\tilde{R}(f_2, h_0) = \frac{\partial f_2}{\partial h_0} h_0 f_2$$

We already know that $f_2 = 0.41$. We need to find h_0 and $\frac{\partial f_2}{\partial h_0}$.

$$h = \begin{bmatrix} h_0 \\ h_1 \\ h_2 \end{bmatrix} = \text{ReLU}\left(\begin{bmatrix} 0.8 & 0.3 & 0.5 \\ -0.4 & -0.9 & -0.5 \\ -0.2 & -0.8 & 0.7 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right)$$

Keeping just the first row of that equation, we have

$$h_0 = \max(0, 0.8 - 0.5) = 0.3$$

Now for the derivative. We have that

$$\begin{aligned}
\frac{\partial f_2}{\partial h_0} &= \sum_j \frac{\partial\, \text{softmax}_2(e)}{\partial e_j} \frac{\partial e_j}{\partial h_0} \\
&= \sum_j \left(\underset{2}{\text{softmax}}(e) \delta[2-j] - \underset{2}{\text{softmax}}(e) \underset{j}{\text{softmax}}(e) \right) w_{1,j,0} \\
&= \sum_j (f_2 \delta[2-j] - f_2 f_j) w_{1,j,0}
\end{aligned}$$

Since the exam instructions tell us our answer can't include summations, we need to simplify that as

$$\begin{aligned}
\frac{\partial f_2}{\partial h_0} &= \left(f_2 - f_2^2\right) w_{1,2,0} - f_2 f_1 w_{1,1,0} - f_2 f_0 w_{1,0,0} \\
&= \left(0.41 - (0.41)^2\right)(0.9) - (0.41)(0.26)(-0.7) - (0.41)(0.33)(04)
\end{aligned}$$

Putting it all together,

$$\tilde{R}(f_2, h_0) = \left(\left(0.41 - (0.41)^2\right)(0.9) - (0.41)(0.26)(-0.7) - (0.41)(0.33)(04)\right)(0.3)(0.41)$$

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

You are a martial arts master, practicing blindfolded. Your opponent starts out either in front of you, to your left, behind you, or to your right. With each step, they either stay where they are, move 90° clockwise, or move 90° counter-clockwise, with equal probability. Each step (even if they don't move) makes a sound, but because of the echos, you only hear the correct direction of the sound with probability p; with probability $1-p$, you hear the sound from one of the three incorrect directions (each with equal probability). Suppose you hear the sound to your left, and then behind you ($X_0 = L, X_1 = B$). What is the probability that these observations occurred, and that your opponent is now behind you? That is, if your opponent's actual position is Y_t and their apparent position is X_t, what is $P(Y_1 = B, X_0 = L, X_1 = B)$?

Solution:

$$\begin{aligned}
P(Y_1 = B, X_0 = L, X_1 = B) &= \sum_{y_0} P(Y_0 = y_0, X_0 = L, Y_1 = B, X_1 = B) \\
&= \sum_{y_0} \pi_{y_0} b_{y_0}(L) a_{y_0,B} b_B(B) \\
&\quad P(X_0 = L, Y_0 = y_0) \cdot P(X_1 = B, Y_1 = B \mid Y_0 = y_0) \\
&= \left(\frac{1}{4}\right) p \left(\frac{1}{3}\right) p + \left(\frac{1}{4}\right)\left(\frac{1-p}{3}\right)\left(\frac{1}{3}\right) p + \left(\frac{1}{4}\right)\left(\frac{1-p}{3}\right)\left(\frac{1}{3}\right) p,
\end{aligned}$$

where the three possibilities correspond to $Y_0 = L$, $Y_0 = B$, and $Y_0 = R$; from the position $Y_0 = F$, it is not possible to move to $Y_1 = B$.

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

Suppose the the input to a transformer is the sequence of scalar values $v_t = \cos\left(\frac{t}{1000}\right)$, 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

$$q_i@k_t = \begin{cases} 0 & t \in \{250, 251, 252\} \\ -\infty & \text{otherwise} \end{cases}$$

Find the numerical value of c_i.

Solution: The transformer computes

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

The t^{th} element of the softmax output is

$$\begin{aligned} \underset{t}{\text{softmax}}(q_i@k) &= \frac{\exp(q_i@k_t)}{\sum_{t'}\exp(q_i@k_{t'})} \\ &= \begin{cases} \frac{\exp(0)}{3\exp(0)+997\exp(-\infty)} & t \in \{250, 251, 252\} \\ \frac{\exp(-\infty)}{3\exp(0)+997\exp(-\infty)} & \text{otherwise} \end{cases} \\ &= \begin{cases} \frac{1}{3} & t \in \{250, 251, 252\} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

The context vector is therefore

$$c_i = \frac{1}{3}\cos\left(\frac{250}{1000}\right) + \frac{1}{3}\cos\left(\frac{251}{1000}\right) + \frac{1}{3}\cos\left(\frac{252}{1000}\right)$$

This page is scratch paper