University of Illinois at Urbana-Champaign
CS440/ECE448 Artificial Intelligence
Exam 3
Spring 2026

April 6, 2026

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.
- No electronic devices (phones, tablets, calculators, computers etc.) are allowed.
- Make sure that your answer includes only the variables that it should include, but DO NOT simplify explicit numerical expressions, including expressions involving standard functions (exp, log, sin, cos) and expressions involving the sum, product, maximum, or argmax of a complete list of alternatives. For example, the answer $x = \max\left(\sin(0.3), \frac{1}{1+\exp(-0.1)}\right)$ is MUCH preferred (much easier for us to grade) than the answer $x = 0.524979$.

Possibly Useful Formulas

$$\sigma(z) = \frac{1}{1+e^{-z}}, \quad \sigma'(z) = \sigma(z)\,(1-\sigma(z))$$

$$f = g\left(\mathbf{w}^T\mathbf{x}\right), \quad \frac{\partial f}{\partial \mathbf{w}} = g'\left(\mathbf{w}^T\mathbf{x}\right)\mathbf{x}$$

$$\text{softmax}(\mathbf{z}) = \frac{\exp(\mathbf{z})}{\sum_i \exp(z_i)}, \quad \frac{\partial\, \text{softmax}_k(\mathbf{z})}{\partial \mathbf{z}} = \text{softmax}_k(\mathbf{z})\,(\mathbb{1}_k - \text{softmax}(\mathbf{z}))$$

$$f[k] = \sum_i w[i]x[k-i], \quad \frac{\partial f[k]}{\partial w[i]} = x[k-i]$$

$$\mathbf{C} = \text{softmax}\left(\frac{\mathbf{QK}^T}{\sqrt{d}}\right)\mathbf{V}$$

$$S : \{\mathscr{V}_P, \mathscr{V}_Q\} \rightarrow \{\mathscr{V}_Q, \mathscr{C}\} \text{ s.t. } S(P) = S(Q) = U$$

$$\frac{x'}{x} = \frac{y'}{y} = \frac{-f}{z}$$

$$R_{c,d} = \frac{\frac{\partial z_c}{\partial x_d} x_d}{\sum_{d'} \frac{\partial z_c}{\partial x_{d'}} x_{d'}} R_c$$

$$\mathscr{C}_{\text{obs}} = \{\mathbf{c} : \exists \mathbf{b} : \phi(\mathbf{b}, \mathbf{c}) \in \mathscr{W}_{\text{obs}}\}$$

Admissible: $\hat{h}(n) \leq h(n)$

Consistent: $\hat{h}(n) - \hat{h}(m) \leq d(n,m)$

Bellman: $u(s) = r(s) + \gamma \max_a \sum_{s'} P(s'|s,a)u(s')$

Policy Eval: $u_i(s) = r(s) + \gamma \sum_{s'} P(s'|s,\pi_i(s))u_i(s')$

Policy Update: $\pi_{i+1}(s) = \arg\max_a \sum_{s'} P(s'|s,a)u_i(s')$

Value Iteration: $u_{i+1}(s) = r(s) + \gamma \max_a \sum_{s'} P(s'|s,a)u_i(s')$

Model-based: $P(s_{t+1}|s_t,a_t) = \frac{N(s_t,a_t,s_{t+1}) + k}{\sum_{s'}(N(s_t,a_t,s') + k)}$

TD-Learning: $q(s_t,a_t) \leftarrow (1-\eta)q(s_t,a_t) + \eta\left(r_t + \gamma \max_a q_t(s_{t+1},a)\right)$

SARSA: $q(s_t,a_t) \leftarrow (1-\eta)q(s_t,a_t) + \eta\,(r_t + \gamma q_t(s_{t+1},a_{t+1}))$

Actor-Critic: $\mathscr{L}_{\text{actor}} = -\sum_a \pi_a(s)q(s,a)$

REINFORCE: $\Delta\mathbf{W} = \eta(r-\mu)\sum_t \frac{\partial \log \pi_{a_t}(s_t)}{\partial \mathbf{W}}$

Question 1 ***(28 points)***

Consider a neural network with the following architecture. The input is a scalar x, and there is a weight vector $\mathbf{w} = [w_1, \ldots, w_n]^T$ and a scalar weight u. The scalar outputs f, g and h are defined as

$$f = \mathbf{w}^T \mathbf{x},$$
$$g = \cos(f),$$
$$h = ug$$

where $\cos(\cdot)$ denotes cosine.

(a) (14 points) Suppose that you know $\frac{\partial \mathscr{L}}{\partial h}$. In terms of $\frac{\partial \mathscr{L}}{\partial h}$ and any of the variables f, g, h, and/or $\mathbf{w}$, find $\frac{\partial \mathscr{L}}{\partial g}$.

Solution:

$$\frac{\partial \mathscr{L}}{\partial g} = u \frac{\partial \mathscr{L}}{\partial h}$$

(b) (14 points) Suppose that you know $\frac{\partial \mathscr{L}}{\partial g}$. In terms of $\frac{\partial \mathscr{L}}{\partial g}$, and in terms of f, g, h, and/or $\mathbf{w}$, find $\frac{\partial \mathscr{L}}{\partial \mathbf{x}}$.

Solution: The equation suggests that $\mathbf{x}$ is a vector, in which case the answer is

$$\frac{\partial \mathscr{L}}{\partial \mathbf{x}} = \frac{\partial \mathscr{L}}{\partial g} \frac{\partial g}{\partial f} \frac{\partial f}{\partial \mathbf{x}} = \mathbf{w} \sin(f) \frac{\partial \mathscr{L}}{\partial g}$$

The text of the problem says that x is scalar, which would make $\mathbf{f}$ and $\mathbf{g}$ both vectors, in which case the answer is

$$\frac{\partial \mathscr{L}}{x} = \sum_j \frac{\partial \mathscr{L}}{\partial g_j} \frac{\partial g_j}{\partial f_j} \frac{\partial f_j}{\partial x} = \sum_j w_j \sin(f_j) \frac{\partial \mathscr{L}}{\partial g_j}$$

Question 2 ***(15 points)***

Consider a one-dimensional convolutional neural network with input image $x[i,j]$, weights $w[i]$, and output image $f[i,j]$ related as:

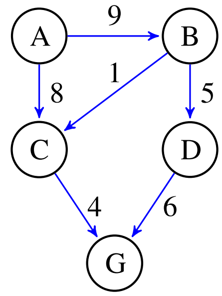
$$f[k,l] = \sum_i x[i,l]w[k-i]$$

In terms of the elements of x, w, and/or f, what is $\frac{\partial f[k,l]}{\partial x[i,j]}$?

Solution:

$$\frac{\partial f[k,l]}{\partial x[i,j]} = \begin{cases} w[k-i] & l = j \\ 0 & \text{otherwise} \end{cases}$$

Question 3 ***(28 points)***



(a) (14 points) The search graph above starts at node A, and ends at node G. Create a table showing, in the first column, the node that is expanded by depth-first search at each step of the search process (starting with A, ending with G), and in the second column, the set of nodes that are in the frontier after the node in the first column has been expanded. Optionally, you may list a priority next to each node in the frontier if you wish. Note that, since DFS is not complete, you might need to put a node on the frontier many times with many different priorities. Ties are broken in alphabetical order.

Solution:

Expand	Frontier
A	B, C
B	C, C, D
C	C, D, G
G	C, D

(b) (14 points) Suppose you're given the following partial heuristic for an A* search: $\hat{h}(C) = 2, \hat{h}(D) = 2, \hat{h}(G) = 0$. You must now choose the heuristics $\hat{h}(A)$ and $\hat{h}(B)$. What is the definition of an **admissible** heuristic? What are the largest values of $\hat{h}(A)$ and $\hat{h}(B)$ that would result in an admissible heuristic?

Solution: An admissible heuristic satisfies $\hat{h}(s) \leq h(s)$, where $h(s)$ is true distance to goal. This is satisfied by $\hat{h}(A) = 12, \hat{h}(B) = 5$.

Question 4 ***(30 points)***

Consider an MDP with two states, $s \in \{0,1\}$, and two actions, $a \in \{0,1\}$. The states have rewards $r(0) = 3$ and $r(1) = 7$, and the transition probabilities are:

s,a	$P(s'=0\|s,a)$	$P(s'=1\|s,a)$
0,0	0.8	0.2
0,1	0.3	0.7
1,0	0.4	0.6
1,1	0.1	0.9

(a) (15 points) Suppose you start value iteration with initial utility estimates of of $u_1(0) = 3$ and $u_1(1) = 7$. Use value iteration, with $\gamma = \frac{1}{2}$, to find an updated utility estimate, $u_2(0)$, for state $s = 0$.

Solution: The formula given on the formula sheet accidentally omitted the γ, so γ in the following answer is optional:

$$u_2(s) = r(s) + \gamma \max_a \sum_{s'} P(s'|s,a) u_1(s')$$

$$\begin{aligned} u_2(0) &= 3 + \frac{1}{2} \max \left(P(0|0,0)u_1(0) + P(1|0,0)u_1(1), \right. \\ &\quad \left. P(0|0,1)u_1(0) + P(1|0,1)u_1(1) \right) \\ &= 3 + \frac{1}{2} \max \left(0.8 \cdot 3 + 0.2 \cdot 7, 0.3 \cdot 3 + 0.7 \cdot 7 \right) \end{aligned}$$

(b) (15 points) Explain in words why reinforcement learning must use a policy that balances the competing demands of **exploration** and **exploitation**.

Solution:

- Exploration is needed to learn the results of all actions, but if you only explore, you won't achieve the best possible rewards.
- Exploitation is needed to achieve good rewards, but if you only exploit, you won't learn the results of all actions, so you may miss rewards that are better than those you've seend.

This page is scratch paper

This page is scratch paper, you may tear it off if you wish