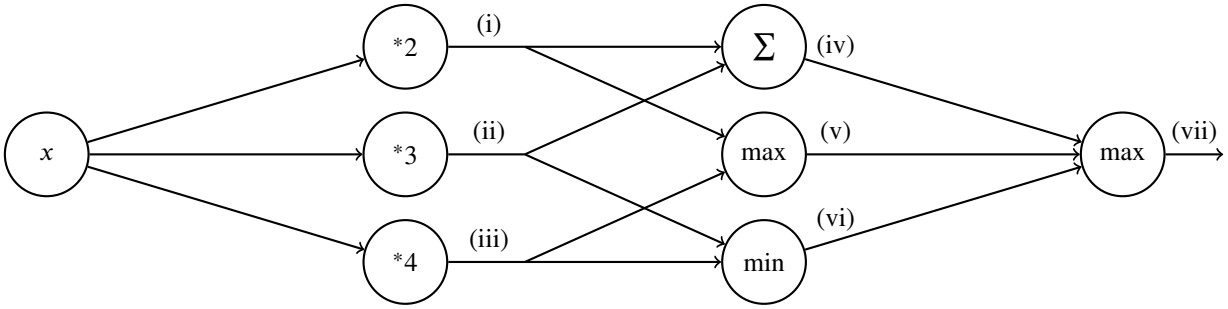
CS 188 Spring 2022 Introduction to Artificial Intelligence

# Exam Prep 14

## Q1. Deep Learning

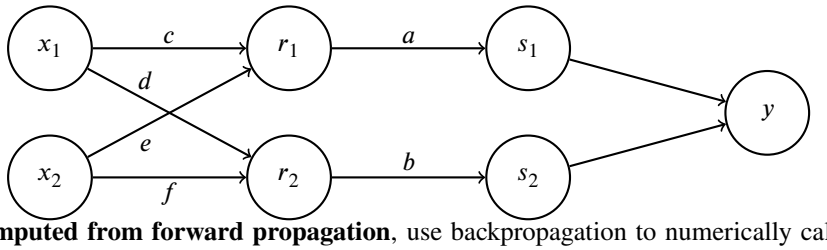
**(a)** Perform forward propagation on the neural network below for $x = 1$ by filling in the values in the table. Note that (i), … , (vii) are outputs after performing the appropriate operation as indicated in the node.

| (i) | (ii) | (iii) | (iv) | (v) | (vi) | (vii) |
|---|---|---|---|---|---|---|
| | | | | | | |




**(b)** **[Optional]** Below is a neural network with weights $a, b, c, d, e, f$. The inputs are $x_1$ and $x_2$.
The first hidden layer computes $r_1 = \max(c \cdot x_1 + e \cdot x_2, 0)$ and $r_2 = \max(d \cdot x_1 + f \cdot x_2, 0)$.
The second hidden layer computes $s_1 = \frac{1}{1+\exp(-a \cdot r_1)}$ and $s_2 = \frac{1}{1+\exp(-b \cdot r_2)}$.
The output layer computes $y = s_1 + s_2$. Note that the weights $a, b, c, d, e, f$ are indicated along the edges of the neural network here.

Suppose the network has inputs $x_1 = 1, x_2 = -1$.
The weight values are $a = 1, b = 1, c = 4, d = 1, e = 2, f = 2$.
Forward propagation then computes $r_1 = 2, r_2 = 0, s_1 = 0.9, s_2 = 0.5, y = 1.4$. Note: some values are rounded.

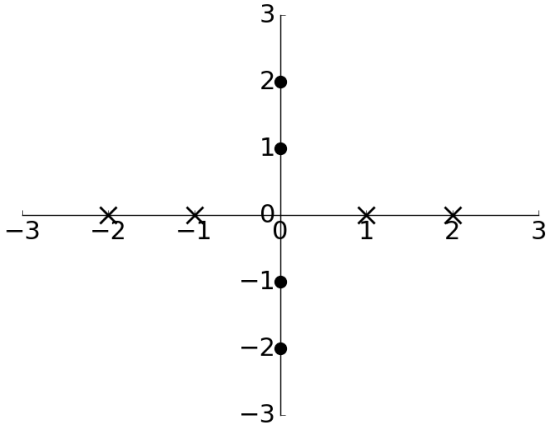



**Using the values computed from forward propagation**, use backpropagation to numerically calculate the following partial derivatives. Write your answers as a single number (not an expression). You do not need a calculator. Use scratch paper if needed.
*Hint:* For $g(z) = \frac{1}{1+\exp(-z)}$, the derivative is $\frac{\partial g}{\partial z} = g(z)(1 - g(z))$.

| $\frac{\partial y}{\partial a}$ | $\frac{\partial y}{\partial b}$ | $\frac{\partial y}{\partial c}$ | $\frac{\partial y}{\partial d}$ | $\frac{\partial y}{\partial e}$ | $\frac{\partial y}{\partial f}$ |
|---|---|---|---|---|---|
| | | | | | |

**(c)** Below are two plots with horizontal axis $x_1$ and vertical axis $x_2$ containing data labelled $\times$ and $\bullet$. For each plot, we wish to find a function $f(x_1, x_2)$ such that $f(x_1, x_2) \geq 0$ for all data labelled $\times$ and $f(x_1, x_2) < 0$ for all data labelled $\bullet$.

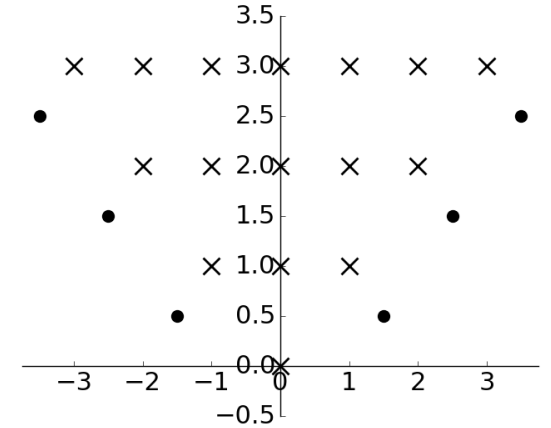
Below each plot is the function $f(x_1, x_2)$ for that specific plot. Complete the expressions such that all the data is labelled correctly. If not possible, mark "No valid combination".




$f(x_1, x_2) = \max(\ \underline{\textbf{(i)}} + \underline{\textbf{(ii)}}\ ,\ \underline{\textbf{(iii)}} + \underline{\textbf{(iv)}}\ ) + \underline{\textbf{(v)}}$

| | | | |
|---|---|---|---|
| **(i)** | ◯ $x_1$ | ◯ $-x_1$ | ◯ 0 |
| **(ii)** | ◯ $x_2$ | ◯ $-x_2$ | ◯ 0 |
| **(iii)** | ◯ $x_1$ | ◯ $-x_1$ | ◯ 0 |
| **(iv)** | ◯ $x_2$ | ◯ $-x_2$ | ◯ 0 |
| **(v)** | ◯ 1 | ◯ $-1$ | ◯ 0 |

◯ No valid combination




$f(x_1, x_2) = \underline{\textbf{(vi)}} - \max(\ \underline{\textbf{(vii)}} + \underline{\textbf{(viii)}}\ ,\ \underline{\textbf{(ix)}} + \underline{\textbf{(x)}}\ )$

| | | | |
|---|---|---|---|
| **(vi)** | ◯ $x_2$ | ◯ $-x_2$ | ◯ 0 |
| **(vii)** | ◯ $x_1$ | ◯ $-x_1$ | ◯ 0 |
| **(viii)** | ◯ $x_2$ | ◯ $-x_2$ | ◯ 0 |
| **(ix)** | ◯ $x_1$ | ◯ $-x_1$ | ◯ 0 |
| **(x)** | ◯ $x_2$ | ◯ $-x_2$ | ◯ 0 |

◯ No valid combination

# Q2. Q-learning

Consider the following gridworld (rewards shown on left, state names shown on right).

Rewards

| | |
|---|---|
| | |
| +10 | +1 |

State names

| | |
|---|---|
| A | B |
| G1 | G2 |

From state A, the possible actions are right(→) and down(↓). From state B, the possible actions are left(←) and down(↓). For a numbered state (G1, G2), the only action is to exit. Upon exiting from a numbered square we collect the reward specified by the number on the square and enter the end-of-game absorbing state $X$. We also know that the discount factor $\gamma = 1$, and in this MDP all actions are **deterministic** and always succeed.

Consider the following episodes:

Episode 1 ($E1$)

| $s$ | $a$ | $s'$ | $r$ |
|---|---|---|---|
| $A$ | ↓ | $G1$ | 0 |
| $G1$ | exit | $X$ | 10 |

Episode 2 ($E2$)

| $s$ | $a$ | $s'$ | $r$ |
|---|---|---|---|
| $B$ | ↓ | $G2$ | 0 |
| $G2$ | exit | $X$ | 1 |

Episode 3 ($E3$)

| $s$ | $a$ | $s'$ | $r$ |
|---|---|---|---|
| $A$ | → | $B$ | 0 |
| $B$ | ↓ | $G2$ | 0 |
| $G2$ | exit | $X$ | 1 |

Episode 4 ($E4$)

| $s$ | $a$ | $s'$ | $r$ |
|---|---|---|---|
| $B$ | ← | $A$ | 0 |
| $A$ | ↓ | $G1$ | 0 |
| $G1$ | exit | $X$ | 10 |

**(a)** Consider using temporal-difference learning to learn $V(s)$. When running TD-learning, all values are initialized to zero. For which sequences of episodes, if repeated infinitely often, does $V(s)$ converge to $V^*(s)$ for all states $s$?

(Assume appropriate learning rates such that all values converge.)
Write the correct sequence under "Other" if no correct sequences of episodes are listed.

☐ $E1, E2, E3, E4$ ☐ $E1, E2, E1, E2$ ☐ $E1, E2, E3, E1$ ☐ $E4, E4, E4, E4$
☐ $E4, E3, E2, E1$ ☐ $E3, E4, E3, E4$ ☐ $E1, E2, E4, E1$

☐ Other _____

**(b)** Consider using Q-learning to learn $Q(s, a)$. When running Q-learning, all values are initialized to zero.
For which sequences of episodes, if repeated infinitely often, does $Q(s, a)$ converge to $Q^*(s, a)$ for all state-action pairs $(s, a)$

(Assume appropriate learning rates such that all Q-values converge.)
Write the correct sequence under "Other" if no correct sequences of episodes are listed.

☐ $E1, E2, E3, E4$ ☐ $E1, E2, E1, E2$ ☐ $E1, E2, E3, E1$ ☐ $E4, E4, E4, E4$
☐ $E4, E3, E2, E1$ ☐ $E3, E4, E3, E4$ ☐ $E1, E2, E4, E1$

☐ Other _____