CSE 415, Winter 2025

Assignment 6

Last name:_____ First name:_____ UWNetID:_____

Due Monday night March 10 via Gradescope at 11:59 PM. (Note: Due to the tight schedule at the end of the quarter, there will be no grace days on this assignment, and two days only to turn in an assignment late – for a 25 percent penalty per day.) You may turn in either of the following types of PDFs: (1) Scans of these pages that include your answers (handwriting is OK, if it's clear), or (2) Documents you create with the answers, saved as PDFs. When you upload to Gradescope, you'll be prompted to identify where in your document your answer to each question lies.

Perform and provide answers for each of the seven exercises. Each TA on the staff has contributed or helped on one or more of the questions. These are intended to take 15-40 minutes each if you know how to do them. Each is worth between 15 points and 20 points. The total possible points for completing the exercises is 150 (not including the extra credit in the formatting bonus, described below).

If any corrections have to be made to this assignment, these will be posted in ED.

This is an individual-work assignment. Do not collaborate on this assignment. Do not use AI systems, such as ChatGPT.

Prepare your answers in a neat, easy-to-read PDF. Our grading rubric will be set up such that when a question is not easily readable or not correctly tagged or with pages repeated or out of order, then points will be be deducted. However, if all answers are clearly presented, in proper order, and tagged correctly when submitted to Gradescope, we will award a 5-point bonus. (Restated, one could lose the bonus due to poor photo contrast, handwriting that graders have difficulty reading, or mess-ups with the tagging of the answers when submitting to Gradescope.)

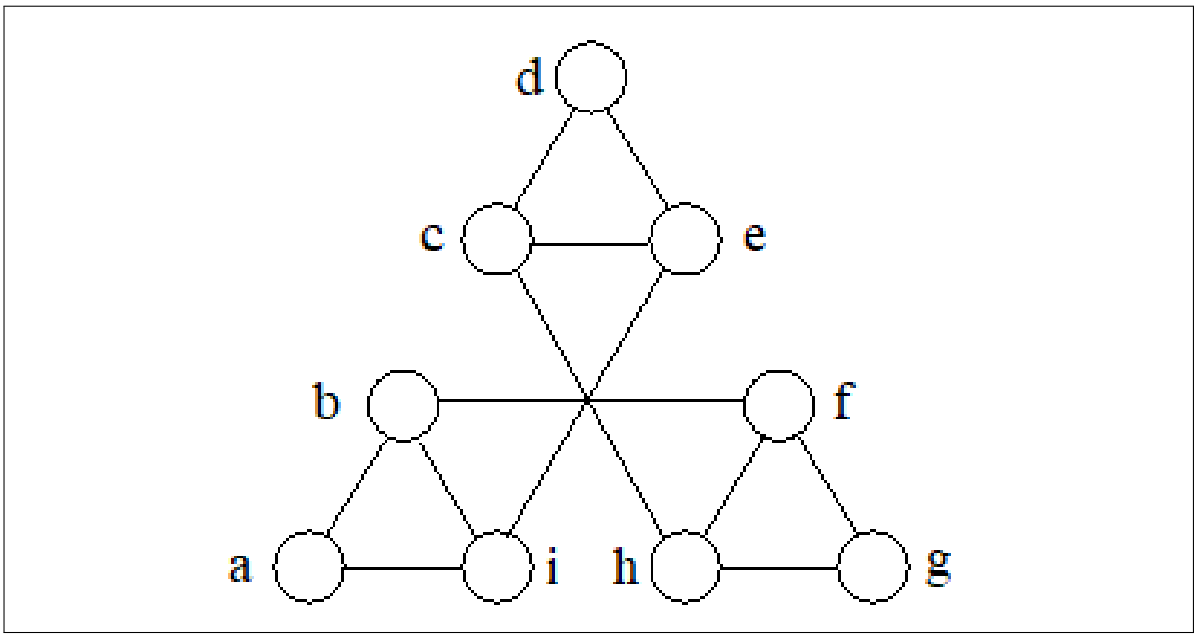
If you choose to typeset your answers in Latex using the template file for this document, please put your answers in blue while leaving the original text black. Using Latex is not required for the bonus points, but it might help in achieving a neat, easy-to-read PDF.

Version 250302. Has corrections to the numbers of points on the subproblems in Q4, so points add up to 20 there.

1 MDP Basics (Xinming)

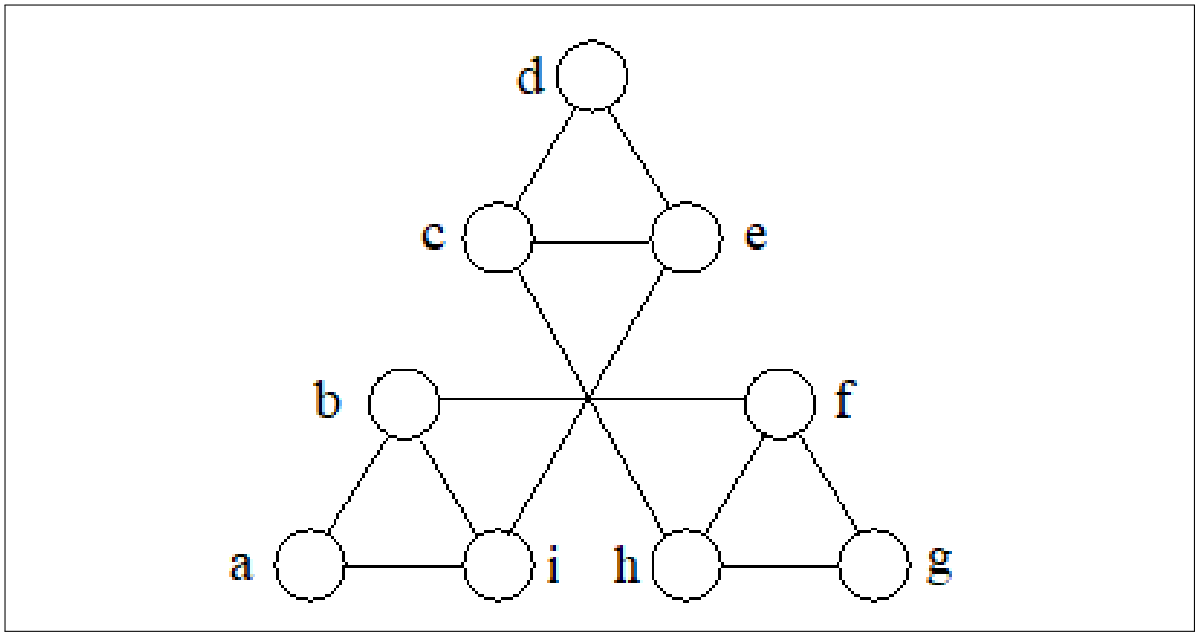
(20 points) This exercise covers basic properties of an MDP.

(a) (4 pts) Suppose that the transition function of TOH2-World is deterministic. Also, suppose that all transitions have a reward of -1, except the transition from g (the "goal") to the Terminal state, which has a reward of 100. Next assume that an agent, finding itself in an arbitrary state of this world, not using any discounting, seeks to maximize its total reward in whatever time remains in its episode. (Say it takes one second per transition in this world.) Assuming it acts optimally, there will be a specific sum of rewards that it can get from each state. Write down that value in each state below.



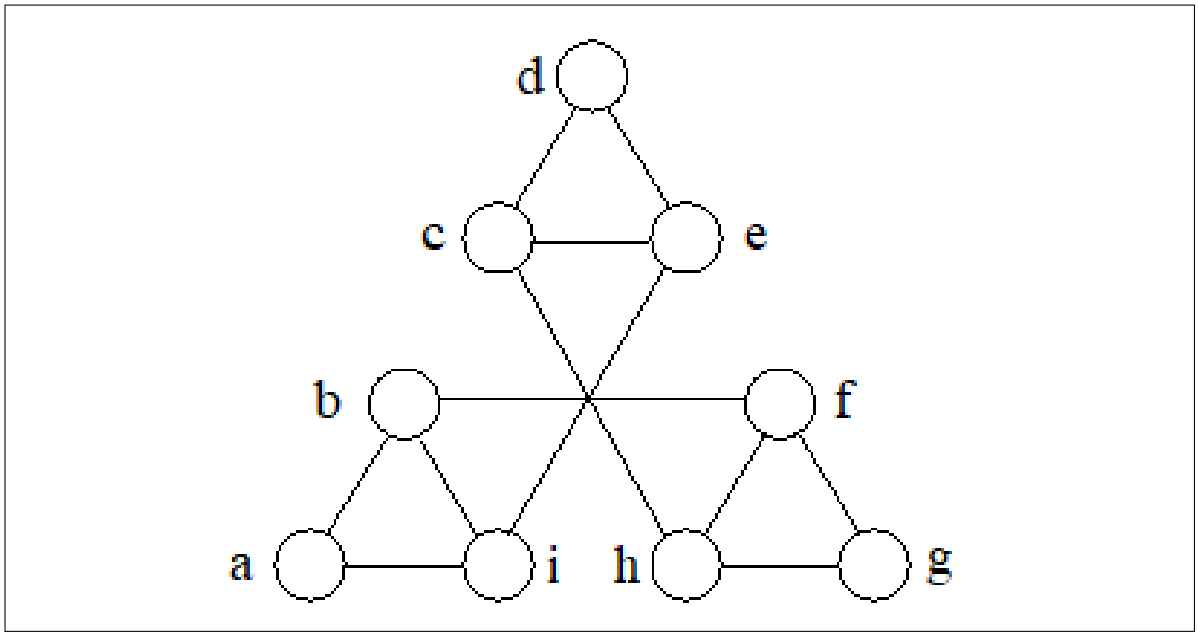
(b) (4 pts.) Drawing on the same diagram, use arrowheads on some of the edges to indicate an optimal policy.

(c) (3 pts.) Suppose state i is a secondary goal node. When the agent is at i, the only legal action is Exit, which transitions the agent to the Terminal state and yields a reward of 10. Using the copy of the graph below, write the value of each state assuming a time horizon of 2 moves. Except for the addition of the new goal assume that the reward function is as before. (and the main goal is still the main goal).



(d) (3 pts.) Also, indicate the optimal policy using arrowheads on this graph.

(e) (3 pts.) Now assume the agent uses discounting of future rewards, with $\gamma = 0.5$. Show the value of each node now.



(f) (3 pts.) Indicate the optimal policy in this case using arrowheads on this graph.

2 Basic Q-Learning in an MDP (Sahil)

(15 points) Use the Markov Decision Process (MDP) diagram below to help you as you answer the following question Q Learning. The states are A, B, C, D, E, F, G, and an unshown Terminal state.

A	B (-10)	C	D	E	F	G (+10)

Use Q Learning to estimate the true Q values of this MDP given the observed state transitions below. For each state transition, indicate which Q value is changing by giving the state and action associated with that value and giving its new value. You may wish to keep your own diagram of the current values while you work through the given transitions.

You may assume that at states $\{A, C, D, E, F\}$ there are two allowed actions, {Left, Right}, at states $\{B, G\}$, the only allowable action is the Exit action, and at the Terminal state there are no valid actions. Make sure to process the transitions in the order provided. All Q values are initialized to 0. Use a discount factor of $\gamma = 0.8$ and a learning rate of $\alpha = 0.5$.

State Transitions				New Q Values		
State (s)	Action (a)	New State (s')	Reward (r)	State	Action	Q(S,A)
A	Right	B	-1			
B	Exit	Terminal	-10			
E	Right	D	-1			
D	Right	E	-1			
E	Right	F	-1			
F	Right	G	-1			
G	Exit	Terminal	10			

3 Joint Distributions and Inference (Emilia)

(15 points) Let C represent the proposition that it is cloudy in Seattle. Let R represent the proposition that it is raining in Seattle. Consider the table given below.

C	R	$P(C, R)$
cloudy	*rain*	0.53
cloudy	*sun*	0.13
clear	*rain*	0.02
clear	*sun*	0.32

(a) (2 points) Compute the marginal distribution $P(C)$ and express it as a table.

(b) (2 points) Similarly, compute the marginal distribution $P(R)$ and express it as a table.

(c) (2 points) Compute the conditional distribution $P(R|C = \textit{cloudy})$ and express it as a table. Show your work/calculations.

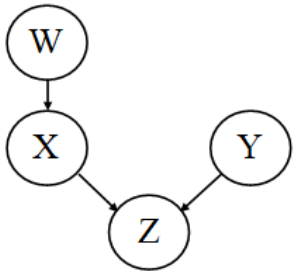
(d) (2 points) Compute the conditional distribution $P(C|R = \textit{sun})$ and express it as a table. Show your work/calculations.

(e) (3 points) Is it true that $C \perp\!\!\!\perp R$? (i.e., are they statistically independent?) Explain your reasoning.

(f) (4 points) Suppose you decide to track additional weather patterns of Seattle such as temperature (hot/cold), humidity (humid/dry), and wind (windy/calm) denoted as the random variables T, W, H respectively. Is it possible to compute $P(C, R, T, W, H)$ as a product of five terms? If so, show your work. What assumptions need to be made, if any? Otherwise, explain why it is not possible.

4 Bayes Net Structure and Meaning (Krish)

(20 points) Consider a Bayes net whose graph is shown below.



Random variable W has a domain with two values $\{w_1, w_2\}$; the domain for X has three values: $\{x_1, x_2, x_3\}$; Y's domain has three values: $\{y_1, y_2, y_3\}$; and Z's domain has two values: $\{z_1, z_2\}$.

(a) (6 points) Give a formula for the joint distribution of all four random variables, in terms of the marginals (e.g., $P(W = w_i)$), and conditionals that must be part of the Bayes net (e.g., $P(Z = z_m | X = X_j, Y = y_k)$).

(b) (2 point) How many probability values belong in the (full) joint distribution table for this set of random variables?

(c) (4 points) For each random variable: give the number of probability values in its marginal (for W) or conditional distribution table (for the others).

W:

X:

Y:

Z:

(d) (8 points) For each random variable, give the number of *non-redundant* probability values in its table from (c).

W:

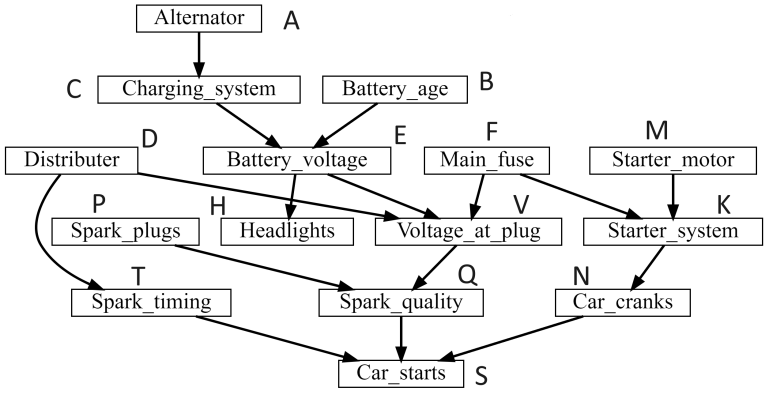
X:

Y:

Z:

5 D-Separation (Krish)

(20 points) Consider the automobile engine-diagnosis Baves Net below.



(a) (1 point) Assuming no observations have been made, are A and B independent?

(b) (1 point) From the diagram can we assume that, are A and C independent?

(c) (5 points) List all "undirected" paths from C to S.

(d) (3 points) Which of the above are active paths?

(e) (2 points) Is it guaranteed that $A \perp\!\!\!\perp S|Q, C$? If not, give an active path from A to S.

(f) (2 points) Is it guaranteed that $B \perp\!\!\!\perp S|V, K, T$? If not, give an active path from B to S.

(g) (2 points) Is it guaranteed that $P \perp\!\!\!\perp B|S$? If not, give an active path from P to B.

(h) (2 points) Is it guaranteed that $A \perp\!\!\!\perp M|F, V, S$? If not, give an appropriate active path.

(i) (2 points) What is the longest loop-free undirected path you can find in this graph? What nodes would need to be observed to make it an active path?

6 Choosing Hidden Sequences (Xinming)

(20 points) Consider the Jones family – Mom, Dad, Jim, Sally. Jim and Sally are twin siblings, and both 16 years old. Once a month, the family goes out for ethnic food – it's either Thai or Indian. Jim has his list of favorite restaurants, and Sally has her own list of her favorites. Each month, Dad uses one of the lists to choose what type of food the family will eat during their outing. He tends to stick with the same list more often than switch, and he tends to slightly favor Jim's list, maybe because of his own preferences.

With the current list, Dad selects a restaurant at random, assuming the restaurants on the list are equally likely to be picked. But they don't necessarily go to that restaurant. If the restaurant is Indian, Jim gets to pick any Indian restaurant from his list, and they go there. If the restaurant is Thai, then Sally gets to pick any Thai restaurant from her list, and they go there.

```
Jim's list:
  My-Pad-Thai (Thai), Punjabi Kitchen (Indian), Delhi Curry (Indian).
Sally's list:
  Bangkok Bites (Thai), Mango Mansion (Thai),
  Siam Shack (Thai), Delhi Curry (Indian).
```

Then after Dad makes his random selection from the current list, he uses some kind of spinning device that we don't understand, and he determines whether to switch lists for next time according to the following transition probabilities.

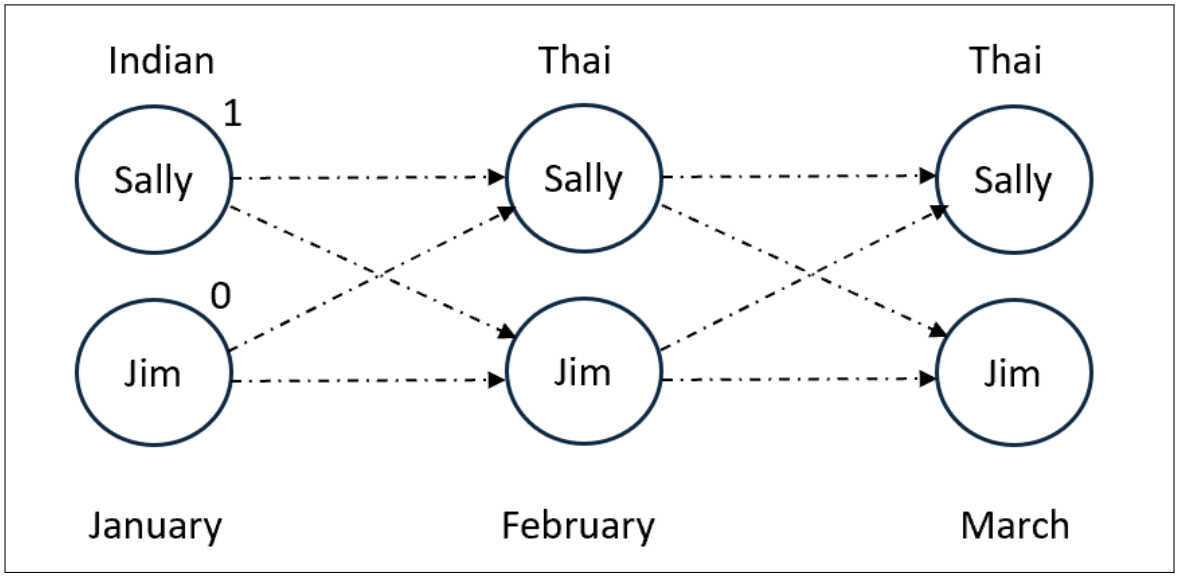
If the current list is Jim's, the probably of staying with Jim's is 0.7, and switching to Sally's is 0.3. If the current list is Sally's the probably of staying with Sally's is 0.6, and switching to Jim's is 0.4.

(a) (2 points) Assume Dad starts the year using Sally's list for January's outing. He's public about that, but in subsequent months, he does not tell anyone which list he is using.

One possible 3-month sequence is list uses is: Sally's, Sally's, Jim's (for January, February, March). What is the probability of that sequence of list usages, without considering the ethnicities of the restaurants they visited?

(b) (2 points) Suppose that quarter of the year they end up going to an Indian restaurant in January, a Thai restaurant in February, and a Thai restaurant in March. Taking that additional information into consideration. what is the probability of that same list-usage sequence and seeing that sequence of cuisine choices?
(i.e., compute P(X_1 =Sally, X_2 =Sally, X_3 =Jim, E_1 =Indian, E_2 =Thai, E_3 =Thai).

(c) (6 points) What is the most likely sequence of list usages by Dad given that same restaurant ethnicity sequence? (Indian, Thai, Thai). Use the diagram template below to create a trellis diagram for this problem. If the probability coming into a dashed arrow is 0, leave it dashed. Otherwise, write over the dashed line to make it a more solid arrow. Before identifying the most likely list-usage sequence, compute, at each node, the probability of reaching that node along a most likely path from the Sally-in-January node. The probability value 1 is provided on that starting node, since it's a given in this exercise. Naturally, the probability of getting to the Jim-in-January node is 0, because we're given that Sally's list is the one being used in the first month (January). Use the Viterbi algorithm to get these probabilities at the other 4 nodes of the diagram. Finally, highlight the most probable path that starts at Sally-in-January and ends at one of the nodes in the March column. (Hint: the HMM worksheet involves a similar computation.) Show the factors you are multiplying on the appropriate edges of the trellis. For the March column, at least, use a calculator to show the two probabilities to 4 decimal places.



(d) (2 points) Suppose Mom somehow gets suspicious that Sally's preferences are not being considered fairly, in comparison with Jim's. What is a possible basis for that?

(e) (3 points) Compute the stationary probability of Jim's list vs Sally's list being used. (Assume that when the sequence extends beyond 12 months, Dad does NOT automatically go back to Sally's list each January. In other words, the Markov Model represented by the transition CPT is not limited to 12 time steps.)

Show the needed equations before you solve them, and show the steps you use in solving them.

(f) (2 points) Compute the marginal probabilities of the family's having Indian and Thai food on their outings, corresponding to the stationary probabilities of list usages.

(g) (3 points) To what extent are these distributions (that you computed in parts e and f) biased against or for each child's preferences? Explain. There are three pertinent points you can describe here.

7 Perceptrons (Sahil)

(20 points) You are planning an exciting road trip across California and need to decide which cities to include in your itinerary. You classify cities into two categories: Must-Visit Cities and Skip Cities. A city is labeled as Must-Visit (+1) if it is highly attractive for tourists, while a city is labeled as Skip (−1) if it does not meet your travel preferences. You decide to use a perceptron model to classify cities based on their features:
Dataset Features and Specification:
Natural Attractions: Represented as low = −2, moderate = 0, high = 2.
Accommodation Cost: Represented as cheap = 1, moderate = 2, expensive = 3.
Travel Distance: Represented as far = −2, near = 2.

Perceptron Parameters

- The perceptron has a weight vector, w, with four components: bias term, natural attractions, accommodation cost, and travel distance.
- Initial weights: $w = [1, 0, 0, 0]$.
- Threshold: 0.
- Learning rate: 1.

Example Number	Natural Attractions	Accommodation Cost	Travel Distance	Label
1	high	expensive	near	Must-Visit
2	moderate	moderate	far	Skip
3	high	moderate	near	Skip
4	moderate	expensive	near	Must-Visit
5	low	cheap	fast	???

Table 1: Road Trip Dataset

(a) (4 points) What would be the updated weights **w** after processing example number 1?

(b) (3 points) What would be the updated weights **w** after processing example number 2?

(c) (3 points) What would be the updated weights **w** after processing example number 3?

(d) (3 points) What would be the updated weights **w** after processing example number 4?

(e) (3 points) What is your prediction for the label of example number 5 based on the final weights?

(f) (4 points) Is convergence guaranteed for this perceptron and data set? Why or why not?

8 AI and the Potential for Harm (Emilia)

(20 points) Isaac Asimov was an American scientist and author who wrote a collection of short stories about robots. In his fictional world, robot behavior was governed by the three laws of robotics:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.

2. A robot must obey orders given it by human beings except where such orders would conflict with the First Law.

3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

The stories often dealt with paradoxical situations arising from the attempted application of these laws. Despite their shortcomings, the laws have been influential in considerations of the ethics of artificial intelligence.

For this question, consider one of the issues listed below. For your selected issue, consider how might Asimov's Laws serve as a starting off point for guidelines to promote ethical use of AI technologies (although the laws are written to apply to robots, consider them as being relevant more generally for artificial intelligence technologies). The links provided lead to optional resources that may help you become familiar with each issue listed.

- Violations of personal intellectual property rights - AI requires vast amounts of training data, often scraped from internet sites. Writers, musicians, and artists have been especially affected by having their work taken without their permission for such training. For more information on this issue, see: Generative AI Has an Intellectual Property Problem

- Deep fakes, bots, and misinformation - Using AI technologies, the dissemination of false information or false interpretations of and responses to events has become widespread. For more information on this issue, see: The Rise of Artificial Intelligence and Deepfakes and What are 'bots' and how can they spread fake news?

- Violations of personal privacy - A large proportion of social interaction currently occurs online. What we view and what we respond to provides insights into our beliefs and values. Combined with other data that is available about us online, it can feel as if the concept of personal privacy is obsolete. For more information on this issue, see: The privacy paradox with AI

- The use of AI in warfare - From automated drones to more sophisticated systems that can analyze and respond to situations in a human-like manner, AI technologies are

likely to be increasingly used in areas of conflict as the technologies become ever more sophisticated. For more information on this issue, see: War, Artificial Intelligence, and the Future of Conflict

- Environmental costs of AI - Training, deploying, and fine-tuning generative AI models requires enormous quantities of electricity and water. For more information on this issue, see: Explained: Generative AI's environmental impact
- An issue of your own choosing - If choosing this selection, please give a short description of the issue and a link to a resource that provides a basic introduction to your issue as part of your response to the first item below

Please answer the following questions:

(a) (1 points) Which issue have you selected to address?

(b) (5 points) Asimov's Laws assume the AI (the robot) is self aware. To date, a truly self-aware AI has not been developed, and clearly the potential for harm does not require self awareness. Given the ways AI is used currently, **how might the laws still contribute to a discussion of the ethical use of AI and to the development of ethical AI technologies**? You might want to consider questions such as: Where is AI used? Who controls it? What are the consequences when it works as intended? When it doesn't work as intended?

(c) (5 points) How might your chosen issue violate the first law and "allow a human being to come to harm"?

(d) (5 points) Suggest at least one guideline for the use of AI that might be helpful in reducing the potential for the harm you described above.

(e) (4 points) Do you think it is possible to regulate the use of AI so that it is predominantly used in ways that are beneficial to humanity? If so, explain why you are optimistic about the use of AI and if not, explain why your views are perhaps more pessimistic.