Assignment 7 in CSE 415, Winter 2023

by the Staff of CSE 415

This is due Tuesday, March 7, via Gradescope at 11:59 PM. Prepare a PDF file with your answers and upload it to Gradescope. The PDF file can created however you like. For example it can be from a scan of a printout of the assignment document onto which you have hand-written your answers. Or it can be from Word or Latex file with your answers. It can even be from photos of your handwritten answers on plain paper. You don't have to include the questions themselves, but it is fine to do so. In any case, it must be very clear to read, and it must be obvious and easy for each grader where to find your solutions to the exercises.

As with Assignment 4, this is an *individual work* assignment. Collaboration is not permitted.

Do the following exercises. These are intended to take 10-30 minutes each if you know how to do them. Each is worth 10 or 20 points. The total of possible points is 120. Names of responsible staff members are given for each question.

If corrections or clarifications to the problems have to be given, this will happen in the ED discussion forum under topic "Assignment 7."

Last name: ____________________, first name: ____________________

Student number: ____________________

1 Joint Distributions and Inference (Kenan)

(10 points) Let D be the random variable that represents whether the food is delicious. Let F be the random variable that represents whether you will choose to eat noodles or steak.

Consider the table given below:

D	F	$\mathbb{P}(D, F)$
delicious	*noodles*	0.35
delicious	*steak*	0.5
horrible	*noodles*	0.1
horrible	*steak*	0.05

(a) (1 point) Compute the marginal distribution $\mathbb{P}(D)$ and express it as a table.

(b) (1 point) Similarly, compute the marginal distribution $\mathbb{P}(F)$ and express it as a table.

(c) (2 points) Compute the conditional distribution $\mathbb{P}(D|F = steak)$ and express it as a table. Show your work/calculations.

(d) (2 points) Compute the conditional distribution $\mathbb{P}(F|D = horrible)$ and express it as a table. Show your work/calculations.

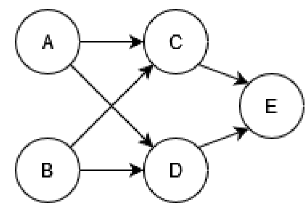
(e) (1 point) Is it true that $D \perp\!\!\!\perp F$? (i.e., are they statistically independent?) Explain your reasoning.

(f) (3 points) After going to the restaurant too many times, you wish to try a new restaurant for noodles and steak. Suppose that the joint probability distribution for D and F for the new restaurant is as follows:

D	F	$\mathbb{P}(D, F)$
delicious	*noodles*	0.45
delicious	*steak*	0.3
horrible	*noodles*	0.05
horrible	*steak*	0.2

Suppose you have a probability of 0.8 of going to the new restaurant and a probability of 0.2 of going to the original restaurant. What's the probability that you will have gone to the new restaurant given that you will have eaten delicious noodles?

2 Bayes Net Structure and Meaning (Phuong)



(10 points) Consider the following Bayes net, where variable A has a domain with 2 values, variable B's domain has 3 values, C's domain has 2 values, D's domain has 4 values, and E's domain has 5 values.

(a) (3 points) Write down the full joint probability distribution associated with the above Bayes net. Express the answer as a product of terms representing individual conditional probabilities tables associated with this Bayes Net:

(b) (1 point) How many probability values (number of entries) belong in the full joint distribution table for this set of random variables?

(c) (1.5 points) For each random variable: give the number of probability values (number of entries) in its marginal (for A and B) or conditional distribution table (for the others).

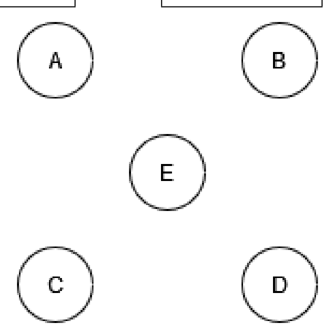
A : B : C : D : E :

(d) (1.5 points) For each random variable, give the number of *non-redundant* probability values in its table from (c).

A : B : C : D : E :

(e) (3 points) Draw the Bayes net associated with the following joint distribution by connecting (directed) arrows between each variable:

$$\mathbb{P}(A|B) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C|A) \cdot \mathbb{P}(D|C, B) \cdot \mathbb{P}(E|A, B, C)$$



3 D-Separation (Steve)

(20 points) Consider the Bayes Net graph β below, which represents the topology of a web-server security model. Here the random variables have the following interpretations:

V = Vulnerability exists in web-server code or configs.

C = Complexity to access the server is high. (Passwords, 2-factor auth., etc.)

S = Server accessibility is high. (Firewall settings, and configs on blocked IPs are permissive).

A = Attacker is active.

L = Logging infrastructure is state-of-the-art.

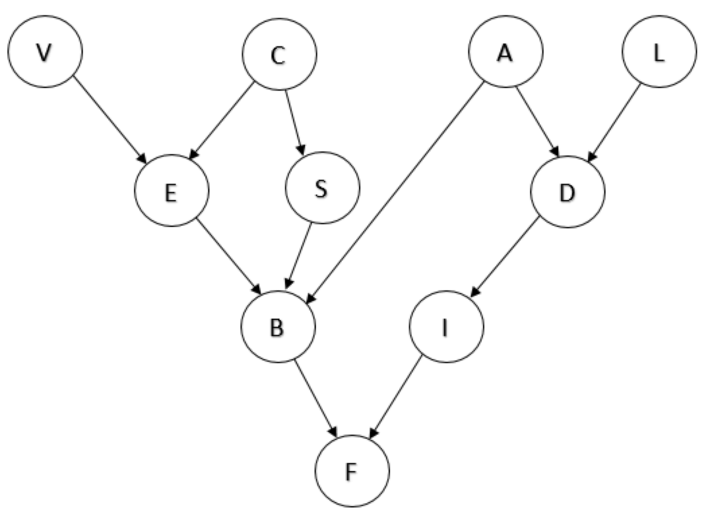
E = Exposure to vulnerability is high.

D = Detection of intrusion attempt.

B = Break-in; the web server is compromised.

I = Incident response is effective.

F = Financial losses are high (due to data loss, customer dissatisfaction, etc).



Let β' be the undirected graph obtained from β by removing the arrowheads from the edges of β. By an "undirected path" in β we mean any path in β'. A "loop-free" path is any path in which no vertex is repeated.

(a) (5 points) List all loop-free undirected paths from C to A in the graph β.

(b) (5 points) Suppose random variable F is observed, and no others are observed. Then which (if any) of those paths would be active paths? Justify your answer.

For each of the following statements, indicate whether (True) or not (False) the topology of the net guarantees that that the statement is true. If False, identify a path ("undirected") through which influence propagates between the two random variables being considered. (Be sure that the path follows the D-Separation rules covered in lecture.) The first one is done for you.

(c) $E \perp\!\!\!\perp S$: False (ECS)

(d) (1 point) $E \perp\!\!\!\perp S \mid C$

(e) (1 point) $V \perp\!\!\!\perp A \mid E, B$

(f) (1 point) $C \perp\!\!\!\perp I \mid A, B, S$

(g) (1 point) $L \perp\!\!\!\perp C \mid D, E, F$

(h) (1 point) $F \perp\!\!\!\perp V \mid B$

(i) (5 points) Suppose that the company hired an outside expert to examine the system and she determines that B and E are true: a Break-in has occurred (the web server is compromised), and the Exposure to vulnerability is high). Given this information, your job is to explain to management why getting additional information about A (Whether an attacker is active) could have an impact on the probability of V (Vulnerability existing in the web-server code or configs). Give your explanation, for the manager of the company, using about 3 to 12 lines of text, which should be based on what you know about D-separation, applied to this situation. However, your explanation should not use the terminology of D-separation but be in plain English. (You can certainly use words like "influence", "probability", "given", but not "active path", "triple", or even "conditionally independent").

4 Perceptrons (Mingyu)

(20 points) For all parts of this question perceptrons should output 1 if $w_n + \sum_{i=0}^{n-1} w_i x_i \geq 0$ and 0 otherwise. The weight w_n is called the bias weight.

(a) (5 points) Assuming there will be two inputs x_0 and x_1, each with possible values in $\{0, 1\}$, give values for a triple of weights $\langle w_0, w_1, w_2 \rangle$ such that the corresponding perceptron would act as a NAND gate for the two inputs. (Weight w_2 is the bias weight.) Note that a NAND gate outputs 1 when at least one of the inputs is 0; it outputs 0 otherwise.

(b) (5 points) Draw a perceptron, with weights, that accepts a single integer x and outputs 1 if and only if the input is greater than or equal to -8. Be sure to include the bias input of value 2 and its weight in your diagram. Draw another perceptron that outputs 1 if and only if the input is less than or equal to 8.

(c) (2 points) Using the previous perceptrons, create a two-layer perceptron that outputs 1 if $|x| \leq 8$, and 0 otherwise.

(d) (5 points) Suppose we want to train a perceptron to compare two numbers x_0 and x_1 and produce output $y = 1$ provided that x_1 exceeds x_0 by at least 5. Assume that the initial weight vector is: $\langle w_0, w_1, w_2 \rangle = \langle 0, 0, 1 \rangle$. Consider a first training example: $(\langle x_0, x_1 \rangle, y) = (\langle 2, 1 \rangle, 0)$. This says that with inputs 1, and 2, the output y should be 0, since 2 exceeds 1 by only 1. What will be the new values of the weights after this training example has been processed one time? Assume the learning rate is 2.

(e) (3 points) Continuing with the last example, now suppose that the next step of training involves a different training example: $(\langle 2, 8 \rangle, 1)$. The output for this example should be 1, since 8 does exceed 2 by at least 5. Starting with the weights already learned in the first step, determine what the adjusted weights should be after this new example has also been processed once.

5 The Laws of Robotics (Emilia)

(20 points) In the 1940's, Isaac Asimov introduced a set of three laws to govern robot behavior:

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 the 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 Laws.

(NOTE: You might also want to take a look at this cartoon `https://xkcd.com/1613/`)

Imagine you live in a world, essentially identical to our own, where personal robotic assistants are a bit more advanced than they are here. They are still not commonplace, but in another 10 years (in our scenario), they could very well be. Further, imagine that although you are a poor college student, your best friend and housemate works in tech and loves being the first to obtain new tech gadgets. Consequently, you have access to a brand new personal robot in your apartment.

You are a computer science student and you're also a science fiction aficionado. Your favorite author – Isaac Asimov. You can't resist hacking your new robot so that it now is programmed to obey Asimov's three laws before considering any other part of its programming. You feel quite pleased with yourself. Now, not only do you have a personal robot, you also have a personal bodyguard.

You've been hearing a lot about advances in AI in the news recently. You're especially interested in algorithms that create artistic masterpieces and you'd also really like to be able to hold conversations with your robot that are more engaging than the interactions it is currently capable of. Given that your first attempt at hacking the robot seems so successful, you are seriously considering trying to "improve" it some more. You are also interested in the ethical applications of AI, so the next change you make is adding a sophisticated Ethics Module to the robot's codebase.The robot is now familiar with the major issues under discussion with regards to the use of AI and its effects on individuals and societies.

Considering the above information, please answer the following free response questions. Provide a justification for your response based on one (or more) of Asimov's Laws, referring to them by number as appropriate (e.g. "According to Asimov's 1st Law, ..."). Note that there are not specific "correct" answers to these questions. We are interested in your thoughts on current issues and your reasoning.

(a) (2 points) Assuming the same issues that concern us are relevant in our imagined world, what might be some of the issues related to AI that the robot is now aware of, thanks to its Ethics Module? Please mention at least two different issues.

(b) (5 points) Encouraged by your successes installing Asimov's Laws and Ethics, you next want to give your robot artistic abilities. Surprisingly, the installation fails, even though you try the installation several times. In exasperation, you moan, "Why, why is this happening?" Although you weren't actually talking to the robot, it responds and tells you what the problem is. Apparently, the robot views the installation as potentially harmful. You ask it to explain itself – after all, how can making beautiful pictures harm anyone? What might the robot's answer be?

(c) (5 points) You're feeling a bit discouraged, but you decide to try to make the robot a better conversationalist by installing a ChatGPT Module. This time, you are not really surprised when this installation also fails. Instead of wasting more time trying to install it, you just ask your robot, "What is it this time?" What might the robot tell you?

(d) (5 points) Giving up on the robot, you go to your room to work on your computer. At least it doesn't argue with you about everything. You start typing on your keyboard, searching for how to uninstall Asimov's Laws. Suddenly, you realize that the robot is standing behind you, looking at your computer screen over your shoulder. What do you think the robot says or does next (and why)?

(e) (3 points) Do you think the robot was correct in its responses to scenarios (a) and (b) above? Please explain your response.

6 Markov Models (Mingyu)

(20 points) According to an unnamed source, the stock market can be modeled using a Markov model, where there are two states "bull" and "bear." The dynamics of the model are given:

S_{t-1}	S_t	$P(S_t \mid S_{t-1})$
bull	bull	0.8
bull	bear	0.2
bear	bull	0.3
bear	bear	0.7

(a) (2 points) Suppose it's given that S_0 = bull. Compute the probability that S_2 = bull.

(b) (4 points) Compute the stationary probabilities for bull and bear.

(c) (2 points)Now suppose that whenever it's a bull market, a certain company's (Acme, Inc) stock stays the same or rises in value with probability 0.8 and falls in value with probably 0.2.

When it's a bear market Acme's stock value stays the same or rises with probability 0.4 and falls with probability 0.6.

Suppose an observer cannot directly tell whether the state of the stock market is bull or bear, but can only see whether Acme's stock is "rising" or "falling."

State S	Observation Q	$P(Q \mid S)$
bull	rising	0.7
bull	falling	0.3
bear	rising	0.2
bear	falling	0.8

Suppose $P(S_0 = \text{bull}) = 0.5$. If the observation at time 1 is "rising," what is the belief in S=bull right after the observation?

(d) (2 points) Suppose at time 2, the observation is "falling". what is the belief in $S =$ bull right after that observation? (This belief will take into consideration the previous belief you computed above.)

(e) (2 points) Suppose that the actual state sequence for the first four time steps is bear, bear, bear, bull.

What is the probably of observing the sequence (starting at $t = 1$) rising, falling, rising?

(f) (3 points) Now, what if the state sequence was bull, bear, bear, bull. What is the probability of observing the same sequence of stock changes as above?

(g) (2 points) Which of these two state sequences is more likely, given that sequence of observations?

(h) (3 points) Is there another state sequence that is even more likely? Explain.

7 Probabilistic Context-Free Grammars (Kenan)

(20 points) Consider the sentence, "Duplicate codes duplicate codes." Here there is ambiguity at multiple levels: lexical (word parts of speech and meanings), and syntactic (phrase structure). The semantics also vary.

The sentence could mean that copied codes (referred to by the first occurrence of "duplicate codes") now make a copy of (i.e., they "duplicate") other codes.

It could also mean that a programmer named "Duplicate" programs ("codes") other codes that themselves are duplicates of some codes.

With the probabilistic context-free grammar given below, find two parses, and compute a score for each one. Then identify the most probable parse using the scores. Assume the number at the right of a production is its conditional probability of being applied, given that the symbol to be expanded is that production's left-hand side. The probabilities for all the given productions for a specific left-hand-side non-terminal might not sum to one here, because we are showing a relevant subset of a larger set of productions.

(a) (7 points) Convert each probability into a score by taking score $= -\log_{10}(p)$. Round scores to 2 decimal places of accuracy. Write the production scores in the "__.__" blanks.

S	::= NP VP	0.9	0.05	VB	::= codes	0.08	__.__
NP	::= ADJ NP	0.3	__.__	VB	::= duplicate	0.02	__.__
NP	::= NN	0.2	__.__	NNS	::= codes	0.06	__.__
NP	::= NNS	0.4	__.__	NN	::= duplicate	0.07	__.__
VP	::= VB NP	0.1	__.__	ADJ	::= duplicate	0.05	__.__
VP	::= VBG NP	0.8	__.__				
VP	::= VB S	0.5	__.__				

(b) (5 points) Give a first parse for the sentence. This parse should correspond to the interpretation, "Copied codes ("duplicate codes") now make a copy of ("duplicate") other codes ("codes") " Compute the (total) score for this parse, showing the production scores at each internal node.

`duplicate` `codes` `duplicate` `codes`

(c) (5 points) Give the second parse. This parse should correspond to the interpretation, "A programmer named 'duplicate' programs ("codes") other codes that are copied ("duplicate codes")" Compute its total score, also showing the production scores at each internal node.

`duplicate` `codes` `duplicate` `codes`

(d) (3 points) Convert each score back to a probability and write them here as P1 and P2. Then tell which parse is more probable.