CMPUT 366, Winter 2022
Assignment #2

Due: Monday, Feb. 28, 2022, 11:59pm
Total points: 98

For this assignment use the following consultation model:

1. you can discuss assignment questions and exchange ideas with other *current* CMPUT 366 students;
2. you must list all members of the discussion in your solution;
3. you may **not** share/exchange/discuss written material and/or code;
4. you must write up your solutions individually;
5. you must fully understand and be able to explain your solution in any amount of detail as requested by the instructor and/or the TAs.

Anything that you use in your work and that is not your own creation must be properly cited by listing the original source. Failing to cite others' work is plagiarism and will be dealt with as an academic offence.

First name: ______________________________

Last name: ______________________________

CCID: ________________________@ualberta.ca

Collaborators: ____________________________

These are example solutions for assignment 2.

1. **(Probability theory)**

   (a) **[20 points]**

   Consider the following scenario. 2% of the people who walk through a specific metal detector at YEG are carrying a gun. 30% of the people who walk through the same metal detector are carrying coins. The remaining 68% are carrying nothing made of metal. Everyone carries either nothing, coins, or a gun through the detector; never both coins and a gun.

   If someone carries a gun through this metal detector, it will beep with probability 95%. If someone carries coins through this same metal detector, it will beep with probability 80%. If someone carries nothing made of metal through the detector, it will still beep about 25% of the time.

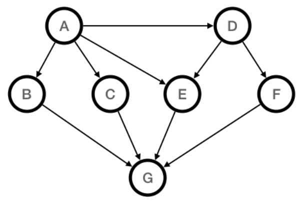
   Suppose that the metal detector beeps when someone walks through it. With what probability is that person carrying a gun? Show how you calculated your answer.

   We can answer this question using Bayes' rule. First, specify the probabilities. Let C be a random variable with domain $\{gun, coins, nothing\}$. Let B be a binary random variable that is true when the machine beeps. We are asked to compute $P(C = gun \mid B = true)$:

$$
\begin{aligned}
P(C = gun \mid B = true) &= \frac{P(B = true \mid C = gun)P(C = gun)}{P(B = true)} \\
&= \frac{P(B = true \mid C = gun)P(C = gun)}{\sum_{x \in dom(C)} P(B = true \mid C = x)P(C = x)} \\
&= \frac{.95 \times .02}{(.95 \times .02) + (.80 \times .30) + (.68 \times .25)} \\
&= \frac{.019}{.429} \\
&\approx .044.
\end{aligned}
$$

2. **Belief networks**

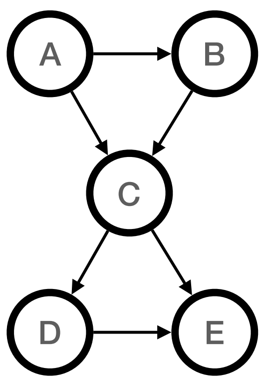
   (a) **[5 points]** What factorization of the joint distribution $P(A, B, C, D, E, F, G)$ does the network below represent?

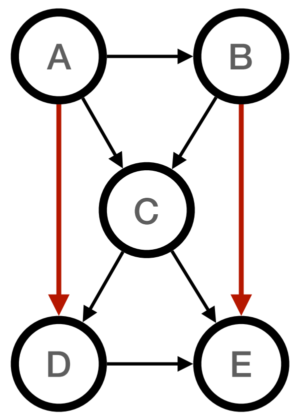


$P(G|B, C, E, F)P(B|A)P(C|A)P(E|A, D)P(F|D)P(A)P(D|A)$

Note: The order of the factors does not matter.

   (b) **[5 points]** Draw a belief network that is consistent with a joint distribution that factors as $P(A, B, C, D, E) = P(E|C, D)P(D|C)P(C|A, B)P(B|A)P(A)$.
   For **5 bonus marks**, draw another, different belief network that is *also* consistent with this factoring.





(For bonus marks, any acyclic graph that adds extra edges to the real graph will be consistent.)

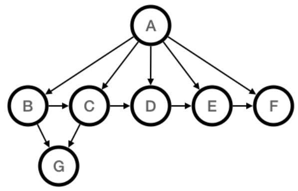
   (c) **[3 points]** Suppose that every random variable in the joint distribution of question (2b) has a domain containing 10 elements. How many rows are needed to list the full joint distribution in an explicit table?
   $10^5 - 1$
   (10^5 is also acceptable for full marks)

   (d) **[7 points]** Suppose that every random variable in the joint distribution of question (2b) has a domain containing 10 elements. How many rows in total are needed to list the conditional probability tables for your belief network representation?

$P(E\|C, D)$:	9×10^2	(10^3 also acceptable)
$P(D\|C)$:	9×10	(10^2 also acceptable)
$P(C\|A, B)$:	9×10^2	(10^3 also acceptable)
$P(B\|A)$:	9×10	(10^2 also acceptable)
$P(A)$:	9	(10 also acceptable).

Total of $900 + 90 + 900 + 90 + 9 = 1989$.
(2210 also acceptable)

3. **(Variable Elimination)** Consider the belief network below.



(a) **[15 points]** List the factors that would be created, and the operations used to create them, by running the variable elimination algorithm on this belief network to answer the query $P(B|G,E)$. Use the variable ordering G, E, A, B, C, D, F.

(Same factors will be created regardless of the values queried against for B and E; suppose for now that they are b and e.)

i. Construct factors for each conditional probability table:

$$\mathcal{F} = \{f_0(A), f_1(A,B), f_2(A,B,C), f_3(A,C,D), f_4(A,D,E), f_5(A,E,F), f_6(B,C,G)\}$$

ii. Condition on observed G: $f_7 = (f_6)_{G=g}$

$$\mathcal{F} = \{f_0(A), f_1(A,B), f_2(A,B,C), f_3(A,C,D), f_4(A,D,E), f_5(A,E,F), f_7(B,C)\}$$

iii. Condition on observed E: $f_8 = (f_4)_{G=g}, f_9 = (f_5)_{G=g}$

$$\mathcal{F} = \{f_0(A), f_1(A,B), f_2(A,B,C), f_3(A,C,D), f_8(A,D), f_9(A,F), f_7(B,C)\}$$

iv. Sum out A from product of $f_0, f_1, f_2, f_3, f_8, f_9$: $f_{10} = \sum_A (f_0 \times f_1 \times f_2 \times f_3 \times f_8 \times f_9)$

$$\mathcal{F} = \{f_{10}(B,C,D,F), f_7(B,C)\}$$

v. Sum out C from product of f_{10}, f_7: $f_{11} = \sum_C (f_{10} \times f_7)$

$$\mathcal{F} = \{f_{11}(B,D,F)\}$$

vi. Sum out D from f_{11}: $f_{12} = \sum_D f_{11}$

$$\mathcal{F} = \{f_{12}(B,F)\}$$

vii. Sum out F from f_{12}: $f_{13} = \sum_F f_{12}$

$$\mathcal{F} = \{f_{13}(B)\}$$

viii. Normalize by division: $query(B) = f_{13}(B) \Big/ \sum_B f_{13}(B)$.

(b) **[15 points]** List the factors that would be created, and the operations used to create them, by running the variable elimination algorithm on this belief network to answer the query $P(B|G,E)$. Use the variable ordering G, E, F, D, C, B, A

i. Construct factors for each conditional probability table:

$$\mathcal{F} = \{f_0(A), f_1(A,B), f_2(A,B,C), f_3(A,C,D), f_4(A,D,E), f_5(A,E,F), f_6(B,C,G)\}$$

ii. Condition on observed G: $f_7 = (f_6)_{G=g}$

$$\mathcal{F} = \{f_0(A), f_1(A,B), f_2(A,B,C), f_3(A,C,D), f_4(A,D,E), f_5(A,E,F), f_7(B,C)\}$$

iii. Condition on observed E: $f_8 = (f_4)_{G=g}, f_9 = (f_5)_{G=g}$

$$\mathcal{F} = \{f_0(A), f_1(A,B), f_2(A,B,C), f_3(A,C,D), f_8(A,D), f_9(A,F), f_7(B,C)\}$$

iv. Sum out F from f_9: $f_{10} = \sum_F f_9$

$$\mathcal{F} = \{f_0(A), f_1(A,B), f_2(A,B,C), f_3(A,C,D), f_8(A,D), f_9(A), f_7(B,C)\}$$

v. Sum out D from product of f_3, f_8: $f_{11} = \sum_D (f_3 \times f_8)$

$$\mathcal{F} = \{f_0(A), f_1(A,B), f_2(A,B,C), f_{11}(A,C), f_9(A), f_7(B,C)\}$$

vi. Sum out C from product of f_2, f_{11}, f_7: $f_{12} = \sum_C (f_2 \times f_{11} \times f_7)$

$$\mathcal{F} = \{f_0(A), f_1(A,B), f_{12}(A,B), f_9(A)\}$$

vii. Sum out A from product of f_0, f_1, f_{12}: $f_{13} = \sum_A (f_0 \times f_1 \times f_{12} \times f_9)$

$$\mathcal{F} = \{f_{13}(B)\}$$

viii. Normalize by division: $query(B) = f_{13}(B) \Big/ \sum_B f_{13}(B)$.

(c) [**5 points**] Which of the two given variable orderings is more efficient for this query? Justify your answer. You may assume that the domain of each variable is the same size.

The ordering in question (3b) is the most efficient. Assuming that the domain of every variable is the same size, we can see this without even computing the exact number of operations required. Both orderings induce 8 steps:

i. Identical number of operations between the two orderings.
ii. Identical number of ops
iii. Identical number of ops
iv. (a) sum out of factor with **5** variables
   (b) sum out of factor with **2** variables
v. (a) sum out of factor with **4** variables
   (b) sum out of factor with **3** variables
vi. (a) sum out of factor with 3 variables
   (b) sum out of factor with 3 variables
vii. (a) sum out of factor with 2 variables
   (b) sum out of factor with 2 variables

In step (iv), the ordering in question (3a) requires the construction, by multiplying 6 factors, of a factor with 5 variables (which must then be summed out of). This step will dominate the rest in both time and space requirements. In step (v), the ordering in question (3a) constructs and sums out of a factor with 4 variables, compared to a factor with 3 variables in ordering (3b). In all other steps, the factors created have the same number of variables.

Computing the actual operation counts is also acceptable for full marks.

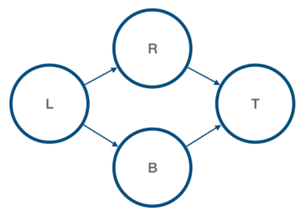
4. **(Causal inference)**

   Consider the following causal model containing random variables {L, R, B, T}, with $\text{dom}(T) = \{high, low\}$, $\text{dom}(B) = \{many, few\}$, and all other variables having domain $\{true, false\}$.

   The variable L indicates that the parents in a house like to read. The variable R indicates that the parents in a house read to the children in the house. The variable B indicates whether there are few or many books in the house. The variable T indicates whether the children in the house get high or low scores on reading tests.

   Parents who like to read (L) are more likely to read to (R). Parents who like to read (L) are also more likely to have lots of books (B) in their house. Both of being read to (R) and having lots of books (B) in the house have a causal influence on a child's performance on reading tests (T).

   (a) **[10 points]** Draw a directed graph representing the causal model.



   (b) **[3 points]** What factorization is represented by the causal model of question (4a)?

$$P(T, R, B, L) = P(T \mid B, R)P(B \mid L)P(R \mid L)P(L)$$

   Note: The order of the factors does not matter.

   (c) **[2 points]** Give an expression for the observational query

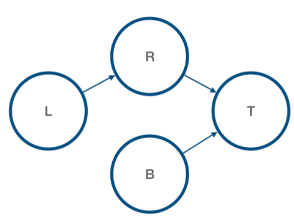
$$P(T = high \mid B = many)$$

   using the factors listed in question (4b).

$$P(T = high \mid B = many) = \frac{\sum_{r,\ell} P(T = high \mid B = many, R = r)P(B = many \mid L = \ell)P(R = r \mid L = \ell)P(L = \ell)}{\sum_{r,\ell,t} P(T = t \mid B = many, R = r)P(B = many \mid L = \ell)P(R = r \mid L = \ell)P(L = \ell)}$$

   (d) **[5 points]** Draw a directed graph representing the post-intervention distribution for the causal query

$$P(T = high \mid do(B = many)).$$



(e) **[3 points]** Give an expression for the causal query

$$P(T = high \mid do(B = many))$$

using the factors listed in question (4b).

$$P(T = high \mid do(B = many)) = \frac{\sum_{r,\ell} P(T = high \mid B = many, R = r)P(R = r \mid L = \ell)P(L = \ell)}{\sum_{r,\ell,t} P(T = t \mid B = many, R = r)P(R = r \mid L = \ell)P(L = \ell)}$$

Submission

The assignment you downloaded from eClass is a single ZIP archive which includes this document as a PDF *and* its LaTeX source.

Each assignment is to be submitted electronically via eClass by the due date. **Your submission must be a a single PDF file containing your answers.**

To generate the PDF file with your answers you can do any of the following:

- insert your answers into the provided LaTeX source file between `\begin{answer}` and `\end{answer}`. Then run the source through LaTeX to produce a PDF file;
- print out the provided PDF file and legibly write your answers in the blank spaces under each question. Make sure you write as legibly as possible for we cannot give you any points if we cannot read your hand-writing. Then scan the pages and include the scan in your ZIP submission to be uploaded on eClass;
- use your favourite text processor and type up your answers there. Make sure you number your answers in the same way as the questions are numbered in this assignment.