
CMPUT 261, Fall 2026
Assignment #2
SOLUTIONS GUIDE

1. (Probability theory; 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 25% of the time.

- (a) [15 points] 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 $P(C = gun \mid B = true)$:

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

- (b) [5 points] Suppose that the metal detector beeps when someone walks through it. With what probability is that person carrying coins?

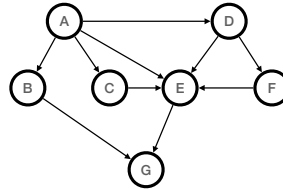
This question can be answered using the same approach as in 1a.

We now ask $P(C = coins \mid B = true)$. Note that the denominator will be identical; we only change the numerator to be $P(B = true \mid C = coins)P(C = coins) = .80 \times .30$.

Thus, $P(C = coins \mid B = true) = \frac{.80 \times .30}{.429} \approx .559$.

2. (Belief networks; 25 points)

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



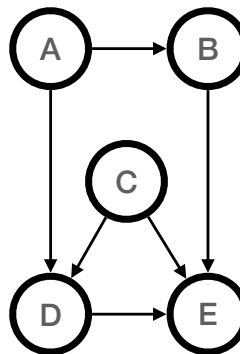
In this network,

- A has no parents
- B has parent A
- C has parent A
- D has parent A
- E has parents A, C, D , and F
- F has parent D
- G has parents B and E

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

Note: The order of the factors does not matter.

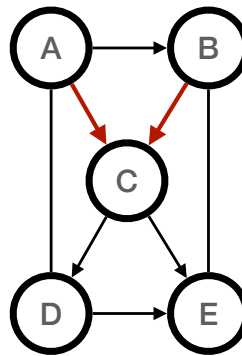
- (b) [5 points] Draw a belief network that is consistent with a joint distribution that factors as $P(A, B, C, D, E) = P(A)P(B|A)P(C)P(D|A, C)P(E|B, C, D)$, or list each of the nodes of the network along with its parents.



Or, equivalently:

- A has no parents
- B has parent A
- C has no parents
- D has parents A and C
- E has parents B, C , and D

- (c) [5 points] Draw (or list nodes and parents for) another belief network that is *different* from the network of question (2b) but *also* consistent with the factoring $P(A, B, C, D, E) = P(A)P(B|A)P(C)P(D|A, C)P(E|B, C, D)$. (i.e., a different belief network that is guaranteed to be correct for any joint distribution that can be correctly represented by this factoring).



Or, equivalently:

- *A* has no parents
- *B* has parent *A*
- *C* has parents *A* and *B*
- *D* has parents *A* and *C*
- *E* has parents *B*, *C*, and *D*

(Any acyclic graph that adds extra edges to the original graph will be consistent. Any belief network that represents a chain-rule factoring of the distribution is also consistent.)

- (d) [3 points] Suppose that every random variable in the joint distribution of question (2b) has a domain containing 8 elements. How many rows are needed to list the full joint distribution in an explicit table?

8^5

($8^5 - 1$ is also acceptable for full marks)

- (e) [7 points] Suppose that every random variable in the joint distribution of question (2b) has a domain containing 8 elements. How many rows in total are needed to list the conditional probability tables for your belief network representation?

$$P(E|B, C, D) : 8^4 \quad (7 \times 8^3 \text{ also acceptable})$$

$$P(D|A, C) : 8^3 \quad (7 \times 8^2 \text{ also acceptable})$$

$$P(C) : 8 \quad (7 \text{ also acceptable})$$

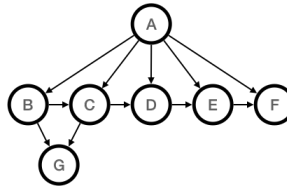
$$P(B|A) : 8^2 \quad (7 \times 8 \text{ also acceptable})$$

$$P(A) : 8 \quad (7 \text{ also acceptable}).$$

Total of $4096 + 512 + 8 + 64 + 8 = 4688$.

($3584 + 448 + 7 + 56 + 7 = 4102$ also acceptable)

3. (**Variable Elimination**; 75 points) Consider the belief network below.



In this belief network,

- A has no parents
- B has parent A
- C has parents A and B
- D has parents A and C
- E has parents A and D
- F has parents A and E
- G has parents B and C

The graph has the following properties:

- $\text{dom}(V) = \{0, 1\}$ for each variable V
 - $\Pr(A = 1) = 0.75$
 - For all $V \neq A$, the probability that $V = 1$ is $1 - 2^{-q}$, where $q = \sum_{W \in \text{parents}(V)} W$.
(In other words, the more of V 's parents have value 1, the more likely $V = 1$).
- (a) [5 points] Write a table that fully specifies the conditional probability distribution for variable D in the above belief network.

D	A	C	$\Pr(D \mid A, C)$
0	0	0	1.0
0	0	1	0.5
0	1	0	0.5
0	1	1	0.25
1	0	0	0.0
1	0	1	0.5
1	1	0	0.5
1	1	1	0.75

(b) [60 points] Implement variable elimination in Python 3 by editing the provided `ve.py` as follows:

- Edit the `Factor` class to implement missing functionality.
- Edit the `var_elimination` function to implement variable elimination.
- Edit the function `example_bn` to return a collection of factor objects that represent the belief network above.

Run the code with the command `python3 ve.py` to compute the results of the query $\Pr(B \mid G = 0, E = 1)$ using variable order G, E, A, B, C, D, F .

What is the probability $\Pr(B = 1 \mid G = 0, E = 1)$?

Approximately 0.2994350282485876.

Note: Different rounding choices are acceptable here.

- (c) [10 points] Give an example of a variable order that would be more efficient. Explain why your ordering is more efficient.

The order G, E, B, C, D, F, A would be more efficient; it is identical except that it puts A at the end instead of the beginning of the un-conditioned variables. This is because the original factors for B, C, D, F all involve A , so eliminating A early means constructing a factor that multiplies several factors together, yielding an intermediate factor $f(A, D, F, C, B)$, which has 2^5 rows. If we instead do A at the end, it will only require multiplying together a bunch of 2-row tables instead.

Any order that puts A later is sufficient for full marks if it comes with a plausible explanation. I want the students to be thinking about the size of the factors being created and the number of operations being performed; note that they do *not* have to actually tally up the number of operations in order to get full marks.