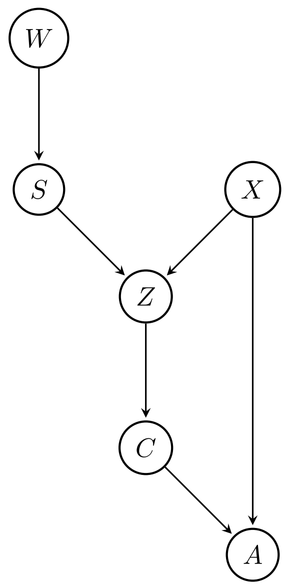
CS 188 Spring 2024 | Introduction to Artificial Intelligence | HW 4B Sol.

Solutions for HW 4B

Q1. [30 pts] Bayes Net

Flying a quadcopter can be modeled using a Bayes Net with the following variables:

- W (weather) $\in$ {clear, cloudy, rainy}
- S (signal strength) $\in$ {strong, medium, weak}
- X (true position) $= (x, y, z, \theta)$ where x, y, z **each** can take on values $\in \{0, 1, 2, 3, 4\}$ and θ can take on values $\in \{0°, 90°, 180°, 270°\}$
- Z (reading of the position) $= (x, y, z, \theta)$ where x, y, z **each** can take on values $\in \{0, 1, 2, 3, 4\}$ and θ can take on values $\in \{0°, 90°, 180°, 270°\}$
- C (control from the pilot) $\in$ {forward, backward, rotate left, rotate right, ascend, descend} (6 controls in total)
- A (smart alarm to warn pilot if that control could cause a collision) $\in$ {bad, good}



(a) Representation

(i) [3 pts] What is N_x, where N_x is the domain size of the variable X? Please explain your answer.

Answer: $N_x = 500$

Explanation: $5 * 5 * 5 * 4 = 500$

(ii) [4 pts] Please list **all** of the Conditional Probability Tables that are needed in order to represent the Bayes Net above. Note that there are 6 of them.

$P(W), P(S|W), P(X), P(Z|S, X), P(C|Z), P(A|C, X)$

(iii) [3 pts] What is the size of the Conditional Probability Table for Z? You may use N_x in your answer.

$P(Z|S, X)$ has size 750000 or $3 * (N_x)^2$

Now, assume that we look at this setup from the perspective of Spencer – a spectator who can observe A and W. Spencer observes A=bad and W=clear, and he now wants to infer the signal strength. In BN terminology, he wants to calculate $P(S|A = \text{bad}, W = \text{clear})$.

(b) [5 pts] Inference by Enumeration

If Spencer chooses to solve for this quantity using inference by enumeration, what are the different probability terms that need to be multiplied together in the summation?

Solution 1: There are six product terms in the sum.
$P(S|A,W) = \frac{P(S,A,W)}{P(A,W)}$

We know $P(S,A,W) = \sum_c \sum_x \sum_z P(W,S,z,x,A,c)$
$= \sum_c \sum_x \sum_z P(W)P(S|W)P(Z|S,x)P(c|z)P(A|c,x)P(x)$

And $P(A,W) = \sum_s P(s,A,W)$.

So all 6 factors need to be multiplied together for Spencer to calculate $P(S|A,W)$.

Solution 2: We can alternatively solve $P(S|A,W) = \frac{P(S,A|W)}{P(A|W)}$.

We can do $P(S,A|W) = \sum_c \sum_x \sum_z P(S,z,x,A,c|W)$
$= \sum_c \sum_x \sum_z P(S|W)P(Z|S,x)P(c|z)P(A|c,x)P(x)$

And $P(A|W) = \sum_s P(s,A|W)$

Using this formulation, you would only need to multiply 5 factors together to calculate $P(S|A,W)$.

(c) [15 pts] Inference by Variable Elimination

Spencer chooses to solve for this quantity by performing variable elimination in the order of $Z-X-C$. Answer the following prompts to work your way through this procedure.

(1a) First, we need to eliminate Z. Which factors (from the 6 CPTs above) are involved?
$P(Z|S,X), P(C|Z)$

(1b) Describe how you eliminate the variable of interest by multiplication of those factors. What conditional probability **factor** results from this step?
Eliminating Z out of $P(z|S,X) * P(C|z)$ results in $P(C|S,X)$

(2a) Second, we need to eliminate X. Which factors are involved?
$P(A = bad|C,X), P(X), P(C|S,X)$

(2b) Describe how you eliminate the variable of interest by multiplication of those factors. What conditional probability **factor** results from this step?
Eliminating X out of $P(A = bad|C,X) * P(X) * P(C|S,X)$ results in $P(A = bad, C|S)$

(3a) Third, we need to eliminate C. Which factor/s are involved?
$P(A = bad, C|S)$

(3b) Describe how you eliminate the variable of interest by multiplication of those factors. What conditional probability **factor** results from this step?
Eliminating C out of $P(A = bad, C|S)$ results in $P(A = bad|S)$

(4) List the 3 conditional probability factors that you calculated as a result of the 3 elimination steps above, along with their domain sizes. You may use N_x in your answer. Which factor is the biggest? Is this bigger or smaller than the biggest factor from the "inference by enumeration" approach?
The 3 conditional probability factors are $P(C|S,X)$ with size $18 * N_x$, $P(A = bad, C|S)$ with size 18, $P(A = bad|S)$ with size 3.
Biggest is $P(C|S,X)$, whose size is $6 * 3 * N_x = 18 * N_x$. This is smaller than the biggest factor from inference by enumeration.

(5) List the **1** unused conditional probability factor from the 3 that you calculated above, and also list the **2** remaining conditional probability factors from the 6 original CPTs.
$P(A = bad|S), P(S|W = clear), P(W = clear)$

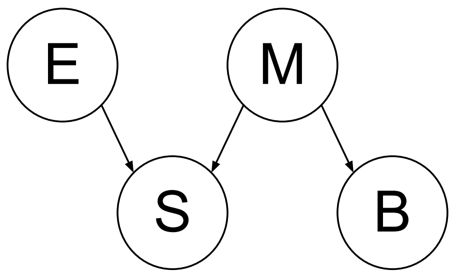
(6) Finally, let's solve for the original quantity of interest: $P(S|A = \text{bad}, W = \text{clear})$. After writing the equations to show how to use the factors from (5) in order to solve for $f(S|A = bad, W = clear)$, don't forget to write how to turn that into a probability $P(S|A = bad, W = clear)$.
Hint: use the definition of conditional probability, and use the 3 resulting factors that you listed in the previous question.
By definition of conditional probability, $P(S|A = bad, W = clear) = \frac{P(A=bad,S,W=clear)}{P(A=bad,W=clear)}$.
To get the numerator, we multiply the three resulting factors from previous part $P(A = bad|S)P(S|W = clear)P(W = clear)$ to get the joint factor $P(A = bad, S, W = clear)$. We then normalize the factor over S to get $P(A = bad, S, W = clear)$.
To get the denominator we marginalize out S from the numerator to get $P(A = bad, W = clear)$.

(Alternatively, we can do $P(S|A = bad, W = clear) = \frac{f(A=bad,S,W=clear)}{\sum_s f(A=bad,s,W=clear)}$ and not need to normalize anything.)

(d) In the following five questions, do not just answer with a numerical value. Write explicitly what probability expression you are calculating. For instance you should write $\mathbf{P(A)P(B) = 0.2}$ instead of just 0.2.

$P(E)$	
$+e$	0.4
$-e$	0.6

$P(S\|E, M)$			
$+e$	$+m$	$+s$	1.0
$+e$	$+m$	$-s$	0.0
$+e$	$-m$	$+s$	0.8
$+e$	$-m$	$-s$	0.2
$-e$	$+m$	$+s$	0.3
$-e$	$+m$	$-s$	0.7
$-e$	$-m$	$+s$	0.1
$-e$	$-m$	$-s$	0.9



$P(M)$	
$+m$	0.1
$-m$	0.9

$P(B\|M)$		
$+m$	$+b$	1.0
$+m$	$-b$	0.0
$-m$	$+b$	0.1
$-m$	$-b$	0.9

Figure 1: Bayes Net and probability tables.

10.1) (2 pts) Write an expression to compute the value of the joint distribution $P(-e, -s, -m, -b)$. What is its value?

$P(-e, -s, -m, -b) = P(-e)P(-m)P(-s| - e, -m)P(-b| - m) = (0.6)(0.9)(0.9)(0.9) = 0.4374$ by expanding the joint according to the chain rule of conditional probability.

10.2) (2 pts) Write an expression to compute the value of $P(+b)$. What is its value?

$P(+b) = P(+b| + m)P(+m) + P(+b| - m)P(-m) = (1.0)(0.1) + (0.1)(0.9) = 0.19$ by marginalizing out m according to the law of total probability.

10.3) (2 pts) Write an expression to compute the value of $P(+m| + b)$. What is its value?

$P(+m| + b) = \frac{P(+b|+m)P(+m)}{P(+b)} = \frac{(1)(0.1)}{0.19} \approx 0.5263$ by marginalizing out m according to the law of total probability.

10.4) (2 pts) Write an expression to compute the value of $P(+m| + s, +b, +e)$. What is its value?

$P(+m| + s, +b, +e) = \frac{P(+m,+s,+b,+e)}{\sum_m P(m,+s,+b,+e)} = \frac{P(+e)P(+m)P(+s|+e,+m)P(+b|+m)}{\sum_m P(+e)P(m)P(+s|+e,m)P(+b|m)} \frac{(0.4)(0.1)(1.0)(1.0)}{(0.4)(0.1)(1.0)(1.0)+(0.4)(0.9)(0.8)(0.1)} \approx 0.5814$.

10.5) (2 pts) Write an expression to compute the value of $P(+e| + m)$. What is its value?

$P(+e| + m) = P(+e) = 0.4$ The first equality holds true as E is independent of M, which can be inferred from the graph of the Bayes net.

(e) Now consider the following Bayes Net in which all the variables are **binary**. We want to compute the query $P(B, D|+f)$, so we run variable elimination with the variable elimination ordering being A, C, E, G.

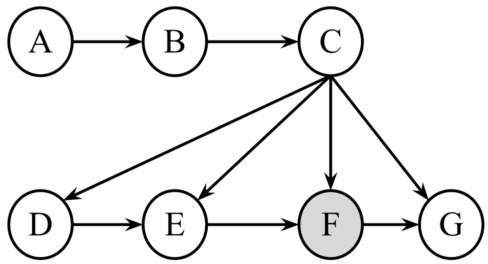


Figure 2: Bayes Net.

After observing evidence on F, we have the following factors

$$P(A),\ P(B|A),\ P(C|B),\ P(D|C),\ P(E|C,D),\ P(+f|C,E),\ P(G|C,+f)$$

When we eliminate the variable A, we then create a new factor $f_1(B) = \sum_a P(B|a)P(a)$ and the remaining factors are

$$f_1(B),\ P(C|B),\ P(D|C),\ P(E|C,D),\ P(+f|C,E),\ P(G|C,+f)$$

10.6) (2 pts) When we eliminate C next, what is the new factor f_2 we obtain? Furthermore, list the leftover factors.

$f_2(B, D, E, +f, G) = \sum_c P(c|B)P(D|c)P(E|c, D)P(+f|c, E)P(G|c, +f)$.
The remaining factors are: $f_1(B), f_2(B, D, E, +f, G)$.

10.7) (2 pts) When we eliminate E next, what is the new factor f_3 we obtain? Furthermore, list the leftover factors.

$f_3(B, D, +f, G) = \sum_e f_2(B, D, E, +f, G)$.
The remaining factors are: $f_1(B), f_3(B, D, +f, G)$.

10.8) (2 pts) When we eliminate G next, what is the new factor f_4 we obtain? Furthermore, list the leftover factors.

$f_4(B, D, +f) = \sum_g f_3(B, D, +f, G)$.
The remaining factors are: $f_1(B), f_4(B, D, +f)$.

10.9) (2 pts) How can we compute $P(B, D|+f)$ using the factors form **10.8)**? Explain in a couple of sentences.

Join $f_1 f_4$ to obtain $P(B, D, +f)$ and normalize it to get $P(B, D| + f)$. More concretely, $P(b, d| + f) = \frac{f_1(b)f_4(b,d,+f)}{\sum_{b',d'} f_1(b')f_4(b',d',+f)}$.

10.10) (1 pts) Between f_1, f_2, f_3 and f_4 which is the largest factor (or equivalently whose table has the most rows)?

$f_2(B, D, E, +f, G)$ is the largest factor generated. It has 4 variables, hence $2^4 = 16$ entries.

10.11) (4 pts) For the same query $P(B,\ D| + f)$, find a new variable elimination ordering that minimizes the size of the largest factor created during the process. Fill in your ordering and the generated factors on the table below. For instance, if we were to follow the original ordering you would write on the left and right columns of the first row of the table A and $f_1(B)$ respectively, and so on. *Hint: Your largest factor should be of size* 4*, i.e. have only two variables.*

Variable Eliminated	Factor Generated
A	$f_1(B)$
G	$f_2(C, +f)$
E	$f_3(C, D, +f)$
C	$f_4(B, D, +f)$

Note: multiple orderings are possible. In particular in this case all orderings with E and G before C are correct.