Feedback — Bayesian Network Fundamentals

You achieved a score of 10.00 out of 11.00

Please check our grading policy under "Course Logistics" before submitting the quiz. The quiz isn timed - you can save your answers halfway and come back again later.

Question 1

Factor product. Let X, Y be binary variables, and let Z be a variable that takes on values 1, 2, or 3.

If $\phi_1(X,Y)$ and $\phi_2(Y,Z)$ are the factors shown below, compute the selected entries (marked by a " the factor $\psi(X,Y,Z) = \phi_1(X,Y) \cdot \phi_2(Y,Z)$, giving your answer according to the ordering of assignments to variables as shown below.

You may separate the 3 entries of the factor by commas, e.g., an answer of "0.1, 0.2, 0.3" means tha $\psi(1,1,2) = 0.1$, $\psi(1,2,1) = 0.2$ and $\psi(2,1,3) = 0.3$

X	Y	$\phi_1(X,Y)$
1	1	0.7
1	2	0.1
2	1	0.4
2	2	0.1

$\times$

Y	Z	$\phi_2(Y,Z)$
1	1	0.2
1	2	0.8
1	3	0.5
2	1	0.0
2	2	0.9
2	3	0.3

=

X	Y	Z	$\psi(X,Y,Z)$
1	1	1	
1	1	2	?
1	1	3	
1	2	1	?
1	2	2	
1	2	3	
2	1	1	
2	1	2	
2	1	3	?
2	2	1	
2	2	2	
2	2	3	

0.56, 0.00, 0.20

Your Answer	Score	Explanation

0.56	✔	0.33	
0.00	✔	0.33	
0.20	✔	0.33	
Total		1.00	

Question 2

Factor marginalization. Let X, Z be binary variables, and let Y be a variable that takes on values 1 or 3.

If $\phi(X, Y, Z)$ is the factor shown below, compute the entries of the factor $\psi(X, Z) = \sum_Y \phi(X, Y, Z$ giving your answer according to the ordering of assignments to variables as shown below.

As before, you may separate the 4 entries of the factor by commas.

X	Y	Z	$\phi(X, Y, Z)$
1	1	1	14
1	1	2	60
1	2	1	40
1	2	2	27
1	3	1	42
1	3	2	85
2	1	1	4
2	1	2	59
2	2	1	54
2	2	2	3
2	3	1	96
2	3	2	30

X	Z	$\psi(X, Z)$
1	1	?
1	2	?
2	1	?
2	2	?

96, 172, 154, 92

Your Answer		Score	Explanation
96	✔	0.25	
172	✔	0.25	
154	✔	0.25	

92	✔	0.25
Total		1.00

Question 3
Factor reduction. Let X, Z be binary variables, and let Y be a variable that takes on values 1, 2, or

Now say we observe $Y = 3$. If $\phi(X, Y, Z)$ is the factor shown below, compute the missing entries of reduced factor $\psi(X, Z)$ given that $Y = 3$, giving your answer according to the ordering of assignmer variables as shown below.

As before, you may separate the 4 entries of the factor by commas.

X	Y	Z	$\phi(X,Y,Z)$
1	1	1	14
1	1	2	60
1	2	1	40
1	2	2	27
1	3	1	42
1	3	2	85
2	1	1	4
2	1	2	59
2	2	1	54
2	2	2	3
2	3	1	96
2	3	2	30

X	Z	$\psi(X,Z)$
1	1	?
1	2	?
2	1	?
2	2	?

42, 85, 96, 30

Your Answer		Score	Explanation
42	✔	0.25	
85	✔	0.25	
96	✔	0.25	
30	✔	0.25	
Total		1.00	

Question 4

Properties of independent variables. Assume that A and B are independent random variables. Wh

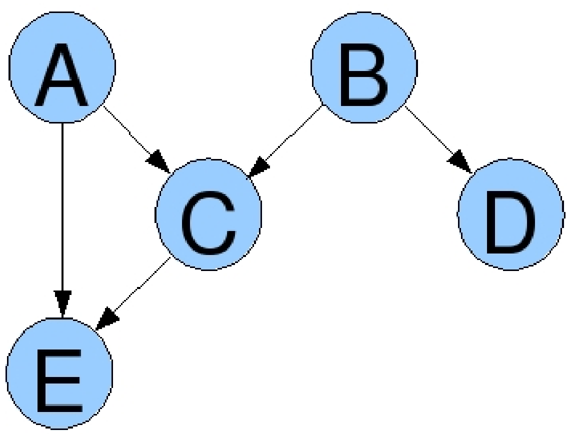
of the following options are always true? You may choose more than one option.

Your Answer		Score	Explanation
$P(A,B) = P(A) + P(B)$	✔	0.25	
✔ $P(A\|B) = P(A)$	✔	0.25	In intuitive terms, this means that the value of A is not value of B. We can derive this from $P(A,B) = P(A)$ $$\begin{aligned} P(A,B) &= P(A) \times P(B) && \text{(by definiti} \\ &= P(A\|B) \times P(B) && \text{(by chain} \\ \implies P(A\|B) &= P(A). \end{aligned}$$
✔ $P(A,B) = P(A) \times P(B)$	✔	0.25	This is the standard definition of independence.
$P(A) + P(B) = 1$	✔	0.25	
Total		1.00	

Question 5

Independencies in a graph. Which pairs of variables are independent in the graphical model below

given that none of them have been observed?

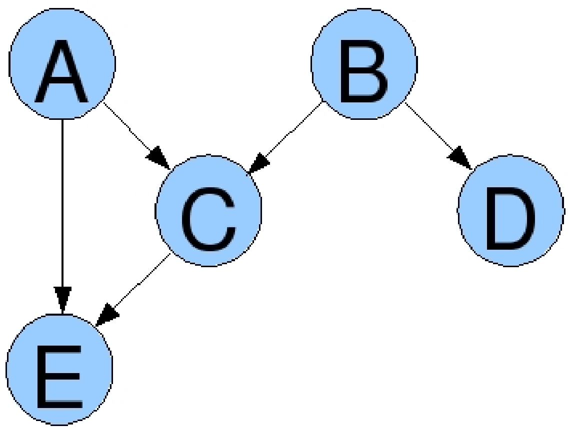


Your Answer		Score	Explanation
None - there are no pairs of independent variables.	✔	0.20	
C, D	✔	0.20	There is an active trail connecting C and D that goes through B.
✔ A, B	✔	0.20	There are no active trails between A and B so they are independent.
A, C	✔	0.20	There is a directed edge from A to C.

B, E	✔	0.20	There is an active trail connecting B and E that goes through C.
Total		1.00	

Question 6

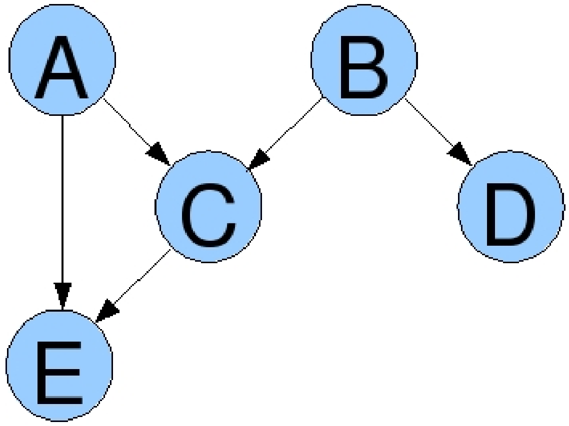
***Independencies in a graph.** (An asterisk marks a question that is more challenging. Congratulatio you get it right!) Now assume that the value of E is known. (E is observed. A, B, C, and D are not observed.) Which pairs of variables (not including E) are independent in the same graphical model, g E?



Your Answer		Score	Explanation
A, C	✔	0.14	There is a directed edge from A to C.
A, B	✔	0.14	Observing E activates the V-structures around C and E Hence, influence can flow from A to B through C; know B will now affect the probabilities of A taking on each of values.
B, D	✔	0.14	There is a directed edge from B to D.
✔ None - given E, there are no pairs of variables that are independent.	✔	0.14	Observing E activates the V-structures around C and E giving rise to active trails between every pair of variable the network.
B, C	✔	0.14	There is a directed edge from B to C.
D, C	✔	0.14	Influence can flow along the active trail $D \leftarrow B \rightarrow C$.
A, D	✔	0.14	Observing E activates the V-structures around C and E Hence, influence can flow from A to B through C, and therefore from A to D through C and B.
Total		1.00	

Question 7

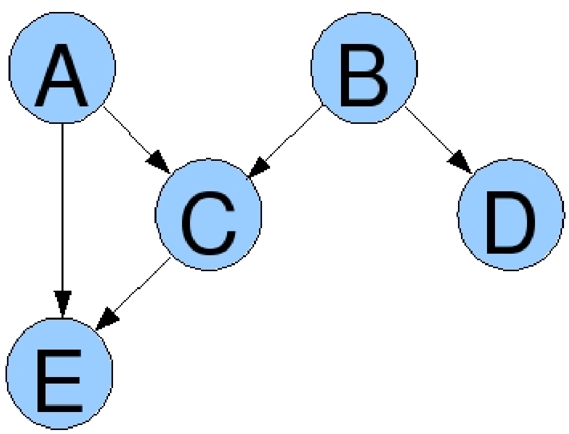
Factorization. Given the same model as above, which of these is an appropriate decomposition of t joint distribution $P(A, B, C, E)$?



Your Answer		Score	Explanation
$P(A, B, C, E) = P(A)P(B)P(C\|A, B)P(E\|A, C)$	✔	1.00	We can read off the appropri the graph by examining the p in the graph: A and B have child of A, B and E is a chilc $P(A, B, C, E) = P(A)P(B$
Total		1.00	

Question 8

Independent parameters. How many independent parameters are required to uniquely define the C of E (the conditional probability distribution associated with the variable E) in the same graphical moc as above, if A, B, and D are binary, and C and E have three values each?



Your Answer		Score	Explanation
8	✖	0.00	In a Bayesian network, the conditional probability distribution associated with a variable is the conditional probability distribution of that variable gi its parents. When calculating the number of free parameters, the number possible values for the different variables should be multiplied, not addec
Total		0.00	

Question 9

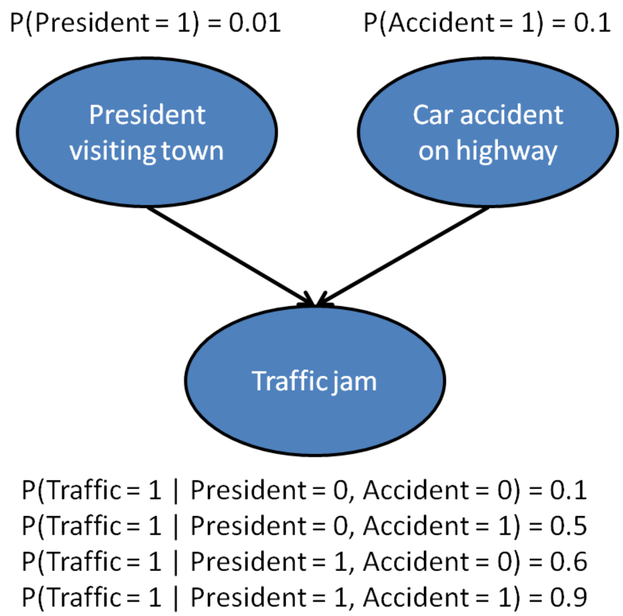
I-maps. I-maps can also be defined directly on graphs as follows. Let $I(G)$ be the set of independen
encoded by a graph G. Then G_1 is an I-map for G_2 if $I(G_1) \subseteq I(G_2)$.

Which of the following statements about I-maps are true? You may select more than one option.

Your Answer		Score	Explanation
☐ An I-map is a function f that maps a graph G to itself, i.e., $f(G) = G$.	✔	0.25	This is an identity function, not an I-map.
☐ A graph K is an I-map for a graph G if and only if K and G are identical, i.e., they have exactly the same nodes and edges.	✔	0.25	K is an I-map for G if K and G are identical, bu there might be other I-maps for G.
☐ A graph K is an I-map for a graph G if and only if K encodes exactly the same independencies as G.	✔	0.25	This is the definition of a perfect map. All perf maps are I-maps, but the converse need not hold.
☑ A graph K is an I-map for a graph G if and only if all of the independencies encoded by K are also encoded by G.	✔	0.25	K is an I-map for G if K does not make independence assumptions that are not true i An easy way to remember this is that the complete graph, which has no independencie an I-map of all distributions.
☐ I-maps are Apple's answer to Google Maps.	✘	0.00	This is not true -- yet!
Total		1.00	

Question 10

***Inter-causal reasoning.** Consider the following model for traffic jams in a small town, which we ass
can be caused by a car accident, or by a visit from the president (and the accompanying security motorcade).



Calculate P(Accident = 1 | Traffic = 1) and P(Accident = 1 | Traffic = 1, President = 1). Separate your answers with a comma, e.g., an answer of "0.15, 0.25" means that P(Accident = 1 | Traffic = 1) = 0.1 P(Accident = 1 | Traffic = 1, President = 1) = 0.25. Round your answer to two decimal places.

0.34, 0.14

Your Answer		Score	Explanation
0.34	✔	0.50	
0.14	✔	0.50	
Total		1.00	

To calculate the required values, we can apply Bayes' rule. For instance,

$$P(A=1|T=1,P=1) = \frac{P(A=1,T=1,P=1)}{P(T=1,P=1)}$$
$$= \frac{P(A=1,T=1,P=1)}{P(A=0,T=1,P=1)+P(A=1,T=1,P=1)}.$$

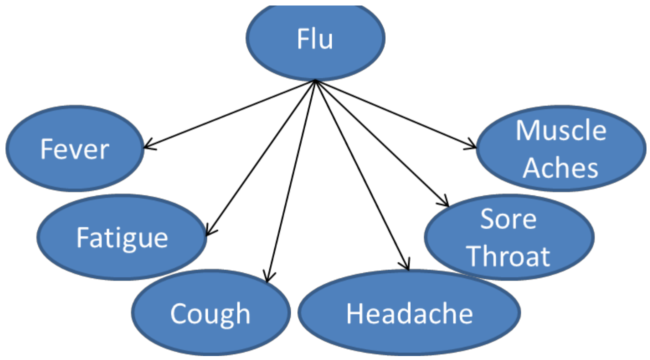
We can then use the chain rule of Bayesian networks to substitute the correct values in, e.g.,

$$P(A=1,T=1,P=1) = P(P=1)\times P(A=1)\times P(T=1|P=1,A=1)$$

This example of inter-causal reasoning meshes well with common sense: if we see a traffic jam, the probability that there was a car accident is relatively high. However, if we also see that the presiden visiting town, we can reason that the president's visit is the cause of the traffic jam; the probability th there was a car accident therefore drops correspondingly.

Question 11

***Naive Bayes.** Consider the following Naive Bayes model for flu diagnosis:



Which of the following statements are true in this model?

Your Answer		Score	Explanation
Say we observe that 1000 people have a headache (and possibly other symptoms), out of which 500 people have the flu (and possibly other symptoms), and 500 people have a fever (and possibly other symptoms). We would expect that approximately 250 people with a headache also have both the flu and a fever.	✔	0.25	Even after observing the Headache variable, there is still an active probability of someone with a headache also having a flu is depend a fever as well. For example, if someone has a flu, he could be mo irrespective of whether he has a headache or not. We therefore cannot estimate $P(Flu = 1, Fever = 1 \| Headache$ probabilities $P(Flu = 1 \| Headache = 1)$ and $P(Fever = 1 \| Hea$
✔ Say we observe that 500 people have a headache (and possibly other symptoms) and 500 people have a fever (and possibly other symptoms). Without more information, we cannot estimate how many people have both a	✔	0.25	Without having observed the Flu variable, there is an active trail fro probability of someone having a headache (without observing flu s probability of the same person having a fever. For example, if som more likely to have the flu, which would correspondingly increase t as well. We therefore cannot estimate $P(Headache = 1, Fever = 1)$ fron $P(Headache = 1)$ and $P(Fever = 1)$.

have both a headache and fever.			
Say we observe that 1000 people have the flu, out of which 500 people have a headache (and possibly other symptoms) and 500 have a fever (and possibly other symptoms). We would expect that approximately 250 people with the flu also have both a headache and fever.	✔	0.25	Given that someone has the flu, whether he has a headache is ind fever. We can thus calculate: $P(Headache = 1, Fever = 1 \| Flu = 1) = P(Headache = 1 \| F$ $\approx 0.5 * 0.5$ $= 0.25.$ Since 1000 people have the flu, we can estimate that 250 of these and fever. Note that this is only an estimate: we can assert with high confiden $P(Headache = 1, Fever = 1 \| Flu = 1)$ is near to 0.25, but in ge Moreover, even if it is exactly 0.25, the number of people with the f not be exactly 250 all the time. Think of this as analogous to flippin probability of seeing a heads is exactly 0.5, in any given sequence exactly half of the coins turning up heads.
Say we observe that 500 people have a headache (and possibly other symptoms) and 500 people have a fever (and possibly other symptoms). We would expect that approximately 250 people have both a headache and fever.	✔	0.25	Without having observed the Flu variable, there is an active trail fro probability of someone having a headache (without observing flu s probability of the same person having a fever. For example, if som more likely to have the flu, which would correspondingly increase t as well. We therefore cannot estimate $P(Headache = 1, Fever = 1)$ fron $P(Headache = 1)$ and $P(Fever = 1)$.
Total		1.00	