Bayesian Network Fundamentals

测验, 3 个问题

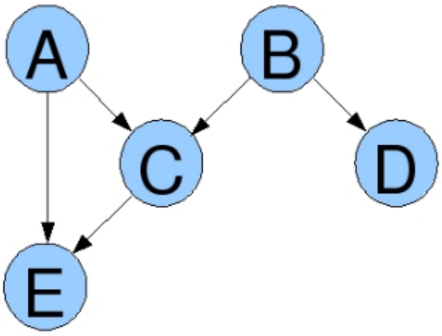
恭喜！您通过了！

下一项

1 / 1
分数

1.
Factorization.

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



- $P(A, B, C, E) = P(A)P(B)P(C|A)P(C|B)P(E|A)P(E|C)$
- $P(A, B, C, E) = P(A)P(B)P(C)P(E)$
- $P(A, B, C, E) = P(A)P(B)P(A, B|C)P(A, C|E)$
- $P(A, B, C, E) = P(A)P(B)P(C|A, B)P(E|A, C)$

正确
We can read off the appropriate factorization from the graph by examining the parents of each variable in the graph: A and B have no parents, while C is a child of A, B and E is a child of A, C.

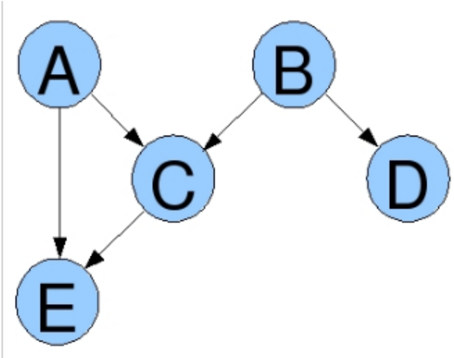
This gives us $P(A, B, C, E) = P(A)P(B)P(C|A, B)P(E|A, C)$.

1 / 1
分数

2.

Independent parameters.

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



If you haven't come across the term before, here's a brief explanation: A multinomial distribution over m possibilities $x_1, \ldots, x_m$ has m parameters, but $m-1$ independent parameters, because we have the constraint that all parameters must sum to 1, so that if you specify $m-1$ of the parameters, the final one is fixed. In a CPD $P(X|Y)$, if X has m values and Y has k values, then we have k distinct multinomial distributions, one for each value of Y, and we have $m-1$ independent parameters in each of them, for a total of $k(m-1)$. More generally, in a CPD $P(X|Y_1, \ldots, Y_r)$, if each Y_i has k_i values, we have a total of $k_1 \times \ldots \times k_r \times (m-1)$ independent parameters.

Example: Let's say we have a graphical model that just had $X \to Y$, where both variables are binary. In this scenario, we need 1 parameter to define the CPD of X. The CPD of X contains two entries $P(X=0)$ and $P(X=1)$. Since the sum of these two entries has to be equal to 1, we only need one parameter to define the CPD.

Now we look at Y. The CPD for Y contains 4 entries which correspond to: $P(Y=0|X=0), P(Y=1|X=0), P(Y=0|X=1), P(Y=1|X=1)$. Note that $P(Y=0|X=0)$ and $P(Y=1|X=0)$ should sum to one, so we need 1 independent parameter to describe those two entries; likewise, $P(Y=0|X=1)$ and $P(Y=1|X=1)$ should also sum to 1, so we need 1 independent parameter for those two entries.

Therefore, we need 1 independent parameter to define the CPD of X and 2 independent parameters to define the CPD of Y.

- 3
- 18
- 8
- 17
- 11
- 6
- 12

正确

In a Bayesian network, the conditional probability distribution associated with a variable is the conditional probability distribution of that variable given its parents. There are 6 possibilities for the values of E's parents (A is binary and C takes on 3 values). For each of these possibilities, there are 3 possible values for E, which corresponds to 2 free parameters (since the 3 numbers have to sum to 1). So there are 6*2=12 total free parameters.

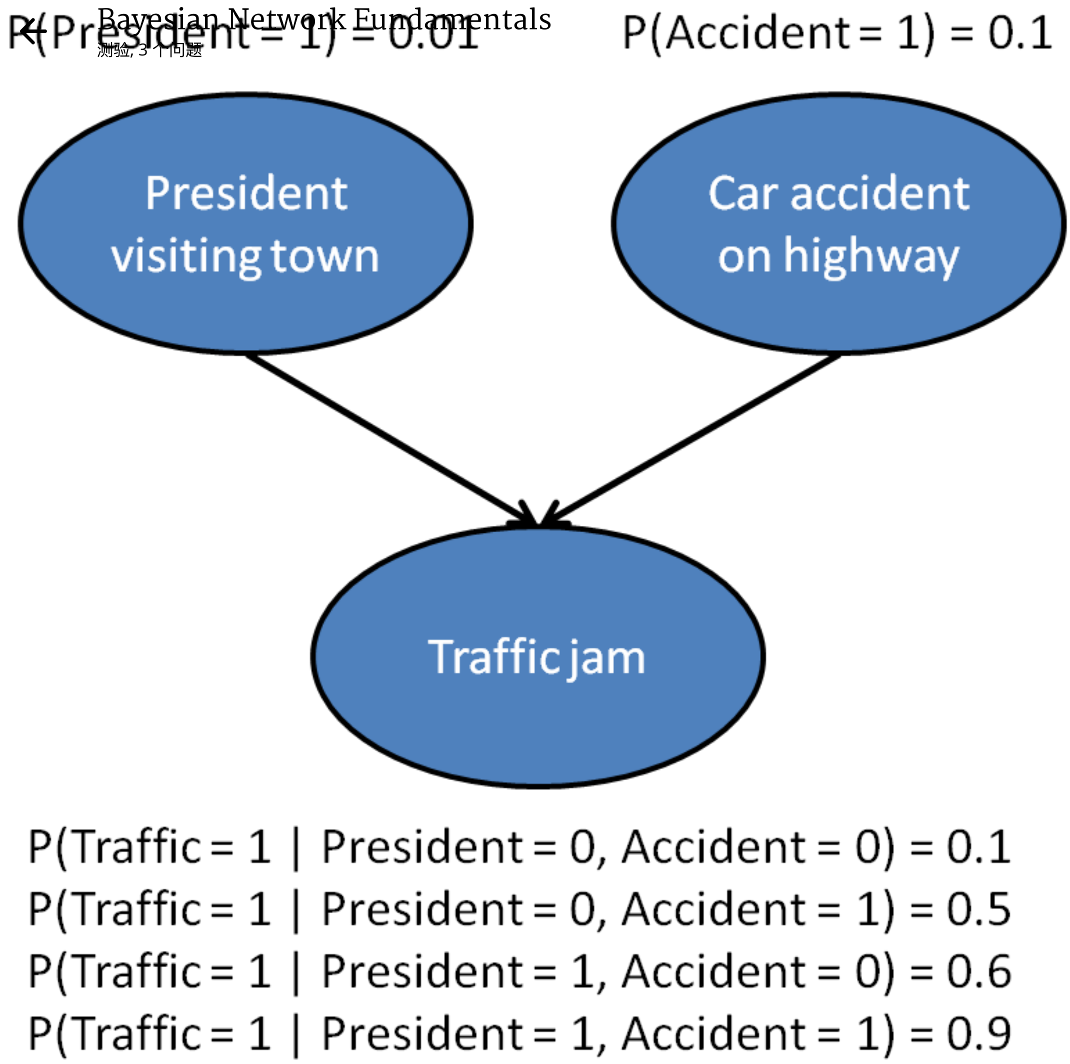
---

1 / 1
分数

3.

***Inter-causal reasoning.**

Consider the following model for traffic jams in a small town, which we assume 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 space, e.g., an answer of

0.15 0.25

means that P(Accident = 1 | Traffic = 1) = 0.15 and P(Accident = 1 | Traffic = 1, President = 1) = 0.25. Round your answers to two decimal places and write a leading zero, like in the example above.

0.35 0.14

正确回答

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

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

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 president is visiting town, we can reason that the president's visit is the cause of the traffic jam; the probability that there was a car accident therefore drops correspondingly.

Bayesian Network Fundamentals
测验, 3 个问题