

## Bayesian Network Quiz

1. Why is a Bayesian Network a Directed Graph?
  - A. edges in our graph represent the probability of a node depends on its parent. Within Bayes net parents are not dependent on their children's value, so our edge must have a strict direction from parent to child.
2. What is the reason Bayesian Networks aren't allowed to form cycles?
  - A. When performing inference we take the product of all the conditional probabilities given a data sample. If we were then to allow cycle in our Bayes Net when performing inference, we would find that we are finding the  $P[A|B]$  and  $P[B|A]$  which result in circular logic making our calculation impossible.
3. Described uses of conditional Probability table.
  - A. Stores the likelihood values conditioned for every possible combination of values of the parent, which is essential for calculating fast inference.
4. What are the benefits of variable elimination?
  - A. Variable elimination allows us to dramatically increase the speed of our inference calculation but removing nondependent information.
5. Write down the steps for the Parameter learning algorithm.
  - A.
    - For every node
    - For the values of the node
    - For all combinations of parent values
    - Look at each sample in the training set
    - If the sample satisfies N and S increment your numerator and denominator by +1
    - Else if the sample only satisfies S increment your denominator by +1

6. What technique should you use when your conditional Probability table is empty and you do not have a large data set?

A. Laplace Smoothing

7. Write the equation for determining the branch weights in the chow-liu algorithm.

A. 
$$I(X,Y) = \sum_{x \in \text{values}(X)} \sum_{y \in \text{values}(Y)} P(x,y) \log_2 \frac{P(x,y)}{P(x)P(y)}$$

8. What is the time complexity for complete search and score, structure learning?  
(Super exponential)(exponential)(linear)

A. (Super exponential)

9. What are the necessary steps for determining the values of a missing sample given a trained Bayes Net?

A. We calculate the maximum likelihood estimation using

$$\mathbb{P}[S] = \prod_{i=1}^x \mathbb{P}[n_i]$$

This allows us to test every value which the missing data would take on and find the one with the highest probability.

10. If a node that could take on 5 different values has 4 parents which each can take on three possible values what is the size of the conditional Probability table for that node?

A.  $5 \cdot 3^4$