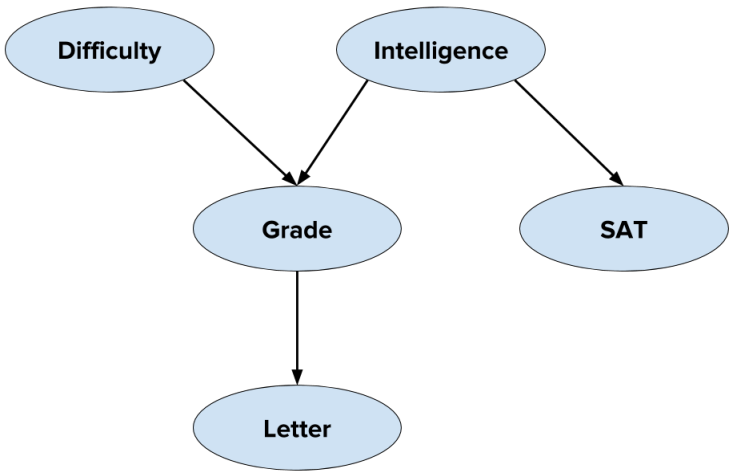
Assignment 9 (Sol.)

Introduction to Machine Learning
Prof. B. Ravindran

1. Here is a popular toy graphical model. It models the grades obtained by a student in a course and it's implications. Difficulty represents the difficulty of the course and intelligence is an indicator of how intelligent the student is, SAT represents the SAT scores of the student and Letter presents the event of the student receiving a letter of recommendation from the faculty teaching the course.



Given this graphical model, which of the following statements are true? Note - More than one can be correct.

(a) Given the grade, difficulty and letter are indpendent variables.
(b) Given grade, difficulty and intelligence are independent
(c) Without knowing any information, Difficulty and Intelligence are independent.
(d) Given the intelligence, SAT and grades are independent.

Solution: A, C, D
To check independence between pairs of variables, first check all the paths between the pair of nodes. We have to ensure that all the paths should be blocked between the nodes. We call a path blocked in the following cases

- The nodes which occur on the path with head to tail or tail to tail should be known.
- The nodes which occur on the path with head to head shouldn't be known.

Using this strategy we will try to evaluate each of the option next.

For option A, there is only one path between D and L, which passes through G. Since G is a node with head to tail, and G is known, hence the path is blocked which makes D, L independent. Similarly you can evaluate for all the options and reach the given solution.

2. The random variables given in the previous model are modeled as discrete variables and the corresponding CPDs are as below.

d^0	d^1
0.6	0.4

i^0	i^1
0.7	0.3

	g^1	g^2	g^3
i^0, d^0	0.3	0.4	0.3
i^0, d^1	0.05	0.25	0.7
i^1, d^0	0.9	0.08	0.02
i^1, d^1	0.5	0.3	0.2

	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

	l^0	l^1
g^1	0.1	0.9
g^2	0.4	0.6
g^3	0.99	0.01

What is the probability of i^1, d^0, g^2, s^1, l^0 occuring?

(a) 0.004608

(b) 0.003872

(c) 0.001

(d) 0.0078

Solution: A
This is straight forward computation.

$$P(i^1, d^0, g^2, s^1, l^0) = P(i^1)P(d^0)P(g^2|i^1, d^0)P(s^1|i^1)P(l^0|g^2)$$

3. Using the given example and CPD's compute the probability of following assignment, i^1, g^1, s^1, l^1 irrespective of the difficulty of the course?

(a) 0.16

(b) 0.2

(c) 0.5

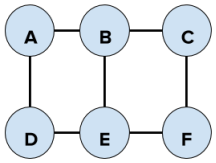
(d) 0.6

Solution: A

$$P(i^1, g^1, s^1, l^1) = P(i^1)P(s^1|i^1)P(l^1|g^1) \sum_{d}^{d=0,1} (P(d)P(g^1|i^1, d))$$

$$P(i^1, g^1, s^1, l^1) = 0.3 \times 0.8 \times 0.9 \times (0.9 \times 0.6 + 0.5 \times 0.4)$$

4. Which of the following is a valid Gibbs distribution over this graph?



   (a) $\frac{\phi(A)\phi(B)\phi(C)\phi(D)\phi(E)\phi(F)}{Z}$, where Z is the partition function.

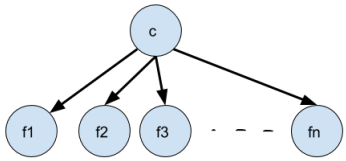
   (b) $\frac{\phi(A,B,D)\phi(C,E,F)}{Z}$, where Z is the partition function

   (c) There is no Gibbs distribution for this Markov network

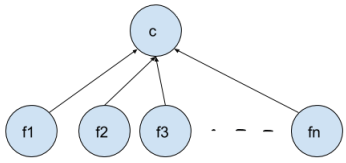
   (d) $\phi(A)\phi(B)\phi(C)\phi(D)\phi(E)\phi(F)$

   Solution: A

5. Which of the following graphical models capture the Naive Bayes assumption, where c represents the class label and f_i are the features?



   (a)



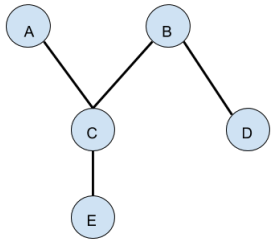
   (b)

   (c) It cannot be captured by a graphical model.

   (d) Graphical model can caputre the assumption, but the given models don't do it.

   Solution: A
   The bayes assumption states that given the class label, the features are independent. This is captured when the class label is the parent node for all the feature nodes.

6. Consider the following graphical model and identify which of the pairs of random variables are independent?



(a) A, B

(b) C, D

(c) E, D

(d) None of them.

Solution: D
All the variables are connected, and thus they are dependednt.

7. What is the value obtained on marginalizing all the variables in a graphical model?

(a) Depends on the model

(b) Some value less than 1

(c) 1

(d) Some value greater than 1

Solution: C
The graphical model represents a probability distribution which should and will sum up to 1.