Assignment - 4

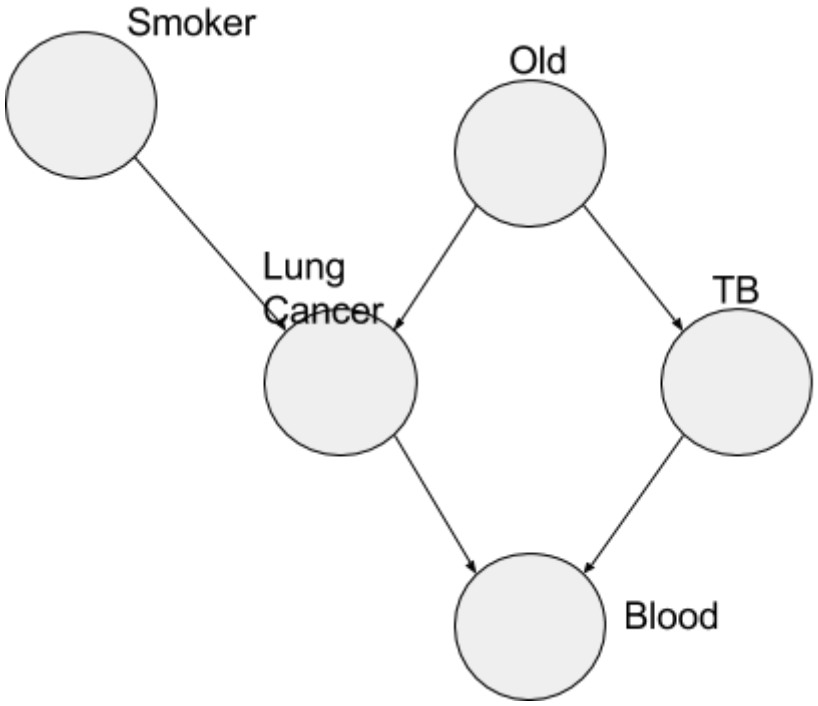
Bayesian Networks

Submission Deadline - November 3rd,11.55 PM

Name- Roll No -

1. What is the maximum number of edges in a Bayesian network with n nodes?

2.



Consider the above graph and the following given probabilities and answer the following questions.

P(Smoker) = 0.1
P(old) = 0.2
P(tb | old) = 0.05
P(tb | ~old) = 0.01
P(lungcancer | old ^ smoker) = 0.3
P(lungcancer | ~old ^ smoker) = 0.1
P(lungcancer | old ^ ~smoker) = 0.2
P(lungcancer | ~old ^ ~smoker) = 0.01
P(blood | lungcancer ^ tb) =1.0
P(blood | lungancer ^ ~tb) = 0.9
P(blood | ~lungcancer ^ tb) = 0.8
P(blood | ~lungcancer ^~tb) = 0.01

i) What is P(Smoker | lungcancer) ?
ii) What is P(blood | Smoker) ?
iii) What is P(lungcancer | old ^ smoker ^ blood)?

3. For each of the networks given in Figure 1 (a,b), do the following statements hold? Please explain your reasoning.

a) A⊥C | B, D
b) B⊥D | A, C

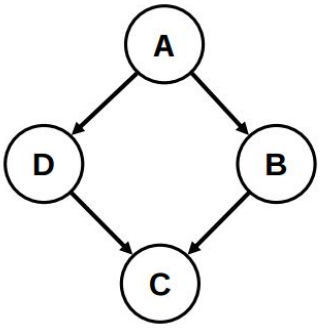


Fig 1 a

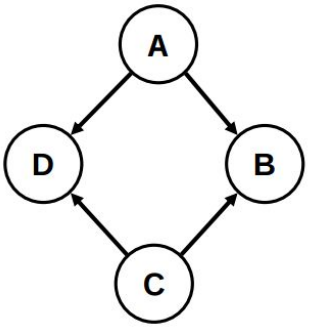


Fig 1 b

4. Construct two different Bayesian network which encode exactly the following conditional independence assumptions

A⊥C | B
A⊥D | B
C⊥D | B

5. Assume the following Bayes net and the corresponding distributions over the variables in the Bayes net

P(+a) = ¼
P(-a) = ¾

P(-b | -a) = ⅔
P(+b | -a) = ⅓
P(-b | +a) = ⅘
P(+b | +a) = ⅕

P(-d | -c) = 1/8
P(+d | -c) = 7/8
P(-d | +c) = 5/6
P(+d | +c) = 1/6

P(-c | -b) = ¼
P(+c | -b) = ¾
P(-c | +b) = ½
P(+c | +b) = ½

a)You are given the following samples

+a +b -c -d
+a -b +c -d
-a +b +c -d
-a -b +c -d
+a -b -c +d
+a +b +c -d
-a +b -c +d
-a -b +c -d

i) Assume that these samples came from performing Prior Sampling, calculate the sample estimate of P(+c)
ii) Which samples would not be used when doing Rejection sampling for the task estimating P(+c | +a,-d) . What is the sample of estimate of P(+c | +a,-d)

b) Using Likelihood Weighting Sampling to estimate P(-a | +b,-d) the following samples were obtained. Calculate the weight of each sample

i) -a +b +c -d
ii) +a +b +c -d
iii) +a +b -c -d
iv) -a +b -c -d

c) From the weighted samples in the previous question estimate P(-a | +b,-d)

d)Which query is better suited for likelihood weighting, P(D|A) or P(A|D)? Justify.

e) Recall that during Gibbs Sampling, samples are generated through an iterative process. Assume that the only evidence that is available is A=+a. Which of the following sequences below could have been generated by Gibbs sampling. Justify.

i) Sequence 1:

1) +a -b -c +d
2) +a -b -c +d
3) +a -b +c +d

ii) Sequence 2:

1) +a -b -c +d
2) +a -b -c -d
3) -a -b -c +d

iii) Sequence 3:

1) +a -b -c +d
2) +a -b -c -d
3) +a +b -c -d

iv) Sequnce 4:

1) +a -b -c +d
2) +a -b -c -d
3) +a +b -c +d