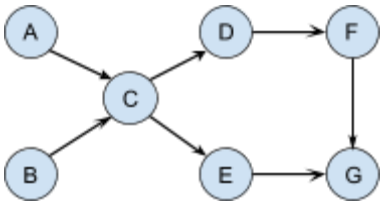
Homework 4

Edited on 12/2 in red color. Changes are marked in red color.

1. Given the following Bayes Net, answer the independence questions. If the (conditional) independence cannot be guaranteed, please write down an active path.




   a. (5 pts) Could we guarantee the independence between A and B?

   b. (5 pts) Could we guarantee the conditional independence between A and F given E?

   c. (5 pts) Could we guarantee the conditional independence between E and F, given C?

   d. (5 pts) Could we guarantee the conditional independence between E and F, given G?

   e. (5 pts) Could we guarantee the conditional independence between E and F, given C and G?

   f. (5 pts) Could we guarantee the conditional independence between E and F, given A?

2. Over the patients of COVID-19 (V) who do not have other diseases, 80% have the symptom of "shortness of breath" (S), 80% have cough (C), and 70% have the symptom of fever (H, short for high temperature). In comparison, for the patients of flu (F) who do not have other diseases, 80% have the symptom of fever (H), 60% have cough (C), and 70% have the symptom of body aches (B). Patients who have both V and F are more likely to have symptom C (90%) and symptom H (90%).

   The probabilities that a healthy person has shortness of breath, cough, fever, and body aches are 0.05, 0.10, 0.02, and 0.10 respectively. Over the whole population, infection rates of V and F are 1% and 2%

   I made up those numbers for the purposes of this HW.

a. (10 pts) Model this domain using a Bayes network, and provide a complete definition of it (graph and CPTs).

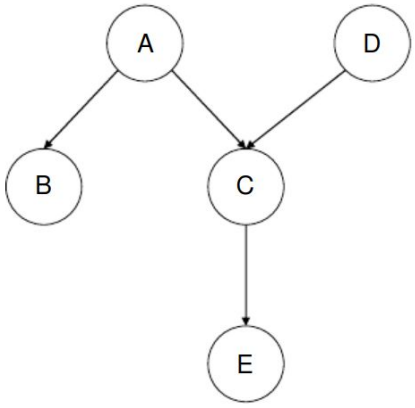
b. (5 pts) Could we guarantee the independence of V and F given the symptom of S? In other words, a patient who has the shortness of breath symptom tells a doctor that he/she has the flu. If this "flu" information adds anything to the diagnosis of COVID-19, we cannot guarantee independence.

c. (5 pts) Could we guarantee the independence of V and F given the symptom of C?

d. (5 pts) Could we guarantee the independence of V and F given the symptom of S and C?

e. (10 pts) A patient has the symptoms of cough (C) and fever (H), and has no symptom of shortness of breath (S). The doctor did not ask about body aches (B), so its value is unknown. Compute the probability that the patient has COVID-19.

3. Given the following Bayes network and conditional probabilities,




| A | D | C | P(C\|A,D) |
|---|---|---|---|
| false | false | false | 0.1 |
| false | false | true | 0.9 |
| false | true | false | 0.2 |
| false | true | true | 0.8 |
| true | false | false | 0.3 |
| true | false | true | 0.7 |
| true | true | false | 0.4 |
| true | true | true | 0.6 |

| A | B | P(B\|A) |
|---|---|---|
| false | false | 0.1 |
| false | true | 0.9 |
| true | false | 0.9 |
| true | true | 0.1 |

| C | E | P(E\|C) |
|---|---|---|
| false | false | 0.8 |
| false | true | 0.2 |
| true | false | 0.2 |
| true | true | 0.8 |

| A | P(A) |
|---|---|
| false | 0.5 |
| true | 0.5 |

| D | P(D) |
|---|---|
| false | 0.25 |
| True | 0.75 |

a. (10 pts) Compute the exact probability of P(B=true | D=true, E=false)

b. Now you are given a “sampling machine”. This machine can be used for generating a sample from distribution (p, 1-p). The machine is deterministically unstable: it returns the most likely value twice, the least value once, and repeats this process forever. In case of sampling from a uniform distribution, the sampling machine is triggered, though it always outputs “true”. For instance, if we use this machine to sample from P(A), ~~P(D), P(A),~~ P(A), P(D), P(D), P(A), P(D), P(B|A=true), P(B|A=true), P(B|A=true) in a row, the samples are true, true, ~~true,~~ false, true, true, false, false, false, true. Nodes are selected in alphabet order to break a tie: A-D-B-C-E in our case.

   i. (15 pts) Use this machine and Prior Sampling to generate 5 samples. Show the samples and then use the samples to compute P(A=true, D=true, E=true) and P(B=true | D=true, E=false).

   ii. (10 pts) Use this machine and Rejection Sampling to generate 5 samples (rejected samples do not count). Show the samples, and use the samples to compute P(B=true | D=true, E=false).