CS 188
Fall 2022

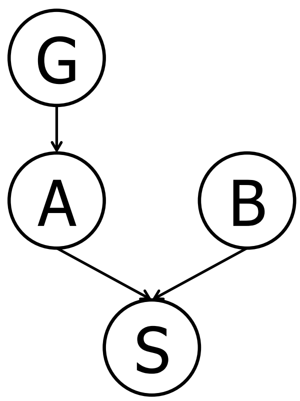
# Regular Discussion 6 Solutions

## 1 Bayes' Nets Representation and Probability

Suppose that a patient can have a symptom ($S$) that can be caused by two different diseases ($A$ and $B$). It is known that the variation of gene $G$ plays a big role in the manifestation of disease $A$. The Bayes' Net and corresponding conditional probability tables for this situation are shown below. For each part, you may leave your answer as an arithmetic expression.

| $P(G)$ | |
|---|---|
| $+g$ | 0.1 |
| $-g$ | 0.9 |

| $P(A\|G)$ | | |
|---|---|---|
| $+g$ | $+a$ | 1.0 |
| $+g$ | $-a$ | 0.0 |
| $-g$ | $+a$ | 0.1 |
| $-g$ | $-a$ | 0.9 |




| $P(B)$ | |
|---|---|
| $+b$ | 0.4 |
| $-b$ | 0.6 |

| $P(S\|A,B)$ | | | |
|---|---|---|---|
| $+a$ | $+b$ | $+s$ | 1.0 |
| $+a$ | $+b$ | $-s$ | 0.0 |
| $+a$ | $-b$ | $+s$ | 0.9 |
| $+a$ | $-b$ | $-s$ | 0.1 |
| $-a$ | $+b$ | $+s$ | 0.8 |
| $-a$ | $+b$ | $-s$ | 0.2 |
| $-a$ | $-b$ | $+s$ | 0.1 |
| $-a$ | $-b$ | $-s$ | 0.9 |

**(a)** Compute the following entry from the joint distribution:

$P(+g,+a,+b,+s) =$ $P(+g)P(+a|+g)P(+b)P(+s|+b,+a) = (0.1)(1.0)(0.4)(1.0) = 0.04$

**(b)** What is the probability that a patient has disease $A$?

$P(+a) =$ $P(+a|+g)P(+g) + P(+a|-g)P(-g) = (1.0)(0.1) + (0.1)(0.9) = 0.19$

**(c)** What is the probability that a patient has disease $A$ given that they have disease $B$?

$P(+a|+b) =$ $P(+a) = 0.19$ The first equality holds true since $A \perp\!\!\!\perp B$, which can be inferred from the graph of the Bayes' net.

**(d)** What is the probability that a patient has disease $A$ given that they have symptom $S$ and disease $B$?

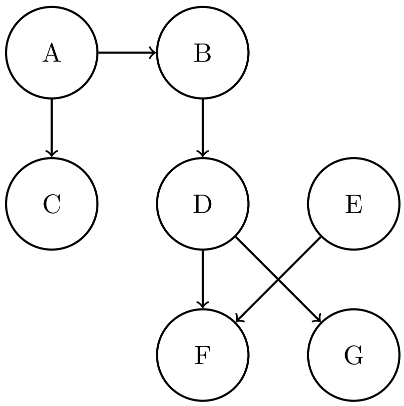
$$P(+a|+s,+b) = \frac{P(+a,+b,+s)}{P(+a,+b,+s)+P(-a,+b,+s)} = \frac{P(+a)P(+b)P(+s|+a,+b)}{P(+a)P(+b)P(+s|+a,+b)+P(-a)P(+b)P(+s|-a,+b)}$$
$$= \frac{(0.19)(0.4)(1.0)}{(0.19)(0.4)(1.0)+(0.81)(0.4)(0.8)} = \frac{0.076}{0.076+0.2592} \approx 0.2267$$

**(e)** What is the probability that a patient has the disease carrying gene variation $G$ given that they have disease $A$?

$$P(+g|+a) = \frac{P(+g)P(+a|+g)}{P(+g)P(+a|+g)+P(-g)P(+a|-g)} = \frac{(0.1)(1.0)}{(0.1)(1.0)+(0.9)(0.1)} = \frac{0.1}{0.1+0.09} = 0.5263$$

# 2 Bayes Nets: Representation

Parts (a), (b), and (c) pertain to the following Bayes' Net.




**(a)** Express the joint probability distribution as a product of terms from the Bayes Nets CPTs.

$P(A)P(C|A)P(B|A)P(D|B)P(E)P(F|D,E)P(G|D)$

**(b)** Assume each node can take on 4 values. How many entries do the factors at A, D, and F have?

A: 4 D: $4^2$ F: $4^3$

**(c)** Mark all that are guaranteed to be true:

- ☐ $B \perp\!\!\!\perp C$
- ☐ $A \perp\!\!\!\perp F$
- ☐ $D \perp\!\!\!\perp E|F$
- ■ $E \perp\!\!\!\perp A|D$
- ■ $F \perp\!\!\!\perp G|D$
- ■ $B \perp\!\!\!\perp F|D$
- ☐ $C \perp\!\!\!\perp G$
- ■ $D \perp\!\!\!\perp E$

Parts (d) and (e) pertain to the following CPTs.

| A | $P(A)$ |
|---|---|
| +a | 0.8 |
| −a | 0.2 |

| A | B | $P(B\|A)$ |
|---|---|---|
| +a | +b | 0.9 |
| +a | −b | 0.1 |
| −a | +b | 0.6 |
| −a | −b | 0.4 |

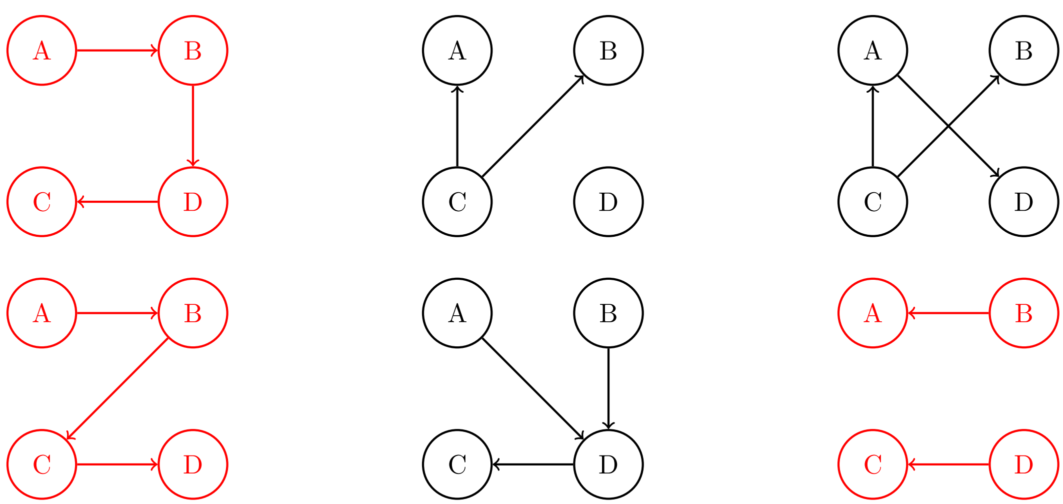
| B | C | $P(C\|B)$ |
|---|---|---|
| +b | +c | 0.8 |
| +b | −c | 0.2 |
| −b | +c | 0.8 |
| −b | −c | 0.2 |

| C | D | $P(D\|C)$ |
|---|---|---|
| +c | +d | 0.25 |
| +c | −d | 0.75 |
| −c | +d | 0.5 |
| −c | −d | 0.5 |

**(d)** State all non-conditional independence assumptions that are implied by the probability distribution tables.

From the tables, we have $A \not\perp\!\!\!\perp B$ and $C \not\perp\!\!\!\perp D$, and $B \perp\!\!\!\perp C$.

**(e)** Circle all the Bayes net(s) that can represent a distribution that is consistent with the tables given.



The question asks for Bayes Nets that **can** represent the distribution in the tables. So, in the nets we circle, the only requirement must be that A and B must not be independent, and C and D must not be independent.

The top left, bottom left, and bottom right nets have arrows between the A-B nodes and the C-D nodes, so we can circle those.

The top middle net has C and D as independent (D is not connected to anything), so we cannot circle it. The bottom middle net has A and B as independent (common effect), so we cannot circle it.

The top right net seems like it could represent the distribution, because D-separation finds that: A and B are not guaranteed to be independent (common cause), and C and D are not guaranteed to be independent (causal chain). However, since $B \perp\!\!\!\perp C$, the arrow between B and C is vacuous, thus A and B cannot actually be dependent.