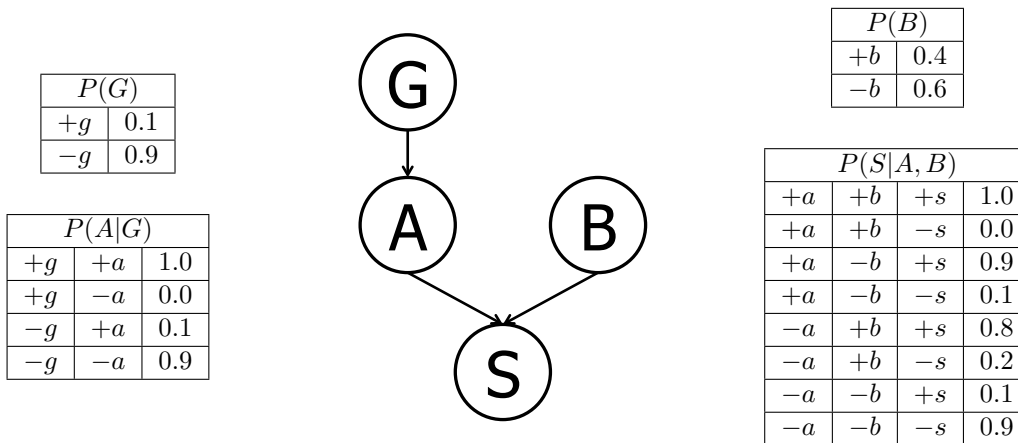


1 Bayes' Nets Representation and Probability

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



(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 ?

$$P(+a|+b) = P(+a) = 0.19$$

This equality holds true as we have $A \perp\!\!\!\perp B$ which is given in the problem. You can also infer this from the graph of the Bayes' net via d-separation, which is covered in future lectures.

(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 ?

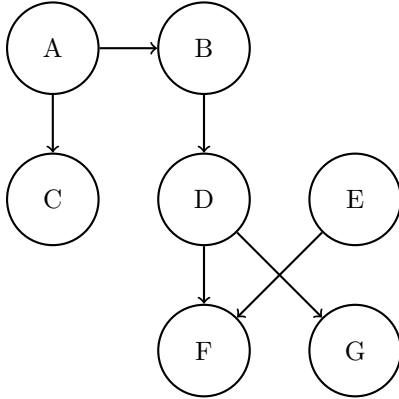
$$P(+g|+a) = \frac{P(+g)P(+a|+g)}{P(+g)P(+a|+g)+P(-g)P(+a|-g)} = \frac{(0.1)(1.0)}{(0.1)(1.0)+(0.9)(0.1)} = \frac{0.1}{0.1+0.09} = 0.5263$$

- (f) What is the probability that a patient has the disease carrying gene variation G given that they have disease B ?

$P(+g|+b) = P(+g) = 0.1$ The first equality holds true as we have $G \perp\!\!\!\perp B$, which can be inferred from the graph of the Bayes' net.

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?

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

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

- | | |
|--|--|
| <input type="checkbox"/> $B \perp\!\!\!\perp C$ | <input checked="" type="checkbox"/> $F \perp\!\!\!\perp G D$ |
| <input type="checkbox"/> $A \perp\!\!\!\perp F$ | <input checked="" type="checkbox"/> $B \perp\!\!\!\perp F D$ |
| <input type="checkbox"/> $D \perp\!\!\!\perp E F$ | <input type="checkbox"/> $C \perp\!\!\!\perp G$ |
| <input checked="" type="checkbox"/> $E \perp\!\!\!\perp A D$ | <input checked="" type="checkbox"/> $D \perp\!\!\!\perp E$ |

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

A	B	$P(B A)$	B	C	$P(C B)$	C	D	$P(D C)$
+a	+b	0.9	+b	+c	0.8	+c	+d	0.25
+a	-b	0.1	+b	-c	0.2	+c	-d	0.75
-a	+b	0.6	-b	+c	0.8	-c	+d	0.5
-a	-b	0.4	-b	-c	0.2	-c	-d	0.5

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

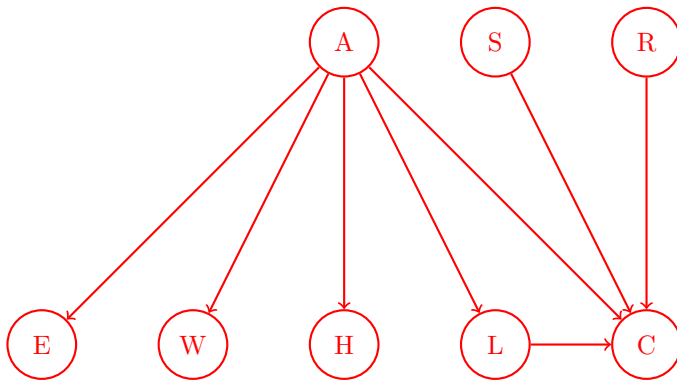
From the tables, we have $A \not\perp\!\!\!\perp B$ and $C \not\perp\!\!\!\perp D$, and $B \perp\!\!\!\perp C$.

You are building advanced safety features for cars that can warn a driver if they are falling asleep (A) and also calculate the probability of a crash (C) in real time. You have at your disposal 6 sensors (random variables):

- E : whether the driver's eyes are open or closed
- W : whether the steering wheel is being touched or not
- L : whether the car is in the lane or not
- S : whether the car is speeding or not
- H : whether the driver's heart rate is somewhat elevated or resting
- R : whether the car radar detects a close object or not

A influences $\{E, W, H, L, C\}$. C is influenced by $\{A, S, L, R\}$.

- (e) Draw the Bayes Net associated with the description above by adding edges between the provided nodes where appropriate.



Q3. Bayes Nets: Variable Elimination

	$P(A)$		$P(B A)$	$+b$	$-b$		$P(C A)$	$+c$	$-c$
$+a$	0.25		$+a$	0.5	0.5		$+a$	0.2	0.8
$-a$	0.75		$-a$	0.25	0.75		$-a$	0.6	0.4

	$P(D B)$	$+d$	$-d$		$P(E B)$	$+e$	$-e$
$+b$	0.6	0.4		$+b$	0.25	0.75	
$-b$	0.8	0.2		$-b$	0.1	0.9	

(a) Using the Bayes' Net and conditional probability tables above, calculate the following quantities:

(i) $P(+b | +a) = 0.5$

(ii) $P(+a, +b) =$
 $0.25 * 0.5 = 0.125 = \frac{1}{8}$

(iii) $P(+a | +b) =$
 $\frac{0.25 * 0.5}{0.25 * 0.5 + 0.25 * 0.75} = 0.4 = \frac{2}{5}$

(b) Now we are going to consider variable elimination in the Bayes' Net above.

(i) Assume we have the evidence $+c$ and wish to calculate $P(E | +c)$. What factors do we have initially?
 $P(A), P(B | A), P(+c | A), P(D | B), P(E | B)$

(ii) If we eliminate variable B, we create a new factor. What probability does that factor correspond to?
 $P(D, E | A)$

(iii) What is the equation to calculate the factor we create when eliminating variable B?
 $f(A, D, E) = \sum_B P(B | A) * P(D | B) * P(E | B)$

(iv) After eliminating variable B, what are the new set of factors? As in (ii), write the probabilities that the factors represent. For each factor, also provide its size.

Factor	Size after elimination
$P(A)$	2
$P(+c A)$	2
$P(D, E A)$	2^3

(v) Now assume we have the evidence $-c$ and are trying to calculate $P(A | -c)$. What is the most efficient elimination ordering? If more than one ordering is most efficient, provide any one of them. **E, D, B** or **D, E, B**

(vi) Once we have run variable elimination and have $f(A, -c)$ how do we calculate $P(+a | -c)$?
 $\frac{f(+a, -c)}{f(+a, -c) + f(-a, -c)}$ or note that elimination is unnecessary - just use Bayes' rule