

# **Section 1**

## **Causal and Bayesian Networks**

# The car start problem (causally)

---

In the morning, my car will not start. I can hear the starter turn, but nothing happens. The most probable causes are that the fuel has been stolen over night or that the spark plugs are dirty. I look at the fuel meter. It shows *half full*, so I decide to clean the spark plugs.

## Events:

- Fuel? $\{y,n\}$
- Clean spark plugs? $\{y,n\}$
- Start? $\{y,n\}$
- Fuel meter $\{full, \frac{1}{2}, empty\}$ .

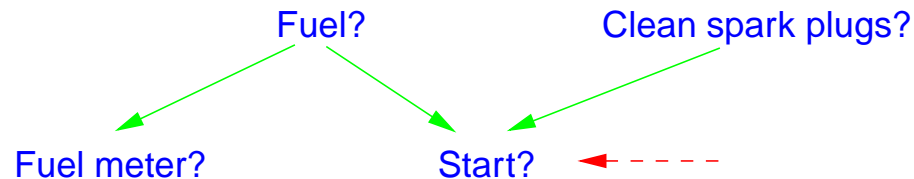
# The car start problem (causally)

In the morning, my car will not start. I can hear the starter turn, but nothing happens. The most probable causes are that the fuel has been stolen over night or that the spark plugs are dirty. I look at the fuel meter. It shows *half full*, so I decide to clean the spark plugs.

## Events:

- Fuel?{y,n}
- Clean spark plugs?{y,n}
- Start?{y,n}
- Fuel meter{full,  $\frac{1}{2}$ , empty}.

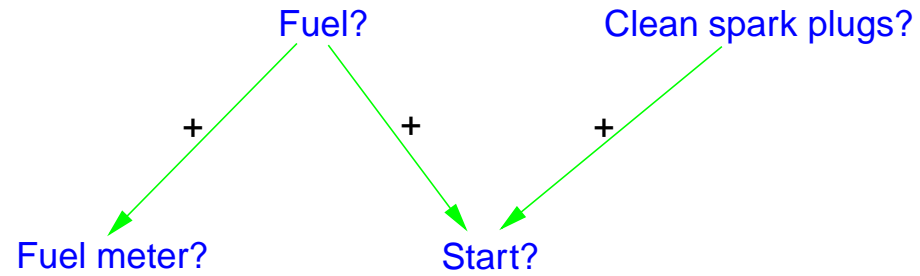
## Causal relations:



When I enter the car I have some prior belief on the various events but then start=n.

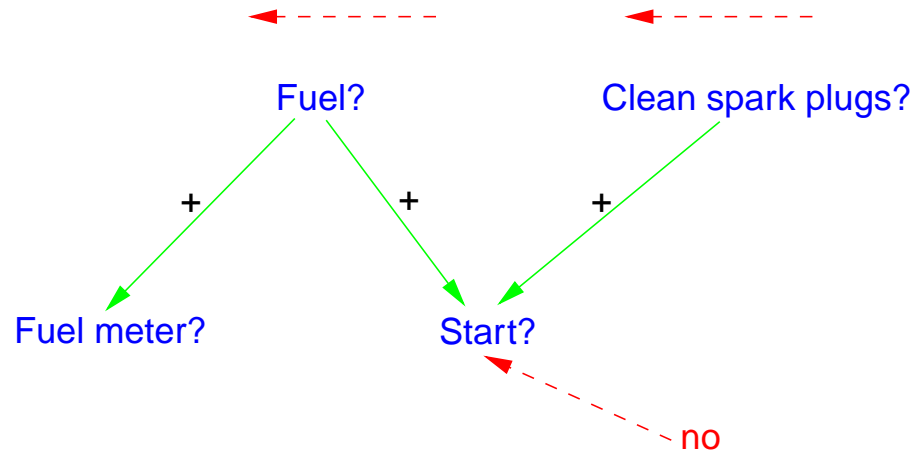
# The reasoning (explaining away)

---



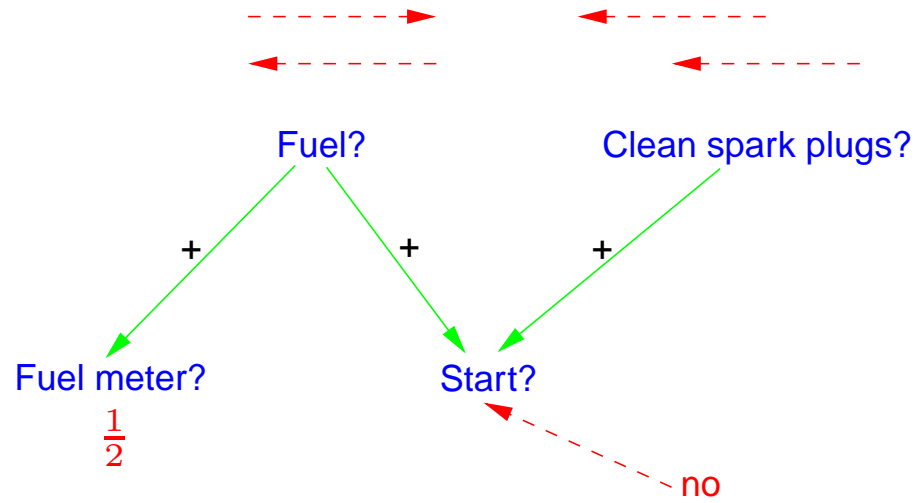
# The reasoning (explaining away)

---



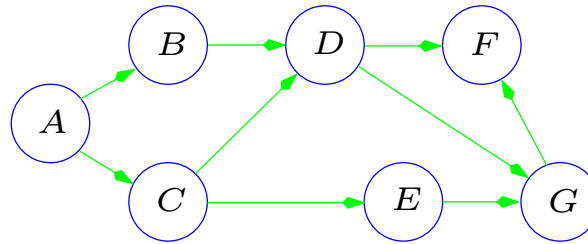
# The reasoning (explaining away)

---



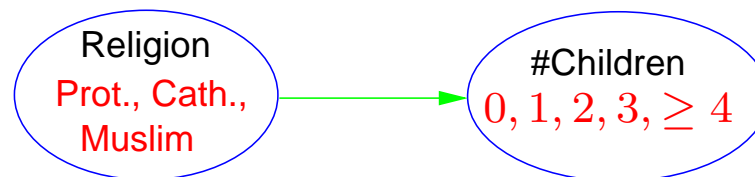
# Causal networks

A causal network is a directed acyclic graph:



- The nodes are variables with a finite set of states that are mutually exclusive and exhaustive:
  - For example  $\{y,n\}$ ,  $\{\text{red, blue, green}\}$ ,  $\{0,1,2,3,42\}$ .
- The links represent **cause – effect** relations.

For example:



All variables are in exactly one state, but we may not know which one.

# Reasoning under uncertainty 1

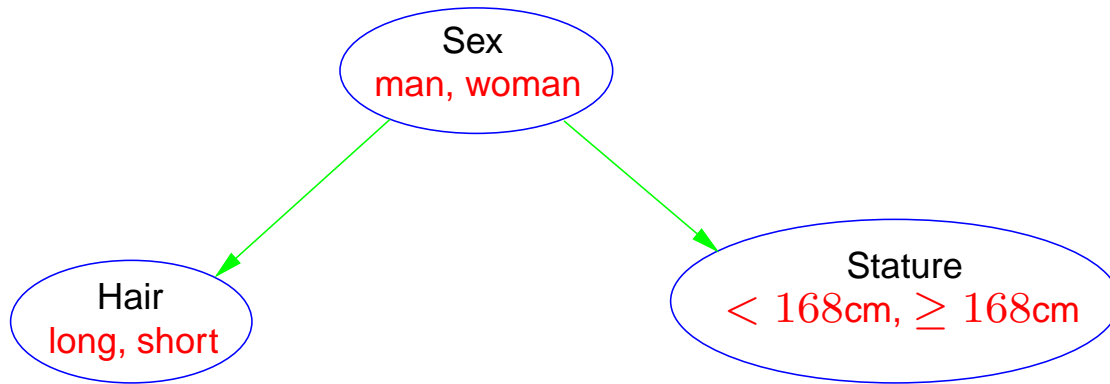
---



- Information on flooding may change the belief of *Rainfall*.
- If we know that the water level is high, then information on flooding will not change the belief of *Rainfall*.

# Reasoning under uncertainty 2

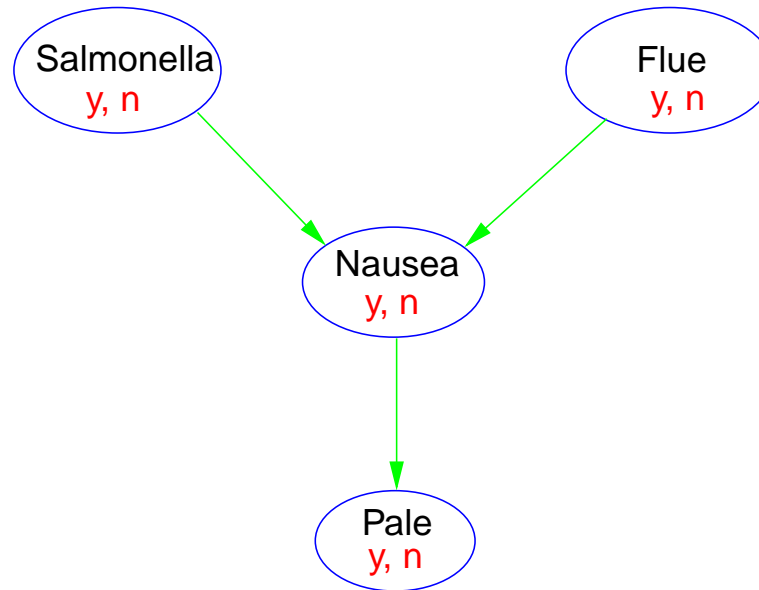
---



- Knowledge of a person's hair length may change the belief of the stature.
- If we know that it is a women, then information of hair length has no impact on the belief of her stature.

# Reasoning under uncertainty 3

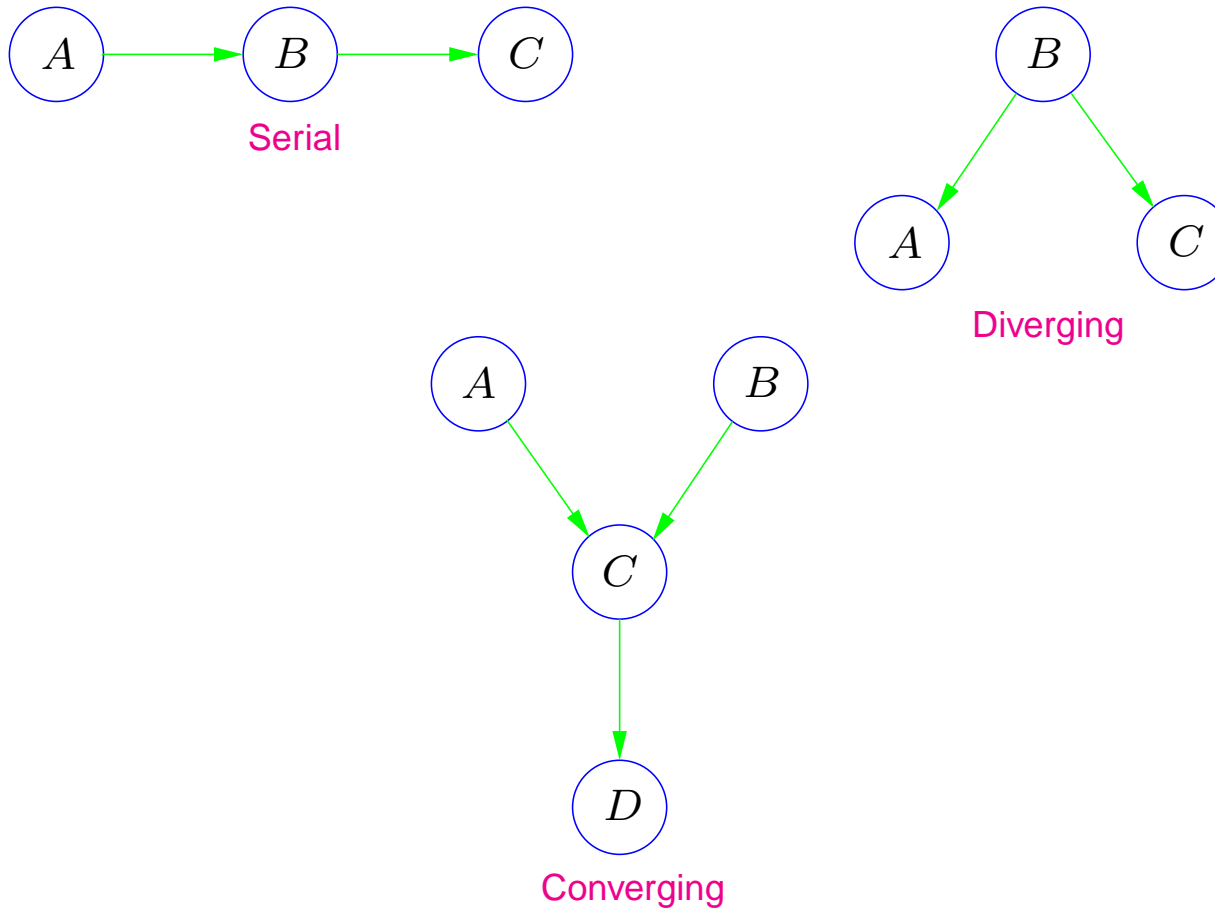
---



- Salmonella has no impact on Flue?
- If a person is **Pale**, then knowledge of *Salmonella* may change the belief on *Flue* (explaining away).

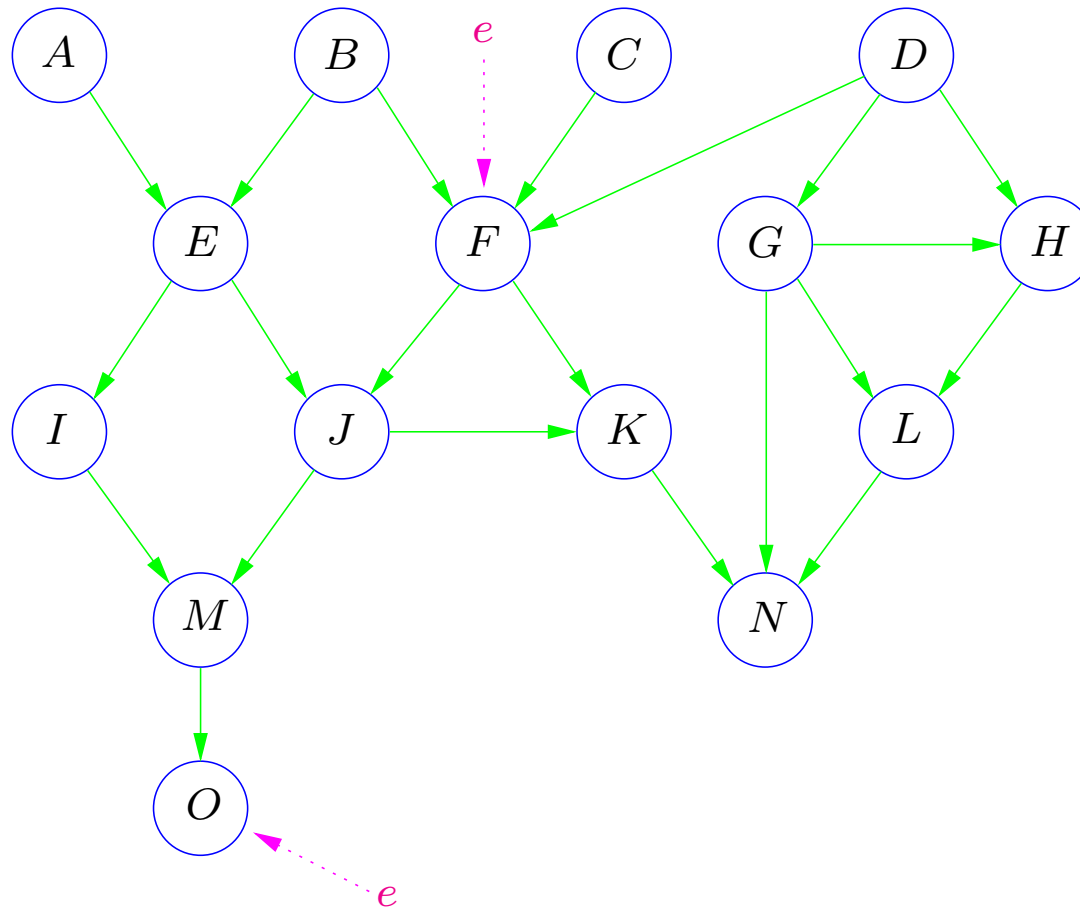
# Rules of d-separation

Relevance changes with evidence



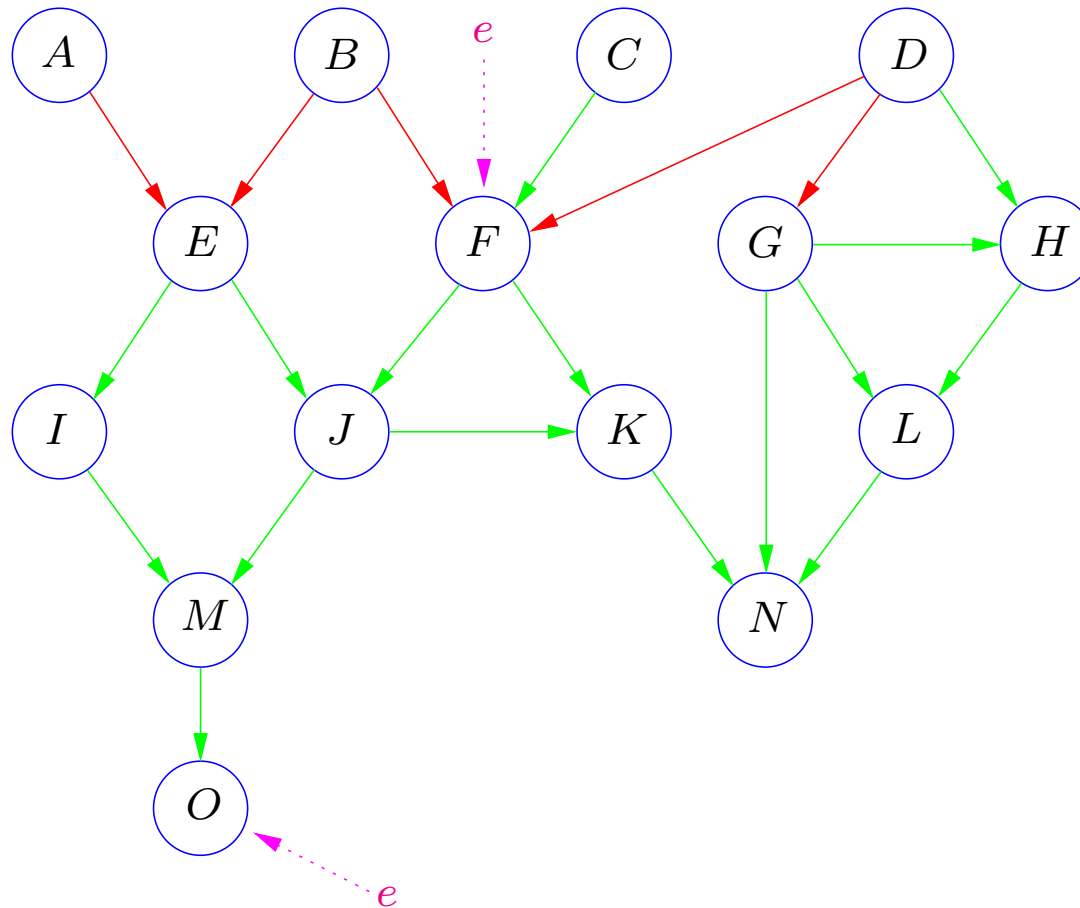
**Note:** These are general rules about reasoning under uncertainty on causal models.

# Transmission of evidence



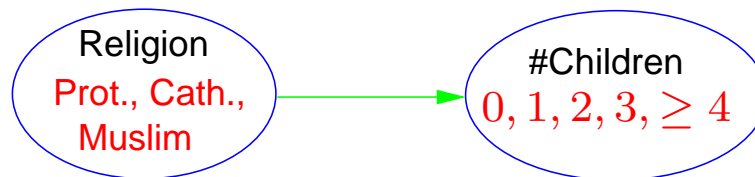
Can knowledge of  $A$  have an impact on our knowledge of  $G$ ?

# Transmission of evidence



Can knowledge of  $A$  have an impact on our knowledge of  $G$ ? yes!

# Quantification of causal networks



The strength of the **link** is represented by probabilities:

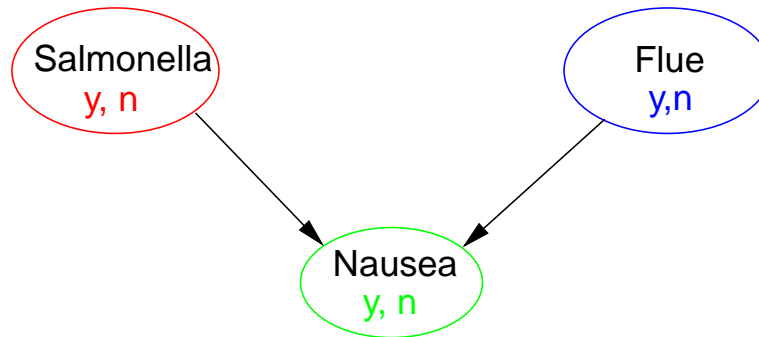
$$\begin{array}{ccc}
 P(0|p) & P(0|c) & P(0|m) \\
 P(1|p) & P(1|c) & P(1|m) \\
 P(2|p) & P(2|c) & P(2|m) \\
 P(3|p) & P(3|c) & P(3|m) \\
 P(\geq 4|p) & P(\geq 4|c) & P(\geq 4|m)
 \end{array}$$



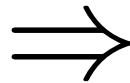
		Religion		
		p	c	m
#Children	0	0.15	0.05	0.05
	1	0.2	0.1	0.1
	2	0.4	0.2	0.1
	3	0.2	0.4	0.1
	≥ 4	0.05	0.25	0.35

$P(\#Children|Religion)$

# Several parents



$$\begin{array}{ll}
 P(y|y, y) & P(y|y, n) \\
 P(n|y, y) & P(n|y, n) \\
 P(y|n, y) & P(y|n, n) \\
 P(n|n, y) & P(n|n, n)
 \end{array}$$

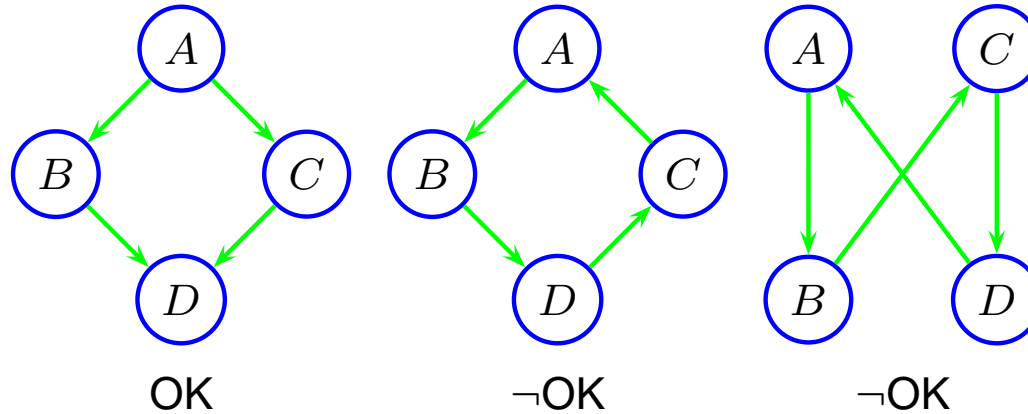


		Salmonella	
		y	n
Flue	y	(0.9, 0.1)	(0.6, 0.4)
	n	(0.8, 0.2)	(0.1, 0.9)

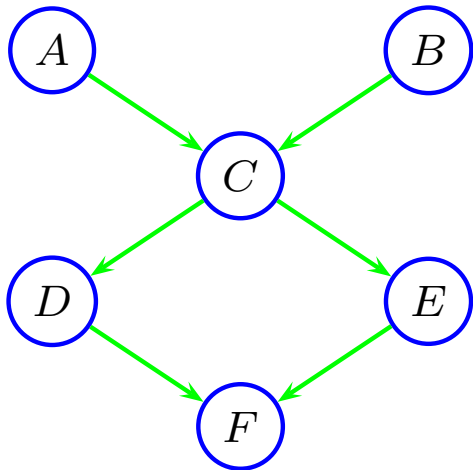
$P(\text{Nausea}|\text{Salmonella}, \text{Flue})$

# Bayesian networks

A causal network without directed cycles:



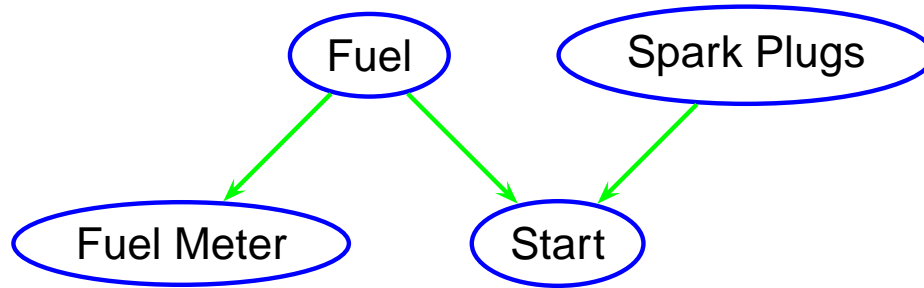
For each variable  $A$  with parents  $B_1, \dots, B_n$  there is a conditional probability table  $P(A|B_1, \dots, B_n)$ .



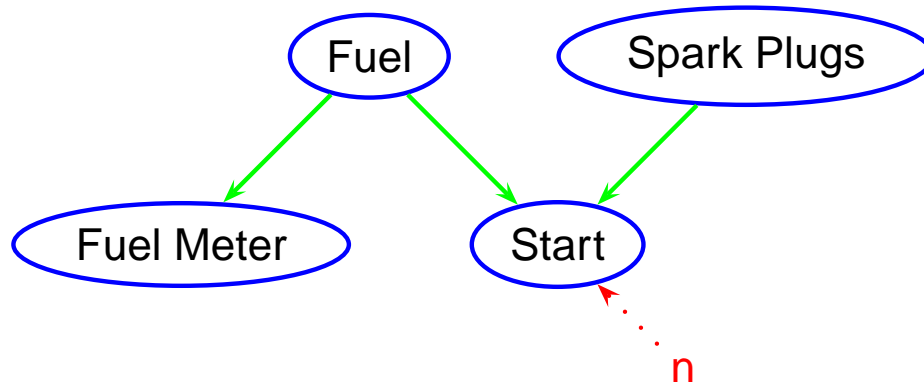
**Note:** Nodes without parents receive a prior distribution.

# Bayesian networks as a tool for reasoning

---



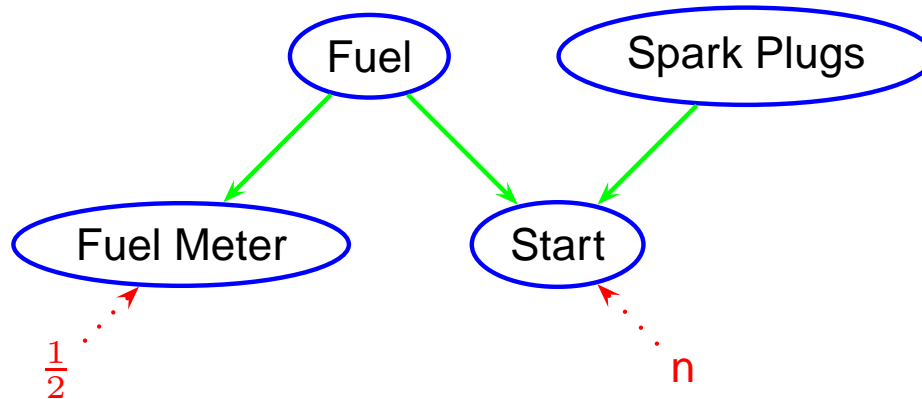
# Bayesian networks as a tool for reasoning



Consider evidence  $e_1 = (\text{Start}=n)$  and find:

- $P(\text{Spark Plugs}|e_1) = ??$
- $P(\text{Fuel}|e_1) = ??$
- $P(\text{Fuel Meter}|e_1) = ??$

# Bayesian networks as a tool for reasoning



Consider evidence  $e_1 = (\text{Start} = n)$  and find:

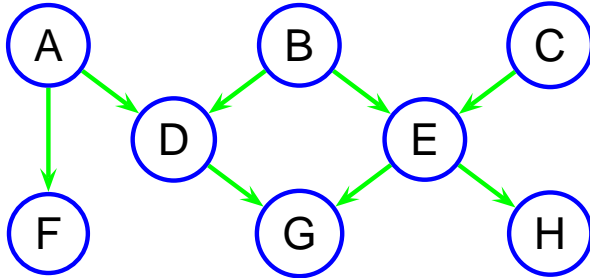
- $P(\text{Spark Plugs} | e_1) = ??$
- $P(\text{Fuel} | e_1) = ??$
- $P(\text{Fuel Meter} | e_1) = ??$

If we also have evidence  $e_2 = (\text{Fuel Meter} = \frac{1}{2})$  what is:

- $P(\text{Spark Plugs} | e_1, e_2) = ??$
- $P(\text{Fuel} | e_1, e_2) = ??$

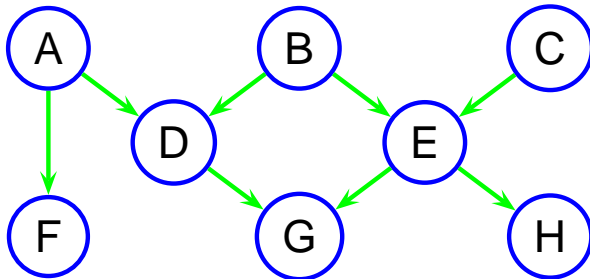
# Bayesian belief updating

---



Find  $P(B|a, f, g, h)$

# Bayesian belief updating



Find  $P(B|a, f, g, h)$

We can if we have access to  $P(a, B, C, D, E, f, g, h)$ :

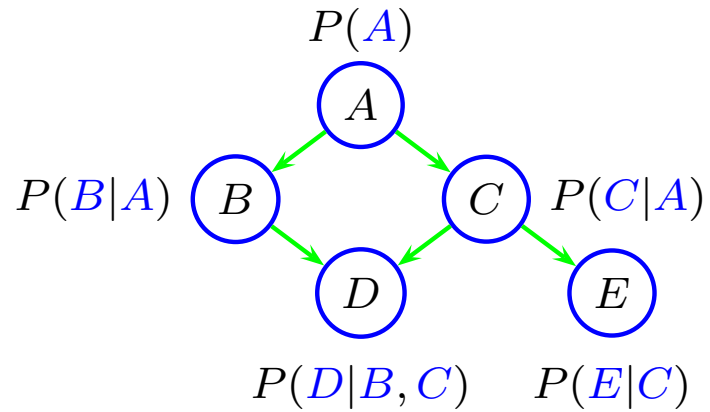
$$P(B, a, f, g, h) = \sum_{C, D, E} P(a, B, C, D, E, f, g, h)$$

$$P(B|a, f, g, h) = \frac{P(B, a, f, g, h)}{P(a, f, g, h)},$$

where

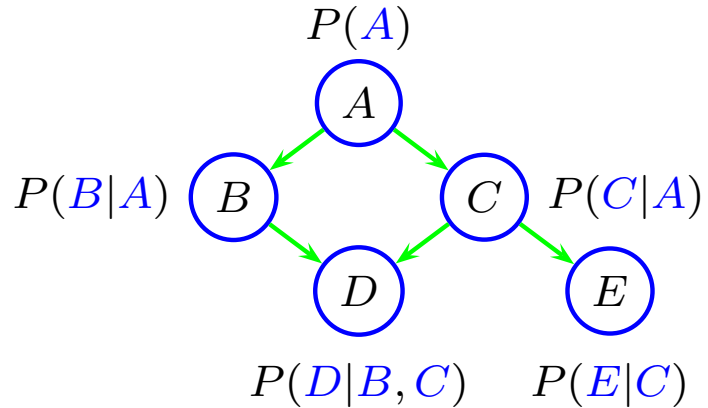
$$P(a, f, g, h) = \sum_B P(B, a, f, g, h)$$

# Joint probabilities



Calculate  $P(A, B, C, D, E)$

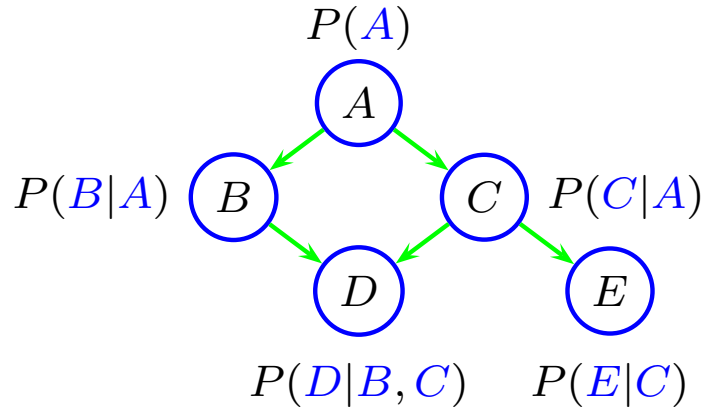
# Joint probabilities



Calculate  $P(A, B, C, D, E)$

$$P(A, B, C, D, E) = P(E|A, B, C, D)P(A, B, C, D)$$

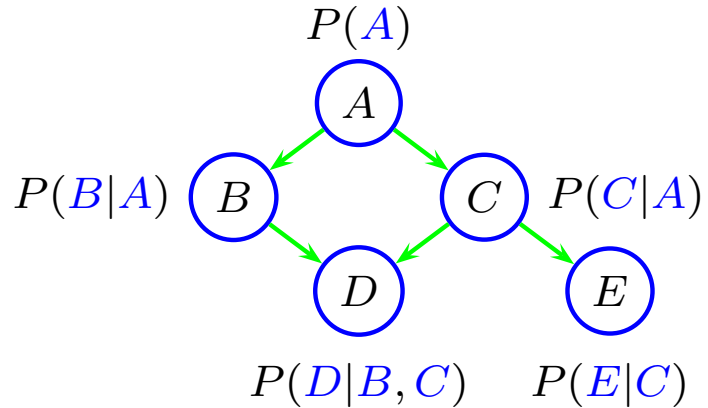
# Joint probabilities



Calculate  $P(A, B, C, D, E)$

$$\begin{aligned} P(A, B, C, D, E) &= P(E|A, B, C, D)P(A, B, C, D) \\ &= P(E|C)P(A, B, C, D) \end{aligned}$$

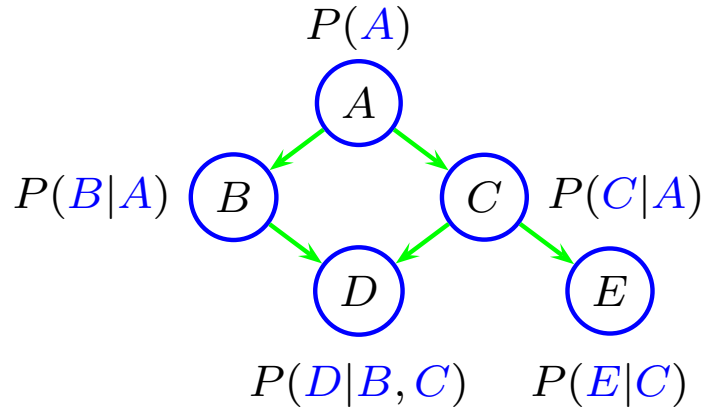
# Joint probabilities



Calculate  $P(A, B, C, D, E)$

$$\begin{aligned} P(A, B, C, D, E) &= P(E|A, B, C, D)P(A, B, C, D) \\ &= P(E|C)P(A, B, C, D) \\ &= P(E|C)P(D|A, B, C)P(A, B, C) \end{aligned}$$

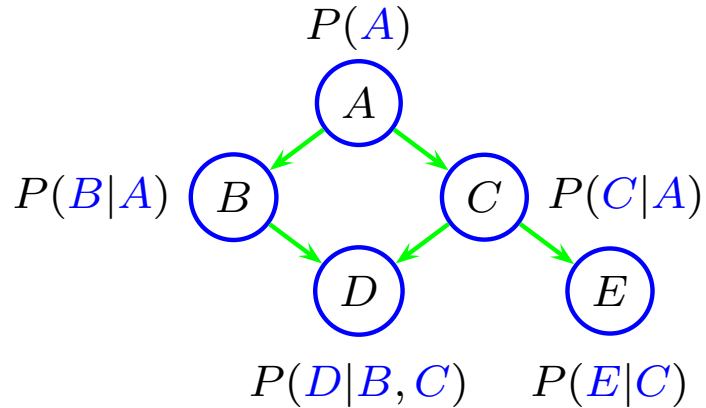
# Joint probabilities



Calculate  $P(A, B, C, D, E)$

$$\begin{aligned} P(A, B, C, D, E) &= P(E|A, B, C, D)P(A, B, C, D) \\ &= P(E|C)P(A, B, C, D) \\ &= P(E|C)P(D|A, B, C)P(A, B, C) \\ &= P(E|C)P(D|B, C)P(A, B, C) \end{aligned}$$

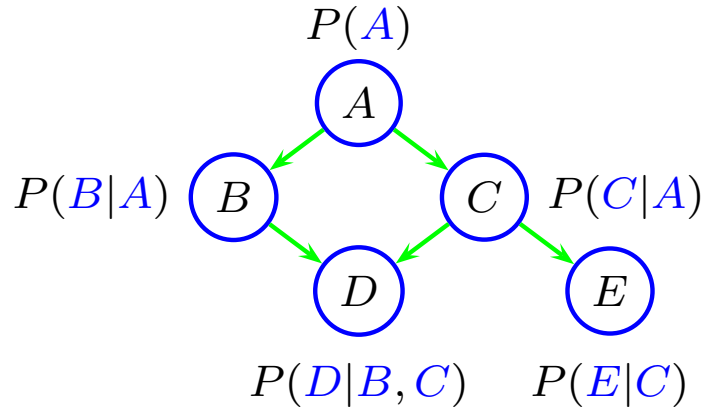
# Joint probabilities



Calculate  $P(A, B, C, D, E)$

$$\begin{aligned} P(A, B, C, D, E) &= P(E|A, B, C, D)P(A, B, C, D) \\ &= P(E|C)P(A, B, C, D) \\ &= P(E|C)P(D|A, B, C)P(A, B, C) \\ &= P(E|C)P(D|B, C)P(A, B, C) \\ &= P(E|C)P(D|B, C)P(C|A, B)P(A, B) \end{aligned}$$

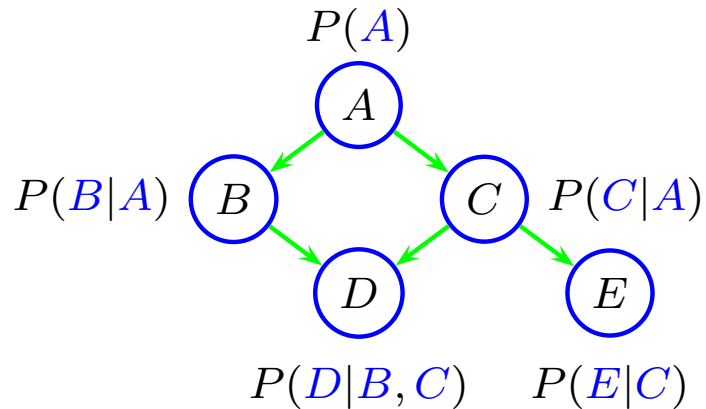
# Joint probabilities



Calculate  $P(A, B, C, D, E)$

$$\begin{aligned} P(A, B, C, D, E) &= P(E|A, B, C, D)P(A, B, C, D) \\ &= P(E|C)P(A, B, C, D) \\ &= P(E|C)P(D|A, B, C)P(A, B, C) \\ &= P(E|C)P(D|B, C)P(A, B, C) \\ &= P(E|C)P(D|B, C)P(C|A, B)P(A, B) \\ &= P(E|C)P(D|B, C)P(C|A)P(B, A) \end{aligned}$$

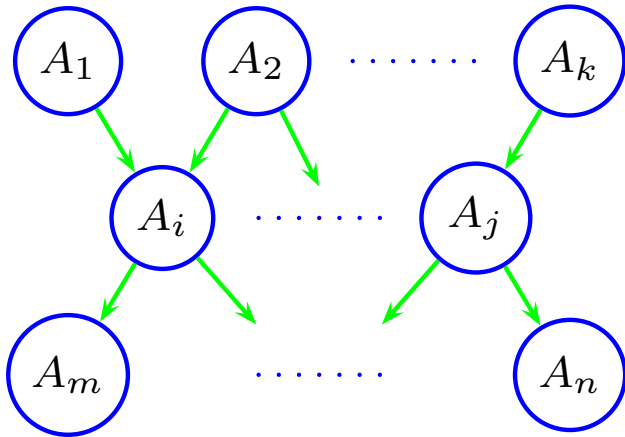
# Joint probabilities



Calculate  $P(A, B, C, D, E)$

$$\begin{aligned} P(A, B, C, D, E) &= P(E|A, B, C, D)P(A, B, C, D) \\ &= P(E|C)P(A, B, C, D) \\ &= P(E|C)P(D|A, B, C)P(A, B, C) \\ &= P(E|C)P(D|B, C)P(A, B, C) \\ &= P(E|C)P(D|B, C)P(C|A, B)P(A, B) \\ &= P(E|C)P(D|B, C)P(C|A)P(B, A) \\ &= P(E|C)P(D|B, C)P(C|A)P(B|A)P(A) \end{aligned}$$

# The chain rule



Let BN be a Bayesian network over  $\mathcal{U} = \{A_1, \dots, A_n\}$

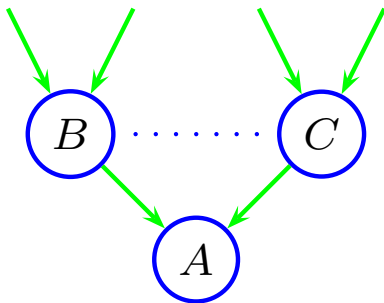
Then:

$$P(\mathcal{U}) = \prod_i P(A_i | \text{Pa}(A_i)),$$

where  $\text{Pa}(A_i)$  are the parents of  $A_i$ .

- $P(\mathcal{U})$  is the product of the potentials specified in BN.
- BN is a compact representation of  $P(\mathcal{U})$ .

.....



$$\begin{aligned} P(\mathcal{U}) &= P(A | \mathcal{U} \setminus \{A\}) P(\mathcal{U} \setminus \{A\}) \\ &= P(A | B, \dots, C) \prod_{X \in \mathcal{U} \setminus \{A\}} P(X | \text{Pa}(X)) \end{aligned}$$

# Evidence I

Consider a variable  $A$  with five states  $a_1, a_2, a_3, a_4, a_5$  and with probability:

$$P(A) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}, \quad \sum_{i=1}^5 x_i = 1$$

Assume that we get the evidence  $e$ : “ $A$  is either in state  $a_2$  or  $a_4$ ”. Then:

$$P(A, e) = \begin{pmatrix} 0 \\ x_2 \\ 0 \\ x_4 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Thus,  $e$  can be represented by a potential  $\bar{e} = (0, 1, 0, 1, 0)^T$  and:

$$P(A, e) = P(A) \cdot \bar{e}$$

# Evidence II

**Definition:** Let  $A$  be a variable with  $n$  states. A **finding** on  $A$  is an  $n$ -dimensional table with 0s and 1s.

**Semantics:** The states marked with a 0 are impossible.

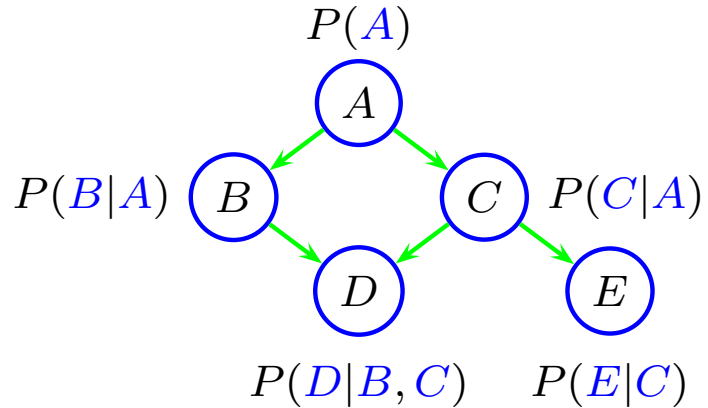
**Theorem:** Let BN be a Bayesian network over the universe  $\mathcal{U} = \{A_1, \dots, A_n\}$ , and let  $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_m$  be findings. Then:

$$\begin{aligned} P(\mathcal{U}, e) &= P(\mathcal{U}) \cdot \prod_{i=1}^m \bar{e}_i \\ &= \prod_{i=1}^n P(A_i | \text{Pa}(A_i)) \prod_{j=1}^m \bar{e}_j. \end{aligned}$$

Hence, to find  $P(A|e)$  we use:

$$P(A|e) = \frac{\sum_{\mathcal{U} \setminus \{A\}} P(\mathcal{U}, e)}{P(e)}.$$

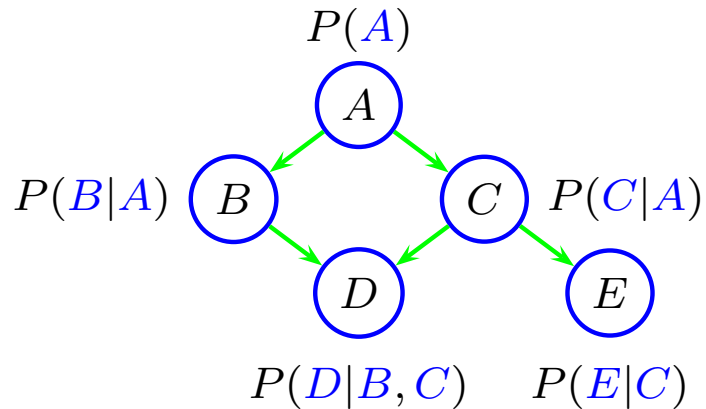
# Variable elimination



Do we need  $P(\mathcal{U}) = P(A, B, C, D, E)$  in order to calculate  $P(A|c, e)$ ?

Note: 
$$P(A|c, e) = \frac{\sum_B \sum_D P(A, B, c, D, e)}{P(c, e)}.$$

# Variable elimination

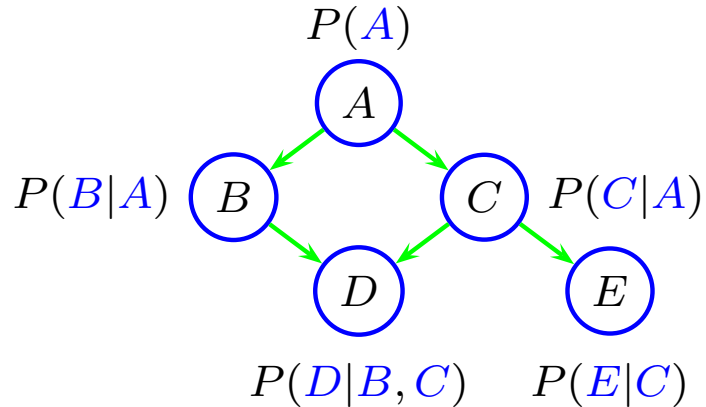


Do we need  $P(\mathcal{U}) = P(A, B, C, D, E)$  in order to calculate  $P(A|c, e)$ ?

Note:  $P(A|c, e) = \frac{\sum_B \sum_D P(A, B, c, D, e)}{P(c, e)}$ .

$$\sum_B \sum_D P(A, B, c, D, e) = \sum_B \sum_D P(e|c)P(c|A)P(D|c, B)P(A)P(B|A)$$

# Variable elimination

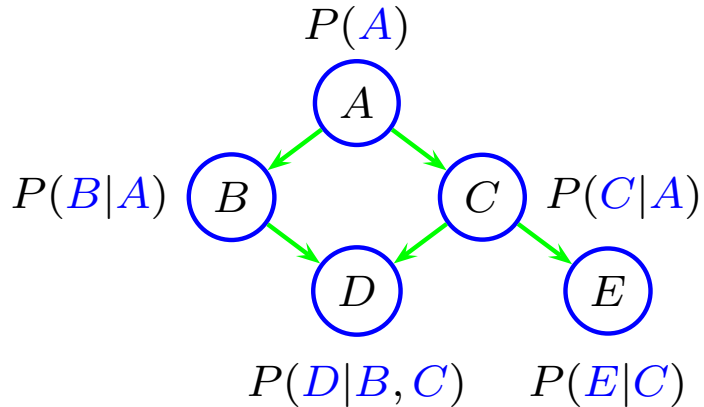


Do we need  $P(\mathcal{U}) = P(A, B, C, D, E)$  in order to calculate  $P(A|c, e)$ ?

Note:  $P(A|c, e) = \frac{\sum_B \sum_D P(A, B, c, D, e)}{P(c, e)}$ .

$$\begin{aligned} \sum_B \sum_D P(A, B, c, D, e) &= \sum_B \sum_D P(e|c)P(c|A)P(D|c, B)P(A)P(B|A) \\ &= P(e|c)P(c|A)P(A) \sum_B \sum_D P(D|c, B)P(B|A) \end{aligned}$$

# Variable elimination

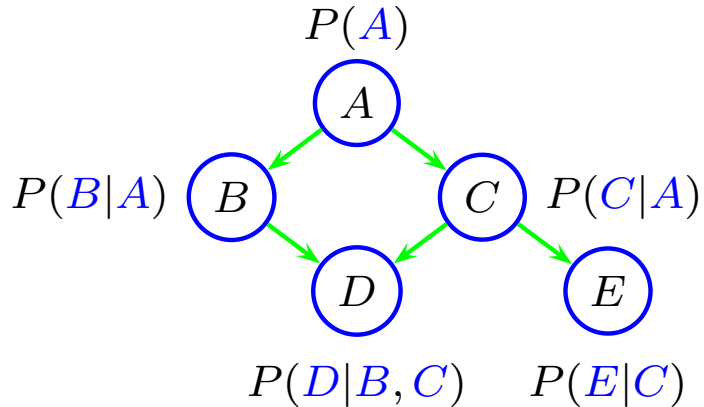


Do we need  $P(\mathcal{U}) = P(A, B, C, D, E)$  in order to calculate  $P(A|c, e)$ ?

Note:  $P(A|c, e) = \frac{\sum_B \sum_D P(A, B, c, D, e)}{P(c, e)}$ .

$$\begin{aligned}
 \sum_B \sum_D P(A, B, c, D, e) &= \sum_B \sum_D P(e|c)P(c|A)P(D|c, B)P(A)P(B|A) \\
 &= P(e|c)P(c|A)P(A) \sum_B \sum_D P(D|c, B)P(B|A) \\
 &= P(e|c)P(c|A)P(A) \sum_B P(B|A) \sum_D P(D|c, B)
 \end{aligned}$$

# Variable elimination



Do we need  $P(\mathcal{U}) = P(A, B, C, D, E)$  in order to calculate  $P(A|c, e)$ ?

$$\text{Note: } P(A|c, e) = \frac{\sum_B \sum_D P(A, B, c, D, e)}{P(c, e)}.$$

$$\begin{aligned} \sum_B \sum_D P(A, B, c, D, e) &= \sum_B \sum_D P(e|c)P(c|A)P(D|c, B)P(A)P(B|A) \\ &= P(e|c)P(c|A)P(A) \sum_B \sum_D P(D|c, B)P(B|A) \\ &= P(e|c)P(c|A)P(A) \sum_B P(B|A) \sum_D P(D|c, B) \\ &= P(e|c)P(c|A)P(A) \end{aligned}$$

So instead of constructing a table with  $2^5$  entries we only need 2 numbers!