CS5201: Advanced Artificial Intelligence

Probabilistic Reasoning

Arijit Mondal
Dept of Computer Science and Engineering
Indian Institute of Technology Patna
www.iitp.ac.in/~arijit/

Need for probabilistic model

- **Till now, we have considered that intelligent agent has**
  - **Known environment**
  - **Full observability**
  - **Deterministic world**

- **Reasons for using probability**
  - **Example:** $Toothache \implies Cavity$
    - **Toothache can be caused by many other ways also, eg.** $Toothache \implies GumDisease \lor Cavity \lor WisdomTeeth \lor \ldots$
  - **Specifications become too large**
  - **Theoretical ignorance**
  - **Practical ignorance**

Probabilistic reasoning

- Useful for prediction
  - Analyzing causal effect, predicting outcome
    - Given that I have cavity, what is the chance that I will have toothache?
- Useful for diagnosis
  - Analyzing causal effect, finding out reasons for a given effect
    - Given that I have toothache, what is the chance that it is caused by a cavity?
- Require a methodology to analyze both the scenarios

Axioms of probability

- **All probabilities lie between 0 and 1, ie.,** $0 \leq P(A) \leq 1$
- $P(True) = 1$ **and** $P(False) = 0$
- $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

- $P(A \wedge B) = P(A|B) \times P(B) = P(B|A) \times P(A)$
  - $P(A) = \dfrac{P(A|B) \times P(B)}{P(B|A)}$
  - **Bayes' rule**

Joint probability

Color \ Type	EV	SUV	Sedan	Truck
Red	0.05	0.20	0.00	0.10
Green	0.10	0.00	0.10	0.00
Black	0.00	0.10	0.05	0.10
White	0.10	0.00	0.10	0.00

- **What is the probability of a vehicle to be EV and Red?**

Joint probability

Color \ Type	EV	SUV	Sedan	Truck
Red	0.05	0.20	0.00	0.10
Green	0.10	0.00	0.10	0.00
Black	0.00	0.10	0.05	0.10
White	0.10	0.00	0.10	0.00

- **What is the probability of a vehicle to be EV and Red?**
- **What is the probability of a Red vehicle?**

Joint probability

Color \ Type	EV	SUV	Sedan	Truck
Red	0.05	0.20	0.00	0.10
Green	0.10	0.00	0.10	0.00
Black	0.00	0.10	0.05	0.10
White	0.10	0.00	0.10	0.00

- **What is the probability of a vehicle to be EV and Red?**
- **What is the probability of a Red vehicle?**
  - **Marginal probability:** $\sum_{T} P(C = Red \wedge T = *)$

Joint probability

Type / Color	EV	SUV	Sedan	Truck
Red	0.05	0.20	0.00	0.10
Green	0.10	0.00	0.10	0.00
Black	0.00	0.10	0.05	0.10
White	0.10	0.00	0.10	0.00

- **What is the probability of a vehicle to be EV and Red?**
- **What is the probability of a Red vehicle?**
  - **Marginal probability:** $\sum_{T} P(C = Red \wedge T = *)$
- **What is the probability of a vehicle to be EV given that it is Red?**

Joint probability

Color \ Type	EV	SUV	Sedan	Truck
Red	0.05	0.20	0.00	0.10
Green	0.10	0.00	0.10	0.00
Black	0.00	0.10	0.05	0.10
White	0.10	0.00	0.10	0.00

- **What is the probability of a vehicle to be EV and Red?**
- **What is the probability of a Red vehicle?**
  - **Marginal probability:** $\sum_{T} P(C = Red \wedge T = *)$
- **What is the probability of a vehicle to be EV given that it is Red?**
  - **Conditional probability:** $\dfrac{P(C = Red \wedge T = EV)}{P(C = Red)}$

Independence

- **Two variables are independent if $P(X, Y) = P(X) \times P(Y)$ holds**
  - **It means that their joint distribution factors into a product two distributions**
  - **This can be expressed as $P(X|Y) = P(X)$**
- **Independence is a simplifying modeling assumption**
  - **Empirical joint distributions can be at best close to independent**

Example: Independence

T	W	P(T,W)
Hot	Sun	0.4
Hot	Rain	0.1
Cold	Sun	0.2
Cold	Rain	0.3

Example: Independence

T	W	P(T,W)
Hot	Sun	0.4
Hot	Rain	0.1
Cold	Sun	0.2
Cold	Rain	0.3

- **Are T and W independent?**

Example: Independence

T	W	P(T,W)
Hot	Sun	0.4
Hot	Rain	0.1
Cold	Sun	0.2
Cold	Rain	0.3

- **Are *T* and *W* independent?**
- **Find marginal probabilities**

Example: Independence

T	W	P(T,W)
Hot	Sun	0.4
Hot	Rain	0.1
Cold	Sun	0.2
Cold	Rain	0.3

- **Are T and W independent?**
- **Find marginal probabilities**
  - $P(T = Hot) = ?$, $P(T = Cold) = ?$
  - $P(W = Sun) = ?$, $P(W = Rain) = ?$

Example: Independence

T	W	P(T,W)
Hot	Sun	0.4
Hot	Rain	0.1
Cold	Sun	0.2
Cold	Rain	0.3

- **Are T and W independent?**
- **Find marginal probabilities**
  - $P(T = Hot) =?$, $P(T = Cold) =?$
  - $P(W = Sun) =?$, $P(W = Rain) =?$
- **Now check for independence**

Example: Independence

- **Tossing of N fair coins**

$P(X_1)$	
H	0.5
T	0.5

$P(X_2)$	
H	0.5
T	0.5

$\ldots\ldots$

$P(X_n)$	
H	0.5
T	0.5

$P(X_1, X_2, \ldots, X_n)$

X_1	X_2	X_3	$\ldots$	$\ldots$	X_n	P
H					H	
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	
T					H	

2^n

Chain rule

- **Product rule**

$$P(X, Y) = P(X) \times P(Y|X)$$

- **Chain rule**

$$\begin{aligned} P(X_1, X_2, \ldots, X_n) &= P(X_1) \times P(X_2|X_1) \times P(X_3|X_1, X_2) \times \ldots \times P(X_N|X_1, \ldots, X_{n-1}) \\ &= \prod_i P(X_i|X_1, \ldots, X_{i-1})) \end{aligned}$$

Conditional independence

- $P(Toothache, Cavity, Catch)$
- **If I have a cavity, the probability that the probe catches in it doesn't depend on whether I have a toothache:**

$$P(+catch| + toothache, +cavity) = P(+catch| + cavity)$$

- **The same independence holds if I don't have a cavity:**

$$P(+catch| + toothache, -cavity) = P(+catch| - cavity)$$

- **Catch is conditionally independent of Toothache given Cavity:**

$$P(Catch|Toothache, Cavity) = P(Catch|Cavity)$$

Conditional independence

- Absolute independence is very rare
- Conditional independence is the most basic and robust form of knowledge about un-certain environments
- Mathematically it is defined as - X is conditionally independent of Y given Z

$$P(X, Y|Z) = P(X|Z) \times P(Y|Z)$$

- It can also be shown that

$$P(X|Z, Y) = P(X|Z)$$

Belief network

- Qualitative information
  - A belief network is a graph consist of the following:
  - Nodes – Set of random variables
  - Edges – Dependency of nodes. $X \rightarrow Y$ means X has direct influence on Y
- Quantitative information
  - Each node has conditional probability table that quantifies the effects the parents have on the node

- The network is a directed acyclic graph (DAG), ie., without any cycle

Example

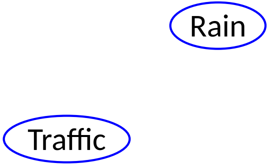
- **Consider the following knowledge base**
  - *Rain* **- Raining**
  - *Traffic* **- There may be traffic if it rains**
  - *Umbrella* **- People may carry umbrella if it rains**

Example

- **Consider the following knowledge base**
  - *Rain* **- Raining**
  - *Traffic* **- There may be traffic if it rains**
  - *Umbrella* **- People may carry umbrella if it rains**

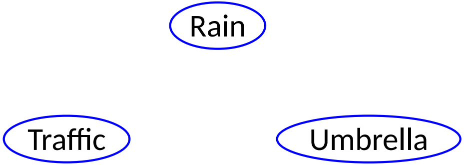
Example

- **Consider the following knowledge base**
  - *Rain* **- Raining**
  - *Traffic* **- There may be traffic if it rains**
  - *Umbrella* **- People may carry umbrella if it rains**



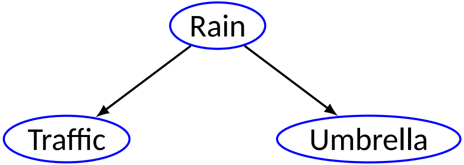
Example

- **Consider the following knowledge base**
  - *Rain* **- Raining**
  - *Traffic* **- There may be traffic if it rains**
  - *Umbrella* **- People may carry umbrella if it rains**



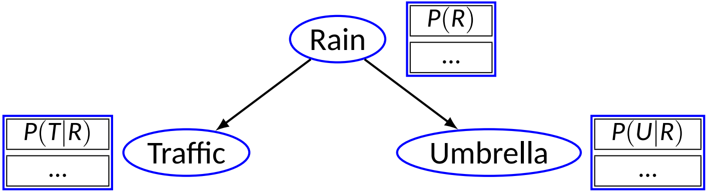
Example

- **Consider the following knowledge base**
  - *Rain* **- Raining**
  - *Traffic* **- There may be traffic if it rains**
  - *Umbrella* **- People may carry umbrella if it rains**

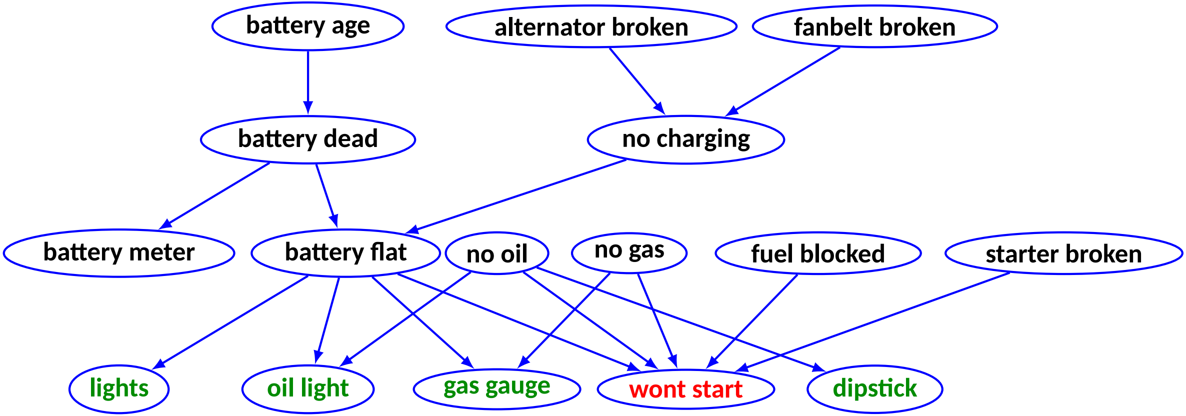


Example

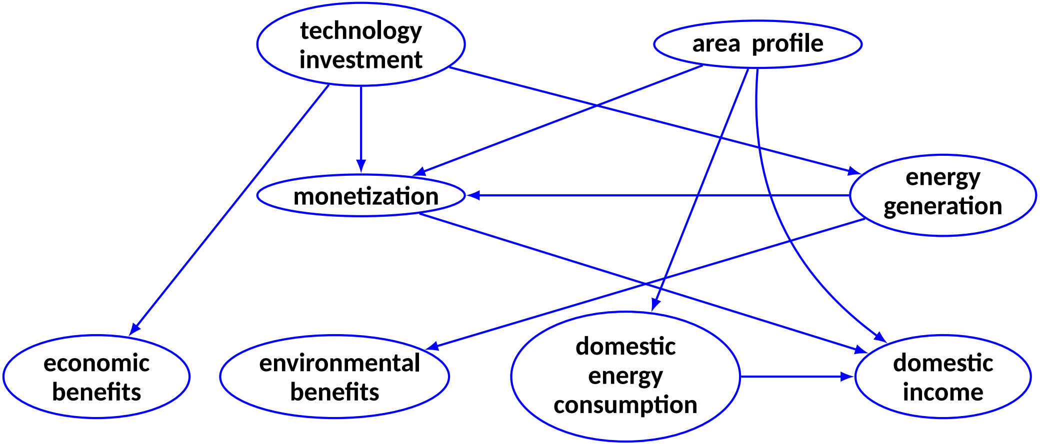
- **Consider the following knowledge base**
  - *Rain* - **Raining**
  - *Traffic* - **There may be traffic if it rains**
  - *Umbrella* - **People may carry umbrella if it rains**



Belief network example: car



Belief network example: renewable energy



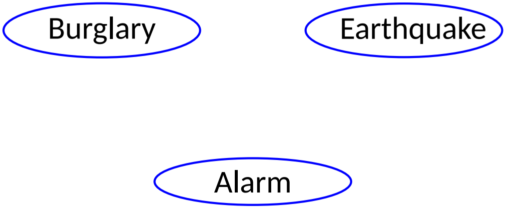
Classical example

- Burglar alarm at a home
  - It works almost perfectly
  - Sometime alarm goes off if there is an earthquake
- John calls police when he hears the alarm but sometimes he misinterpret telephone ringing as alarm and calls police too.
- Mary also calls police. But she loves loud music, therefore, she misses the alarm sometime.

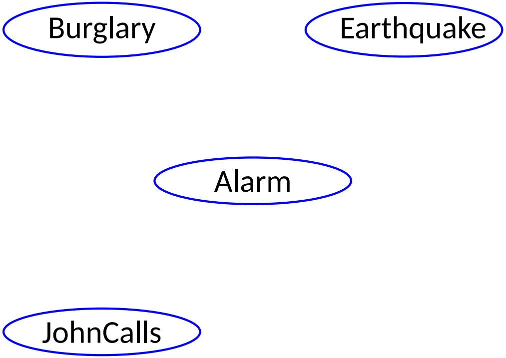
Belief network example

Belief network example

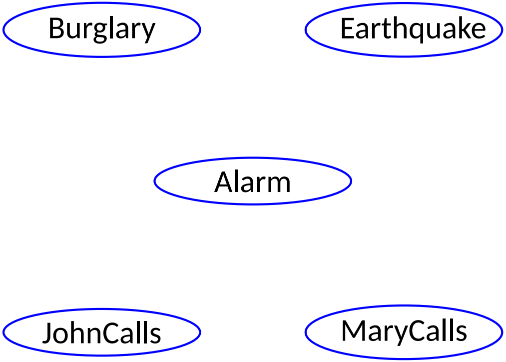
Belief network example



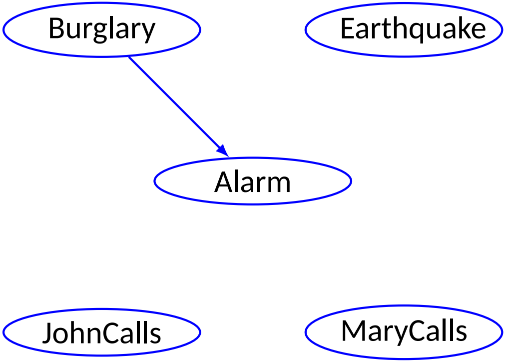
Belief network example



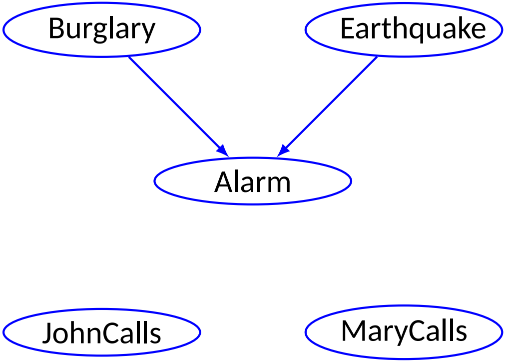
Belief network example



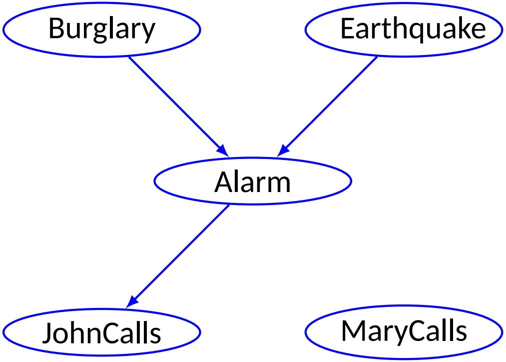
Belief network example



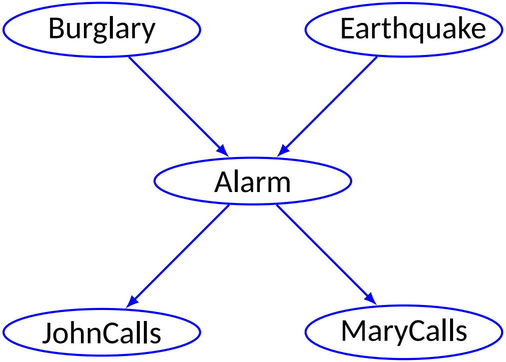
Belief network example



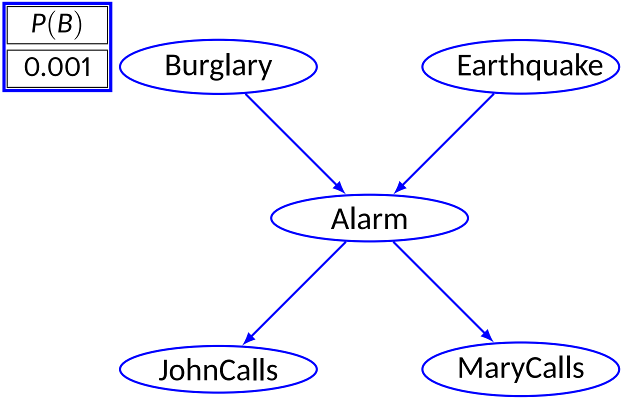
Belief network example



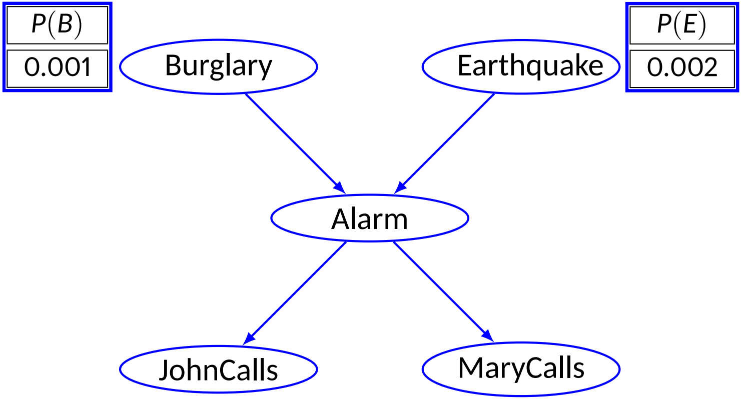
Belief network example



Belief network example



Belief network example

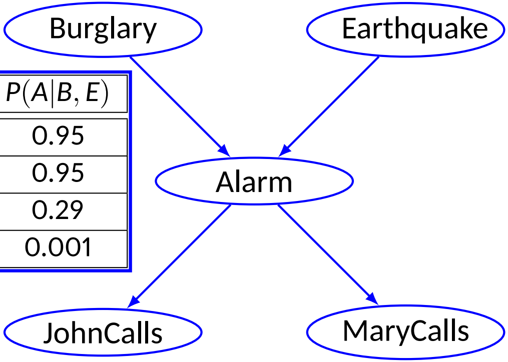


Belief network example

$P(B)$
0.001

$P(E)$
0.002

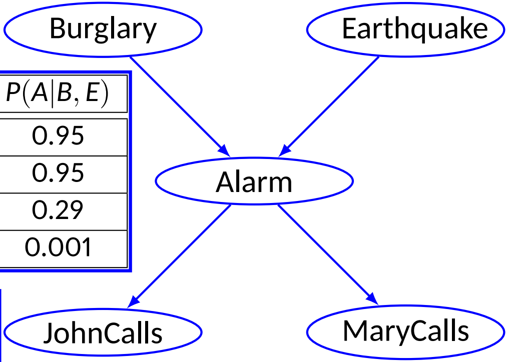
B	E	$P(A \mid B, E)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001



Belief network example

$P(B)$
0.001

$P(E)$
0.002



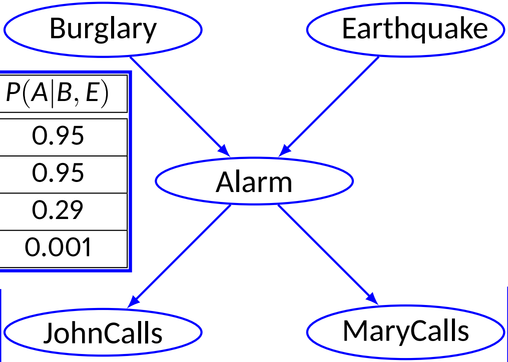
B	E	$P(A\|B, E)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J\|A)$
T	0.90
F	0.05

Belief network example

$P(B)$
0.001

$P(E)$
0.002



B	E	$P(A\|B,E)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J\|A)$
T	0.90
F	0.05

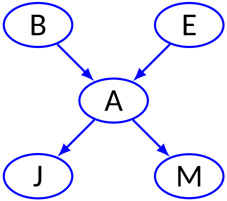
A	$P(M\|A)$
T	0.70
F	0.01

Probabilities in Belief Nets

- **Bayes' net implicitly encode joint distributions**
  - **This can be determined from local conditional distributions**
  - **Need to multiply relevant conditional probabilities**
  - $P(X_1, X_2, \ldots, X_n) = \prod_{i}^{n} P(X_i | parents(X_i))$

Joint probability distribution

- **Probability of the event that the alarm has sounded but neither a burglary nor an earthquake has occurred, and both Mary and John call:**



$P(B)$
0.001

$P(E)$
0.002

B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

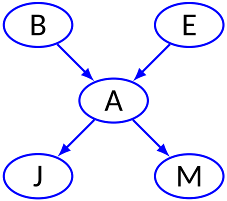
A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution

- **Probability of the event that the alarm has sounded but neither a burglary nor an earthquake has occurred, and both Mary and John call:**

  $P(J \wedge M \wedge A \wedge \neg B \wedge \neg E)$



$P(B)$
0.001

$P(E)$
0.002

B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

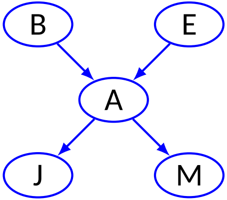
A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution

- **Probability of the event that the alarm has sounded but neither a burglary nor an earthquake has occurred, and both Mary and John call:**

$$
\begin{aligned}
&P(J \land M \land A \land \neg B \land \neg E) \\
&\quad = P(J|A) \times P(M|A) \times P(A|\neg B \land \neg E) \times P(\neg B) \times P(\neg E) \\
&\quad = 0.9 \times 0.7 \times 0.001 \times 0.999 \times 0.998 \\
&\quad = 0.00062
\end{aligned}
$$



$P(B)$
0.001

$P(E)$
0.002

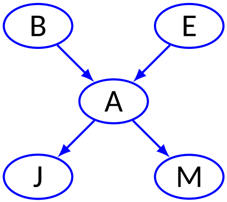
B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(A)$

$P(A)$



$P(B)$
0.001

$P(E)$
0.002

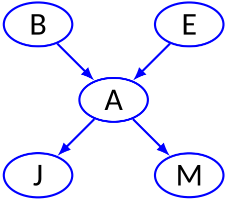
B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(A)$

$$
\begin{aligned}
&P(A) \\
&\quad = P(A\bar{B}\bar{E}) + P(A\bar{B}E) + P(AB\bar{E}) + P(ABE)
\end{aligned}
$$



$P(B)$
0.001

$P(E)$
0.002

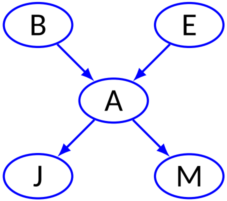
B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(A)$

$$
\begin{aligned}
P(A) & \\
&= P(A\bar{B}\bar{E}) + P(A\bar{B}E) + P(AB\bar{E}) + P(ABE) \\
&= P(A|\bar{B}\bar{E})\times P(\bar{B}\bar{E}) + P(A|\bar{B}E)\times P(\bar{B}E) + P(A|B\bar{E})\times P(B\bar{E}) + P(A|BE)\times P(BE)
\end{aligned}
$$



$P(B)$
0.001

$P(E)$
0.002

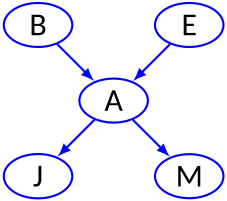
B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(A)$

$$
\begin{aligned}
&P(A)\\
&\quad = P(A\bar{B}\bar{E}) + P(A\bar{B}E) + P(AB\bar{E}) + P(ABE)\\
&\quad = P(A|\bar{B}\bar{E})\times P(\bar{B}\bar{E}) + P(A|\bar{B}E)\times P(\bar{B}E) + P(A|B\bar{E})\times P(B\bar{E}) + P(A|BE)\times P(BE)\\
&\quad = 0.001\times 0.999\times 0.998 + 0.29\times 0.999\times 0.002 + 0.95\times 0.001\times 0.998\\
&\qquad\quad +0.95\times 0.001\times 0.002\\
&\quad = 0.0025
\end{aligned}
$$



$P(B)$
0.001

$P(E)$
0.002

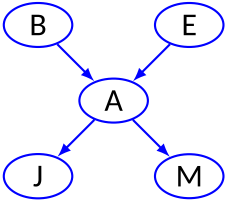
B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(J)$

$P(J)$



$P(B)$
0.001

$P(E)$
0.002

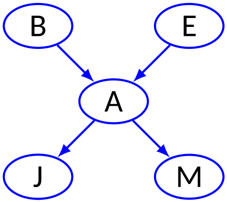
B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(J)$

$$P(J)$$
$$= P(JA) + P(J\bar{A})$$



$P(B)$
0.001

$P(E)$
0.002

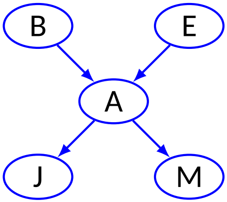
B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(J)$

$$
\begin{aligned}
& P(J) \\
& \quad = P(JA) + P(J\bar{A}) \\
& \quad = P(J|A) \times P(A) + P(J|\bar{A}) \times P(\bar{A}) \\
& \quad = 0.9 \times 0.0025 + 0.05 \times (1 - 0.0025) \\
& \quad = 0.052125
\end{aligned}
$$



$P(B)$
0.001

$P(E)$
0.002

B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

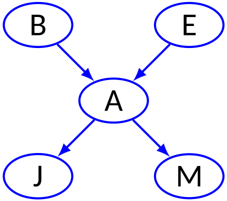
A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(J)$

$$
\begin{aligned}
P(J) & \\
&= P(JA) + P(J\bar{A}) \\
&= P(J|A)\times P(A) + P(J|\bar{A})\times P(\bar{A}) \\
&= 0.9\times 0.0025 + 0.05\times(1 - 0.0025) \\
&= 0.052125
\end{aligned}
$$

$P(AB)$



$P(B)$
0.001

$P(E)$
0.002

B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

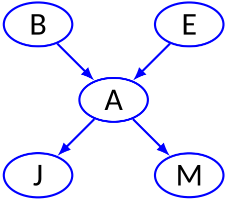
A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(J)$

$$
\begin{aligned}
&P(J)\\
&\quad = P(JA) + P(J\bar{A})\\
&\quad = P(J|A)\times P(A) + P(J|\bar{A})\times P(\bar{A})\\
&\quad = 0.9\times 0.0025 + 0.05\times(1 - 0.0025)\\
&\quad = 0.052125
\end{aligned}
$$

$$
\begin{aligned}
&P(AB)\\
&\quad = P(ABE) + P(AB\bar{E})
\end{aligned}
$$



$P(B)$
0.001

$P(E)$
0.002

B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

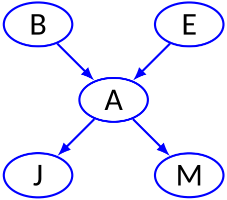
A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(J)$

$$P(J)$$
$$= P(JA) + P(J\bar{A})$$
$$= P(J|A)\times P(A) + P(J|\bar{A})\times P(\bar{A})$$
$$= 0.9\times 0.0025 + 0.05\times(1 - 0.0025)$$
$$= 0.052125$$

$$P(AB)$$
$$= P(ABE) + P(AB\bar{E})$$
$$= 0.95\times 0.001\times 0.002 + 0.95\times 0.001\times 0.998$$
$$= 0.00095$$



$P(B)$
0.001

$P(E)$
0.002

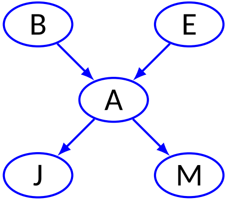
B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(JB)$

$P(JB)$



$P(B)$
0.001

$P(E)$
0.002

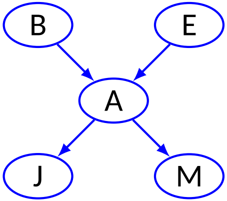
B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(JB)$

$$P(JB)$$
$$= P(JBA) + P(JB\bar{A})$$



$P(B)$
0.001

$P(E)$
0.002

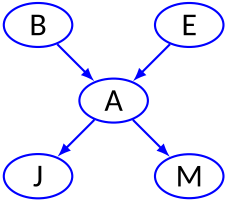
B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(JB)$

$$P(JB)$$
$$= P(JBA) + P(JB\bar{A})$$
$$= P(J|AB) \times P(AB) + P(J|\bar{A}B) \times P(\bar{A}B)$$



$P(B)$
0.001

$P(E)$
0.002

B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

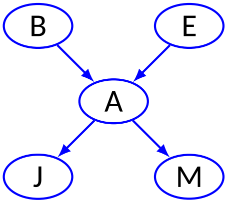
A	$P(M)$
T	0.70
F	0.01

Joint probability distribution: $P(JB)$

$P(JB)$

$$
\begin{aligned}
&= P(JBA) + P(JB\bar{A}) \\
&= P(J|AB)\times P(AB) + P(J|\bar{A}B)\times P(\bar{A}B) \\
&= P(J|A)\times P(AB) + P(J|\bar{A})\times P(\bar{A}B) \\
&= 0.9\times 0.00095 + 0.05\times 0.00005 \\
&= 0.00086
\end{aligned}
$$

$P(J|B)$



$P(B)$
0.001

$P(E)$
0.002

B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

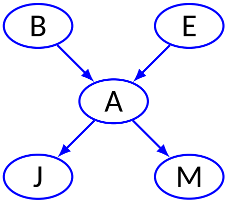
Joint probability distribution: $P(JB)$

$P(JB)$

$$= P(JBA) + P(JB\bar{A})$$
$$= P(J|AB)\times P(AB) + P(J|\bar{A}B)\times P(\bar{A}B)$$
$$= P(J|A)\times P(AB) + P(J|\bar{A})\times P(\bar{A}B)$$
$$= 0.9\times 0.00095 + 0.05\times 0.00005$$
$$= 0.00086$$

$P(J|B)$

$$= \frac{P(JB)}{P(B)} = \frac{0.00086}{0.001} = 0.86$$



$P(B)$
0.001

$P(E)$
0.002

B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

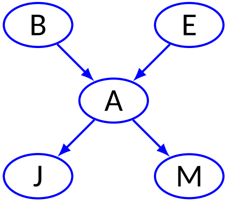
Inferences using belief networks

- **Diagnostic inferences (effects to causes)**
- **Given that JohnCalls, infer that** $P(Burglary|JohnCalls) = 0.016$

- **Causal inferences (causes to effects)**
- **Given Burglary, infer that**

  $P(JohnCalls|Burglary) = 0.86$

  $P(MaryCalls|Burglary) = 0.67$



$P(B)$
0.001

$P(E)$
0.002

B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

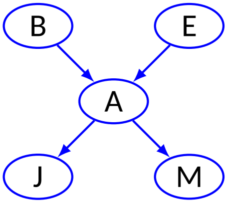
A	$P(M)$
T	0.70
F	0.01

Inferences using belief networks

- **Inter-causal inferences (between causes to a common effect)**
- **Given Alarm, we have** $P(Burglary|Alarm) = 0.376$
- **Also, if it is given that Earthquake is true, then**
  $P(Burglary|Alarm \wedge Earthquake) = 0.003$

- **Mixed inferences**

  $P(Alarm|JohnCalls \wedge \neg Earthquake) = 0.003$



$P(B)$
0.001

$P(E)$
0.002

B	E	$P(A)$
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	$P(J)$
T	0.90
F	0.05

A	$P(M)$
T	0.70
F	0.01

Example: traffic

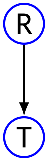
- Consider the following variables
  - R - It is Raining
  - T - There is traffic

Example: traffic

- Consider the following variables
  - R - It is Raining
  - T - There is traffic

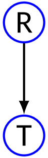
Example: traffic

- Consider the following variables
  - R - It is Raining
  - T - There is traffic



Example: traffic

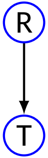
- **Consider the following variables**
  - R **- It is Raining**
  - T **- There is traffic**



$P(R)$
0.25

Example: traffic

- Consider the following variables
  - R - It is Raining
  - T - There is traffic

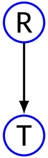


$P(R)$
0.25

	$P(T \mid R)$
R	0.75
$\neg R$	0.50

Example: traffic

- Consider the following variables
  - R - It is Raining
  - T - There is traffic



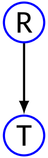
$P(R)$
0.25

	$P(T\|R)$
R	0.75
$\neg R$	0.50

		$P(T,R)$
R	T	3/16
R	$\neg T$	1/16
$\neg R$	T	6/16
$\neg R$	$\neg T$	6/16

Example: traffic

- Consider the following variables
  - R - It is Raining
  - T - There is traffic



$P(R)$
0.25

	$P(T\|R)$
R	0.75
$\neg R$	0.50

		$P(T,R)$
R	T	3/16
R	$\neg T$	1/16
$\neg R$	T	6/16
$\neg R$	$\neg T$	6/16

$P(T)$
9/16

Example: reverse traffic

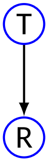
- **Consider the following variables**
  - R **- It is Raining**
  - T **- There is traffic**

Example: reverse traffic

- **Consider the following variables**
  - R **- It is Raining**
  - T **- There is traffic**

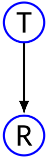
Example: reverse traffic

- Consider the following variables
  - R - It is Raining
  - T - There is traffic



Example: reverse traffic

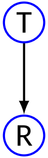
- **Consider the following variables**
  - R - **It is Raining**
  - T - **There is traffic**



$P(T)$
$9/16$

Example: reverse traffic

- **Consider the following variables**
  - R **- It is Raining**
  - T **- There is traffic**

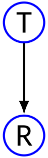


$P(T)$
9/16

	$P(R\|T)$
T	1/3
$\neg T$	1/7

Example: reverse traffic

- **Consider the following variables**
  - R - **It is Raining**
  - T - **There is traffic**



$P(T)$
$9/16$

	$P(R\|T)$
T	$1/3$
$\neg T$	$1/7$

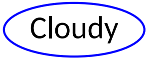
		$P(T,R)$
R	T	$3/16$
R	$\neg T$	$1/16$
$\neg R$	T	$6/16$
$\neg R$	$\neg T$	$6/16$

Example: Sprinkler

- There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.

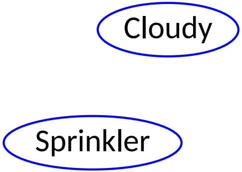
Example: Sprinkler

- **There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.**



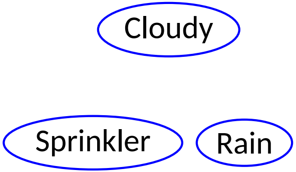
Example: Sprinkler

- **There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.**



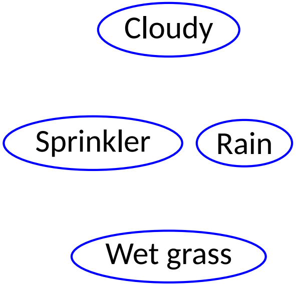
Example: Sprinkler

- **There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.**



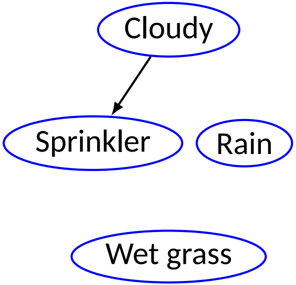
Example: Sprinkler

- **There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.**



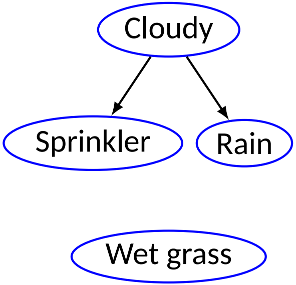
Example: Sprinkler

- There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.



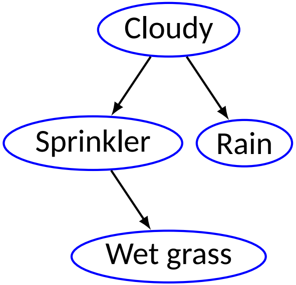
Example: Sprinkler

- **There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.**



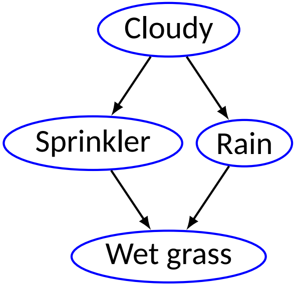
Example: Sprinkler

- **There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.**



Example: Sprinkler

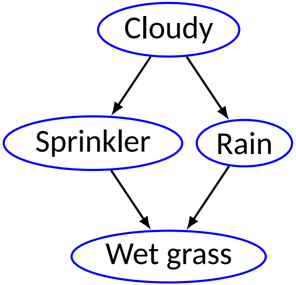
- **There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.**



Example: Sprinkler

- **There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.**

$P(C)$
0.50

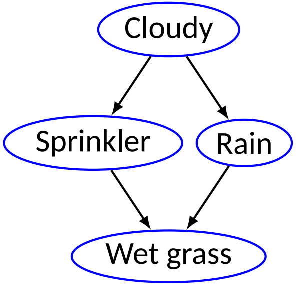


Example: Sprinkler

- There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50



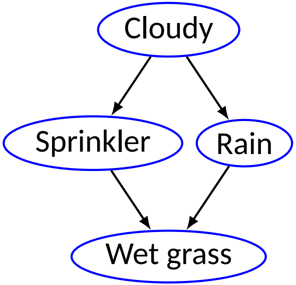
Example: Sprinkler

- **There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.**

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

	$P(R)$
C	0.80
$\neg C$	0.20



Example: Sprinkler

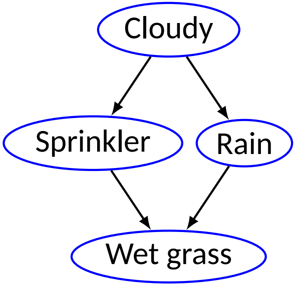
- **There is some chance that tomorrow will be cloudy. If it is cloudy, then chances of rain is high. Also, there is a sprinkler to water the grass. When there is a prediction of cloud, it is less likely to run sprinkler. We want to model whether grass is wet or not.**

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

	$P(R)$
C	0.80
$\neg C$	0.20

		$P(W)$
S	R	0.99
s	$\neg R$	0.90
$\neg S$	R	0.90
$\neg S$	$\neg R$	0.00



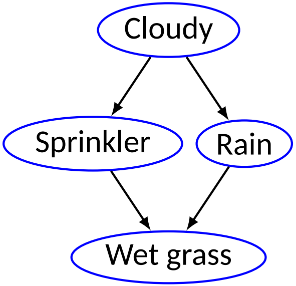
Example: Sprinkler

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

	$P(R)$
C	0.80
$\neg C$	0.20

		$P(W)$
S	R	0.99
S	$\neg R$	0.90
$\neg S$	R	0.90
$\neg S$	$\neg R$	0.00



- **Find** $P(W, S, R, C)$

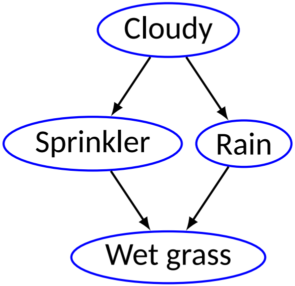
Example: Sprinkler

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

	$P(R)$
C	0.80
$\neg C$	0.20

		$P(W)$
S	R	0.99
S	$\neg R$	0.90
$\neg S$	R	0.90
$\neg S$	$\neg R$	0.00



- **Find** $P(W, S, R, C)$

$$= P(C) \times P(S|C) \times P(R|C) \times P(W|S, R)$$

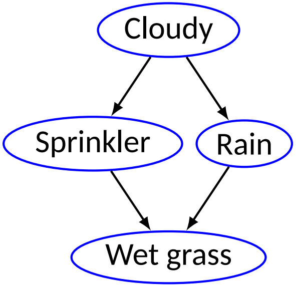
Example: Sprinkler

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

	$P(R)$
C	0.80
$\neg C$	0.20

		$P(W)$
S	R	0.99
S	$\neg R$	0.90
$\neg S$	R	0.90
$\neg S$	$\neg R$	0.00



- **Find** $P(W, S, R, C)$

$$= P(C) \times P(S|C) \times P(R|C) \times P(W|S, R)$$

- **Find** $P(W = T|C = T)$

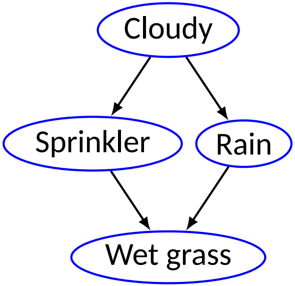
Example: Sprinkler

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

	$P(R)$
C	0.80
$\neg C$	0.20

		$P(W)$
S	R	0.99
S	$\neg R$	0.90
$\neg S$	R	0.90
$\neg S$	$\neg R$	0.00



- **Find** $P(W, S, R, C)$
$$= P(C) \times P(S|C) \times P(R|C) \times P(W|S, R)$$
- **Find** $P(W = T|C = T)$
$$= \frac{P(W = T, S, R, C = T)}{P(C = T)}$$

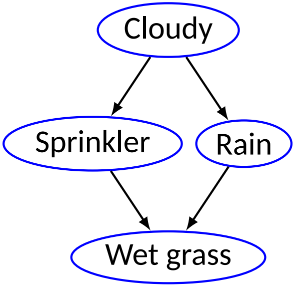
Example: Sprinkler

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

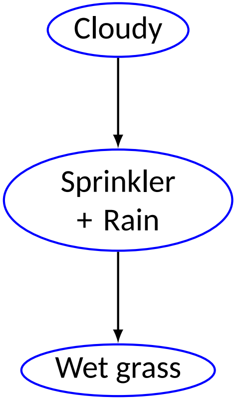
	$P(R)$
C	0.80
$\neg C$	0.20

		$P(W)$
S	R	0.99
S	$\neg R$	0.90
$\neg S$	R	0.90
$\neg S$	$\neg R$	0.00



- **Find** $P(W, S, R, C)$
$$= P(C) \times P(S|C) \times P(R|C) \times P(W|S, R)$$
- **Find** $P(W = T|C = T)$
$$= \frac{P(W = T, S, R, C = T)}{P(C = T)} = 0.7452$$

Example: Sprinkler - merging nodes



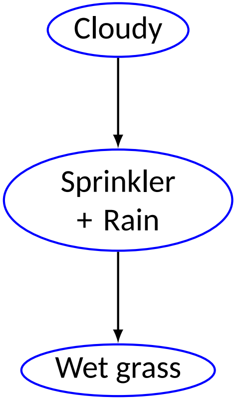
Example: Sprinkler - merging nodes

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

	$P(R)$
C	0.80
$\neg C$	0.20

		$P(W)$
S	R	0.99
s	$\neg R$	0.90
$\neg S$	R	0.90
$\neg S$	$\neg R$	0.00

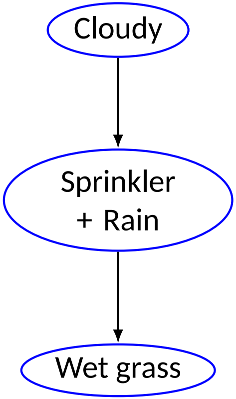


Example: Sprinkler - merging nodes

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

	$P(R)$
C	0.80
$\neg C$	0.20



$P(S \vee R)$

		$P(W)$
S	R	0.99
s	$\neg R$	0.90
$\neg S$	R	0.90
$\neg S$	$\neg R$	0.00

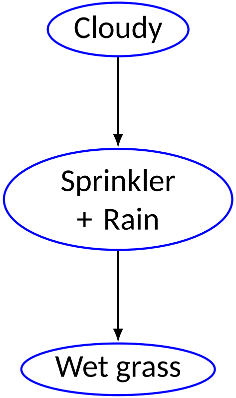
Example: Sprinkler - merging nodes

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

	$P(R)$
C	0.80
$\neg C$	0.20

		$P(W)$
S	R	0.99
s	$\neg R$	0.90
$\neg S$	R	0.90
$\neg S$	$\neg R$	0.00



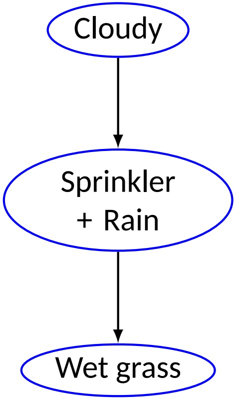
$P(S \vee R)$				
S+R / C	TT	TF	FT	FF
C				

Example: Sprinkler - merging nodes

$P(C)$
0.50

	$P(S)$
C	0.10
$\neg C$	0.50

	$P(R)$
C	0.80
$\neg C$	0.20



$P(S \vee R)$				
S+R / C	TT	TF	FT	FF
C	0.08	0.02	0.72	0.18
$\neg C$	0.40	0.10	0.40	0.10

		$P(W)$
S	R	0.99
s	$\neg R$	0.90
$\neg S$	R	0.90
$\neg S$	$\neg R$	0.00

Incremental network construction

- Select the set of random variables (X_i) to describe the domains
- Select an appropriate ordering of the variables
- Steps:
  - Select a variable (X) and add a node for it
  - Set $Parents(X)$ to some minimal set of existing nodes such that the conditional independence property is satisfied
  - Define the conditional probability table for X

Burglar-Alarm

- Variable ordering
  - MaryCalls, JohnCalls, Alarm, Burglary, Earthquake

Burglar-Alarm

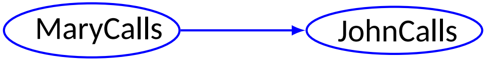
- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**

Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**

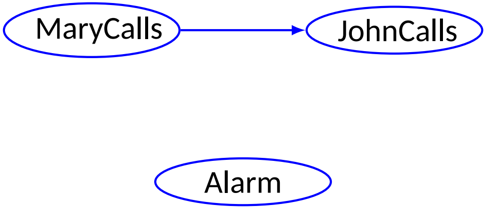
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**



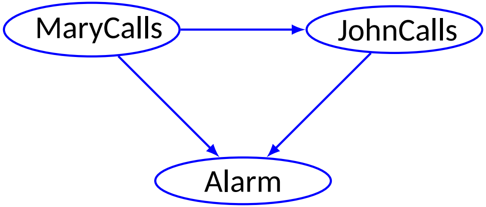
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**



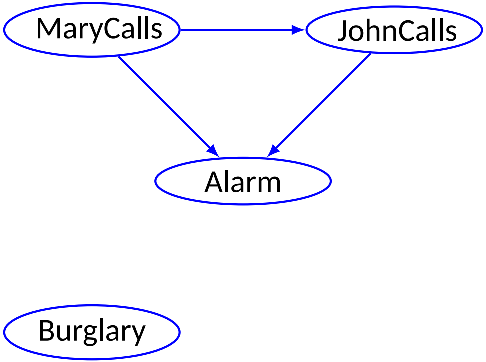
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**



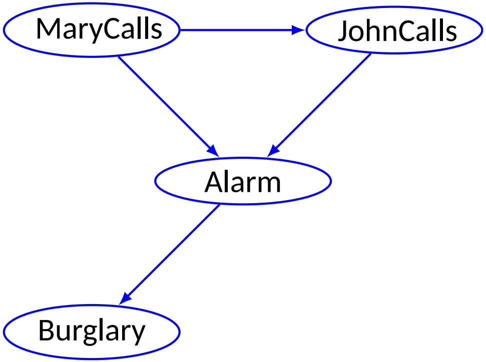
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**



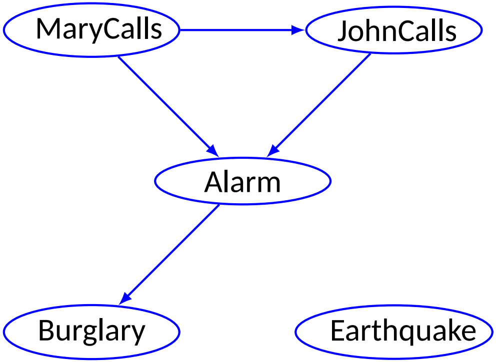
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**



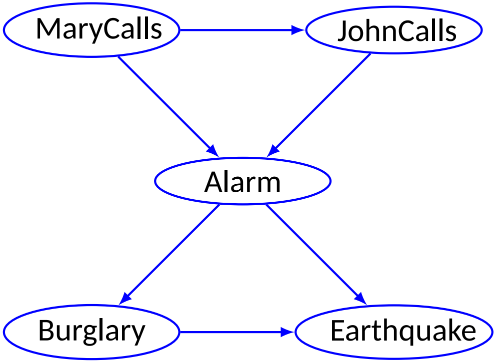
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**



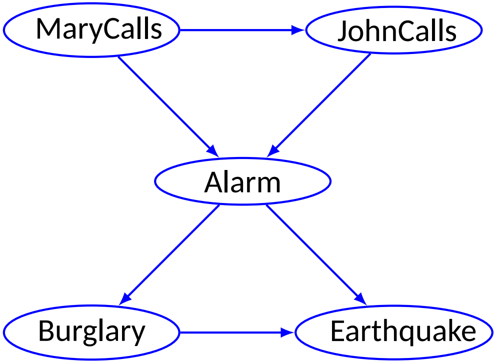
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**



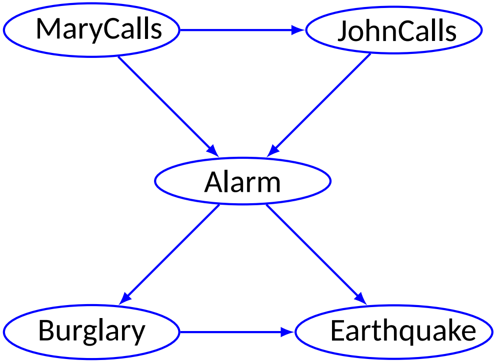
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**
  - **MaryCalls, JohnCalls, Earthquake, Burglary, Alarm**



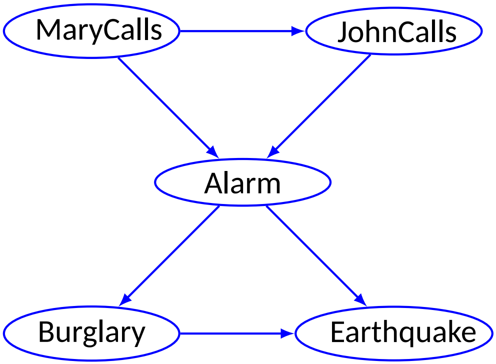
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**
  - **MaryCalls, JohnCalls, Earthquake, Burglary, Alarm**



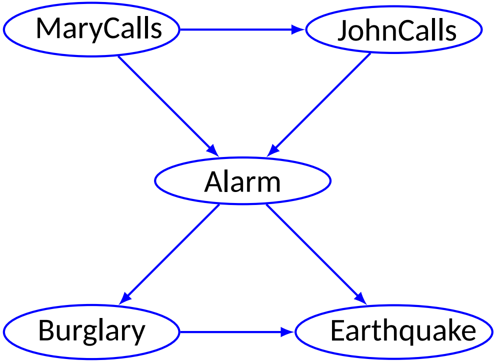
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**
  - **MaryCalls, JohnCalls, Earthquake, Burglary, Alarm**



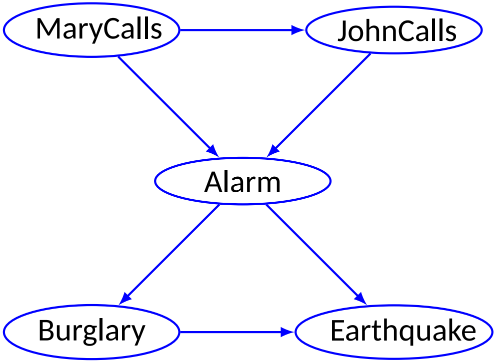
Burglar-Alarm

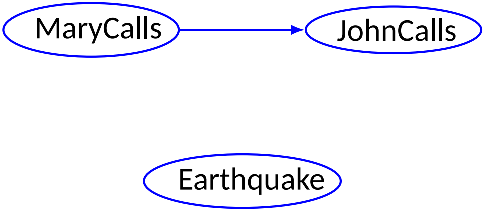
- Variable ordering
  - MaryCalls, JohnCalls, Alarm, Burglary, Earthquake
  - MaryCalls, JohnCalls, Earthquake, Burglary, Alarm



Burglar-Alarm

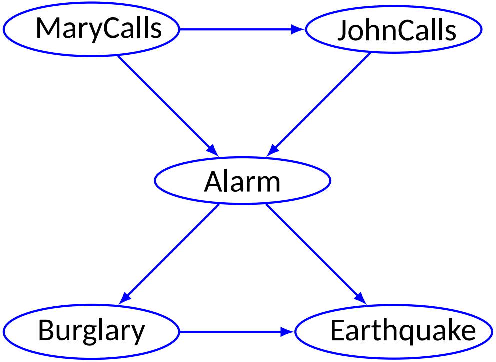
- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**
  - **MaryCalls, JohnCalls, Earthquake, Burglary, Alarm**

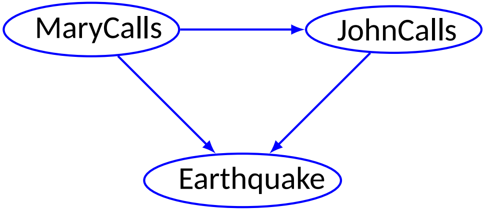




Burglar-Alarm

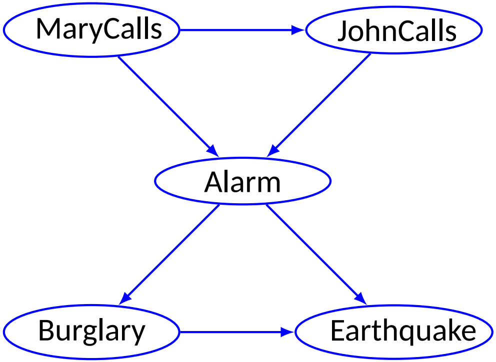
- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**
  - **MaryCalls, JohnCalls, Earthquake, Burglary, Alarm**

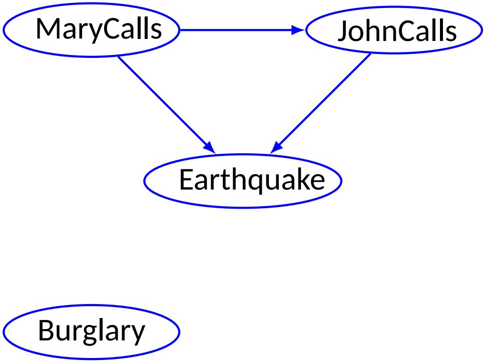




Burglar-Alarm

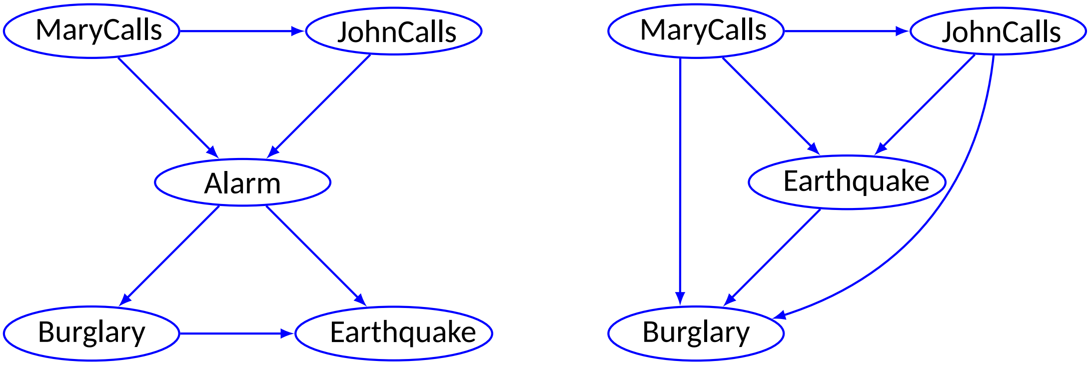
- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**
  - **MaryCalls, JohnCalls, Earthquake, Burglary, Alarm**





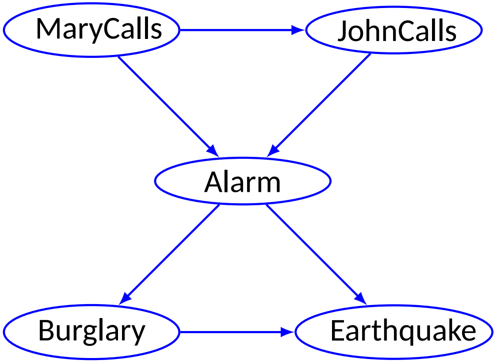
Burglar-Alarm

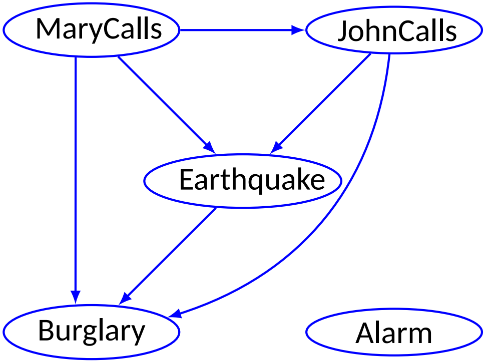
- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**
  - **MaryCalls, JohnCalls, Earthquake, Burglary, Alarm**



Burglar-Alarm

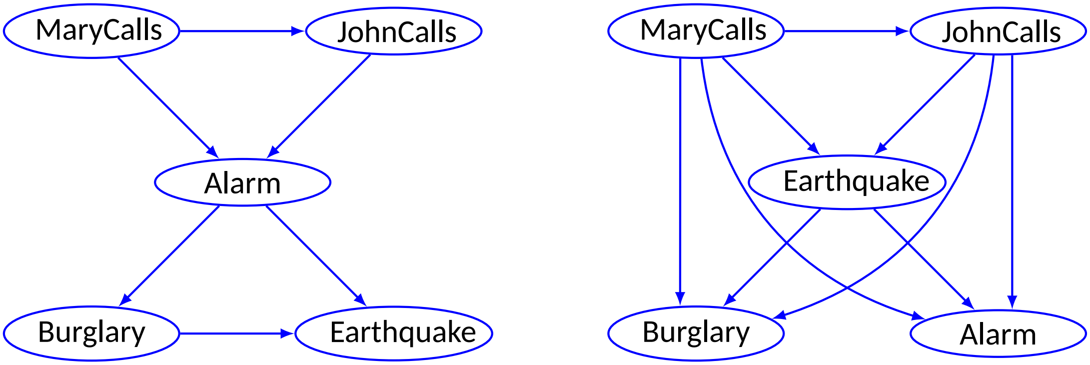
- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**
  - **MaryCalls, JohnCalls, Earthquake, Burglary, Alarm**





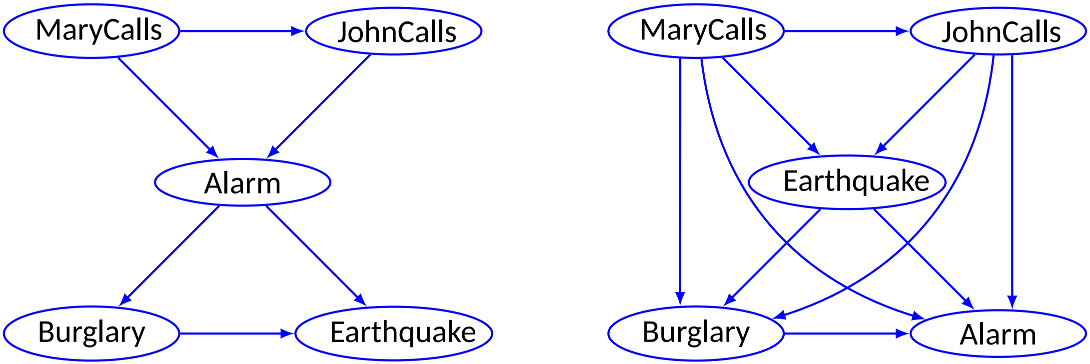
Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**
  - **MaryCalls, JohnCalls, Earthquake, Burglary, Alarm**

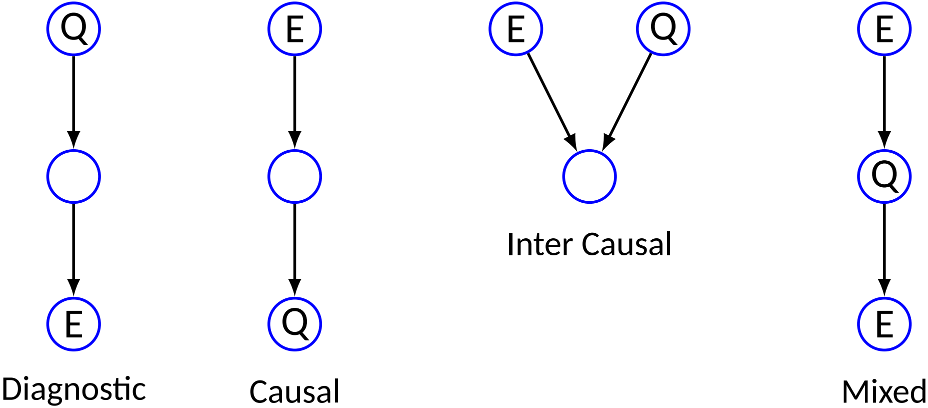


Burglar-Alarm

- **Variable ordering**
  - **MaryCalls, JohnCalls, Alarm, Burglary, Earthquake**
  - **MaryCalls, JohnCalls, Earthquake, Burglary, Alarm**

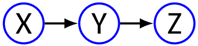


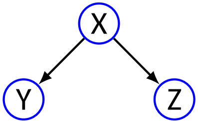
The four patterns

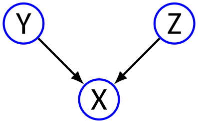


Conditional independence in DAG

- Given a DAG model, how to determine conditional independent relations among nodes/variables
- DAG has three basic graphical patterns, also known as canonical networks
  - Chain - three nodes connected in a line along the same direction
  - Fork - two children of a node
  - V-structure - two parents of a node





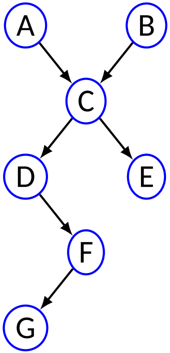


D-separation

- **Input: X, Y and Z**
- **Output: whether X is conditionally independent of Y given Z**

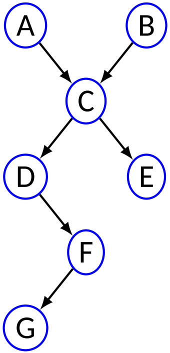
1. **[Ancestral graph] Consider ancestral subgraph using X, Y, Z and their ancestors**
2. **[Moralize] Add links between any unlinked pair of nodes that share a common child; now we have the so-called moral graph.**
3. **[Disorient] Replace all directed links by undirected links**
4. **[Delete givens] If Z blocks all paths between X and Y in the resulting graph, then Z d-separates X and Y. In that case, X is conditionally independent of Y, given Z.**

D-separation example

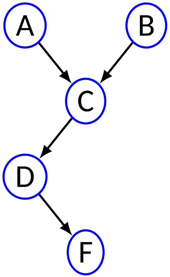


- **Are A and B conditionally independent, given D and F?**

D-separation example

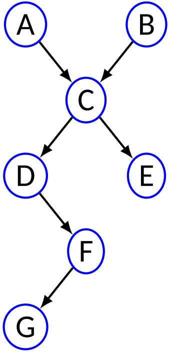


Ancestor graph

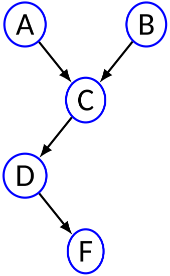


- **Are A and B conditionally independent, given D and F?**

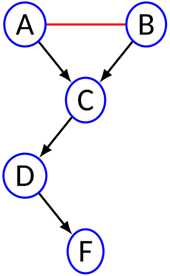
D-separation example



Ancestor graph

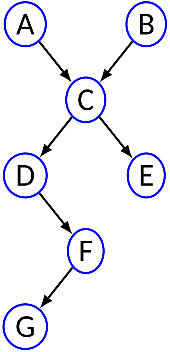


Moral graph

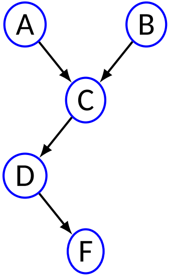


- **Are A and B conditionally independent, given D and F?**

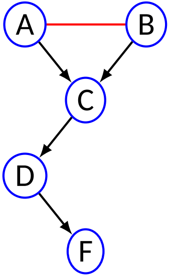
D-separation example



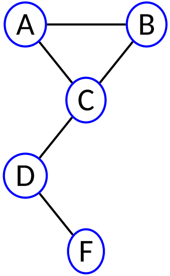
Ancestor graph



Moral graph

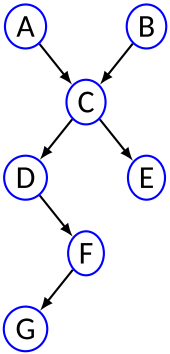


Disorient

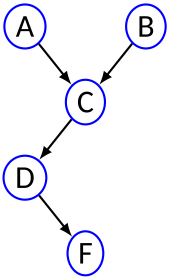


- **Are A and B conditionally independent, given D and F?**

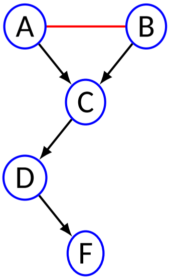
D-separation example



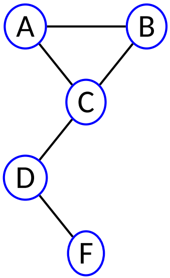
Ancestor graph



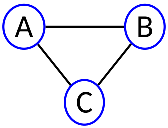
Moral graph



Disorient

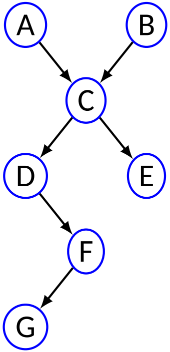


Delete givens



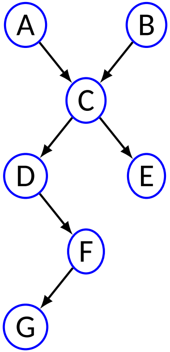
- **Are A and B conditionally independent, given D and F?**

D-separation example

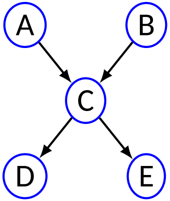


- **Are D and E conditionally independent, given C?**

D-separation example

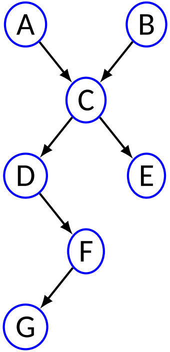


Ancestor graph

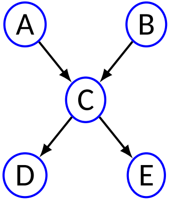


- **Are D and E conditionally independent, given C?**

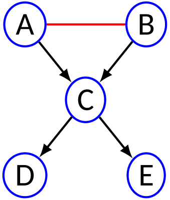
D-separation example



Ancestor graph

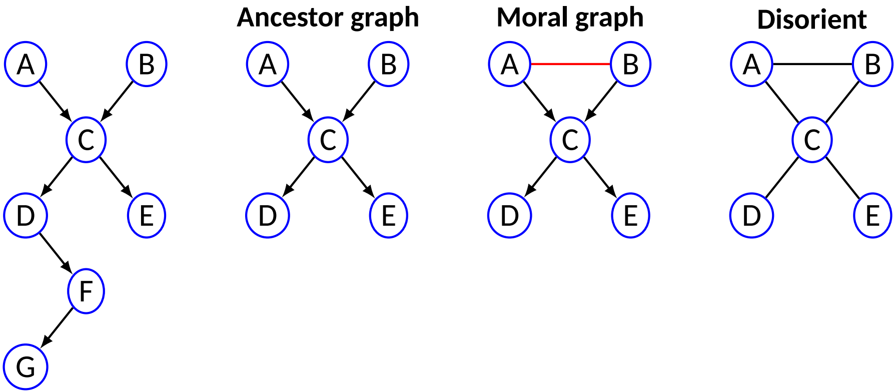


Moral graph



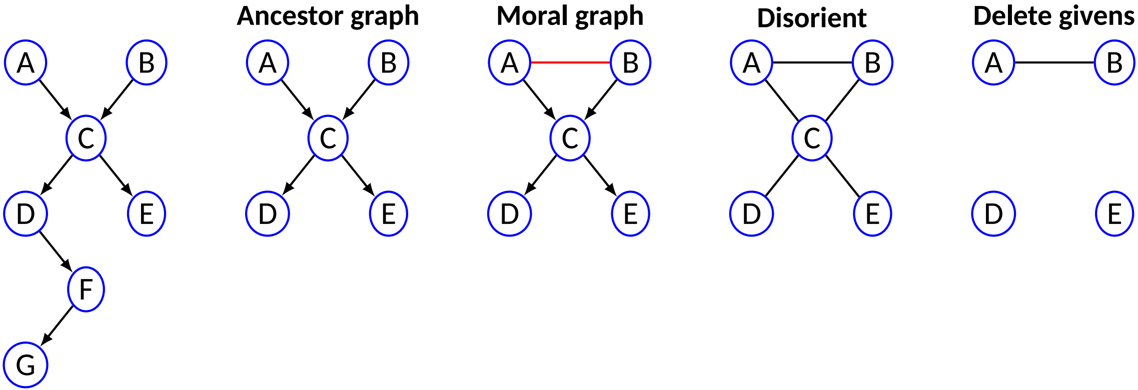
- **Are D and E conditionally independent, given C?**

D-separation example



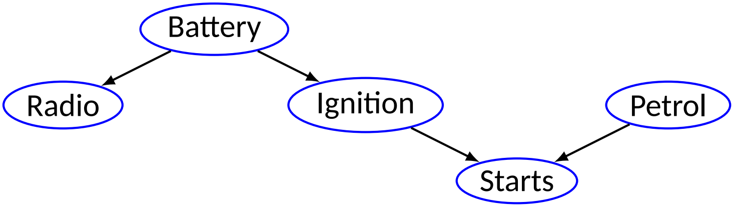
- **Are D and E conditionally independent, given C?**

D-separation example

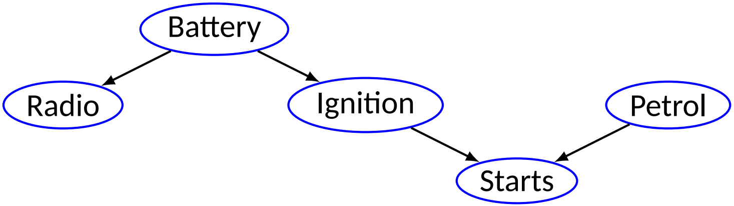


- **Are D and E conditionally independent, given C?**

Conditional independence in belief network

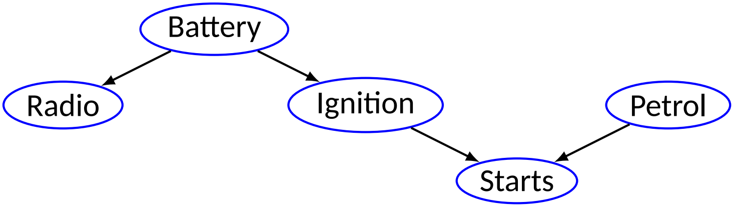


Conditional independence in belief network



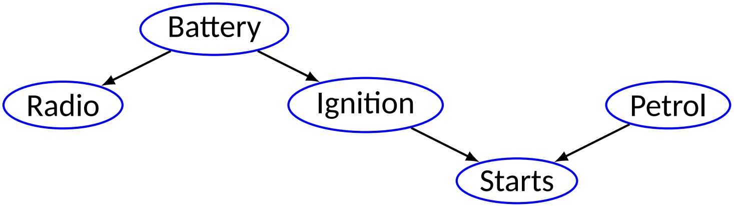
- **Whether there is petrol and whether the radio plays are independent given evidence about whether the ignition takes place**

Conditional independence in belief network



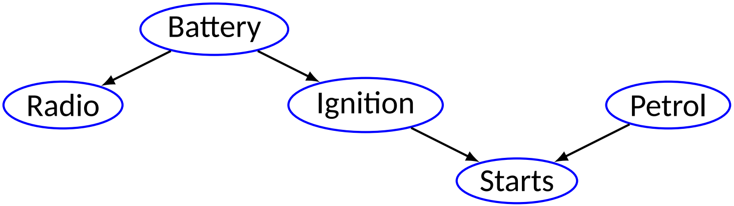
- **Whether there is petrol and whether the radio plays are independent given evidence about whether the ignition takes place**
- **Petrol and Radio are independent if it is known whether the battery works**

Conditional independence in belief network



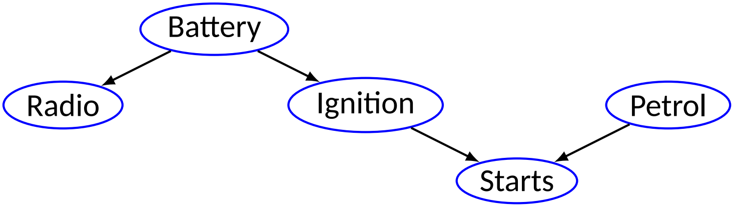
- Whether there is petrol and whether the radio plays are independent given evidence about whether the ignition takes place
- Petrol and Radio are independent if it is known whether the battery works
- Petrol and Radio are independent given no evidence at all.

Conditional independence in belief network



- Whether there is petrol and whether the radio plays are independent given evidence about whether the ignition takes place
- Petrol and Radio are independent if it is known whether the battery works
- Petrol and Radio are independent given no evidence at all.
- But they are dependent given evidence about whether the car starts.

Conditional independence in belief network

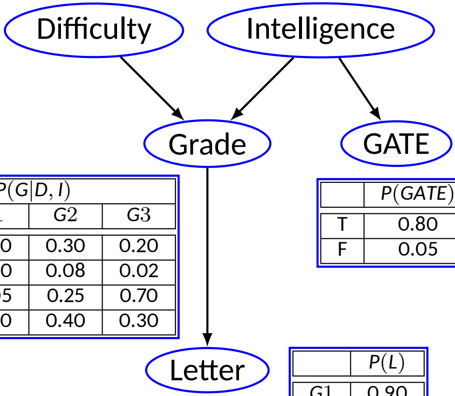


- Whether there is petrol and whether the radio plays are independent given evidence about whether the ignition takes place
- Petrol and Radio are independent if it is known whether the battery works
- Petrol and Radio are independent given no evidence at all.
- But they are dependent given evidence about whether the car starts.
- If the car does not start, then the radio playing is increased evidence that we are out of petrol.

Exercise: Grade

$P(D)$
0.40

$P(I)$
0.30



$P(G \mid D, I)$			
D, I	$G1$	$G2$	$G3$
T, T	0.50	0.30	0.20
T, F	0.90	0.08	0.02
F, T	0.05	0.25	0.70
F, F	0.30	0.40	0.30

	$P(GATE)$
T	0.80
F	0.05

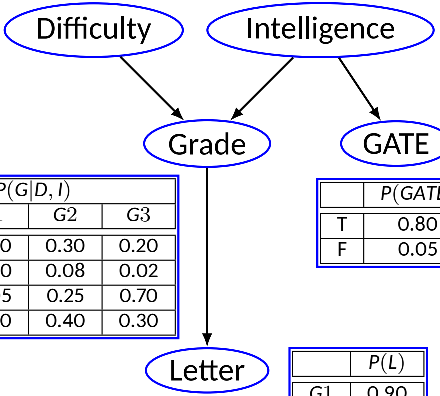
	$P(L)$
$G1$	0.90
$G2$	0.60
$G3$	0.01

Exercise: Grade

- **What is the probability of getting excellent recommendation letter given the semester question paper was very easy?**

$P(D)$
0.40

$P(I)$
0.30



$P(G \mid D, I)$			
D, I	$G1$	$G2$	$G3$
T, T	0.50	0.30	0.20
T, F	0.90	0.08	0.02
F, T	0.05	0.25	0.70
F, F	0.30	0.40	0.30

	$P(GATE)$
T	0.80
F	0.05

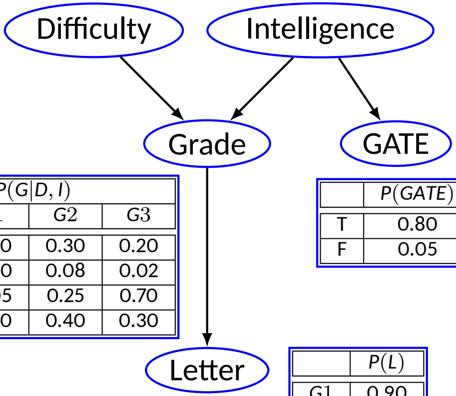
	$P(L)$
$G1$	0.90
$G2$	0.60
$G3$	0.01

Exercise: Grade

- **What is the probability of getting excellent recommendation letter given the semester question paper was very easy?**
- **What is the probability of a student being very intelligent given he gets a poor recommendation letter?**

$P(D)$
0.40

$P(I)$
0.30



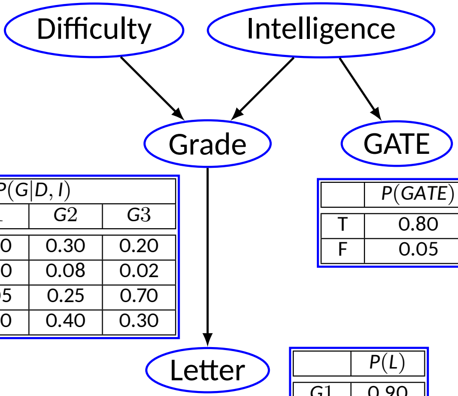
$P(G \mid D, I)$			
D, I	$G1$	$G2$	$G3$
T, T	0.50	0.30	0.20
T, F	0.90	0.08	0.02
F, T	0.05	0.25	0.70
F, F	0.30	0.40	0.30

	$P(GATE)$
T	0.80
F	0.05

	$P(L)$
$G1$	0.90
$G2$	0.60
$G3$	0.01

Exercise: Grade

- **What is the probability of getting excellent recommendation letter given the semester question paper was very easy?**
- **What is the probability of a student being very intelligent given he gets a poor recommendation letter?**
- **What is the probability of getting a poor GATE score given the semester grade was very good?**



$P(D)$
0.40

$P(I)$
0.30

$P(G \mid D, I)$			
D, I	$G1$	$G2$	$G3$
T, T	0.50	0.30	0.20
T, F	0.90	0.08	0.02
F, T	0.05	0.25	0.70
F, F	0.30	0.40	0.30

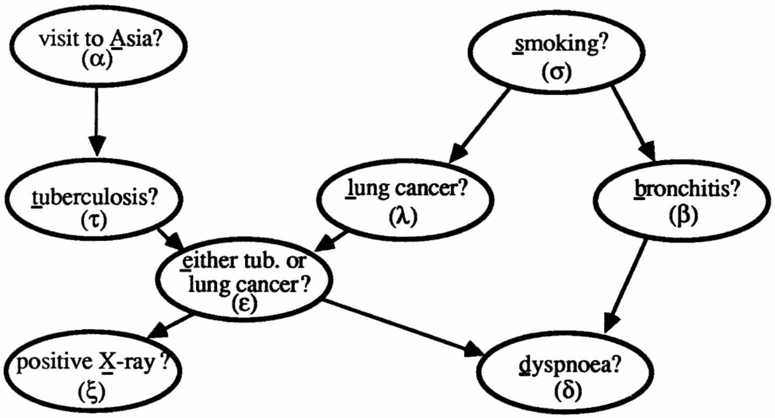
	$P(GATE)$
T	0.80
F	0.05

	$P(L)$
$G1$	0.90
$G2$	0.60
$G3$	0.01

Belief network: medical diagnosis

- **Construct a belief network for the following: [Lauritzen & Spiegelhalter, 1988]**
  - **Shortness-of-breath (dyspnoea) may be due to tuberculosis, lung cancer or bronchitis, or none of them, or more than one of them. A recent visit to Asia increases the chances of tuberculosis, while smoking is known to be a risk factor both lung cancer and bronchitis. The results of a single chest X-ray do not discriminate between lung cancer and tuberculosis, as neither does the presence or absence of dyspnoea.**

Belief network: medical diagnosis - BN



source: Lauritzen & Spiegelhalter, 1988

Belief network: medical diagnosis - CPT

α:	$p(a) = .01$		ε:	$p(e \mid l, t) = 1$	
				$p(e \mid l, \bar{t}) = 1$	
τ:	$p(t \mid a) = .05$			$p(e \mid \bar{l}, t) = 1$	
	$p(t \mid \bar{a}) = .01$			$p(e \mid \bar{l}, \bar{t}) = 0$	
σ:	$p(s) = .50$		ξ:	$p(x \mid e) = .98$	
				$p(x \mid \bar{e}) = .05$	
λ:	$p(l \mid s) = .10$				
	$p(l \mid \bar{s}) = .01$		δ:	$p(d \mid e, b) = .90$	
				$p(d \mid e, \bar{b}) = .70$	
β:	$p(b \mid s) = .60$			$p(d \mid \bar{e}, b) = .80$	
	$p(b \mid \bar{s}) = .30$			$p(d \mid \bar{e}, \bar{b}) = .10$	

- **A patient presents at a chest clinic with dyspnoea, and has recently visited Asia**
- **The doctor would like to know the chance that each of the diseases is present**
- **Suppose information on smoking has been provided now, repeat the above step**

source: Lauritzen & Spiegelhalter, 1988

Belief network: phd admission

- **Construct a belief network for the following**
  - **Let us assume that you are interested in pursing phd at IIT-Patna!!**
  - **To take admission in the phd program, you may need to provide at most two recommendation letters from the course instructors or to appear for PMRF interview.**
  - **A course instructor can provide at most one recommendation letter based on the your grade on that subject.**
  - **Moreover, the grade depends on the attendance, marks obtained in the assignments and in the final examination.**
  - **Intelligence of the student can also play a major role in the performance of a subject or in the interview.**
  - **You may feel happy if you perform well in the subjects or you get admission for the phd program.**

Thank you!