# Probabilistic Graphical Models

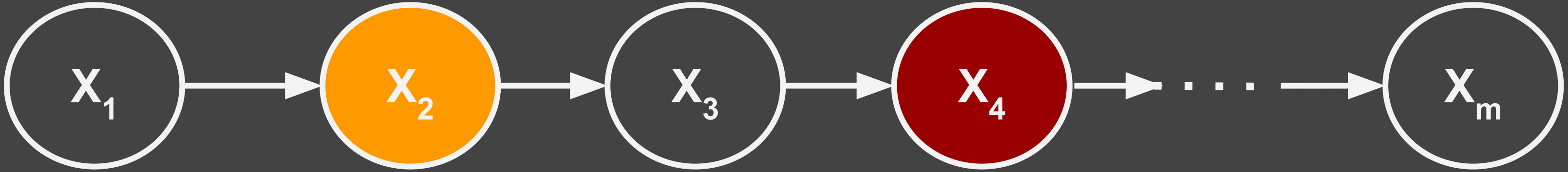
## Lectures 7,8

I-MAPs, Flow of influence, separation

# Questions

- What are all independence relations encoded by a Graphical Model?
- Which one is stronger? BNs or MRFs?

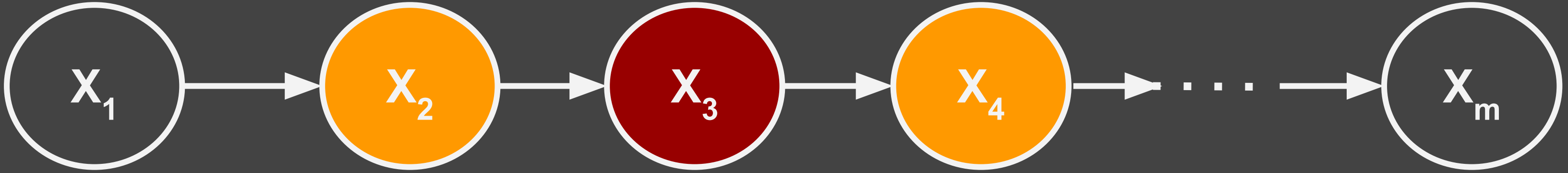
# Example



Markov Property: Given $X_2$, is $X_3$ independent of $X_1$.

Given $X_2$, is $X_4$ independent of $X_1$?

# Example



Given $X_2$ and $X_4$, is $X_3$ independent of the rest of the nodes?

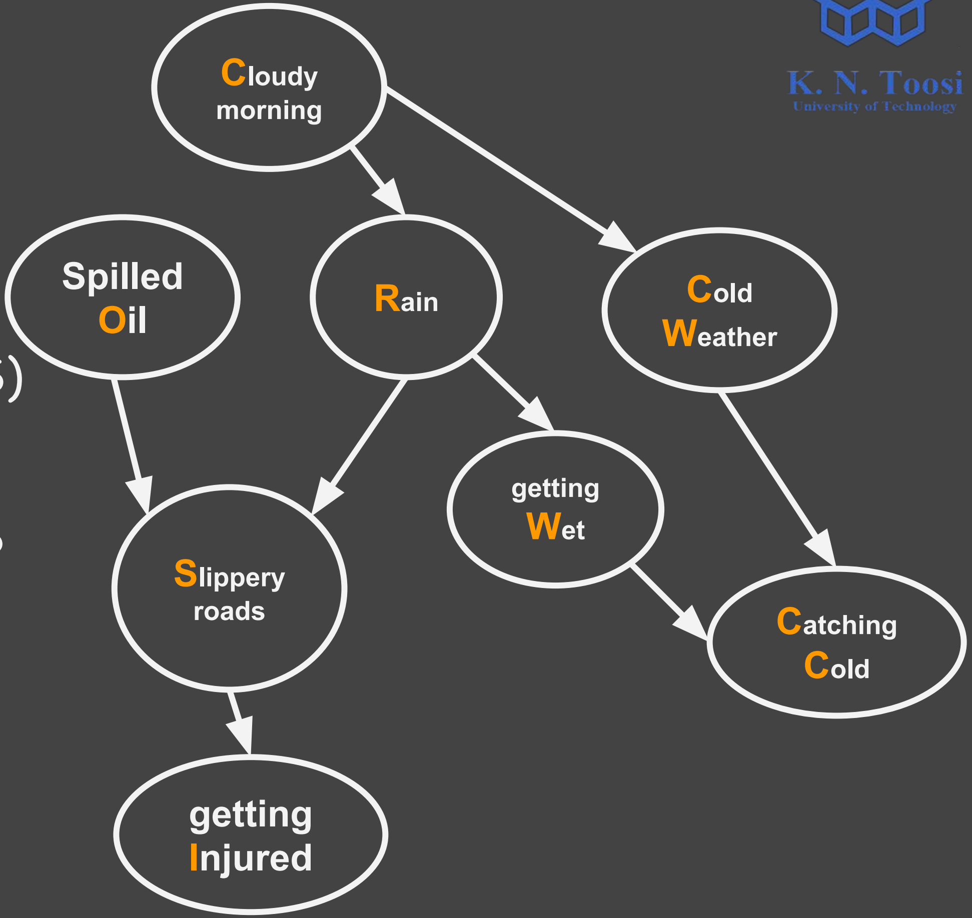
# Example

X, Y: nodes

S: a subset of nodes

An independence relation: $(X \perp Y \mid S)$

What are all the independence relations one can read from a graph?



# Example

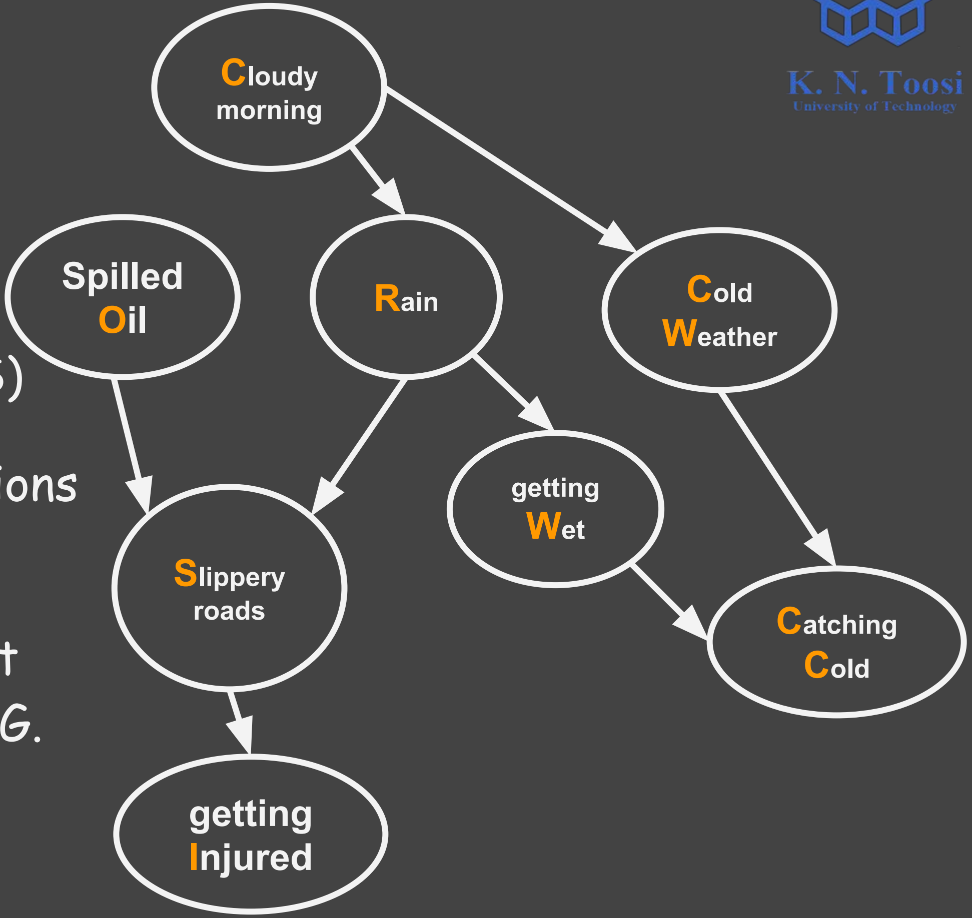
X, Y: nodes

S: a subset of nodes

An independence relation: $(X \perp Y \mid S)$

What are all the independence relations one can read from a graph?

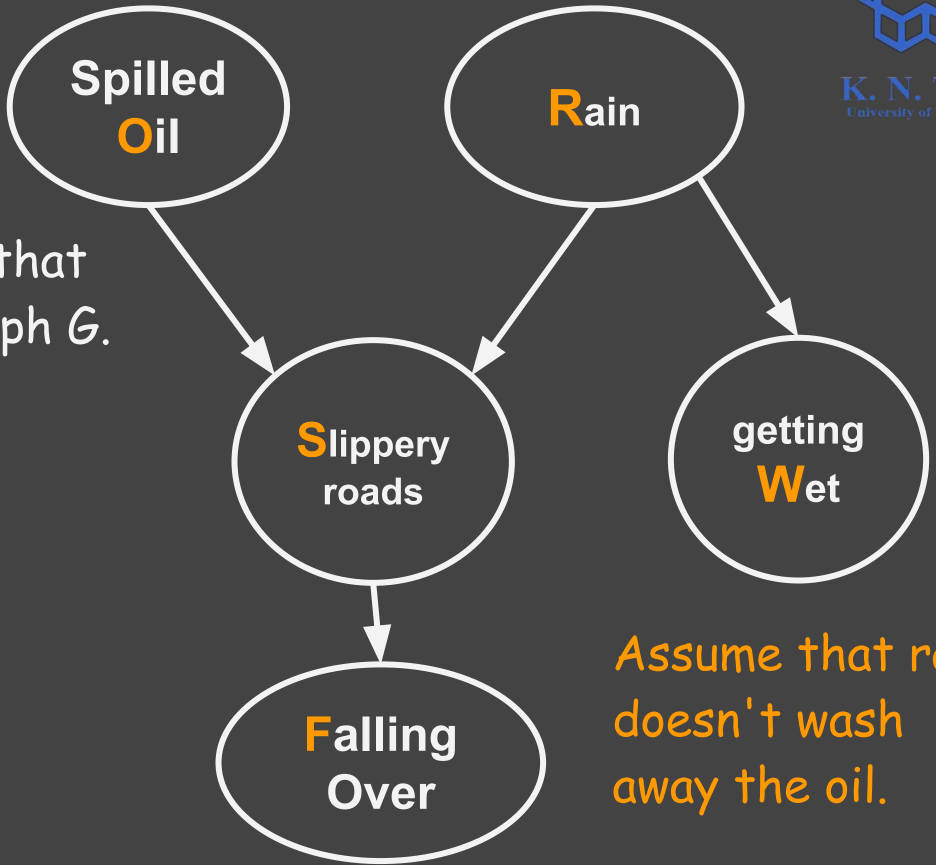
I(G) = idependence relationships that
could be read from the graph G.



# I-Map

I(G) = idependence relationships that could be read from the graph G.

I(G) = {

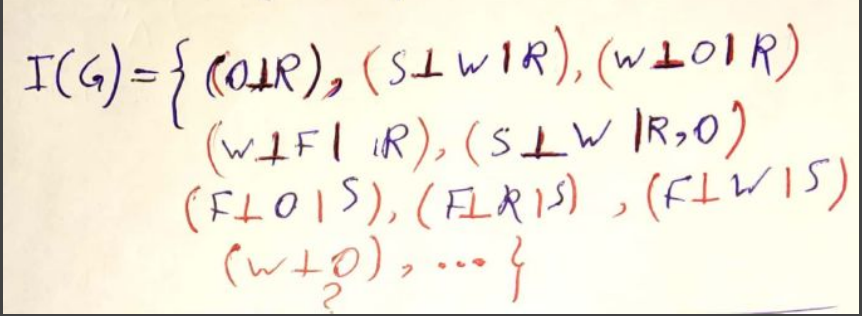


Assume that rain doesn't wash away the oil.

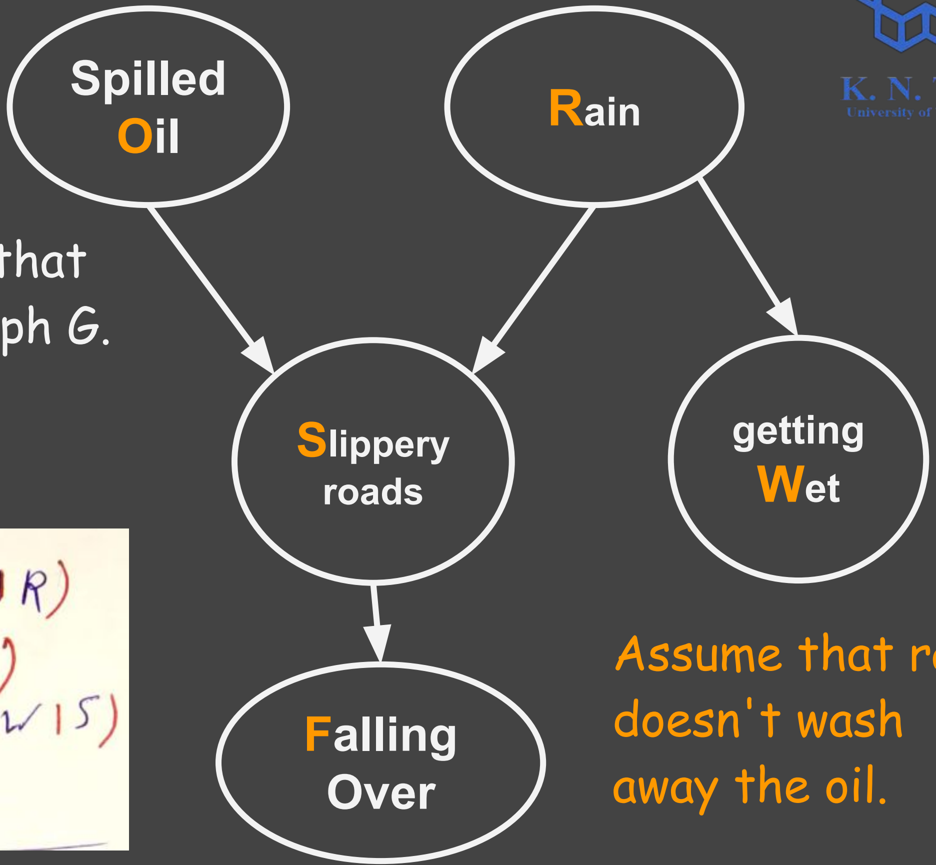
# I-Map

I(G) = idependence relationships that could be read from the graph G.

I(G) = {



$I(G) = \{ (O \perp R), (S \perp W | R), (W \perp O | R)$
$(W \perp F | R), (S \perp W | R, O)$
$(F \perp O | S), (F \perp R | S), (F \perp W | S)$
$(W \perp \underset{?}{O}), \ldots \}$



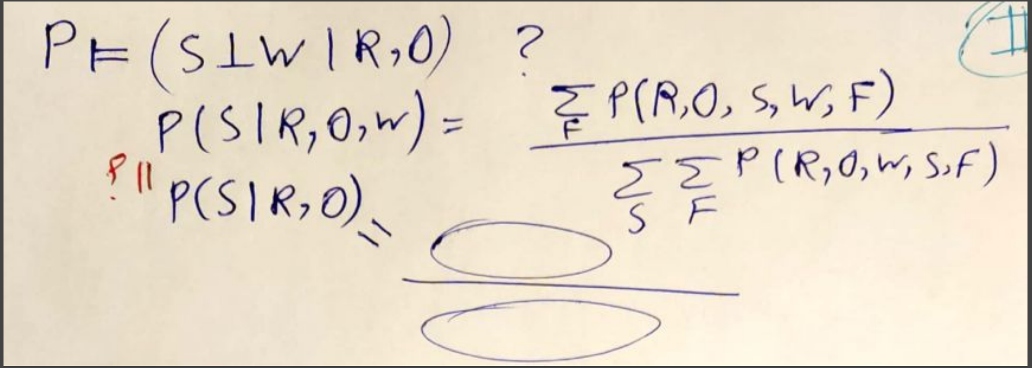
Assume that rain doesn't wash away the oil.

# I-Map

X, Y: nodes

S: a subset of nodes

I(P) = { (X ⊥ Y | S) | P ⊨ (X ⊥ Y | S) }
= independence relations the joint distribution P satisfies.

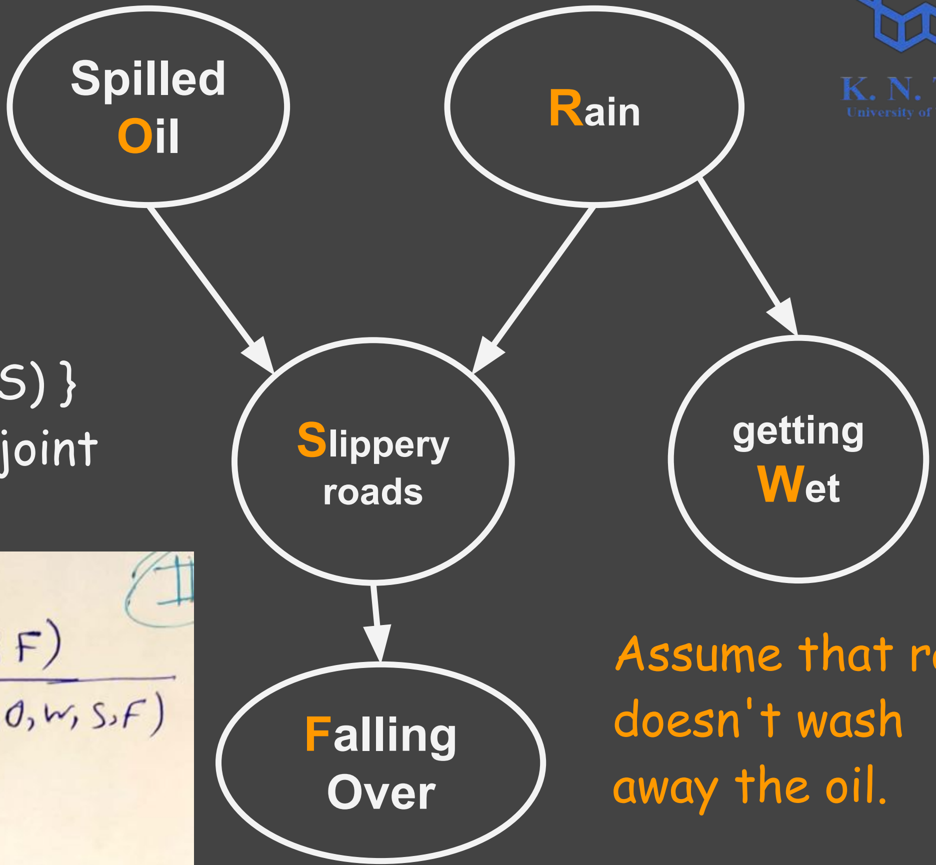


$P \models (S \perp W \mid R, O)$ ?

$$P(S \mid R,O,W) = \frac{\sum_F P(R,O,S,W,F)}{\sum_S \sum_F P(R,O,W,S,F)}$$

$\overset{?}{=}$

$$P(S \mid R,O) =$$



Spilled **O**il

**R**ain

**S**lippery roads

getting **W**et

**F**alling Over

Assume that rain doesn't wash away the oil.

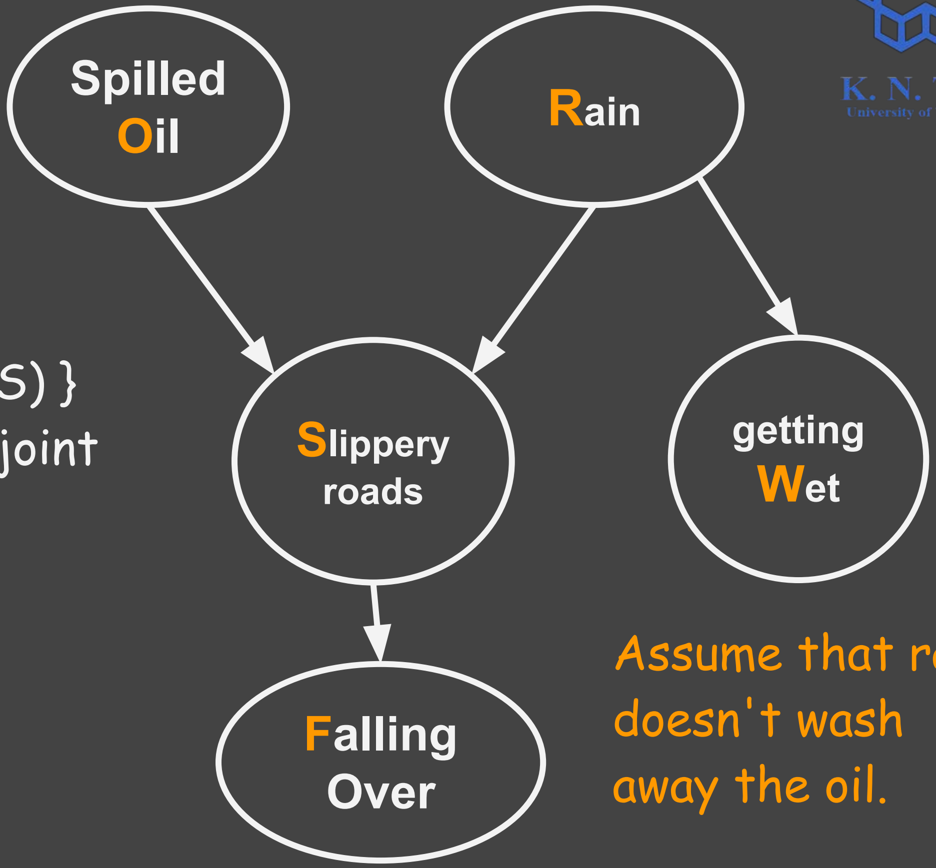
# I-Map

X, Y: nodes

S: a subset of nodes

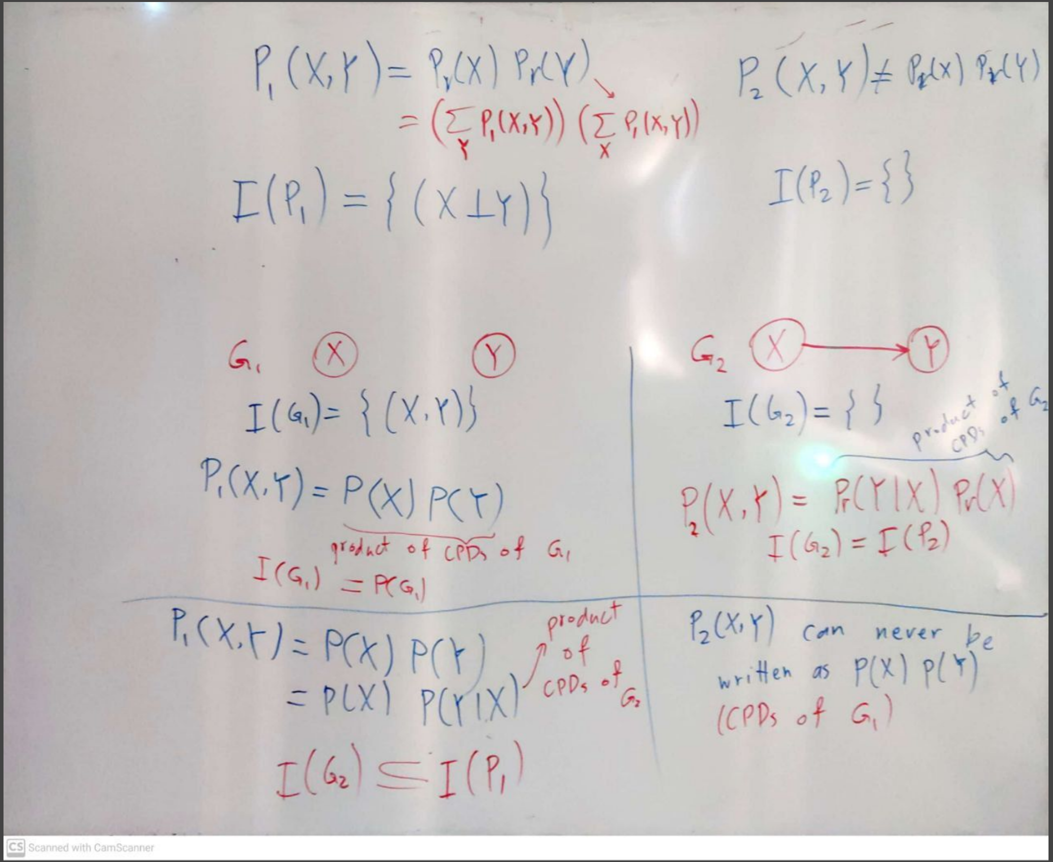
$I(P) = \{ (X \perp Y \mid S) \mid P \models (X \perp Y \mid S) \}$

= independence relations the joint distribution P satisfies.



Assume that rain doesn't wash away the oil.

# Example



$P_1(X,Y) = P_r(X)\,P_r(Y)$

$= \left(\sum_Y P_1(X,Y)\right)\left(\sum_X P_1(X,Y)\right)$

$I(P_1) = \{(X \perp Y)\}$

$P_2(X,Y) \neq P_r(X)\,P_r(Y)$

$I(P_2) = \{\}$

$G_1$ $X$ $Y$

$I(G_1) = \{(X, Y)\}$

$P_1(X,Y) = \underbrace{P(X)\,P(Y)}_{\text{product of CPDs of } G_1}$

$I(G_1) = P(G_1)$

$G_2$ $X \rightarrow Y$

$I(G_2) = \{\}$

$P_2(X,Y) = \overbrace{P_r(Y|X)\,P_r(X)}^{\text{product of CPDs of } G_2}$

$I(G_2) = I(P_2)$

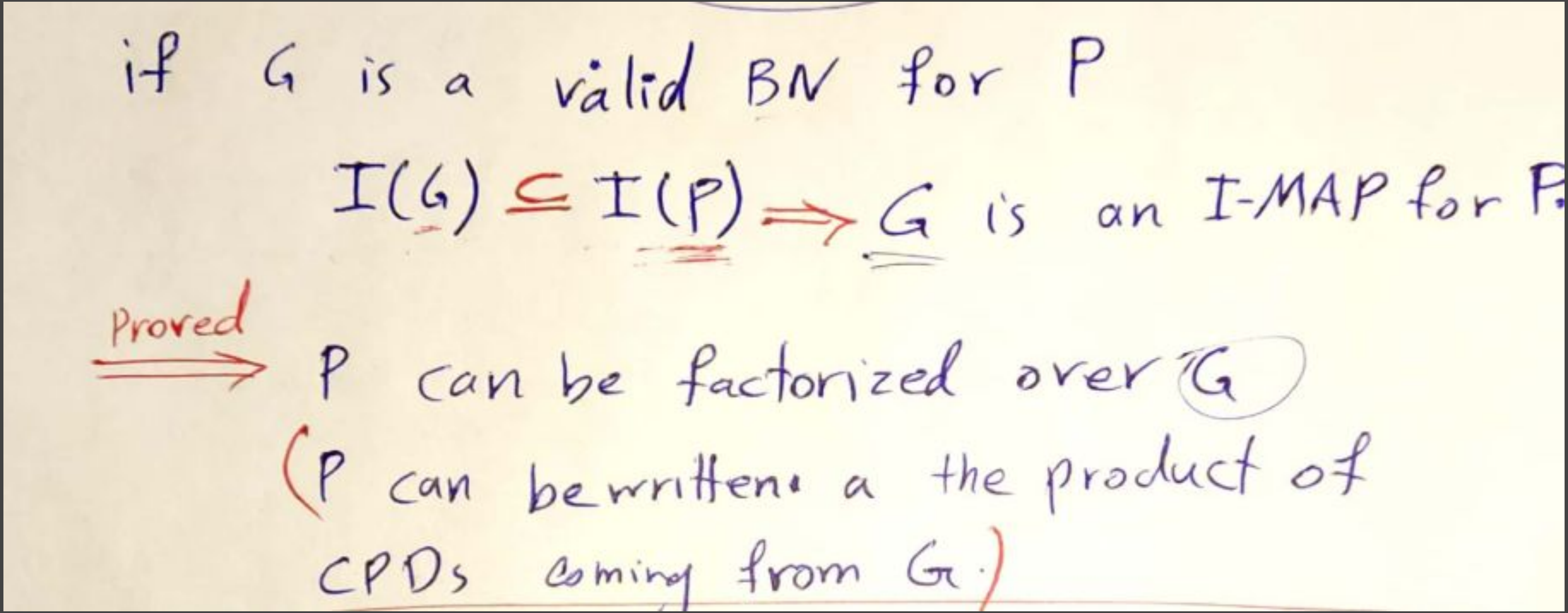
$P_1(X,Y) = P(X)\,P(Y)$

$= P(X)\,P(Y|X)$ product of CPDs of $G_2$

$I(G_2) \subseteq I(P_1)$

$P_2(X,Y)$ can never be written as $P(X)\,P(Y)$ (CPDs of $G_1$)

# I-Map



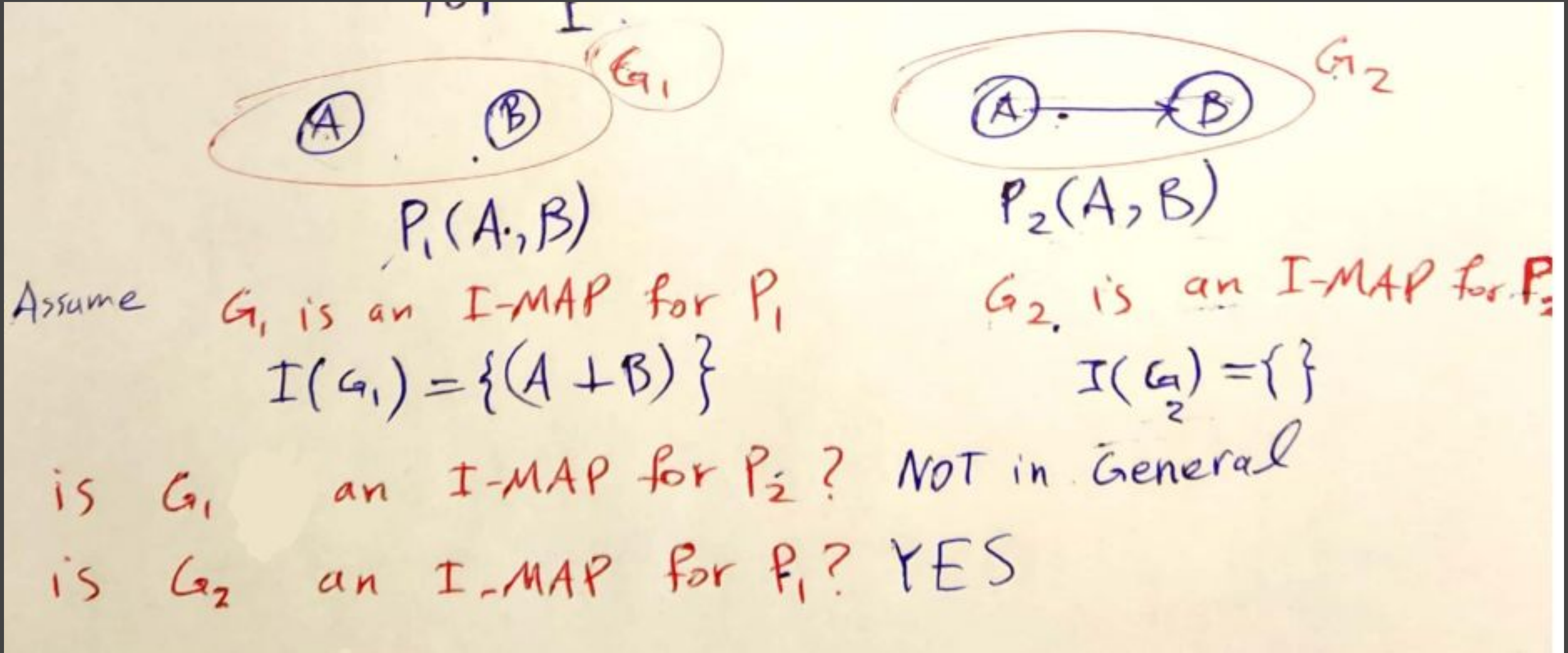
if G is a valid BN for P

$I(G) \subseteq I(P) \Rightarrow$ G is an I-MAP for P.

$\xRightarrow{\text{Proved}}$ P can be factorized over G

(P can be written a the product of CPDs coming from G.)

# Example



$P_1(A, B)$

$P_2(A, B)$

Assume $G_1$ is an I-MAP for $P_1$

$G_2$ is an I-MAP for P

$I(G_1) = \{(A \perp B)\}$

$I(G_2) = \{\}$

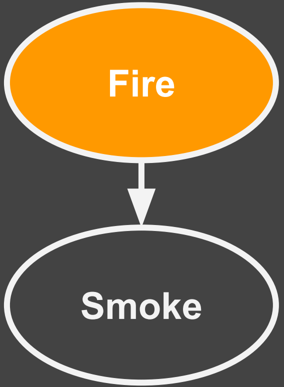
is $G_1$ an I-MAP for $P_2$? NOT in General

is $G_2$ an I-MAP for $P_1$? YES

# Perfect Map

The graph G is a **Perfect Map** for the joint distribution P if **I(G) = I(P)**.

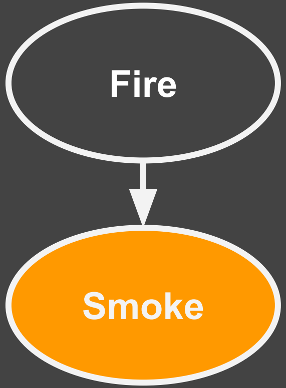
# Causal Reasoning



Fire

Smoke

# Evidential Reasoning

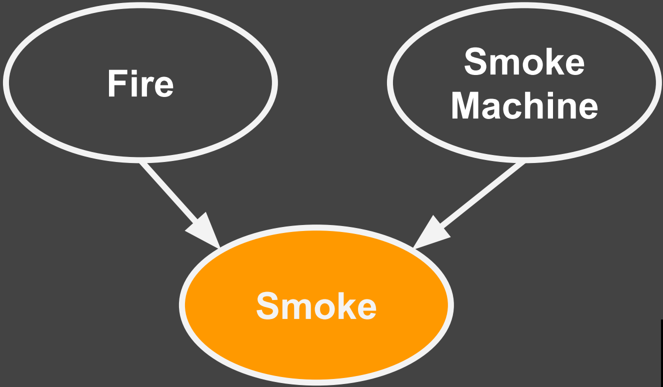


Fire

Smoke

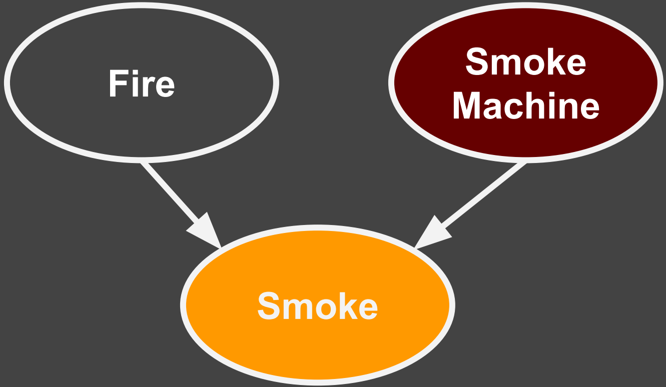
# Intercausal Reasoning

K. N. Toosi
University of Technology



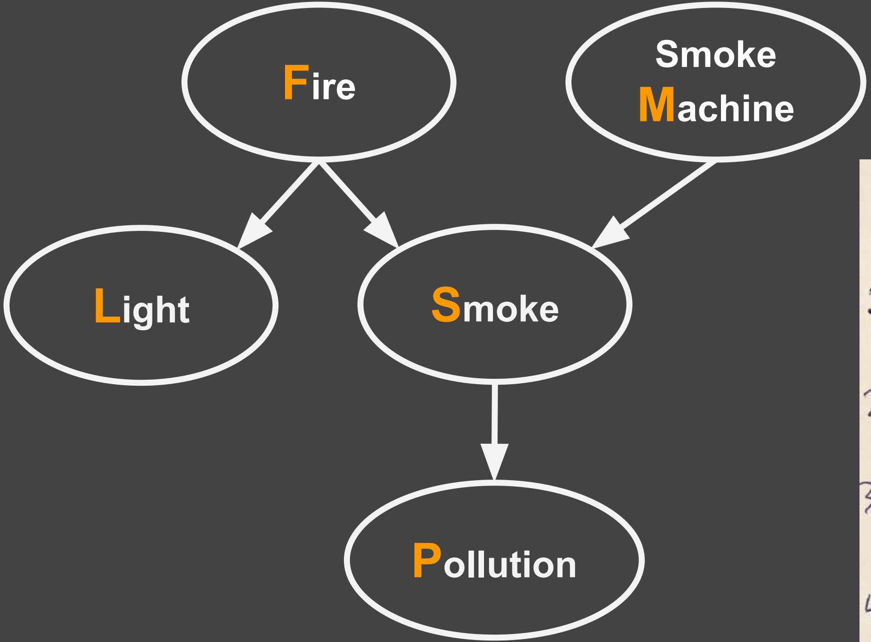
P(Fire | Smoke)

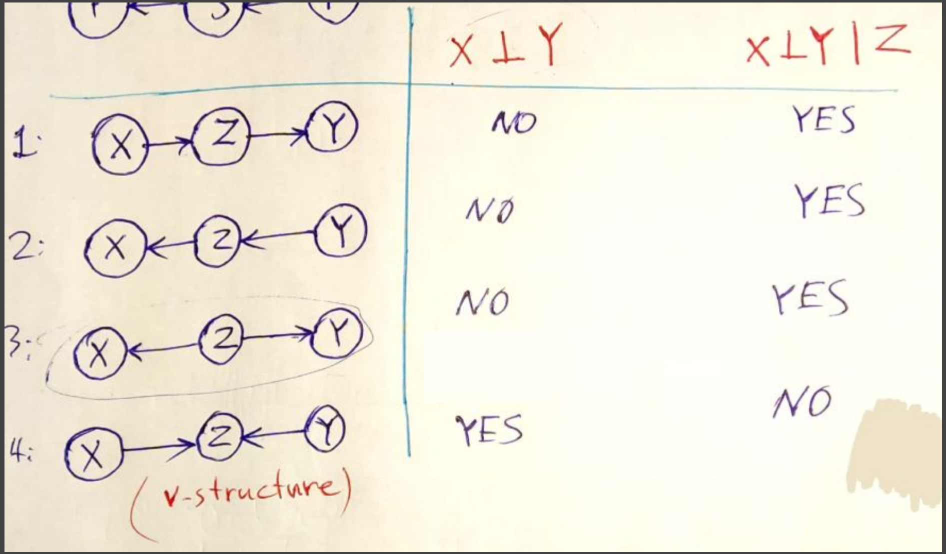
# Intercausal Reasoning



$P(\text{Fire}=1 \mid \text{Smoke}=1, \text{S-Machine}=1)$
$< P(\text{Fire}=1 \mid \text{Smoke}=1)$

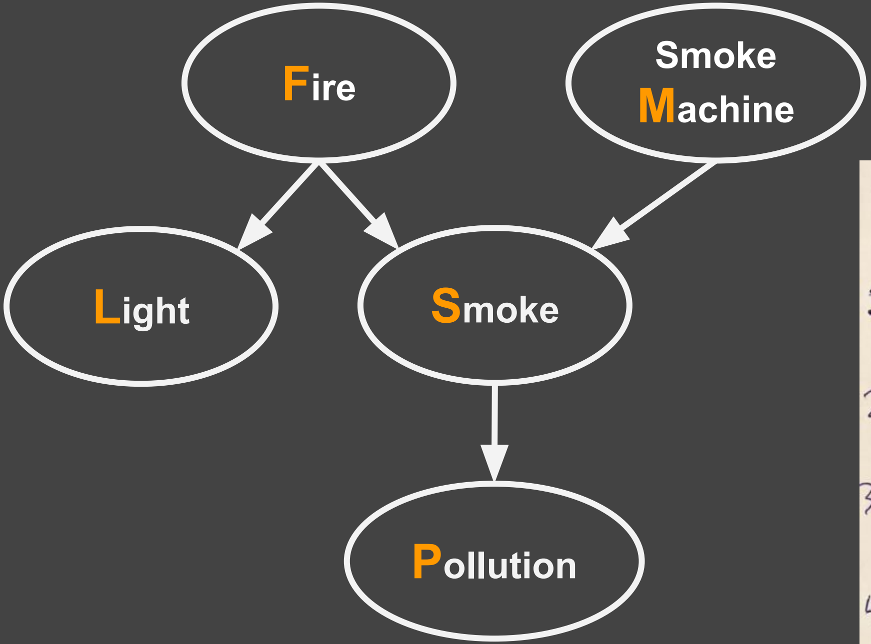
# Flow of influence





| | $X \perp Y$ | $X \perp Y \mid Z$ |
|---|---|---|
| 1: $X \rightarrow Z \rightarrow Y$ | NO | YES |
| 2: $X \leftarrow Z \leftarrow Y$ | NO | YES |
| 3: $X \leftarrow Z \rightarrow Y$ | NO | YES |
| 4: $X \rightarrow Z \leftarrow Y$ (v-structure) | YES | NO |

# Flow of influence



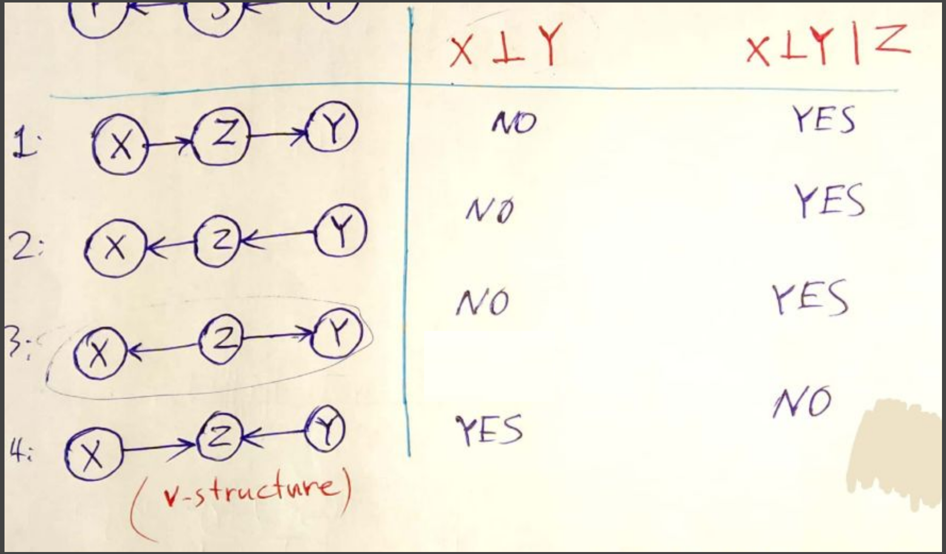
Fire

Smoke Machine

Light

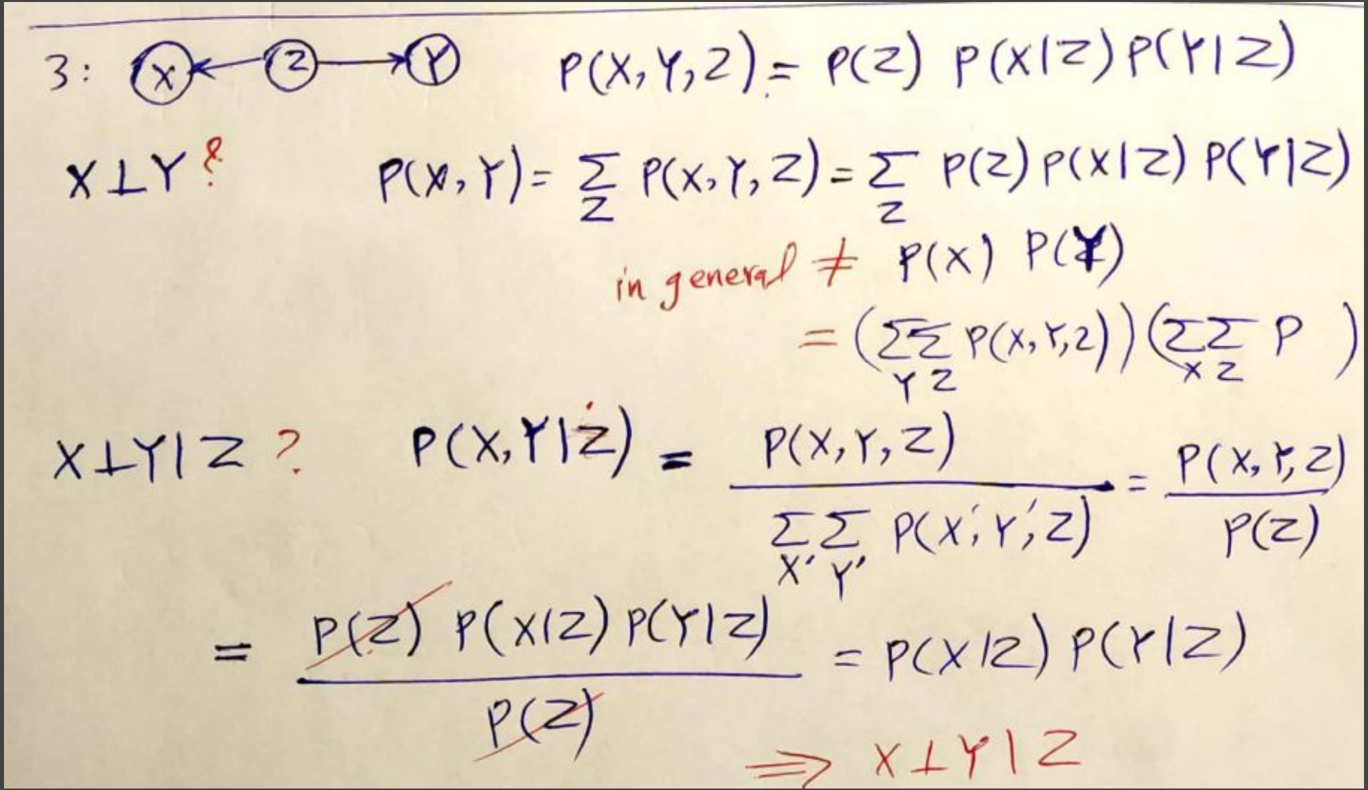
Smoke

Pollution



| | $X \perp Y$ | $X \perp Y \mid Z$ |
|---|---|---|
| 1: $X \rightarrow Z \rightarrow Y$ | NO | YES |
| 2: $X \leftarrow Z \leftarrow Y$ | NO | YES |
| 3: $X \leftarrow Z \rightarrow Y$ | NO | YES |
| 4: $X \rightarrow Z \leftarrow Y$<br>(v-structure) | YES | NO |

# Example



3: $X \leftarrow Z \rightarrow Y$ $\qquad P(X,Y,Z) = P(Z)\, P(X|Z)\, P(Y|Z)$

$X \perp Y$ ?

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

in general $\neq P(X)\ P(Y)$

$$= \Big(\sum_Y \sum_Z P(X,Y,Z)\Big)\Big(\sum_X \sum_Z P \quad \Big)$$

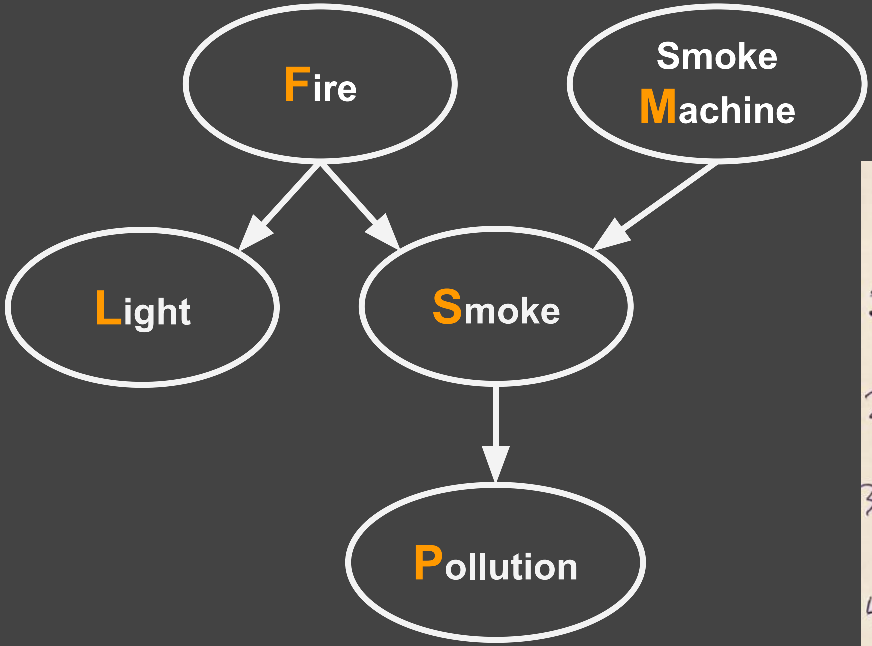
$X \perp Y | Z$ ?

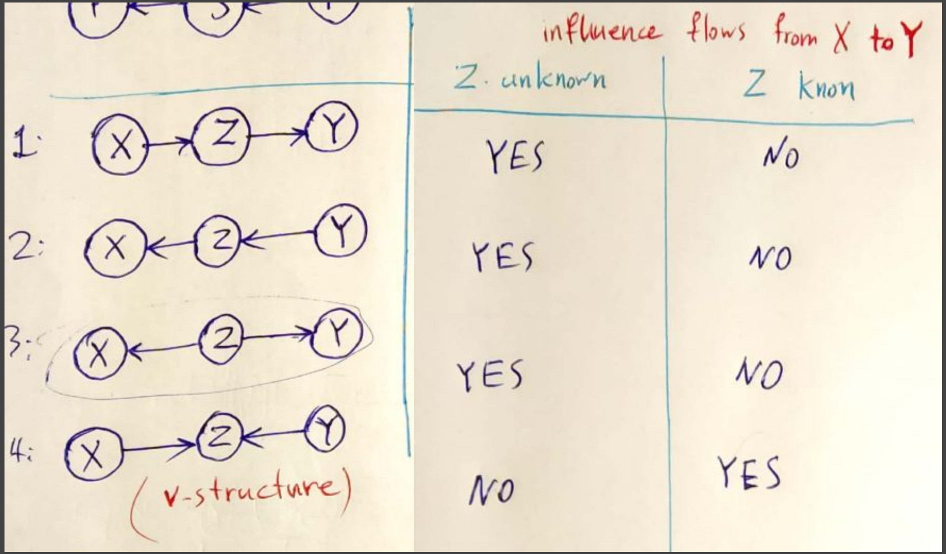
$$P(X,Y|Z) = \frac{P(X,Y,Z)}{\sum_{X'}\sum_{Y'} P(X',Y',Z)} = \frac{P(X,Y,Z)}{P(Z)}$$

$$= \frac{\cancel{P(Z)}\ P(X|Z)\,P(Y|Z)}{\cancel{P(Z)}} = P(X|Z)\,P(Y|Z)$$

$\Rightarrow X \perp Y | Z$

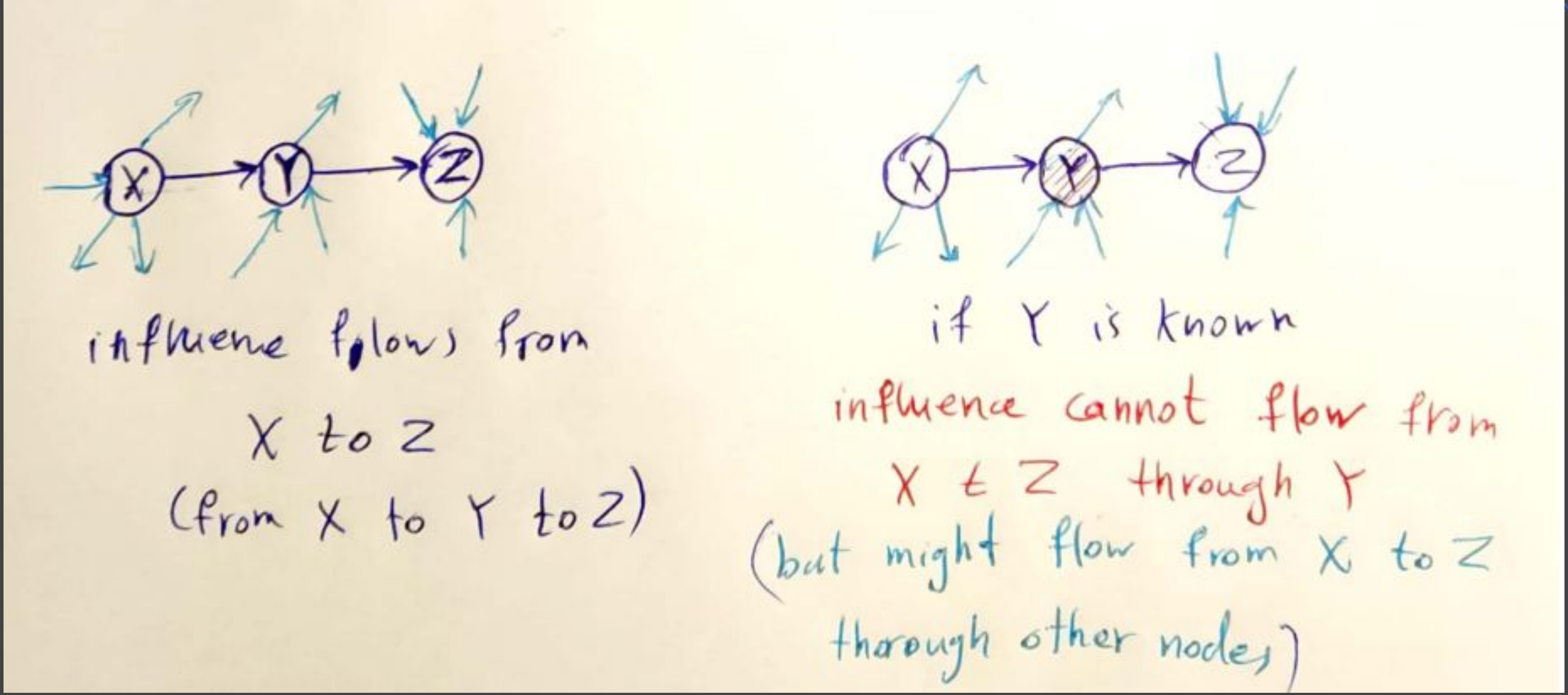
# Flow of influence





| | influence flows from X to Y | |
|---|---|---|
| | Z. unknown | Z knon |
| 1: $X \rightarrow Z \rightarrow Y$ | YES | No |
| 2: $X \leftarrow Z \leftarrow Y$ | YES | NO |
| 3: $X \leftarrow Z \rightarrow Y$ | YES | NO |
| 4: $X \rightarrow Z \leftarrow Y$ (v-structure) | NO | YES |

# What if there are more than 3 nodes?



influene flows from

X to Z

(from X to Y to Z)

if Y is known

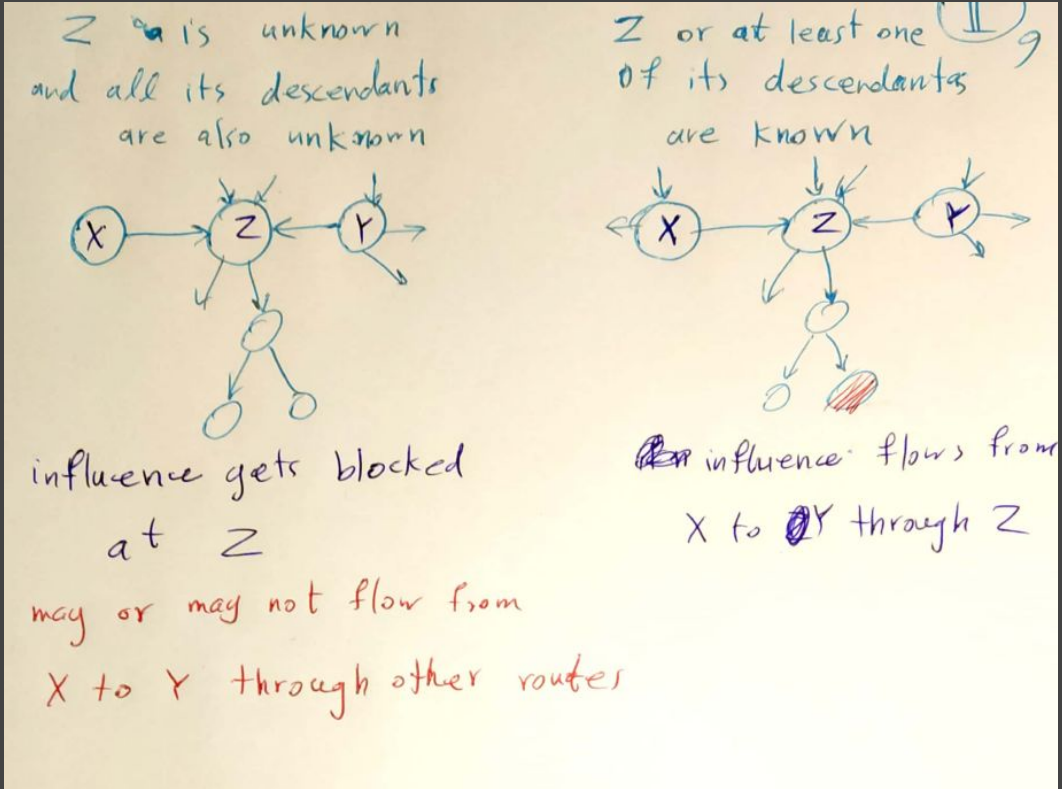
influence cannot flow from

X t Z through Y

(but might flow from X to Z

through other nodes)

# What if there are more than 3 nodes?



Z is unknown
and all its descendants
are also unknown

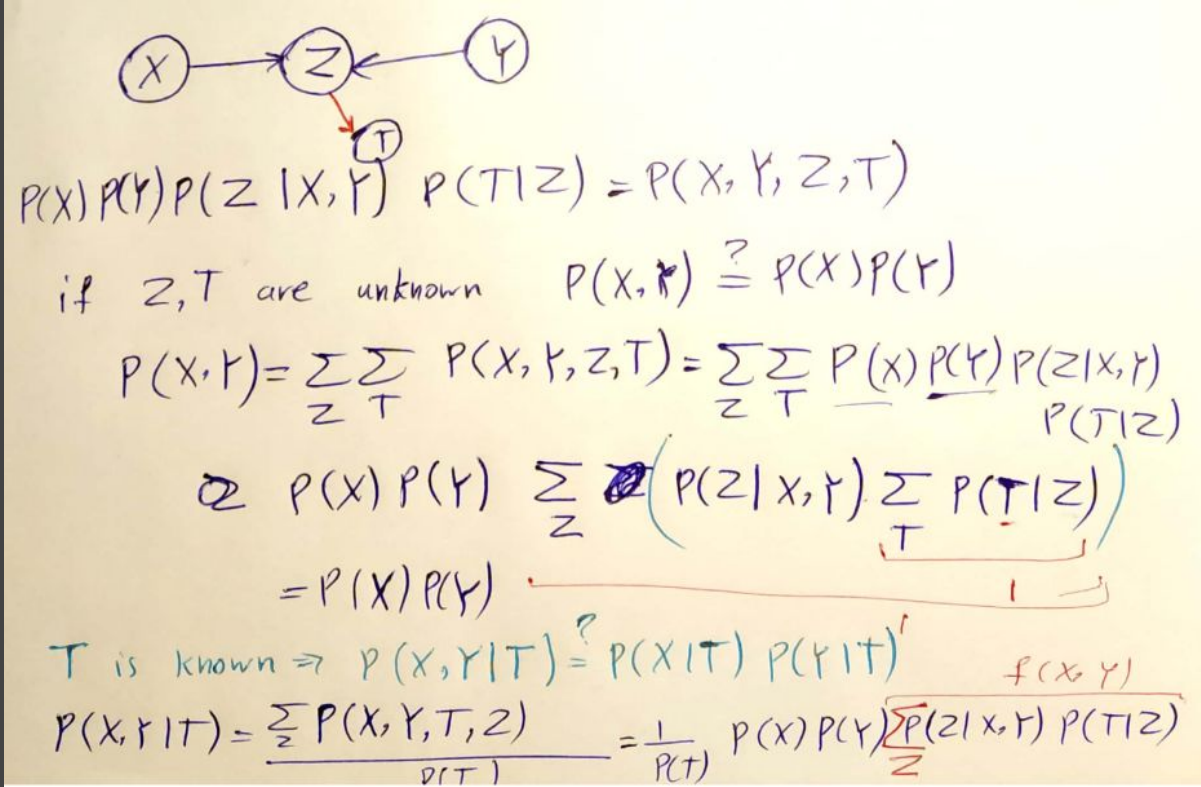
influence gets blocked
at Z

may or may not flow from
X to Y through other routes

Z or at least one
of its descendants
are known

influence flows from
X to Y through Z

# Example



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

if Z,T are unknown $P(X,Y) \overset{?}{=} P(X)P(Y)$

$$P(X,Y) = \sum_Z \sum_T P(X,Y,Z,T) = \sum_Z \sum_T P(X)P(Y)P(Z|X,Y)\,P(T|Z)$$

$$= P(X)P(Y) \underbrace{\sum_Z \left( P(Z|X,Y) \underbrace{\sum_T P(T|Z)}_{1} \right)}_{1}$$

$$= P(X)P(Y)$$

T is known ⇒ $P(X,Y|T) \overset{?}{=} P(X|T)\,P(Y|T)$

$$P(X,Y|T) = \frac{\sum_Z P(X,Y,T,Z)}{P(T)} = \frac{1}{P(T)} P(X)P(Y) \overbrace{\sum_Z P(Z|X,Y)\,P(T|Z)}^{f(X,Y)}$$

# Active trail

Given a set of observed nodes in a Bayesian network, a route $X_1, X_2, \ldots, X_n$ between two nodes $X_1$ and $X_n$ is an active trail if for all $1 < i < n$

- if $X_{i-1}, X_i, X_{i+1}$ is a v-structure then $X_i$ or one of its descendants are observed (known)
- if $X_{i-1}, X_i, X_{i+1}$ is not a v-structure, $X_i$ is not observed (is unknown).

# d-separation

Two nodes X, Y in a Bayesian Network are d-separated if there is no active trail between X,Y. => There is no route from X to Y through which the influence can flow.

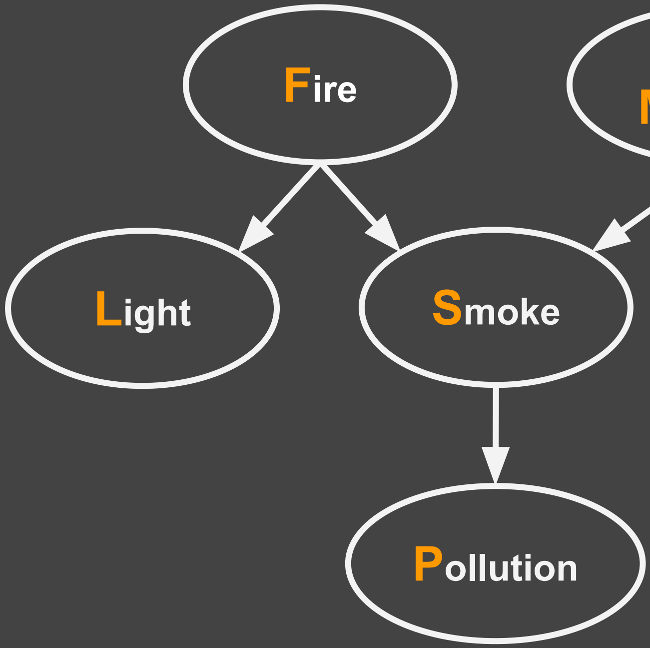
If X,Y are d-separeted given a set of observations S (subset of V) =>

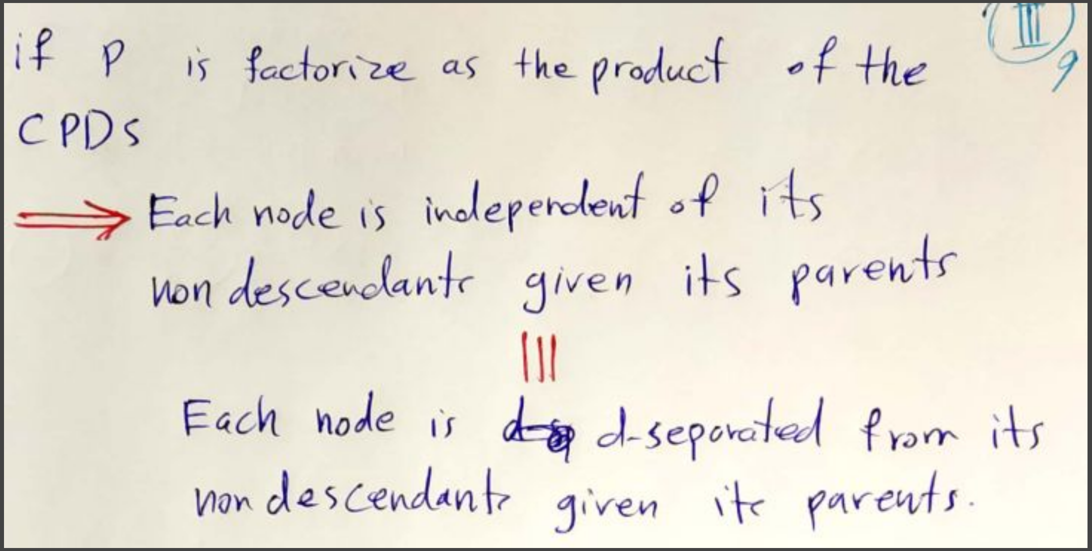
Then, in the corresponding joint distribution X and Y are independent given S

# Markov Independence => Factorization

Proved before!

# Factorization => Independence





if p is factorize as the product of the CPDs

⟹ Each node is independent of its non descendants given its parents

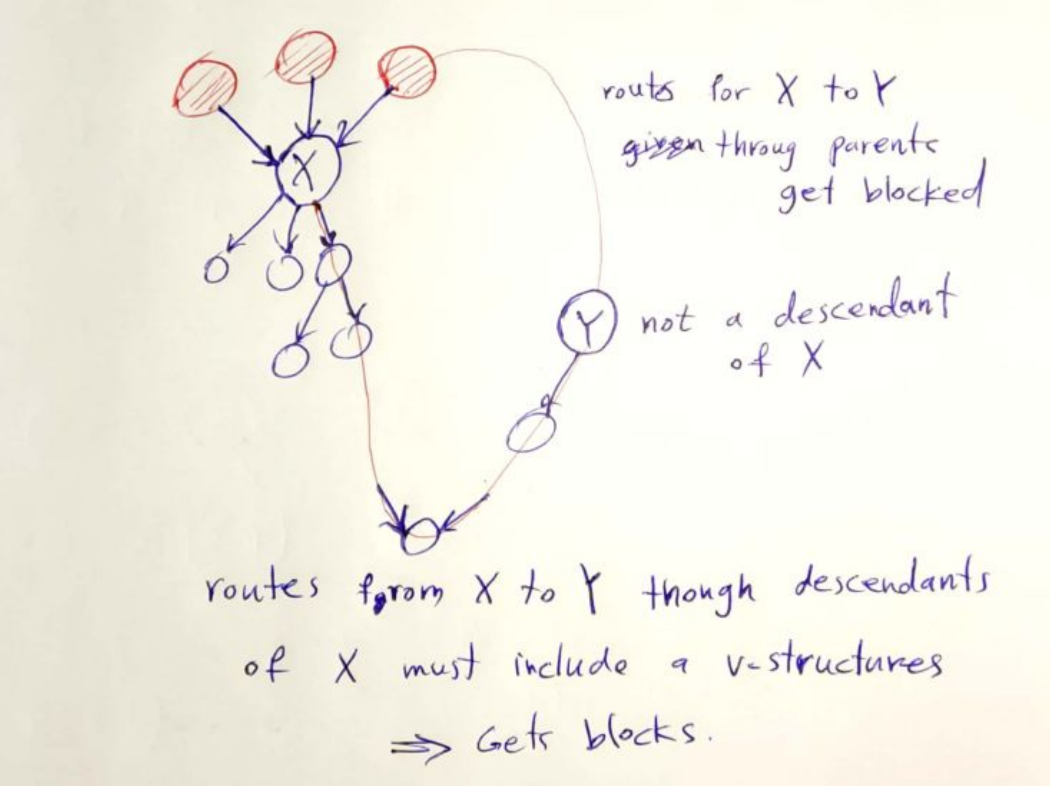
|||

Each node is ~~d-s~~ d-separated from its non descendants given its parents.

III

9

# Factorization => Independence



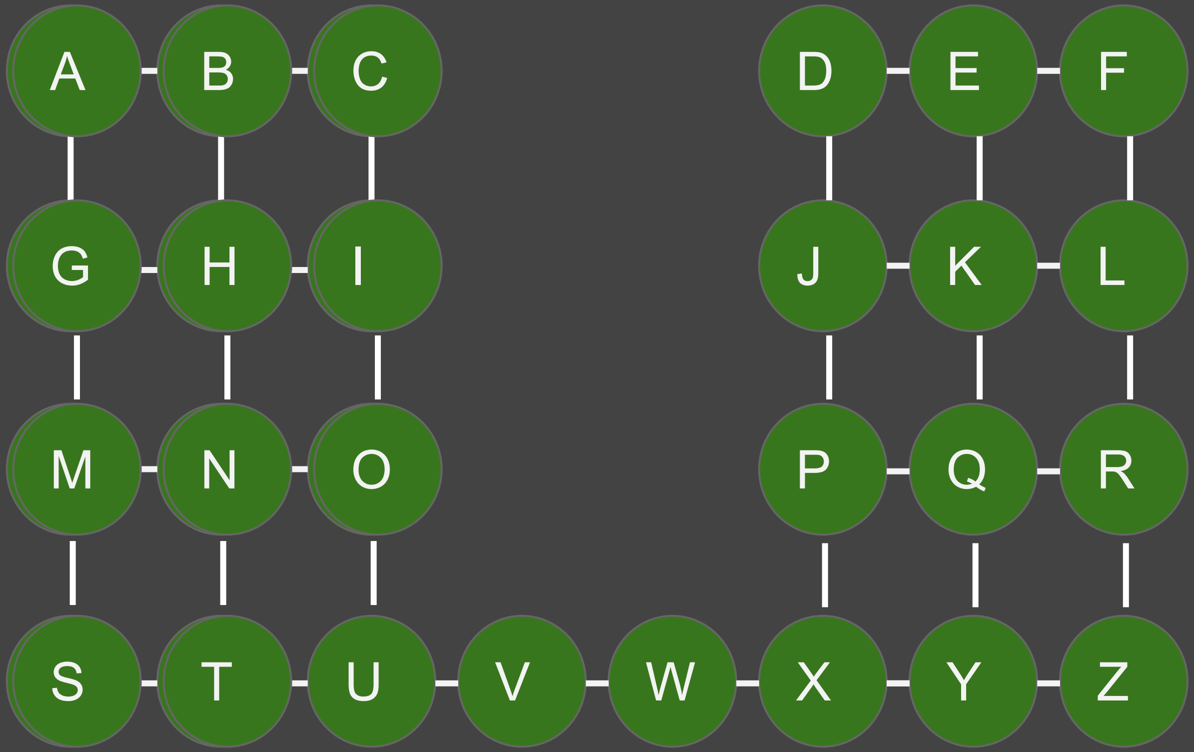
routs for X to Y ~~given~~ throug parents get blocked

not a descendant of X

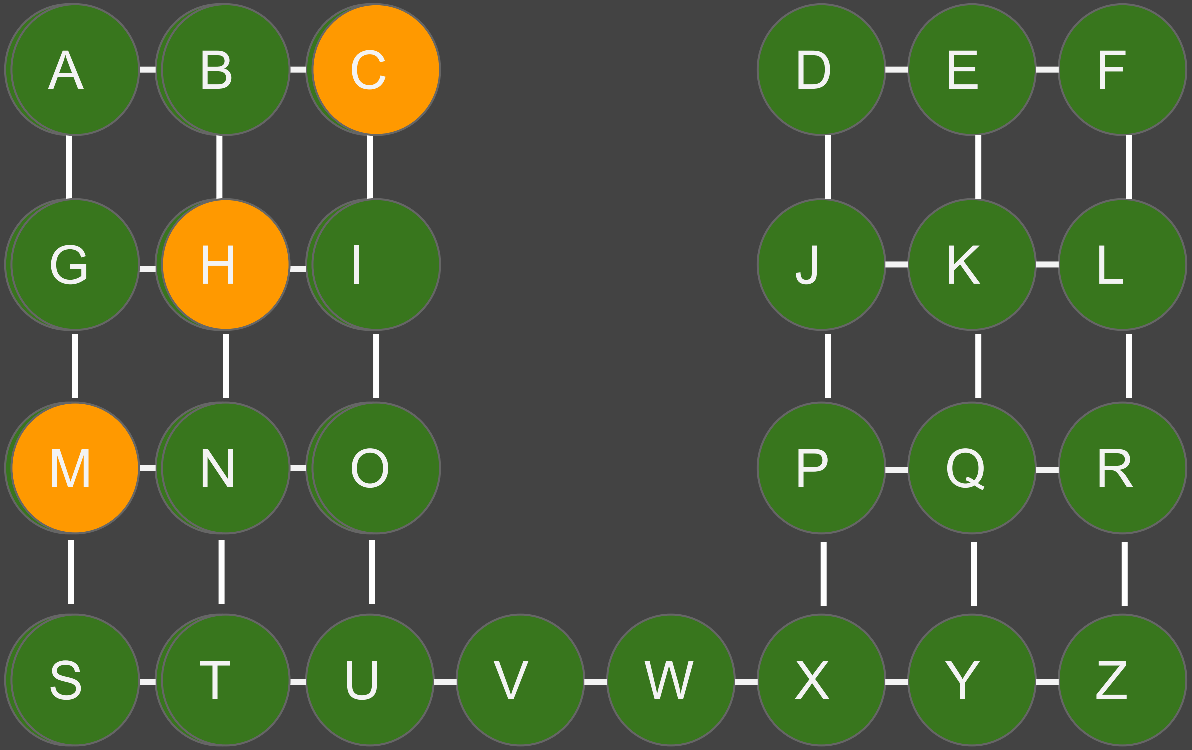
routes from X to Y though descendants of X must include a v-structures

⇒ Gets blocks.

# Markov Networks



# Markov Networks



# MRFs - Active trail

A route $X_1, X_2, \ldots, X_n$ between $X_1$ and $X_n$ is an active trail if for all $1 < i < n$, $X_i$ is not observed (is unknown).

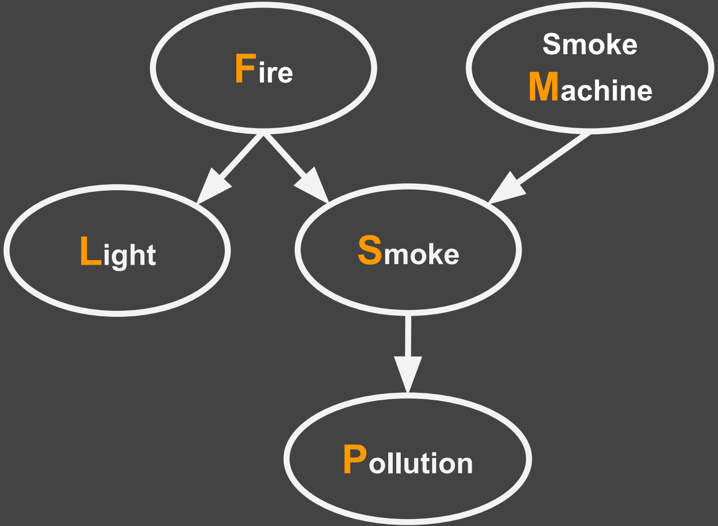
# MRFs - Sepration

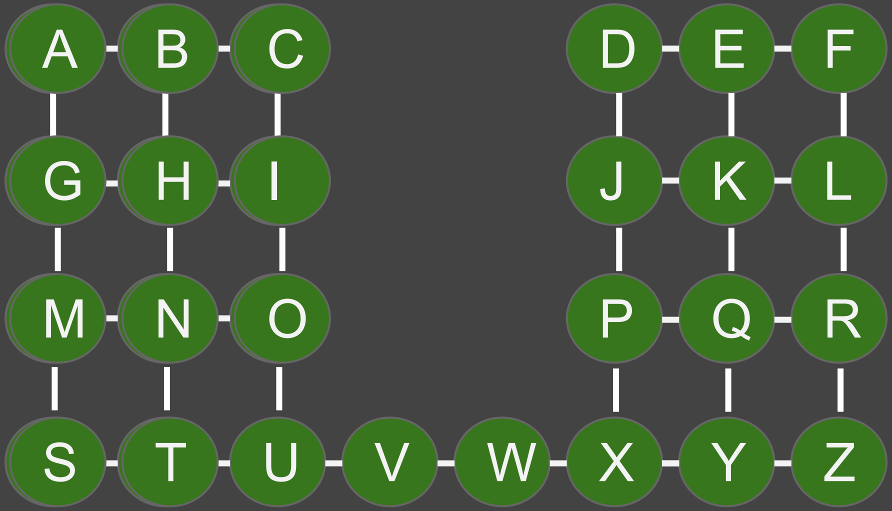
Two nodes X, Y in a Markov Network are **separated** if there is no active trail between X,Y.

Remember the three types of Markov property (pariwise, local, and global). The above is related to the Global Markov Property. We assume that the joint distribution is strictly positive, thus the three properties are equivalent.

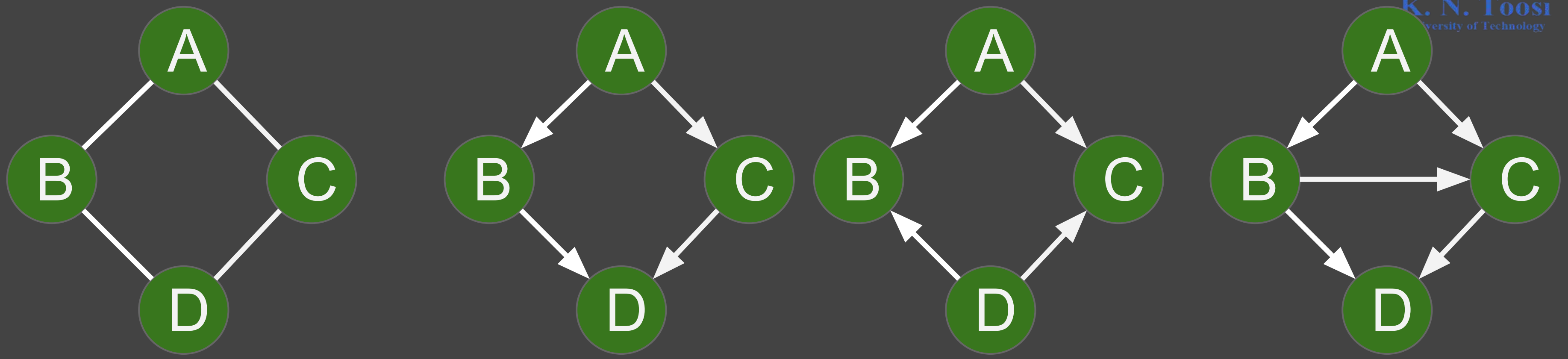
# Which one is more general? MN or BN?

K. N. Toosi
University of Technology



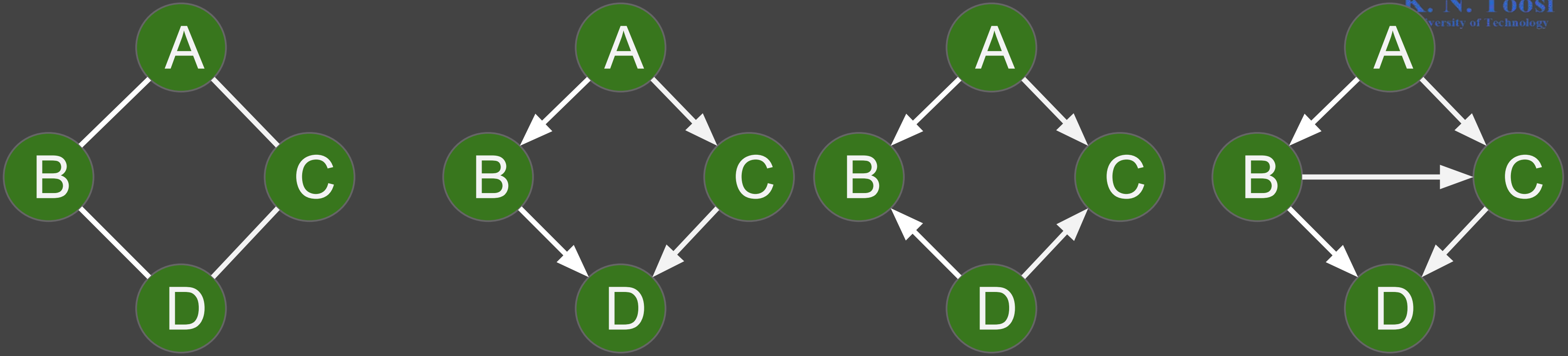


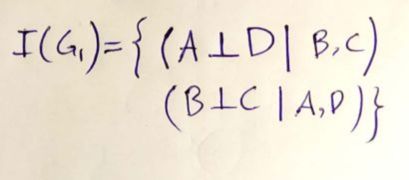
# Which one is more general? MN or BN?



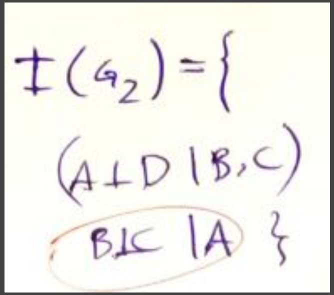
Koller

# Which one is more general? MN or BN?

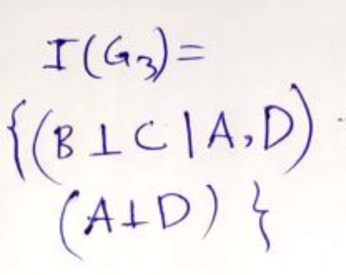




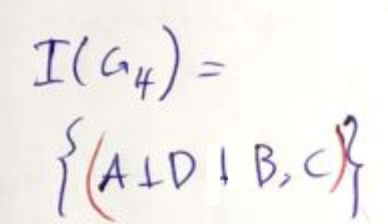
$I(G_1) = \{ (A \perp D \mid B, C)$
$(B \perp C \mid A, D) \}$



$I(G_2) = \{$
$(A \perp D \mid B, C)$
$\boxed{B \perp C \mid A} \}$



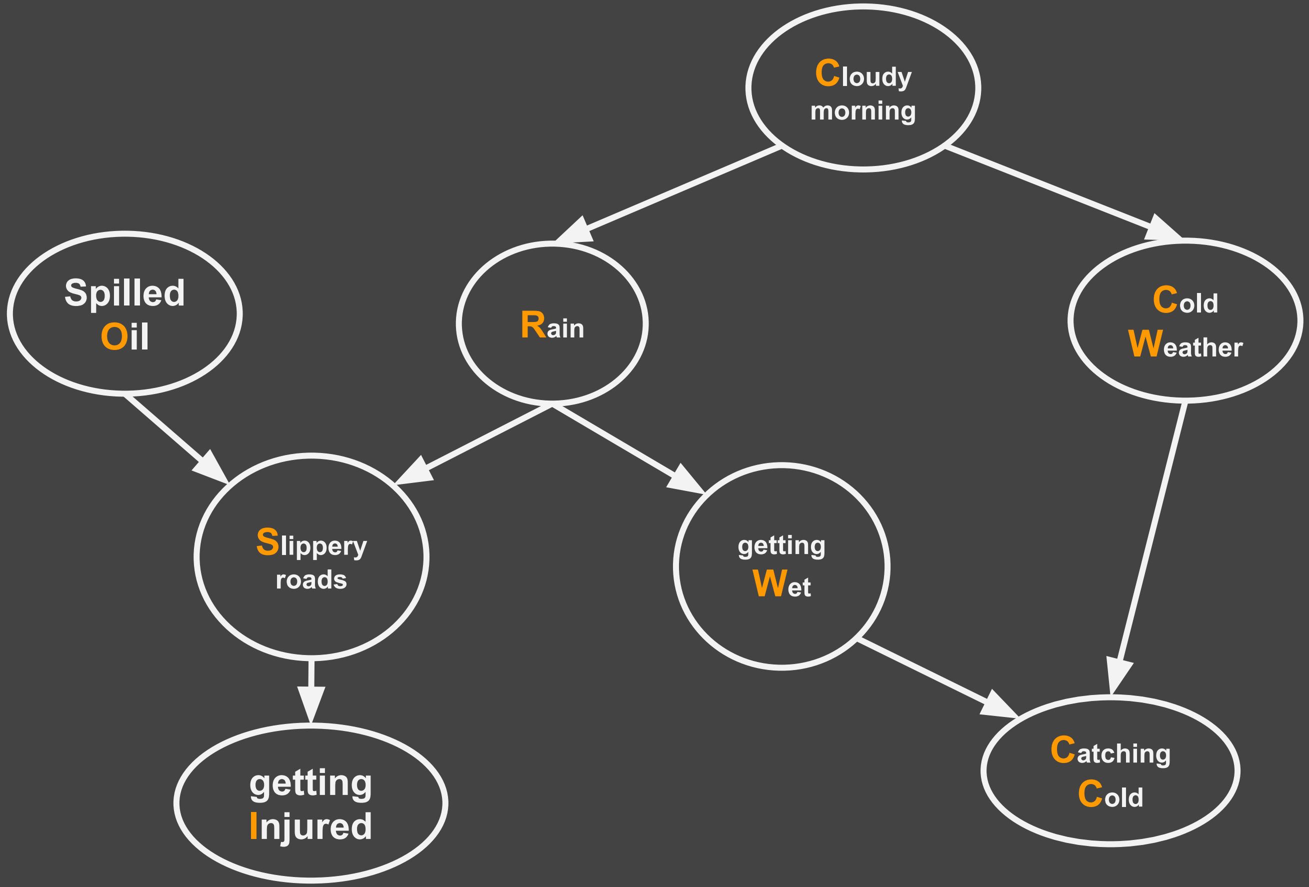
$I(G_3) =$
$\{ (B \perp C \mid A, D)$
$(A \perp D) \}$



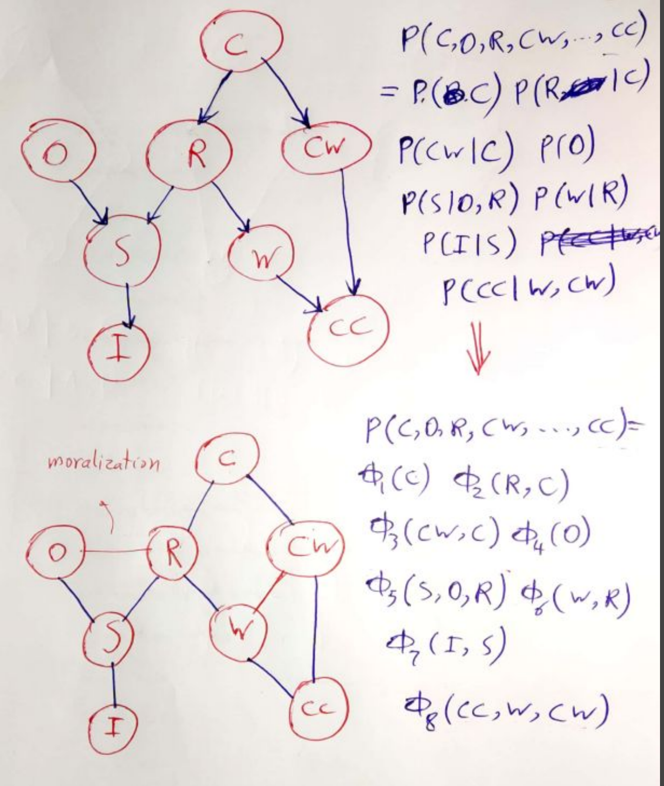
$I(G_4) =$
$\{ (A \perp D \mid B, C) \}$

Koller

# BN as MRFs, Moralization



# BN as MRFs, Moralization



$P(C, O, R, CW, \dots, CC)$

$= P(C)\, P(R|C)$

$P(CW|C)\; P(O)$

$P(S|O,R)\; P(W|R)$

$P(I|S)$ ~~P(CC|W,C~~

$P(CC|W,CW)$

$\Downarrow$

$P(C, O, R, CW, \dots, CC) =$

$\Phi_1(C)\; \Phi_2(R, C)$

$\Phi_3(CW, C)\; \Phi_4(O)$

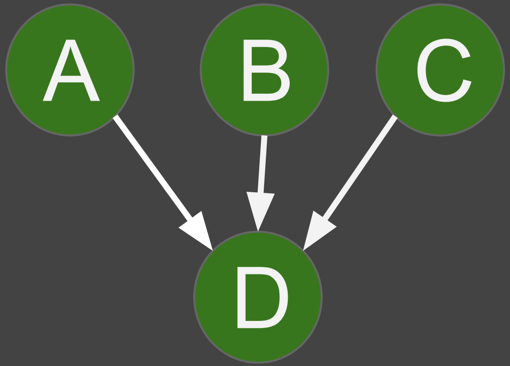
$\Phi_5(S, O, R)\; \Phi_6(W, R)$

$\Phi_7(I, S)$

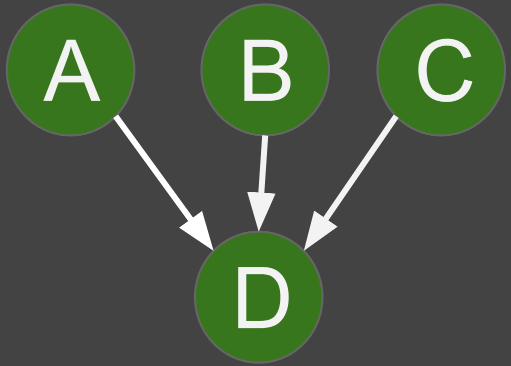
$\Phi_8(CC, W, CW)$

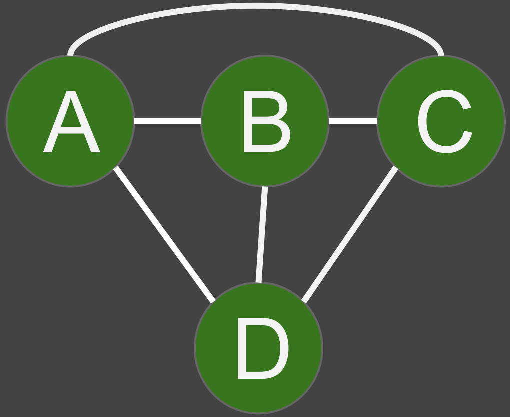
# BN as MRFs, Moralization

# BN as MRFs, Moralization

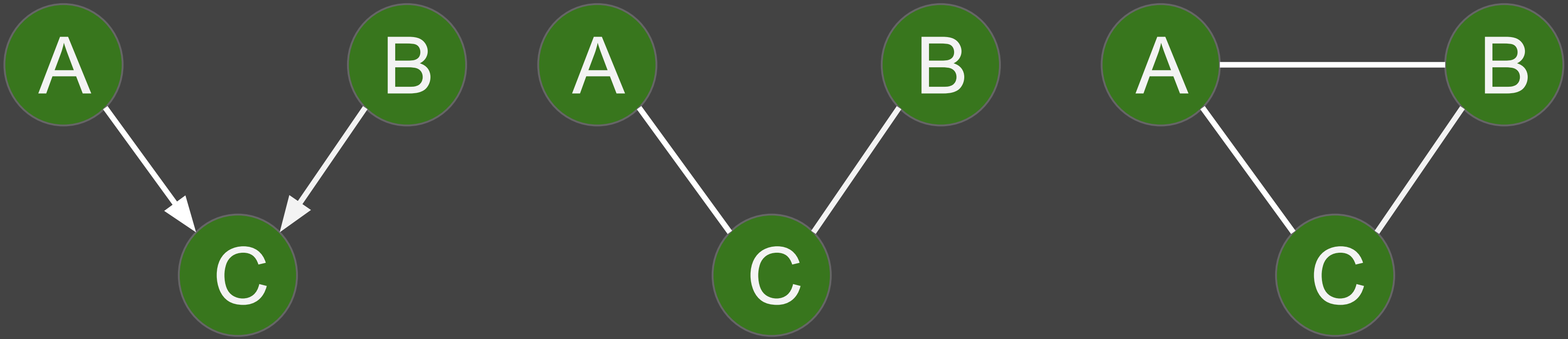


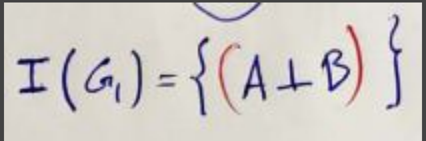
# BN as MRFs, Moralization

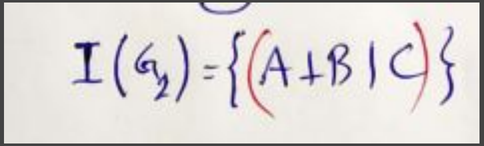


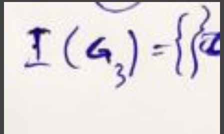


# BN as MRFs, Moralization

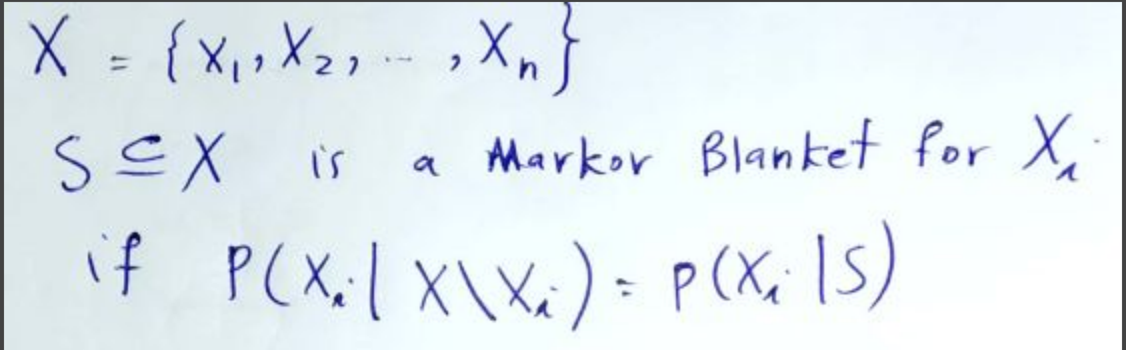



$I(G_1) = \{(A \perp B)\}$


$I(G_2) = \{(A \perp B \mid C)\}$


$I(G_3) = \{\,\}$

# Markov Blanket



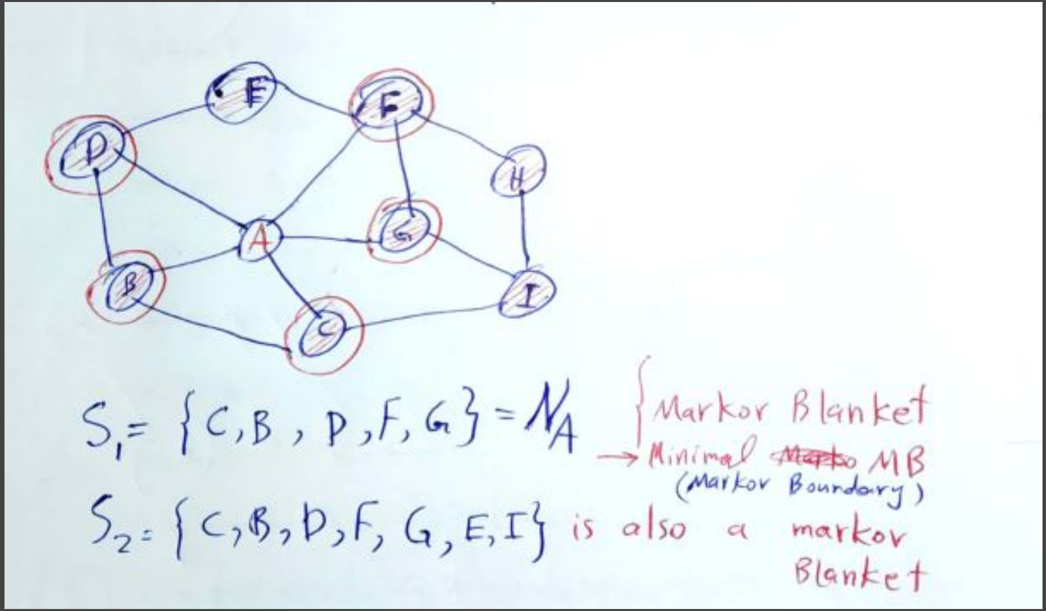
$X = \{X_1, X_2, \cdots, X_n\}$

$S \subseteq X$ is a Markov Blanket for $X_i$.

if $P(X_i \mid X \setminus X_i) = P(X_i \mid S)$

# Markov Blanket - Markov Nets

Minimal Markov Blanket: Neighbours



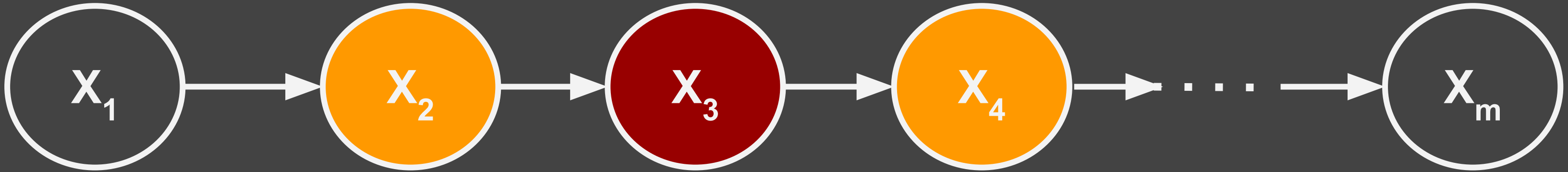
$S_1 = \{C, B, D, F, G\} = N_A$ Markov Blanket

→ Minimal ~~Marko~~ MB (Markov Boundary)

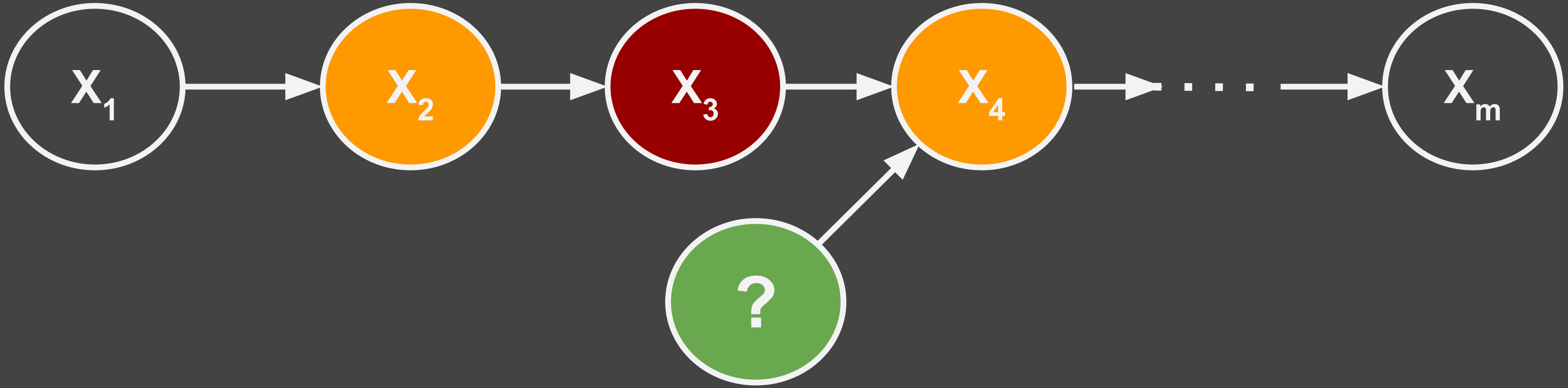
$S_2 = \{C, B, D, F, G, E, I\}$ is also a markov Blanket

# Markov Blanket - Bayesian Nets

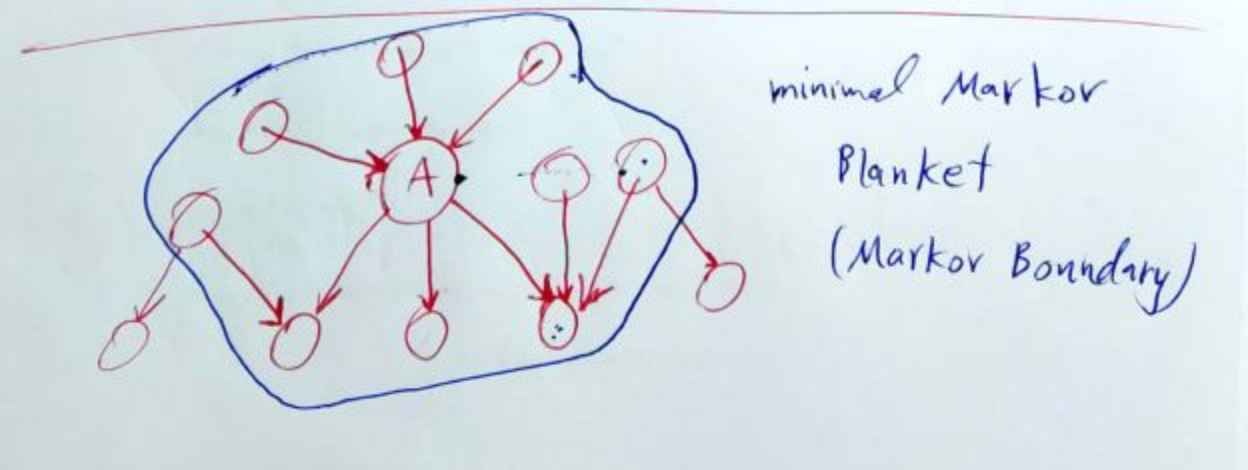
K. N. Toosi
University of Technology



# Markov Blanket - Bayesian Nets



# Markov Blanket - Bayesian Nets



Minimal Markov Blanket: Parents, Children, Parents of children