CS 295: Causal Reasoning

Rina Dechter

More on Structural Causal Models
Definition and distributions

Primer, (chapters 1, 2) PCH 1.2)

Slides by Elias Bareinboim

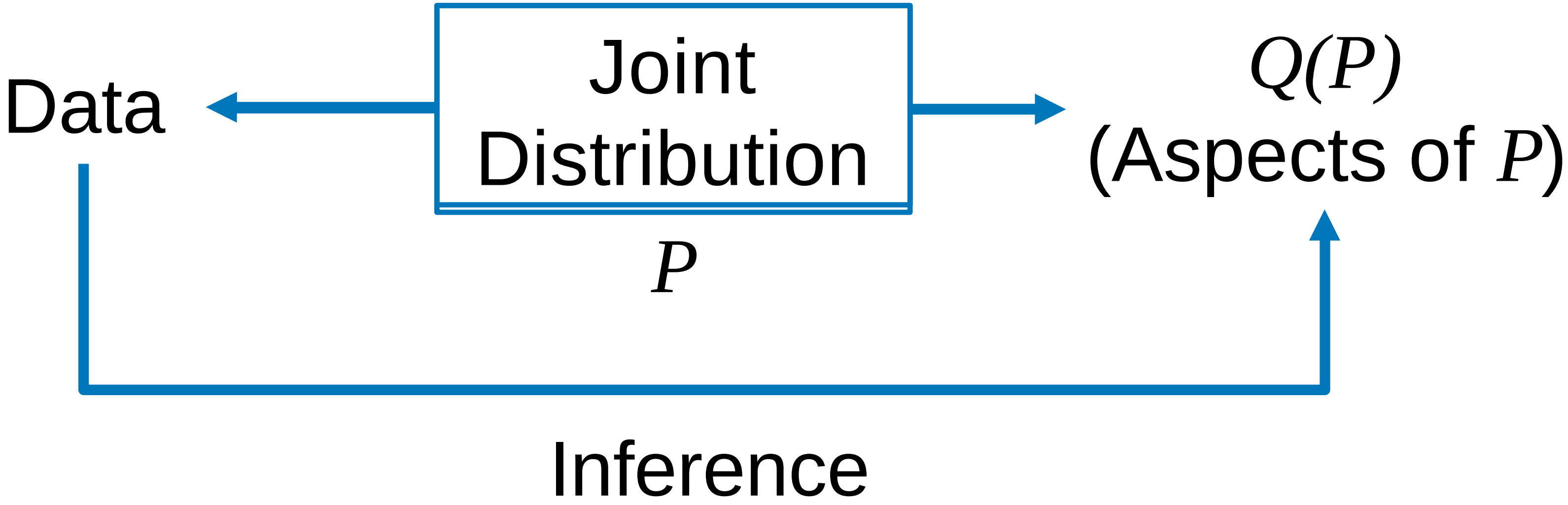
Winter 2022

Outline

- Structural Causal Models
- Product form of Markov SCM
- d-seperation
- Bayesian networks

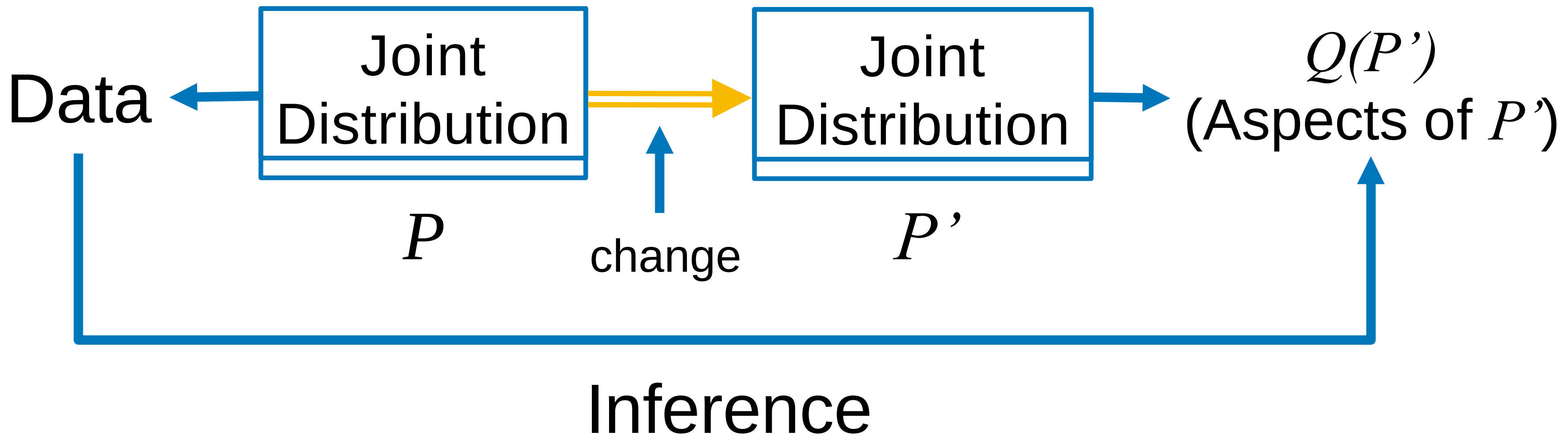
Traditional Stats-ML Inferential Paradigm

- **Approach:** Find a good representation for the data.



e.g., Infer whether customers who bought product A would also buy product B — or, compute $Q = P(B \mid A)$.

From Statistical to Causal Analysis



e.g., Estimate *P'(sales)* if we double the price

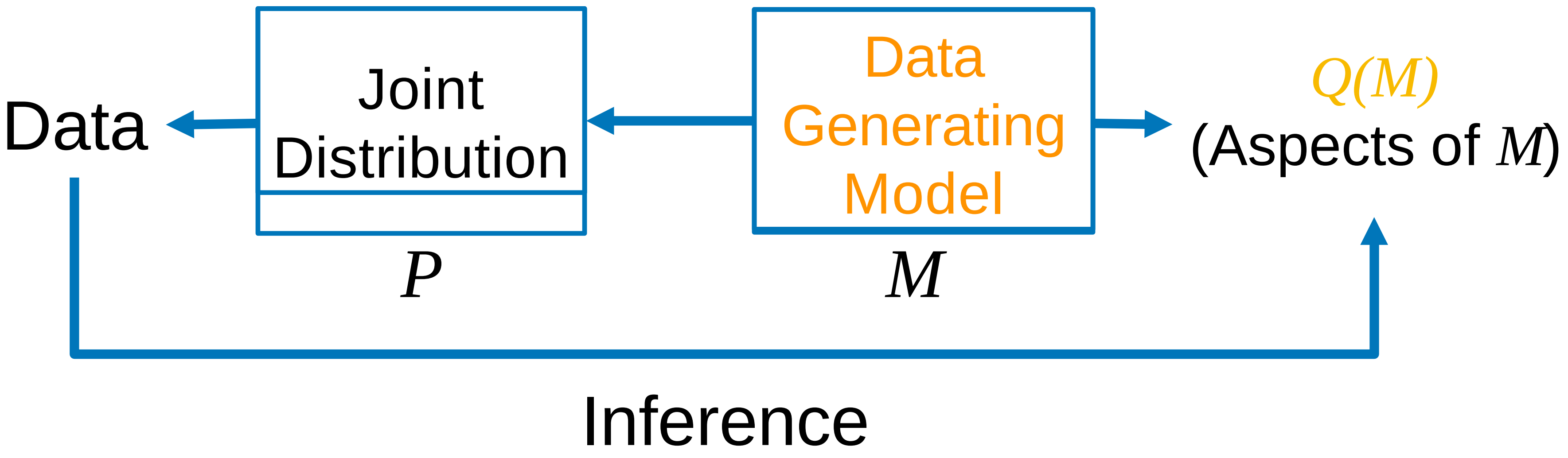
Estimate *P'(cancer)* if we ban smoking

Q: How does *P* (factual) changes to *P'(hypothetical)*?

Needed: New formalism to represent both P & P'.

P is tied to the data; P' is never observed, no data.

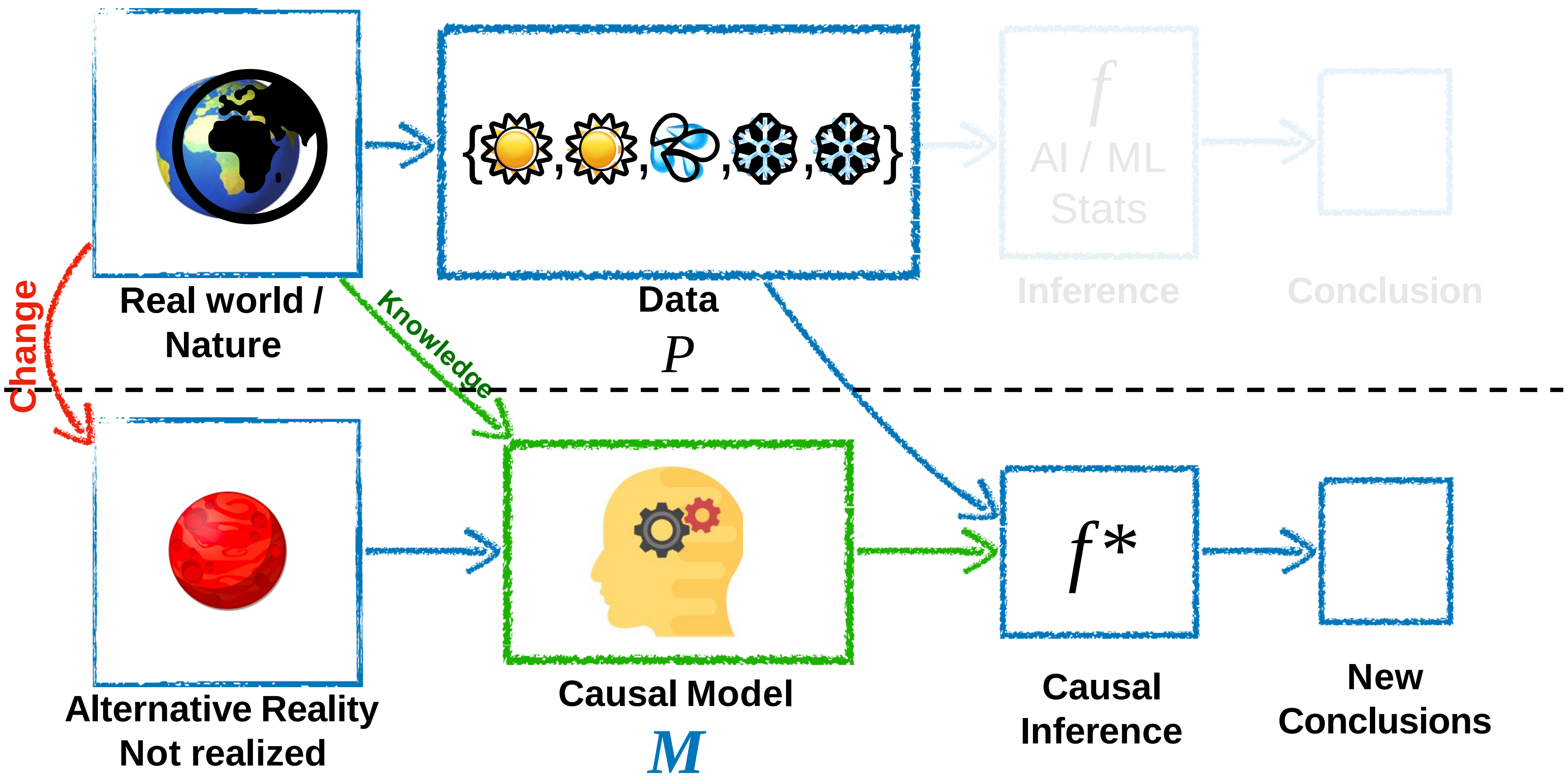
New Oracle -
The Structural Causal Model Paradigm



M – Invariant strategy (mechanism, recipe, law, protocol) by which Nature assigns values to variables in the analysis.

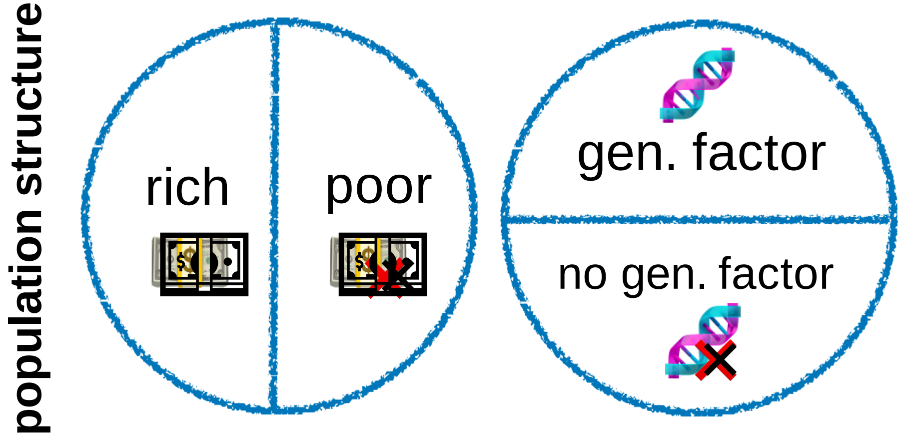
P - model of data, M - model of reality

Back to the Big Picture



Modeling Reality with SCM

- The population of a certain city is falling ill from a contagious disease. There is a drug believed to help patients survive the infection.
- Unknown to the physicians, folks with good living conditions (rich) will always survive.
- While some people have a gene that naturally fights the disease and don't require treatment, they will develop an allergic reaction if treated, which is fatal under poor living conditions.



Being rich and having the genetic factor are independent events.

<u>Reality (unknown to physicians):</u>

rich = alive anyways
$poor_1$ = die anyways (no gene)
$poor_2$ = die iff take the drug (gene)

$\Pi = rich \cup poor_1 \cup poor_2$

$P(rich) = P(poor)$
$P(poor_1) = P(poor_2)$

Modeling Reality in our Example

Variables we observe (**V**):

R ($R=1$ for rich, $=0$ for poor)

D ($D=1$ for taking the drug)

A ($A=1$ if person ends up alive)

Modeling Reality in Our Example

Variables we observe (**V**):

R ($R=1$ for rich, $=0$ for poor)
D ($D=1$ for taking the drug)
A ($A=1$ if person ends up alive)

Variables that are unobserved (**U**):

U_g ($U_g=1$ has genetic factor, $=0$ o/w)
U_r (Other factors affecting Wealth)

Modeling Reality in Our Example

Variables we observe (**V**):

R ($R=1$ for rich, $=0$ for poor)

D ($D=1$ for taking the drug)

A ($A=1$ if person ends up alive)

How are the observed variables determined?

$$R \leftarrow U_r$$
$$D \leftarrow R$$
$$A \leftarrow R \vee (U_g \wedge \neg D)$$

Variables that are unobserved (**U**):

U_g ($U_g=1$ has genetic factor, $=0$ o/w)

U_r (Other factors affecting Wealth)

Modeling Reality in our Example

Variables we observe (**V**):

R ($R=1$ for rich, $=0$ for poor)
D ($D=1$ for taking the drug)
A ($A=1$ if person ends up alive)

How are the observed variables determined?

$$R \leftarrow U_r$$
$$D \leftarrow R$$
$$A \leftarrow R \vee (U_g \wedge \neg D)$$

Variables that are unobserved (**U**):

U_g ($U_g=1$ has genetic factor, $=0$ o/w)
U_r (Other factors affecting Wealth)

- Rich is always alive.
- Poor will survive only if they have the gene and don't take the drug.

Modeling Reality in our Example

Variables we observe (**V**):

R ($R=1$ for rich, $=0$ for poor)

D ($D=1$ for taking the drug)

A ($A=1$ if person ends up alive)

- How are the observed variables determined?

- $R \leftarrow U_r$
  $D \leftarrow R$
- $A \leftarrow R \vee (U_g \wedge \neg D)$

Variables that are unobserved (**U**):

U_g ($U_g=1$ has genetic factor, $=0$ o/w)
U_r (Other factors affecting Wealth)

- What is the randomness over the unobserved vars:

- $P(U_g=1)=1/2$, $P(U_r=1)=1/2$

Modeling Reality in our Example

Variables we observe (**V**):

R ($R=1$ for rich, $R=0$ for poor)

D ($D=1$ for taking the drug

A ($A=1$ if person en

How are the observed variables determined?

Variables that are unobserved (**U**):

U_g ($U_g=1$ has genetic factor, =0 o/w)

U_r (Other factors affecting Wealth)

What is the randomness over the hidden variab

This is a fully specified Model of Reality!

It implies both P and P' (more soon).

This will be our new, almighty Oracle,
which is known as **Structural Causal Model**.

(Now, let's generalize this object…)

Outline

- Structural Causal Models
- Product form of Markov SCM
- d-seperation
- Bayesian networks

The New Oracle: Structural Causal Models

Definition: A structural causal model (SCM) M is a 4-tuple $<\boldsymbol{V}, \boldsymbol{U}, \mathcal{F}, P(\boldsymbol{u})>$, where

- $\boldsymbol{V} = \{V_1,\ldots,V_n\}$ are endogenous variables;
- $\boldsymbol{U} = \{U_1,\ldots,U_m\}$ are exogenous variables;
- $\mathcal{F} = \{f_1,\ldots, f_n\}$ are functions determining $\boldsymbol{V}$,
  $v_i \leftarrow f_i(pa_i, u_i)$, $Pa_i \subset V_i, U_i \subset U$;
- $P(\boldsymbol{u})$ is a distribution over $\boldsymbol{U}$

Not regression!!

e.g. $y = \alpha + \beta X + U_Y$

Axiomatic Characterization:

(Galles-Pearl, 1998; Halpern, 1998).

1. SCM induces distribution $P(\boldsymbol{v})$

- $\mathcal{F}$ can be seen as a mapping from $\boldsymbol{U} \longrightarrow \boldsymbol{V}$

$$(u_1, u_2, \ldots, u_k) \longrightarrow \boxed{\mathcal{F}} \longrightarrow (v_1, v_2, \ldots, v_n)$$

- When the input $\boldsymbol{U}$ is a set of random vars, then the output $\boldsymbol{V}$ also becomes a set of r.v's.
- P(v) is the layer 1 of the PCH, known as the observational (or passive) prob. distribution.
- Each event, person, observation, etc… corresponds to a instantiation of $\boldsymbol{U}=\boldsymbol{u}$.

1. SCM induces distribution $P(\mathbf{v})$

Example: (Drug, Rich, Alive)

- Each citizen follows in one of four groups according to the unobservables in the model:

$$\mathcal{F} = \begin{cases} f_R : U_r \\ f_D : R \\ f_A : R \vee (U_g \wedge \neg D) \end{cases}$$

$$\mathcal{F}$$

$$(U_r{=}1, U_g{=}1) \longrightarrow (R{=}1, D{=}1, A{=}1)$$

$$(U_r{=}1, U_g{=}0) \longrightarrow (R{=}1, D{=}1, A{=}1)$$

$$(U_r{=}0, U_g{=}1) \longrightarrow (R{=}0, D{=}0, A{=}1)$$

$$(U_r{=}0, U_g{=}0) \longrightarrow (R{=}0, D{=}0, A{=}0)$$

1. SCM induces distribution $\boldsymbol{P(v)}$

In our example:

- Events in the $\boldsymbol{U}$-space translate into events in the space of $\boldsymbol{V}$.

$$\mathcal{F} = \begin{cases} f_R : U_r \\ f_D : R \\ f_A : R \vee (U_g \wedge \neg D) \end{cases}$$

$P(\mathbf{u})$			$P(\mathbf{v})$
1/4	$(U_r=1, U_g=1)$	$\longrightarrow$ $(R=1, D=1, A=1)$	1/2 (rows 1–2)
1/4	$(U_r=1, U_g=0)$	$\longrightarrow$ $(R=1, D=1, A=1)$	
1/4	$(U_r=0, U_g=1)$	$\longrightarrow$ $(R=0, D=0, A=1)$	1/4
1/4	$(U_r=0, U_g=0)$	$\longrightarrow$ $(R=0, D=0, A=0)$	1/4

1. SCM induces distribution $P(v)$

In our example:

- Events in the ***U***-space translate into events in the space of ***V***.

$$\mathscr{F} = \begin{cases} f_R : U_r \\ f_D : R \\ f_A : R \vee (U_g \wedge \neg D) \end{cases}$$

$P(\mathbf{u})$

1/4 — $(U_r=1, U_g=1) \longrightarrow$

1/4 — $(U_r=1, U_g=0) \longrightarrow$

1/4 — $(U_r=0, U_g=1) \longrightarrow$

1/4 — $(U_r=0, U_g=0) \longrightarrow$

R	D	A	P(r,d,a)
0	0	0	1/4
0	0	1	1/4
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1/2

$P(\mathbf{v})$

1/2

1/4

1/4

1. SCM induces distribution $P(\boldsymbol{v})$

- [Def. 2, PCH chapter] An SCM $M = <\boldsymbol{V}, \boldsymbol{U}, \mathscr{F}, P(\boldsymbol{u})>$ defines a joint probability distribution $P^M(V)$ s.t. *for each* $Y \subseteq V$:

$$P^M(y) = \sum_{u \mid Y(u)=y} P(u)$$

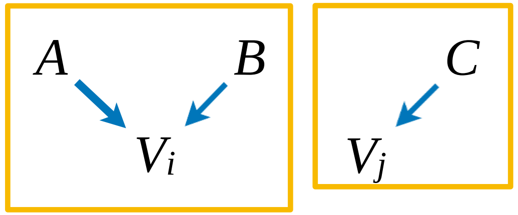
- $\mathscr{F}$ can be seen as a mapping from $\boldsymbol{U} \longrightarrow \boldsymbol{V}$

$$(u_1, u_2, \ldots, u_k) \longrightarrow \boxed{\mathcal{F}} \longrightarrow (v_1, v_2, \ldots, v_n)$$

2. SCM $\longrightarrow$ Causal Diagram

- Every SCM M induces a causal diagram
- Represented as a DAG where:
  - Each $V_i \in \boldsymbol{V}$ is a node,
  - There is $W \longrightarrow V_i$ if for $W \in Pa_i$,

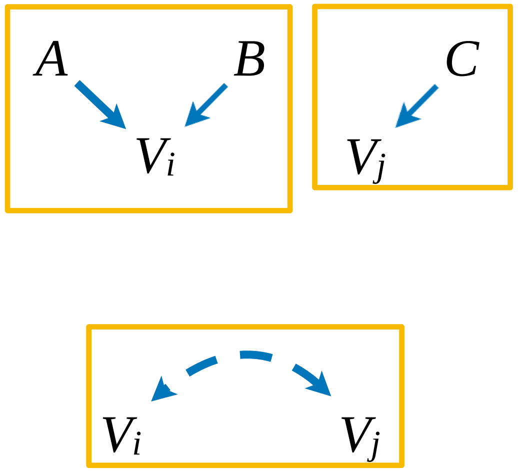
$V_i \leftarrow f_i(A,B,U)$
$V_j \leftarrow f_j(C,U)$



2. SCM → Causal Diagram

- Every SCM M induces a causal diagram
- Represented as a DAG where:
  - Each $V_i \in \boldsymbol{V}$ is a node,
  - There is $W \longrightarrow V_i$ if for $W \in Pa_i$,
  - There is $V_i \leftarrow\cdots\rightarrow V_j$ whenever $U_i \cap U_j \neq \varnothing$.

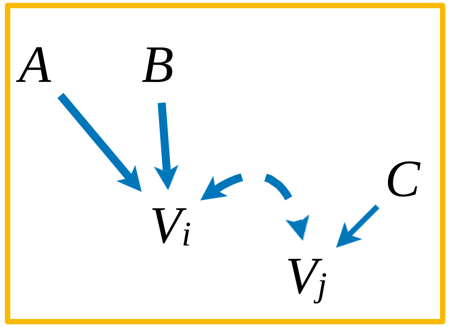
$V_i \leftarrow f_i\,(A,B,U)$
$V_j \leftarrow f_j(C,U)$



2. SCM → Causal Diagram

- Every SCM M induces a causal diagram
- Represented as a DAG where:
  - Each $V_i \in \boldsymbol{V}$ is a node,
  - There is $W \longrightarrow V_i$ if for $W \in Pa_i$,
  - There is $V_i \leftarrow\cdots\cdots\rightarrow V_j$ whenever $U_i \cap U_j \neq \emptyset$.

$V_i \leftarrow f_i\,(A,B,U)$
$V_j \leftarrow f_j(C,U)$



G

Causal Diagram — Definition

- ***Causal Diagram*** [Def. 13, PCH chapter] — Consider an SCM M = $<\boldsymbol{V}, \boldsymbol{U}, \mathcal{F}, P(\boldsymbol{u})>$. Then G is said to be a causal diagram (of M) if constructed as follows:

1. add vertex for every endogenous variable $\boldsymbol{V_i} \in \boldsymbol{V}$.
2. add edge ($\boldsymbol{V_j} \rightarrow \boldsymbol{V_i}$) for every $\boldsymbol{V_i}$, $\boldsymbol{V_j}$ $\subset$V if $\boldsymbol{V_j}$ appears as argument of $\boldsymbol{f_i} \in \mathcal{F}$.
3. add a bidirected edge ($\boldsymbol{V_j} \leftarrow\!\!\!-\!\!-\!\!\!\rightarrow \boldsymbol{V_i}$) for every $\boldsymbol{V_i}$, $\boldsymbol{V_j}$ $\subset$V if $\boldsymbol{U_i}$, $\boldsymbol{U_j} \subset \boldsymbol{U}$ are correlated or the corresponding functions $\boldsymbol{f_i}$, $\boldsymbol{f_j}$ share some $U \in \mathbf{U}$ as argument.

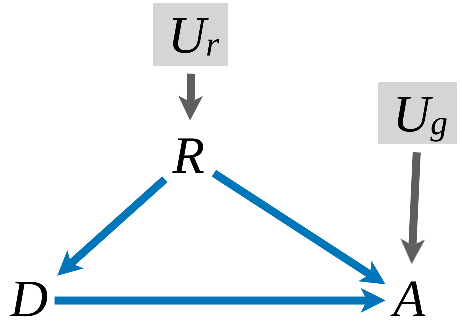
2. SCM → Causal Diagram

Recall our medical example:

- Endogenous (observed) variables $\boldsymbol{V}$:
  - R ($R=1$ for rich, $=0$ for poor)
  - D ($D=1$ for taking the drug, $D=0$ o/w)
  - A ($A=1$ if person ends up alive, $=0$ o/w)

$$\mathcal{F} = \begin{cases} f_R \leftarrow U_r \\ f_D \leftarrow R \\ f_A \leftarrow R \vee (U_g \wedge \neg D) \end{cases}$$

- Exogenous (unobserved) Variables $\boldsymbol{U}$:
  - U_r (Wealthiness factors)
  - U_g ($=1$ has the genetic factor, $=0$ o/w)
- Distribution over $\boldsymbol{U}$: $P(U_r)=1/2$, $P(U_g)=1/2$



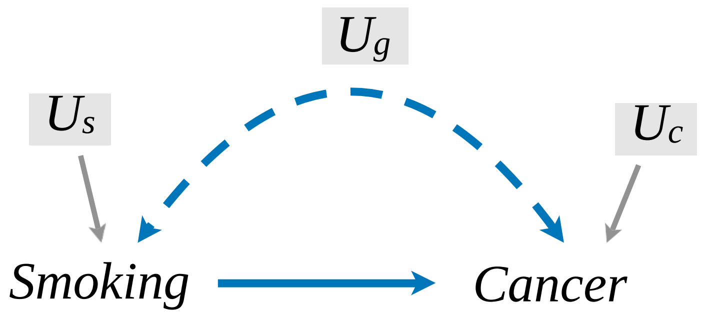
2. SCM → Causal Diagram

Another example:

- $\boldsymbol{V} = \{ Smoking, Cancer \}$
- $\boldsymbol{U} = \{ U_s, U_c, U_g \}$
- $\mathcal{F}$:

unobserved genotype

$Smoking \leftarrow \mathrm{f}_{\mathrm{Smoking}}(U_s, U_g)$
$Cancer \leftarrow \mathrm{f}_{\mathrm{Cancer}}(Smoking, U_c, U_g)$

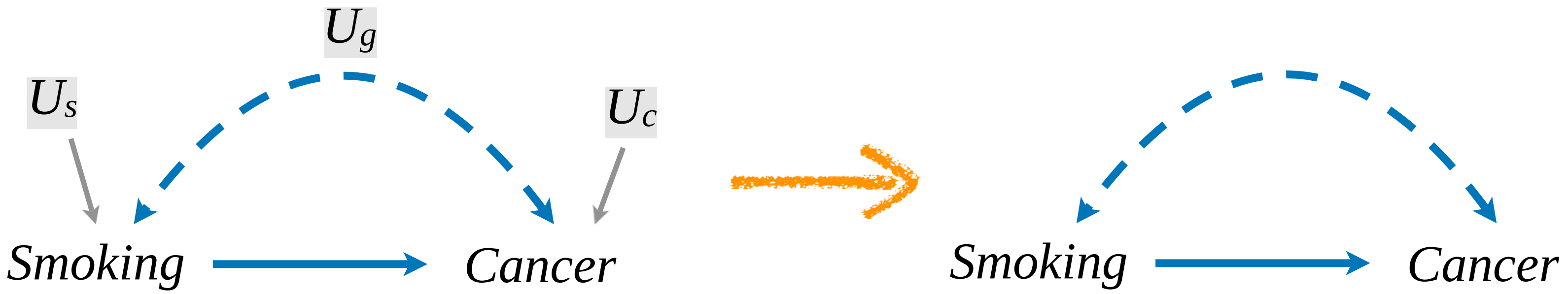


Remark 1. The mapping is just 1-way (i.e., from a SCM to a causal graph) since the graph itself is compatible with infinitely many SCMs with the same scope (the same functions signatures and compatible exogenous distributions).

Remark 2. This observation will be central to causal inference since, in most practical settings, researchers may know the scope of the functions, for example, but not the details about the underlying mechanisms.

Causal Diagrams

- Convention. The unobserved variables are left implicit in the graph.



Food for thought

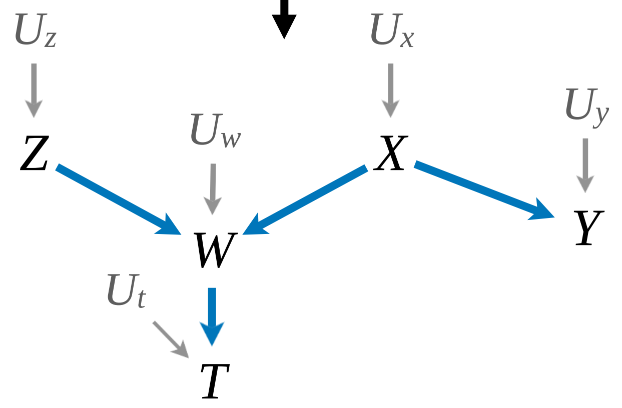
Does the causal diagram give us any clues about the (in)dependence relations in the obs. distribution P(V)?

- Is T independent of W?
- Is W independent of T?
- Is Z independent of T?
- Is Z independent of X?
- Is Y independent of W?
- Is Y independent of W if we know the value of X?

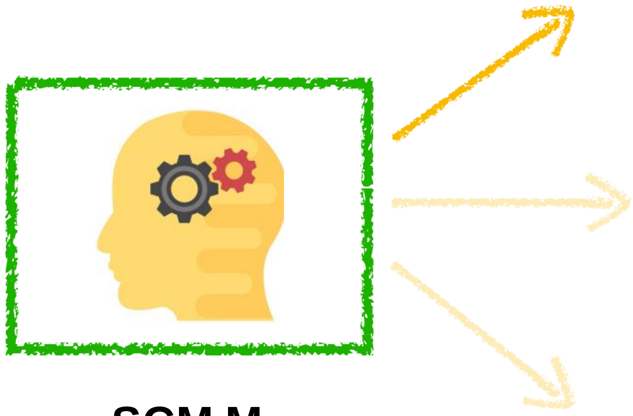
Prop. 1

$$M = \begin{cases} Z \leftarrow f_Z(u_z) \\ X \leftarrow f_x(u_x) \\ W \leftarrow f_W(z, x, u_w) \\ Y \leftarrow f_Y(x, u_y) \\ T \leftarrow f_T(w, u_t) \end{cases}$$

Prop. 2



3. SCM → Pearl's Causal Hierarchy



SCM M

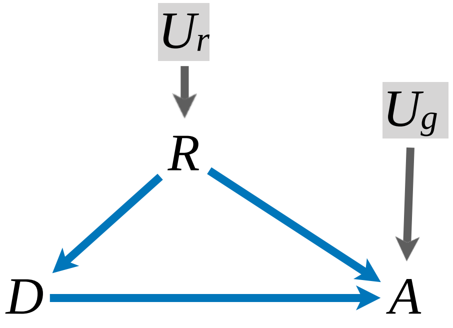
	Level (Symbol)	Typical Activity
1	Association $P(y \mid x)$	Seeing ML - (Un)Supervised Learning
2	Intervention $P(y \mid do(x), c)$	Doing ML - Reinforcement Learning
3	Counterfactual $P(y_x \mid x', y')$	Imagining, Retrospection

Outline

- Structural Causal Models
- Product form of Markov SCM
- d-seperation
- Bayesian networks

The Emergence of the First Layer

In our example,



The joint distribution over the observables ***P(v)*** is equal to:

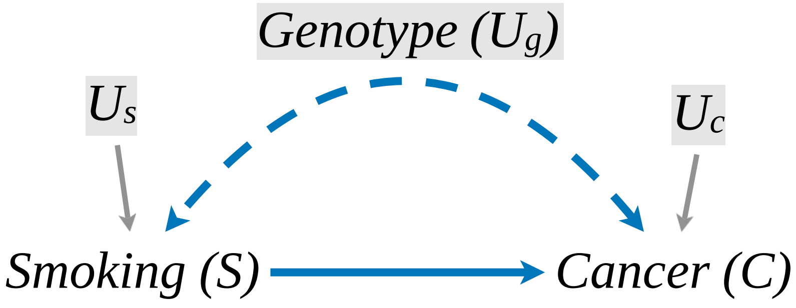
$$P(R=r, D=d, A=a) = \sum_{u_r, u_g} P(R=r, D=d, A=a, U_r=u_r, U_g=g)$$

For short,

$$P(r,d,a) = \sum_{u_r, u_g} P(r,d,a,u_r,u_g)$$

The Emergence of the First Layer

In the second example,



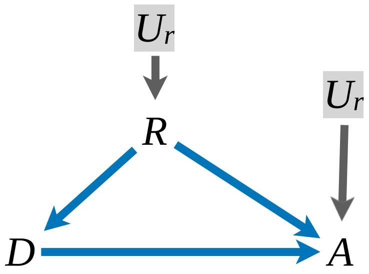
The joint probability distribution over the observed variables (V), *Smoking* and *Cancer*, is given by

$$P(s, c) = \sum_{u_s, u_g, u_c} P\left(s, c, u_s, u_g, u_c\right)$$

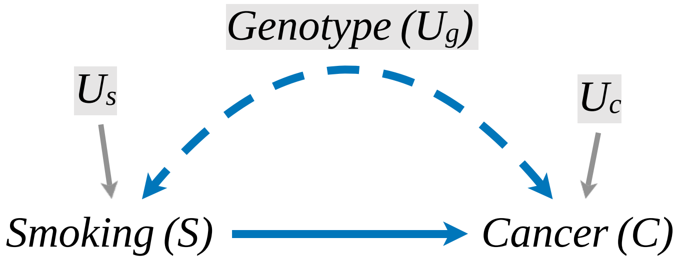
Recall, this distribution is called observational distribution. Sometimes, it's also called passive or non-experimental distribution.

What the Diagram Encodes

- Since G is a directed acyclic graph, there exists a topological order over $\boldsymbol{V}$ such that every variable goes after its parents, i.e., $Pa_i < V_i$.



$$R < D < A$$



$$S < C$$

What the Diagram Encodes

- M induces $P(\mathbf{V})$:

$$P(\mathbf{v}) = \sum_{\mathbf{u}} P(\mathbf{v}, \mathbf{u}),$$

  - Using the chain rule and following a topological order,

$$P(\mathbf{v}) = \sum_{\mathbf{u}} P(\mathbf{u}) \prod_{V_i \in \mathbf{V}} P\left(v_i \mid v_1, \ldots, v_{i-1}, \mathbf{u}\right),$$

- An observed variable is fully determined by its observed and unobserved parents; also $\{pa_i, u_i\} \subseteq \{v_1, \ldots, v_{i-1}, \mathbf{u}\}$, then

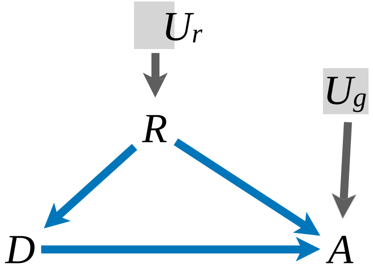
$$P(v_i \mid v_1, \ldots, v_{i-1}, \mathbf{u}) = P(v_i \mid pa_i, u_i)$$

What the Diagram Encodes

- The distribution $P(\mathbf{V})$ decomposes as:

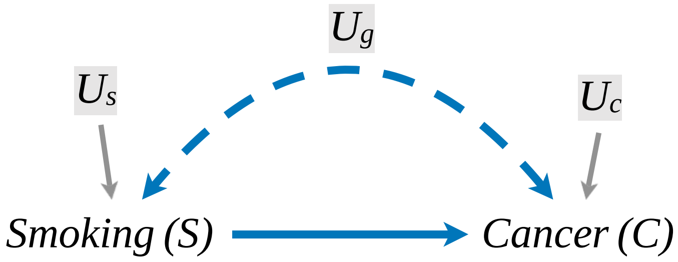
$$P(\mathbf{v}) = \sum_{\mathbf{u}} P(\mathbf{u}) \prod_{V_i \in \mathbf{V}} P\left(v_i \mid v_1, \ldots, v_{i-1}, \mathbf{u}\right)$$

$$= \sum_{\mathbf{u}} P(\mathbf{u}) \prod_{V_i \in \mathbf{V}} P\left(v_i \mid pa_i, u_i\right)$$



U r

$$P(r, d, a) = \sum_{u_r, u_g} P(u_r, u_g) P(r \mid u_r) P(d \mid r) P(a \mid r, d, u_g)$$



$$P(s, c) = \sum_{u_s, u_g, u_c} P(u_s, u_g, u_c) P(s \mid u_g, u_s) P(c \mid s, u_g, u_c)$$

Conditional Independences

- If knowing that variable X = x doesn't change the belief in Y = y, then X and Y are said to be probabilistically independent. This is written as X ⫫ Y.
  - $X \perp\!\!\!\perp Y \equiv P(Y = y \mid X = x) = P(Y = y)$

$P(Y = y, X = x) = P(Y = y)\, P(X = x)$

- More generally, once we know the value of a third variable Z = z, if knowing that X = x doesn't affect the belief of Y = y, X and Y are conditionally independent given Z, i.e., X ⫫ Y | Z.
  - $X \perp\!\!\!\perp Y \mid Z \equiv P(X = x \mid Z = z, Y = y) = P(X = x \mid Z = z).$

Lack of functional dependence → probabilistic independence.

Markovian Factorization

- Suppose no variable in $\boldsymbol{U}$ is a parent of two variables in $\boldsymbol{V}$ *(observables)* (i.e., $\forall_{i,j}\ U_i \cap U_j = \emptyset$), then the model is called Markovian. We have:

$$P(\mathbf{v}) = \sum_{\mathbf{u}} P(\mathbf{u}) \prod_{V_i \in \mathbf{V}} P\left(v_i \mid pa_i, u_i\right)$$

$$= \sum_{\mathbf{u}} \prod_{V_i \in \mathbf{V}} P\left(v_i \mid pa_i, u_i\right) P\left(u_i\right)$$

$$= \sum_{\mathbf{u}} \prod_{V_i \in \mathbf{V}} P\left(v_i \mid pa_i, u_i\right) P\left(u_i \mid pa_i\right)$$

$$= \prod_{V_i \in \mathbf{V}} \sum_{u_i} P\left(v_i, u_i \mid pa_i\right) = \boxed{\prod_{V_i \in \mathbf{V}} P\left(v_i \mid pa_i\right)}$$

Bayesian Factorization

In Markovian models
SCM yields a Bayesian network
Over the visible variables

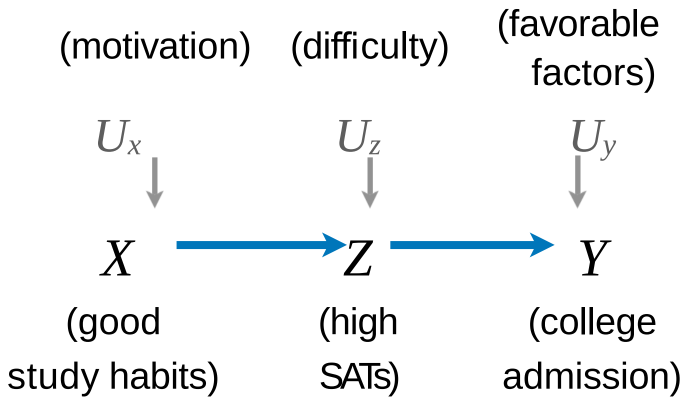
Local Markovian Condition

$$\boxed{(V_i \perp\!\!\!\perp Nd_i \setminus Pa_i \mid Pa_i)}$$

Outline

- Structural Causal Models
- Product form of Markov SCM
- d-separation: reading independencies from the DAG
- Bayesian networks

Causal Chains



- Are X and Y independent given Z?
- Yes,
- e.g., knowing $Z=1$ (high SAT scores), the probability of being admitted ($Y=1$) does not change if we know the student has good study habits ($X=1$) or not($X=0$).

$$P(x,y|z) = \frac{P(x,z,y)}{P(z)} = \frac{P(x)\,P(z|x)\,P(y|z)}{P(z)}$$

$$= \frac{P(x,z)}{P(z)}P(y|z)$$

$$= \boxed{P(x|z)\,P(y|z)}$$

Bayes Factorization

M: $X \leftarrow U_x$

$Z \leftarrow X \vee \neg U_z$

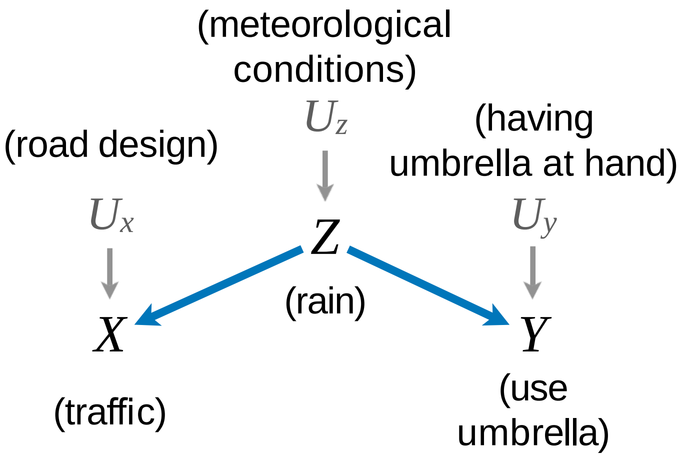
$Y \leftarrow Z \wedge U_y$

$P(U_x, U_z, U_y)$

Graphically, observing Z

"blocks" the influence from X to Y.

Common Cause



- Are X and Y independent?
- No,
- e.g., seeing someone coming in with an umbrella in hand ($Y=1$) rises the probability of rain ($Z=1$), which increases the likelihood of bad traffic ($X=1$).

$$\exists_{x,y} P(X = x, Y = y) \neq P(X = x)P(Y = y)$$

try it out!

Graphically, information “flows" from Y going through the common cause Z and down to X.

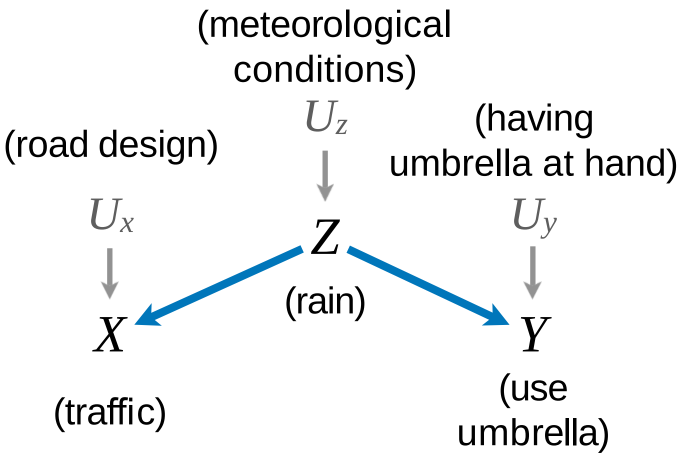
M: $Z \leftarrow U_z$

$X \leftarrow Z \oplus \neg U_x$

$Y \leftarrow Z \vee U_y$

$P(U_x, U_z, U_y)$

Common Cause



Are X and Y independent given Z?

Yes,

e.g., if we know it is raining ($Z=1$), observing people with umbrellas ($Y=1$) tell us nothing about the traffic (X).

Bayes Factorization

$$P(x, y \mid z) = \frac{P(x,z,y)}{P(z)} = \frac{P(z)\,P(x \mid z)\,P(y \mid z)}{P(z)} = \boxed{P(x \mid z)\,P(y \mid z)}$$

Graphically, observing Z "blocks" the influence from X to Y.

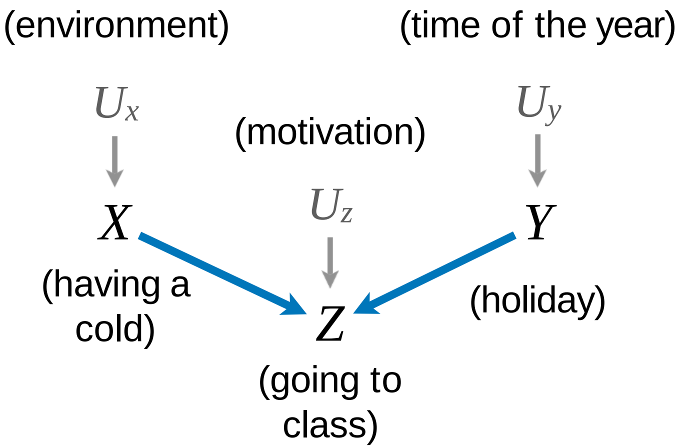
M: $Z \leftarrow U_z$

$X \leftarrow Z \oplus \neg U_x$

$Y \leftarrow Z \vee U_y$

$P(U_x, U_z, U_y)$

Common Effect



- Are X and Y independent?

Yes!,

- e.g., having a cold $X=1$ is independent of being on holiday $Y=1$.

$$P(x,y) = \sum_z P(x)P(y)P(z \mid x,y)$$
$$= P(x)P(y)\sum_z P(z \mid x,y)$$
$$= P(x)P(y)$$

Graphically, influence from X reaches Z but does not "go up" to Y.

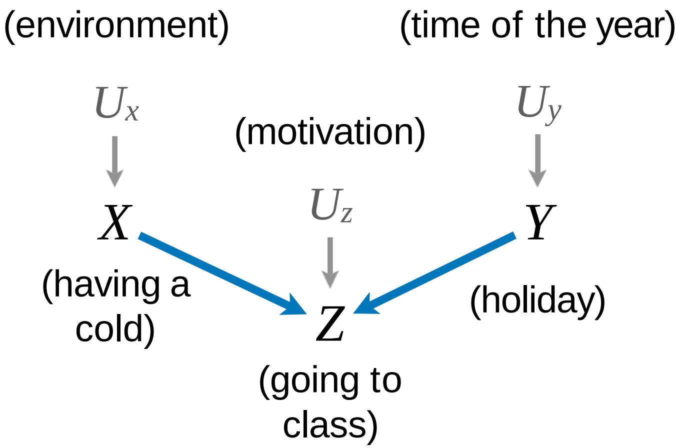
M: $X \leftarrow U_x$

$Y \leftarrow U_y$

$Z \leftarrow \neg Y \wedge (\neg X \oplus U_z)$

$P(U_x, U_z, U_y)$

Common Effect



Are X and Y independent given Z?

No!

e.g., if we observe that a student didn't go to class ($Z=0$) and today is not a holiday ($Y=0$), it is more likely that she may have a cold ($X=1$).

$$\exists_{x,y,z} P(X = x, Y = y \mid Z = z) \neq P(X = x \mid Z = z) P(Y = y \mid Z = z)$$

try it out!

M: $X \leftarrow U_x$

$Y \leftarrow U_y$

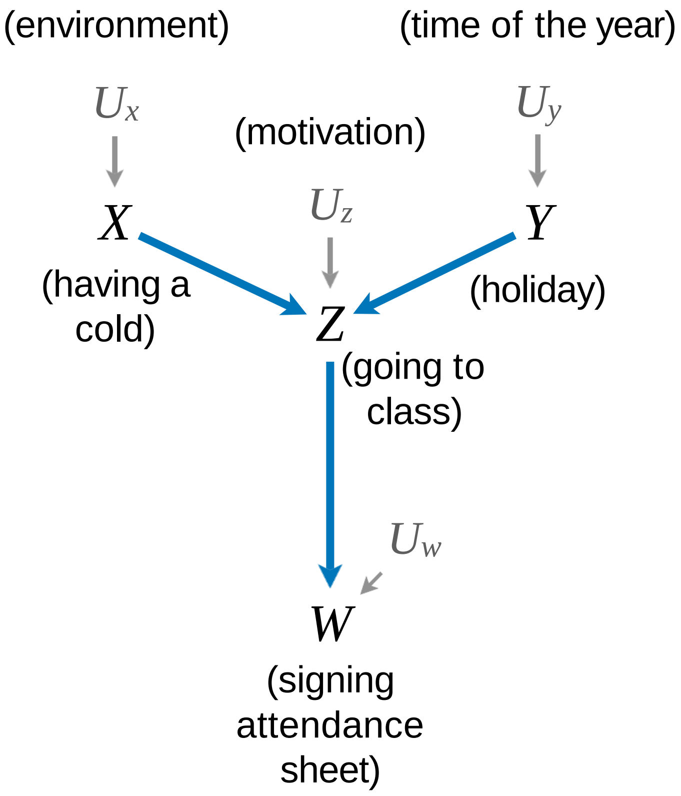
$Z \leftarrow \neg Y \wedge (\neg X \oplus U_z)$

$P(U_x, U_z, U_y)$

Graphically, influence from X reaching Z (when Z is observed) bumps "back up" to Y.

This behavior is opposite to the previous cases.

Common Effect

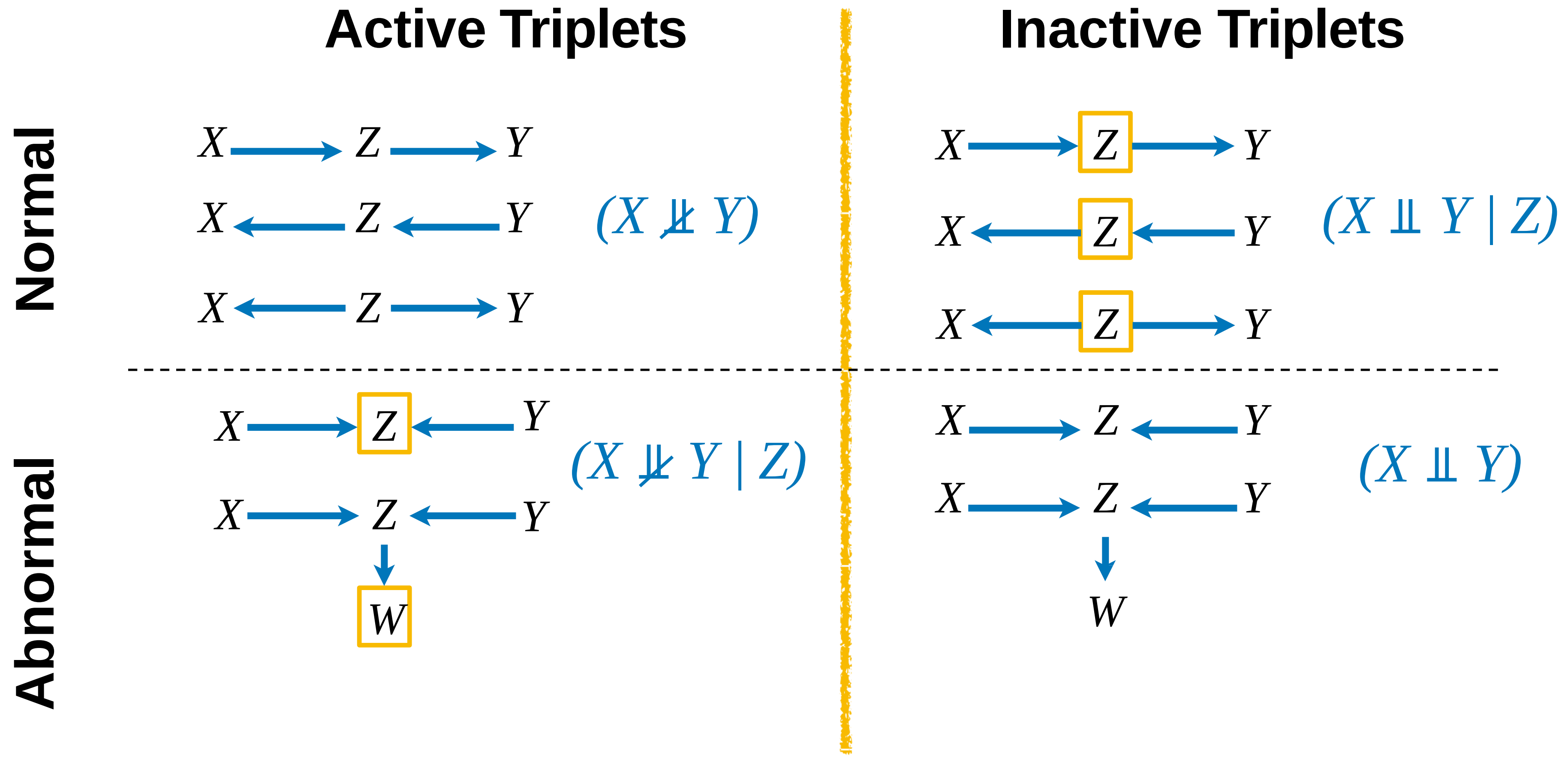


- Are X and Y independent given W? No!, again!,
- e.g., observing that the student didn't sign the assistance sheet ($W=0$) increases the likelihood of the student being absent ($Z=0$), that as we said, make X and Y dependent.

Graphically, influence from X reaching W (when W is observed) "bumps back up" to Z, and then Y.

Watch out for the descendants of the colliders!

Summary



What about larger graphical structures?

Graph Separation (d-Separation)

- Consider the question of whether X and Y are independent given Z.

1. Look at every path from X to Y in the graph.

2. A path is active if every triplet in it is active (given Z).

3. If any path is active, X and Y are not independent.

Graph Separation (d-Separation)

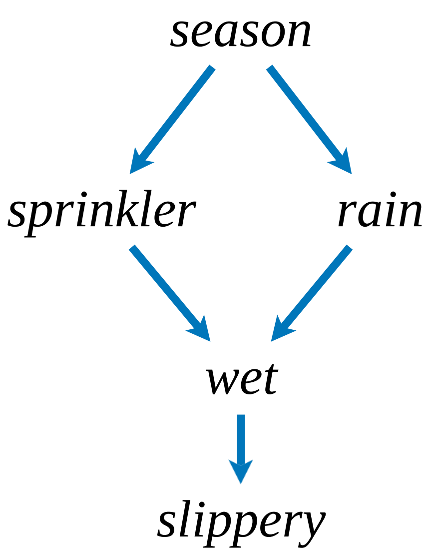
CI$_1$: $(Wet \perp\!\!\!\perp Sprinkler)$

CI$_2$: $(Wet \perp\!\!\!\perp Season \mid Sprinkler)$

CI$_3$: $(Rain \perp\!\!\!\perp Slippery \mid Wet)$

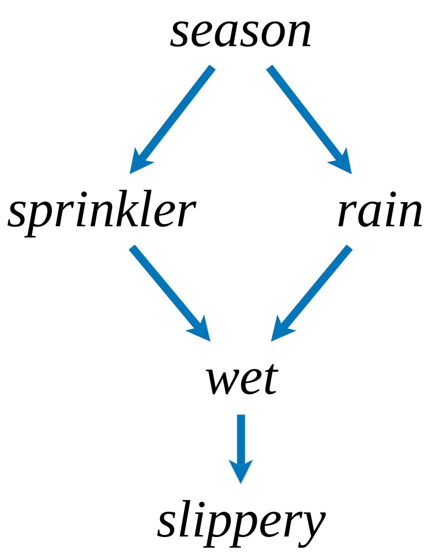
CI$_4$: $(Season \perp\!\!\!\perp Wet \mid Sprinkler, Rain)$

CI$_5$: $(Sprinkler \perp\!\!\!\perp Rain \mid Season, Wet)$



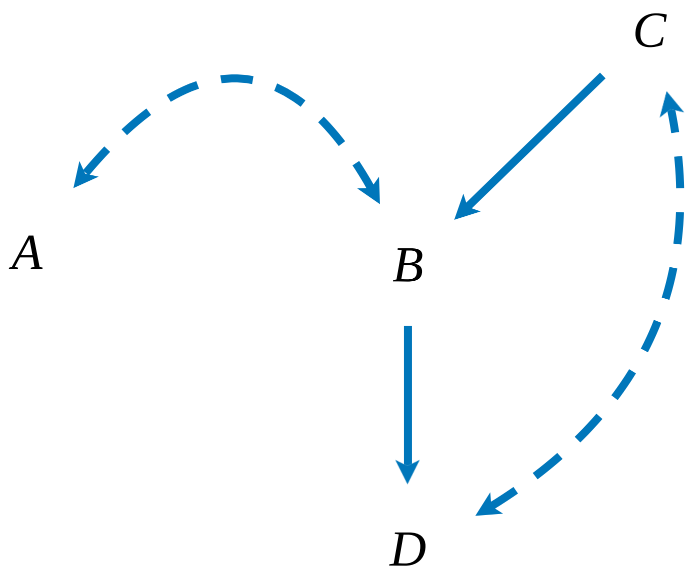
Graph Separation (d-Separation)

- ✓ CI_1: $(Wet \perp\!\!\!\perp Sprinkler)$
- ✓ CI_2: $(Wet \perp\!\!\!\perp Season \mid Sprinkler)$
- ✔ CI_3: $(Rain \perp\!\!\!\perp Slippery \mid Wet)$
- ✔ CI_4: $(Season \perp\!\!\!\perp Wet \mid Sprinkler, Rain)$
- ✓ CI_5: $(Sprinkler \perp\!\!\!\perp Rain \mid Season, Wet)$



d-Separation (food for thought)

- Is A independent of D?
- Is A independent of C?
- Is A independent of C given D?
- Is D independent of C given B?



We want to be able to answer all these questions just from the DAG

Outline

- Structural Causal Models
- Product form of Markov SCM
- d-seperation
- Bayesian networks
- More of d-seperation

DECOMPOSITION BY BAYESIAN NETWORKS

Given a distribution P, on n discrete variables, $X_1, X_2, \ldots, X_n$. Decompose P by the chain rule:

$$P(x_1, \ldots, x_n) = \prod_j P(x_j | x_1, \ldots, x_{j-1}). \qquad \textbf{(1.30)}$$

Suppose X_j is independent of all other predecessors, once we know the value of a select group of predecessors called PA_j. Simplification:

$$P(x_j | x_1, \ldots, x_{j-1}) = P(x_j | pa_j) \qquad \textbf{(1.31)}$$

PA_j : *Markovian parents* of X_j, relative to a given ordering.

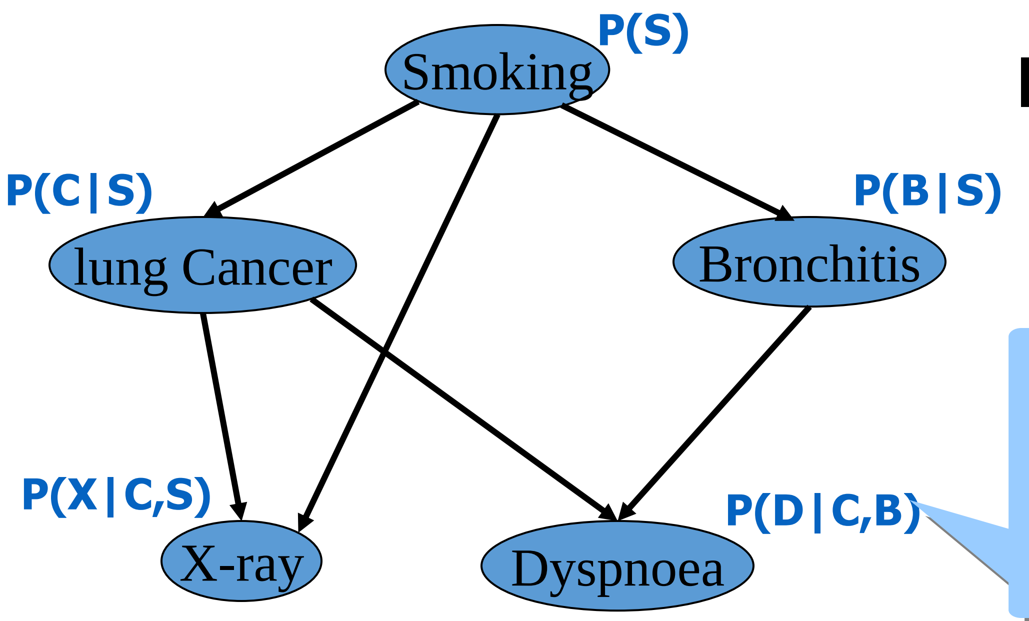
Causality, Pearl ch1

Formal Definition

A **Bayesian network** is:

- An directed acyclic graph (DAG), where
- Each node represents a random variable
- And is associated with the conditional probability of the node given its parents.

Bayesian Networks: Representation



BN = (G, Θ)

CPD:

C	B	D=0	D=1
0	0	0.1	0.9
0	1	0.7	0.3
1	0	0.8	0.2
1	1	0.9	0.1

$$P(S, C, B, X, D) = P(S)\, P(C|S)\, P(B|S)\, P(X|C,S)\, P(D|C,B)$$

Conditional Independencies ➡ Efficient Representation

Definition 1.2.1 (Markovian Parents)
Let $V = \{X_1, \ldots, X_n\}$ be an ordered set of variables, and let $P(v)$ be the joint probability distribution on these variables. A set of variables PA_j is said to be **Markovian parents** *of X_j if PA_j is a minimal set of predecessors of X_j that renders X_j independent of all its other predecessors. In other words, PA_j is any subset of $\{X_1, \ldots, X_{j-1}\}$ satisfying*

$$P(x_j|pa_j) = P(x_j|x_1, \ldots, x_{j-1}) \qquad \textbf{(1.32)}$$

and such that no proper subset of PA_j satisfies **(1.32).**

Interpretation:
Knowing the values of other preceding variables is redundant once we know the values pa_j of the parent set PA_j.

CONSTRUCTING A BAYESIAN NETWORK

Given: P, and an ordering of the variables.

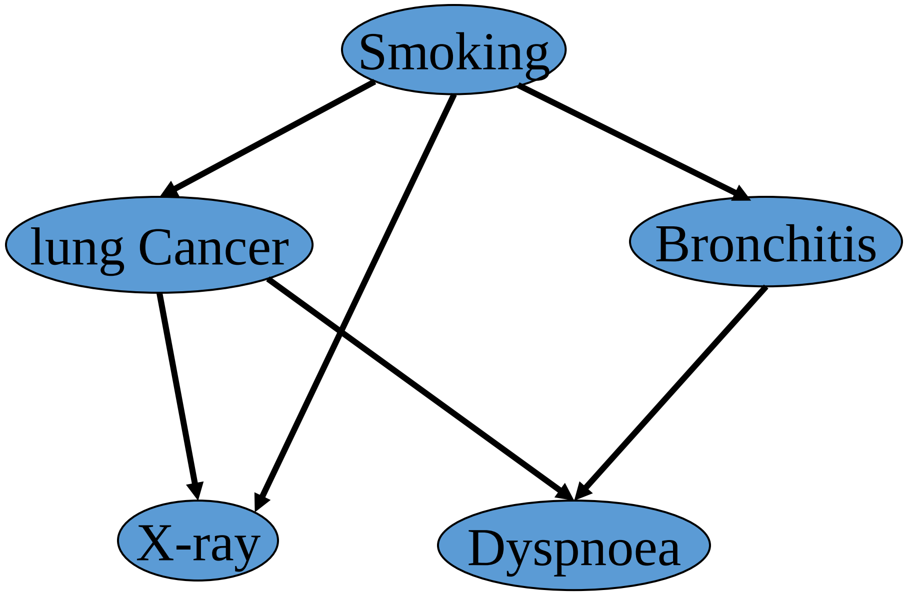
At the jth stage, select any minimal set of X_j's predecessors that screens off X_j from its other predecessors.

Call this set PA_j, and draw an arrow from each member in PA_j to X_j.

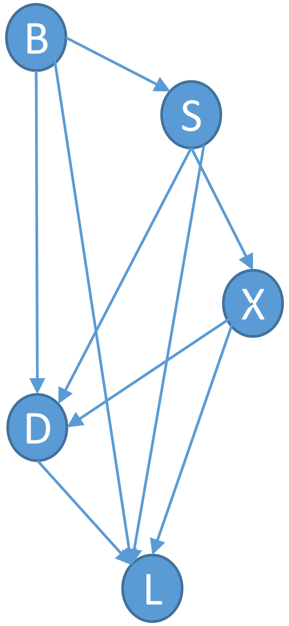
The result is a directed acyclic graph, called a **Bayesian network**, in which an arrow from X_i to X_j assigns X_i as a Markovian parent of X_j, consistent with Definition 1.2.1

The resulting network is unique given the ordering of the variables, whenever the distribution $P(v)$ is strictly positive.

Bayesian Networks: Representation



P(S, C, B, X, D)



Is X independent of B given S?

MARKOV COMPATIBILITY

Definition 1.2.2 (Markov Compatibility)
If a probability function P *admits the factorization of* (1.33) *relative to DAG* G, *we say that* G **represents** P, *that* G *and* P *are* **compatible**, *or that* P *is* **Markov relative** *to* G.

Compatibility implies that G can "explain" the generation of the data represented by P.

THE d-SEPARATION CRITERION

Definition 1.2.3 (d-Separation)
A path p is said to be **d-separated** (*or* **blocked**) *by a set of nodes Z if and only if*

1. *p contains a chain $i \rightarrow m \rightarrow j$ or a fork $i \leftarrow m \rightarrow j$ such that the middle node m is in Z, or*

2. *p contains an inverted fork (or* **collider***) $i \rightarrow m \leftarrow j$ such that the middle node m is not in Z and such that no descendant of m is in Z.*

A set Z is said to d-separate X from Y if and only if Z blocks every path from a node in X to a node in Y.

Theorem 1.2.4
(Probabilistic Implications of d-Separation)
If sets X and Y are d-separated by Z in a DAG G, then X is independent of Y conditional on Z in every distribution compatible with G. Conversely, if X and Y are **not** *d-separated by Z in a DAG G, then X and Y are dependent conditional on Z in at least one distribution compatible with G.*

Theorem 1.2.5
For any three disjoint subsets of nodes (X, Y, Z) in a DAG G and for all probability functions P, we have:

(i) $(X \perp\!\!\!\perp Y | Z)_G \Longrightarrow (X \perp\!\!\!\perp Y | Z)_P$ *whenever G and P are compatible, and*

(ii) *if $(X \perp\!\!\!\perp Y | Z)_P$ holds in all distributions compatible with G, it follows that $(X \perp\!\!\!\perp Y | Z)_G$.*

G is an **Independency map (IMAP)** of any compatible P relative to d-separation.

Theorem 1.27
(Parental Markov Condition)
A necessary and sufficient condition for a probability distribution P to be Markov relative a DAG G is that every variable be independent of all its nondescendants (in G), conditional on its parents.

Theorem 1.28
(Observational Equivalence)
Two DAGs are observationally equivalent if and only if they have the same skeletons and the same sets of v-structures, that is, two converging arrows whose tails are not connected by an arrow (Verma and Pearl 1990).

Will discuss later

Outline

- Structural Causal Models
- Product form of Markov SCM
- d-seperation
- Bayesian networks
- More of d-seperation

Triplets - Summary

Active Triplets

$X \rightarrow Z \rightarrow Y$

$X \leftarrow Z \rightarrow Y$

Triplets - Summary

Active Triplets

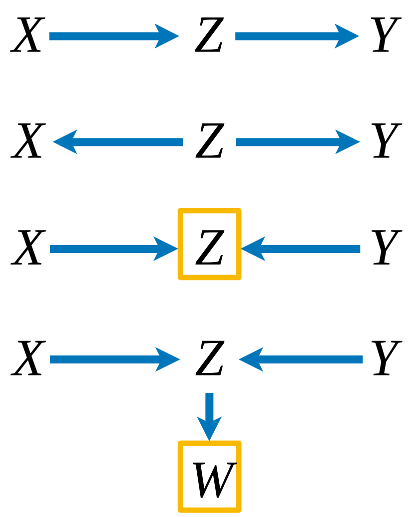
$$X \longrightarrow Z \longrightarrow Y$$

$$X \longleftarrow Z \longrightarrow Y$$

$$X \longrightarrow \boxed{Z} \longleftarrow Y$$

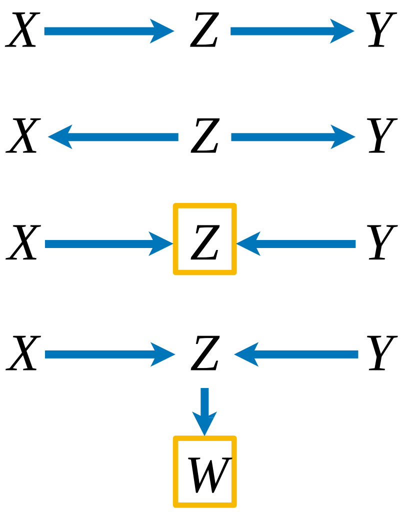
Triplets - Summary

Active Triplets



Triplets - Summary

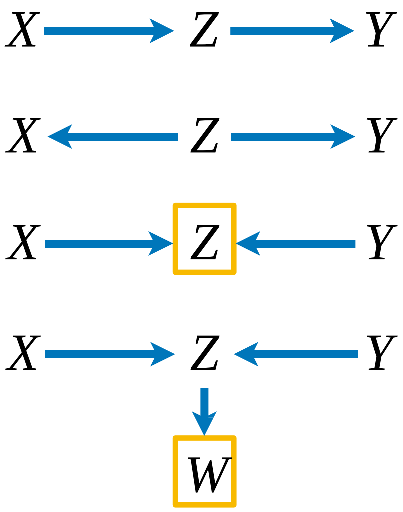
Active Triplets



$$(X \not\perp\!\!\!\perp Y \mid Z)$$

Triplets - Summary

Active Triplets

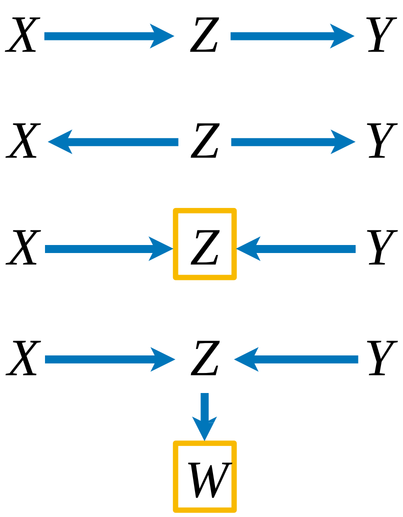


$(X \not\perp\!\!\!\perp Y \mid Z)$

Inactive Triplets

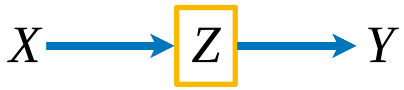
Triplets - Summary

Active Triplets



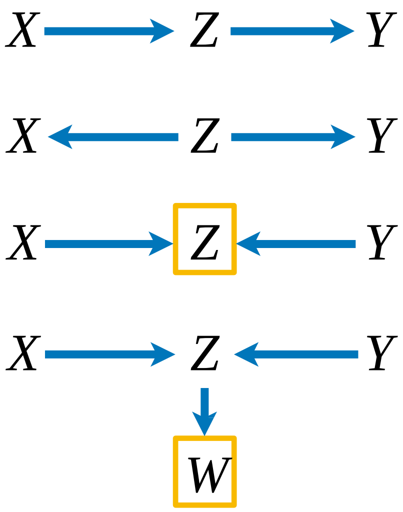
$$(X \not\perp\!\!\!\perp Y \mid Z)$$

Inactive Triplets



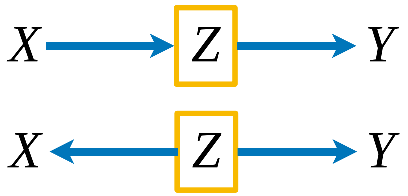
Triplets - Summary

Active Triplets



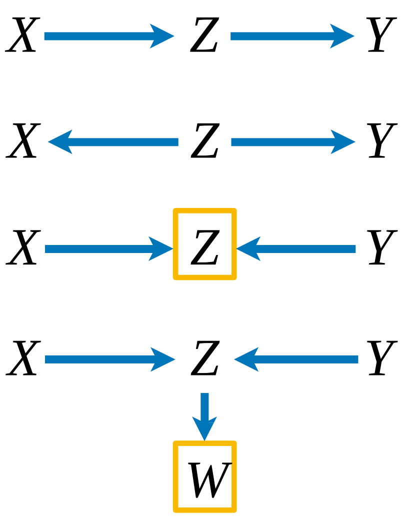
$$(X \not\perp\!\!\!\perp Y \mid Z)$$

Inactive Triplets



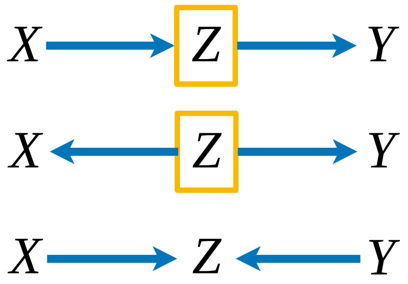
Triplets - Summary

Active Triplets



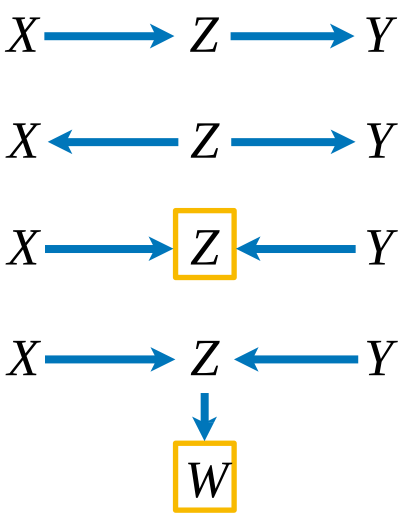
$$(X \not\perp\!\!\!\perp Y \mid Z)$$

Inactive Triplets



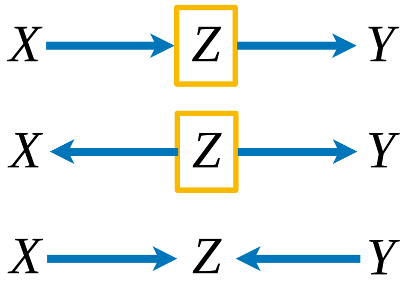
Triplets - Summary

Active Triplets



$(X \not\perp\!\!\!\perp Y \mid Z)$

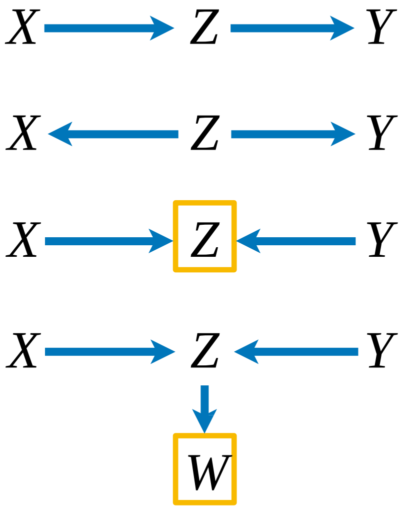
Inactive Triplets



$(X \perp\!\!\!\perp Y \mid Z)$

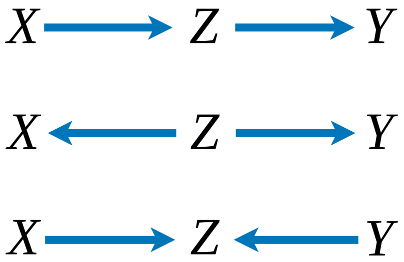
Triplets - Summary

Active Triplets



$(X \not\perp\!\!\!\perp Y \mid Z)$

Inactive Triplets



$(X \perp\!\!\!\perp Y \mid Z)$

What about larger graphical structures?

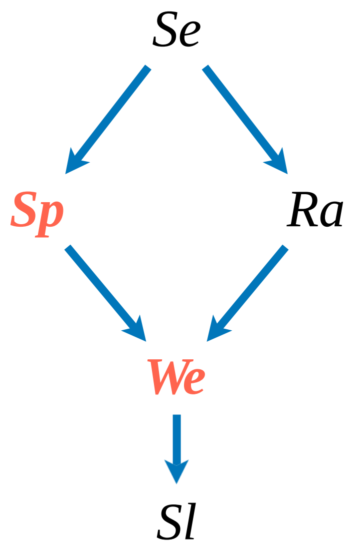
Graph Separation (d-Separation)

- Consider the question of whether $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent given $\boldsymbol{Z}$.

1. Look at every path from $\boldsymbol{X}$ to $\boldsymbol{Y}$ in the graph.
2. A path is active if every triplet in it is active (given $\boldsymbol{Z}$).
3. If any path is active $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not d-separated.

Graph Separation (d-Separation)

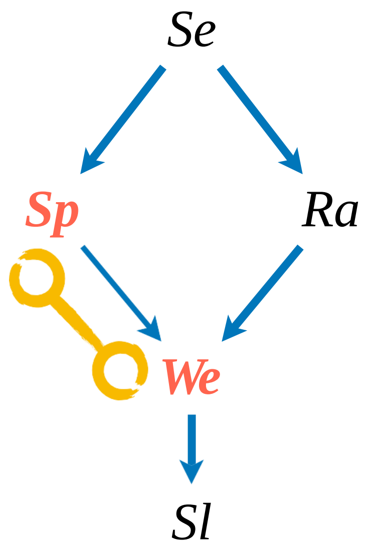
$(Wet \perp\!\!\!\perp Sprinkler)$?



Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Sprinkler)$?

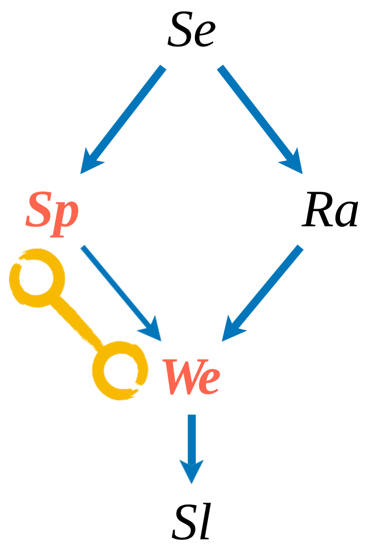
Path 1: $Sp \longrightarrow We$



Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Sprinkler)$?

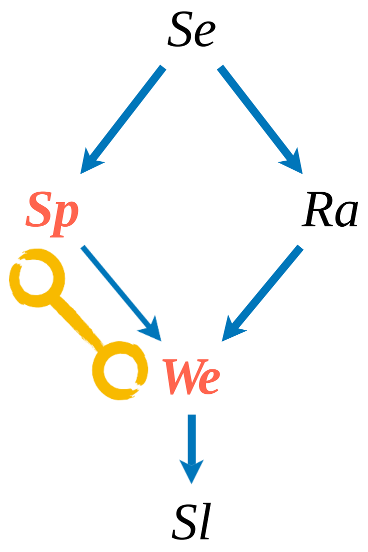
Path 1: $Sp \longrightarrow We$



Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Sprinkler)$?

Path 1: $Sp \longrightarrow We$ always active

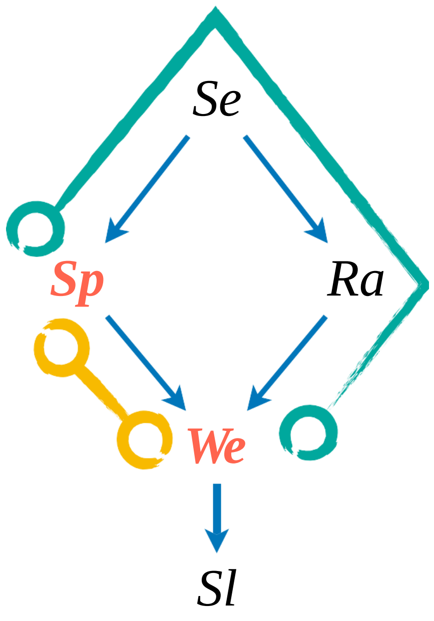


Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Sprinkler)$?

Path 1: $Sp \longrightarrow We$ always active

Path 2: $Sp \longleftarrow Se \longrightarrow Ra \longrightarrow We$

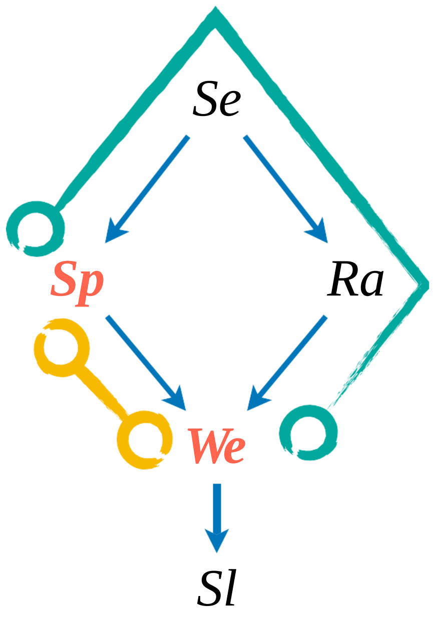


Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Sprinkler)$?

Path 1: $Sp \longrightarrow We$ always active

Path 2: $Sp \longleftarrow Se \longrightarrow Ra \longrightarrow We$

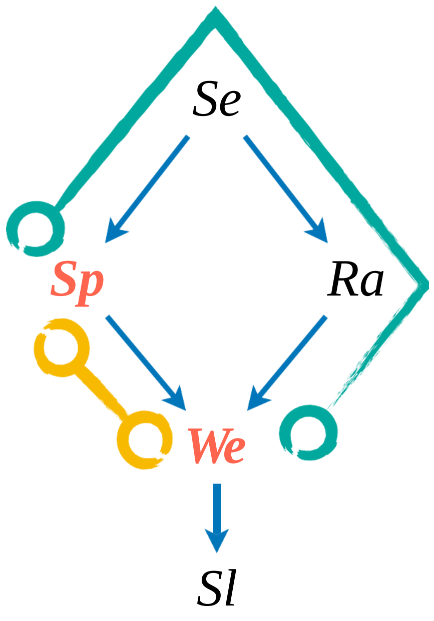


Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Sprinkler)$?

Path 1: $Sp \longrightarrow We$ always active

Path 2: $Sp \longleftarrow Se \longrightarrow Ra \longrightarrow We$



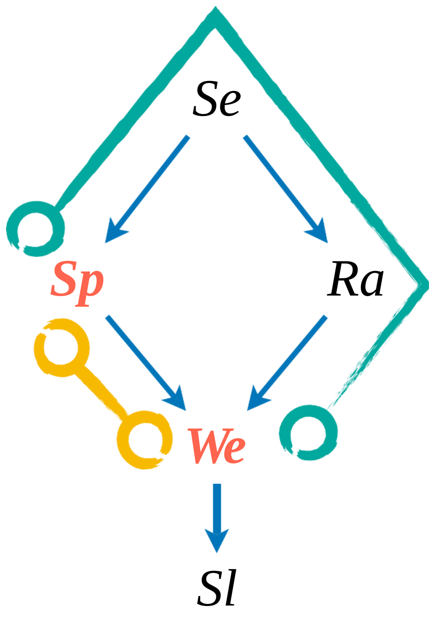
Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Sprinkler)$?

Path 1: $Sp \longrightarrow We$ always active

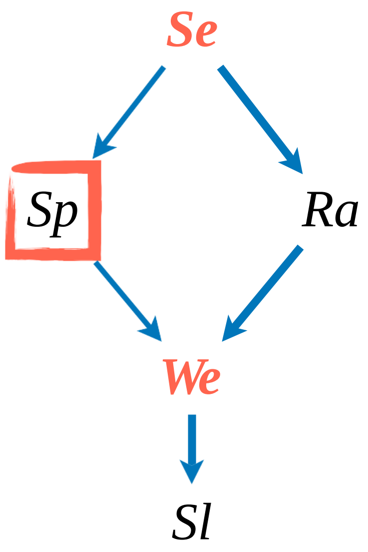
Path 2: $Sp \longleftarrow Se \longrightarrow Ra \longrightarrow We$

There exists a path (actually two) that is active, hence *Sprinkler* and *Wet* are not d-separated.



Graph Separation (d-Separation)

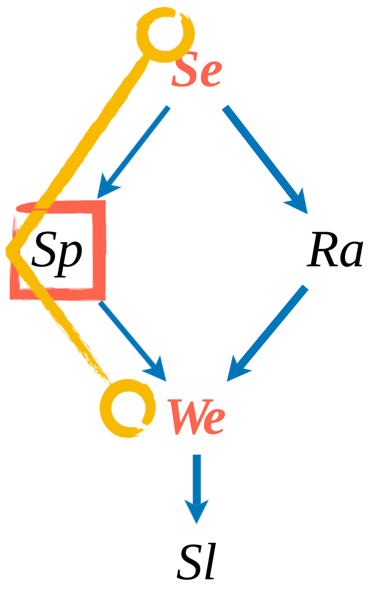
$(Wet \perp\!\!\!\perp Season \mid Sprinkler)$?



Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Season \mid Sprinkler)$?

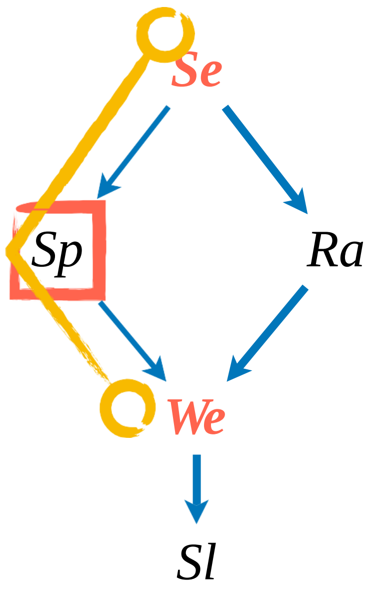
Path 1: $Se \longrightarrow Sp \longrightarrow We$



Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Season \mid Sprinkler)$?

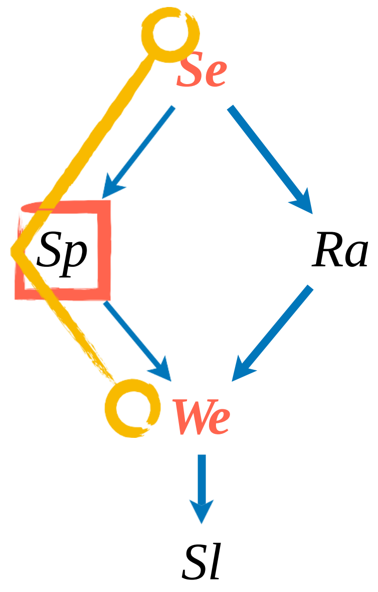
Path 1: $Se \longrightarrow Sp \longrightarrow We$



Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Season \mid Sprinkler)$?

Path 1: $Se \longrightarrow Sp \longrightarrow We$

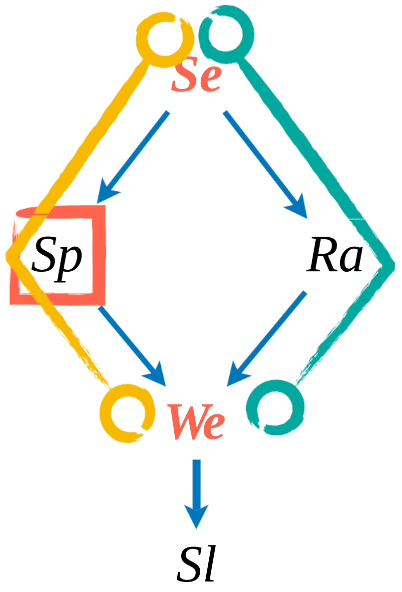


Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Season \mid Sprinkler)$?

Path 1: $Se \longrightarrow Sp \longrightarrow We$

Path 2: $Se \longrightarrow Ra \longrightarrow We$

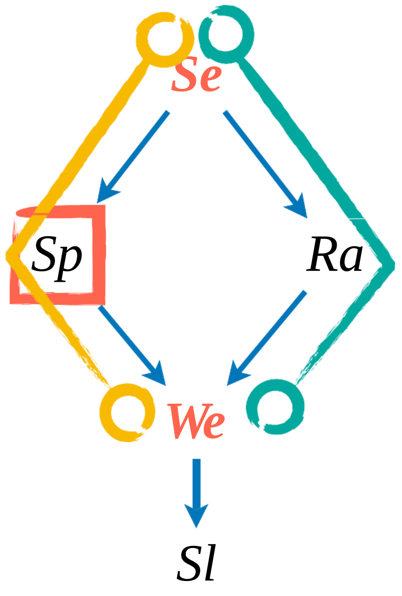


Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Season \mid Sprinkler)$?

Path 1: $Se \longrightarrow Sp \longrightarrow We$

Path 2: $Se \longrightarrow Ra \longrightarrow We$

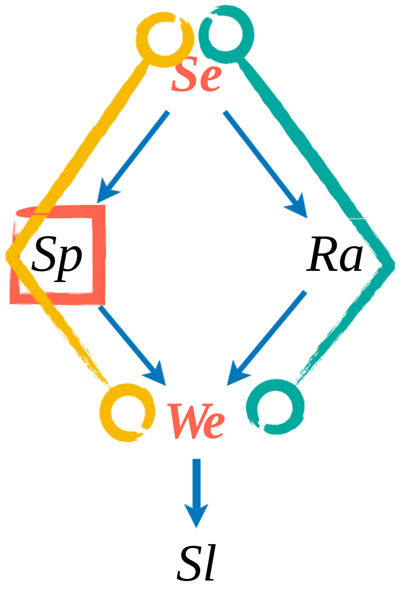


Graph Separation (d-Separation)

$(Wet \perp\!\!\!\perp Season \mid Sprinkler)$?

Path 1: $Se \longrightarrow Sp \longrightarrow We$

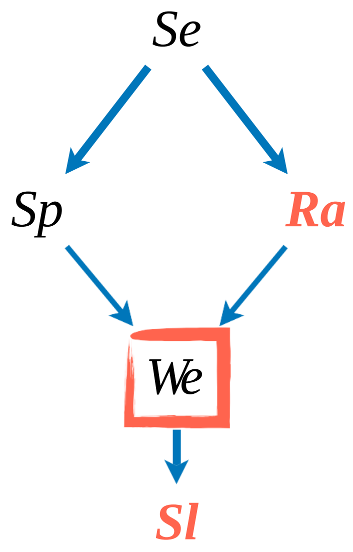
Path 2: $Se \longrightarrow Ra \longrightarrow We$



There exists a path that is active, hence *Wet* and *Season* are not d-separated given *Sprinkler*.

Graph Separation (d-Separation)

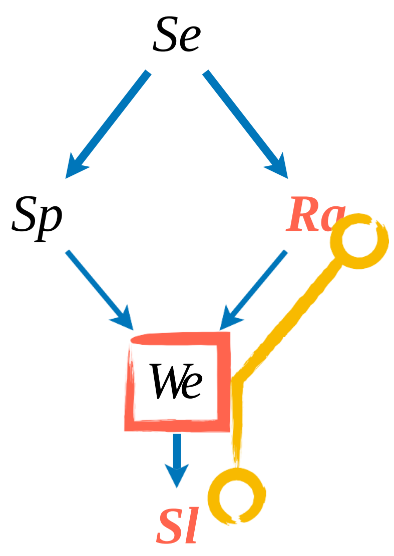
$(Rain \perp\!\!\!\perp Slippery \mid Wet)$?



Graph Separation (d-Separation)

$(Rain \perp\!\!\!\perp Slippery \mid Wet)$?

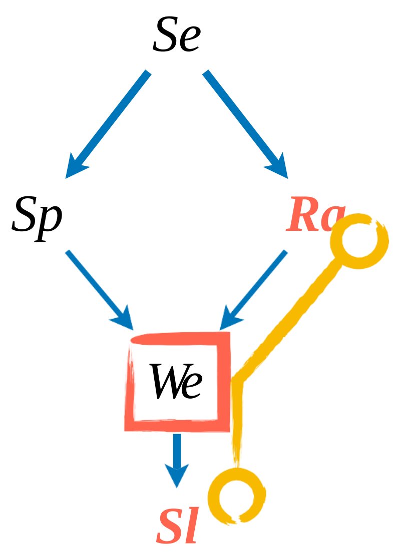
Path 1: $Ra \longrightarrow We \longrightarrow Sl$



Graph Separation (d-Separation)

$(Rain \perp\!\!\!\perp Slippery \mid Wet)$?

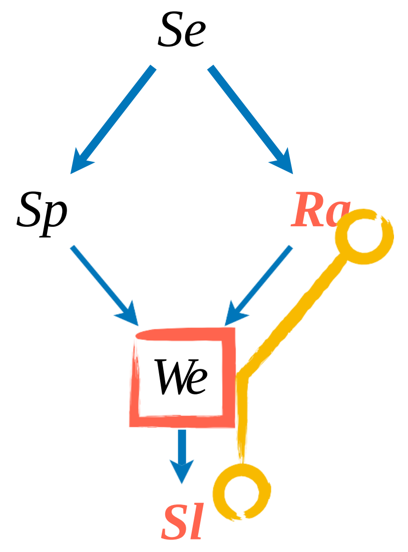
Path 1: $Ra \longrightarrow We \longrightarrow Sl$



Graph Separation (d-Separation)

$(Rain \perp\!\!\!\perp Slippery \mid Wet)$?

Path 1: $Ra \longrightarrow We \longrightarrow Sl$

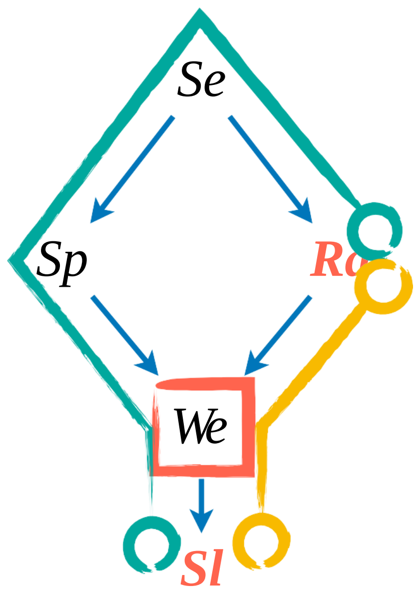


Graph Separation (d-Separation)

$(Rain \perp\!\!\!\perp Slippery \mid Wet)$?

Path 1: $Ra \longrightarrow We \longrightarrow Sl$

Path 2: $Ra \longleftarrow Se \longrightarrow Sp \longrightarrow We \longrightarrow Sl$

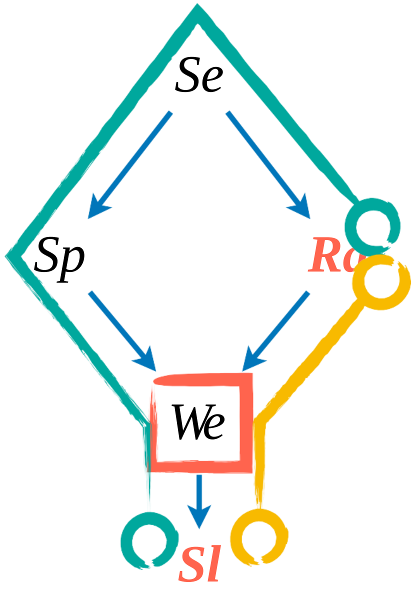


Graph Separation (d-Separation)

$(Rain \perp\!\!\!\perp Slippery \mid Wet)$?

Path 1: $Ra \longrightarrow We \longrightarrow Sl$

Path 2: $Ra \longleftarrow Se \longrightarrow Sp \longrightarrow We \longrightarrow Sl$

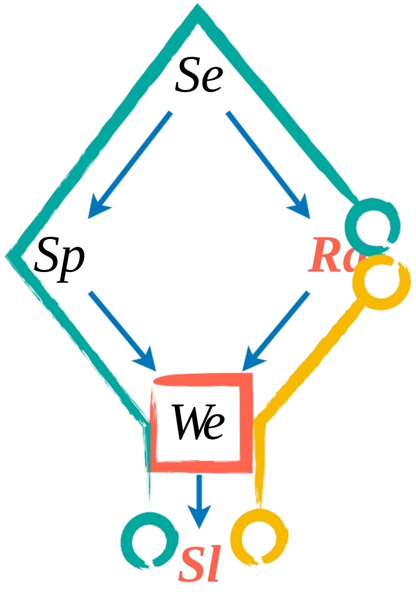


Graph Separation (d-Separation)

$(Rain \perp\!\!\!\perp Slippery \mid Wet)$?

Path 1: $Ra \longrightarrow We \longrightarrow Sl$

Path 2: $Ra \longleftarrow Se \longrightarrow Sp \longrightarrow We \longrightarrow Sl$

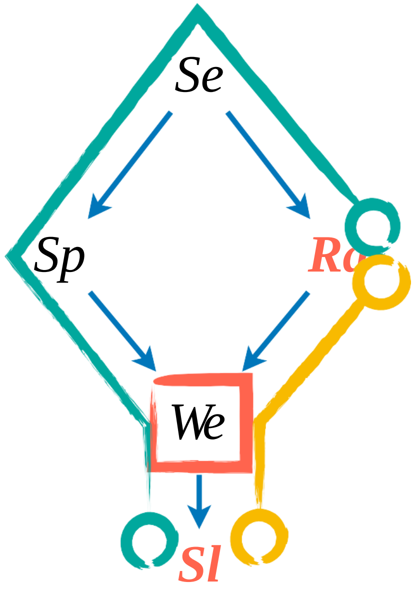


Graph Separation (d-Separation)

$(Rain \perp\!\!\!\perp Slippery \mid Wet)$?

Path 1: $Ra \longrightarrow We \longrightarrow Sl$

Path 2: $Ra \longleftarrow Se \longrightarrow Sp \longrightarrow We \longrightarrow Sl$

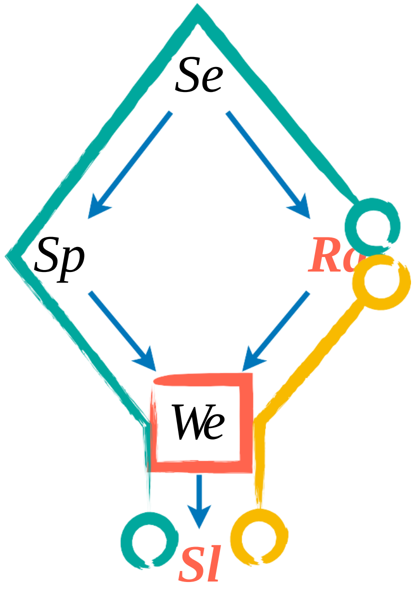


Graph Separation (d-Separation)

$(Rain \perp\!\!\!\perp Slippery \mid Wet)$?

Path 1: $Ra \longrightarrow We \longrightarrow Sl$

Path 2: $Ra \longleftarrow Se \longrightarrow Sp \longrightarrow We \longrightarrow Sl$



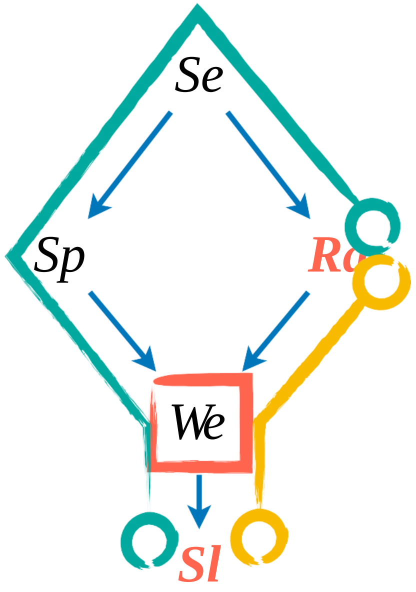
Graph Separation (d-Separation)

$(Rain \perp\!\!\!\perp Slippery \mid Wet)$?

Path 1: $Ra \longrightarrow We \longrightarrow Sl$

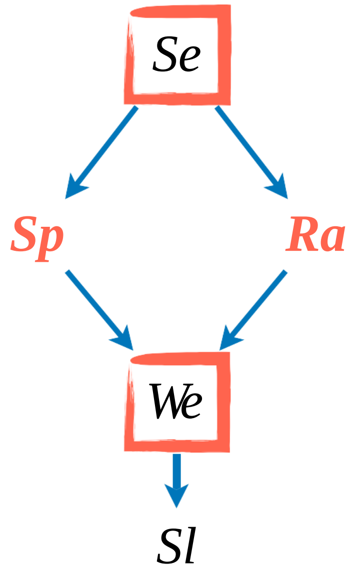
Path 2: $Ra \longleftarrow Se \longrightarrow Sp \longrightarrow We \longrightarrow Sl$

There exists **no** path that is active between *Rain* and *Slippery* given *Wet*, hence they are d-separated.



Graph Separation (d-Separation)

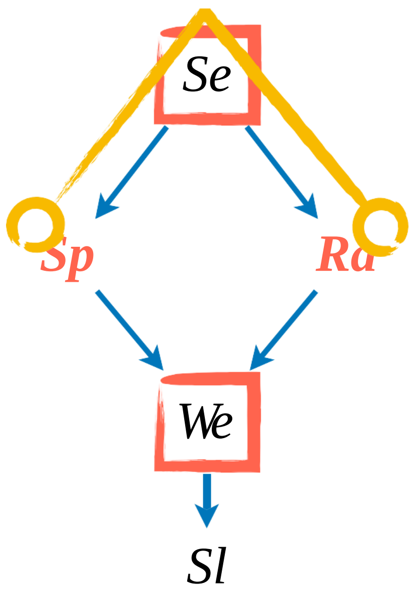
$(Sprinkler \perp\!\!\!\perp Rain \mid Season, Wet)$?



Graph Separation (d-Separation)

$(Sprinkler \perp\!\!\!\perp Rain \mid Season, Wet)$?

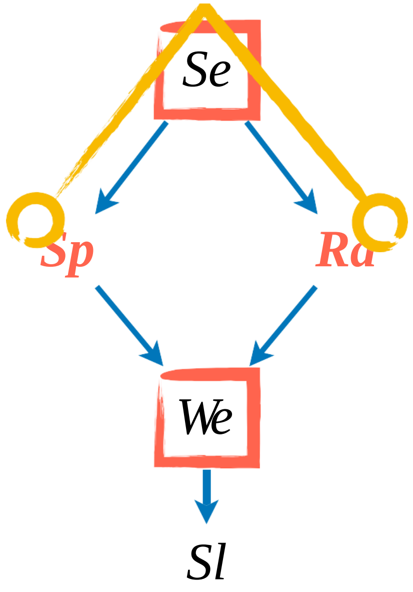
Path 1: $Sp \longleftarrow Se \longrightarrow Ra$



Graph Separation (d-Separation)

$(Sprinkler \perp\!\!\!\perp Rain \mid Season, Wet)$?

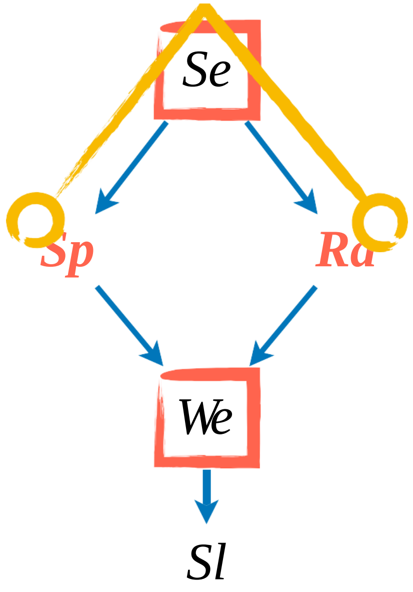
Path 1: $Sp \longleftarrow Se \longrightarrow Ra$



Graph Separation (d-Separation)

$(Sprinkler \perp\!\!\!\perp Rain \mid Season, Wet)$?

Path 1: $Sp \longleftarrow Se \longrightarrow Ra$

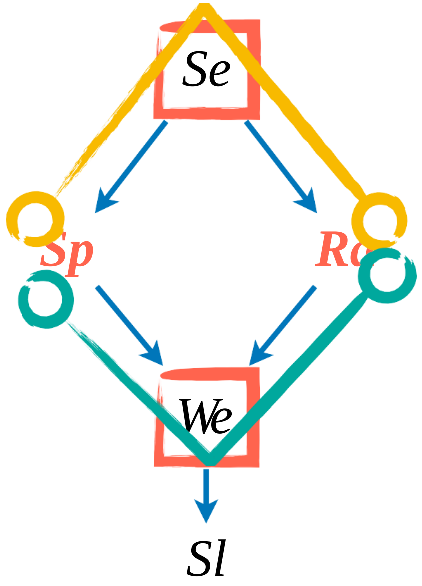


Graph Separation (d-Separation)

$(Sprinkler \perp\!\!\!\perp Rain \mid Season, Wet)$?

Path 1: $Sp \longleftarrow Se \longrightarrow Ra$

Path 2: $Sp \longrightarrow We \longleftarrow Ra$

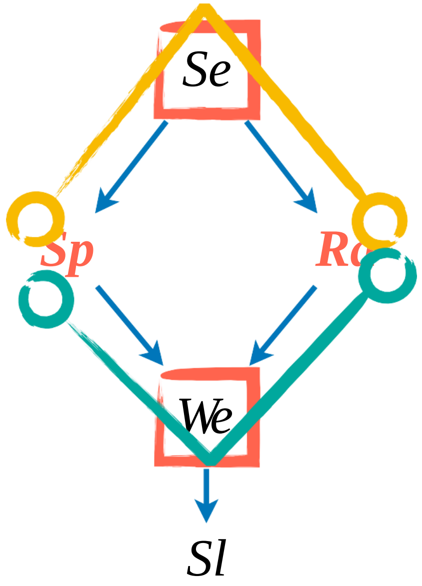


Graph Separation (d-Separation)

(Sprinkler ⫫ *Rain* | *Season, Wet)*?

Path 1: $Sp \longleftarrow Se \longrightarrow Ra$

Path 2: $Sp \longrightarrow We \longleftarrow Ra$

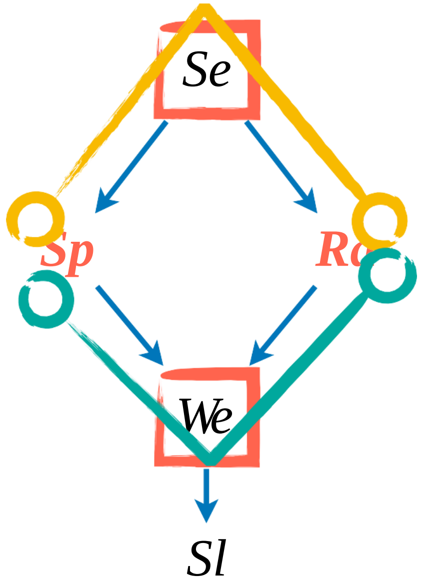


Graph Separation (d-Separation)

(Sprinkler ⫫ Rain | Season, Wet)?

Path 1: $Sp \longleftarrow Se \longrightarrow Ra$

Path 2: $Sp \longrightarrow We \longleftarrow Ra$ becomes active given *We*

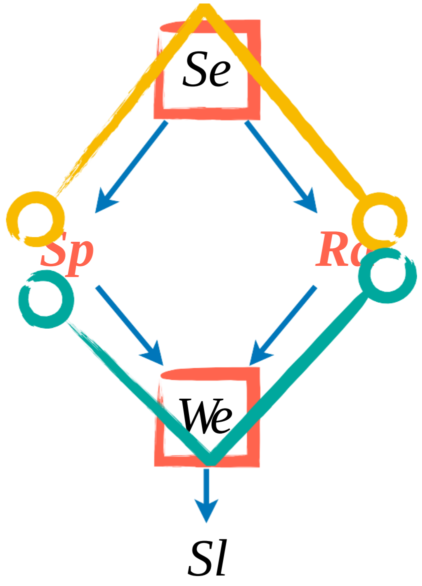


Graph Separation (d-Separation)

(Sprinkler ⫫ *Rain* | *Season, Wet)*?

Path 1: $Sp \longleftarrow Se \longrightarrow Ra$

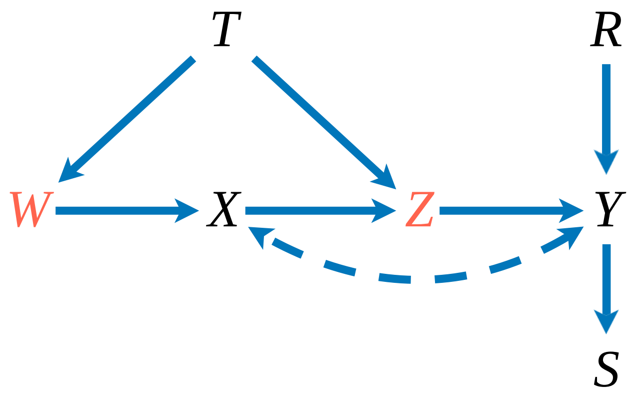
Path 2: $Sp \longrightarrow We \longleftarrow Ra$ becomes active given *We*



There exists a path that is active between *Sprinkler* and *Rain* given *Season* and *Wet*, hence they are **not** d-separated.

Graph Separation (d-Separation)

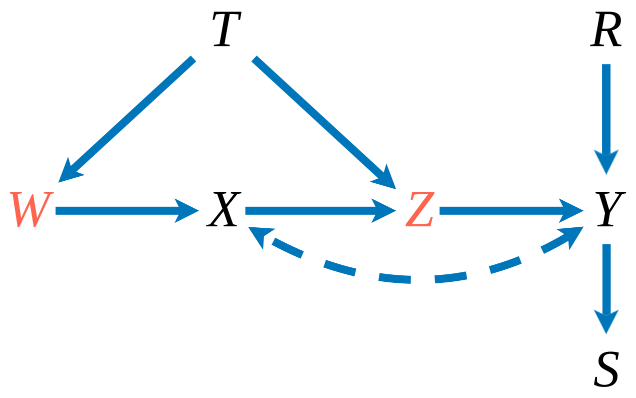
Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Z \mid \boldsymbol{A})$ holds?



Graph Separation (d-Separation)

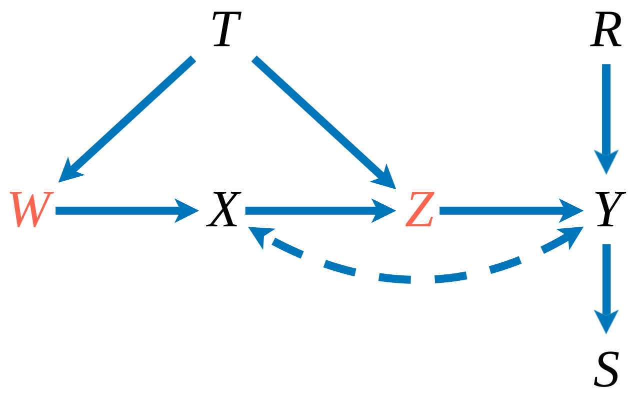
Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Z \mid \boldsymbol{A})$ holds?

Path 1: $W \longleftarrow T \longrightarrow Z$



Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Z \mid \boldsymbol{A})$ holds?

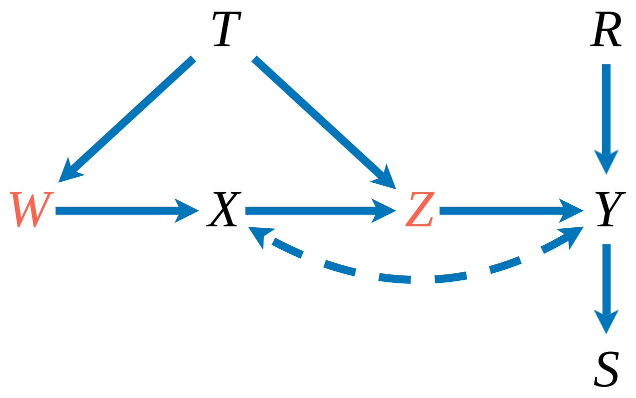


Path 1: $W \longleftarrow \; T \longrightarrow \; Z$

Path 2: $W \longrightarrow \; X \longrightarrow \; Z$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Z \mid \boldsymbol{A})$ holds?



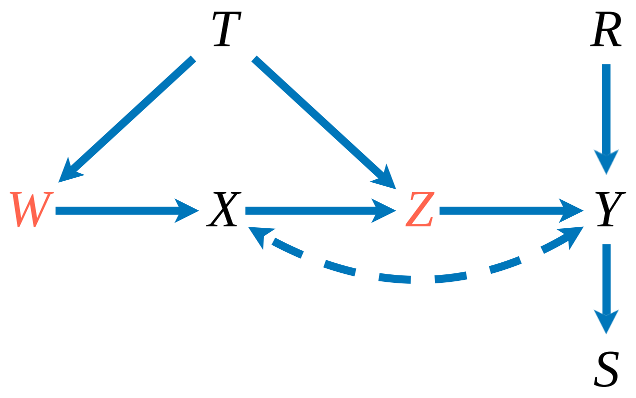
Path 1: $W \longleftarrow T \longrightarrow Z$

Path 2: $W \longrightarrow X \longrightarrow Z$

Path 3: $W \longrightarrow X \longleftrightarrow Y \longleftarrow Z$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Z \mid \boldsymbol{A})$ holds?



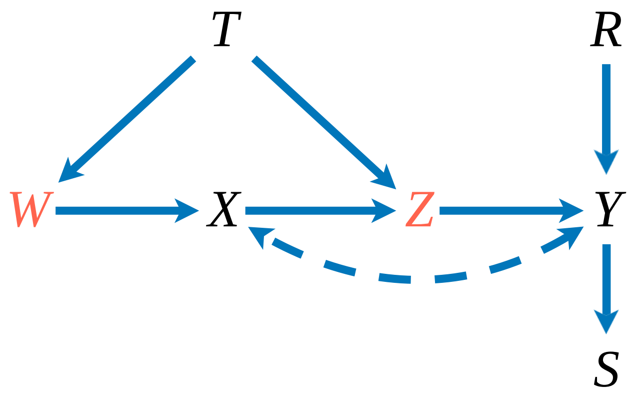
Path 1: $W \longleftarrow T \longrightarrow Z$

Path 2: $W \longrightarrow X \longrightarrow Z$

Path 3: $W \longrightarrow X \longleftrightarrow Y \longleftarrow Z$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Z \mid \boldsymbol{A})$ holds?



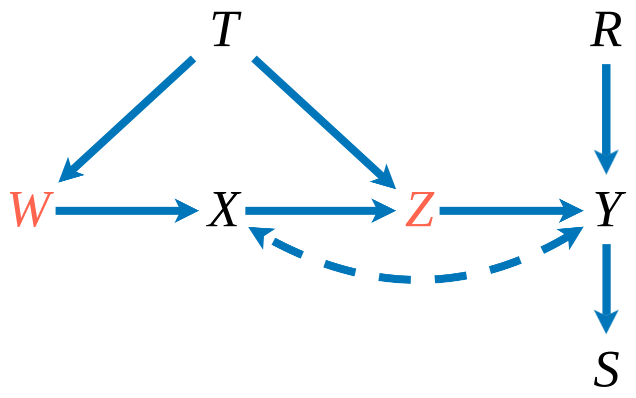
Path 1: $W \longleftarrow T \longrightarrow Z$

Path 2: $W \longrightarrow X \longrightarrow Z$

Path 3: $W \longrightarrow X \longleftrightarrow Y \longleftarrow Z$ = $W \longrightarrow X \longleftarrow U \longrightarrow Y \longleftarrow Z$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Z \mid \boldsymbol{A})$ holds?



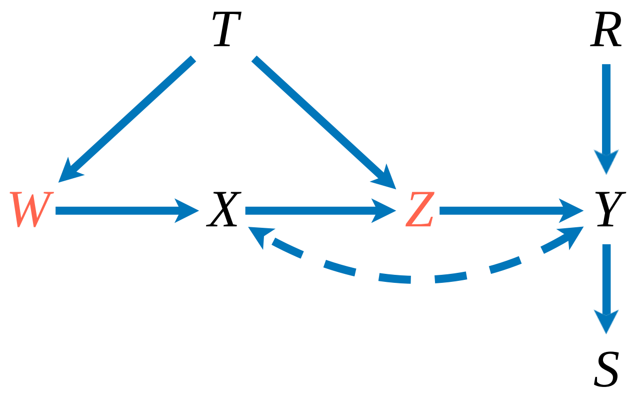
Path 1: $W \longleftarrow T \longrightarrow Z$

Path 2: $W \longrightarrow X \longrightarrow Z$

Path 3: $W \longrightarrow X \longleftrightarrow Y \longleftarrow Z$ $= W \longrightarrow X \longleftarrow U \longrightarrow Y \longleftarrow Z$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Z \mid \boldsymbol{A})$ holds?



Path 1: $W \longleftarrow T \longrightarrow Z$

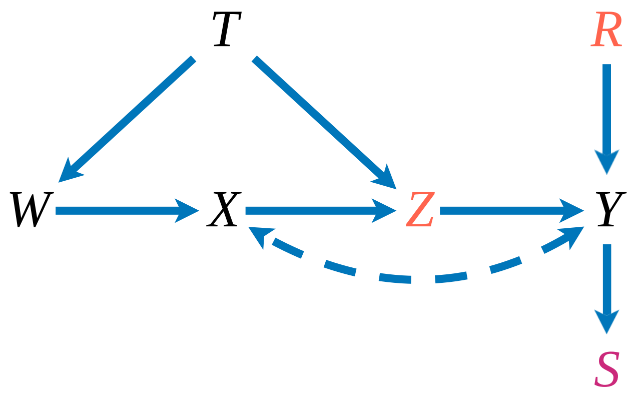
Path 2: $W \longrightarrow X \longrightarrow Z$

Path 1 and 2 need to be blocked, Path 3 is naturally blocked:
$\boldsymbol{A}=\{T, X\}$ suffices.

Path 3: $W \longrightarrow X \longleftrightarrow Y \longleftarrow Z$ $= W \longrightarrow X \longleftarrow U \longrightarrow Y \longleftarrow Z$

Graph Separation (d-Separation)

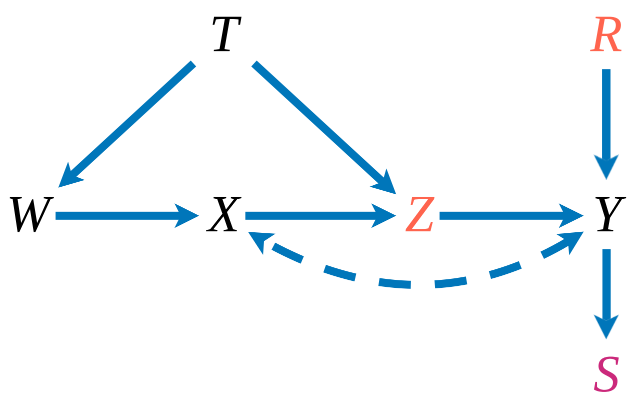
Is there a set $\boldsymbol{A}$ such that the separation statement $(R, Z \perp\!\!\!\perp S \mid \boldsymbol{A})$ holds?



Graph Separation (d-Separation)

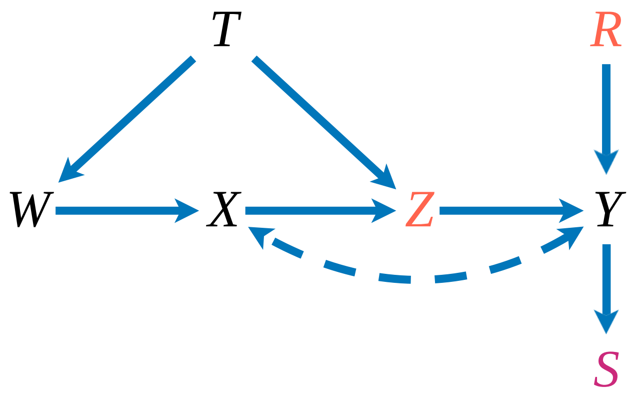
Is there a set $\boldsymbol{A}$ such that the separation statement $(R, Z \perp\!\!\!\perp S \mid \boldsymbol{A})$ holds?

Path 1: $R \longrightarrow Y \longrightarrow S$



Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(R, Z \perp\!\!\!\perp S \mid \boldsymbol{A})$ holds?



Path 1: $R \longrightarrow Y \longrightarrow S$

Path 2: $Z \longrightarrow Y \longrightarrow S$

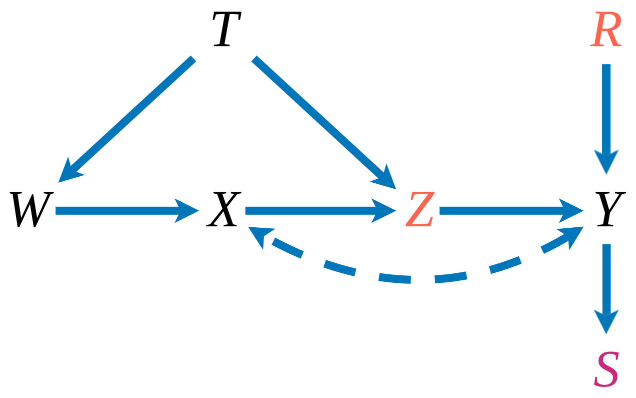
Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(R, Z \perp\!\!\!\perp S \mid \boldsymbol{A})$ holds?

Path 1: $R \longrightarrow Y \longrightarrow S$

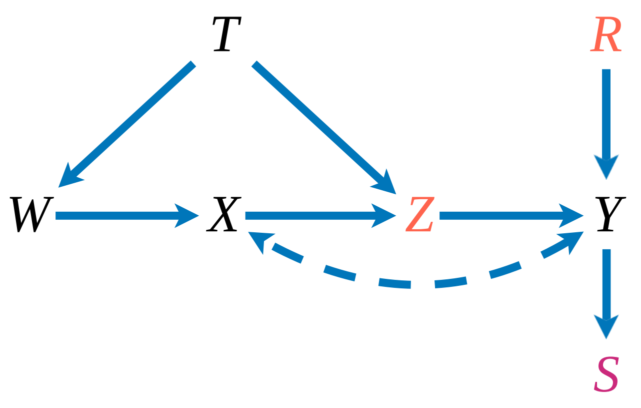
Path 2: $Z \longrightarrow Y \longrightarrow S$

Path 3: $Z \longleftarrow X \longleftrightarrow Y \longrightarrow S$



Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(R, Z \perp\!\!\!\perp S \mid \boldsymbol{A})$ holds?



Path 1: $R \longrightarrow Y \longrightarrow S$

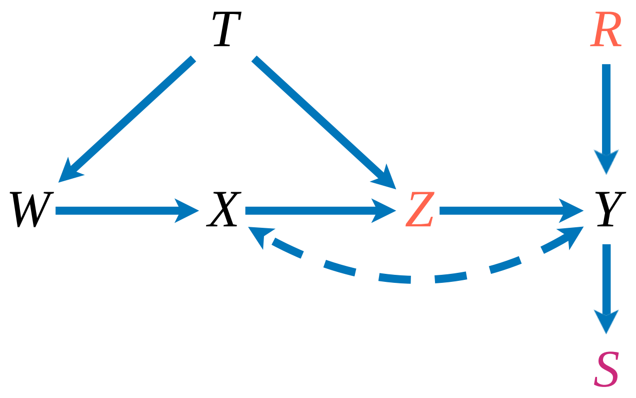
Path 2: $Z \longrightarrow Y \longrightarrow S$

Path 3: $Z \longleftarrow X \longleftrightarrow Y \longrightarrow S$

Path 4: $Z \longleftarrow T \longrightarrow W \longrightarrow X \longleftrightarrow Y \longrightarrow S$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(R, Z \perp\!\!\!\perp S \mid \boldsymbol{A})$ holds?



Path 1: $R \longrightarrow Y \longrightarrow S$

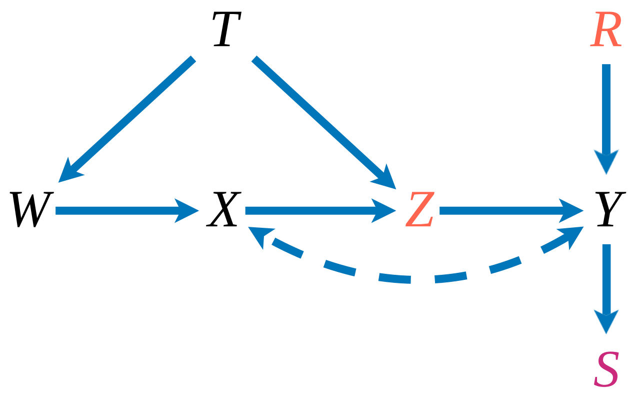
Path 2: $Z \longrightarrow Y \longrightarrow S$

Path 3: $Z \longleftarrow X \longleftrightarrow Y \longrightarrow S$

Path 4: $Z \longleftarrow T \longrightarrow W \longrightarrow X \longleftrightarrow Y \longrightarrow S$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(R, Z \perp\!\!\!\perp S \mid \boldsymbol{A})$ holds?



Path 1: $R \longrightarrow Y \longrightarrow S$

Path 2: $Z \longrightarrow Y \longrightarrow S$

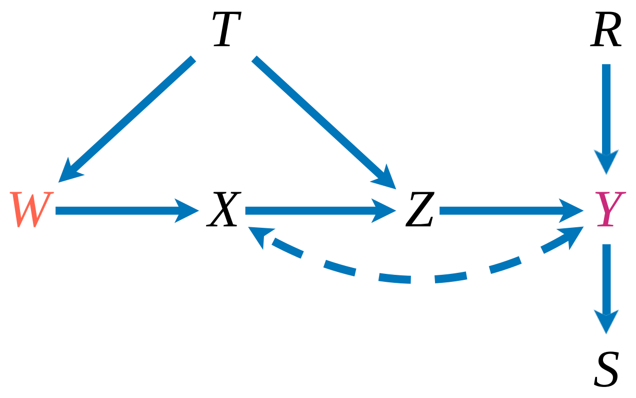
Path 3: $Z \longleftarrow X \longleftrightarrow Y \longrightarrow S$

Path 4: $Z \longleftarrow T \longrightarrow W \longrightarrow X \longleftrightarrow Y \longrightarrow S$

$\boldsymbol{A} = \{Y\}$ suffices.

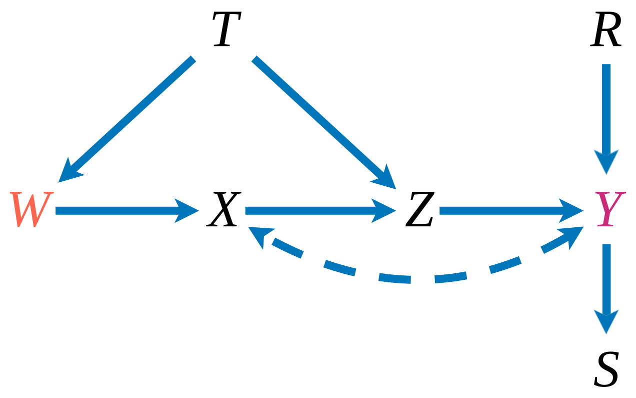
Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Graph Separation (d-Separation)

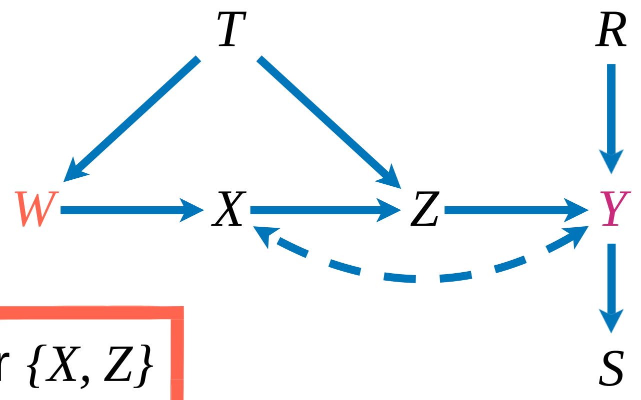
Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?

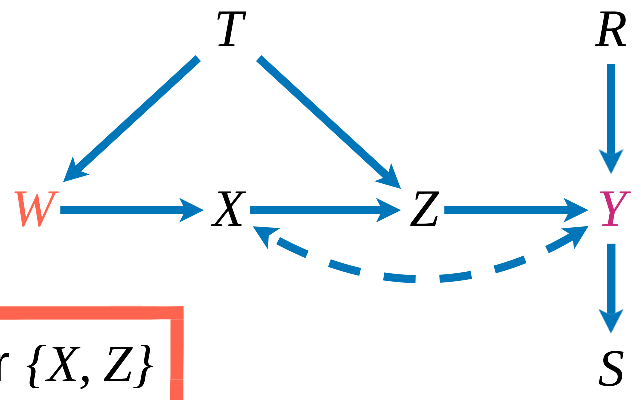


Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$

$\{X\}$ or $\{Z\}$ or $\{X, Z\}$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



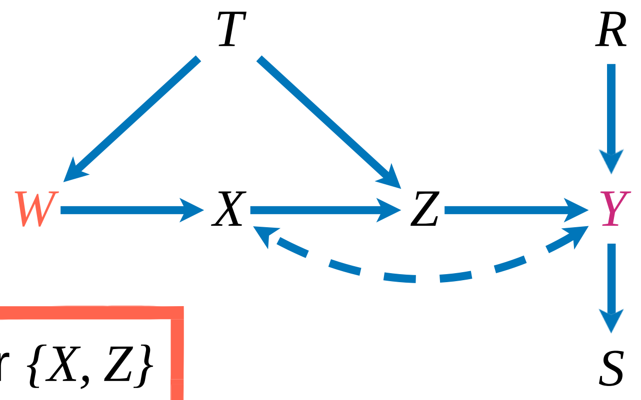
Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$

$\{X\}$ or $\{Z\}$ or $\{X, Z\}$

Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?

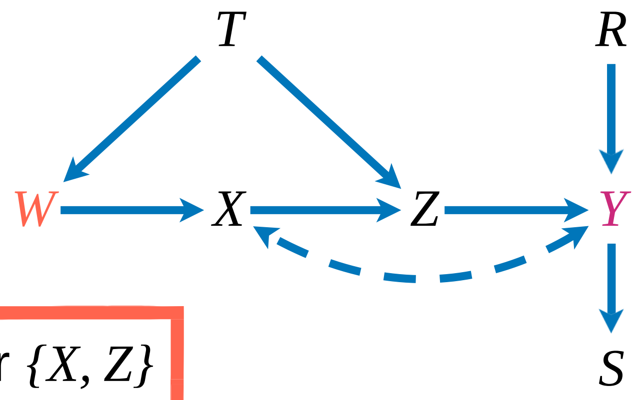


Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ $\{X\}$ or $\{Z\}$ or $\{X, Z\}$

Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



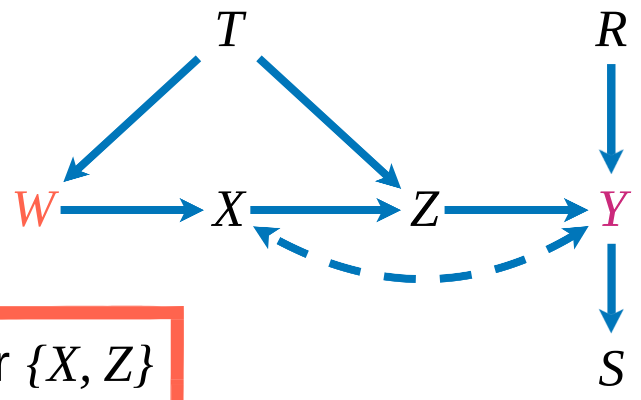
Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ $\{X\}$ or $\{Z\}$ or $\{X, Z\}$

Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Path 3: $W \longrightarrow X \longleftrightarrow Y$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



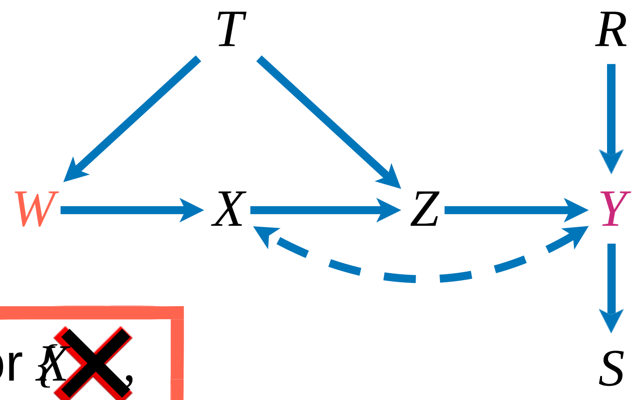
Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — $\{X\}$ or $\{Z\}$ or $\{X, Z\}$

Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Path 3: $W \longrightarrow X \longleftrightarrow Y$ — not X

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



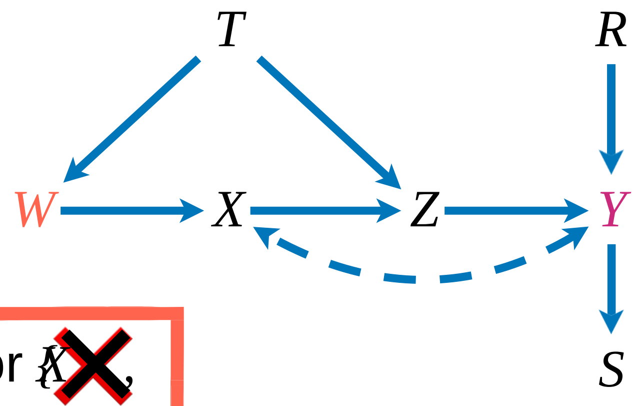
Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — ~~$\{X\}$~~ or $\{Z\}$ or ~~$\{X, Z\}$~~

Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Path 3: $W \longrightarrow X \longleftrightarrow Y$ — not X

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — ~~$\{X\}$~~ or $\{Z\}$ or ~~$\{X, Z\}$~~

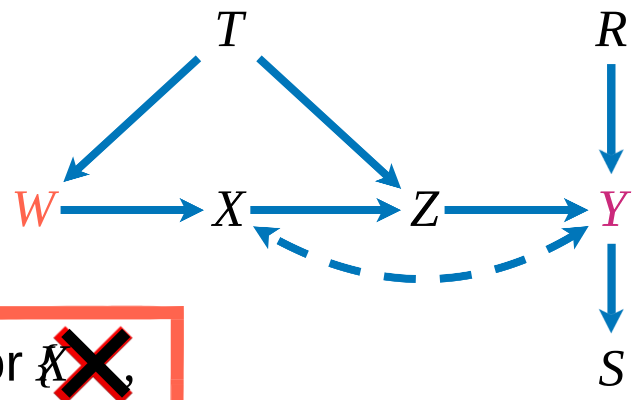
Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Path 3: $W \longrightarrow X \longleftrightarrow Y$ — not X

Path 4: $W \longleftarrow T \longrightarrow Z \longleftarrow X \longleftrightarrow Y$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — ~~$\{X\}$~~ or $\{Z\}$ or ~~$\{X, Z\}$~~

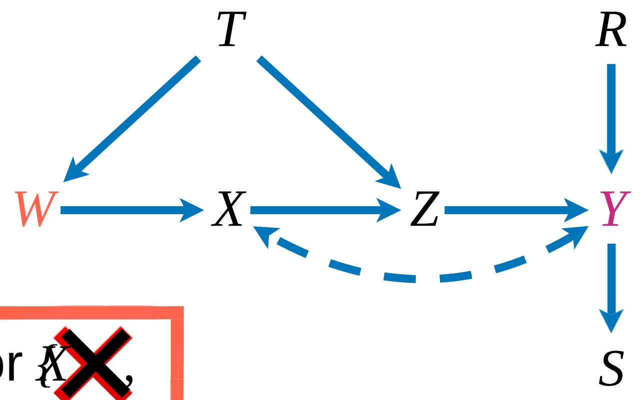
Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Path 3: $W \longrightarrow X \longleftrightarrow Y$ — not X

Path 4: $W \longleftarrow T \longrightarrow Z \longleftarrow X \longleftrightarrow Y$ — $\{T\}$ or $\{X\}$ or $\{T, X\}$ or $\{T, Z\}$ or $\{T,X,Z\}$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — ~~$\{X\}$~~ or $\{Z\}$ or ~~$\{X, Z\}$~~

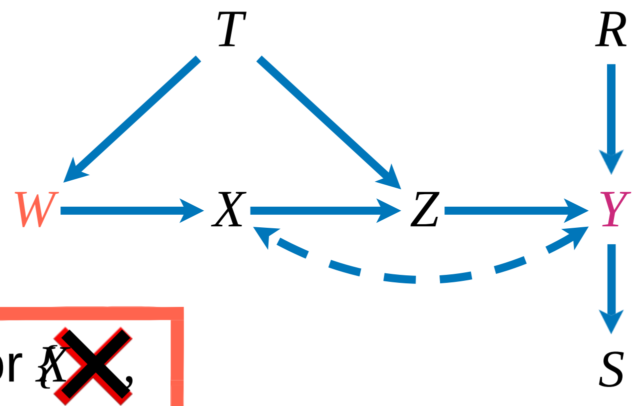
Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Path 3: $W \longrightarrow X \longleftrightarrow Y$ — not X

Path 4: $W \longleftarrow T \longrightarrow Z \longleftarrow X \longleftrightarrow Y$ — $\{T\}$ or ~~$\{Z\}$~~ or ~~$\{T, X\}$~~ or $\{T, Z\}$ or ~~$\{T, X, Z\}$~~

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — ~~$\{X\}$~~ or $\{Z\}$ or ~~$\{X, Z\}$~~

Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

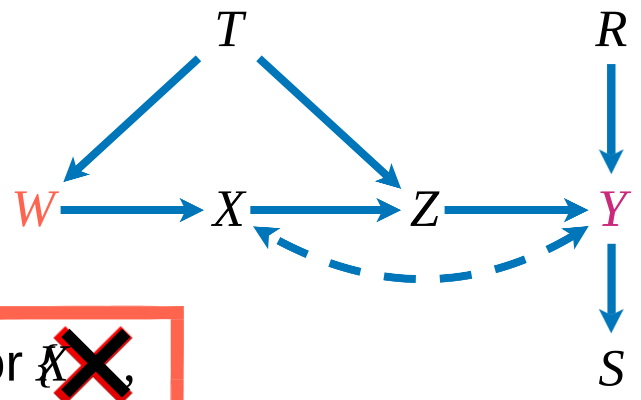
Path 3: $W \longrightarrow X \longleftrightarrow Y$ — not X

Does $\boldsymbol{A}=\{T, Z\}$ suffice?

Path 4: $W \longleftarrow T \longrightarrow Z \longleftarrow X \longleftrightarrow Y$ — $\{T\}$ or ~~$\{X\}$~~ or ~~$\{T, X\}$~~ or $\{T, Z\}$ or ~~$\{T, X, Z\}$~~

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — ~~$\{X\}$~~ or $\{Z\}$ or ~~$\{X, Z\}$~~

Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

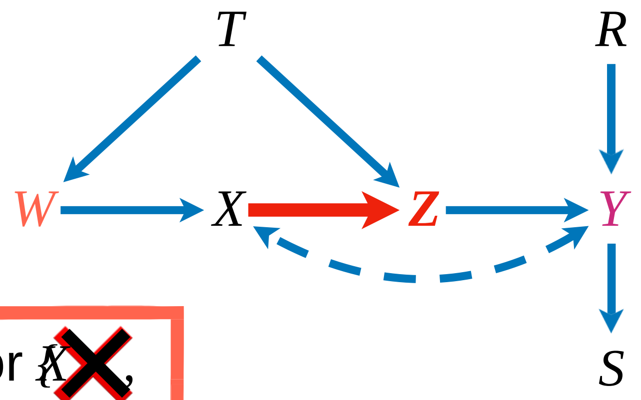
Path 3: $W \longrightarrow X \longleftrightarrow Y$ — not X

Does $\boldsymbol{A}=\{T, Z\}$ suffice?

Path 4: $W \longleftarrow T \longrightarrow Z \longleftarrow X \longleftrightarrow Y$ — $\{T\}$ or ~~$\{X\}$~~ or ~~$\{T, X\}$~~ or $\{T, Z\}$ or ~~$\{T, X, Z\}$~~

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — ~~$\{X\}$~~ or $\{Z\}$ or ~~$\{X, Z\}$~~

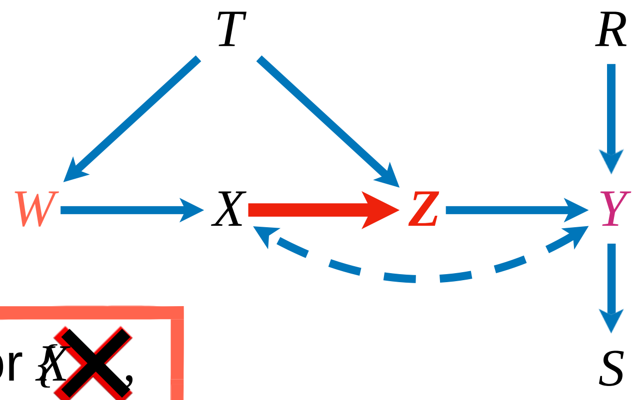
Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Path 3: $W \longrightarrow X \longleftrightarrow Y$ — not X

Path 4: $W \longleftarrow T \longrightarrow Z \longleftarrow X \longleftrightarrow Y$ — $\{T\}$ or ~~$\{X\}$~~ or ~~$\{T, X\}$~~ or $\{T, Z\}$ or ~~$\{T, X, Z\}$~~

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — $\{X\}$ or $\{Z\}$ or $\{X, Z\}$

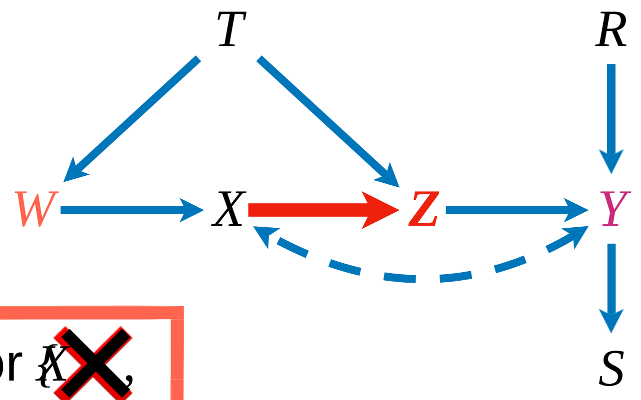
Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Path 3: $W \longrightarrow X \longleftrightarrow Y$, $X \longrightarrow Z$ — not X

Path 4: $W \longleftarrow T \longrightarrow Z \longleftarrow X \longleftrightarrow Y$ — $\{T\}$ or $\{X\}$ or $\{T, X\}$ or $\{T, Z\}$ or $\{T, X, Z\}$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — ~~$\{X\}$~~ or $\{Z\}$ or ~~$\{X, Z\}$~~

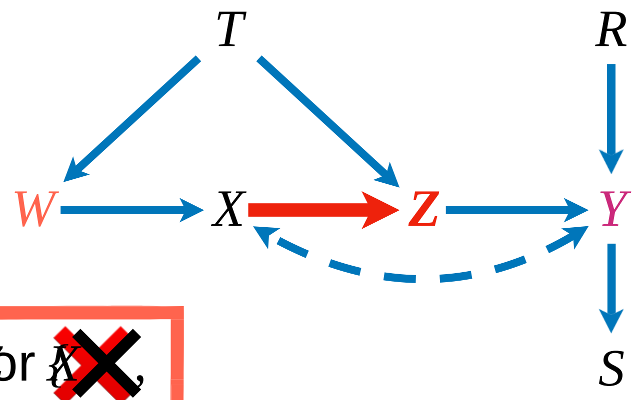
Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Path 3: $W \longrightarrow X \longleftrightarrow Y$, $X \to Z$ — not X | not Z

Path 4: $W \longleftarrow T \longrightarrow Z \longleftarrow X \longleftrightarrow Y$ — $\{T\}$ or ~~$\{X\}$~~ or ~~$\{T, X\}$~~ or $\{T, Z\}$ or ~~$\{T, X, Z\}$~~

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?



Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — $\{X\}$ or $\{Z\}$ or $\{X, Z\}$

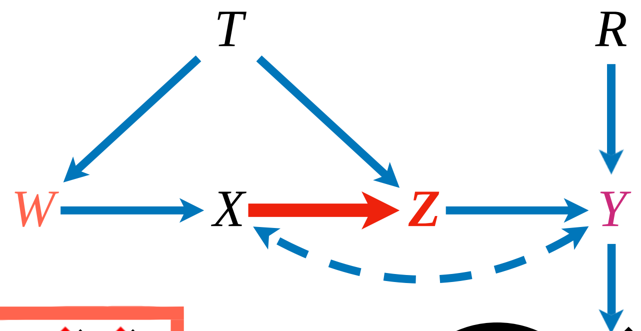
Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, Z\}$

Path 3: $W \longrightarrow X \longleftrightarrow Y$, $X \longrightarrow Z$ — not X | not Z

Path 4: $W \longleftarrow T \longrightarrow Z \longleftarrow X \longleftrightarrow Y$ — $\{T\}$ or $\{Z\}$ or $\{T, X\}$ or $\{X, Z\}$ or $\{T, X, Z\}$

Graph Separation (d-Separation)

Is there a set $\boldsymbol{A}$ such that the separation statement $(W \perp\!\!\!\perp Y \mid \boldsymbol{A})$ holds?

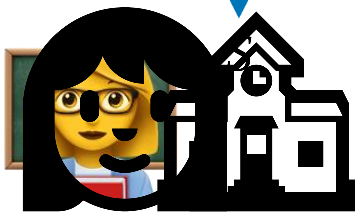


Path 1: $W \longrightarrow X \longrightarrow Z \longrightarrow Y$ — ~~$\{X\}$ or $\{Z\}$ or $\{X, Z\}$~~

Path 2: $W \longleftarrow T \longrightarrow Z \longrightarrow Y$ — $\{T\}$ or ~~$\{Z\}$~~ or ~~$\{T, Z\}$~~

Path 3: $W \longrightarrow X \longleftrightarrow Y$, $X \rightarrow Z$ — not X, not Z

Path 4: $W \longleftarrow T \longrightarrow Z \longleftarrow X \longleftrightarrow Y$ — $\{T\}$ or ~~$\{X\}$~~ or ~~$\{T, X\}$~~ or ~~$\{T, Z\}$~~ or ~~$\{T, X, Z\}$~~



No such $\boldsymbol{A}$!

Don't forget the descendants of the colliders!

d-SEPARATION (EXAMPLE)

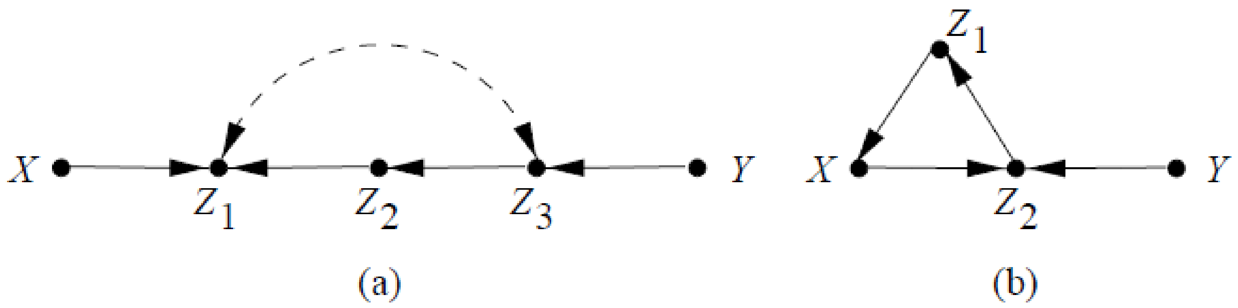


Figure 1.3: Graphs illustrating d-separation. In (a), X and Y are d-separated given Z_2 and d-connected given Z_1. In (b), X and Y cannot be d-separated by any set of nodes.