The Flow of Association and Causation in Graphs

Brady Neal

causalcourse.com

Graph terminology

Bayesian networks and causal graphs

The basic building blocks of graphs

The flow of association and causation

Graph terminology

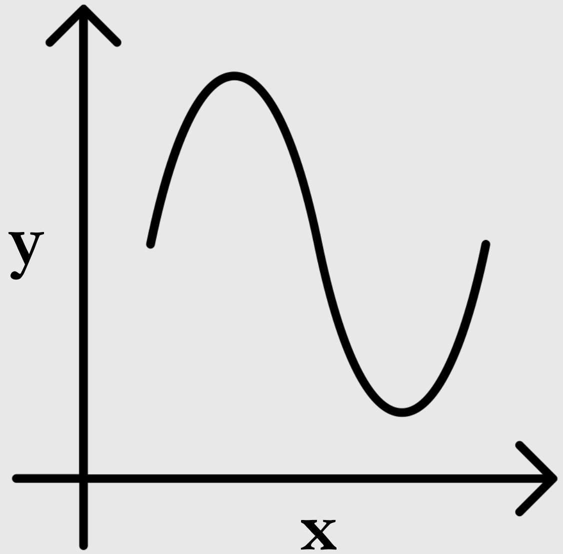
Bayesian networks and causal graphs

The basic building blocks of graphs

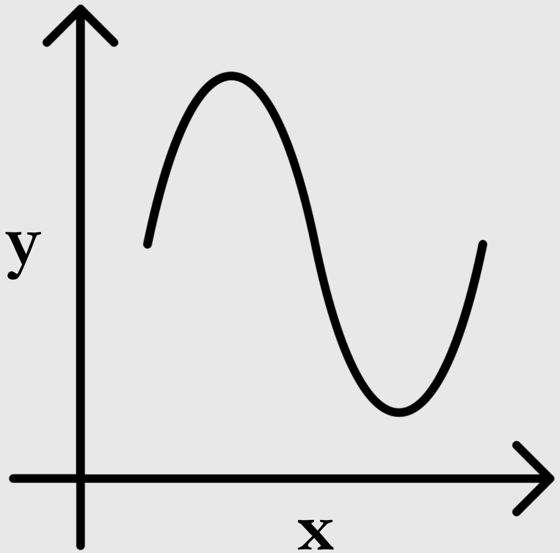
The flow of association and causation

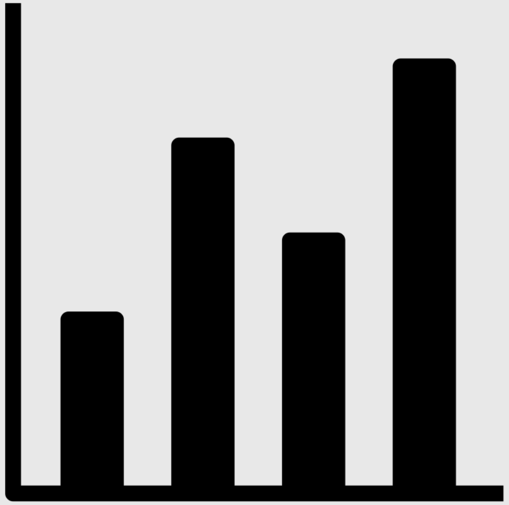
Graph terminology: not a graph

Graph terminology: not a graph

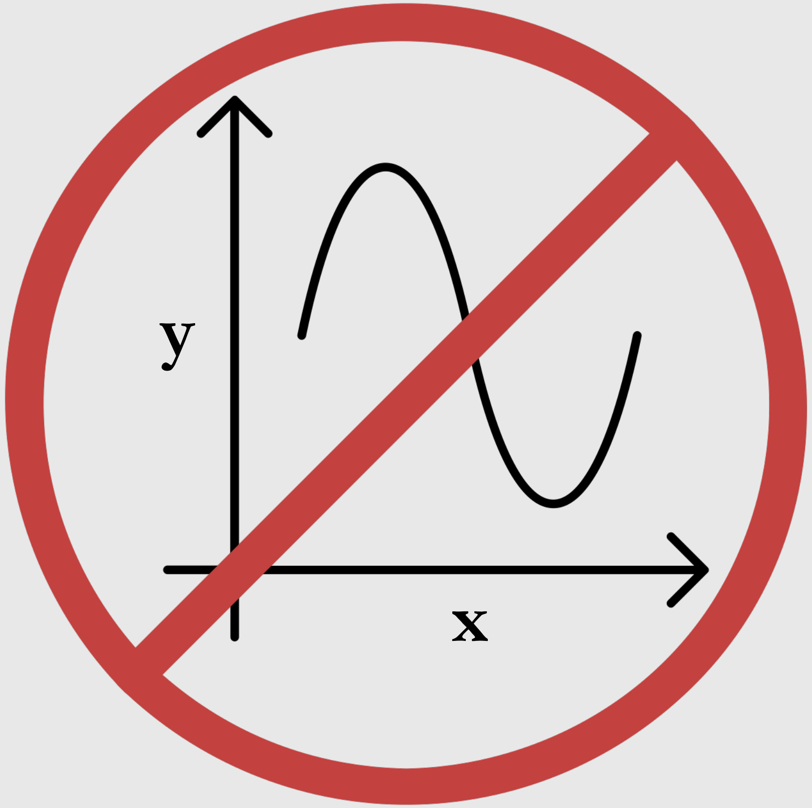


Graph terminology: not a graph

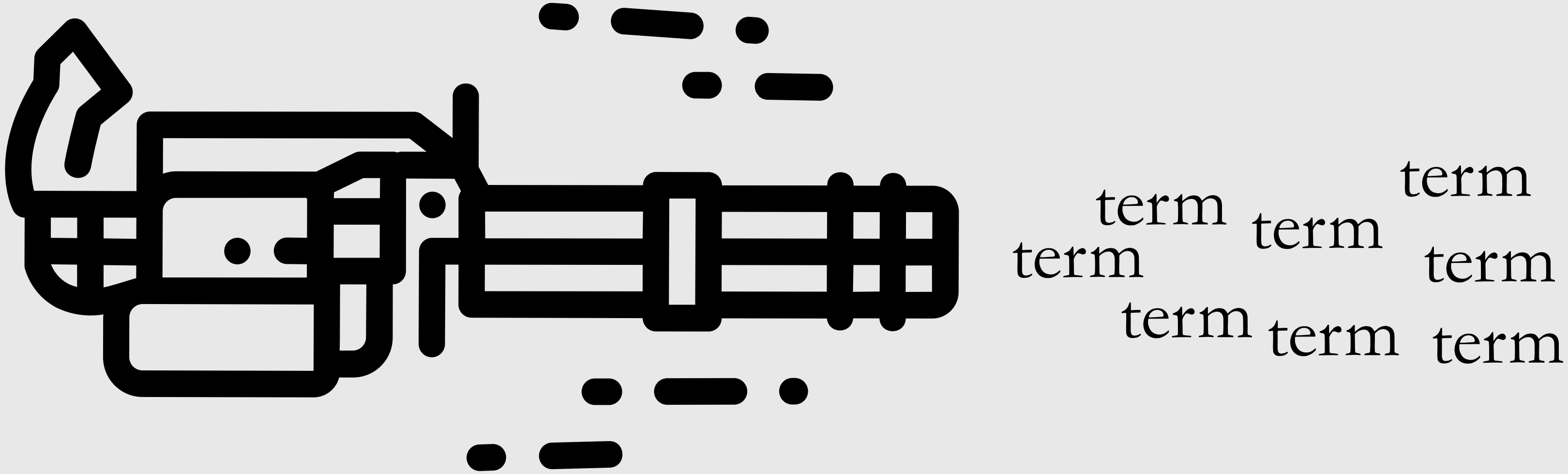




Graph terminology: not a graph

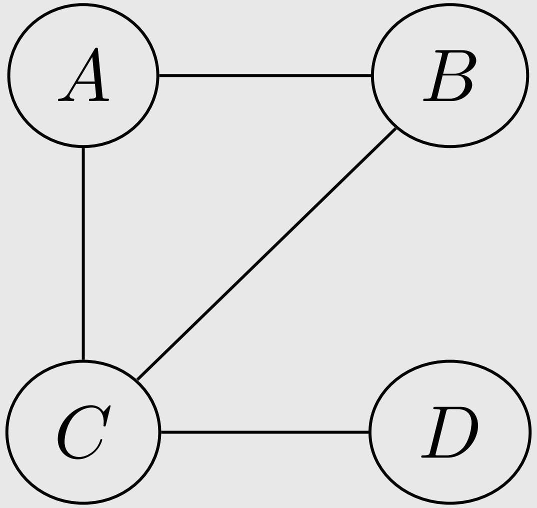


Graph terminology: Terminology Machine Gun



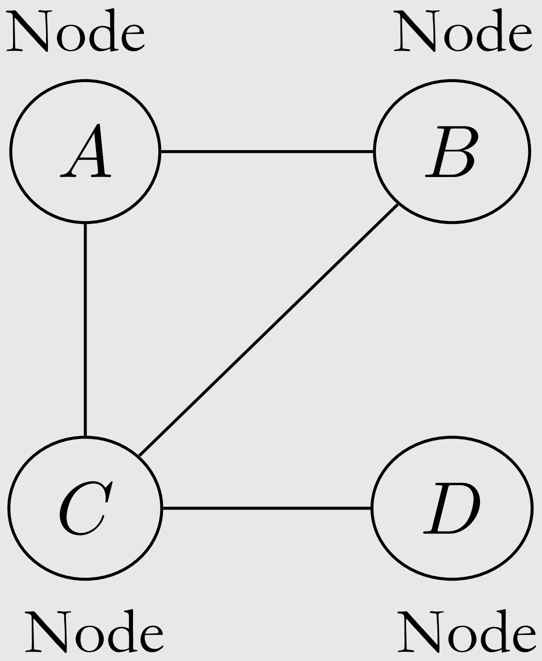
Graph terminology

Graph terminology



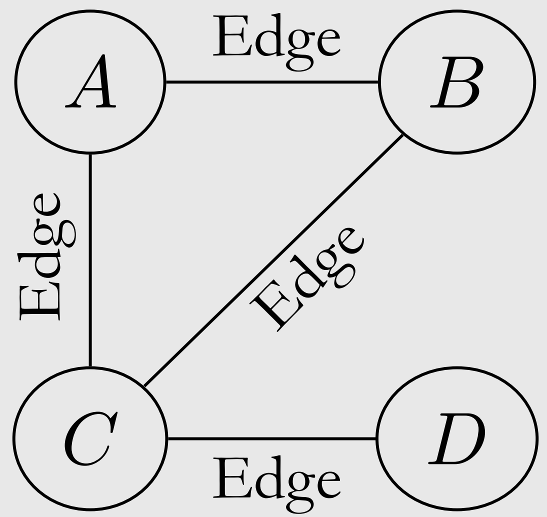
Graph terminology

Nodes



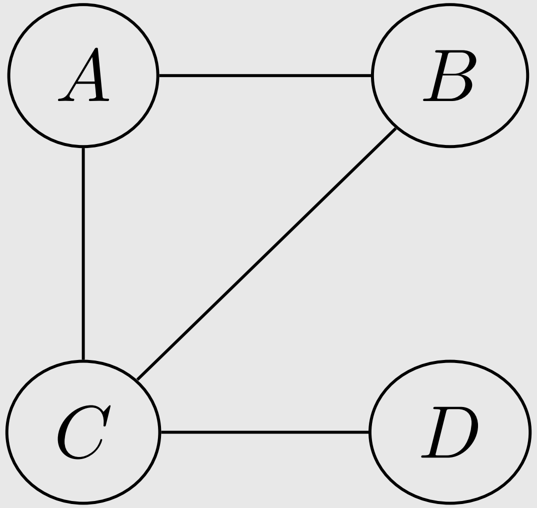
Graph terminology

Edges



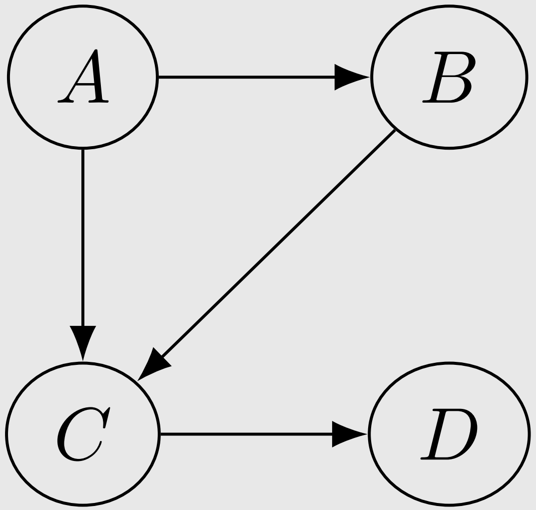
Graph terminology

Undirected Graph



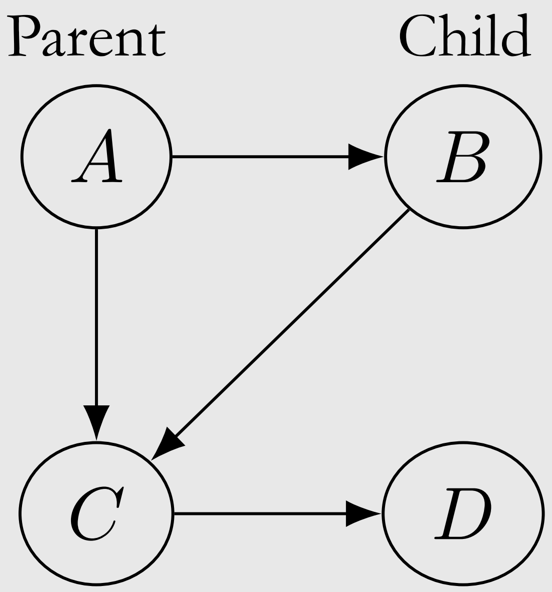
Graph terminology

Directed Graph



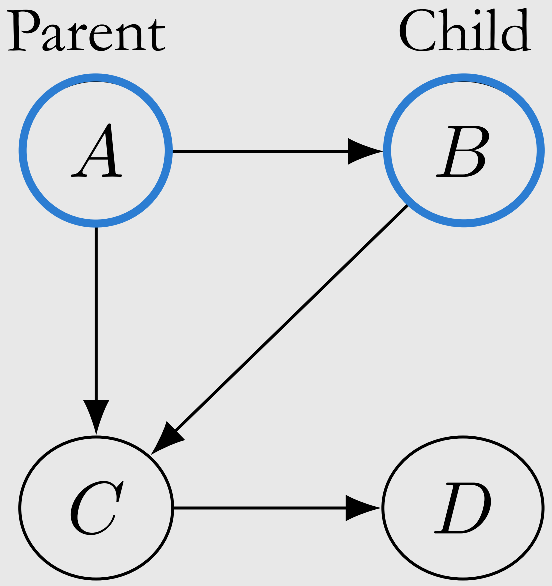
Graph terminology

Directed Graph



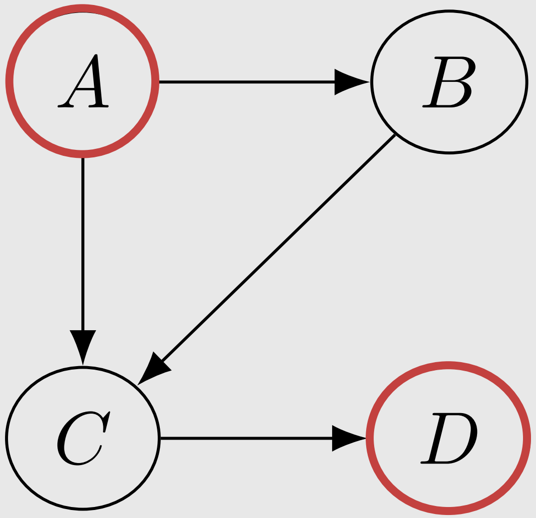
Graph terminology

Adjacent



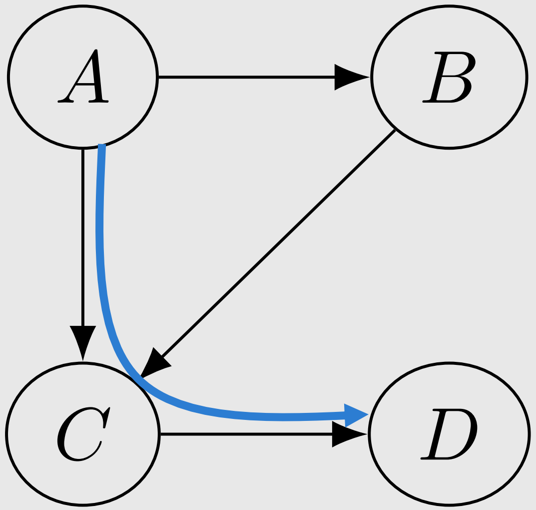
Graph terminology

Not
Adjacent



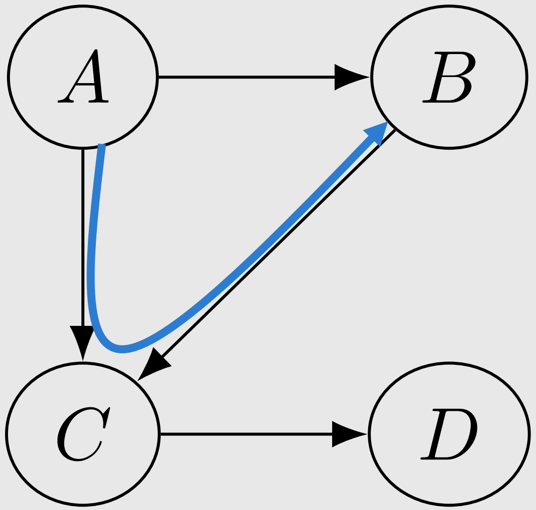
Graph terminology

Path



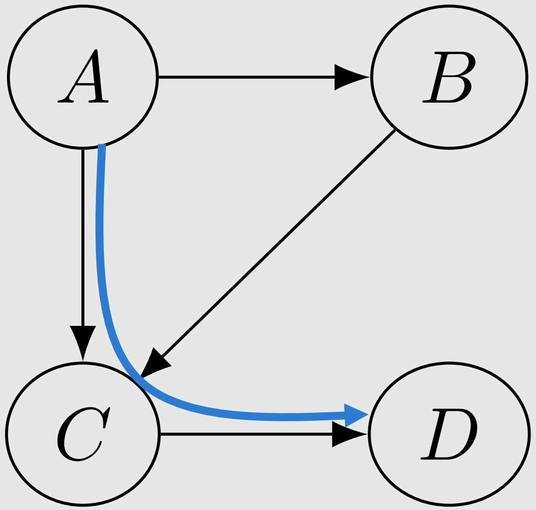
Graph terminology

Path

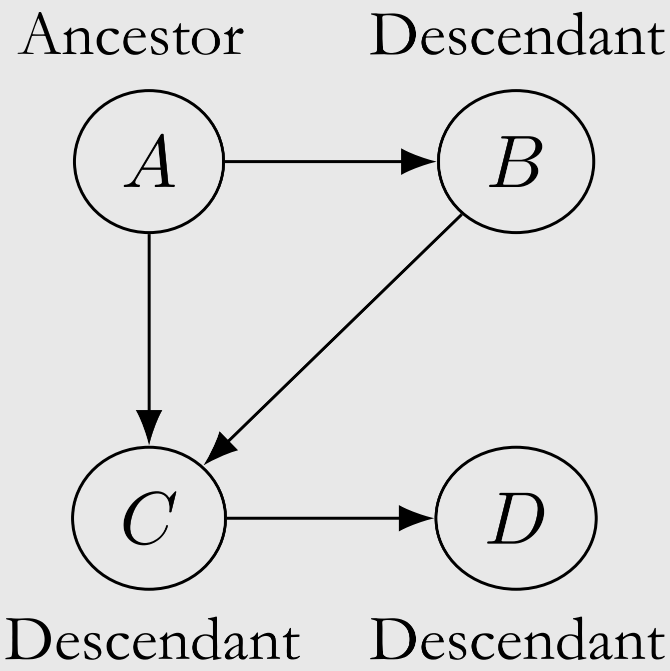


Graph terminology

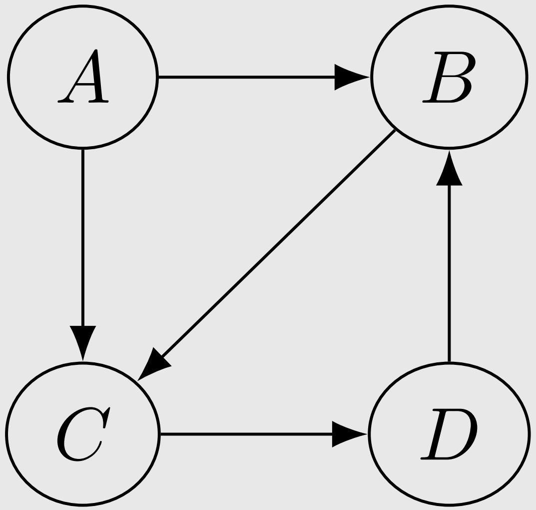
Directed
Path



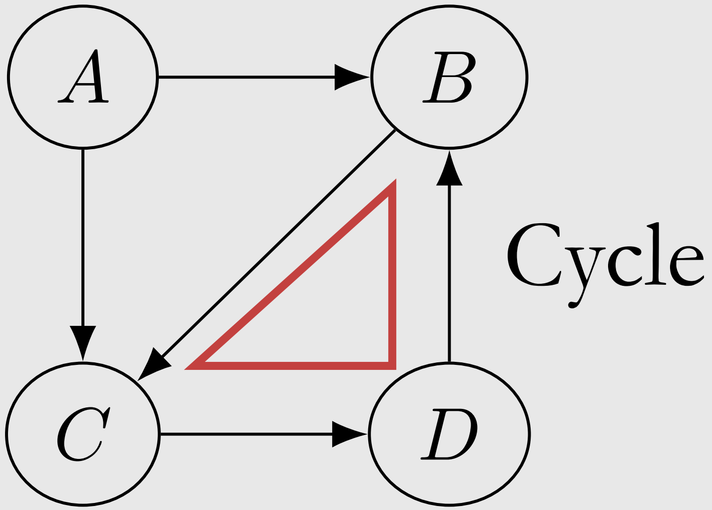
Graph terminology



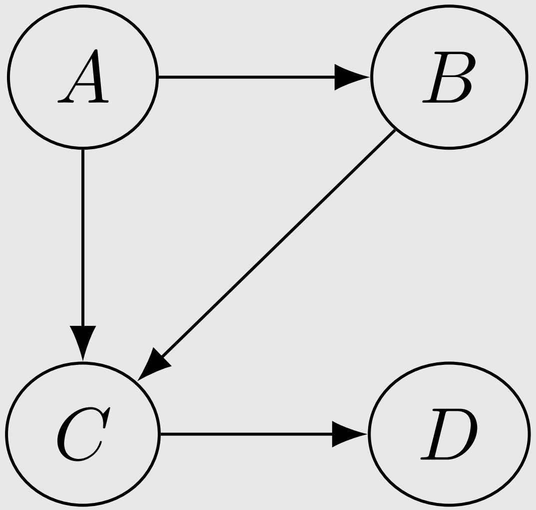
Graph terminology



Graph terminology

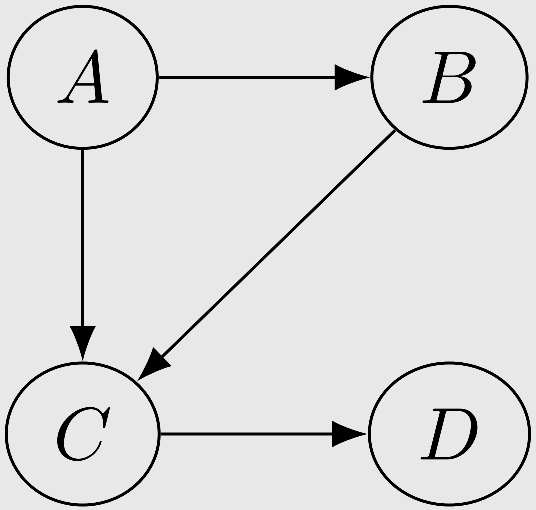


Graph terminology



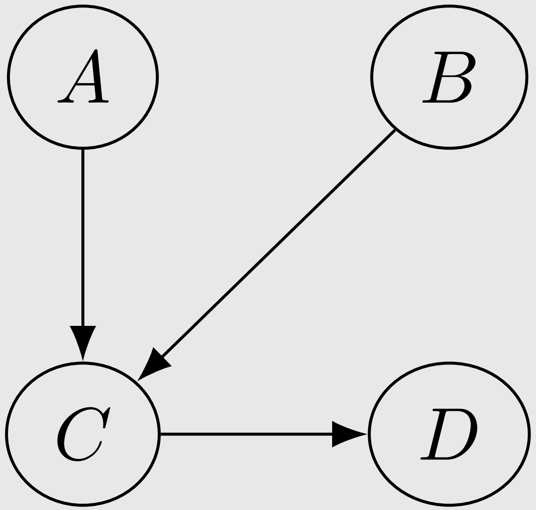
Graph terminology

Directed Acyclic Graph (DAG)



Graph terminology

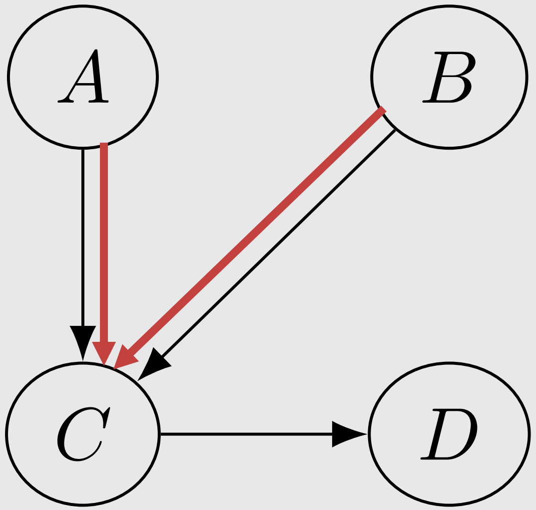
Directed Acyclic Graph (DAG)



Graph terminology

Directed Acyclic Graph (DAG)

Immorality



Naively modeling the joint distribution

Statistical modeling (no causality):

Naively modeling the joint distribution

Statistical modeling (no causality): $P(x_1, x_2, \ldots, x_n)$

Naively modeling the joint distribution

Statistical modeling (no causality): $P(x_1, x_2, \ldots, x_n) = P(x_1) \prod_i P(x_i \mid x_{i-1}, \ldots, x_1)$

Naively modeling the joint distribution

Statistical modeling (no causality): $P(x_1, x_2, \ldots, x_n) = P(x_1) \prod_i P(x_i \mid x_{i-1}, \ldots, x_1)$

$P(x_1, x_2, x_3, x_4) = P(x_1)\, P(x_2 \mid x_1)\, P(x_3 \mid x_2, x_1)\, P(x_4 \mid x_3, x_2, x_1)$

Naively modeling the joint distribution

Statistical modeling (no causality): $P(x_1, x_2, \ldots, x_n) = P(x_1) \prod_i P(x_i \mid x_{i-1}, \ldots, x_1)$

$$P(x_1, x_2, x_3, x_4) = P(x_1)\, P(x_2 \mid x_1)\, P(x_3 \mid x_2, x_1)\, P(x_4 \mid x_3, x_2, x_1)$$

x_1	x_2	x_3	$P(x_4 \mid x_3, x_2, x_1)$
0	0	0	α_1
0	0	1	α_2
0	1	0	α_3
0	1	1	α_4
1	0	0	α_5
1	0	1	α_6
1	1	0	α_7
1	1	1	α_8

Naively modeling the joint distribution

Statistical modeling (no causality): $P(x_1, x_2, \ldots, x_n) = P(x_1) \prod_i P(x_i \mid x_{i-1}, \ldots, x_1)$

$$P(x_1, x_2, x_3, x_4) = P(x_1)\, P(x_2 \mid x_1)\, P(x_3 \mid x_2, x_1)\, P(x_4 \mid x_3, x_2, x_1)$$

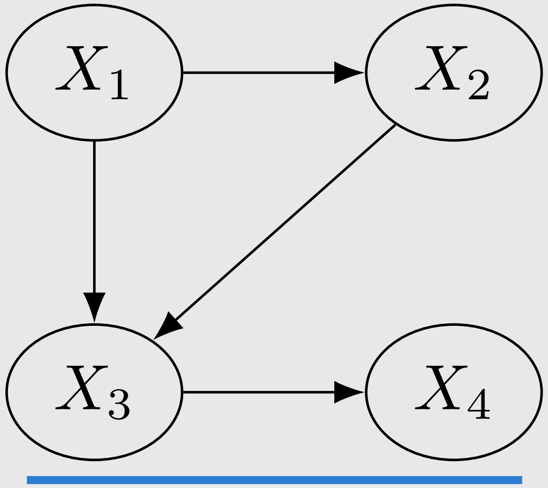
x_1	x_2	x_3	$P(x_4 \mid x_3, x_2, x_1)$
0	0	0	α_1
0	0	1	α_2
0	1	0	α_3
0	1	1	α_4
1	0	0	α_5
1	0	1	α_6
1	1	0	α_7
1	1	1	α_8

2^{n-1} parameters!

Naively modeling the joint distribution

Statistical modeling (no causality): $P(x_1, x_2, \ldots, x_n) = P(x_1) \prod_i P(x_i \mid x_{i-1}, \ldots, x_1)$

$$P(x_1, x_2, x_3, x_4) = P(x_1)\, P(x_2 \mid x_1)\, P(x_3 \mid x_2, x_1)\, P(x_4 \mid x_3, x_2, x_1)$$



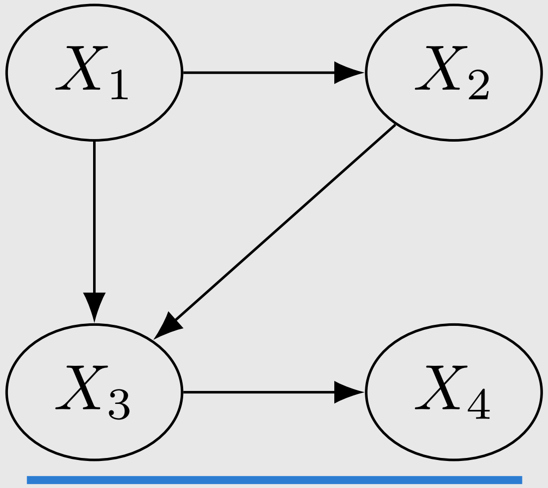
x_1	x_2	x_3	$P(x_4 \mid x_3, x_2, x_1)$
0	0	0	α_1
0	0	1	α_2
0	1	0	α_3
0	1	1	α_4
1	0	0	α_5
1	0	1	α_6
1	1	0	α_7
1	1	1	α_8

2^{n-1} parameters!

Naively modeling the joint distribution

Statistical modeling (no causality): $P(x_1, x_2, \ldots, x_n) = P(x_1) \prod_i P(x_i \mid x_{i-1}, \ldots, x_1)$

$$P(x_1, x_2, x_3, x_4) = P(x_1)\, P(x_2 \mid x_1)\, P(x_3 \mid x_2, x_1)\, P(x_4 \mid x_3, x_2, x_1)$$



x_1	x_2	x_3	$P(x_4 \mid x_3, x_2, x_1)$
0	0	0	α_1
0	0	1	α_2
0	1	0	α_3
0	1	1	α_4
1	0	0	α_5
1	0	1	α_6
1	1	0	α_7
1	1	1	α_8

2^{n-1} parameters!

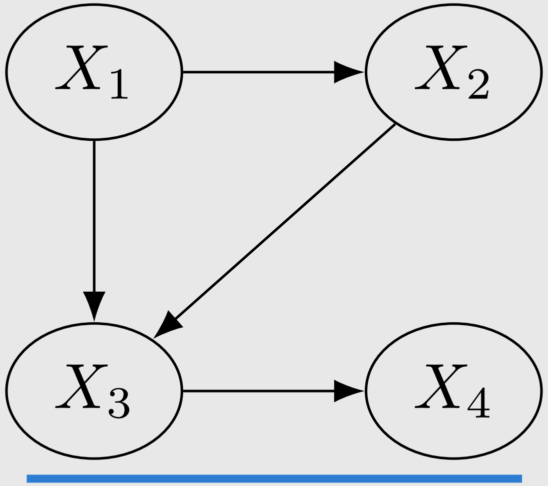
Local Markov assumption

Given its parents in the DAG, a node X is independent of all of its non-descendants.

Local Markov assumption

Given its parents in the DAG, a node X is independent of all of its non-descendants.

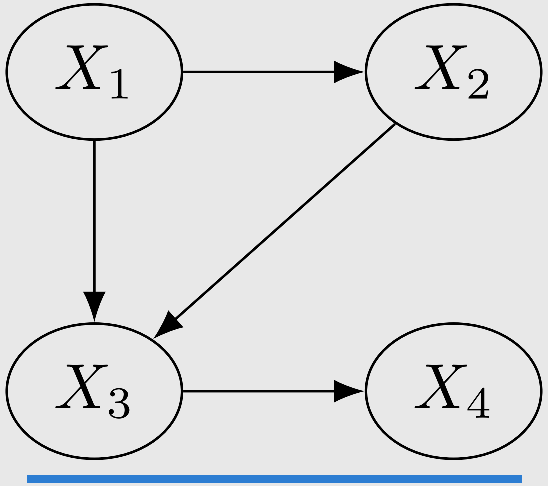
$$P(x_1, x_2, x_3, x_4) = P(x_1)\, P(x_2 \mid x_1)\, P(x_3 \mid x_2, x_1)\, P(x_4 \mid x_3, x_2, x_1)$$



Local Markov assumption

Given its parents in the DAG, a node X is independent of all of its non-descendants.

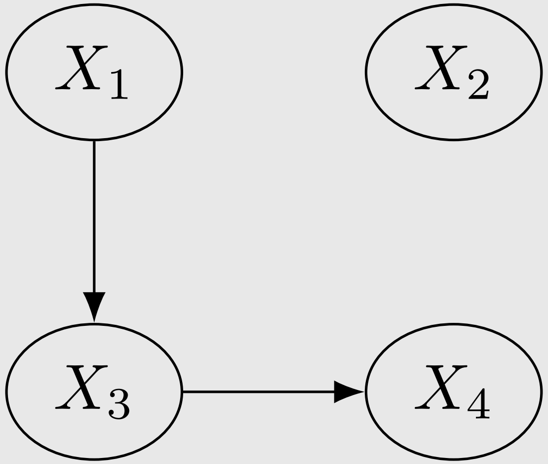
$$P(x_1, x_2, x_3, x_4) = P(x_1)\, P(x_2 \mid x_1)\, P(x_3 \mid x_2, x_1)\, \underline{P(x_4 \mid x_3)}$$



Local Markov assumption

Given its parents in the DAG, a node X is independent of all of its non-descendants.

$$P(x_1, x_2, x_3, x_4) = P(x_1)\, P(x_2 \mid x_1)\, P(x_3 \mid x_2, x_1)\, P(x_4 \mid x_3)$$

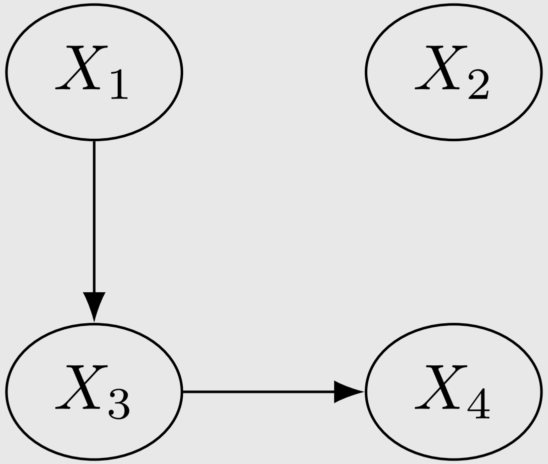


Local Markov assumption

Given its parents in the DAG, a node X is independent of all of its non-descendants.

$$P(x_1, x_2, x_3, x_4) = P(x_1)\, P(x_2 \mid x_1)\, P(x_3 \mid x_2, x_1)\, P(x_4 \mid x_3)$$

Question:
How will the factorization change now?

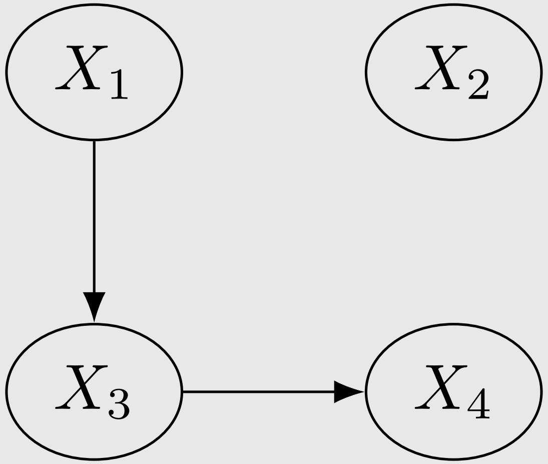


Local Markov assumption

Given its parents in the DAG, a node X is independent of all of its non-descendants.

$$P(x_1, x_2, x_3, x_4) = P(x_1)\, P(x_2)\, P(x_3 \mid x_1)\, P(x_4 \mid x_3)$$

Question:
How will the factorization change now?



Bayesian network factorization

Bayesian network factorization

$$P(x_1, \ldots, x_n) = \prod_i P(x_i \mid \mathrm{pa}_i)$$

Bayesian network factorization

$$P(x_1, \ldots, x_n) = \prod_i P(x_i \mid \text{pa}_i)$$

local Markov assumption $\implies$ Bayesian network factorization

Bayesian network factorization

$$P(x_1, \ldots, x_n) = \prod_i P(x_i \mid \mathrm{pa}_i)$$

local Markov assumption $\implies$ Bayesian network factorization

local Markov assumption $\impliedby$ Bayesian network factorization

Bayesian network factorization

$$P(x_1, \ldots, x_n) = \prod_i P(x_i \mid \text{pa}_i)$$

local Markov assumption $\implies$ Bayesian network factorization

local Markov assumption $\impliedby$ Bayesian network factorization

See Chapter 3 of Koller & Friedman (2009) book for proofs

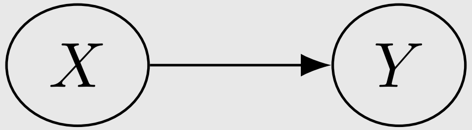
Minimality assumption

Minimality assumption

1. Given its parents in the DAG, a node X is independent of all its non-descendants (local Markov assumption).

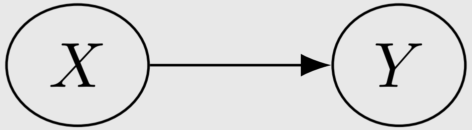
Minimality assumption

1. Given its parents in the DAG, a node X is independent of all its non-descendants (local Markov assumption).



Minimality assumption

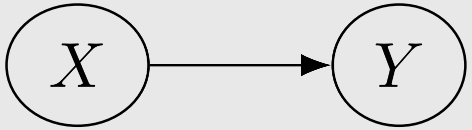
1. Given its parents in the DAG, a node X is independent of all its non-descendants (local Markov assumption).



Permits distributions where $P(x, y) = P(x)\, P(y \mid x)$

Minimality assumption

1. Given its parents in the DAG, a node X is independent of all its non-descendants (local Markov assumption).

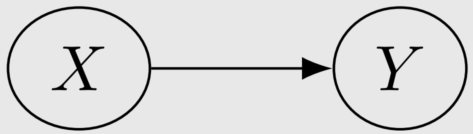


Permits distributions where $P(x, y) = P(x)\, P(y \mid x)$ and also where

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

Minimality assumption

1. Given its parents in the DAG, a node X is independent of all its non-descendants (local Markov assumption).
2. Adjacent nodes in the DAG are dependent.

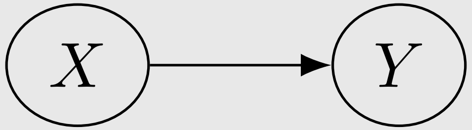


Permits distributions where $P(x, y) = P(x)\, P(y \mid x)$ and also where

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

Minimality assumption

1. Given its parents in the DAG, a node X is independent of all its non-descendants (local Markov assumption).
2. Adjacent nodes in the DAG are dependent.

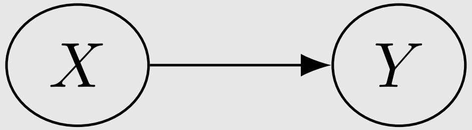


Permits distributions where $P(x, y) = P(x)\, P(y \mid x)$ ~~and also where~~

~~$P(x, y) = P(x)\, P(y)$~~

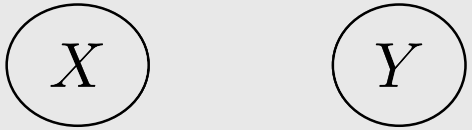
Minimality assumption

1. Given its parents in the DAG, a node X is independent of all its non-descendants (local Markov assumption).
2. Adjacent nodes in the DAG are dependent.

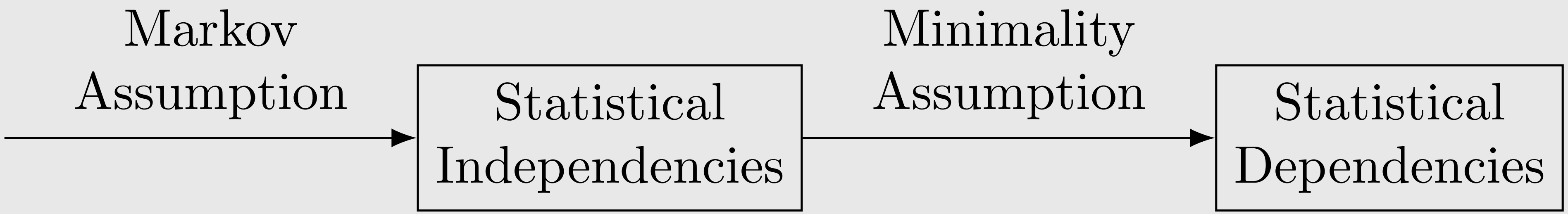


Permits distributions where $P(x, y) = P(x)\, P(y \mid x)$ ~~and also where~~

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



Assumptions flowchart



Recall:

1. How is the local Markov assumption related to the Bayesian network factorization?
2. What are the two parts of the minimality assumption? What do we gain with the second part?

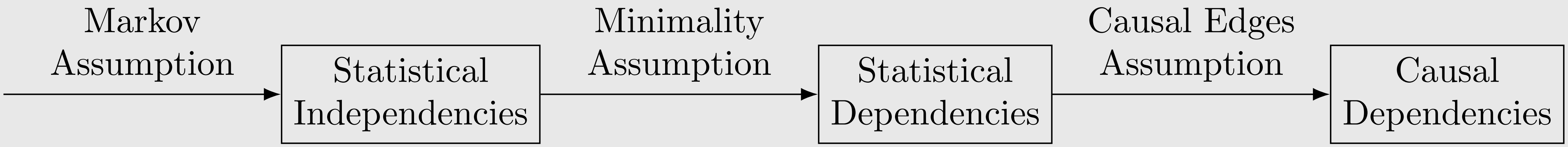
What is a cause?

A variable X is said to be a cause of a variable Y if Y can change in response to changes in X.

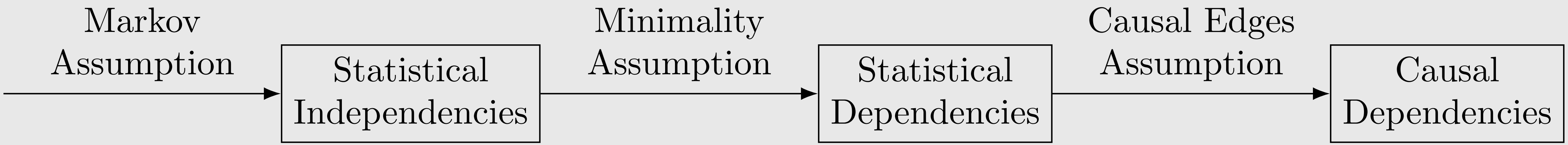
Causal edges assumption

In a directed graph, every parent is a direct cause of all its children.

Assumptions flowchart



Assumptions flowchart



DAG + Two assumptions to give us flow of association and causation in graphs:

1. Markov Assumption
2. Causal Edges Assumption

Graph terminology

Bayesian networks and causal graphs

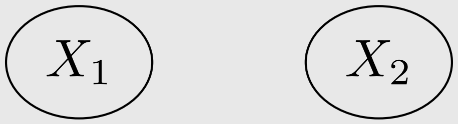
The basic building blocks of graphs

The flow of association and causation

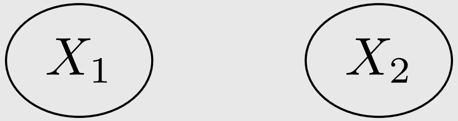
Graphical building blocks

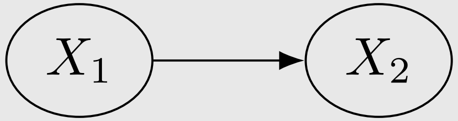
Graphical building blocks

Two nodes:

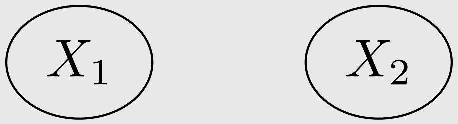


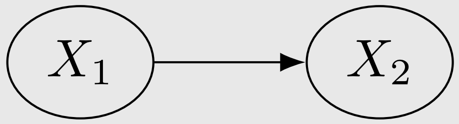
Graphical building blocks

Two nodes: 

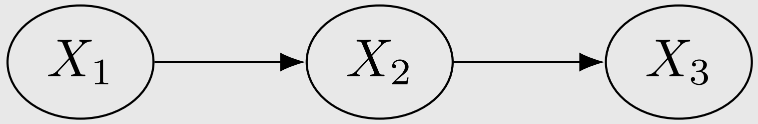
or 

Graphical building blocks

Two nodes: 

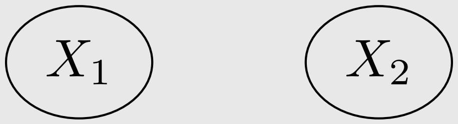
or 

Chain

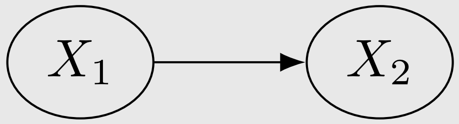


Graphical building blocks

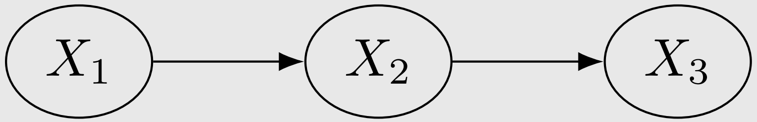
Two nodes:



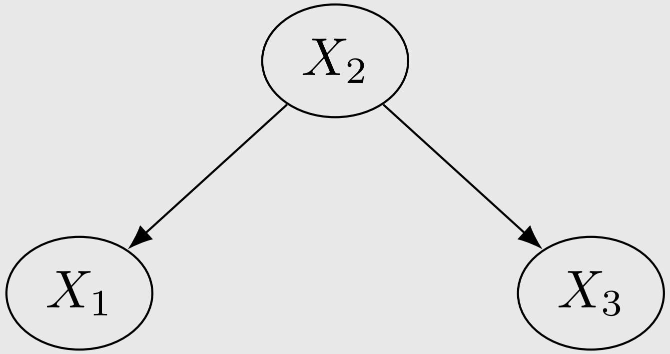
or



Chain

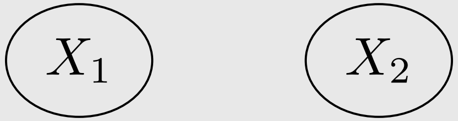


Fork

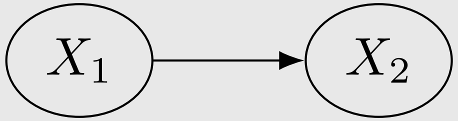


Graphical building blocks

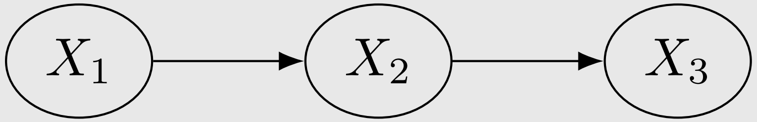
Two nodes:



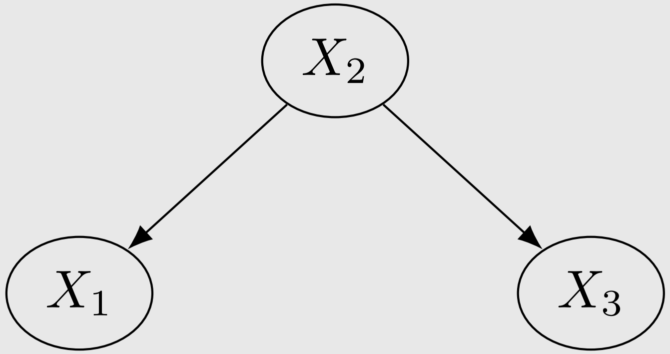
or



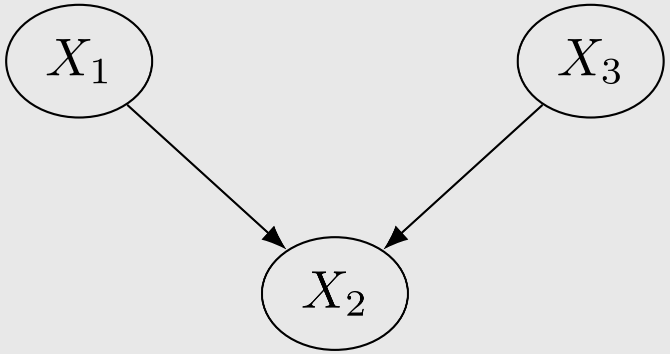
Chain



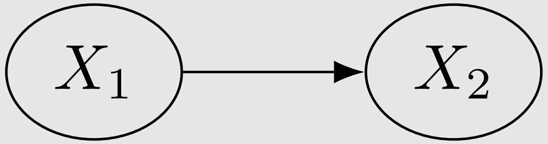
Fork



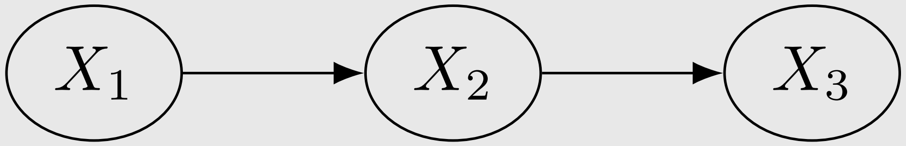
Immorality



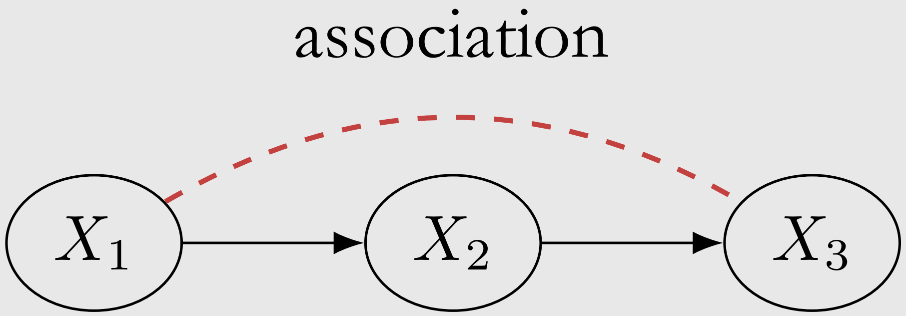
Question:
What assumption tells us that X_1 and X_2 are associated, given the following graph?



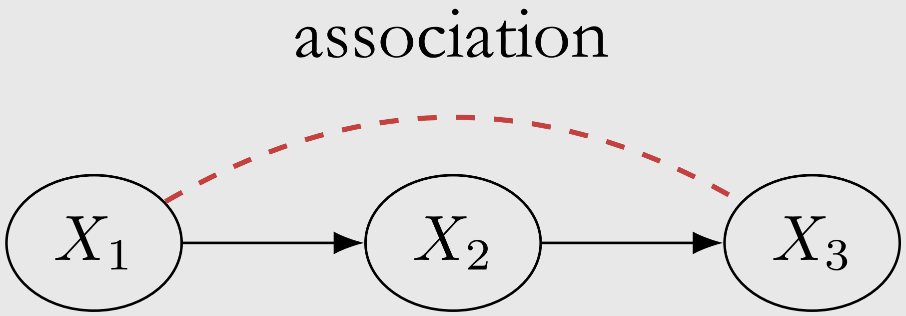
Chains and forks: dependence

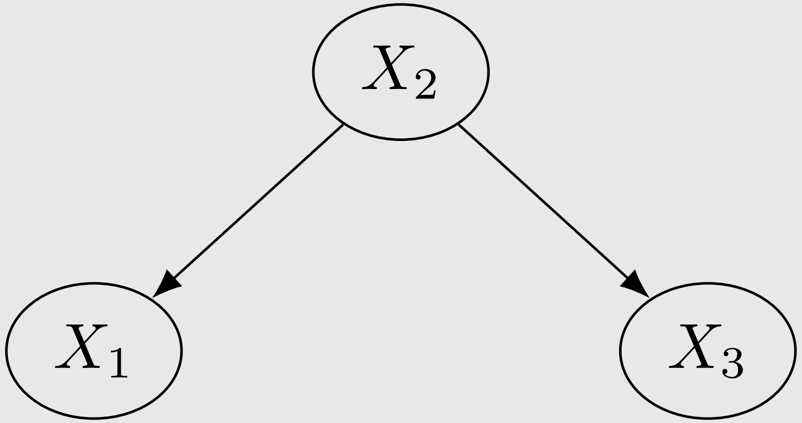


Chains and forks: dependence

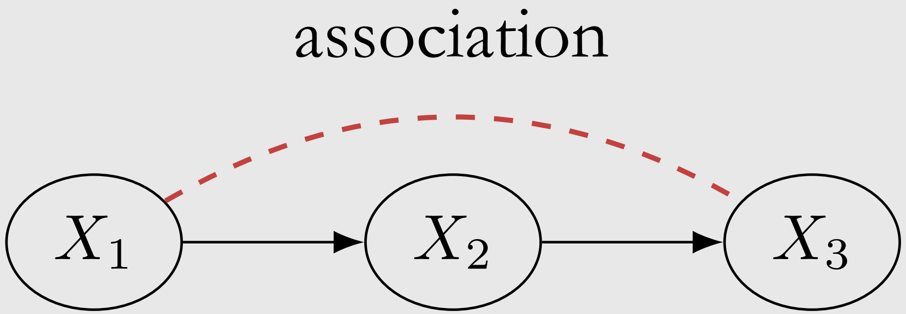


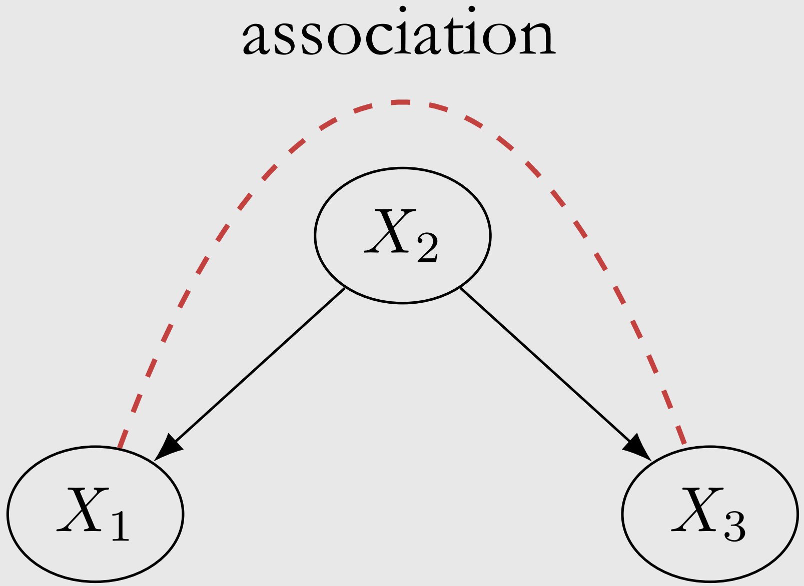
Chains and forks: dependence



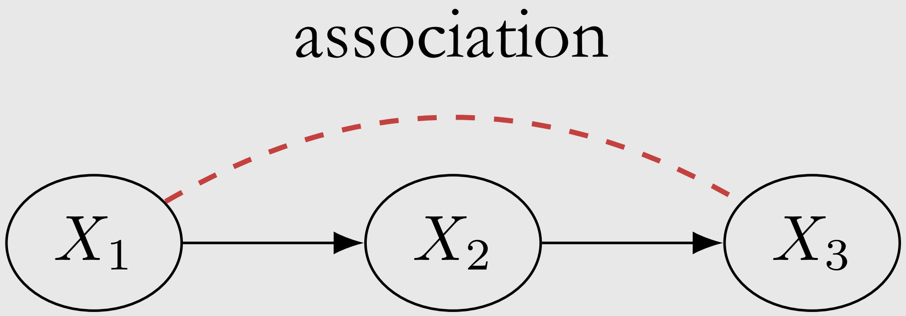


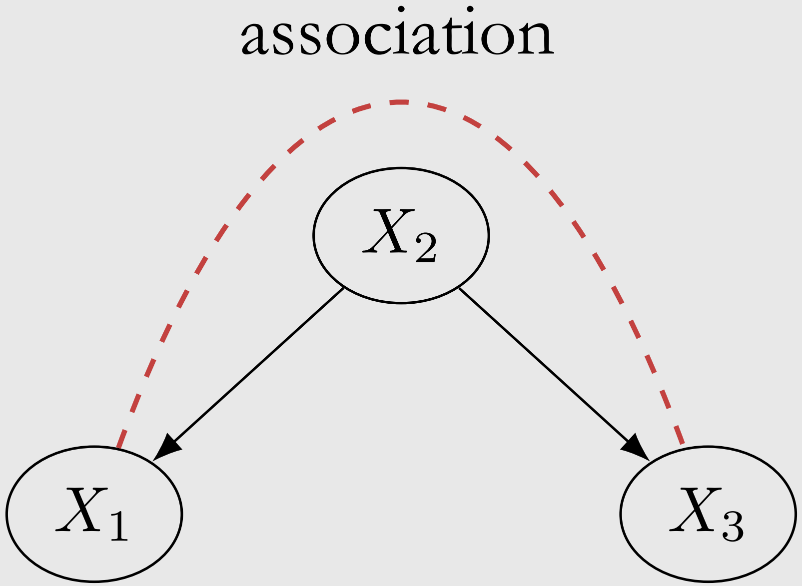
Chains and forks: dependence



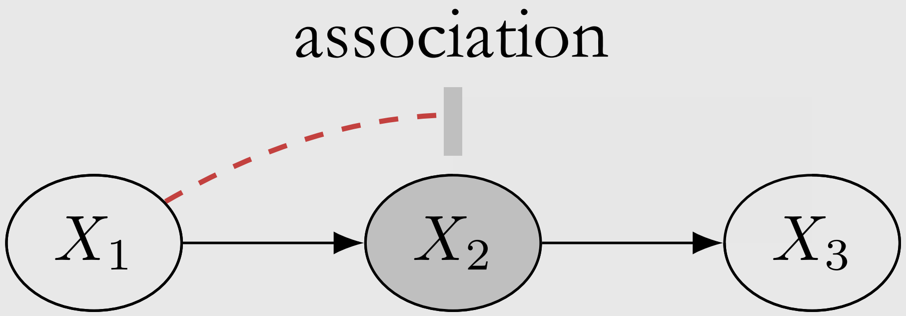


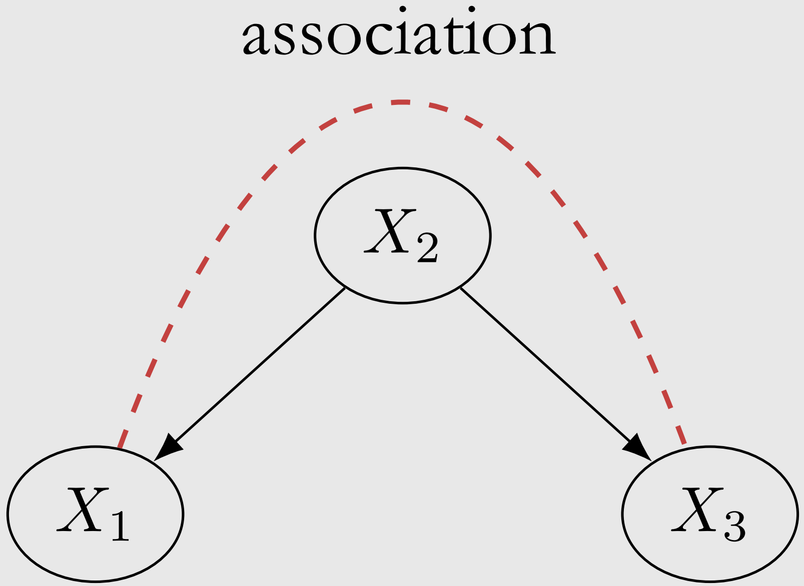
Chains and forks: independence



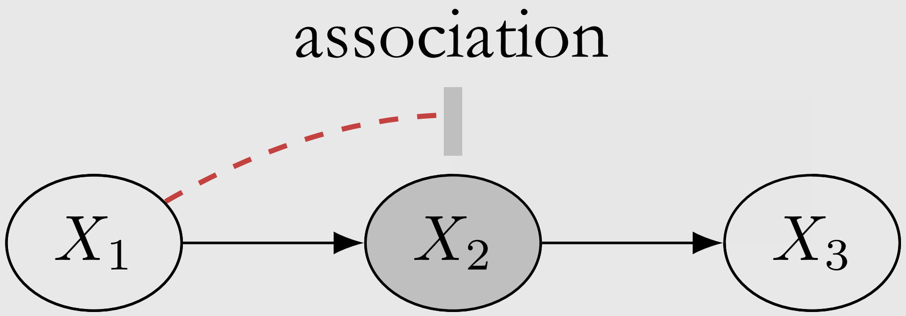


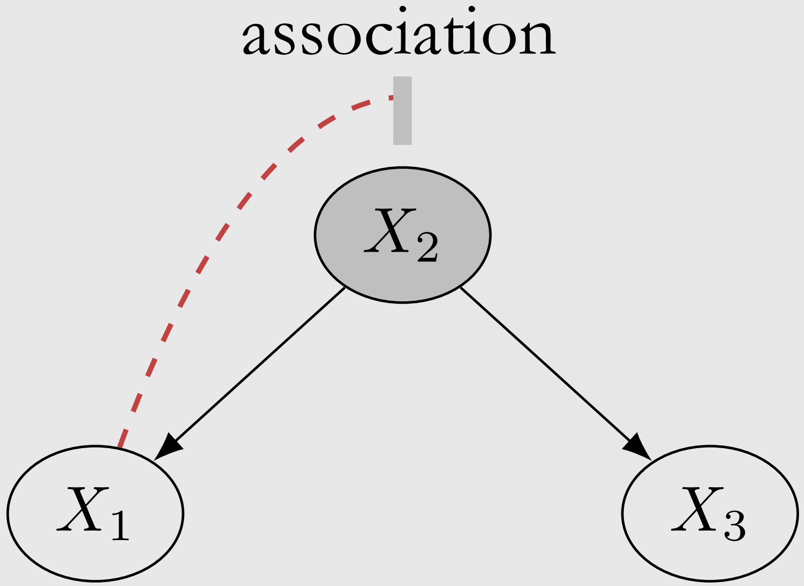
Chains and forks: independence



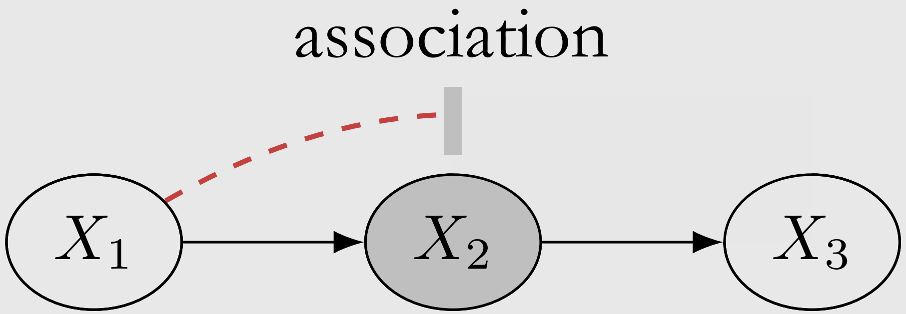


Chains and forks: independence

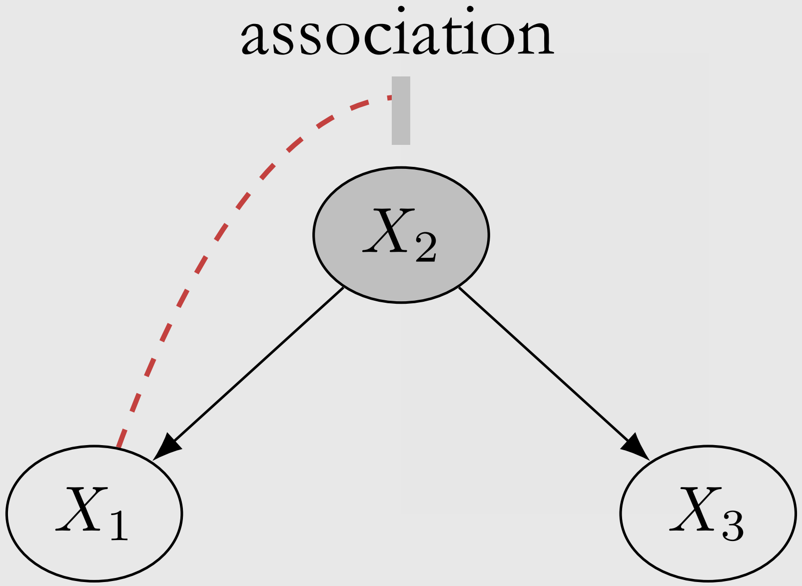




Chains and forks: independence

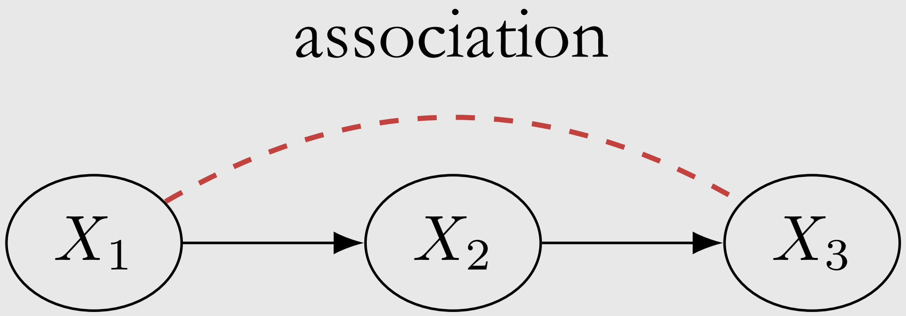


blocked path

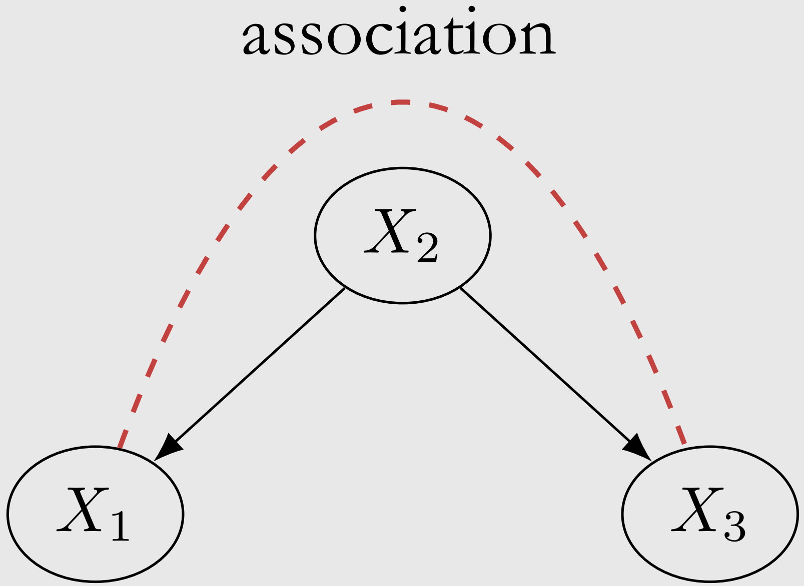


blocked path

Chains and forks: independence

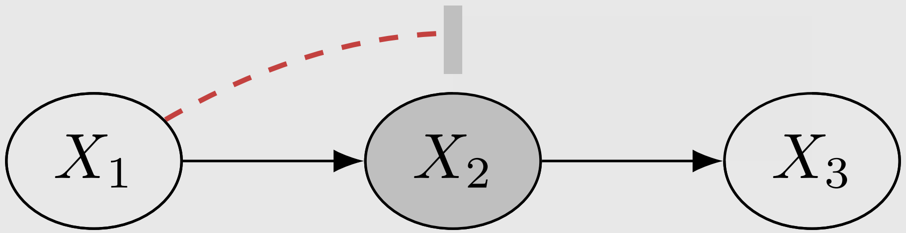


unblocked path



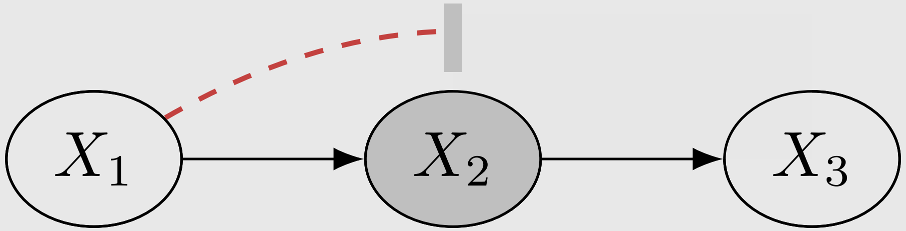
unblocked path

Proof of conditional independence in chains



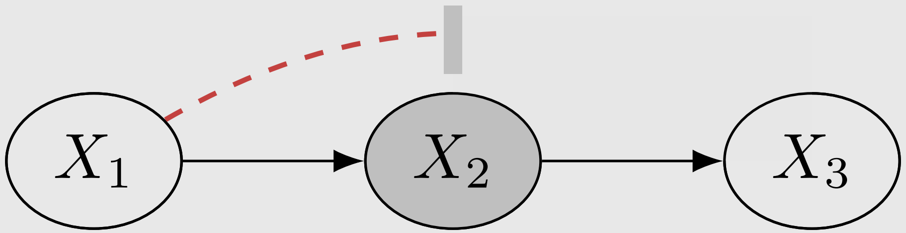
Proof of conditional independence in chains

Goal: show $P(x_1, x_3 \mid x_2) = P(x_1 \mid x_2)\, P(x_3 \mid x_2)$



Proof of conditional independence in chains

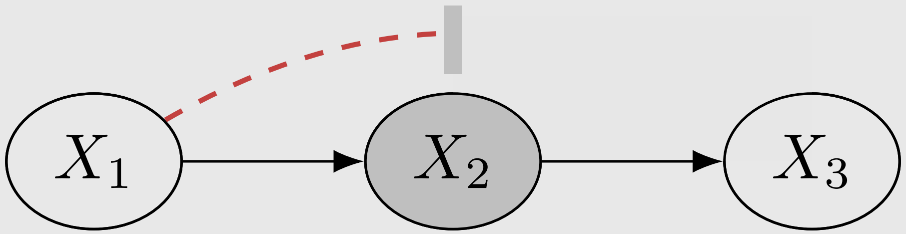
Goal: show $P(x_1, x_3 \mid x_2) = P(x_1 \mid x_2)\, P(x_3 \mid x_2)$



1. Bayesian network factorization:

Proof of conditional independence in chains

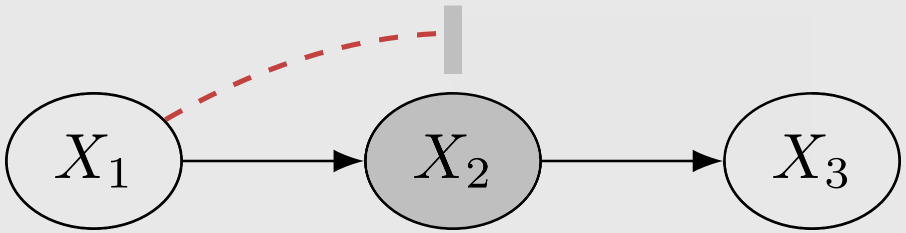
Goal: show $P(x_1, x_3 \mid x_2) = P(x_1 \mid x_2)\, P(x_3 \mid x_2)$



1. Bayesian network factorization: $P(x_1, x_2, x_3) = P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)$

Proof of conditional independence in chains

Goal: show $P(x_1, x_3 \mid x_2) = P(x_1 \mid x_2)\, P(x_3 \mid x_2)$

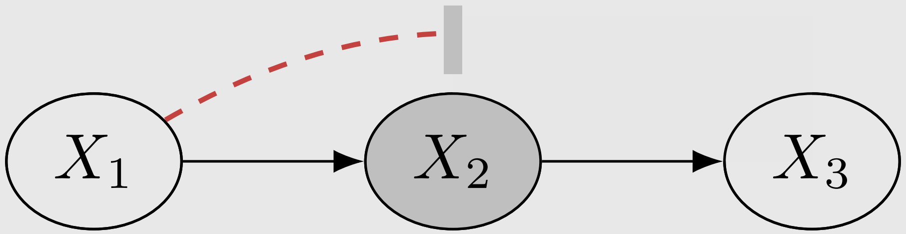


1. Bayesian network factorization: $P(x_1, x_2, x_3) = P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)$

2. Apply Bayes' rule: $$P(x_1, x_3 \mid x_2) = \frac{P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)}{P(x_2)}$$

Proof of conditional independence in chains

Goal: show $P(x_1, x_3 \mid x_2) = P(x_1 \mid x_2)\, P(x_3 \mid x_2)$

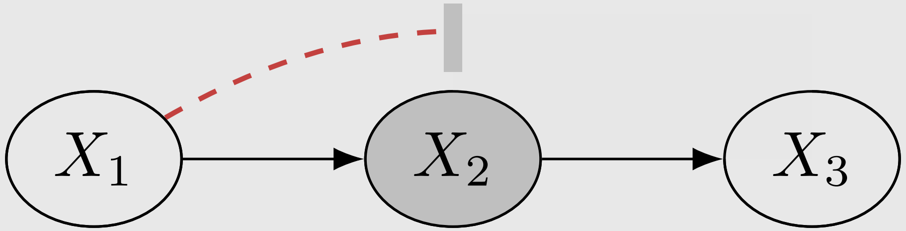


1. Bayesian network factorization: $P(x_1, x_2, x_3) = P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)$

2. Apply Bayes' rule: $$P(x_1, x_3 \mid x_2) = \frac{P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)}{P(x_2)}$$

Proof of conditional independence in chains

Goal: show $P(x_1, x_3 \mid x_2) = P(x_1 \mid x_2)\, P(x_3 \mid x_2)$

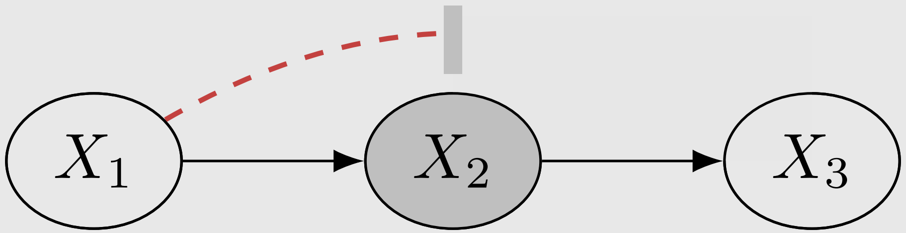


1. Bayesian network factorization: $P(x_1, x_2, x_3) = P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)$

2. Apply Bayes' rule: $$P(x_1, x_3 \mid x_2) = \frac{P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)}{P(x_2)}$$

Proof of conditional independence in chains

Goal: show $P(x_1, x_3 \mid x_2) = \boxed{P(x_1 \mid x_2)} P(x_3 \mid x_2)$



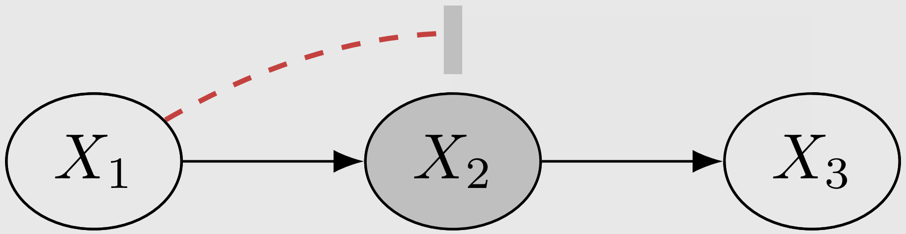
1. Bayesian network factorization: $P(x_1, x_2, x_3) = P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)$

2. Apply Bayes' rule: $P(x_1, x_3 \mid x_2) = \dfrac{P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)}{P(x_2)}$

3. Apply Bayes' rule again:

Proof of conditional independence in chains

Goal: show $P(x_1, x_3 \mid x_2) = \boxed{P(x_1 \mid x_2)} P(x_3 \mid x_2)$



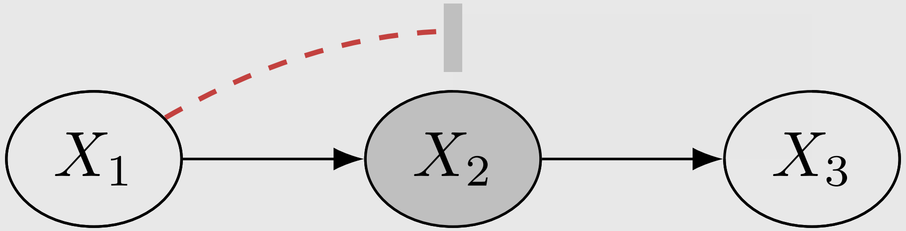
1. Bayesian network factorization: $P(x_1, x_2, x_3) = P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)$

2. Apply Bayes' rule: $$P(x_1, x_3 \mid x_2) = \frac{P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)}{P(x_2)}$$

3. Apply Bayes' rule again: $$P(x_1, x_3 \mid x_2) = \frac{P(x_1, x_2)}{P(x_2)}\, P(x_3|x_2)$$

Proof of conditional independence in chains

Goal: show $P(x_1, x_3 \mid x_2) = \boxed{P(x_1 \mid x_2)} P(x_3 \mid x_2)$



1. Bayesian network factorization: $P(x_1, x_2, x_3) = P(x_1)\, P(x_2|x_1)\, P(x_3|x_2)$

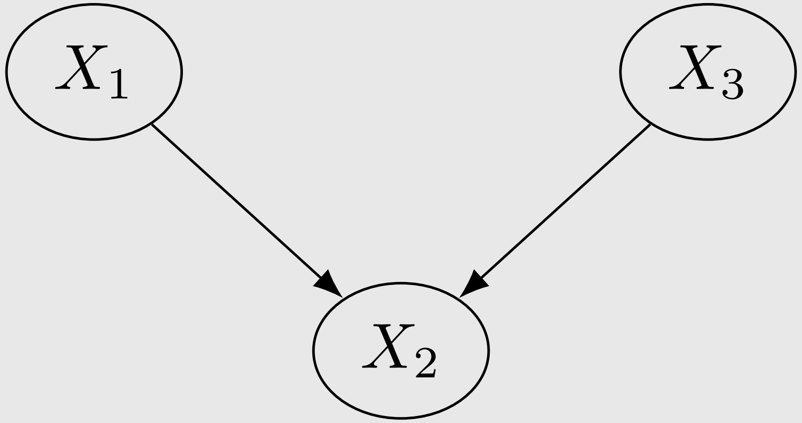
2. Apply Bayes' rule: $$P(x_1, x_3 \mid x_2) = \frac{\boxed{P(x_1)\, P(x_2|x_1)}\, P(x_3|x_2)}{\boxed{P(x_2)}}$$

3. Apply Bayes' rule again: $$P(x_1, x_3 \mid x_2) = \frac{P(x_1, x_2)}{P(x_2)}\, P(x_3|x_2)$$ $$= P(x_1|x_2)\, P(x_3|x_2)$$

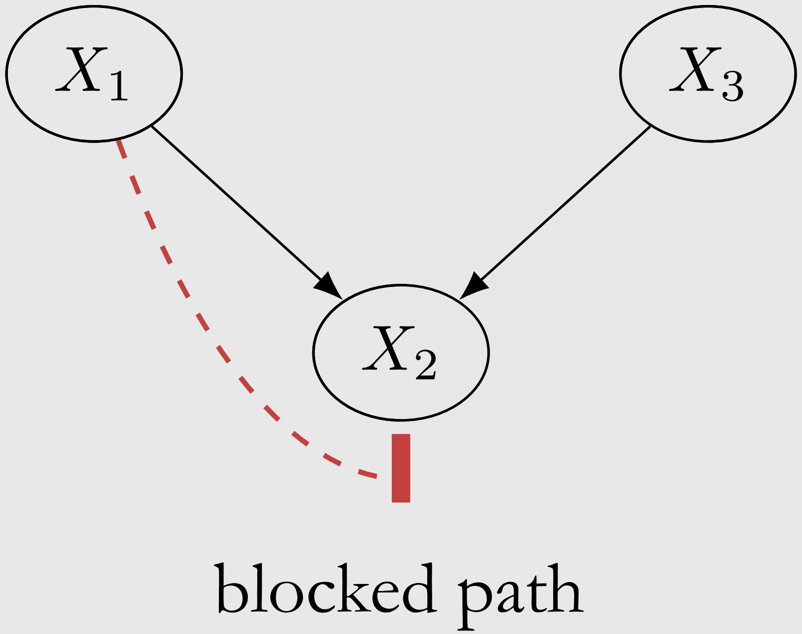
Proof of conditional independence in forks

Your turn ☺

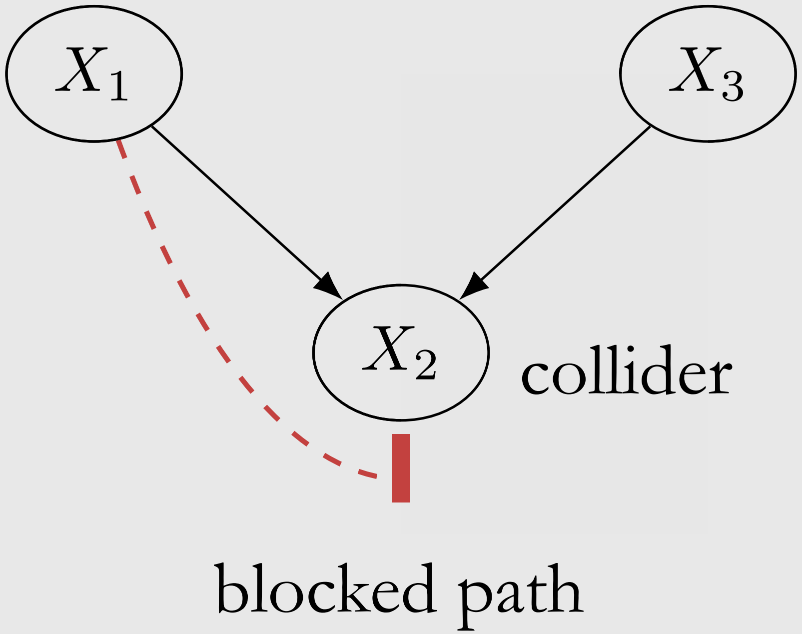
Immoralities



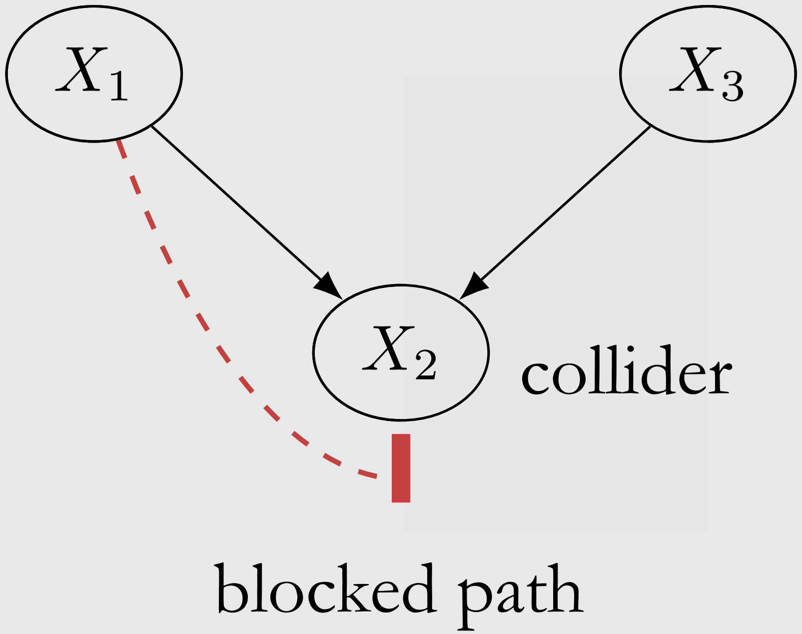
Immoralities



Immoralities

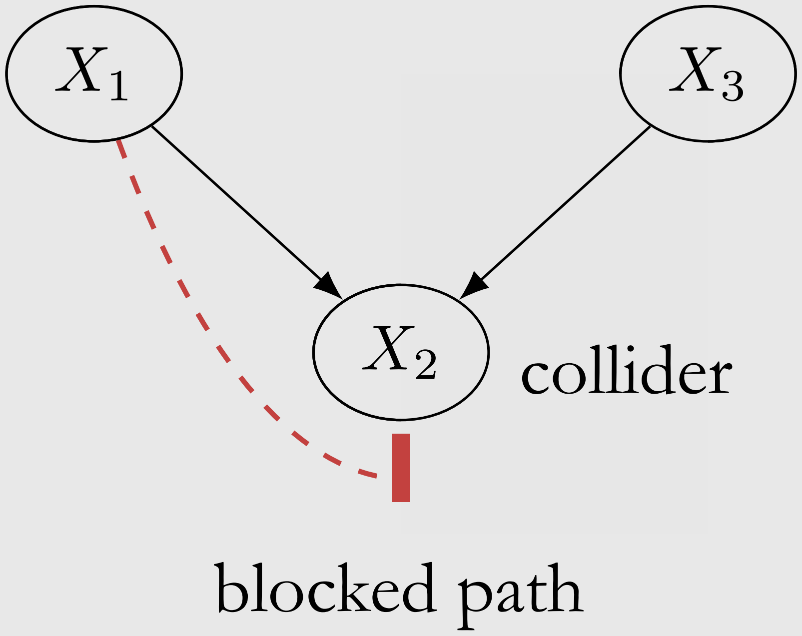


Immoralities



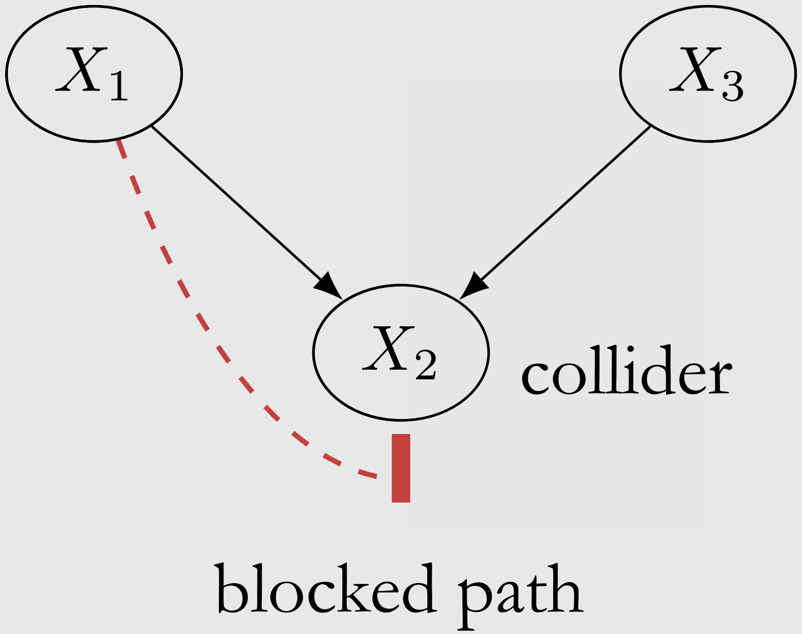
$$P(x_1, x_3) = \sum_{x_2} P(x_1, x_2, x_3)$$

Immoralities



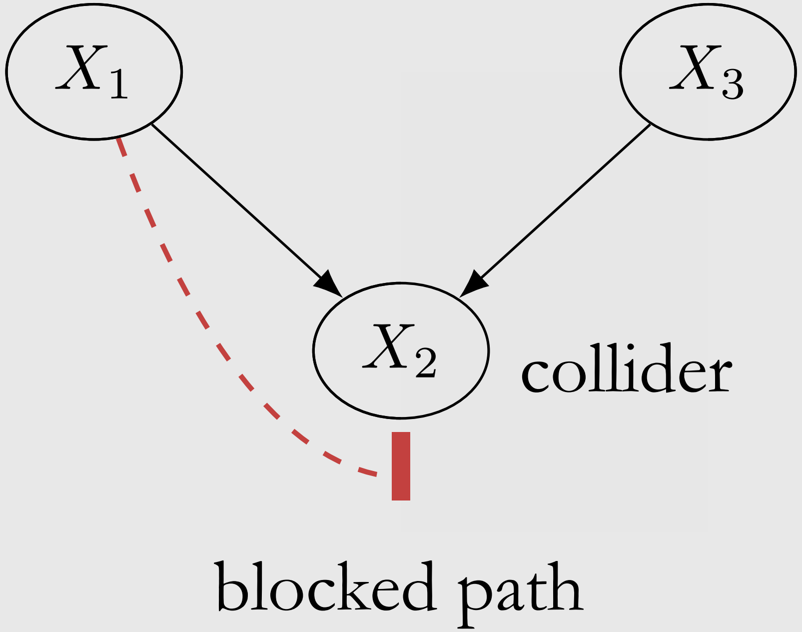
$$P(x_1, x_3) = \sum_{x_2} P(x_1, x_2, x_3)$$
$$= \sum_{x_2} P(x_1)\, P(x_3)\, P(x_2 \mid x_1, x_3)$$

Immoralities



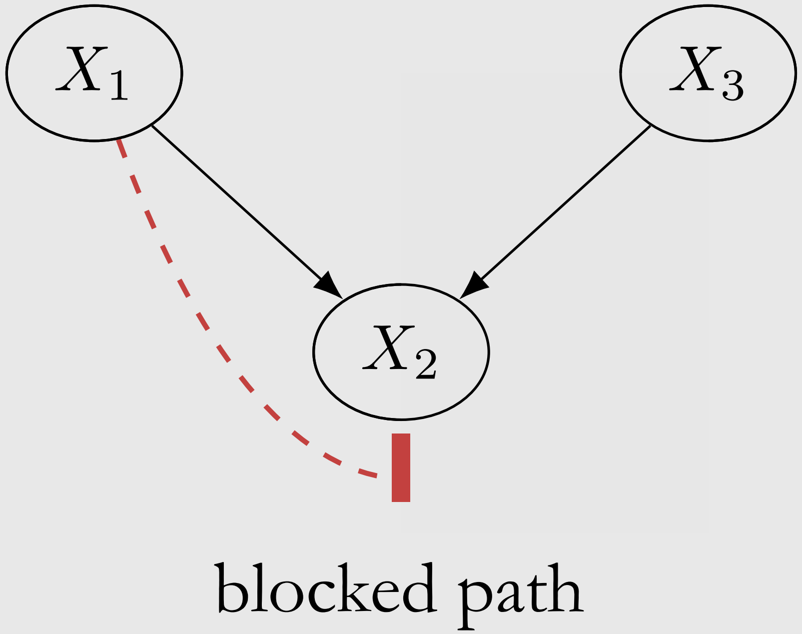
$$P(x_1, x_3) = \sum_{x_2} P(x_1, x_2, x_3)$$
$$= \sum_{x_2} P(x_1)\, P(x_3)\, P(x_2 \mid x_1, x_3)$$
$$= P(x_1)\, P(x_3) \sum_{x_2} P(x_2 \mid x_1, x_3)$$

Immoralities

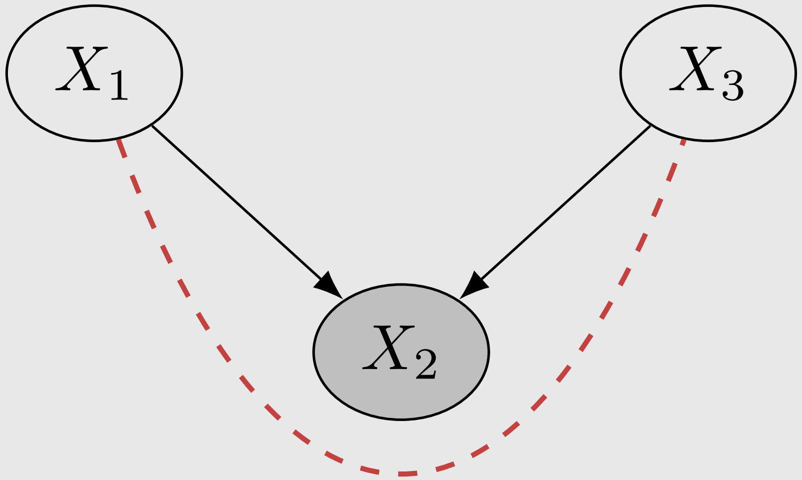


$$P(x_1, x_3) = \sum_{x_2} P(x_1, x_2, x_3)$$
$$= \sum_{x_2} P(x_1)\, P(x_3)\, P(x_2 \mid x_1, x_3)$$
$$= P(x_1)\, P(x_3) \sum_{x_2} P(x_2 \mid x_1, x_3)$$
$$= P(x_1)\, P(x_3)$$

Immoralities: conditioning on the collider

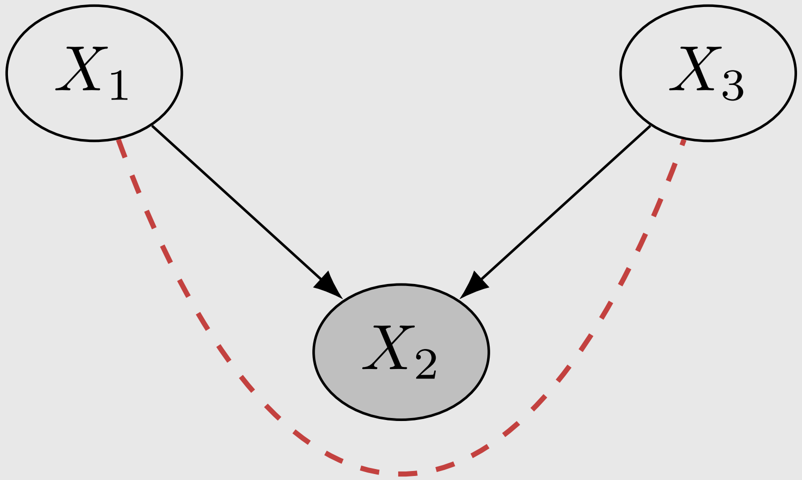


Immoralities: conditioning on the collider



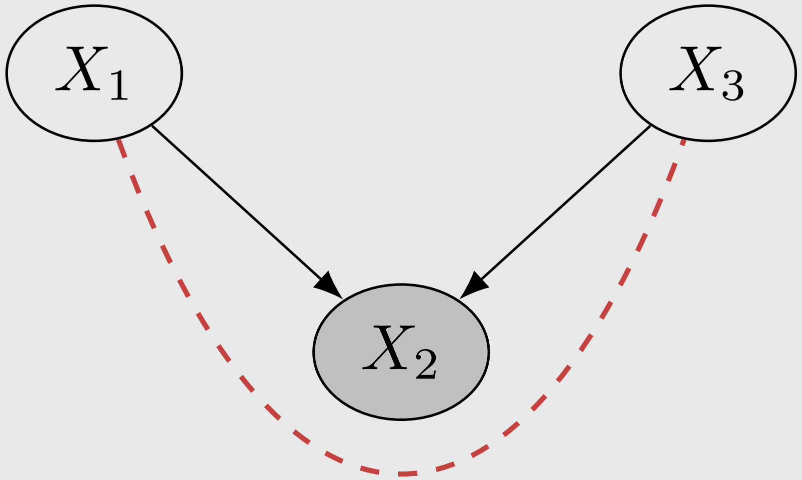
unblocked path

Example: good-looking men are jerks



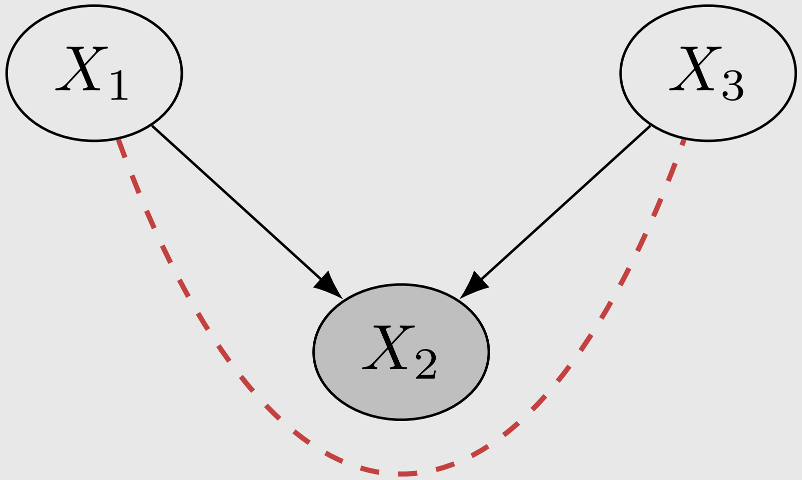
Example: good-looking men are jerks

$$X_1 = \begin{cases} 1 & \text{good-looking} \\ 0 & \text{otherwise} \end{cases}$$



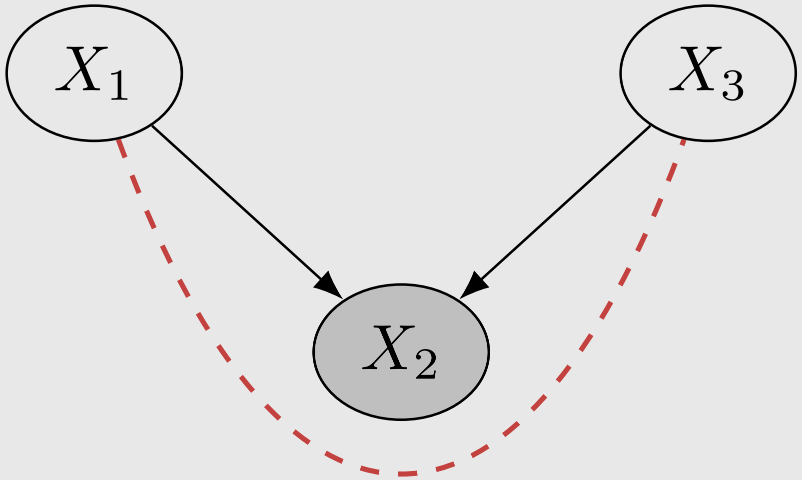
Example: good-looking men are jerks

$$X_1 = \begin{cases} 1 & \text{good-looking} \\ 0 & \text{otherwise} \end{cases} \qquad X_3 = \begin{cases} 1 & \text{kind} \\ 0 & \text{jerk} \end{cases}$$



Example: good-looking men are jerks

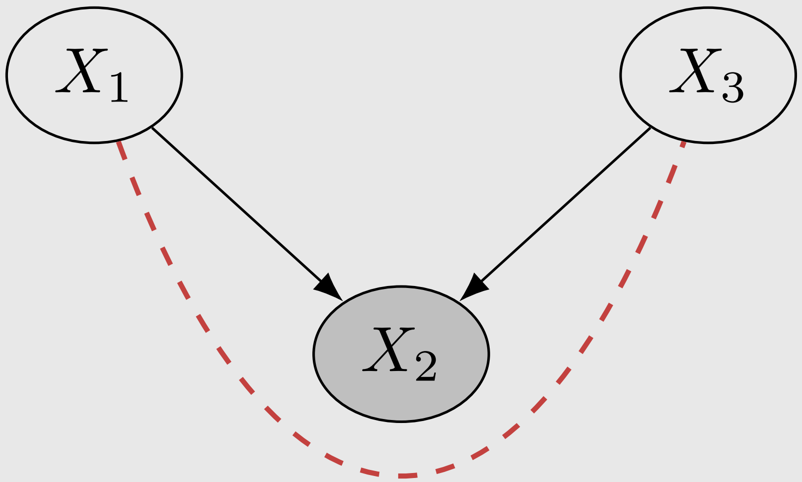
$$X_1 = \begin{cases} 1 & \text{good-looking} \\ 0 & \text{otherwise} \end{cases} \qquad X_3 = \begin{cases} 1 & \text{kind} \\ 0 & \text{jerk} \end{cases}$$



$$X_2 = \begin{cases} 1 & \text{in relationship} \\ 0 & \text{not in relationship} \end{cases}$$

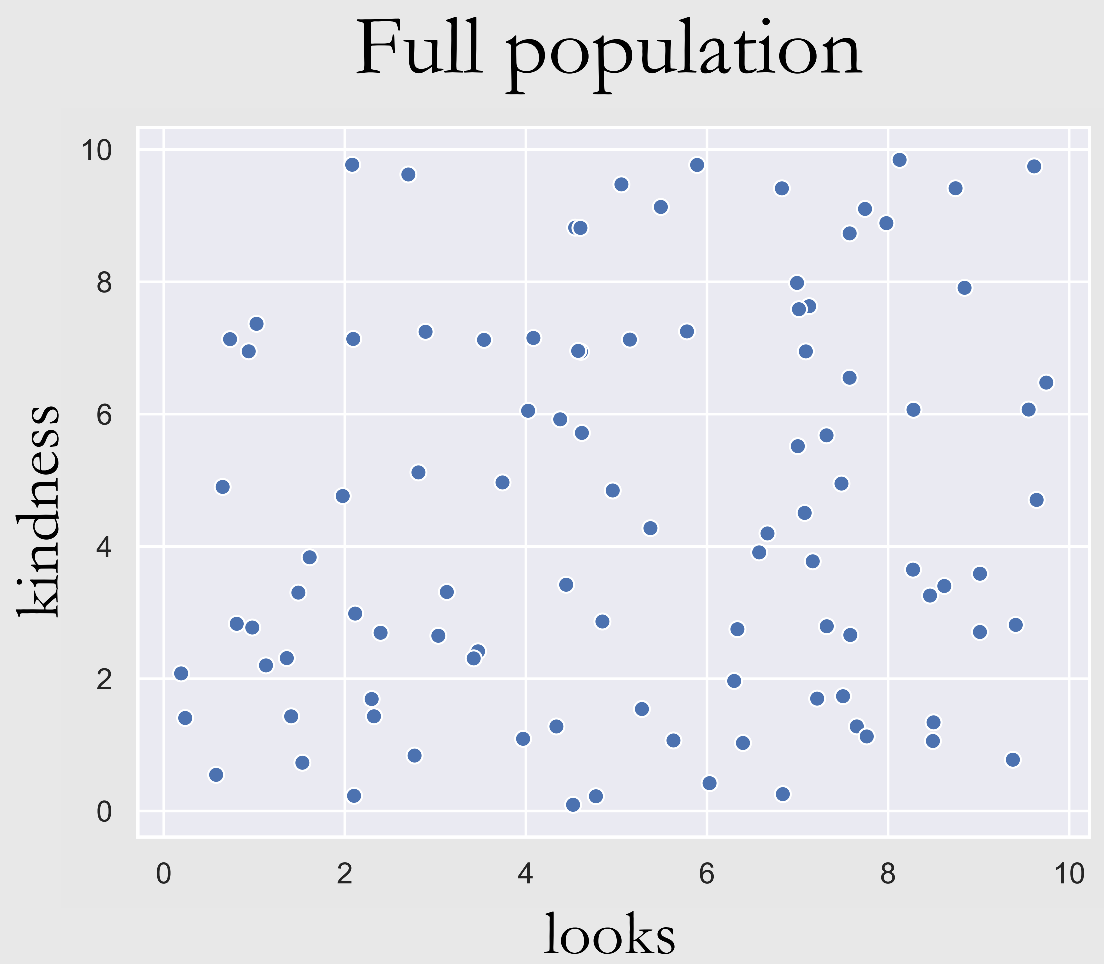
Example: good-looking men are jerks

$$X_1 = \begin{cases} 1 & \text{good-looking} \\ 0 & \text{otherwise} \end{cases} \qquad X_3 = \begin{cases} 1 & \text{kind} \\ 0 & \text{jerk} \end{cases}$$



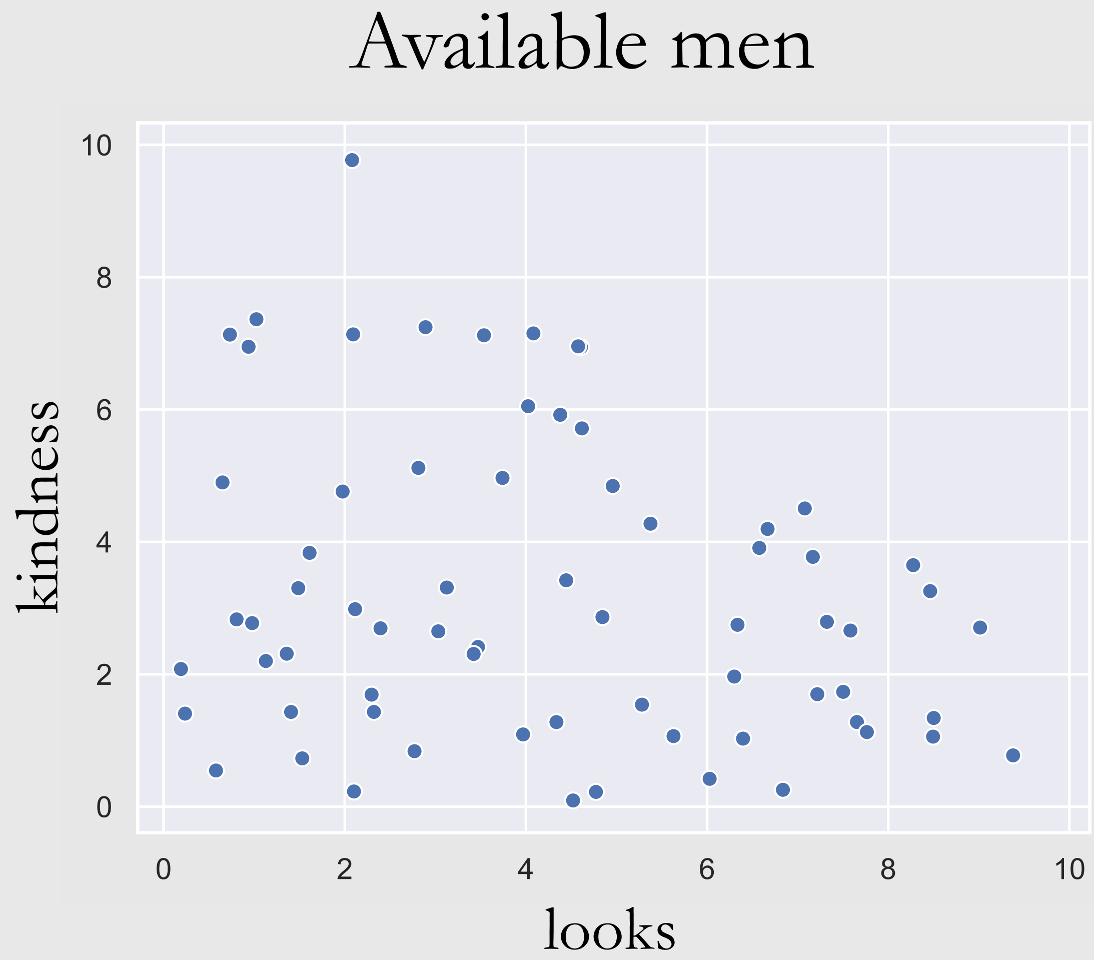
$$X_2 = X_1 \text{ AND } X_3 = \begin{cases} 1 & \text{in relationship} \\ 0 & \text{not in relationship} \end{cases}$$

Good-looking men are jerks scatterplot

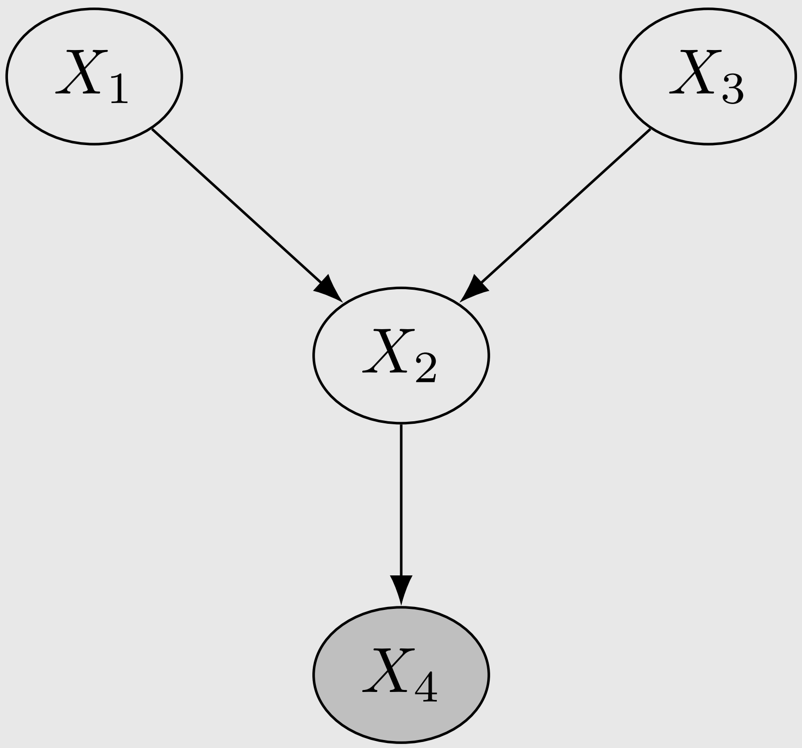


Good-looking men are jerks scatterplot

Good-looking men are jerks scatterplot

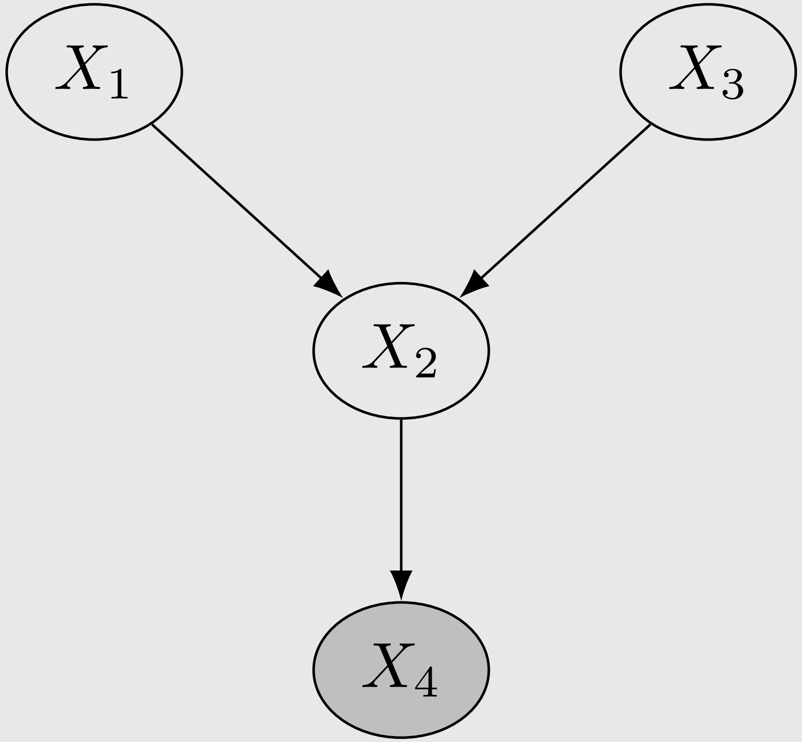


Conditioning on descendants of colliders



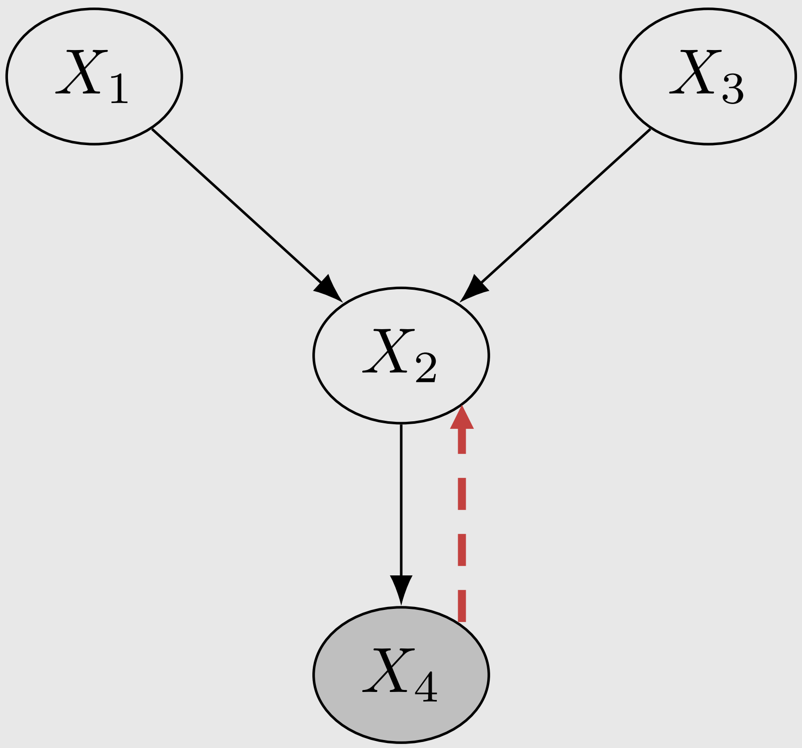
Conditioning on descendants of colliders

$$X_1 \not\perp\!\!\!\perp X_3 \mid X_4$$



Conditioning on descendants of colliders

$$X_1 \not\perp\!\!\!\perp X_3 \mid X_4$$



Question:
In the three different kinds of three-node graphs, what can block a path?

Graph terminology

Bayesian networks and causal graphs

The basic building blocks of graphs

The flow of association and causation

Blocked path definition

Blocked path definition

A path between nodes X and Y is blocked by a (potentially empty) conditioning set Z if either of the following is true:

Blocked path definition

A path between nodes X and Y is blocked by a (potentially empty) conditioning set Z if either of the following is true:

1. Along the path, there is a chain $\cdots \to W \to \cdots$ or a fork $\cdots \leftarrow W \to \cdots$ where W is conditioned on $(W \in Z)$.

Blocked path definition

A path between nodes X and Y is blocked by a (potentially empty) conditioning set Z if either of the following is true:

1. Along the path, there is a chain $\cdots \to W \to \cdots$ or a fork $\cdots \leftarrow W \to \cdots$ where W is conditioned on ($W \in Z$).
2. There is a collider W on the path that is not conditioned on ($W \notin Z$) and none of its descendants are conditioned on ($\text{de}(W) \nsubseteq Z$).

Blocked path definition

A path between nodes X and Y is blocked by a (potentially empty) conditioning set Z if either of the following is true:

1. Along the path, there is a chain $\cdots \rightarrow W \rightarrow \cdots$ or a fork $\cdots \leftarrow W \rightarrow \cdots$ where W is conditioned on ($W \in Z$).
2. There is a collider W on the path that is not conditioned on ($W \notin Z$) and none of its descendants are conditioned on ($\text{de}(W) \not\subseteq Z$).

Unblocked path: a path that is not blocked

d-separation

d-separation

Two (sets of) nodes X and Y are d-separated by a set of nodes Z if all of the paths between (any node in) X and (any node in) Y are blocked by Z.

d-separation

Two (sets of) nodes X and Y are d-separated by a set of nodes Z if all of the paths between (any node in) X and (any node in) Y are blocked by Z.

Theorem: Given that P is Markov with respect to G,

$$X \perp\!\!\!\perp_G Y \mid Z \implies X \perp\!\!\!\perp_P Y \mid Z$$

d-separation

Two (sets of) nodes X and Y are d-separated by a set of nodes Z if all of the paths between (any node in) X and (any node in) Y are blocked by Z.

Theorem: Given that P is Markov with respect to G,

$$X \perp\!\!\!\perp_G Y \mid Z \implies X \perp\!\!\!\perp_P Y \mid Z$$

global Markov assumption

d-separation

Two (sets of) nodes X and Y are d-separated by a set of nodes Z if all of the paths between (any node in) X and (any node in) Y are blocked by Z.

Theorem: Given that P is Markov with respect to G,

$$X \perp\!\!\!\perp_G Y \mid Z \implies X \perp\!\!\!\perp_P Y \mid Z$$

local Markov assumption $\iff$ global Markov assumption

d-separation

Two (sets of) nodes X and Y are d-separated by a set of nodes Z if all of the paths between (any node in) X and (any node in) Y are blocked by Z.

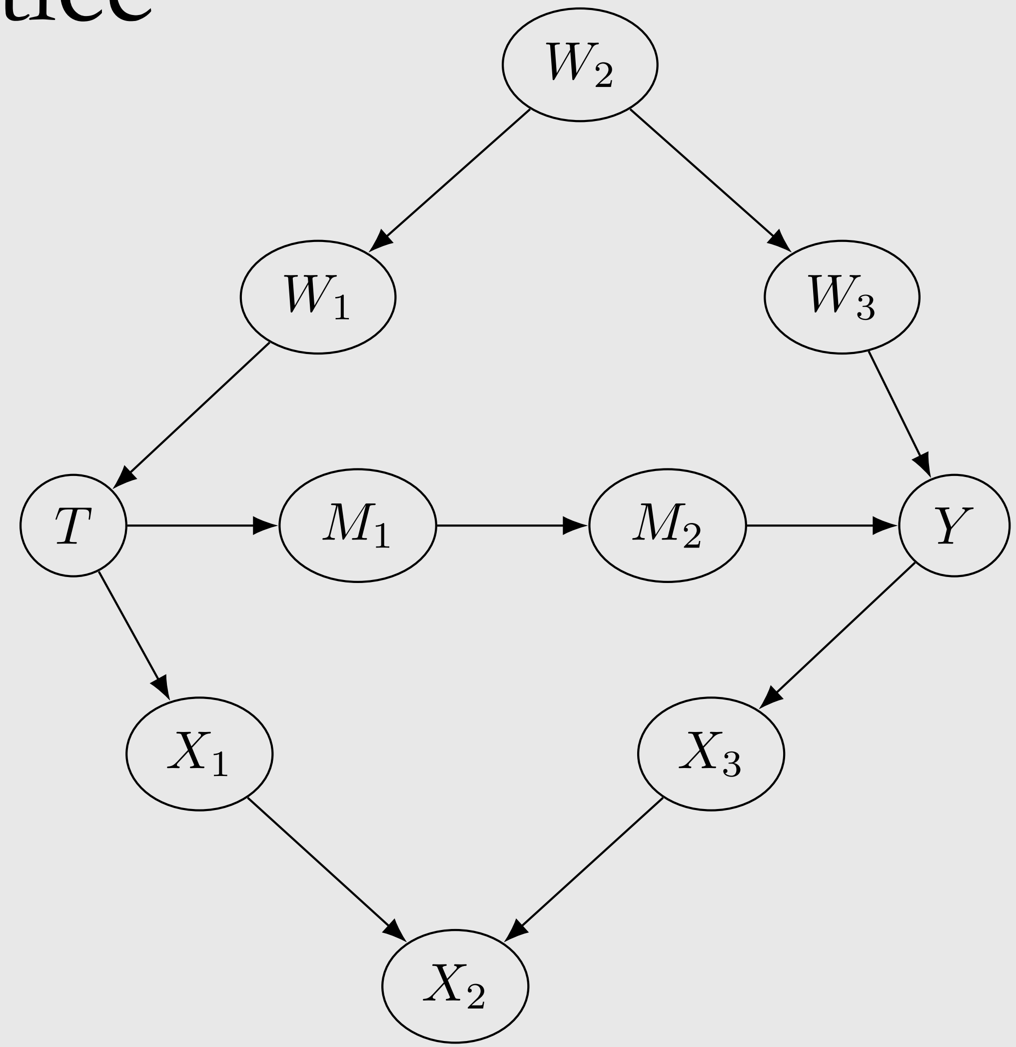
Theorem: Given that P is Markov with respect to G,

$$X \perp\!\!\!\perp_G Y \mid Z \implies X \perp\!\!\!\perp_P Y \mid Z$$

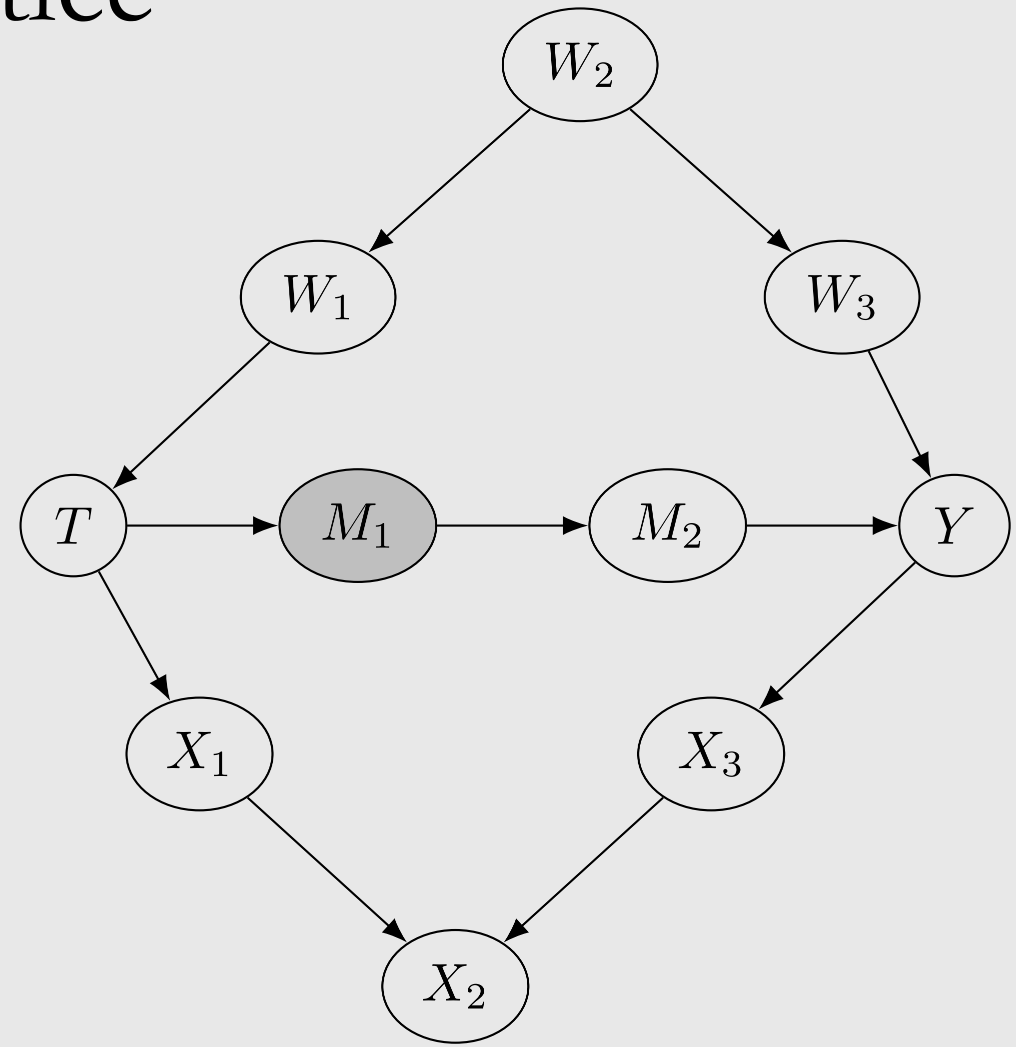
local Markov assumption $\iff$ global Markov assumption

Markov assumption

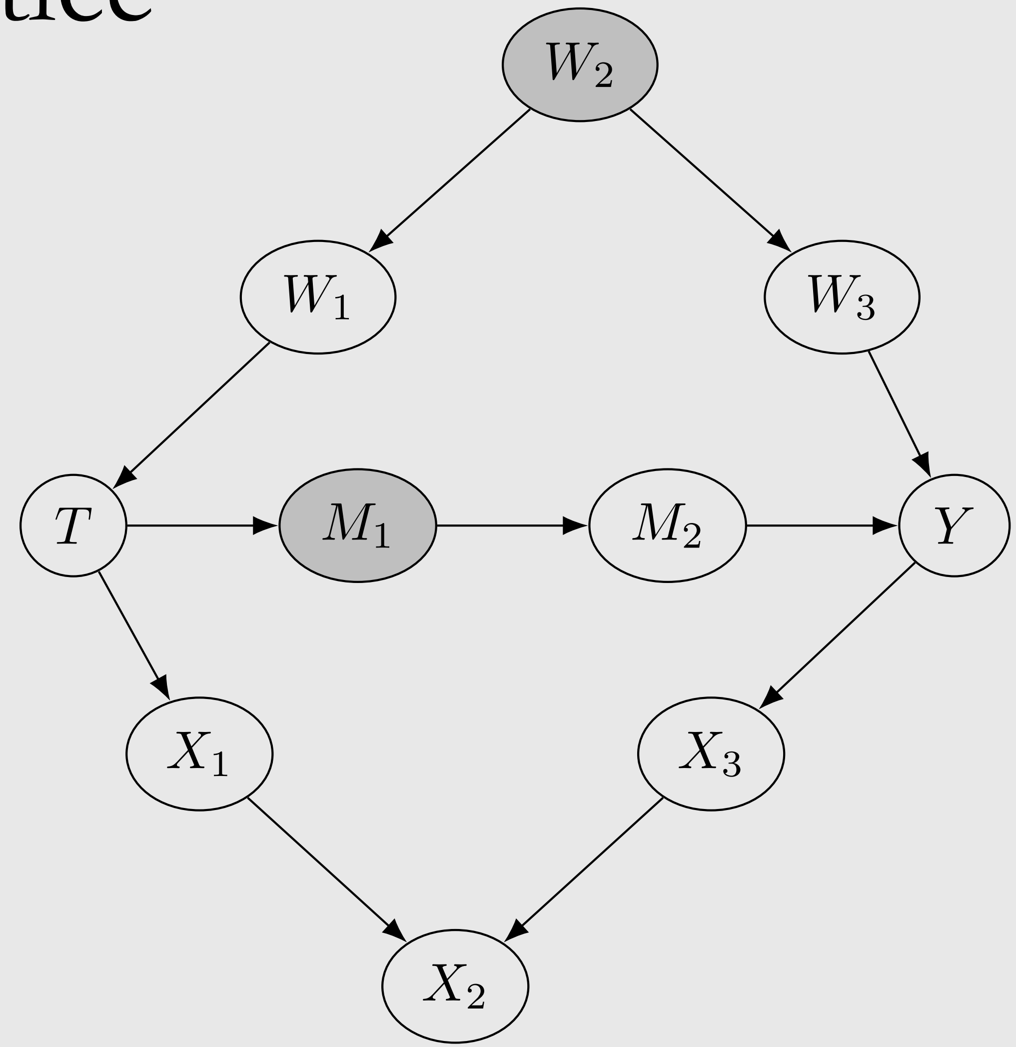
d-separation practice



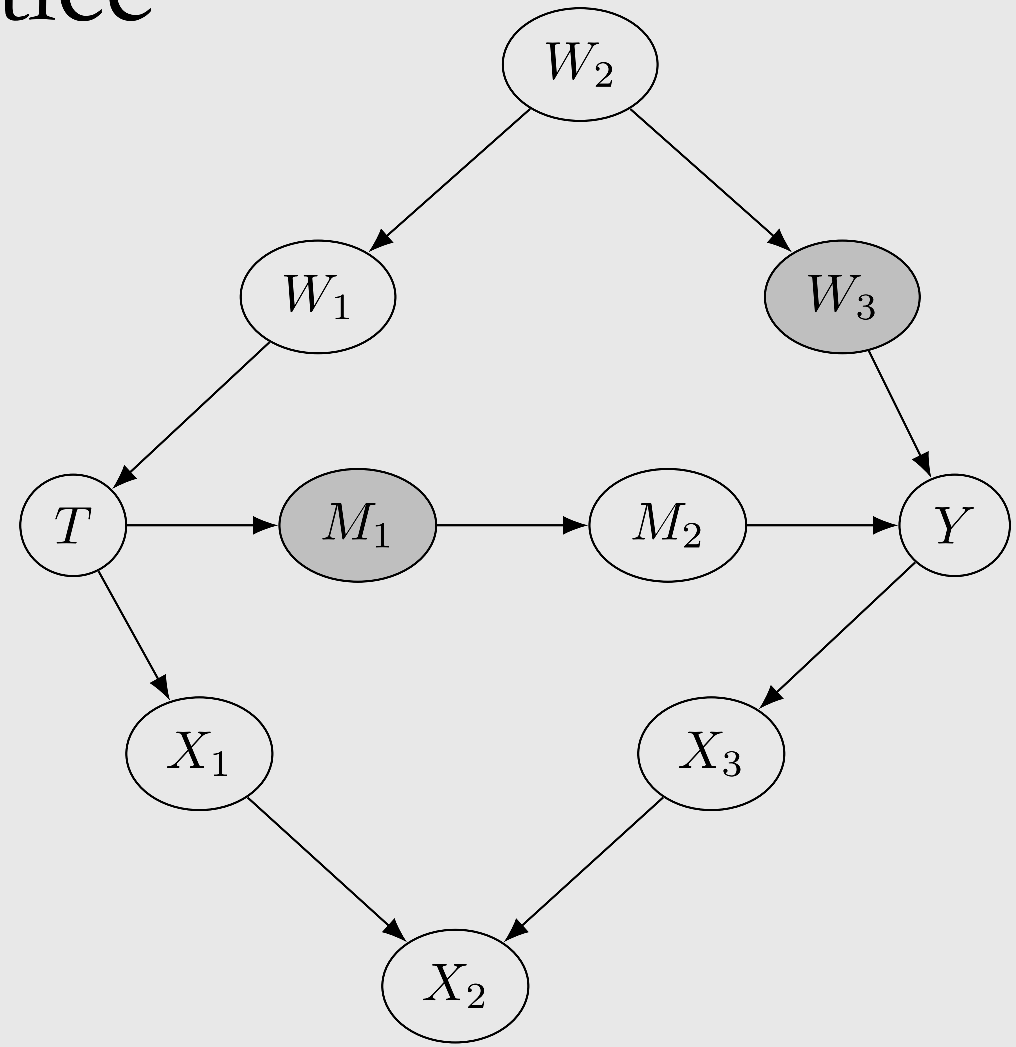
d-separation practice



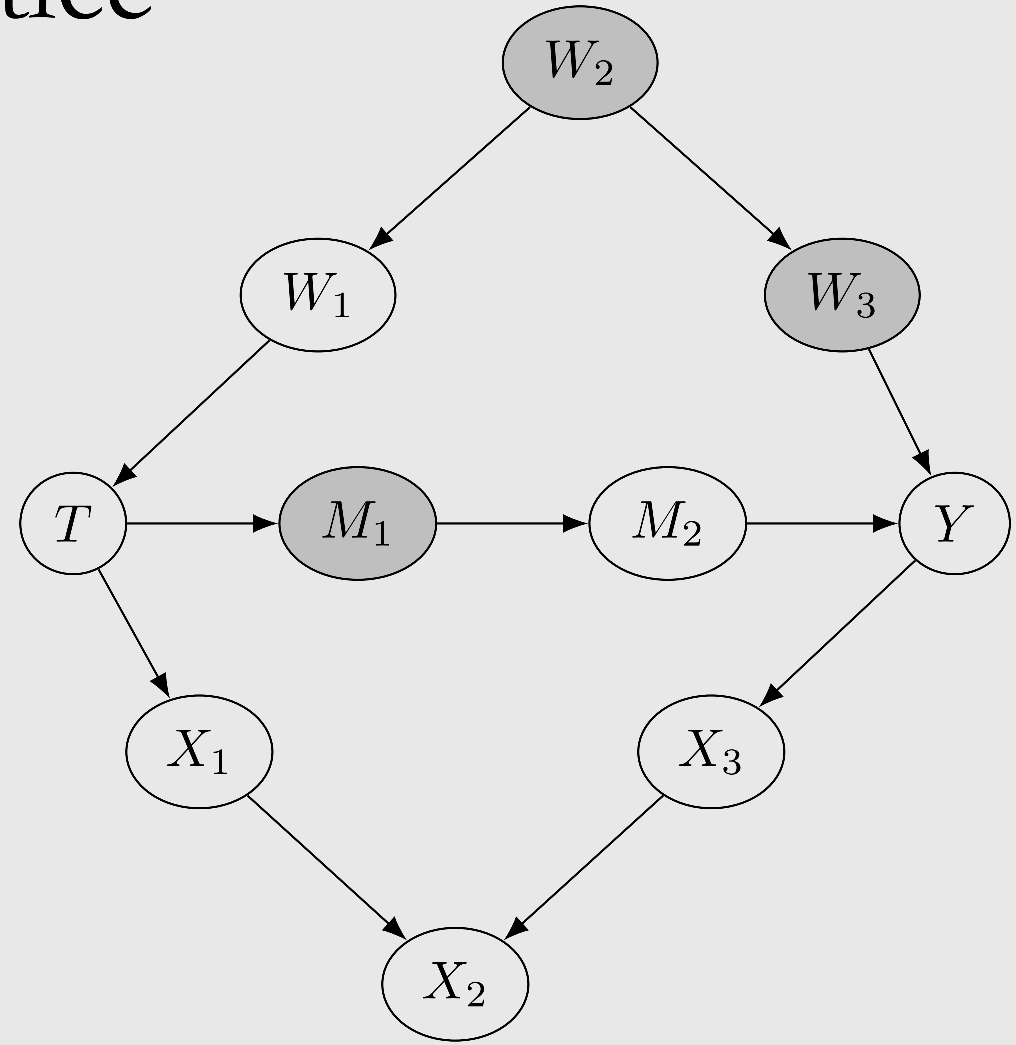
d-separation practice



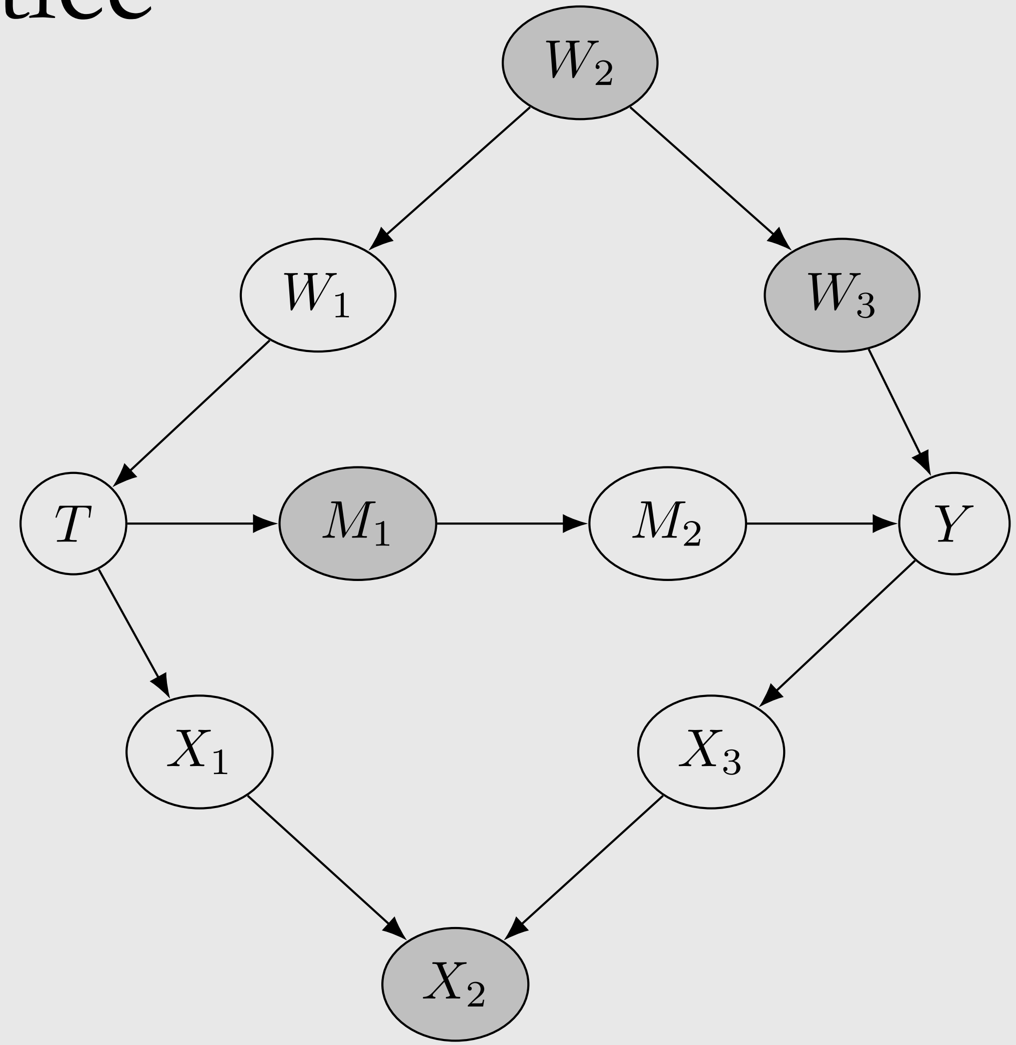
d-separation practice



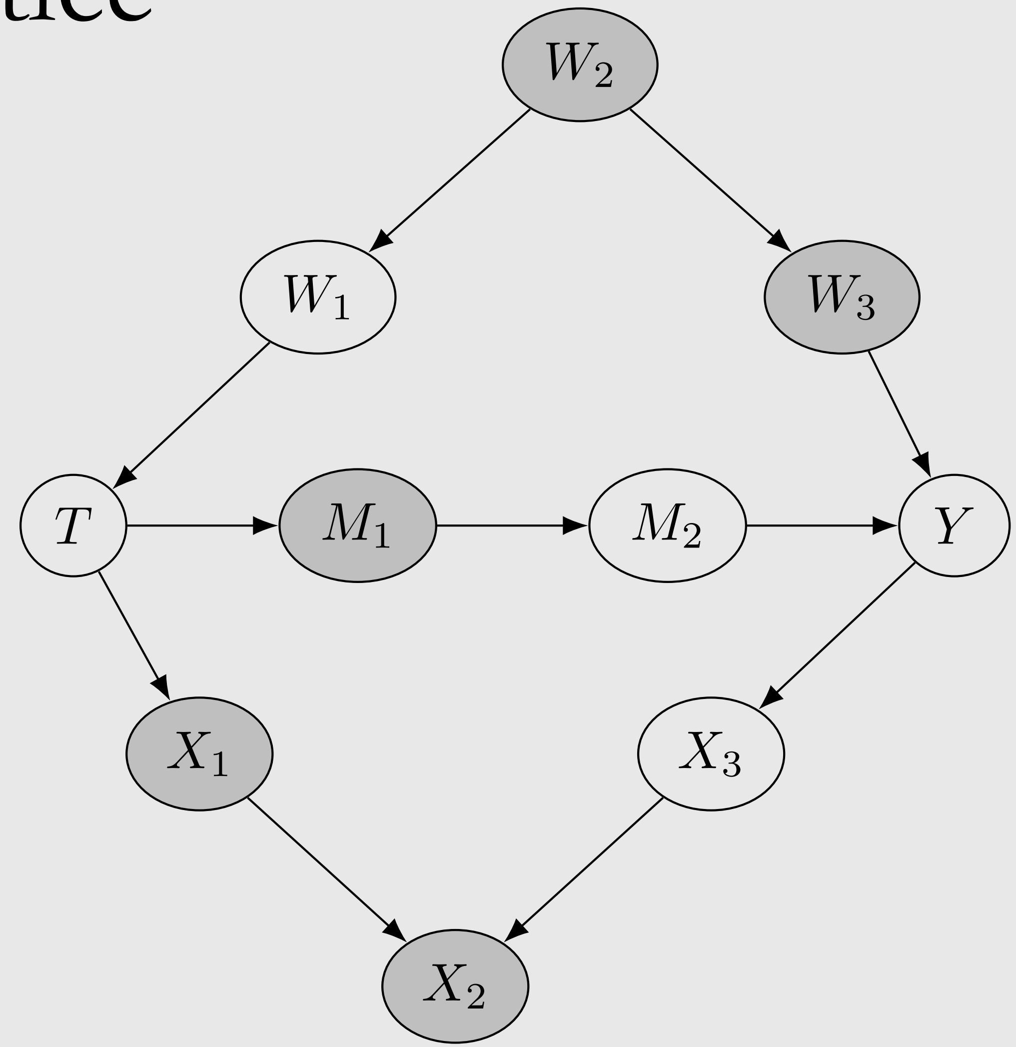
d-separation practice



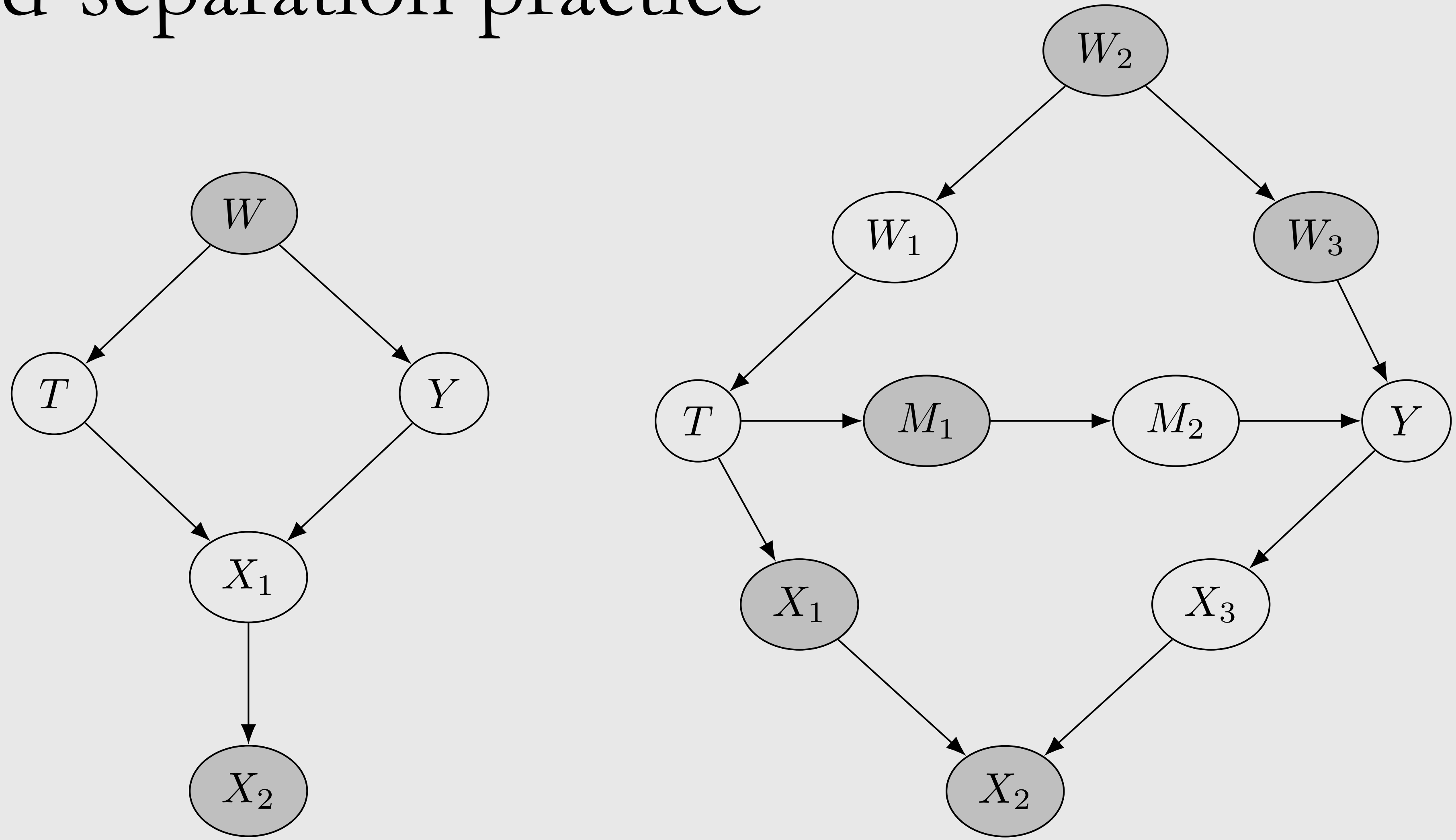
d-separation practice



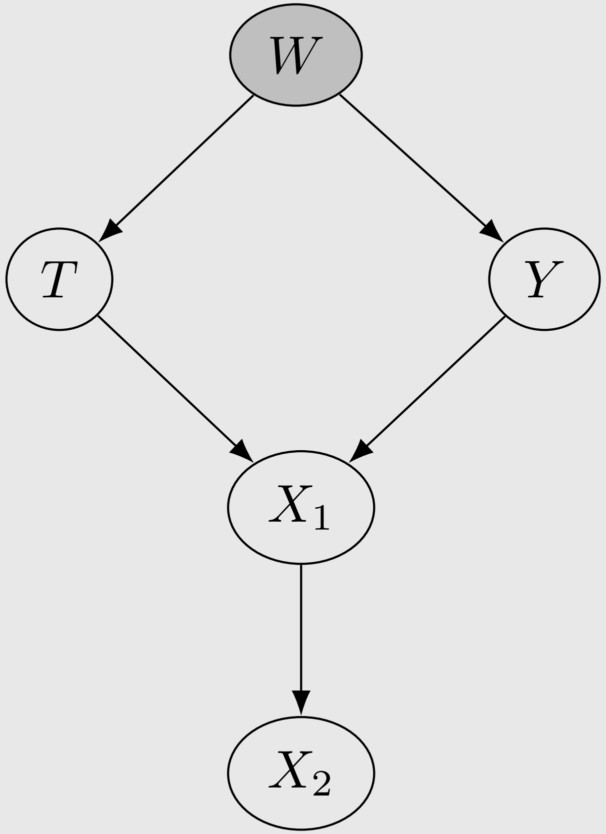
d-separation practice

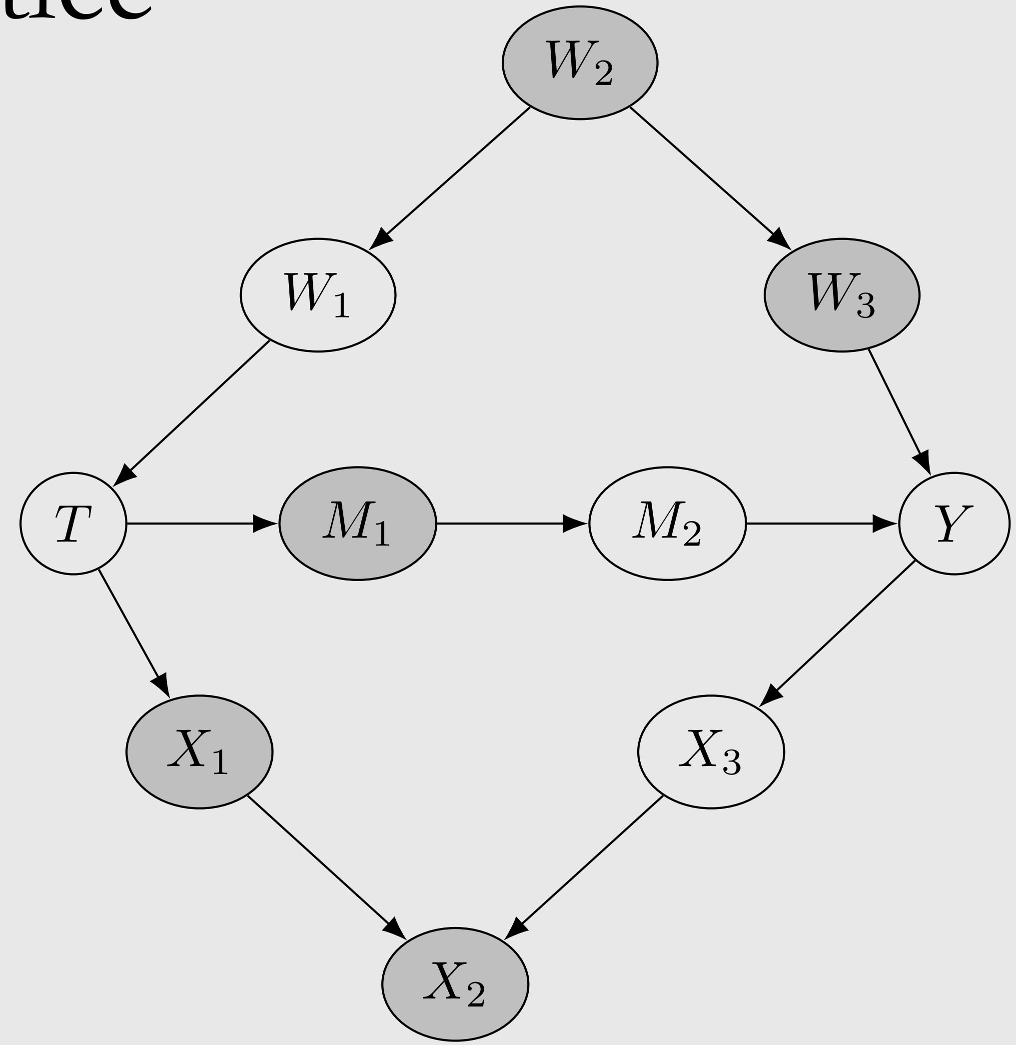


d-separation practice

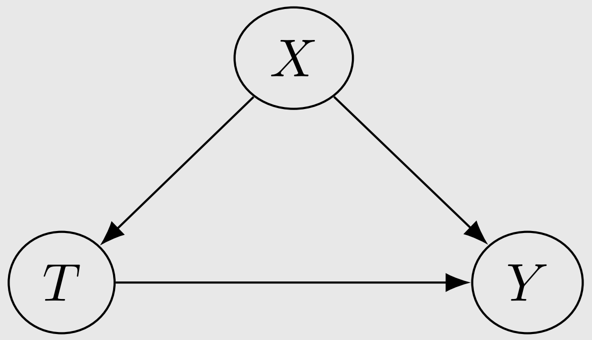


d-separation practice

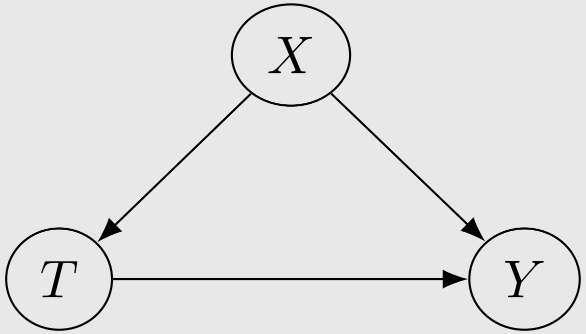


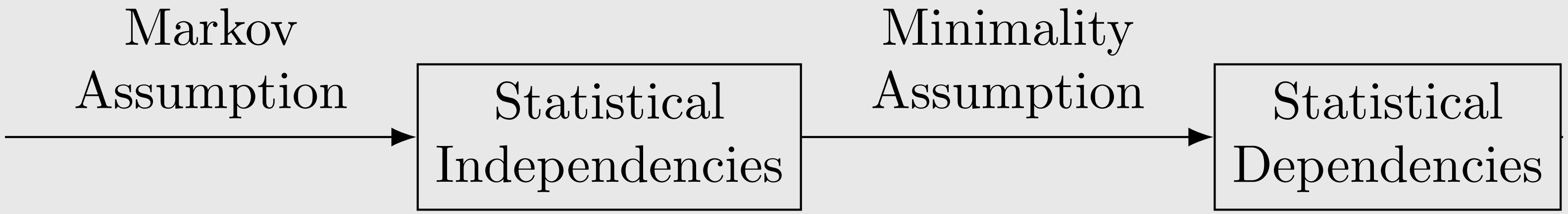


The flow of association and causation

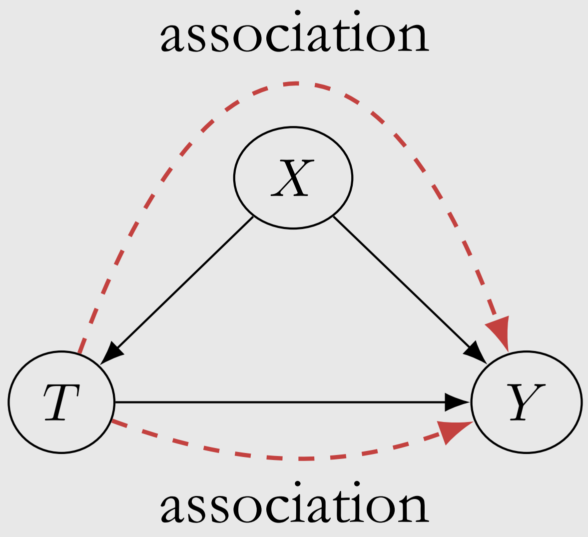


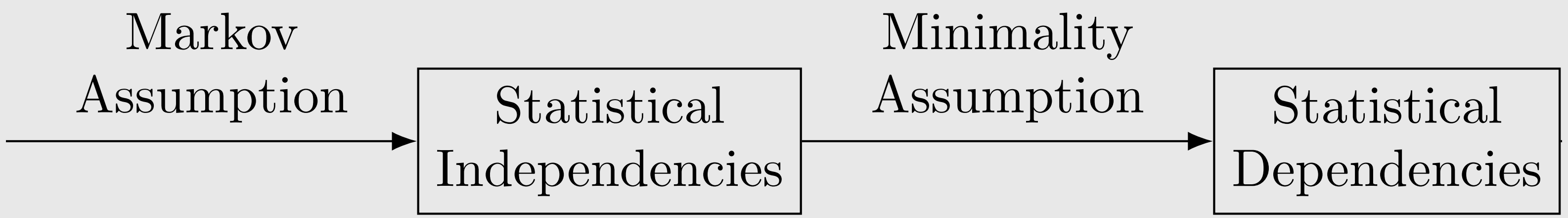
The flow of association and causation



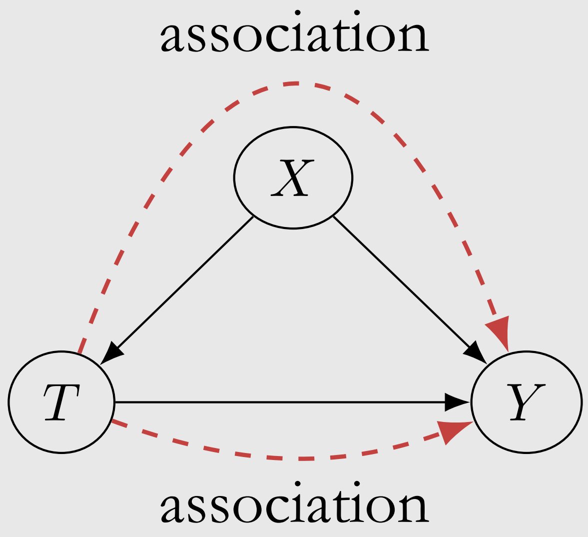


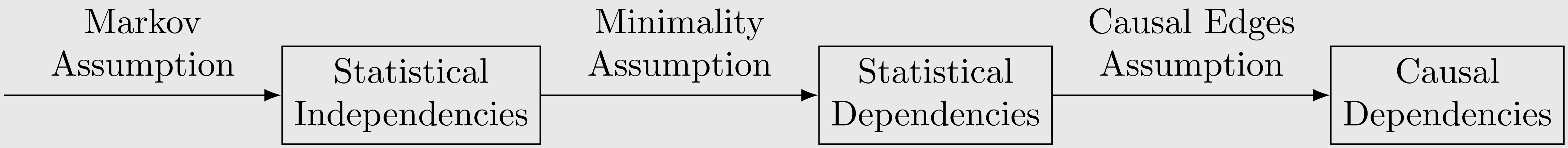
The flow of association and causation



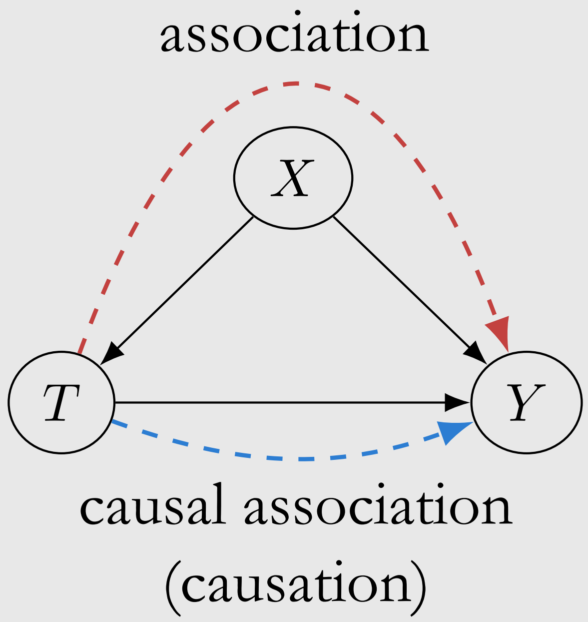


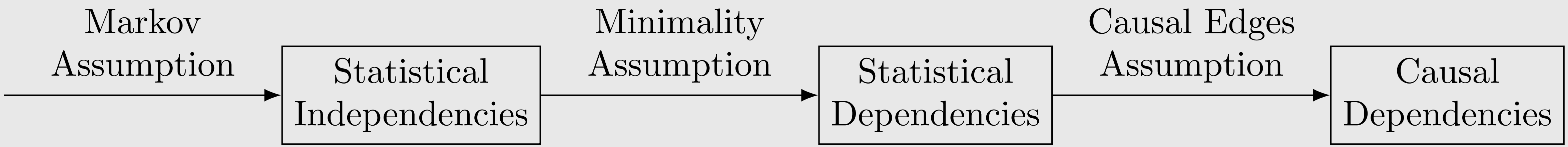
The flow of association and causation



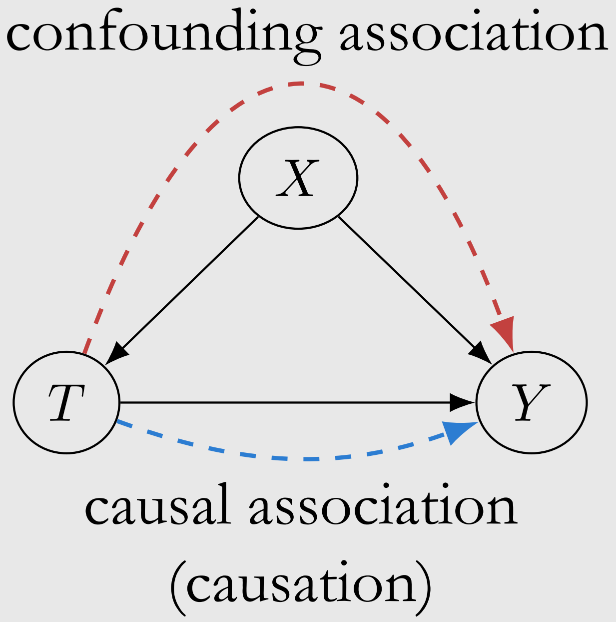


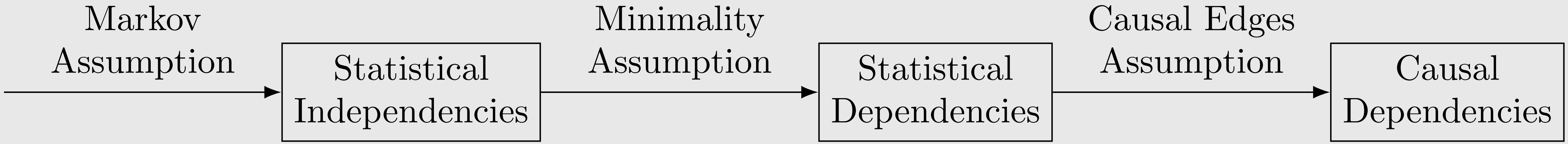
The flow of association and causation

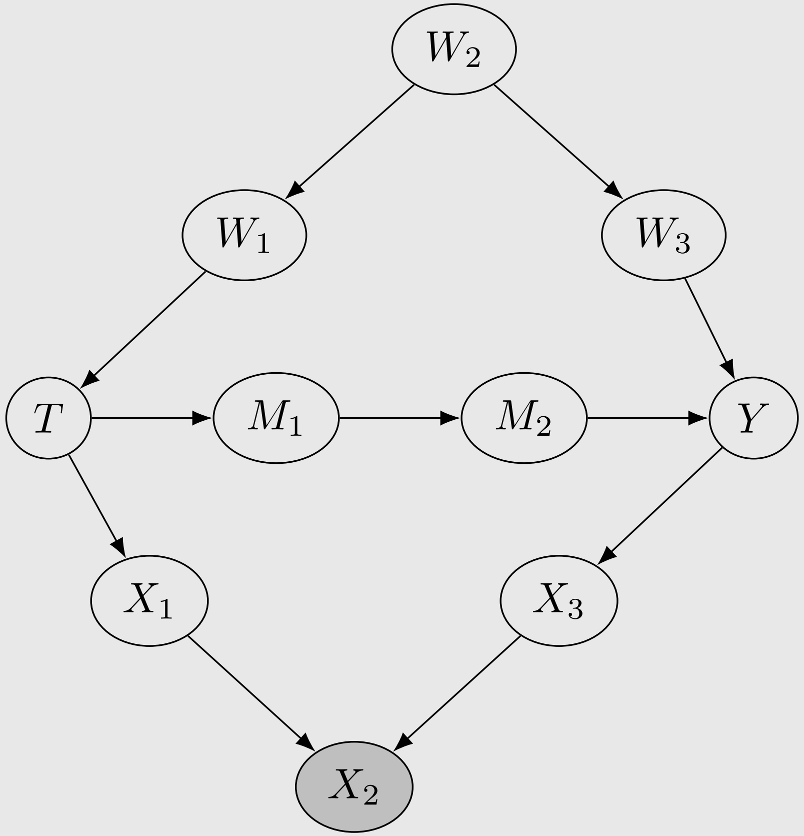


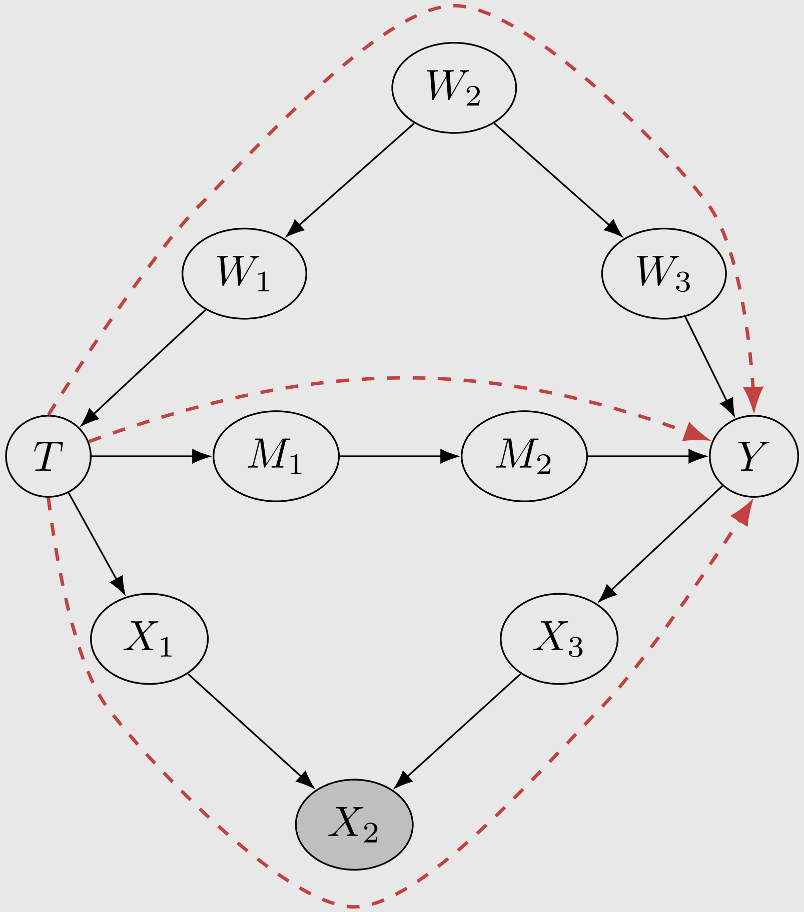


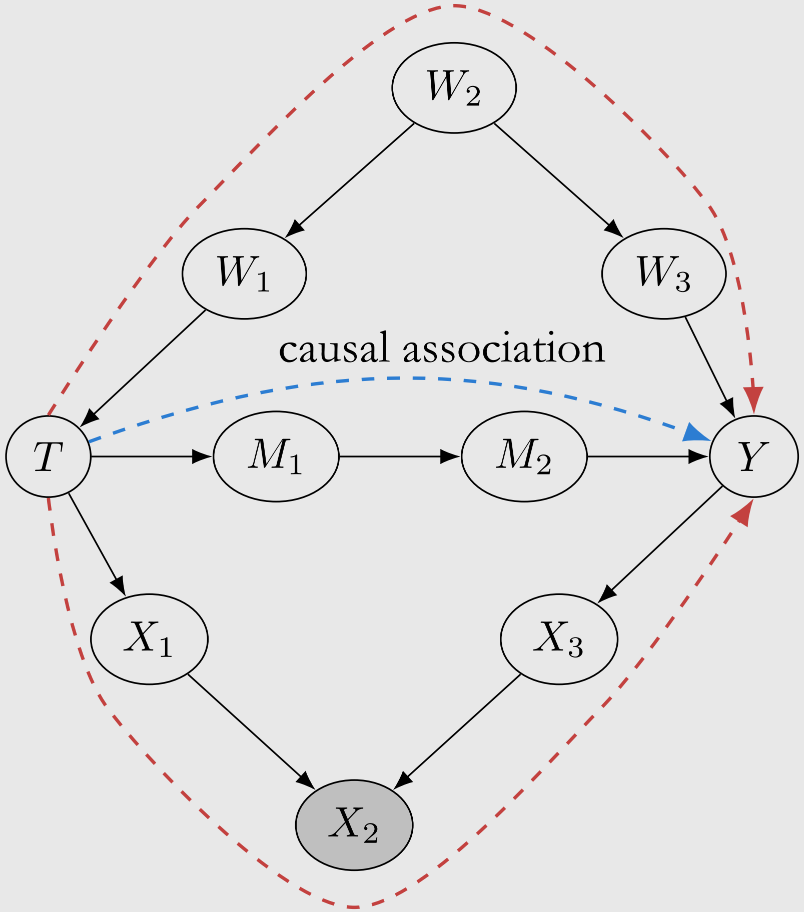
The flow of association and causation



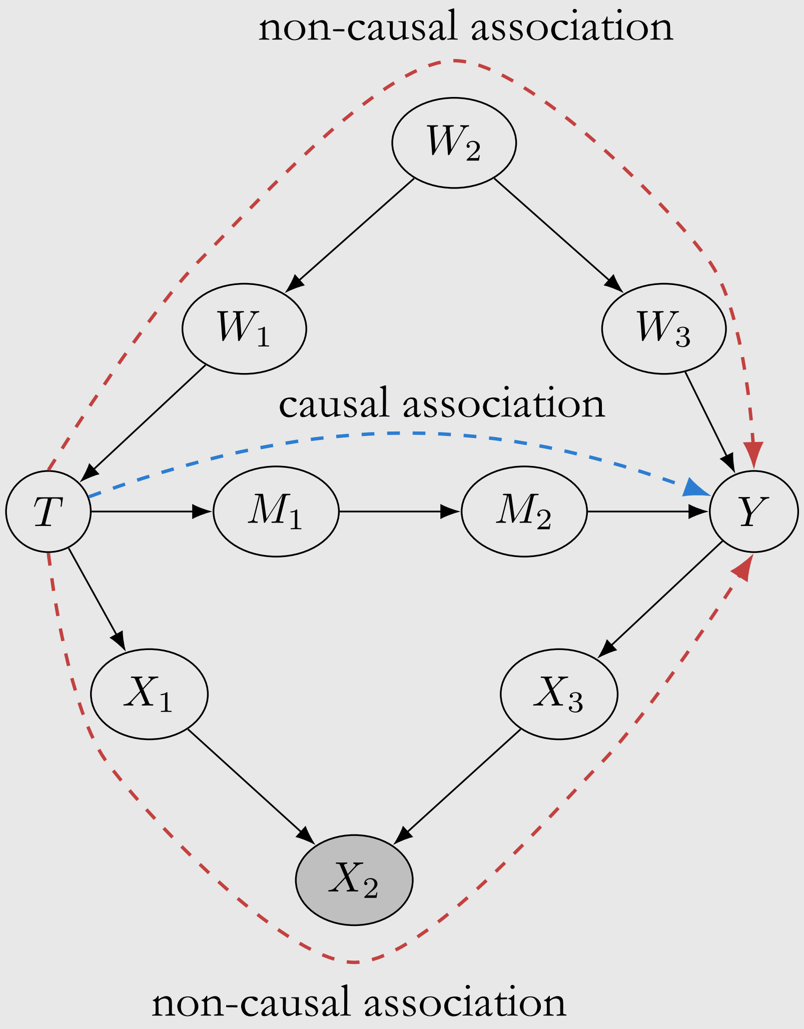



W_2
W_1
W_3
T
M_1
M_2
Y
X_1
X_3
X_2


W_2
W_1
W_3
T
M_1
M_2
Y
X_1
X_3
X_2



W_2
W_1
W_3
causal association
T
M_1
M_2
Y
X_1
X_3
X_2


non-causal association
W2
W1
W3
causal association
T
M1
M2
Y
X1
X3
X2
non-causal association