Bayesian Networks

Today:

- Bayesian Networks
- How do we reason about independence in Bayesian Networks?
- How do we sample from Bayesian Networks?

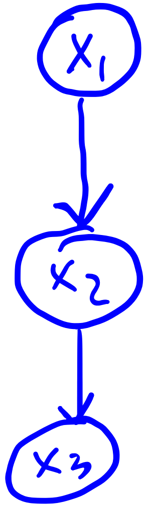
Review of Definitions

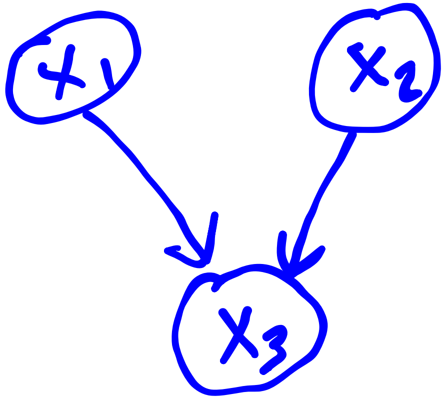
Review of Definitions

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**

Review of Definitions

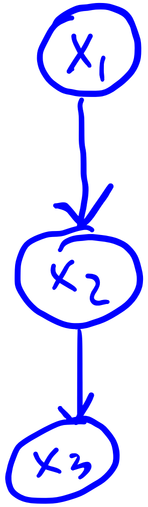
Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**

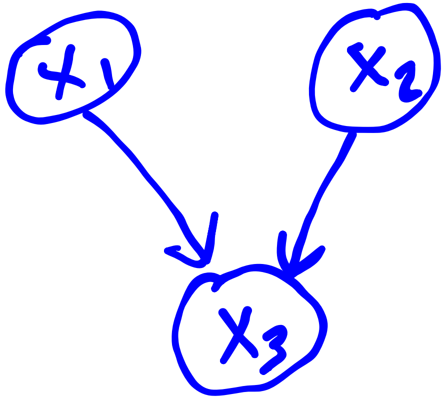




Review of Definitions

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**

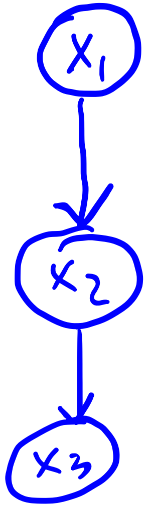


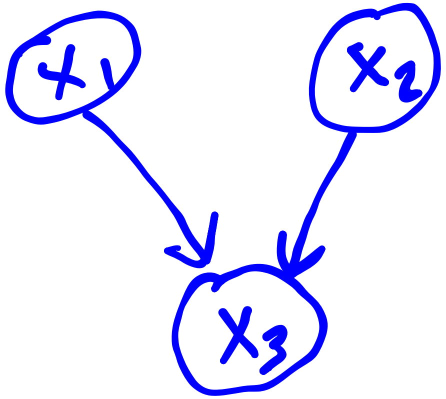


- Node:

Review of Definitions

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**

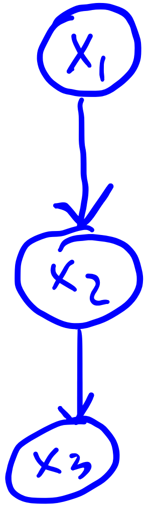


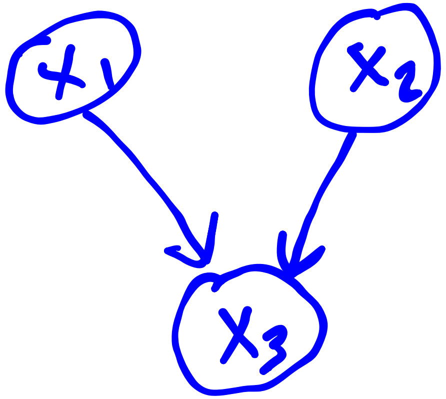


- Node: Random Variable

Review of Definitions

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**



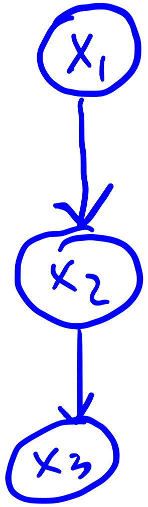


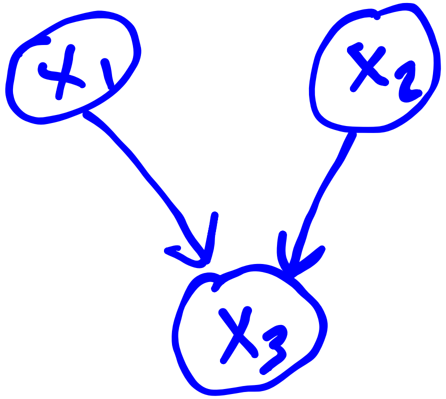
- Node: Random Variable
- Edges encode:

$$P(x_{1:n}) = \prod_{i=1}^{n} P(x_i \mid \mathrm{pa}(x_i))$$

Review of Definitions

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**





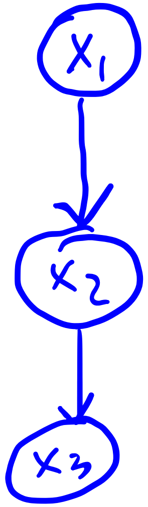
- Node: Random Variable
- Edges encode:

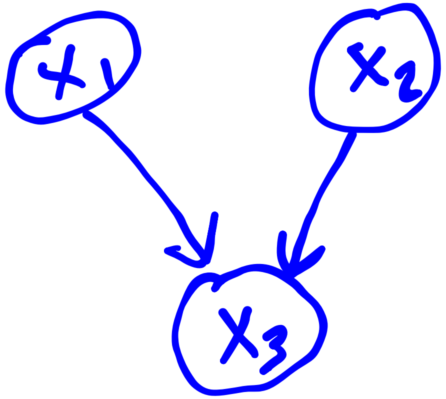
$$P(x_{1:n}) = \prod_{i=1}^{n} P(x_i \mid \mathrm{pa}(x_i))$$

Independence

Review of Definitions

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**





- Node: Random Variable
- Edges encode:

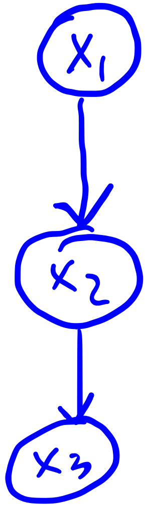
$$P(x_{1:n}) = \prod_{i=1}^{n} P(x_i \mid \mathrm{pa}(x_i))$$

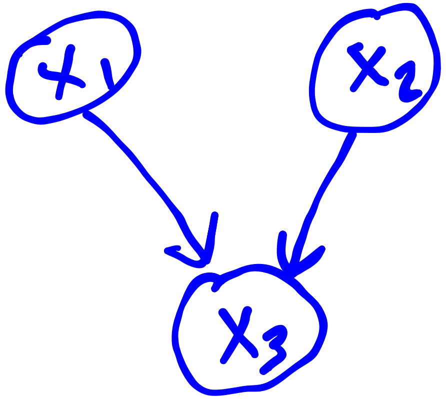
Independence

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

Review of Definitions

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**





- Node: Random Variable
- Edges encode:

$$P(x_{1:n}) = \prod_{i=1}^{n} P(x_i \mid \mathrm{pa}(x_i))$$

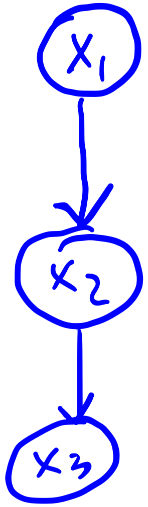
Independence

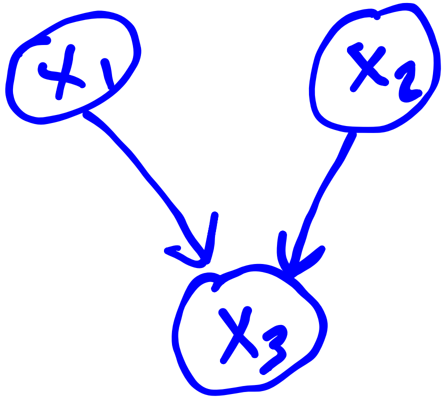
$$P(X, Y) = P(X)\,P(Y)$$

Conditional Independence

Review of Definitions

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**





- Node: Random Variable
- Edges encode:

$$P(x_{1:n}) = \prod_{i=1}^{n} P(x_i \mid \mathrm{pa}(x_i))$$

Independence

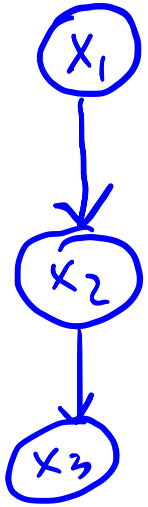
$$P(X, Y) = P(X)\, P(Y)$$

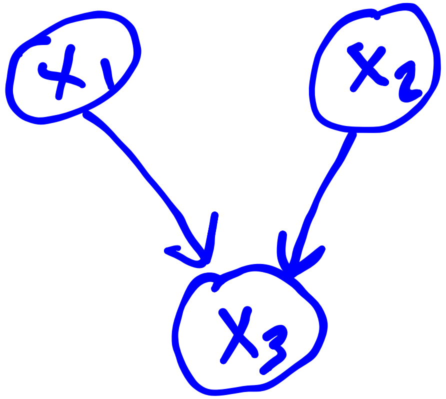
Conditional Independence

$$P(X, Y \mid Z) = P(X \mid Z)\, P(Y \mid Z)$$

Review of Definitions

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**





- Node: Random Variable
- Edges encode:

$$P(x_{1:n}) = \prod_{i=1}^{n} P(x_i \mid \mathrm{pa}(x_i))$$

Independence

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

$X \perp Y$

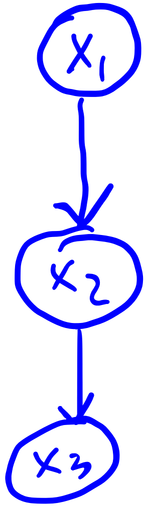
Conditional Independence

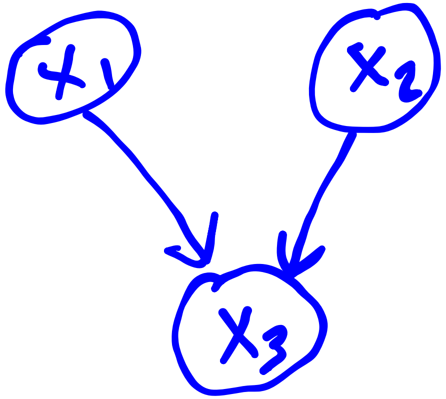
$$P(X, Y \mid Z) = P(X \mid Z)\,P(Y \mid Z)$$

$(X \perp Y \mid Z)$

Review of Definitions

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**





- Node: Random Variable
- Edges encode:

$$P(x_{1:n}) = \prod_{i=1}^{n} P(x_i \mid \mathrm{pa}(x_i))$$

Independence

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

$X \perp Y$

Conditional Independence

$$P(X,Y \mid Z) = P(X \mid Z)\,P(Y \mid Z)$$

$(X \perp Y \mid Z)$

$\Updownarrow$

$P(X \mid Z) = P(X \mid Y,Z)$

What does conditional independence mean?

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z$

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$$X \perp Y \mid Z \qquad \implies$$

https://kunalmenda.com/2019/02/21/causation-and-correlation/

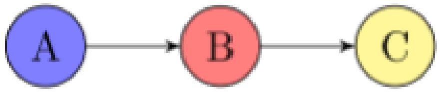
What does conditional independence mean?

$X \perp Y \mid Z \quad \implies \quad$ All of X's dependence on Y comes through Z

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

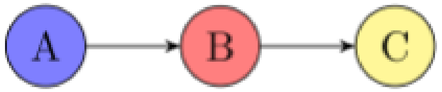
$X \perp Y \mid Z \implies$ All of X's dependence on Y comes through Z



https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z \quad \implies \quad$ All of X's dependence on Y comes through Z

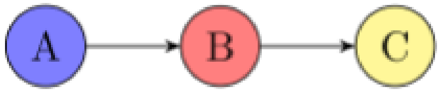


$A \perp C \mid B$?

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z \qquad \Longrightarrow \qquad$ All of X's dependence on Y comes through Z



$A \perp C \mid B$? Yes

$$P(A,B,C) = P(A)P(B|A)\,P(C|B)$$

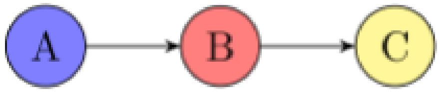
$$A \perp C \mid B : \; P(C,A|B) = P(C|B)P(A|B)$$

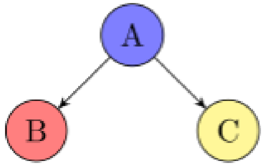
$$P(C,A|B) = \frac{P(A,B,C)}{P(B)} = \frac{P(A)\,P(B|A)\,P(C|B)}{P(B)} = \frac{\cancel{P(A)}\;P(B,A)\,P(C|B)}{\cancel{P(A)}\,P(B)} = P(A|B)P(C|B)$$

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z \quad \Longrightarrow \quad$ All of X's dependence on Y comes through Z



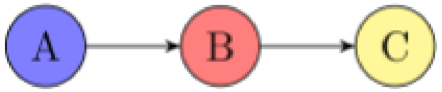


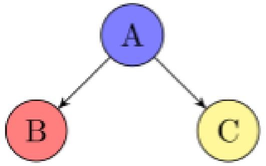
$A \perp C \mid B$? Yes

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z \quad \Longrightarrow \quad$ All of X's dependence on Y comes through Z





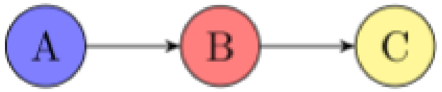
$A \perp C \mid B$? Yes

$B \perp C \mid A$?

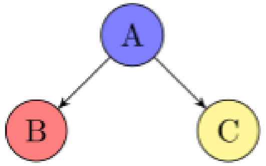
https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z \quad \Longrightarrow \quad$ All of X's dependence on Y comes through Z



$A \perp C \mid B$? Yes

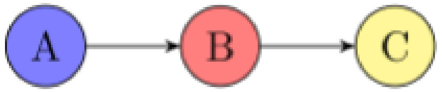


$B \perp C \mid A$? Yes

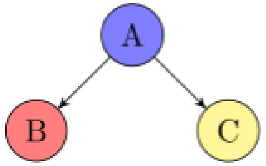
https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

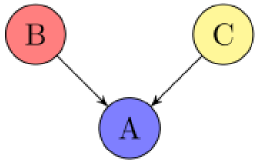
$X \perp Y \mid Z$ $\implies$ All of X's dependence on Y comes through Z



$A \perp C \mid B$? Yes



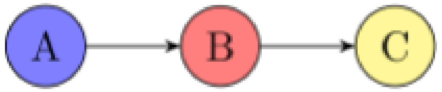
$B \perp C \mid A$? Yes



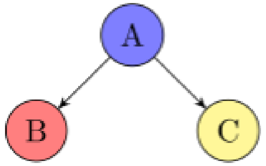
https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

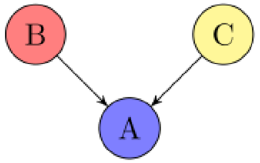
$X \perp Y \mid Z \implies$ All of X's dependence on Y comes through Z



$A \perp C \mid B$? Yes



$B \perp C \mid A$? Yes

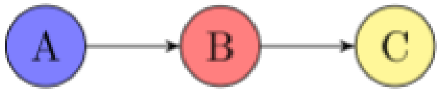


$B \perp C \mid A$?

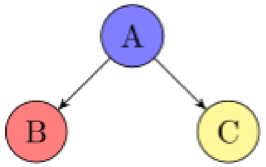
https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

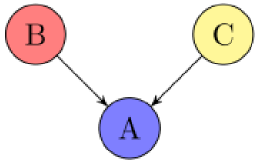
$X \perp Y \mid Z \implies$ All of X's dependence on Y comes through Z



$A \perp C \mid B$? Yes



$B \perp C \mid A$? Yes

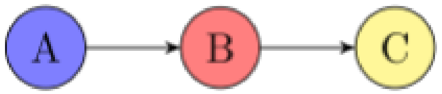


$B \perp C \mid A$? No

https://kunalmenda.com/2019/02/21/causation-and-correlation/

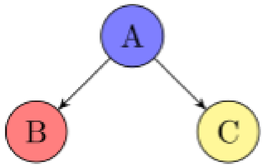
What does conditional independence mean?

$X \perp Y \mid Z \quad \implies$ All of X's dependence on Y comes through Z

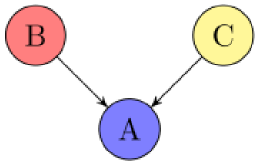


$A \perp C \mid B$? Yes

Mediator

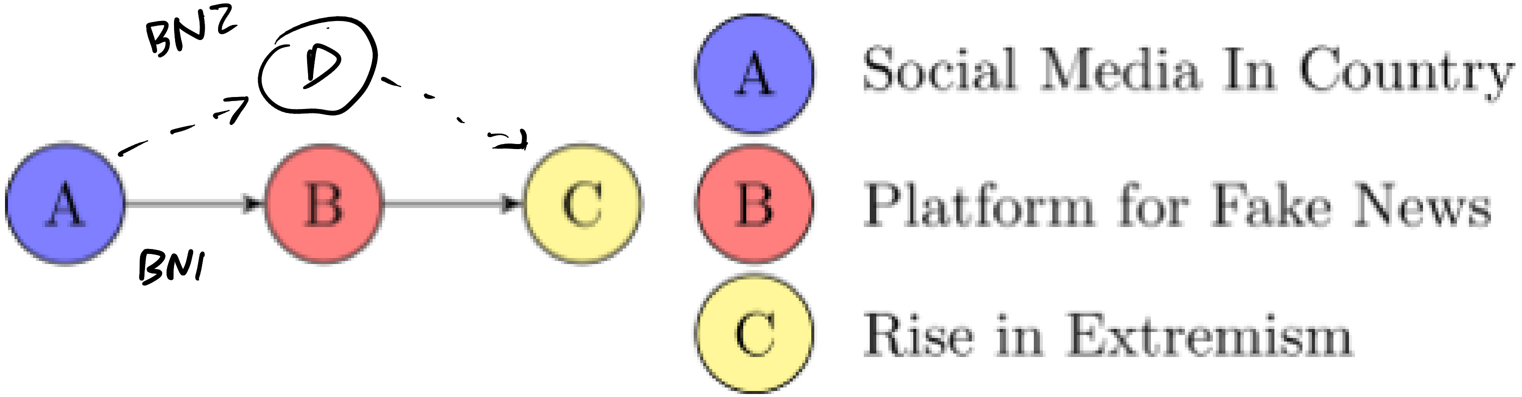


$B \perp C \mid A$? Yes



$B \perp C \mid A$? No

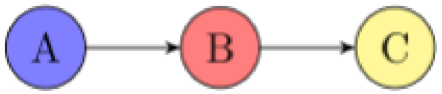
D : Platform for extremists to connect



https://kunalmenda.com/2019/02/21/causation-and-correlation/

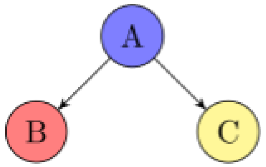
What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's dependence on Y comes through Z



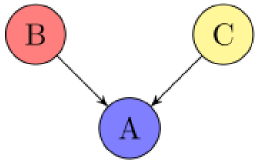
$A \perp C \mid B$? Yes

Mediator

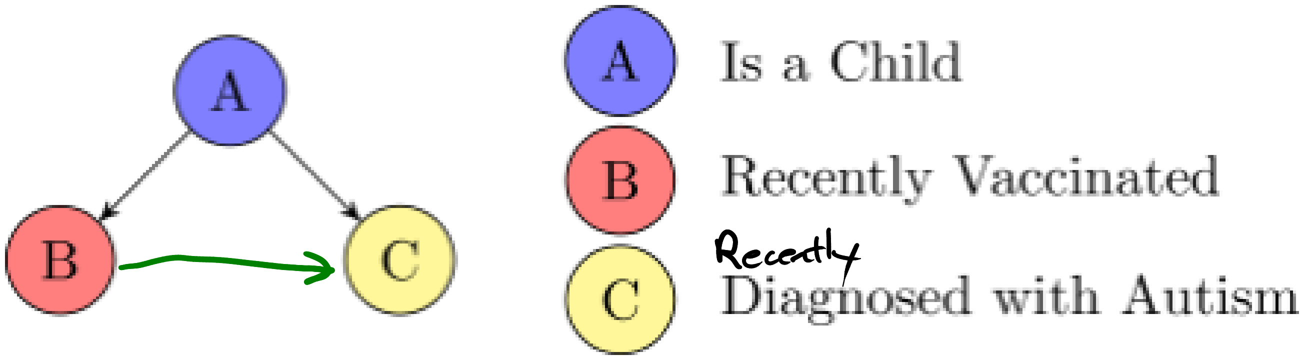


$B \perp C \mid A$? Yes

Confounder



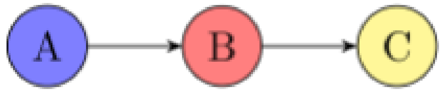
$B \perp C \mid A$? No



https://kunalmenda.com/2019/02/21/causation-and-correlation/

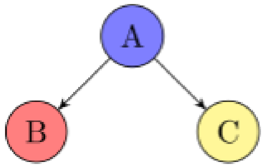
What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's dependence on Y comes through Z



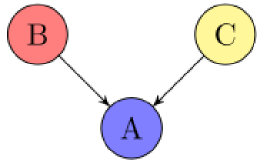
$A \perp C \mid B$? Yes

Mediator



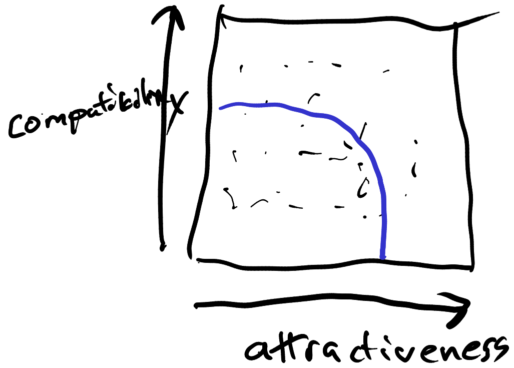
$B \perp C \mid A$? Yes

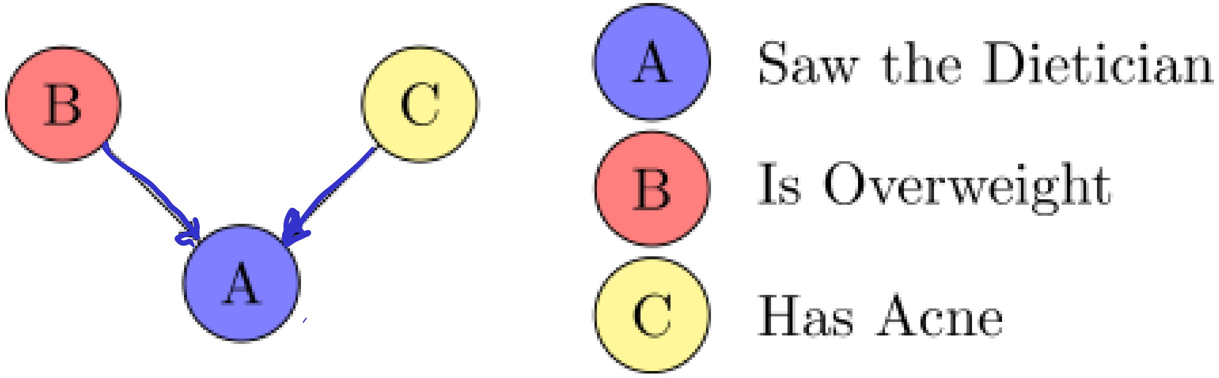
Confounder

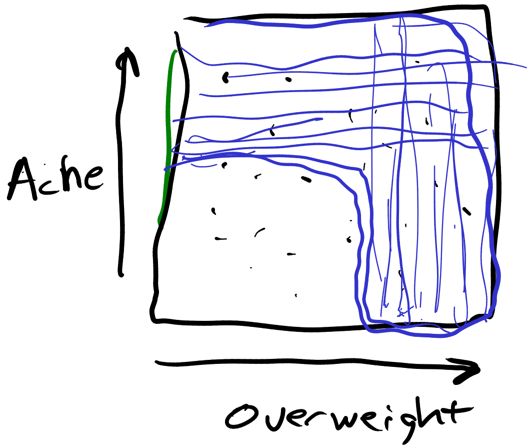


$B \perp C \mid A$? No

Collider



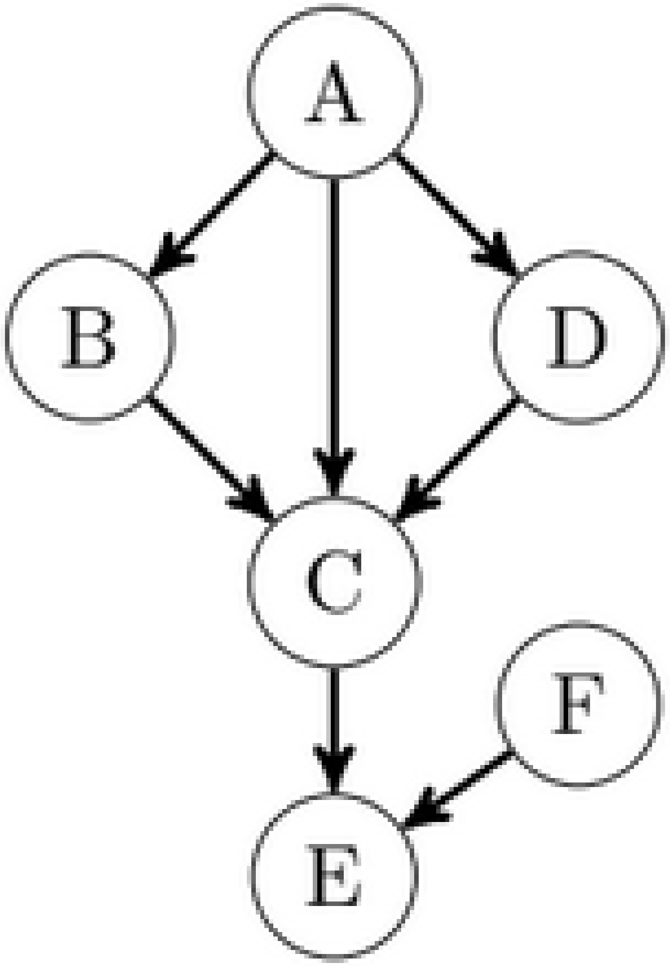




https://kunalmenda.com/2019/02/21/causation-and-correlation/

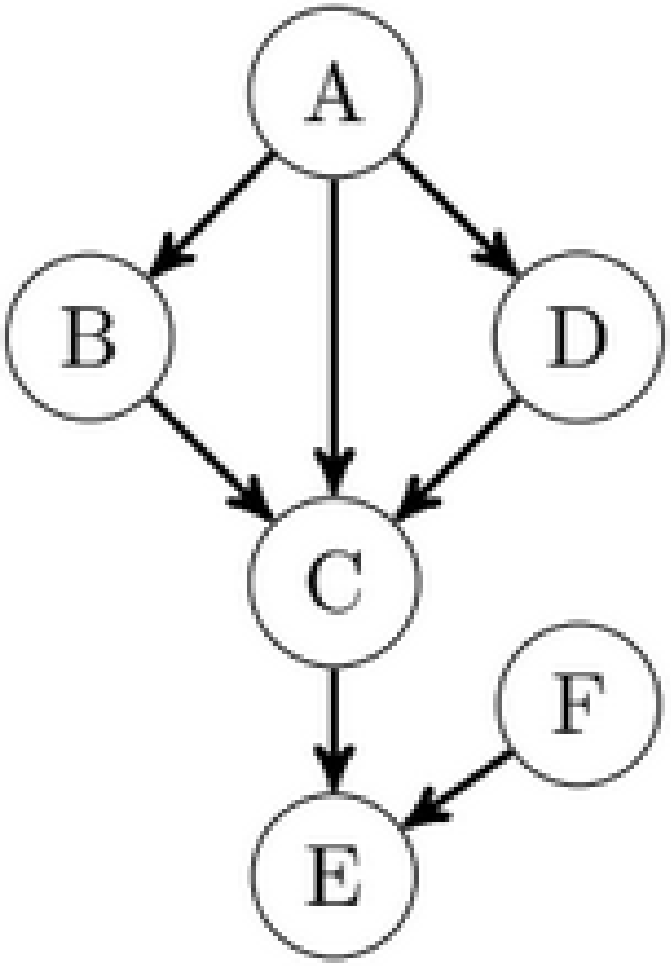
More Complex Example

More Complex Example

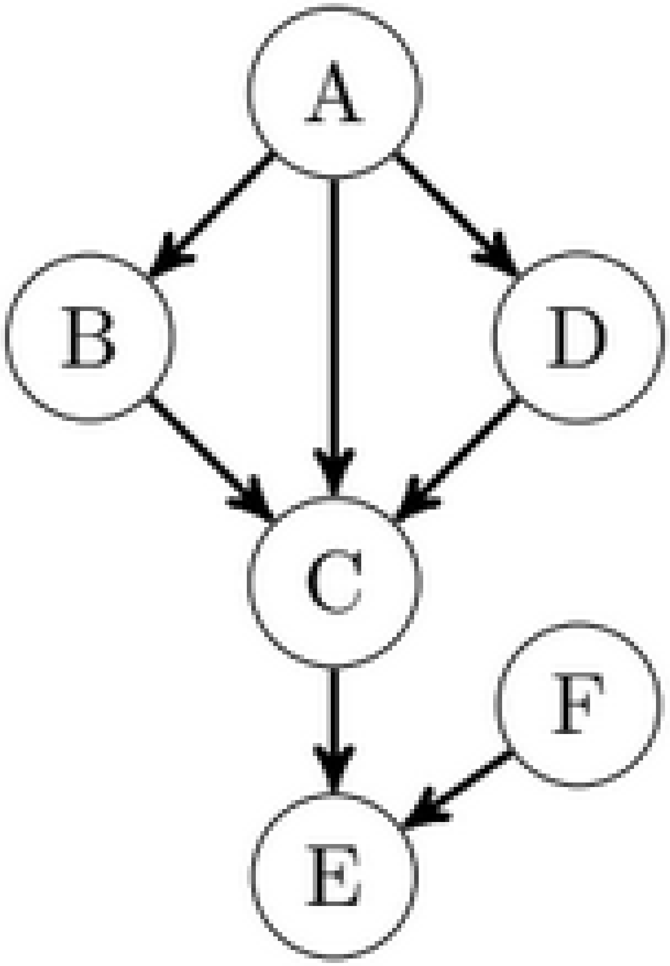


More Complex Example

$(B \perp D \mid A)$?



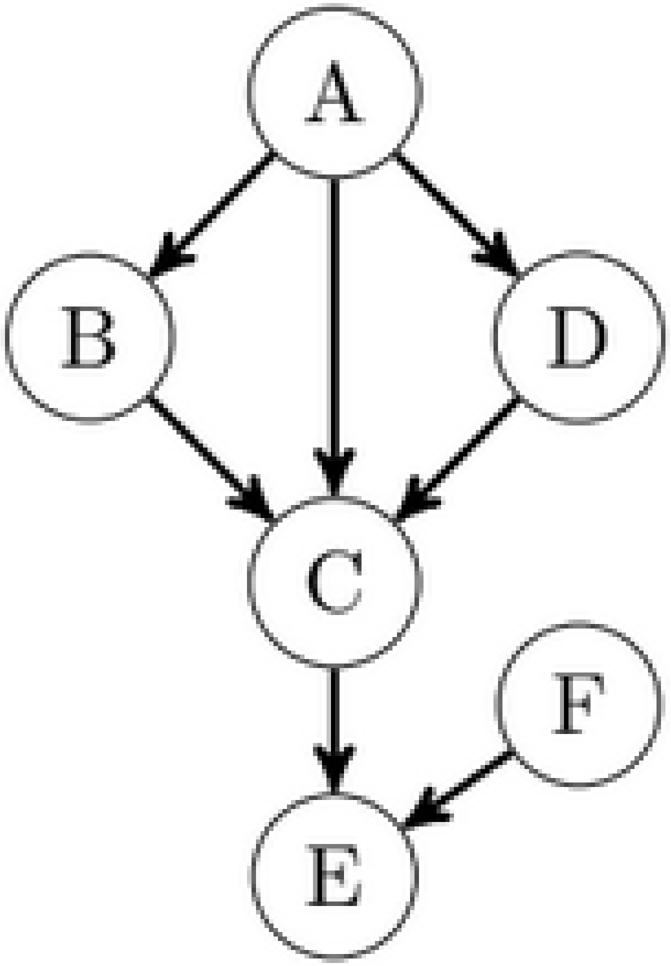
More Complex Example



$(B \perp D \mid A)$?

Yes!

More Complex Example

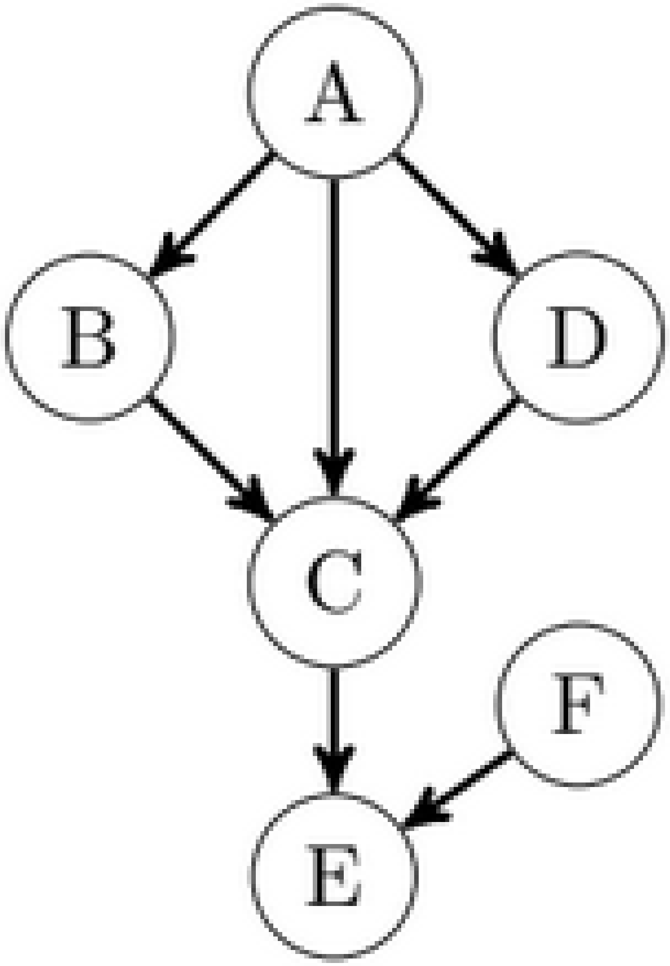


$(B \perp D \mid A)$?

Yes!

$(B \perp D \mid E)$?

More Complex Example



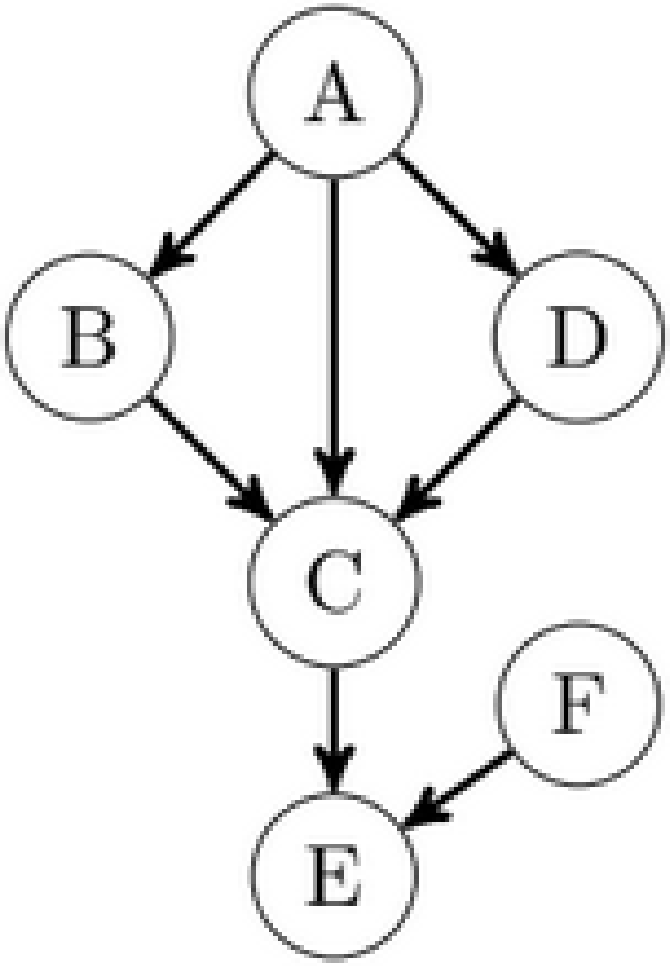
$(B \perp D \mid A)$?

Yes!

$(B \perp D \mid E)$?

No

More Complex Example



$(B \perp D \mid A)$?

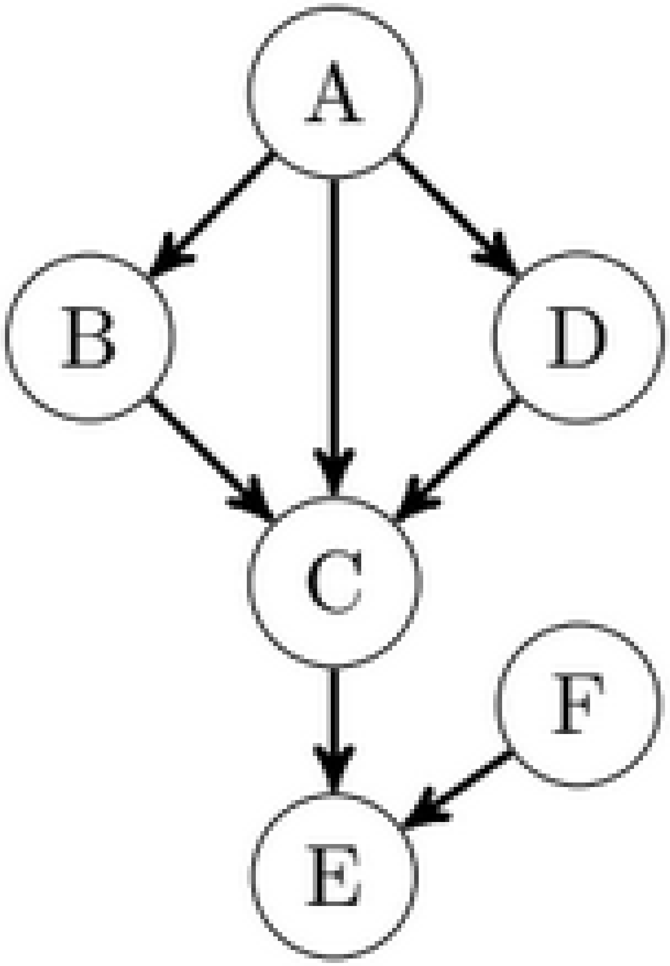
Yes!

$(B \perp D \mid E)$?

No

Why is this relevant?

More Complex Example



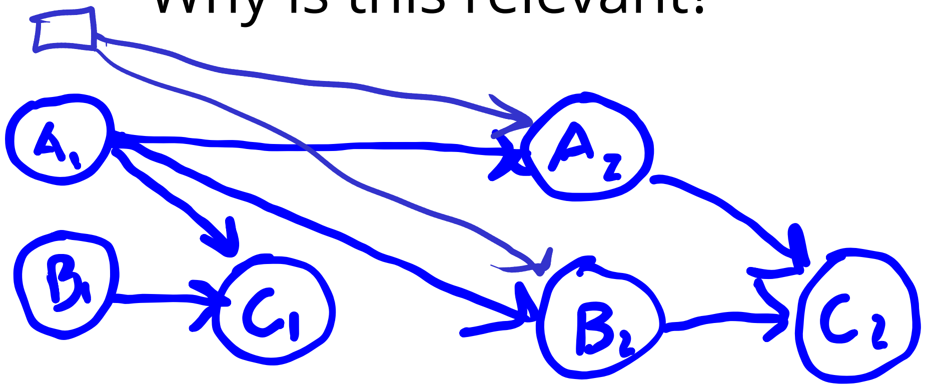
$(B \perp D \mid A)$?

Yes!

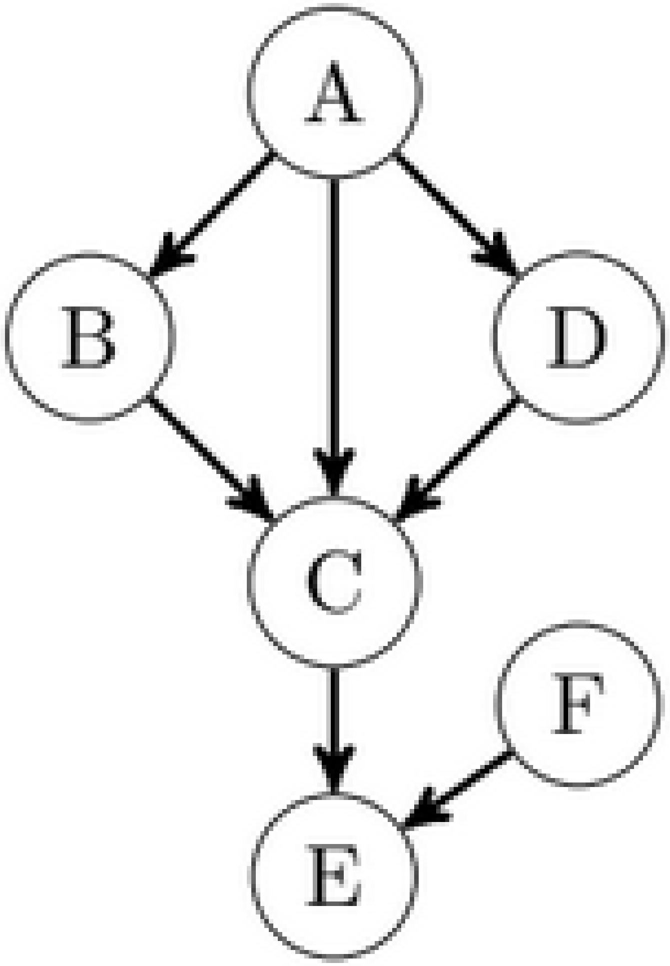
$(B \perp D \mid E)$?

No

Why is this relevant?



More Complex Example



$(B \perp D \mid A)$?

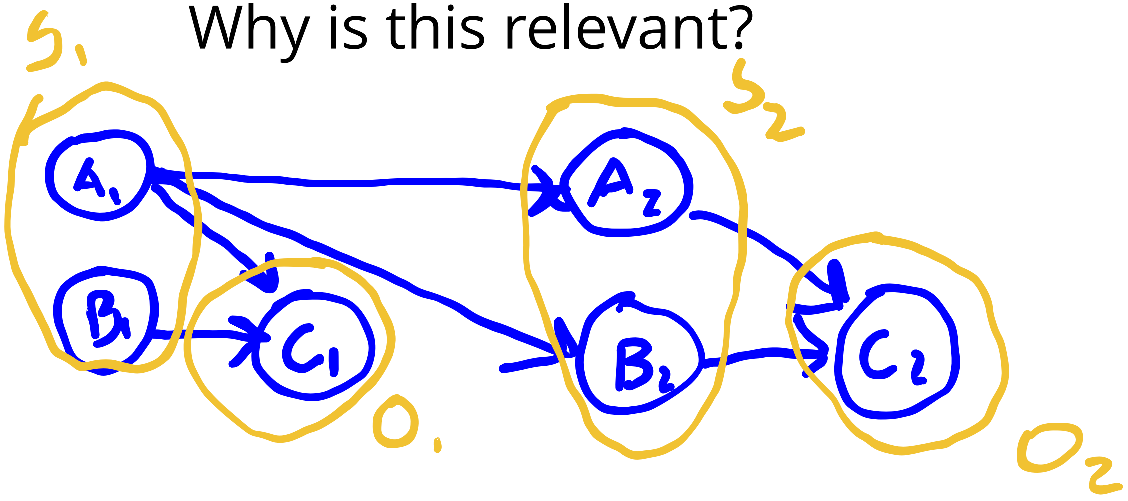
Yes!

$(B \perp D \mid E)$?

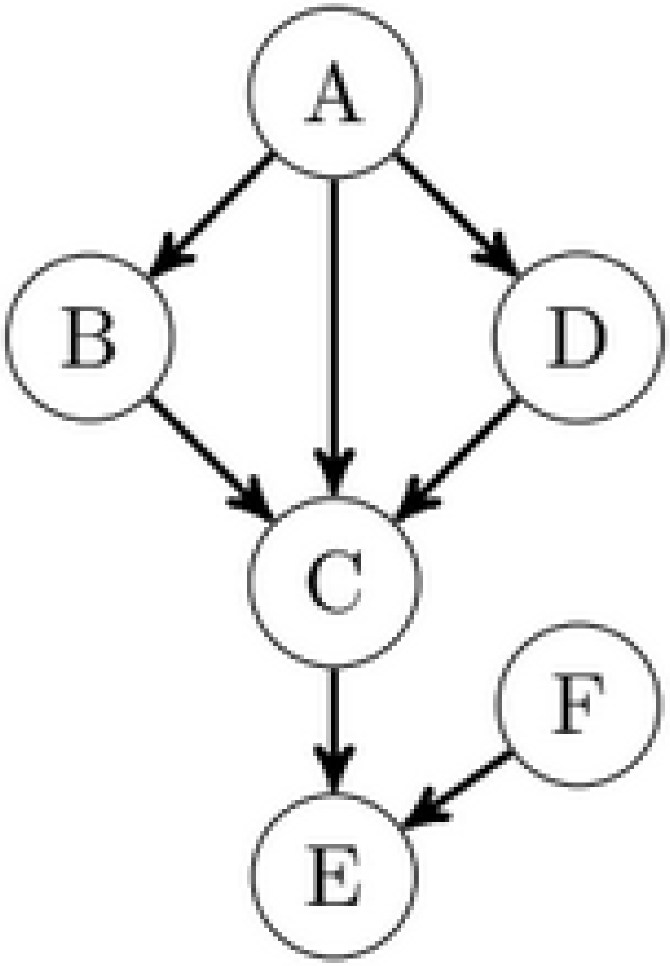
No

$s_{k+1} \perp s_{k-1}, s_{k-2} \ldots \mid s_k$

Why is this relevant?



More Complex Example



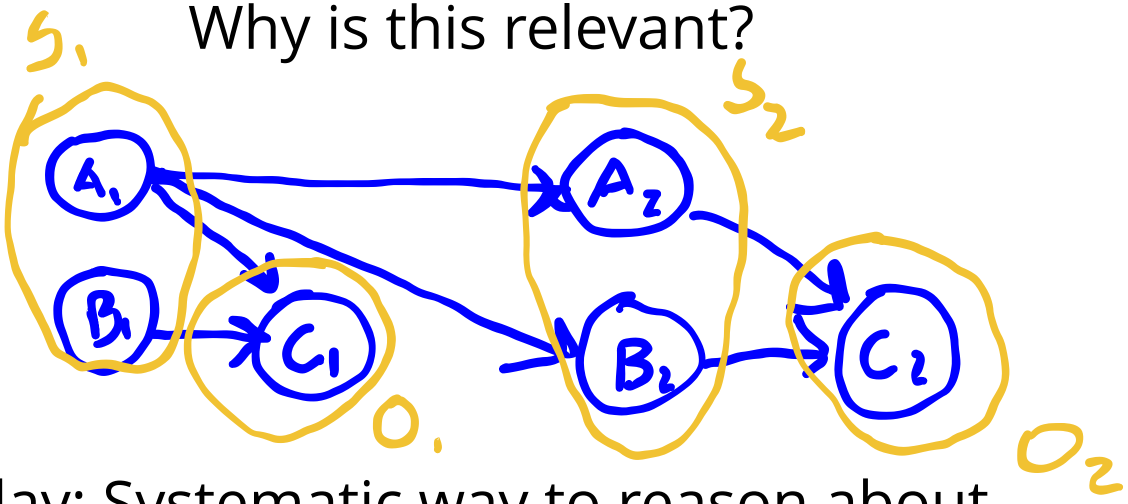
$(B \perp D \mid A)$?

Yes!

$(B \perp D \mid E)$?

No

Why is this relevant?



Today: Systematic way to reason about conditional independence

d-Separation

d-Separation

Let $\mathcal{C}$ be a set of random variables.

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated* by $\mathcal{C}$ if any of the following are true

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated* by $\mathcal{C}$ if any of the following are true

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated* by $\mathcal{C}$ if any of the following are true

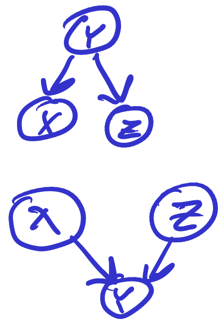
1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated* by $\mathcal{C}$ if any of the following are true

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$



d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated* by $\mathcal{C}$ if any of the following are true

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

We say that A and B are *d-separated* by $\mathcal{C}$ if all paths between A and B are d-separated by $\mathcal{C}$.

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated* by $\mathcal{C}$ if any of the following are true

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

We say that A and B are *d-separated* by $\mathcal{C}$ if all paths between A and B are d-separated by $\mathcal{C}$.

If A and B are d-separated by $\mathcal{C}$ then $A \perp B \mid \mathcal{C}$

Proving Conditional Independence

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

Proving Conditional Independence

1. Enumerate all paths between nodes in question

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

Proving Conditional Independence

1. Enumerate all paths between nodes in question
2. Check all paths for d-separation

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

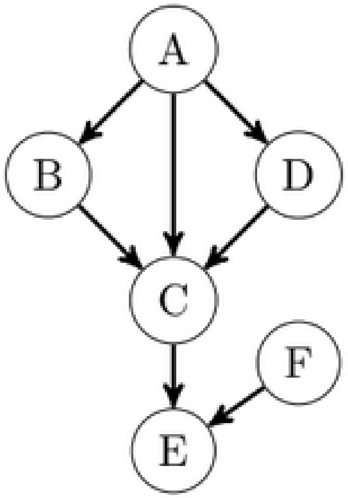
Proving Conditional Independence

1. Enumerate all paths between nodes in question
2. Check all paths for d-separation
3. If all paths d-separated, then CE

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

Proving Conditional Independence

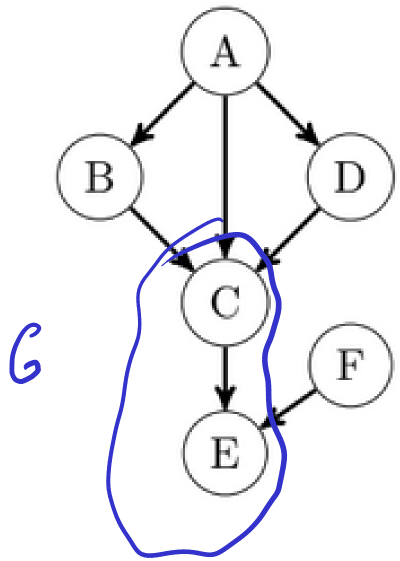
1. Enumerate all paths between nodes in question
2. Check all paths for d-separation
3. If all paths d-separated, then CE



1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

Proving Conditional Independence

1. Enumerate all paths between nodes in question
2. Check all paths for d-separation
3. If all paths d-separated, then CE



$\mathcal{C} = \{C, E\}$

No

Example: $(B \perp D \mid C, E)$?

$B \leftarrow A \rightarrow D$ rule 2: not d-separated

$B \rightarrow C \leftarrow D$ rule 3: not d-separated

$B \leftarrow A \rightarrow C \leftarrow D$ rule 2: not d-separated
rule 3: not d-separated

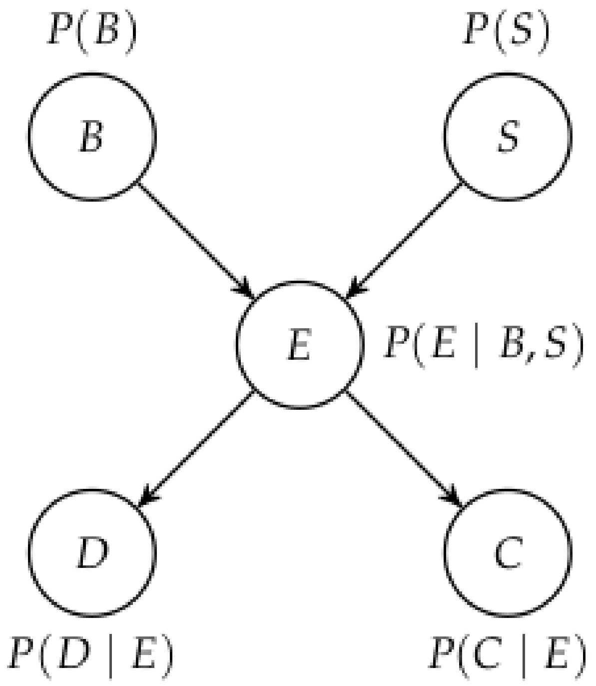
$B \rightarrow C \leftarrow A \rightarrow D$ rule 3: not d-separated
rule 2: not d-separated

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

Exercise

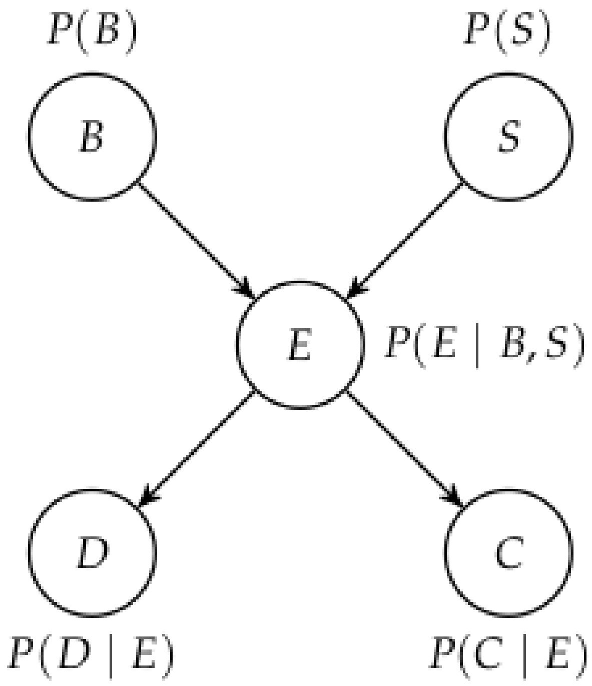
1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

Exercise



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

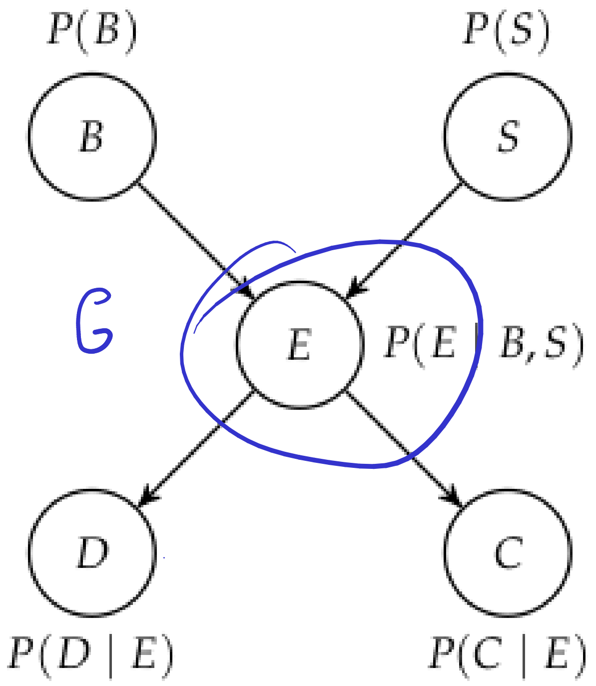


B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Exercise

$$D \perp C \mid B\,?$$

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that $Y \notin \mathcal{C}$



Exercise

$D \perp C \mid B$?

$D \perp C \mid E$? True

$D \leftarrow E \rightarrow C$ ✓ d-separated

B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$ ✓
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$

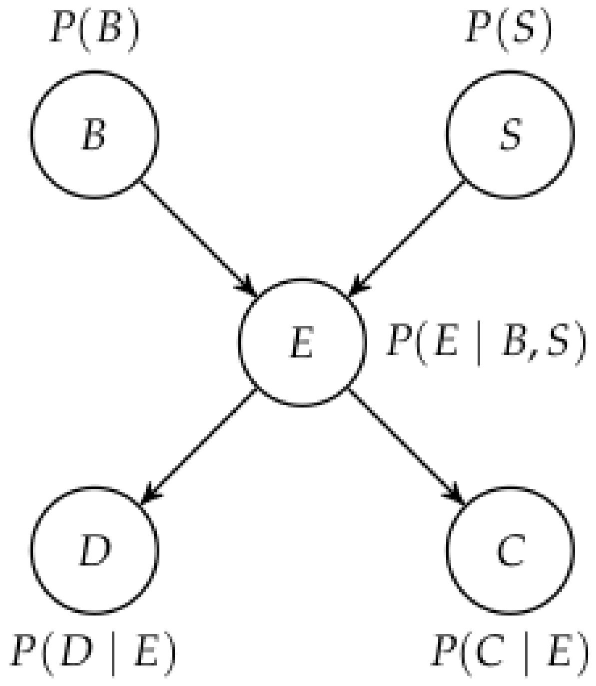
Sampling from a Bayesian Network

Sampling from a Bayesian Network

Given a Bayesian network, how do we sample from the joint distribution it defines?

Sampling from a Bayesian Network

Given a Bayesian network, how do we sample from the joint distribution it defines?

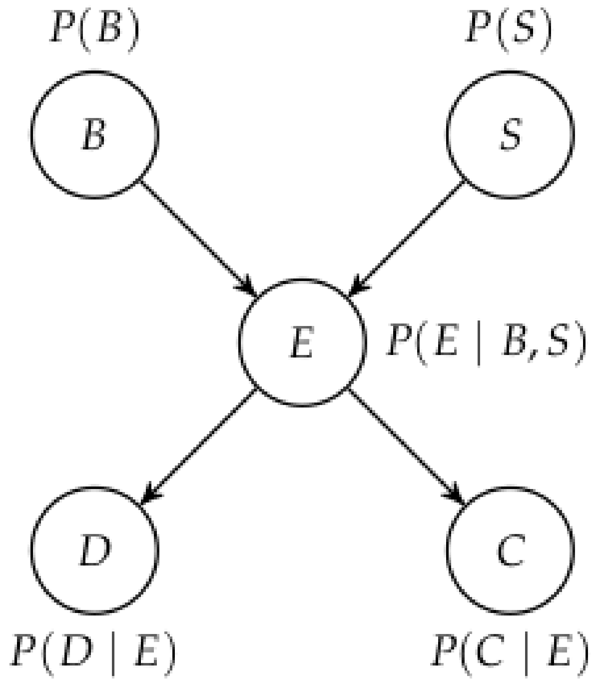


B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Sampling from a Bayesian Network

Given a Bayesian network, how do we sample from the joint distribution it defines?

1. Topoligical Sort (If there is an edge $A \to B$, then A comes before B)

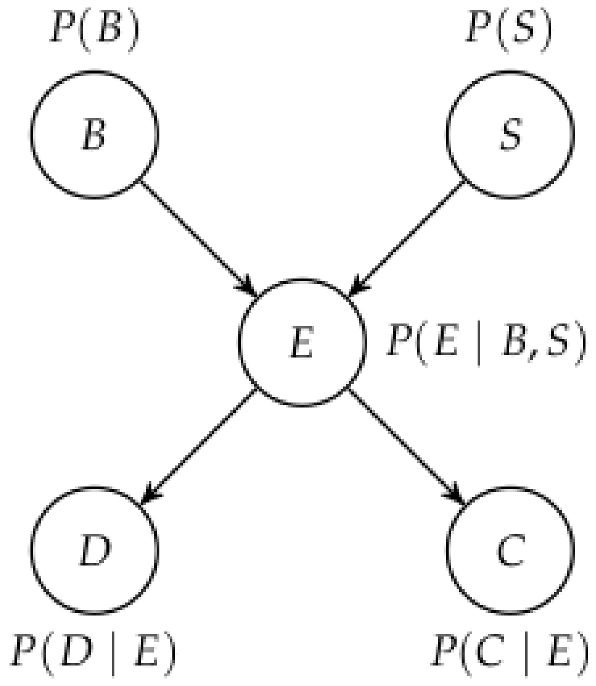


B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Sampling from a Bayesian Network

Given a Bayesian network, how do we sample from the joint distribution it defines?

1. Topoligical Sort (If there is an edge $A \to B$, then A comes before B)
2. Sample from conditional distributions in order of the topological sort



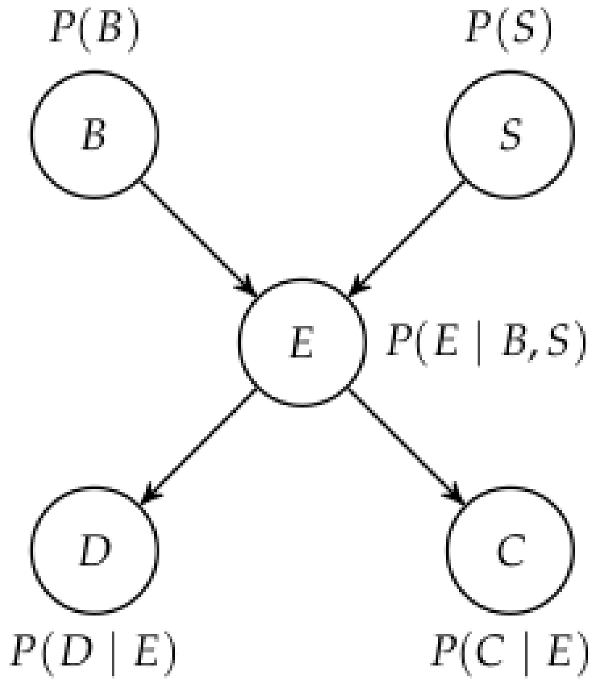
B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Sampling from a Bayesian Network

Given a Bayesian network, how do we sample from the joint distribution it defines?

1. Topoligical Sort (If there is an edge $A \rightarrow B$, then A comes before B)
2. Sample from conditional distributions in order of the topological sort

Analogous to **Simulating** a (PO)MDP



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Recap