UNIVERSITÄT ZU LÜBECK

# Graphical Methods for Finding Instrumental Variables

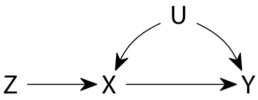
*Benito van der Zander*, Johannes Textor, Maciej Liśkiewicz

IM FOCUS DAS LEBEN

# Graphical Causal Model (Bayesian Network)

Graphical causal models represent random variables and their relationships in a graph.

Example:

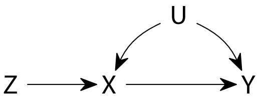



Random variables $X, Y, Z, U$ as nodes
Each variable is a function of (only) its parents

# Graphical Causal Model (Bayesian Network)

Graphical causal models represent random variables and their relationships in a graph.

Example:




Random variables $X, Y, Z, U$ as nodes
Each variable is a function of (only) its parents

Possible problems to research:

| | |
|---|---|
| **Simulation** | Graph + parameters $\Rightarrow$ data |
| **Learning** | Data $\Rightarrow$ graph + parameters |
| **Causal identification** | **Graph + data $\Rightarrow$ parameters** |

# D-separation

Graph reveals conditional independences.

- A path is a sequence of adjacent nodes.
- A node $V$ on a path is a *collider* if the adjacent edges are $\rightarrow V \leftarrow$.
- A path is *blocked* if it contains a collider.
- Two nodes are *d-separated* if all paths between them are blocked.
- Two nodes are *d-connected* if they are not *d*-separated.

# D-separation

Graph reveals conditional independences.

- A path is a sequence of adjacent nodes.
- A node $V$ on a path is a *collider* if the adjacent edges are $\rightarrow V \leftarrow$.
- A path is *blocked* if it contains a collider.
- Two nodes are *d-separated* if all paths between them are blocked.
- Two nodes are *d-connected* if they are not *d*-separated.

# D-separation

Graph reveals conditional independences.

- A path is a sequence of adjacent nodes.
- A node $V$ on a path is a *collider* if the adjacent edges are $\rightarrow V \leftarrow$.
- A path is *blocked* if it contains a collider.
  (i.e., only edges $\rightarrow V \rightarrow$, $\leftarrow V \leftarrow$, and $\leftarrow V \rightarrow$ are allowed)
- Two nodes are *d-separated* if all paths between them are blocked.
- Two nodes are *d-connected* if they are not $d$-separated.

# D-separation

Graph reveals conditional independences.

A path is a sequence of adjacent nodes.

A node $V$ on a path is a *collider* if the adjacent edges are $\rightarrow V \leftarrow$.

A path is *blocked* if it contains a collider.

(i.e., only edges $\rightarrow V \rightarrow$, $\leftarrow V \leftarrow$, and $\leftarrow V \rightarrow$ are allowed)

Two nodes are *d-separated* if all paths between them are blocked.

Two nodes are *d-connected* if they are not *d*-separated.

# D-separation

Graph reveals conditional independences.

A path is a sequence of adjacent nodes.

A node $V$ on a path is a *collider* if the adjacent edges are $\rightarrow V \leftarrow$.

A path is *blocked* if it contains a collider.

(i.e., only edges $\rightarrow V \rightarrow$, $\leftarrow V \leftarrow$, and $\leftarrow V \rightarrow$ are allowed)

Two nodes are *d-separated* if all paths between them are blocked.

Two nodes are *d-connected* if they are not *d*-separated.

(i.e., there exist any unblocked path between them)

# D-separation

Graph reveals conditional independences.

- A path is a sequence of adjacent nodes.
- A node $V$ on a path is a *collider* if the adjacent edges are $\to V \leftarrow$.
- A path is *blocked* if it contains a collider.
  (i.e., only edges $\to V \to$, $\leftarrow V \leftarrow$, and $\leftarrow V \to$ are allowed)
- Two nodes are *d-separated* if all paths between them are blocked.
- Two nodes are *d-connected* if they are not *d*-separated.
  (i.e., there exist any unblocked path between them)

Two random variables are correlated if and only if the corresponding nodes are *d*-connected.

Two random variables are independent if and only if the corresponding nodes are *d*-separated.

# D-separation

Let **W** be a set.

| | $V \notin \mathbf{W}$ | $V \in \mathbf{W}$ |
|---|---|---|
| $\rightarrow V \leftarrow$ | blocked | allowed |
| $\rightarrow V \rightarrow$ | allowed | blocked |
| $\leftarrow V \rightarrow$ | allowed | blocked |
| $\leftarrow V \leftarrow$ | allowed | blocked |

A path is blocked, if any node on it is blocked.

Two variables are independent if all paths between them are blocked .

# D-separation

Conditioning on a set **W** inverts the rules.

| | $V \notin \mathbf{W}$ | $V \in \mathbf{W}$ |
|---|---|---|
| $\rightarrow V \leftarrow$ | blocked<br>unless a descendant is in **W** | allowed |
| $\rightarrow V \rightarrow$ | allowed | blocked |
| $\leftarrow V \rightarrow$ | allowed | blocked |
| $\leftarrow V \leftarrow$ | allowed | blocked |

A path is blocked, if any node on it is blocked.

Two variables are conditionally independent given **W** if all paths between them are blocked given **W**.

# D-separation

Conditioning on a set **W** inverts the rules.

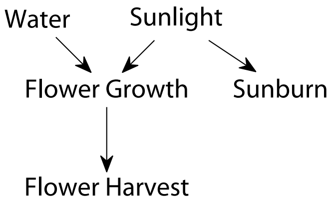
| | $V \notin \mathbf{W}$ | $V \in \mathbf{W}$ |
|---|---|---|
| $\rightarrow V \leftarrow$ | blocked | allowed |
| | unless a descendant is in **W** | |
| $\rightarrow V \rightarrow$ | allowed | blocked |
| $\leftarrow V \rightarrow$ | allowed | blocked |
| $\leftarrow V \leftarrow$ | allowed | blocked |

A path is blocked, if any node on it is blocked.

Two variables are conditionally independent given **W** if all paths between them are blocked given **W**.

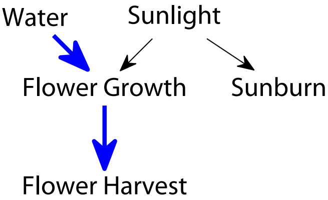
# D-separation

Example:



# D-separation

Example:

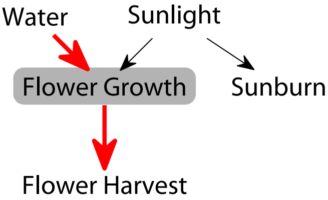



Water is d-connected to Flower Harvest

With more water, more flowers are harvested.

# D-separation

Example:




Water is d-connected to Flower Harvest

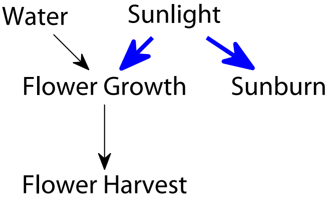
With more water, more flowers are harvested.

Water is d-separated from Flower Harvest given Flower Growth

Knowing the number of grown flowers, the water is uncorrelated to the harvest.

# D-separation

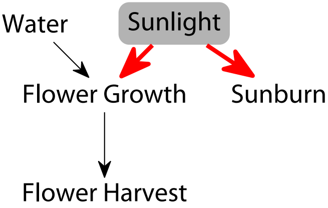
Example:




Flower Growth is d-connected to Sunburn

# D-separation

Example:

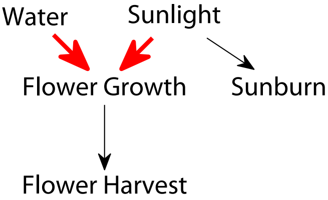



Flower Growth is d-connected to Sunburn

Flower Growth is d-separated from Sunburn given Sunlight

# D-separation

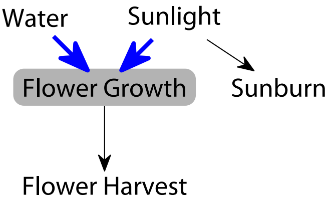
Example:




Water is *d*-separated from Sunlight

# D-separation

Example:




Water is *d*-separated from Sunlight

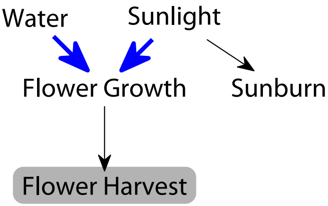
Water is *d*-connected to Sunlight given Flower Growth

There is water and flowers grow: There is sunlight

There is water and flowers do not grow: There is not enough sunlight

# D-separation

Example:




Water is *d*-separated from Sunlight

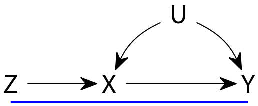
Water is *d*-connected to Sunlight given Flower Harvest

There is water and flowers are harvested: There is sunlight

There is water and flowers are not harvested: There is not enough sunlight

# D-separation

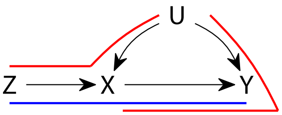
D-separation in the initial example:




*Z* and *Y* are d-connected due to path $Z \rightarrow X \rightarrow Y$

# D-separation

D-separation in the initial example:

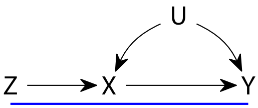



$Z$ and $Y$ are d-connected due to path $Z \rightarrow X \rightarrow Y$

$Z$ and $U$ are d-separated paths $Z \rightarrow X \leftarrow U$
$Z \rightarrow X \rightarrow Y \leftarrow U$ are blocked

# D-separation

D-separation in the initial example:



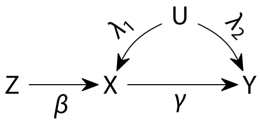

| | |
|---|---|
| $Z$ and $Y$ are d-connected | due to path $Z \rightarrow X \rightarrow Y$ |
| $Z$ and $U$ are d-separated | paths $Z \rightarrow X \leftarrow U$ |
| | $Z \rightarrow X \rightarrow Y \leftarrow U$ are blocked |

Conditioning on $X$ toggles collider $\rightarrow X \leftarrow$

$Z$ and $U$ are d-connected given $X$

# Structural Equation Model

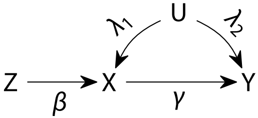
Assumption: All relationships are linear:




$$U := \varepsilon_U$$
$$Z := \varepsilon_Z$$
$$X := \beta Z + \lambda_1 U + \varepsilon_X$$
$$Y := \gamma X + \lambda_2 U + \varepsilon_Y$$

Random variables $X, Y, Z, U$, normalized to mean 0 and variance 1.
Independent random error terms $\varepsilon_X, \varepsilon_Y, \varepsilon_Z, \varepsilon_U$.
Parameters $\beta, \gamma, \lambda_1, \lambda_2$ are called the path coefficients or *causal effects*.

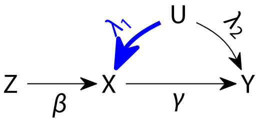
# Identification




$$U := \varepsilon_U$$
$$Z := \varepsilon_Z$$
$$X := \beta Z + \lambda_1 U + \varepsilon_X$$
$$Y := \gamma X + \lambda_2 U + \varepsilon_Y$$

Correlations are the sum over all *d*-connecting paths of the product of all path coefficients.

# Identification




$$U := \varepsilon_U$$
$$Z := \varepsilon_Z$$
$$X := \beta Z + \lambda_1 U + \varepsilon_X$$
$$Y := \gamma X + \lambda_2 U + \varepsilon_Y$$

Correlations are the sum over all *d*-connecting paths of the product of all path coefficients.

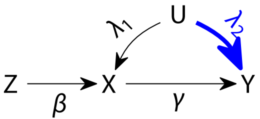
Generating:

$$\mathrm{Cov}(X,U) = \lambda_1$$
$$\mathrm{Cov}(Y,U) = \lambda_2$$
$$\mathrm{Cov}(Z,X) = \beta$$
$$\mathrm{Cov}(X,Y) = \gamma + \lambda_1\lambda_2$$
$$\mathrm{Cov}(Z,Y) = \beta\gamma$$

Causal identification:

$$\lambda_1 = \mathrm{Cov}(X,U)$$
$$\lambda_2 = \mathrm{Cov}(Y,U)$$
$$\beta = \mathrm{Cov}(Z,X)$$
$$\gamma = \mathrm{Cov}(X,Y) - \lambda_1\lambda_2$$
$$\gamma = \mathrm{Cov}(Y,Z)/\mathrm{Cov}(X,Z)$$

# Identification




$$U := \varepsilon_U$$
$$Z := \varepsilon_Z$$
$$X := \beta Z + \lambda_1 U + \varepsilon_X$$
$$Y := \gamma X + \lambda_2 U + \varepsilon_Y$$

Correlations are the sum over all *d*-connecting paths of the product of all path coefficients.

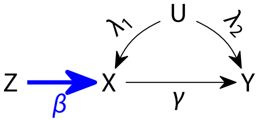
Generating:

$$\mathrm{Cov}(X,U) = \lambda_1$$
$$\mathrm{Cov}(Y,U) = \lambda_2$$
$$\mathrm{Cov}(Z,X) = \beta$$
$$\mathrm{Cov}(X,Y) = \gamma + \lambda_1\lambda_2$$
$$\mathrm{Cov}(Z,Y) = \beta\gamma$$

Causal identification:

$$\lambda_1 = \mathrm{Cov}(X,U)$$
$$\lambda_2 = \mathrm{Cov}(Y,U)$$
$$\beta = \mathrm{Cov}(Z,X)$$
$$\gamma = \mathrm{Cov}(X,Y) - \lambda_1\lambda_2$$
$$\gamma = \mathrm{Cov}(Y,Z)/\mathrm{Cov}(X,Z)$$

# Identification




$$U := \varepsilon_U$$
$$Z := \varepsilon_Z$$
$$X := \beta Z + \lambda_1 U + \varepsilon_X$$
$$Y := \gamma X + \lambda_2 U + \varepsilon_Y$$

Correlations are the sum over all *d*-connecting paths of the product of all path coefficients.

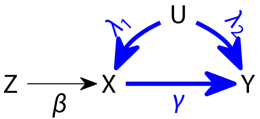
Generating:

$$\mathrm{Cov}(X,U) = \lambda_1$$
$$\mathrm{Cov}(Y,U) = \lambda_2$$
$$\mathrm{Cov}(Z,X) = \beta$$
$$\mathrm{Cov}(X,Y) = \gamma + \lambda_1\lambda_2$$
$$\mathrm{Cov}(Z,Y) = \beta\gamma$$

Causal identification:

$$\lambda_1 = \mathrm{Cov}(X,U)$$
$$\lambda_2 = \mathrm{Cov}(Y,U)$$
$$\beta = \mathrm{Cov}(Z,X)$$
$$\gamma = \mathrm{Cov}(X,Y) - \lambda_1\lambda_2$$
$$\gamma = \mathrm{Cov}(Y,Z)/\mathrm{Cov}(X,Z)$$

# Identification




$$U := \varepsilon_U$$
$$Z := \varepsilon_Z$$
$$X := \beta Z + \lambda_1 U + \varepsilon_X$$
$$Y := \gamma X + \lambda_2 U + \varepsilon_Y$$

Correlations are the sum over all *d*-connecting paths of the product of all path coefficients.

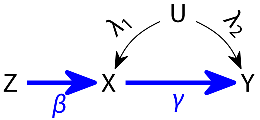
Generating:

$$\mathrm{Cov}(X,U) = \lambda_1$$
$$\mathrm{Cov}(Y,U) = \lambda_2$$
$$\mathrm{Cov}(Z,X) = \beta$$
$$\mathrm{Cov}(X,Y) = \gamma + \lambda_1\lambda_2$$
$$\mathrm{Cov}(Z,Y) = \beta\gamma$$

Causal identification:

$$\lambda_1 = \mathrm{Cov}(X,U)$$
$$\lambda_2 = \mathrm{Cov}(Y,U)$$
$$\beta = \mathrm{Cov}(Z,X)$$
$$\gamma = \mathrm{Cov}(X,Y) - \lambda_1\lambda_2$$
$$\gamma = \mathrm{Cov}(Y,Z)/\mathrm{Cov}(X,Z)$$

# Identification




$$U := \varepsilon_U$$
$$Z := \varepsilon_Z$$
$$X := \beta Z + \lambda_1 U + \varepsilon_X$$
$$Y := \gamma X + \lambda_2 U + \varepsilon_Y$$

Correlations are the sum over all *d*-connecting paths of the product of all path coefficients.

Generating:

$$\mathrm{Cov}(X, U) = \lambda_1$$
$$\mathrm{Cov}(Y, U) = \lambda_2$$
$$\mathrm{Cov}(Z, X) = \beta$$
$$\mathrm{Cov}(X, Y) = \gamma + \lambda_1\lambda_2$$
$$\mathrm{Cov}(Z, Y) = \beta\gamma$$

Causal identification:

$$\lambda_1 = \mathrm{Cov}(X, U)$$
$$\lambda_2 = \mathrm{Cov}(Y, U)$$
$$\beta = \mathrm{Cov}(Z, X)$$
$$\gamma = \mathrm{Cov}(X, Y) - \lambda_1\lambda_2$$
$$\gamma = \mathrm{Cov}(Y, Z)/\mathrm{Cov}(X, Z)$$

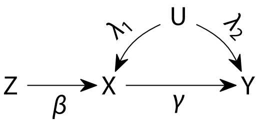
# Latent Variables

Partition variables:

**observed variables** covariances are known (set **M**)

**latent variables** covariances are unknown

Example: $\mathbf{M} = \{X, Y, Z\}$




$$U := \varepsilon_U$$
$$Z := \varepsilon_Z$$
$$X := \beta Z + \lambda_1 U + \varepsilon_X$$
$$Y := \gamma X + \lambda_2 U + \varepsilon_Y$$

Generating:

$$\mathrm{Cov}(X, Y) = \gamma + \lambda_1 \lambda_2$$
$$\mathrm{Cov}(Z, X) = \beta$$
$$\mathrm{Cov}(Z, Y) = \beta\gamma$$

Causal identification:

$$\beta = \mathrm{Cov}(Z, X)$$
$$\gamma = \mathrm{Cov}(X, Y) - \lambda_1 \lambda_2$$
$$\gamma = \mathrm{Cov}(Y, Z)/\mathrm{Cov}(X, Z)$$

# Instrumental variable

**Definition (Instrumental variable)**

Variable $Z$ is an *instrument* relative to edge $X \to Y$, if

(a) $Z$ and $X$ are $d$-connected,

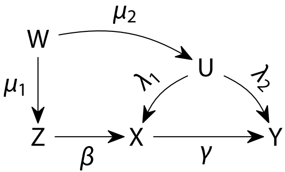
(b) $Z$ and $Y$ are $d$-separated after removal of the edge $X \to Y$,

Given a instrument $Z$ the causal effect of $X \to Y$ is

$$\frac{\mathrm{Cov}(Y, Z)}{\mathrm{Cov}(X, Z)}$$

# Conditional instrumental variables

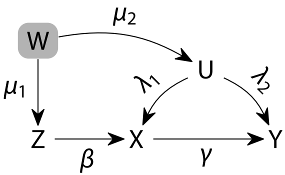
Often no instrumental variable exists:




Condition $Z$ d-separated from $Y$ (after removal of the edge $X \to Y$) cannot be satisfied because of the path $Z \leftarrow W \to U \to Y$.

# Conditional instrumental variables

Often no instrumental variable exists:




Condition $Z$ d-separated from $Y$ (after removal of the edge $X \to Y$) cannot be satisfied because of the path $Z \leftarrow W \to U \to Y$.

Thus change condition to „$Z$ d-separated from $Y$ given $\mathbf{W}$“ for some set $\mathbf{W}$

# Conditional instrumental variables

**Definition (Conditional instrumental variable)**

Variable $Z$ is a *conditional instrument* relative to edge $X \to Y$, if there exists a set $\mathbf{W} \subseteq \mathbf{M}$ such that

(a) $\mathbf{W}$ does not *d*-separate $Z$ and $X$,

(b) $\mathbf{W}$ *d*-separates $Z$ and $Y$ after removal of the edge $X \to Y$,

(c) $\mathbf{W}$ consists of non-descendants of $Y$.

Given a conditional instrument $Z$ the causal effect of $X \to Y$ is

$$\frac{\mathrm{Cov}(Y, Z|\mathbf{W})}{\mathrm{Cov}(X, Z|\mathbf{W})}$$

# Conditional instrumental variables

**Definition (Conditional instrumental variable)**

Variable $Z$ is a *conditional instrument* relative to edge $X \to Y$, if there exists a set $\mathbf{W} \subseteq \mathbf{M}$ such that

(a) $\mathbf{W}$ does not *d*-separate $Z$ and $X$,

(b) $\mathbf{W}$ *d*-separates $Z$ and $Y$ after removal of the edge $X \to Y$,

(c) $\mathbf{W}$ consists of non-descendants of $Y$.

Special cases:

**instrumental variable** $\mathbf{W} = \emptyset$

**ancestral instrumental variable** $\mathbf{W} \subseteq$ Ancestors of $Y$ and $Z$

# NP-completeness

**Theorem**

*Determining if, for given $X, Y, Z \in \mathbf{M}$, node $Z$ is a conditional instrumental variable relative to $X \to Y$ is NP-complete.*

Proof approach: Reduction of 3-SAT to finding the set **W**.

Construct graph with one path between $X$ and $Z$ with a collider for each clause that has 3 descendants.
Choosing a descendant to unblock collider corresponds to choosing satisfied literal.
Choosing too many descendants opens a path between $X$ and $Y$.

# Ancestral instrumental variable

**Theorem**

*Determining if, for given $X, Y, Z \in \mathbf{M}$, node $Z$ is an ancestral instrumental variable relative to $X \to Y$ can be done in $\mathcal{O}(n + m)$ time.*

Need efficient algorithm to find $\mathbf{W} \subseteq$ ancestors of $Y, Z$

# Nearest Separator

**Definition**

A set $\mathbf{W} \subseteq \mathbf{M} \setminus \{Y, Z\}$ is called a *nearest separator* relative to $Y$ and $Z$ if

(a) $\mathbf{W}$ *d*-separates $Y$ and $Z$, and

(b) for any $W \in \mathbf{W}$ and any set $\mathbf{W}' \subseteq \mathbf{M} \setminus \{W, Y, Z\}$ that *d*-separates $Y$ and $Z$, it holds:
$\mathbf{W}'$ *d*-separates $W$ and $Z$.

# Nearest Separator

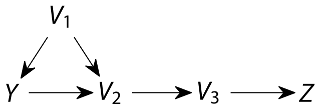
**Definition**

A set $\mathbf{W} \subseteq \mathbf{M} \setminus \{Y, Z\}$ is called a *nearest separator* relative to $Y$ and $Z$ if

(a) $\mathbf{W}$ *d*-separates $Y$ and $Z$, and

(b) for any $W \in \mathbf{W}$ and any set $\mathbf{W}' \subseteq \mathbf{M} \setminus \{W, Y, Z\}$ that *d*-separates $Y$ and $Z$, it holds:
$\mathbf{W}'$ *d*-separates $W$ and $Z$.

Example:



# Nearest Separator

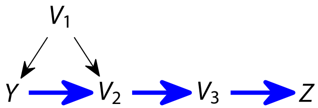
**Definition**

A set $\mathbf{W} \subseteq \mathbf{M} \setminus \{Y, Z\}$ is called a *nearest separator* relative to $Y$ and $Z$ if

(a) $\mathbf{W}$ *d*-separates $Y$ and $Z$, and

(b) for any $W \in \mathbf{W}$ and any set $\mathbf{W}' \subseteq \mathbf{M} \setminus \{W, Y, Z\}$ that *d*-separates $Y$ and $Z$, it holds:
$\mathbf{W}'$ *d*-separates $W$ and $Z$.

Example:



# Nearest Separator

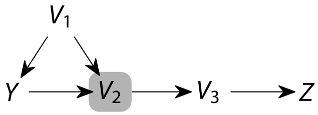
**Definition**

A set $\mathbf{W} \subseteq \mathbf{M} \setminus \{Y, Z\}$ is called a *nearest separator* relative to $Y$ and $Z$ if

(a) $\mathbf{W}$ *d*-separates $Y$ and $Z$, and

(b) for any $W \in \mathbf{W}$ and any set $\mathbf{W}' \subseteq \mathbf{M} \setminus \{W, Y, Z\}$ that *d*-separates $Y$ and $Z$, it holds:
$\mathbf{W}'$ *d*-separates $W$ and $Z$.

Example:




Nearest separator: $\mathbf{W} = \{V_2\}$

# Nearest Separator

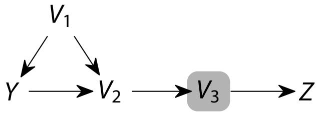
## Definition

A set $\mathbf{W} \subseteq \mathbf{M} \setminus \{Y, Z\}$ is called a *nearest separator* relative to $Y$ and $Z$ if

(a) $\mathbf{W}$ $d$-separates $Y$ and $Z$, and

(b) for any $W \in \mathbf{W}$ and any set $\mathbf{W}' \subseteq \mathbf{M} \setminus \{W, Y, Z\}$ that $d$-separates $Y$ and $Z$, it holds:
$\mathbf{W}'$ $d$-separates $W$ and $Z$.

Example:




Nearest separator: $\mathbf{W} = \{V_2\}$
Not nearest separator: $\mathbf{W}' = \{V_3\}$
Since $\{V_2\}$ does not $d$-separate $V_3$ and $Z$.

# Nearest Separator

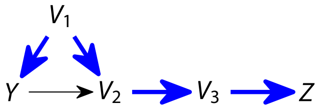
**Definition**

A set $\mathbf{W} \subseteq \mathbf{M} \setminus \{Y, Z\}$ is called a *nearest separator* relative to $Y$ and $Z$ if

(a) $\mathbf{W}$ $d$-separates $Y$ and $Z$, and

(b) for any $W \in \mathbf{W}$ and any set $\mathbf{W}' \subseteq \mathbf{M} \setminus \{W, Y, Z\}$ that $d$-separates $Y$ and $Z$, it holds:
$\mathbf{W}'$ $d$-separates $W$ and $Z$.

Example:




Nearest separator: $\mathbf{W} = \{V_2\}$
Alternative nearest separator: $\mathbf{W}'' = \{\ldots$

# Nearest Separator

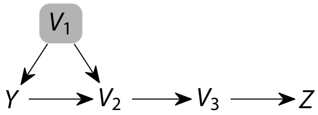
**Definition**

A set $\mathbf{W} \subseteq \mathbf{M} \setminus \{Y, Z\}$ is called a *nearest separator* relative to $Y$ and $Z$ if

(a) $\mathbf{W}$ *d*-separates $Y$ and $Z$, and

(b) for any $W \in \mathbf{W}$ and any set $\mathbf{W}' \subseteq \mathbf{M} \setminus \{W, Y, Z\}$ that *d*-separates $Y$ and $Z$, it holds:
$\mathbf{W}'$ *d*-separates $W$ and $Z$.

Example:




Nearest separator: $\mathbf{W} = \{V_2\}$
Alternative nearest separator: $\mathbf{W}'' = \{V_1,$

# Nearest Separator

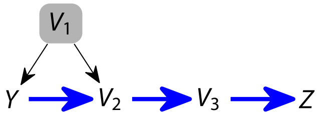
## Definition

A set $\mathbf{W} \subseteq \mathbf{M} \setminus \{Y, Z\}$ is called a *nearest separator* relative to $Y$ and $Z$ if

(a) $\mathbf{W}$ *d*-separates $Y$ and $Z$, and

(b) for any $W \in \mathbf{W}$ and any set $\mathbf{W}' \subseteq \mathbf{M} \setminus \{W, Y, Z\}$ that *d*-separates $Y$ and $Z$, it holds:
$\mathbf{W}'$ *d*-separates $W$ and $Z$.

Example:




Nearest separator: $\mathbf{W} = \{V_2\}$
Alternative nearest separator: $\mathbf{W}'' = \{V_1,$

# Nearest Separator

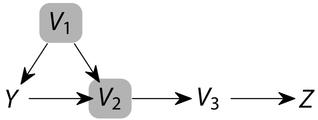
## Definition

A set $\mathbf{W} \subseteq \mathbf{M} \setminus \{Y, Z\}$ is called a *nearest separator* relative to $Y$ and $Z$ if

(a) $\mathbf{W}$ *d*-separates $Y$ and $Z$, and

(b) for any $W \in \mathbf{W}$ and any set $\mathbf{W}' \subseteq \mathbf{M} \setminus \{W, Y, Z\}$ that *d*-separates $Y$ and $Z$, it holds:
$\mathbf{W}'$ *d*-separates $W$ and $Z$.

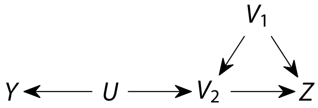
Example:




Nearest separator: $\mathbf{W} = \{V_2\}$
Alternative nearest separator: $\mathbf{W}'' = \{V_1, V_2\}$

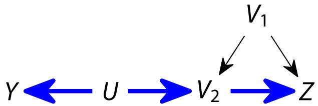
# Nearest Separator

Example:
DAG with unobserved variable $U$



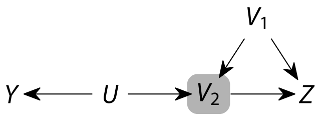
# Nearest Separator

Example:
DAG with unobserved variable $U$



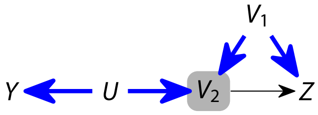
# Nearest Separator

Example:
DAG with unobserved variable $U$



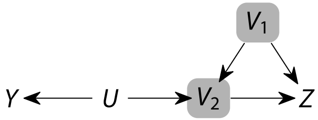
# Nearest Separator

Example:
DAG with unobserved variable $U$



# Nearest Separator

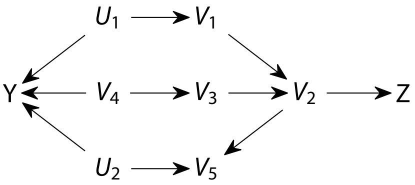
Example:
DAG with unobserved variable $U$




Nearest separator $\mathbf{W} = \{V_1, V_2\}$

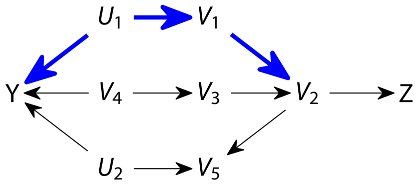
# Nearest Separator

Example:
DAG with unobserved variable $U_1$, $U_2$



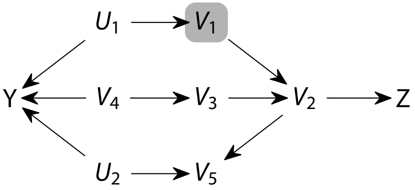
# Nearest Separator

Example:
DAG with unobserved variable $U_1$, $U_2$



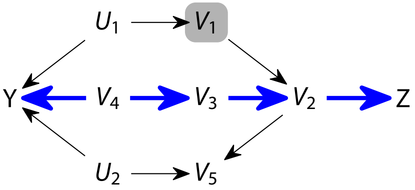
# Nearest Separator

Example:
DAG with unobserved variable $U_1$, $U_2$



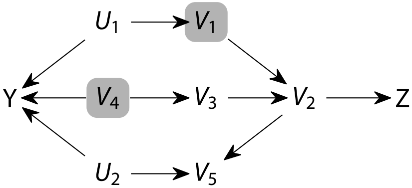
# Nearest Separator

Example:
DAG with unobserved variable $U_1$, $U_2$



# Nearest Separator

Example:
DAG with unobserved variable $U_1$, $U_2$



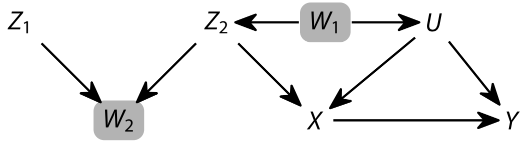

Nearest separator $\mathbf{W} = \{V_1, V_4\}$

# Ancestral Instrumental Variable

**Theorem**

*For given variables $X, Y \in \mathbf{M}$, an ancestral instrumental variable $Z \in \mathbf{M}$ relative to $X \to Y$ exists iff a conditional instrumental variable $Z' \in \mathbf{M}$ relative to $X \to Y$ exists.*

Example:




Conditional instrumental variable: $Z_1$ with $\mathbf{W} = \{W_1, W_2\}$

Ancestral instrumental variable: $Z_2$ with $\mathbf{W} = \{W_1\}$

# Conclusion

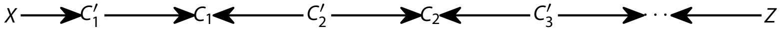
In a linear causal model the causal effect of variable $X$ on variable $Y$ can be calculated from the correlations between $X$ and $Y$ using a conditional instrument.

If a conditional instrument $Z$ exists, a conditional instrument can be found efficiently in $O(n(n+m))$.

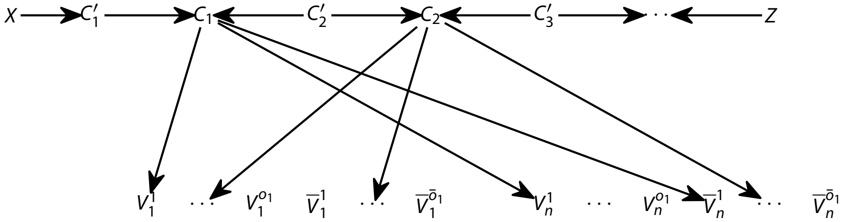
But testing if a given $Z$ is a conditional instrument is NP-complete.

# Appendix

Let $C_1, \ldots, C_m$ be clauses of a 3-SAT-instance with $V_1, \ldots, V_n$ variables. A $V_i$ occurs in $o_i$ different clauses.

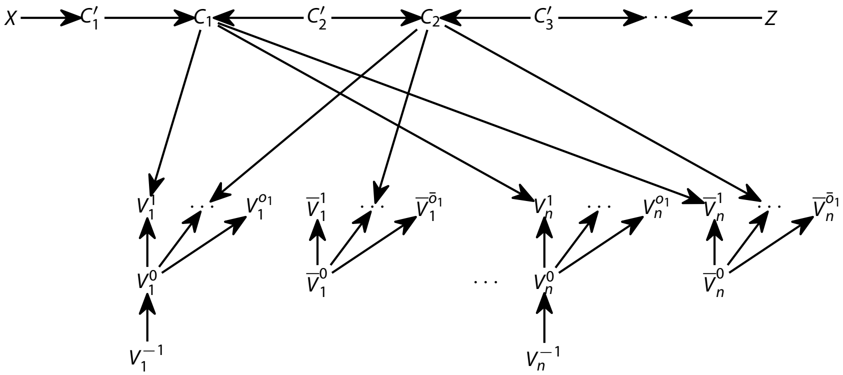


Let $C_1, \ldots, C_m$ be clauses of a 3-SAT-instance with $V_1, \ldots, V_n$ variables. A $V_i$ occurs in $o_i$ different clauses.




...

Let $C_1, \ldots, C_m$ be clauses of a 3-SAT-instance with $V_1, \ldots, V_n$ variables. A $V_i$ occurs in $o_i$ different clauses.



Let $C_1, \ldots, C_m$ be clauses of a 3-SAT-instance with $V_1, \ldots, V_n$ variables. A $V_i$ occurs in $o_i$ different clauses.