# Causation as Production and Dependence

## or, A Model-Invariant Theory of Causation

J. Dmitri Gallow

University of North Carolina, Chapel Hill · February 9th, 2018

**Please interrupt when I stop making sense.**

# Table of contents

# Causal Models

## Causal Models

A causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a 5-tuple of

- A vector, $\mathbb{U} = (U_1, U_2, \ldots, U_M)$, of *exogenous* variables;
- An assignment of values, $\vec{u} = (u_1, u_2, \ldots, u_M)$, to $\mathbb{U}$;
- A vector $\mathbb{V} = (V_1, V_2, \ldots, V_N)$, of *endogenous* variables; and
- A vector $\mathbb{E} = (\phi_{V_1}, \phi_{V_2}, \ldots, \phi_{V_N})$ of *structural equations*, one for each endogenous variable $V_i \in \mathbb{V}$.
- A specification, $\mathscr{D}$, of which variable values are *default, normal, or inertial*, and which values are *deviations* therefrom.

## Causal Models

A causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a 5-tuple of

- A vector, $\mathbb{U} = (U_1, U_2, \ldots, U_M)$, of *exogenous* variables;
- An assignment of values, $\vec{u} = (u_1, u_2, \ldots, u_M)$, to $\mathbb{U}$;
- A vector $\mathbb{V} = (V_1, V_2, \ldots, V_N)$, of *endogenous* variables; and
- A vector $\mathbb{E} = (\phi_{V_1}, \phi_{V_2}, \ldots, \phi_{V_N})$ of *structural equations*, one for each endogenous variable $V_i \in \mathbb{V}$.
- A specification, $\mathscr{D}$, of which variable values are *default, normal, or inertial*, and which values are *deviations* therefrom.

## Causal Models

A causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a 5-tuple of

- A vector, $\mathbb{U} = (U_1, U_2, \ldots, U_M)$, of *exogenous* variables;
- An assignment of values, $\vec{u} = (u_1, u_2, \ldots, u_M)$, to $\mathbb{U}$;
- A vector $\mathbb{V} = (V_1, V_2, \ldots, V_N)$, of *endogenous* variables; and
- A vector $\mathbb{E} = (\phi_{V_1}, \phi_{V_2}, \ldots, \phi_{V_N})$ of *structural equations*, one for each endogenous variable $V_i \in \mathbb{V}$.
- A specification, $\mathscr{D}$, of which variable values are *default, normal, or inertial*, and which values are *deviations* therefrom.

## Causal Models

A causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a 5-tuple of

- A vector, $\mathbb{U} = (U_1, U_2, \ldots, U_M)$, of *exogenous* variables;
- An assignment of values, $\vec{u} = (u_1, u_2, \ldots, u_M)$, to $\mathbb{U}$;
- A vector $\mathbb{V} = (V_1, V_2, \ldots, V_N)$, of *endogenous* variables; and
- A vector $\mathbb{E} = (\phi_{V_1}, \phi_{V_2}, \ldots, \phi_{V_N})$ of *structural equations*, one for each endogenous variable $V_i \in \mathbb{V}$.
- A specification, $\mathscr{D}$, of which variable values are *default, normal, or inertial*, and which values are *deviations* therefrom.

## Causal Models

A causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a 5-tuple of

- A vector, $\mathbb{U} = (U_1, U_2, \ldots, U_M)$, of *exogenous* variables;
- An assignment of values, $\vec{u} = (u_1, u_2, \ldots, u_M)$, to $\mathbb{U}$;
- A vector $\mathbb{V} = (V_1, V_2, \ldots, V_N)$, of *endogenous* variables; and
- A vector $\mathbb{E} = (\phi_{V_1}, \phi_{V_2}, \ldots, \phi_{V_N})$ of *structural equations*, one for each endogenous variable $V_i \in \mathbb{V}$.
- A specification, $\mathscr{D}$, of which variable values are *default, normal, or inertial*, and which values are *deviations* therefrom.

## Causal Models

A causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a 5-tuple of

- A vector, $\mathbb{U} = (U_1, U_2, \ldots, U_M)$, of *exogenous* variables;
- An assignment of values, $\vec{u} = (u_1, u_2, \ldots, u_M)$, to $\mathbb{U}$;
- A vector $\mathbb{V} = (V_1, V_2, \ldots, V_N)$, of *endogenous* variables; and
- A vector $\mathbb{E} = (\phi_{V_1}, \phi_{V_2}, \ldots, \phi_{V_N})$ of *structural equations*, one for each endogenous variable $V_i \in \mathbb{V}$.
- A specification, $\mathscr{D}$, of which variable values are *default, normal, or inertial*, and which values are *deviations* therefrom.

# Causal Models

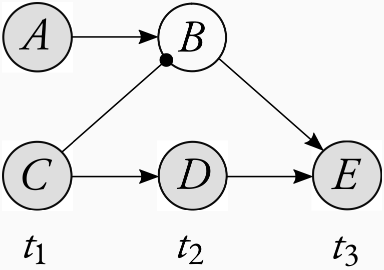



**Figure 1:** Preemptive Overdetermination

# Causal Models

The causal model $\mathbb{M}_1$:

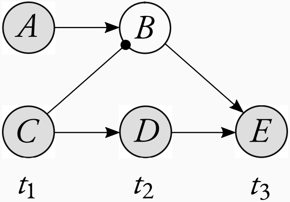
$\mathbb{U} : (A, C)$

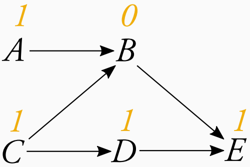
$\vec{u} : (1, 1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee D \\ D & := C \\ B & := A \wedge \neg C \end{pmatrix}$$

$\mathscr{D}$ : 0 is default, 1 is deviant



# Causal Models

The causal model $\mathbb{M}_1$:

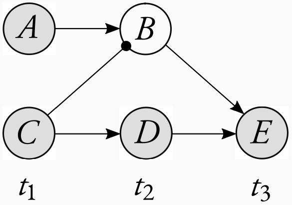
$\mathbb{U} : (A, C)$

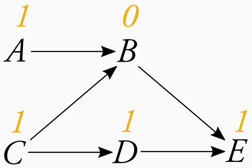
$\vec{u} : (1, 1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee D \\ D & := C \\ B & := A \wedge \neg C \end{pmatrix}$$

$\mathscr{D}$ : 0 is default, 1 is deviant



# Counterfactuals in Causal Models

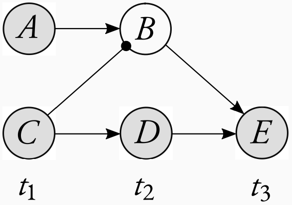
The causal model $\mathbb{M}_1$:

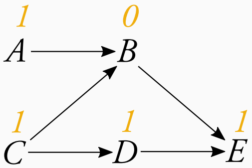
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E := B \vee D \\ D := C \\ B := A \wedge \neg C \end{pmatrix}$$



# Counterfactuals in Causal Models

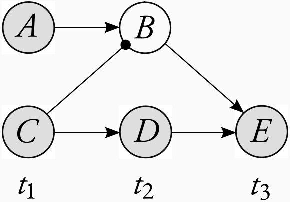
The causal model $\mathbb{M}_1[D \rightarrow 0]$:

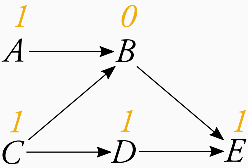
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee D \\ D & := C \\ B & := A \wedge \neg C \end{pmatrix}$$



# Counterfactuals in Causal Models

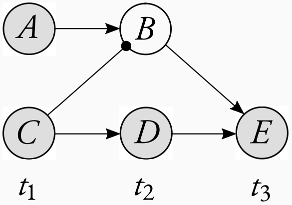
The causal model $\mathbb{M}_1[D \to 0]$:

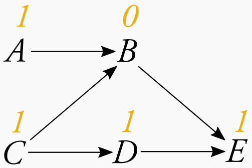
$\mathbb{U} : (A, C, D)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee D \\ D & := C \\ B & := A \wedge \neg C \end{pmatrix}$$



# Counterfactuals in Causal Models

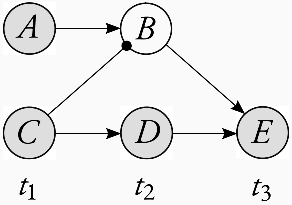
The causal model $\mathbb{M}_1[D \to 0]$:

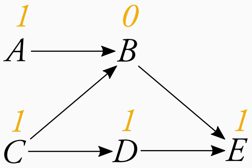
$$\mathbb{U} : (A, C, D)$$

$$\vec{u} : (1, 1, 0)$$

$$\mathbb{V} : (B, E)$$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee D \\ D & := C \\ B & := A \wedge \neg C \end{pmatrix}$$



# Counterfactuals in Causal Models

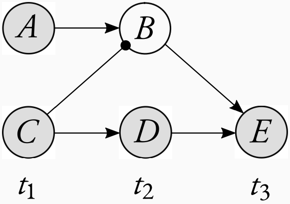
The causal model $\mathbb{M}_1[D \to 0]$:

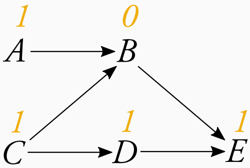
$$\mathbb{U} : (A, C, D)$$

$$\vec{u} : (1, 1, 0)$$

$$\mathbb{V} : (B, E)$$

$$\mathbb{E} : \left( \begin{array}{ll} E & := B \vee D \\ B & := A \wedge \neg C \end{array} \right)$$



# Counterfactuals in Causal Models

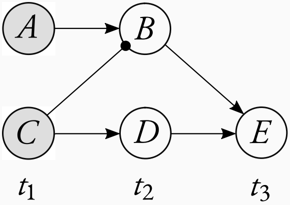
The causal model $\mathbb{M}_1[D \to 0]$:

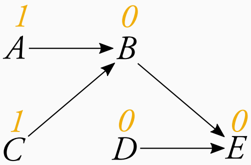
$\mathbb{U} : (A, C, D)$

$\vec{u} : (1, 1, 0)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee D \\ B & := A \wedge \neg C \end{pmatrix}$$



## Counterfactual Models

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, including the variables $\mathbf{V}$, and given the assignment of values $\mathbf{v}$ to $\mathbf{V}$, the counterfactual model $\mathbb{M}[\mathbf{V} \to \mathbf{v}] = (\mathbb{U}[\mathbf{V} \to \mathbf{v}], \vec{u}[\mathbf{V} \to \mathbf{v}], \mathbb{V}[\mathbf{V} \to \mathbf{v}], \mathbb{E}[\mathbf{V} \to \mathbf{v}], \mathscr{D}[\mathbf{V} \to \mathbf{v}])$ is the model such that:

- $\mathbb{V}[\mathbf{V} \to \mathbf{v}] = \mathbb{V} - \mathbf{V}$
- $\mathbb{U}[\mathbf{V} \to \mathbf{v}] = \mathbb{U} \cup \mathbf{V}$
- $\mathbb{E}[\mathbf{V} \to \mathbf{v}] = \mathbb{E} - (\phi_{V_i} \mid V_i \in \mathbf{V})$
- $\vec{u}[\mathbf{V} \to \mathbf{v}] = \vec{u} \cup \mathbf{v}$
- $\mathscr{D}[\mathbf{V} \to \mathbf{v}] = \mathscr{D}$

# Counterfactuals in Causal Models

## Counterfactual Models

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, including the variables $\mathbf{V}$, and given the assignment of values $\mathbf{v}$ to $\mathbf{V}$, the counterfactual model $\mathbb{M}[\mathbf{V} \rightarrow \mathbf{v}] = (\mathbb{U}[\mathbf{V} \rightarrow \mathbf{v}], \vec{u}[\mathbf{V} \rightarrow \mathbf{v}], \mathbb{V}[\mathbf{V} \rightarrow \mathbf{v}], \mathbb{E}[\mathbf{V} \rightarrow \mathbf{v}], \mathscr{D}[\mathbf{V} \rightarrow \mathbf{v}])$ is the model such that:

- $\mathbb{V}[\mathbf{V} \rightarrow \mathbf{v}] = \mathbb{V} - \mathbf{V}$
- $\mathbb{U}[\mathbf{V} \rightarrow \mathbf{v}] = \mathbb{U} \cup \mathbf{V}$
- $\mathbb{E}[\mathbf{V} \rightarrow \mathbf{v}] = \mathbb{E} - (\phi_{V_i} \mid V_i \in \mathbf{V})$
- $\vec{u}[\mathbf{V} \rightarrow \mathbf{v}] = \vec{u} \cup \mathbf{v}$
- $\mathscr{D}[\mathbf{V} \rightarrow \mathbf{v}] = \mathscr{D}$

# Counterfactuals in Causal Models

## Counterfactual Models

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, including the variables $\mathbf{V}$, and given the assignment of values $\mathbf{v}$ to $\mathbf{V}$, the counterfactual model $\mathbb{M}[\mathbf{V} \to \mathbf{v}] = (\mathbb{U}[\mathbf{V} \to \mathbf{v}], \vec{u}[\mathbf{V} \to \mathbf{v}], \mathbb{V}[\mathbf{V} \to \mathbf{v}], \mathbb{E}[\mathbf{V} \to \mathbf{v}], \mathscr{D}[\mathbf{V} \to \mathbf{v}])$ is the model such that:

- $\mathbb{V}[\mathbf{V} \to \mathbf{v}] = \mathbb{V} - \mathbf{V}$
- $\mathbb{U}[\mathbf{V} \to \mathbf{v}] = \mathbb{U} \cup \mathbf{V}$
- $\mathbb{E}[\mathbf{V} \to \mathbf{v}] = \mathbb{E} - (\phi_{V_i} \mid V_i \in \mathbf{V})$
- $\vec{u}[\mathbf{V} \to \mathbf{v}] = \vec{u} \cup \mathbf{v}$
- $\mathscr{D}[\mathbf{V} \to \mathbf{v}] = \mathscr{D}$

# Counterfactuals in Causal Models

## Counterfactual Models

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, including the variables $\mathbf{V}$, and given the assignment of values $\mathbf{v}$ to $\mathbf{V}$, the counterfactual model $\mathbb{M}[\mathbf{V} \to \mathbf{v}] = (\mathbb{U}[\mathbf{V} \to \mathbf{v}], \vec{u}[\mathbf{V} \to \mathbf{v}], \mathbb{V}[\mathbf{V} \to \mathbf{v}], \mathbb{E}[\mathbf{V} \to \mathbf{v}],$ $\mathscr{D}[\mathbf{V} \to \mathbf{v}])$ is the model such that:

- $\mathbb{V}[\mathbf{V} \to \mathbf{v}] = \mathbb{V} - \mathbf{V}$
- $\mathbb{U}[\mathbf{V} \to \mathbf{v}] = \mathbb{U} \cup \mathbf{V}$
- $\mathbb{E}[\mathbf{V} \to \mathbf{v}] = \mathbb{E} - (\phi_{V_i} \mid V_i \in \mathbf{V})$
- $\vec{u}[\mathbf{V} \to \mathbf{v}] = \vec{u} \cup \mathbf{v}$
- $\mathscr{D}[\mathbf{V} \to \mathbf{v}] = \mathscr{D}$

# Counterfactuals in Causal Models

## Counterfactual Models

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, including the variables $\mathbf{V}$, and given the assignment of values $\mathbf{v}$ to $\mathbf{V}$, the counterfactual model $\mathbb{M}[\mathbf{V} \rightarrow \mathbf{v}] = (\mathbb{U}[\mathbf{V} \rightarrow \mathbf{v}], \vec{u}[\mathbf{V} \rightarrow \mathbf{v}], \mathbb{V}[\mathbf{V} \rightarrow \mathbf{v}], \mathbb{E}[\mathbf{V} \rightarrow \mathbf{v}], \mathscr{D}[\mathbf{V} \rightarrow \mathbf{v}])$ is the model such that:

- $\mathbb{V}[\mathbf{V} \rightarrow \mathbf{v}] = \mathbb{V} - \mathbf{V}$
- $\mathbb{U}[\mathbf{V} \rightarrow \mathbf{v}] = \mathbb{U} \cup \mathbf{V}$
- $\mathbb{E}[\mathbf{V} \rightarrow \mathbf{v}] = \mathbb{E} - (\phi_{V_i} \mid V_i \in \mathbf{V})$
- $\vec{u}[\mathbf{V} \rightarrow \mathbf{v}] = \vec{u} \cup \mathbf{v}$
- $\mathscr{D}[\mathbf{V} \rightarrow \mathbf{v}] = \mathscr{D}$

# Counterfactuals in Causal Models

## Counterfactual Models

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, including the variables $\mathbf{V}$, and given the assignment of values $\mathbf{v}$ to $\mathbf{V}$, the counterfactual model $\mathbb{M}[\mathbf{V} \to \mathbf{v}] = (\mathbb{U}[\mathbf{V} \to \mathbf{v}], \vec{u}[\mathbf{V} \to \mathbf{v}], \mathbb{V}[\mathbf{V} \to \mathbf{v}], \mathbb{E}[\mathbf{V} \to \mathbf{v}], \mathscr{D}[\mathbf{V} \to \mathbf{v}])$ is the model such that:

- $\mathbb{V}[\mathbf{V} \to \mathbf{v}] = \mathbb{V} - \mathbf{V}$
- $\mathbb{U}[\mathbf{V} \to \mathbf{v}] = \mathbb{U} \cup \mathbf{V}$
- $\mathbb{E}[\mathbf{V} \to \mathbf{v}] = \mathbb{E} - (\phi_{V_i} \mid V_i \in \mathbf{V})$
- $\vec{u}[\mathbf{V} \to \mathbf{v}] = \vec{u} \cup \mathbf{v}$
- $\mathscr{D}[\mathbf{V} \to \mathbf{v}] = \mathscr{D}$

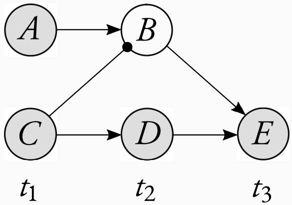
# Counterfactuals in Causal Models

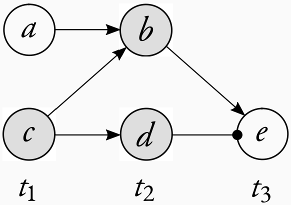
**Causal Counterfactuals**

In a causal model $\mathbb{M}$, containing the variables in $\mathbf{V}$, the causal counterfactual $\mathbf{V} = \mathbf{v} \;\Box\!\!\rightarrow \psi$ is true iff $\psi$ is true in the counterfactual model $\mathbb{M}[\mathbf{V} \rightarrow \mathbf{v}]$,

$$\mathbb{M} \models \mathbf{V} = \mathbf{v} \;\Box\!\!\rightarrow \psi \quad \Longleftrightarrow \quad \mathbb{M}[\mathbf{V} \rightarrow \mathbf{v}] \models \psi$$

# Defaults in Causal Models




$$E := B \vee D$$
$$D := C$$
$$B := A \wedge \neg C$$
$$A = 1$$
$$C = 1$$

$$\overline{e} := \overline{b} \vee d$$
$$d := c$$
$$\overline{b} := \overline{a} \wedge \neg c$$
$$\overline{a} = 1$$
$$c = 1$$

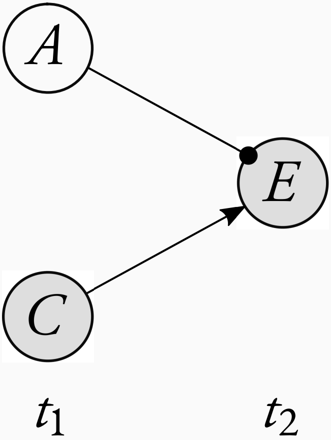
# Model Invariance

# Model Invariance

**Model Invariance**

Given any two causal models, $\mathbb{M}$ and $\mathbb{M}^*$, which both contain the variables $C$ and $E$, if both $\mathbb{M}$ and $\mathbb{M}^*$ are correct, then $C = c$ caused $E = e$ in $\mathbb{M}$ iff $C = c$ caused $E = e$ in $\mathbb{M}^*$.

# Exogenous Reduction



# Exogenous Reduction

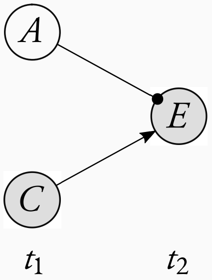
The model $\mathbb{M}$:

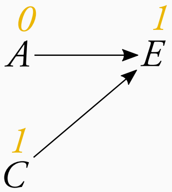
$\mathbb{U} : (A, C)$

$\vec{u} : (0, 1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \wedge \neg A)$



# Exogenous Reduction

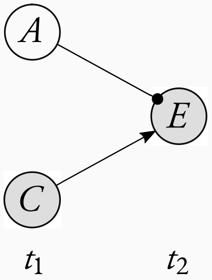
The model $\mathbb{M}$:

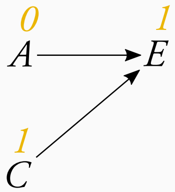
$\mathbb{U} : (A, C)$

$\vec{u} : (0, 1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \land \neg A)$



# Exogenous Reduction

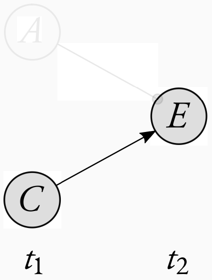
The model $\mathbb{M}^{-A}$:

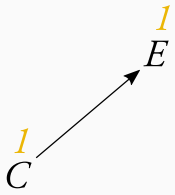
$\mathbb{U} : (A, C)$

$\vec{u} : (0, 1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \wedge \neg A)$



# Exogenous Reduction

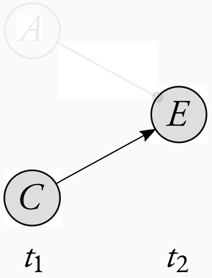
The model $\mathbb{M}^{-A}$:

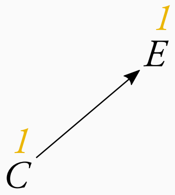
$\mathbb{U} : (A, C)$

$\vec{u} : (0, 1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E \coloneqq C \wedge \neg A)$



# Exogenous Reduction

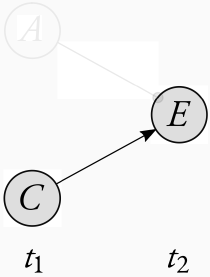
The model $\mathbb{M}^{-A}$:

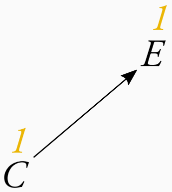
$\mathbb{U} : (C)$

$\vec{u} : (0, 1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \wedge \neg A)$



# Exogenous Reduction

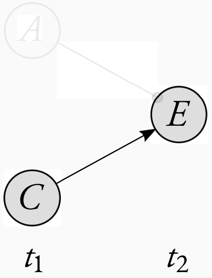
The model $\mathbb{M}^{-A}$:

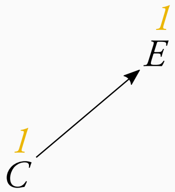
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \wedge \neg A)$



# Exogenous Reduction

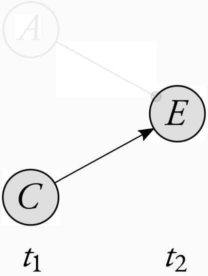
The model $\mathbb{M}^{-A}$:

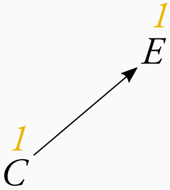
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \land \neg A)$



# Exogenous Reduction

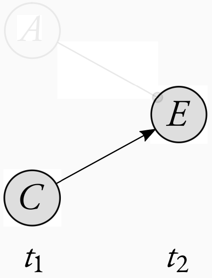
The model $\mathbb{M}^{-A}$:

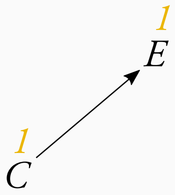
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \wedge \neg 0)$



# Exogenous Reduction

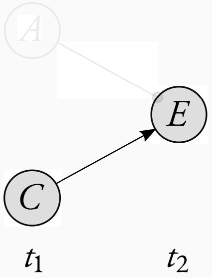
The model $\mathbb{M}^{-A}$:

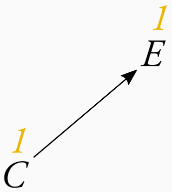
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C)$



# Exogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $U \in \mathbb{U}$, then let $\mathbb{M}^{-U}$ be the model that you get by:
    - Removing $U$ from $\mathbb{U}$
    - Removing $U$'s value from $\vec{u}$
    - Exogenizing any variables in $\mathbb{V}$ whose only parent was $U$
    - Replacing $U$ for its value in every structural equation in $\mathbb{E}$
    - Removing default information about $U$ from $\mathscr{D}$.

# Exogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $U \in \mathbb{U}$, then let $\mathbb{M}^{-U}$ be the model that you get by:
    - Removing $U$ from $\mathbb{U}$
    - Removing $U$'s value from $\vec{u}$
    - Exogenizing any variables in $\mathbb{V}$ whose only parent was $U$
    - Replacing $U$ for its value in every structural equation in $\mathbb{E}$
    - Removing default information about $U$ from $\mathscr{D}$.

# Exogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $U \in \mathbb{U}$, then let $\mathbb{M}^{-U}$ be the model that you get by:
    - Removing $U$ from $\mathbb{U}$
    - Removing $U$'s value from $\vec{u}$
    - Exogenizing any variables in $\mathbb{V}$ whose only parent was $U$
    - Replacing $U$ for its value in every structural equation in $\mathbb{E}$
    - Removing default information about $U$ from $\mathscr{D}$.

# Exogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $U \in \mathbb{U}$, then let $\mathbb{M}^{-U}$ be the model that you get by:
  - Removing $U$ from $\mathbb{U}$
  - Removing $U$'s value from $\vec{u}$
  - Exogenizing any variables in $\mathbb{V}$ whose only parent was $U$
  - Replacing $U$ for its value in every structural equation in $\mathbb{E}$
  - Removing default information about $U$ from $\mathscr{D}$.

# Exogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $U \in \mathbb{U}$, then let $\mathbb{M}^{-U}$ be the model that you get by:
    - Removing $U$ from $\mathbb{U}$
    - Removing $U$'s value from $\vec{u}$
    - Exogenizing any variables in $\mathbb{V}$ whose only parent was $U$
    - Replacing $U$ for its value in every structural equation in $\mathbb{E}$
    - Removing default information about $U$ from $\mathscr{D}$.

# Exogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $U \in \mathbb{U}$, then let $\mathbb{M}^{-U}$ be the model that you get by:
    - Removing $U$ from $\mathbb{U}$
    - Removing $U$'s value from $\vec{u}$
    - Exogenizing any variables in $\mathbb{V}$ whose only parent was $U$
    - Replacing $U$ for its value in every structural equation in $\mathbb{E}$
    - Removing default information about $U$ from $\mathscr{D}$.

# Exogenous Reduction

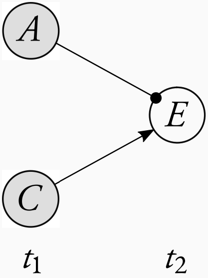
The model $\mathbb{M}$:

$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \wedge \neg A)$



# Exogenous Reduction

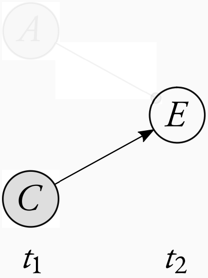
The model $\mathbb{M}^{-A}$:

$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \wedge \neg A)$



# Exogenous Reduction

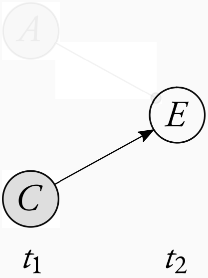
The model $\mathbb{M}^{-A}$:

$\mathbb{U} : (C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \wedge \neg A)$



# Exogenous Reduction

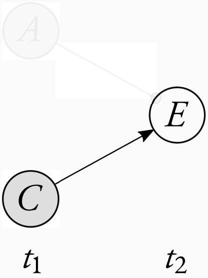
The model $\mathbb{M}^{-A}$:

$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \wedge \neg A)$



# Exogenous Reduction

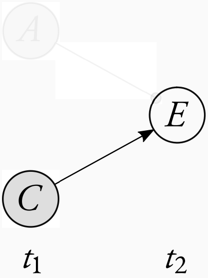
The model $\mathbb{M}^{-A}$:

$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := C \land \neg 1)$



# Exogenous Reduction

- If every equation in $\mathbb{M}^{-U}$ is surjective, then say that $U$ is an *inessential* variable.

**Exogenous Reduction**

If a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is correct, and $U \in \mathbb{U}$ is inessential, then $\mathbb{M}^{-U}$ is also correct.

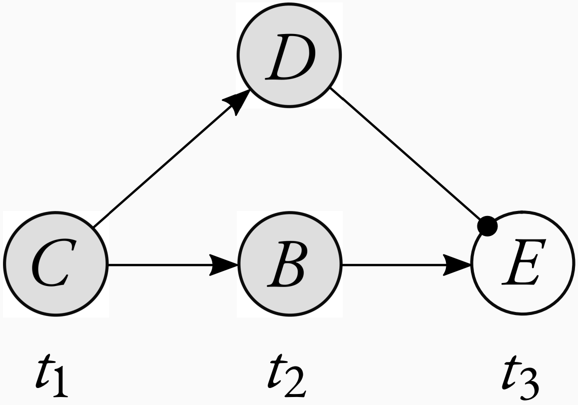
# Exogenous Reduction

- If every equation in $\mathbb{M}^{-U}$ is surjective, then say that $U$ is an *inessential* variable.

**Exogenous Reduction**

If a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is correct, and $U \in \mathbb{U}$ is inessential, then $\mathbb{M}^{-U}$ is also correct.

# Endogenous Reduction



# Model Variance

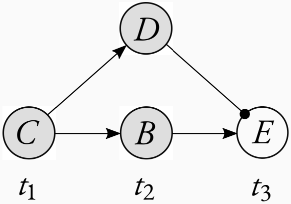
The model $\mathbb{M}$:

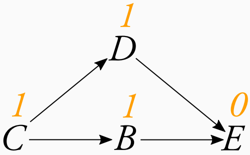
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \wedge \neg D \\ D & := C \\ B & := C \end{pmatrix}$$



# Model Variance

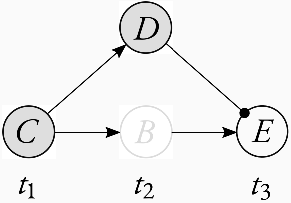
The model $\mathbb{M}^{-B}$:

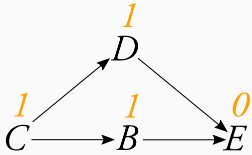
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \wedge \neg D \\ D & := C \\ B & := C \end{pmatrix}$$



# Model Variance

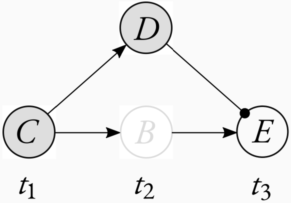
The model $\mathbb{M}^{-B}$:

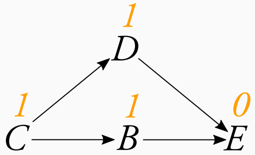
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \land \neg D \\ D & := C \\ B & := C \end{pmatrix}$$



# Model Variance

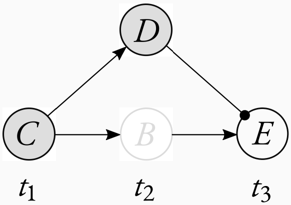
The model $\mathbb{M}^{-B}$:

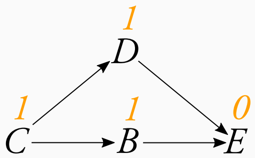
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (D, E)$

$$\mathbb{E} : \begin{pmatrix} E := C \land \neg D \\ D := C \end{pmatrix}$$



# Model Variance

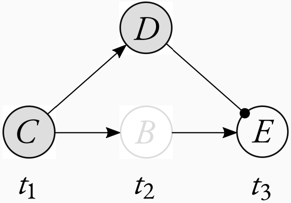
The model $\mathbb{M}^{-B}$:

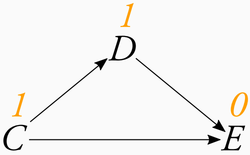
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := C \wedge \neg D \\ D & := C \end{pmatrix}$$



# Endogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $V \in \mathbb{V}$, then let $\mathbb{M}^{-V}$ be the model that you get by:
    - Leaving $\mathbb{U}$ alone
    - Leaving $\vec{u}$ alone
    - Removing $V$ from $\mathbb{V}$
    - Removing $\phi_V$ from $\mathbb{E}$, and replacing $V$ with $\phi_V(\mathbf{PA}(V))$ wherever $V$ appears on the right-hand-side of an equation in $\mathbb{E}$
    - Removing default information about $V$ from $\mathscr{D}$

## Endogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $V \in \mathbb{V}$, then let $\mathbb{M}^{-V}$ be the model that you get by:
  - Leaving $\mathbb{U}$ alone
  - Leaving $\vec{u}$ alone
  - Removing $V$ from $\mathbb{V}$
  - Removing $\phi_V$ from $\mathbb{E}$, and replacing $V$ with $\phi_V(\mathbf{PA}(V))$ wherever $V$ appears on the right-hand-side of an equation in $\mathbb{E}$
  - Removing default information about $V$ from $\mathscr{D}$

# Endogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $V \in \mathbb{V}$, then let $\mathbb{M}^{-V}$ be the model that you get by:
  - Leaving $\mathbb{U}$ alone
  - Leaving $\vec{u}$ alone
  - Removing $V$ from $\mathbb{V}$
  - Removing $\phi_V$ from $\mathbb{E}$, and replacing $V$ with $\phi_V(\mathbf{PA}(V))$ wherever $V$ appears on the right-hand-side of an equation in $\mathbb{E}$
  - Removing default information about $V$ from $\mathscr{D}$

# Endogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $V \in \mathbb{V}$, then let $\mathbb{M}^{-V}$ be the model that you get by:
    - Leaving $\mathbb{U}$ alone
    - Leaving $\vec{u}$ alone
    - Removing $V$ from $\mathbb{V}$
    - Removing $\phi_V$ from $\mathbb{E}$, and replacing $V$ with $\phi_V(\mathbf{PA}(V))$ wherever $V$ appears on the right-hand-side of an equation in $\mathbb{E}$
    - Removing default information about $V$ from $\mathscr{D}$

# Endogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $V \in \mathbb{V}$, then let $\mathbb{M}^{-V}$ be the model that you get by:
  - Leaving $\mathbb{U}$ alone
  - Leaving $\vec{u}$ alone
  - Removing $V$ from $\mathbb{V}$
  - Removing $\phi_V$ from $\mathbb{E}$, and replacing $V$ with $\phi_V(\mathbf{PA}(V))$ wherever $V$ appears on the right-hand-side of an equation in $\mathbb{E}$
  - Removing default information about $V$ from $\mathscr{D}$

# Endogenous Reduction

- If $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is a causal model with $V \in \mathbb{V}$, then let $\mathbb{M}^{-V}$ be the model that you get by:
    - Leaving $\mathbb{U}$ alone
    - Leaving $\vec{u}$ alone
    - Removing $V$ from $\mathbb{V}$
    - Removing $\phi_V$ from $\mathbb{E}$, and replacing $V$ with $\phi_V(\mathbf{PA}(V))$ wherever $V$ appears on the right-hand-side of an equation in $\mathbb{E}$
    - Removing default information about $V$ from $\mathscr{D}$

# Endogenous Reduction

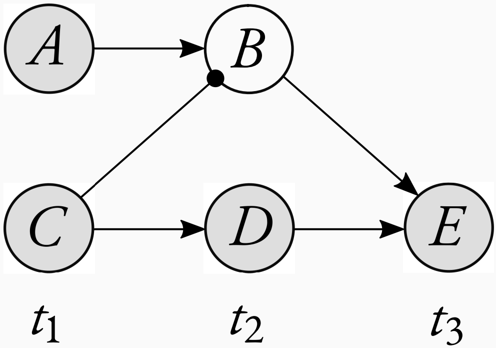



**Figure 1:** *Preemptive Overdetermination*

# Endogenous Reduction

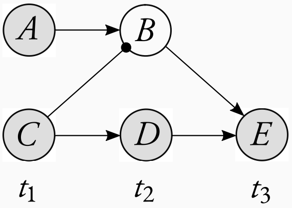
The model $\mathbb{M}_1$:

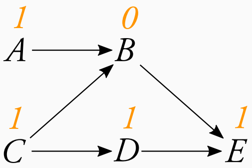
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee D \\ D & := C \\ B & := A \wedge \neg C \end{pmatrix}$$



# Endogenous Reduction

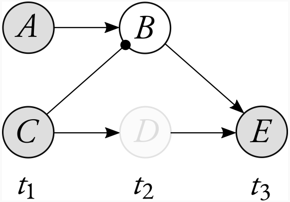
The model $\mathbb{M}_1^{-D}$:

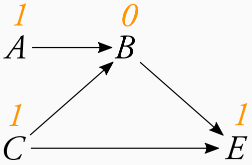
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee C \\ B & := A \wedge \neg C \end{pmatrix}$$



# Endogenous Reduction

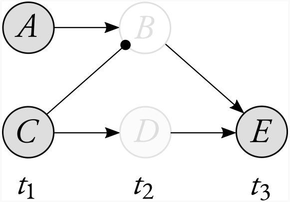
The model $\mathbb{M}_1^{-D,-B}$:

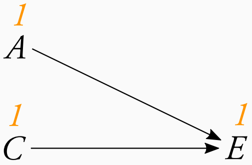
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (E)$

$\mathbb{E} : \left( \; E := (A \wedge \neg C) \vee C \; \right)$



# Endogenous Reduction

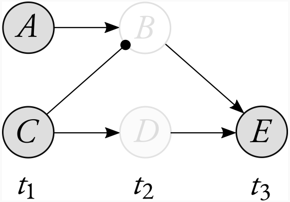
The model $\mathbb{M}_1^{-D,-B}$:

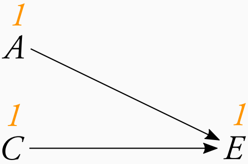
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (E)$

$\mathbb{E} : \left( \ E \ \coloneqq A \vee C \ \right)$



# Endogenous Reduction

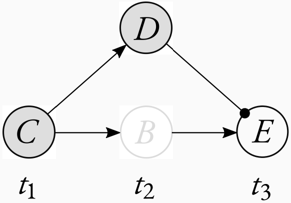
The model $\mathbb{M}_5^{-B}$:

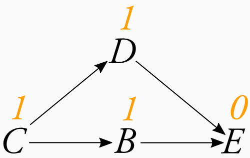
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := C \wedge \neg D \\ D & := C \end{pmatrix}$$



# Endogenous Reduction

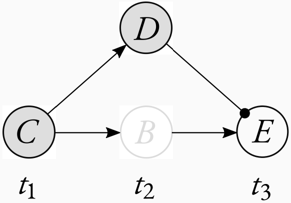
The model $\mathbb{M}_5^{-B}$:

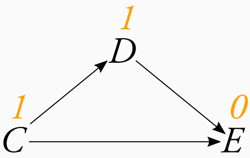
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := C \wedge \neg D \\ D & := C \end{pmatrix}$$



# Endogenous Reduction

- If $V$ has a single parent, $Pa$, and a single child, $Ch$, and if $Pa$ is not also a parent of $Ch$, then say that $V$ is an *interpolated* variable.

$$\ldots Pa \rightarrow V \rightarrow Ch \ldots$$

- If $V$ is interpolated, and the equations in $\mathbb{M}^{-V}$ are surjective, then say that $V$ is *inessential*.

**Endogenous Reduction**

If a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is correct, and $V \in \mathbb{V}$ is an inessential variable, then $\mathbb{M}^{-V}$ is also correct.

# Endogenous Reduction

- If $V$ has a single parent, $Pa$, and a single child, $Ch$, and if $Pa$ is not also a parent of $Ch$, then say that $V$ is an *interpolated* variable.

$$\dots Pa \to V \to Ch \dots$$

- If $V$ is interpolated, and the equations in $\mathbb{M}^{-V}$ are surjective, then say that $V$ is *inessential*.

**Endogenous Reduction**

If a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is correct, and $V \in \mathbb{V}$ is an inessential variable, then $\mathbb{M}^{-V}$ is also correct.

# Endogenous Reduction

- If $V$ has a single parent, $Pa$, and a single child, $Ch$, and if $Pa$ is not also a parent of $Ch$, then say that $V$ is an *interpolated* variable.

$$\ldots Pa \to V \to Ch \ldots$$

- If $V$ is interpolated, and the equations in $\mathbb{M}^{-V}$ are surjective, then say that $V$ is *inessential*.

**Endogenous Reduction**

If a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$ is correct, and $V \in \mathbb{V}$ is an inessential variable, then $\mathbb{M}^{-V}$ is also correct.

# Model Invariance

- The accounts of Hitchcock (2001, 2007), Halpern & Pearl (2001, 2005), Woodward (2003), Halpern (2008), and Weslake (forthcoming) are all inconsistent with **Model Invariance**, **Exogenous Reduction**, and **Endogenous Reduction**.

# A Model Invariant Theory

# A Model Invariant Theory

- I will present a theory of causation in terms of structural equations models which is consistent with **Endogenous Reduction**, **Exogenous Reduction**, and **Model Invariance**.
- I'll build up the theory by progressing through some familiar cases from the literature.

# A Model Invariant Theory

- I will present a theory of causation in terms of structural equations models which is consistent with **Endogenous Reduction**, **Exogenous Reduction**, and **Model Invariance**.
- I'll build up the theory by progressing through some familiar cases from the literature.

# A Model Invariant Theory

## Preemptive Overdetermination

# Preemptive Overdetermination

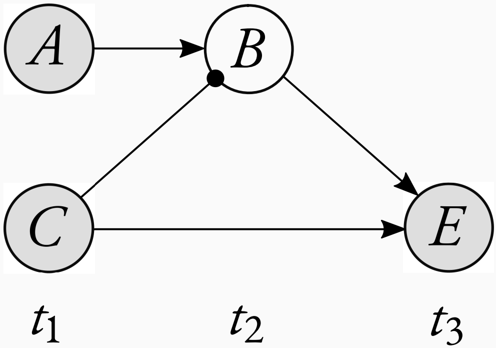



**Figure 2:** *Preemptive Overdetermination*

# Preemptive Overdetermination

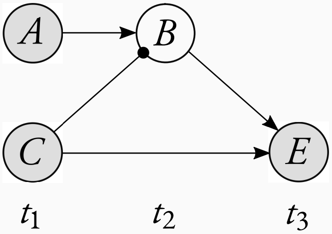
The model $\mathbb{M}_2$:

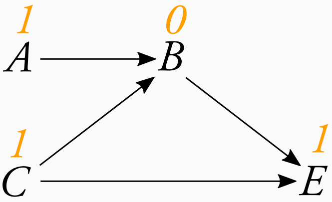
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

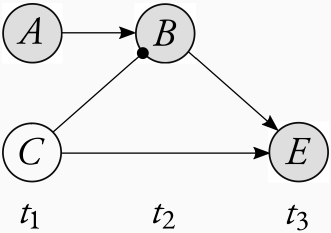
$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee C \\ B & := A \wedge \neg C \end{pmatrix}$$



# Preemptive Overdetermination



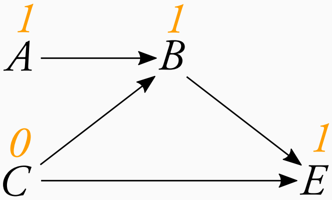

The model $\mathbb{M}_2[C \to 0]$:

$\mathbb{U} : (A, C)$

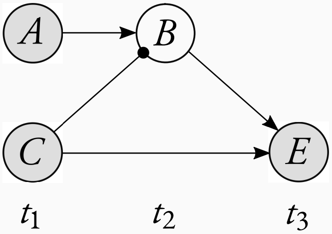
$\vec{u} : (1, 0)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee C \\ B & := A \wedge \neg C \end{pmatrix}$$



# Preemptive Overdetermination



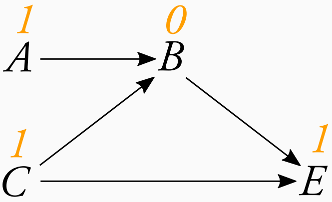

The model $\mathbb{M}_2$:

$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \vee C \\ B & := A \wedge \neg C \end{pmatrix}$$



# Preemptive Overdetermination

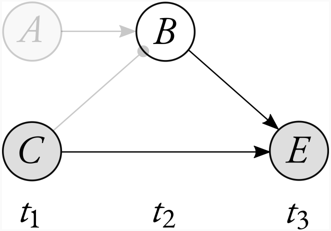
The *local* model $\mathbb{M}_2((E))$:

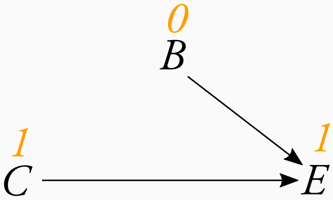
$\mathbb{U} : (C, B)$

$\vec{u} : (1, 0)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := B \vee C)$



# Preemptive Overdetermination

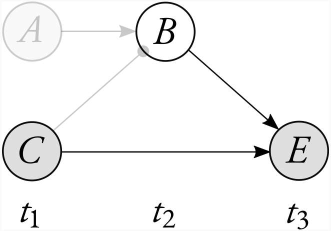
The model $\mathbb{M}_2((E))$:

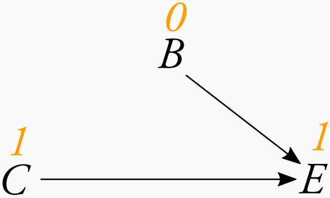
$\mathbb{U} : (C, B)$

$\vec{u} : (1, 0)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := B \vee C)$



# Preemptive Overdetermination

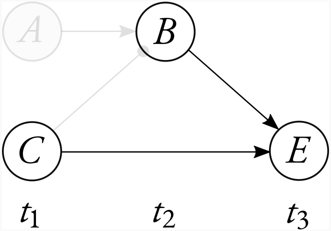
The model $\mathbb{M}_2((E))[C \to 0]$:

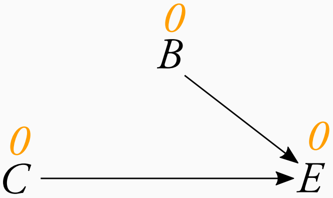
$\mathbb{U} : (C, B)$

$\vec{u} : (0, 0)$

$\mathbb{V} : (E)$

$\mathbb{E} : (E := B \vee C)$



# Local Models

## Local Model

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $E \in \mathbb{V}$, the *local model* at $E$, $\mathbb{M}((E))$, is the causal model in which

(a) The exogenous variables are just the parents of $E$, $\mathbf{PA}(E)$, in the original model $\mathbb{M}$;

(b) The exogenous variables $\mathbf{PA}(E)$ are assigned the values they take on in $\mathbb{M}$;

(c) The sole endogenous variable is $E$;

(d) The sole structural equation is $E$'s structural equation in $\mathbb{M}$, $\phi_E$; and

(e) The defaults for $E$ and $\mathbf{PA}(E)$ are the same as in $\mathbb{M}$.

# Local Models

## Local Model

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $E \in \mathbb{V}$, the *local model* at $E$, $\mathbb{M}(\!(E)\!)$, is the causal model in which

(a) The exogenous variables are just the parents of $E$, $\mathbf{PA}(E)$, in the original model $\mathbb{M}$;

(b) The exogenous variables $\mathbf{PA}(E)$ are assigned the values they take on in $\mathbb{M}$;

(c) The sole endogenous variable is $E$;

(d) The sole structural equation is $E$'s structural equation in $\mathbb{M}$, $\phi_E$; and

(e) The defaults for $E$ and $\mathbf{PA}(E)$ are the same as in $\mathbb{M}$.

# Local Models

**Local Model**

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $E \in \mathbb{V}$, the *local model* at $E$, $\mathbb{M}((E))$, is the causal model in which

(a) The exogenous variables are just the parents of $E$, $\mathbf{PA}(E)$, in the original model $\mathbb{M}$;

(b) The exogenous variables $\mathbf{PA}(E)$ are assigned the values they take on in $\mathbb{M}$;

(c) The sole endogenous variable is $E$;

(d) The sole structural equation is $E$'s structural equation in $\mathbb{M}$, $\phi_E$; and

(e) The defaults for $E$ and $\mathbf{PA}(E)$ are the same as in $\mathbb{M}$.

# Local Models

## Local Model

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $E \in \mathbb{V}$, the *local model* at $E$, $\mathbb{M}(\!(E)\!)$, is the causal model in which

(a) The exogenous variables are just the parents of $E$, $\mathbf{PA}(E)$, in the original model $\mathbb{M}$;

(b) The exogenous variables $\mathbf{PA}(E)$ are assigned the values they take on in $\mathbb{M}$;

(c) The sole endogenous variable is $E$;

(d) The sole structural equation is $E$'s structural equation in $\mathbb{M}$, $\phi_E$; and

(e) The defaults for $E$ and $\mathbf{PA}(E)$ are the same as in $\mathbb{M}$.

# Local Models

## Local Model

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $E \in \mathbb{V}$, the *local model* at $E$, $\mathbb{M}(\!(E)\!)$, is the causal model in which

(a) The exogenous variables are just the parents of $E$, $\mathbf{PA}(E)$, in the original model $\mathbb{M}$;

(b) The exogenous variables $\mathbf{PA}(E)$ are assigned the values they take on in $\mathbb{M}$;

(c) The sole endogenous variable is $E$;

(d) The sole structural equation is $E$'s structural equation in $\mathbb{M}$, $\phi_E$; and

(e) The defaults for $E$ and $\mathbf{PA}(E)$ are the same as in $\mathbb{M}$.

# Local Models

## Local Model

Given a causal model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $E \in \mathbb{V}$, the *local model* at $E$, $\mathbb{M}(\!(E)\!)$, is the causal model in which

(a) The exogenous variables are just the parents of $E$, $\mathbf{PA}(E)$, in the original model $\mathbb{M}$;

(b) The exogenous variables $\mathbf{PA}(E)$ are assigned the values they take on in $\mathbb{M}$;

(c) The sole endogenous variable is $E$;

(d) The sole structural equation is $E$'s structural equation in $\mathbb{M}$, $\phi_E$; and

(e) The defaults for $E$ and $\mathbf{PA}(E)$ are the same as in $\mathbb{M}$.

# Local Counterfactual Dependence

- $E = e$ *locally* counterfactually depends upon $C = c$ iff, *in the local model at E*, $\mathbb{M}(\!(E)\!)$, there's some $c^* \neq c, e^* \neq e$ such that

$$\mathbb{M}(\!(E)\!) \models C = c^* \;\Box\!\!\rightarrow E = e^*$$

- A (preliminary) proposal: either local or global counterfactual dependence suffices for causation.

# Local Counterfactual Dependence

- $E = e$ *locally* counterfactually depends upon $C = c$ iff, *in the local model at E*, $\mathbb{M}(\!(E)\!)$, there's some $c^* \neq c$, $e^* \neq e$ such that

$$\mathbb{M}(\!(E)\!) \models C = c^* \;\Box\!\!\rightarrow E = e^*$$

- A (preliminary) proposal: either local or global counterfactual dependence suffices for causation.

# Local Counterfactual Dependence

- $E = e$ *locally* counterfactually depends upon $C = c$ iff, *in the local model at E*, $\mathbb{M}(\!(E)\!)$, there's some $c^*, e^*$ such that

$$\mathbb{M}(\!(E)\!) \models C = c^* \;\Box\!\!\rightarrow E = e^*$$

- A (preliminary) proposal: either local or global counterfactual dependence suffices for causation.

# Local Counterfactual Dependence

- $E = e$ *locally* counterfactually depends upon $C = c$ iff, *in the local model at E*, $\mathbb{M}(\!(E)\!)$, there's some $c^*, e^*$ such that

$$\mathbb{M}(\!(E)\!) \models C = c^* \mathrel{\Box\!\!\rightarrow} E = e^*$$

- A (preliminary) proposal: either local or global counterfactual dependence suffices for causation.

# Local Counterfactual Dependence

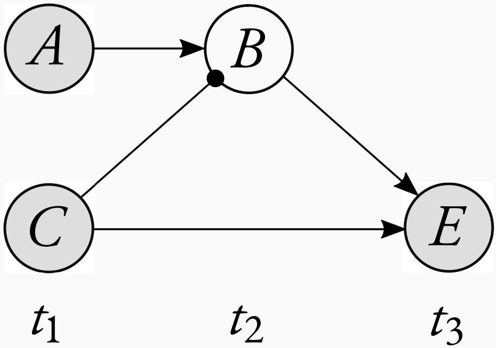



**Figure 2:** *Preemptive Overdetermination*

# Local Counterfactual Dependence

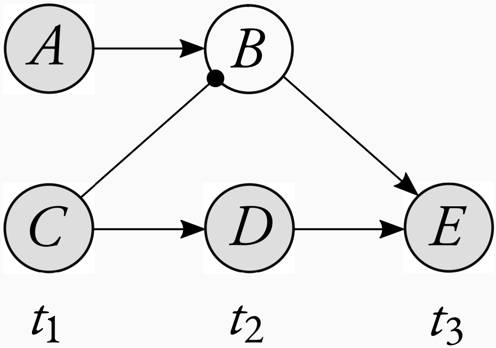



**Figure 1:** *Preemptive Overdetermination*

# A Model Invariant Theory

## Counterexamples to Transitivity

# Transitivity

- Sometimes, we trace out a sequence of causal relations and conclude that the first event in the chain caused the last.
    - If we can do this, then let's say that the chain of causal relations is *transitive*
- When is a chain of causal relations transitive?
    - Lewis said 'Always', but this answer comes at a cost
    - The answer to give: 'Sometimes, but not always'.

# Transitivity

- Sometimes, we trace out a sequence of causal relations and conclude that the first event in the chain caused the last.
    - If we can do this, then let's say that the chain of causal relations is *transitive*
- When is a chain of causal relations transitive?
    - Lewis said 'Always', but this answer comes at a cost
    - The answer to give: 'Sometimes, but not always'.

# Transitivity

- Sometimes, we trace out a sequence of causal relations and conclude that the first event in the chain caused the last.
    - If we can do this, then let's say that the chain of causal relations is *transitive*
- When is a chain of causal relations transitive?
    - Lewis said 'Always', but this answer comes at a cost
    - The answer to give: 'Sometimes, but not always'.

# Transitivity

- Sometimes, we trace out a sequence of causal relations and conclude that the first event in the chain caused the last.
    - If we can do this, then let's say that the chain of causal relations is *transitive*
- When is a chain of causal relations transitive?
    - Lewis said 'Always', but this answer comes at a cost
    - The answer to give: 'Sometimes, but not always'.

# Transitivity

- Sometimes, we trace out a sequence of causal relations and conclude that the first event in the chain caused the last.
    - If we can do this, then let's say that the chain of causal relations is *transitive*
- When is a chain of causal relations transitive?
    - Lewis said 'Always', but this answer comes at a cost
    - The answer to give: 'Sometimes, but not always'.

# Transitive Path

**Transitive Path**

In a causal model $\mathbb{M}$, a directed path $\mathbf{P}$

$$\mathbf{P} : V_1 \to V_2 \to V_3 \to \cdots \to V_N$$

is a *transitive path* iff:

# Counterexamples to Transitivity

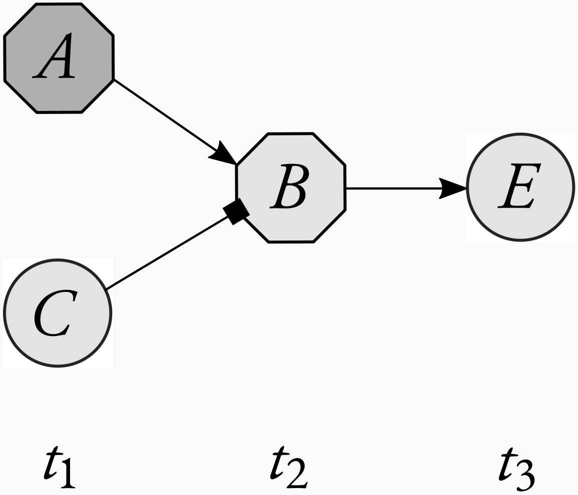



**Figure 3:** *Tampering*

# Counterexamples to Transitivity

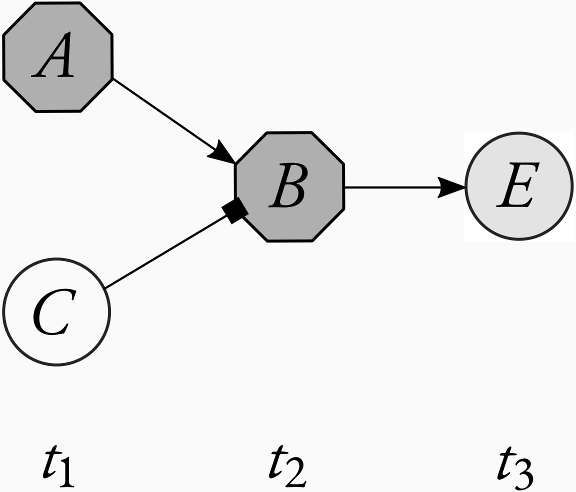



**Figure 3:** *Tampering*

# Counterexamples to Transitivity

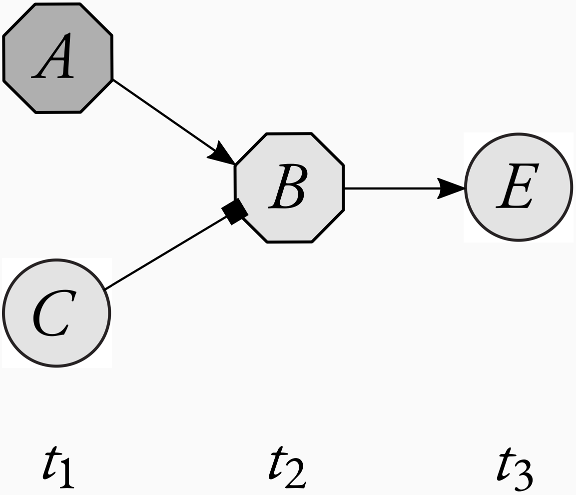



**Figure 3:** *Tampering*

# Counterexamples to Transitivity

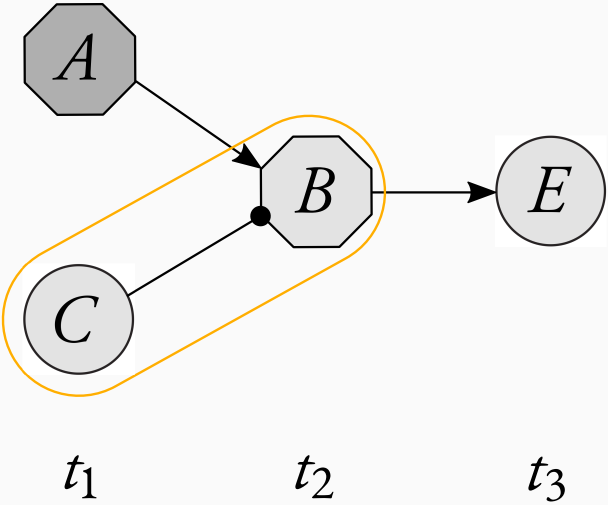



**Figure 3:** *Tampering*

# Counterexamples to Transitivity

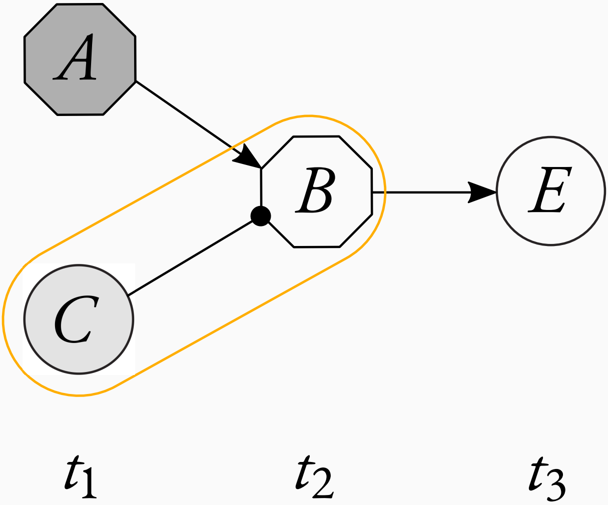



**Figure 3:** *Tampering*

# Counterexamples to Transitivity

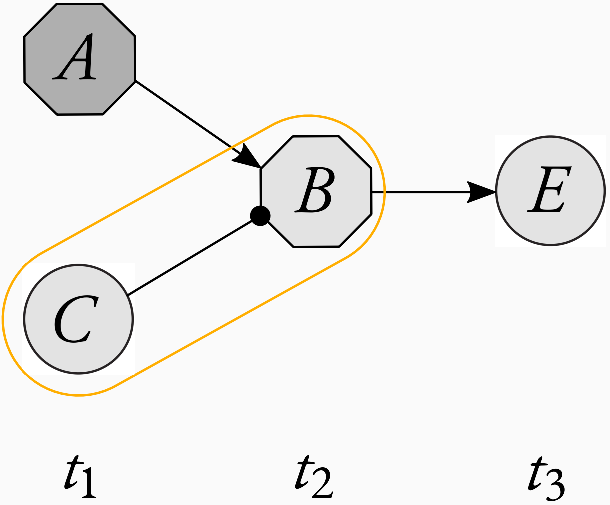



**Figure 3:** *Tampering*

# Counterexamples to Transitivity

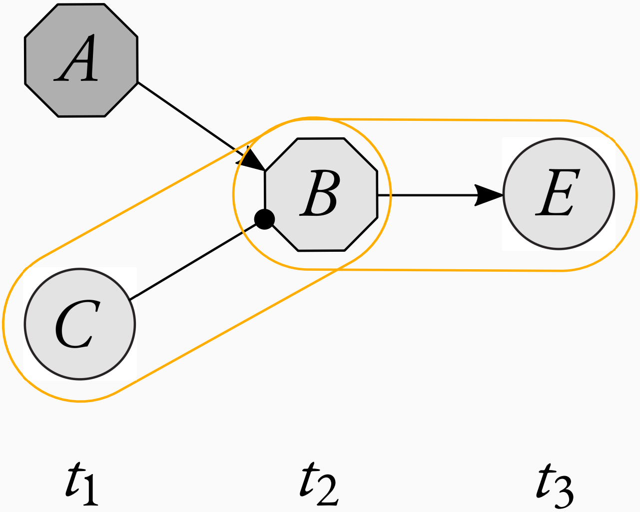



**Figure 3:** *Tampering*

# Transitive Path

## Transitive Path

In a causal model $\mathbb{M}$, a directed path $\mathbf{P}$

$$\mathbf{P} : V_1 \to V_2 \to V_3 \to \cdots \to V_N$$

is a *transitive path* iff:

(a) For each variable $V_i$ along $\mathbf{P}$, there is a pair $(v_i, v_i^*)$ of $V_i$'s actual value $v_i$ in $\mathbb{M}$, and a *contrast* value $v_i^*$,

$$(v_1, v_1^*) \to (v_2, v_2^*) \to (v_3, v_3^*) \to \cdots \to (v_N, v_N^*)$$

such that: for all $j$ between 1 and $N-1$, $V_j$'s taking on the value $v_j$, rather than $v_j^*$, caused $V_{j+1}$ to take on the value $v_{j+1}$, rather than $v_{j+1}^*$;

# Transitive Path

## Transitive Path

In a causal model $\mathbb{M}$, a directed path $\mathbf{P}$

$$\mathbf{P} : V_1 \to V_2 \to V_3 \to \cdots \to V_N$$

is a *transitive path* iff:

(a) For each variable $V_i$ along $\mathbf{P}$, there is a pair $(v_i, v_i^*)$ of $V_i$'s actual value $v_i$ in $\mathbb{M}$, and a *contrast* value $v_i^*$,

$$(v_1, v_1^*) \to (v_2, v_2^*) \to (v_3, v_3^*) \to \cdots \to (v_N, v_N^*)$$

such that: for all $j$ between 1 and $N-1$, $V_j$'s taking on the value $v_j$, rather than $v_j^*$, caused $V_{j+1}$ to take on the value $v_{j+1}$, rather than $v_{j+1}^*$;

# Causes and Contrasts

- Contrastivism gives us a 4-place causal relation:

$$\textsc{Cause}(C = c, C = c^*, E = e, E = e^*)$$

- From this, we may recover a familiar 2-place causal relation:

$$\textsc{Cause}(C = c, E = e) \iff \exists c^* \exists e^* \textsc{Cause}(C = c, C = c^*, E = e, E = e^*)$$

# Causes and Contrasts

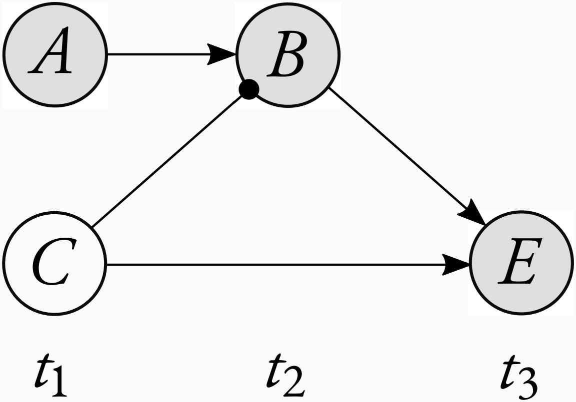
- Contrastivism gives us a 4-place causal relation:

$$\textsc{Cause}(C = c, C = c^*, E = e, E = e^*)$$

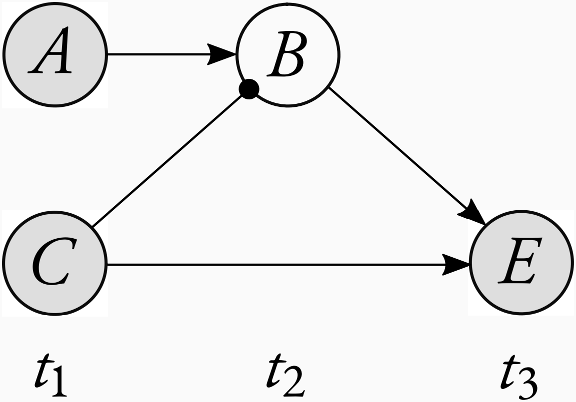
- From this, we may recover a familiar 2-place causal relation:

$$\textsc{Cause}(C = c, E = e) \iff \exists c^* \exists e^* \textsc{Cause}(C = c, C = c^*, E = e, E = e^*)$$

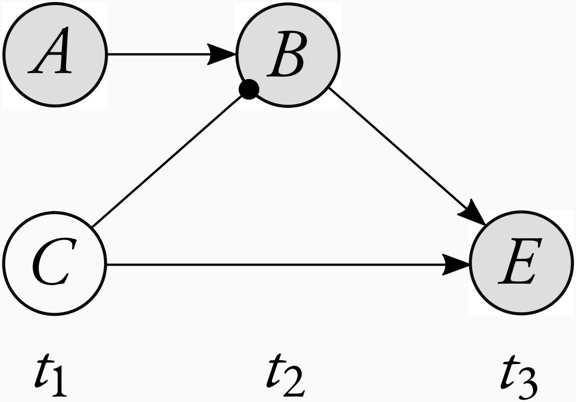
# Counterexamples to Transitivity



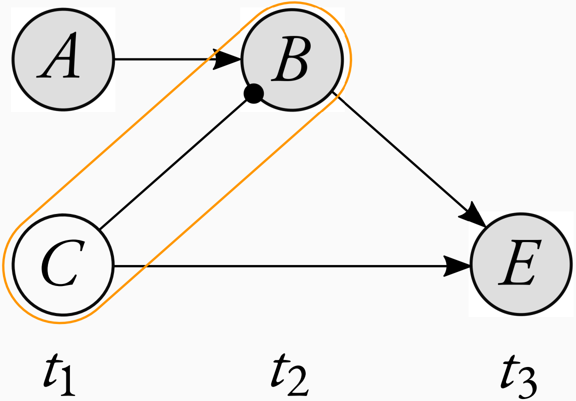
# Counterexamples to Transitivity



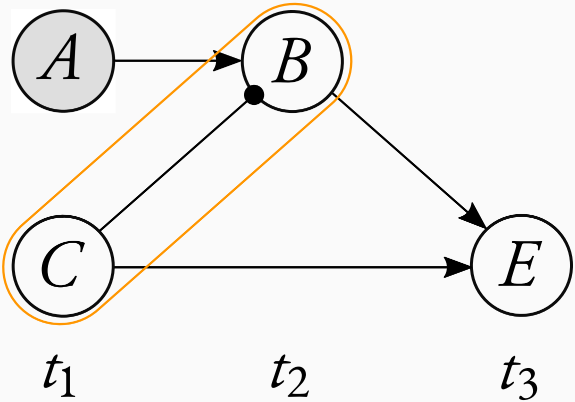
# Counterexamples to Transitivity



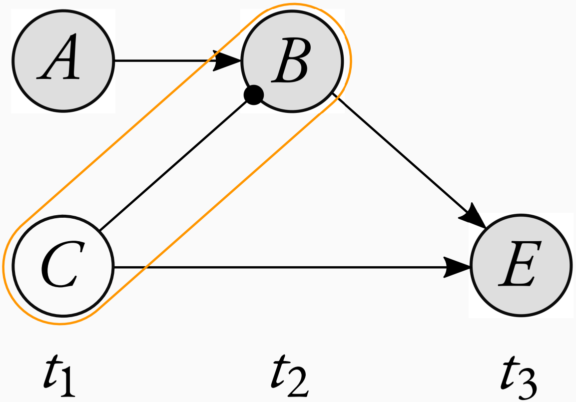
# Counterexamples to Transitivity



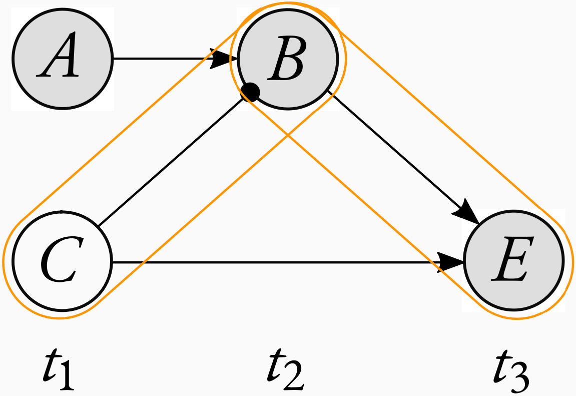
# Counterexamples to Transitivity



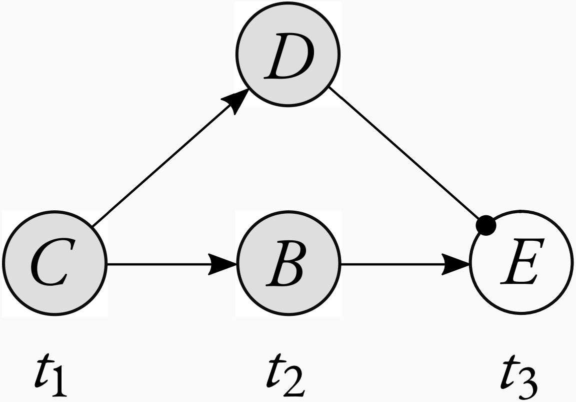
# Counterexamples to Transitivity



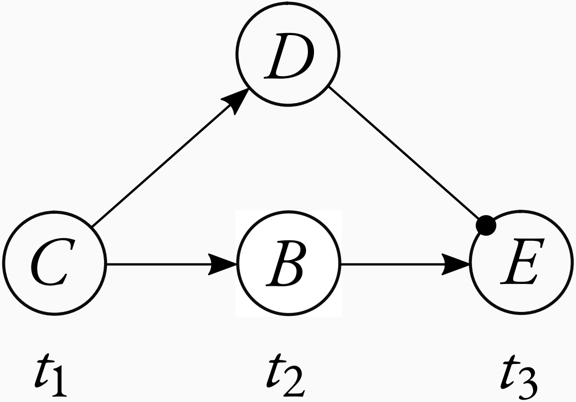
# Counterexamples to Transitivity



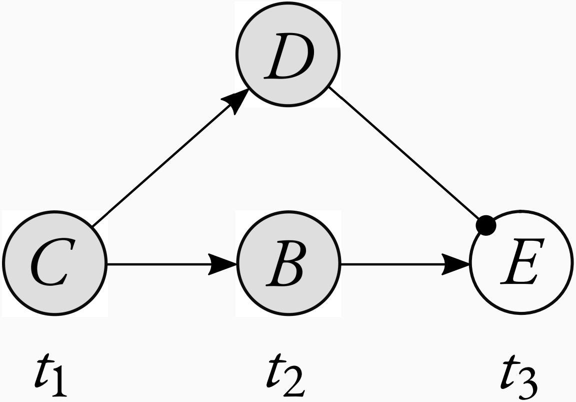
# Counterexamples to Transitivity



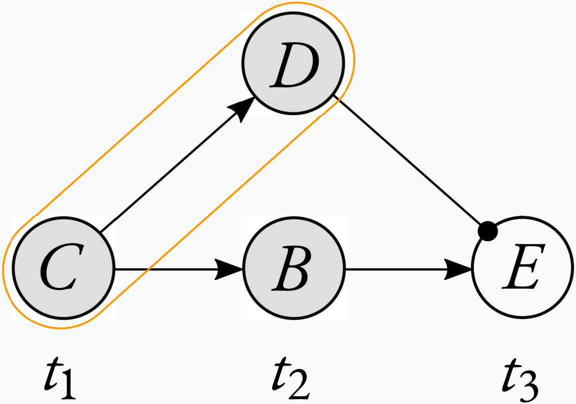
# Counterexamples to Transitivity



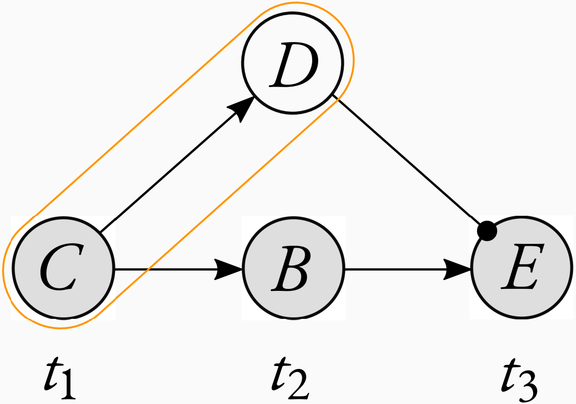
# Counterexamples to Transitivity



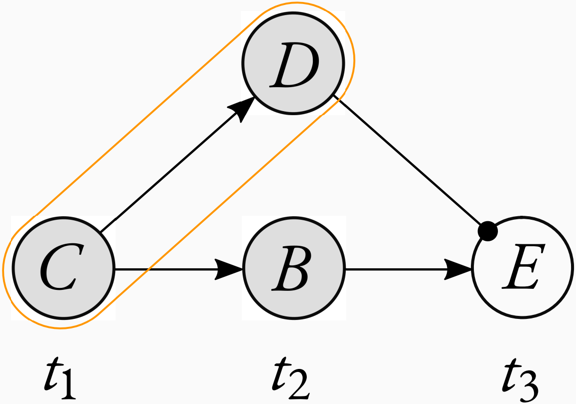
# Counterexamples to Transitivity



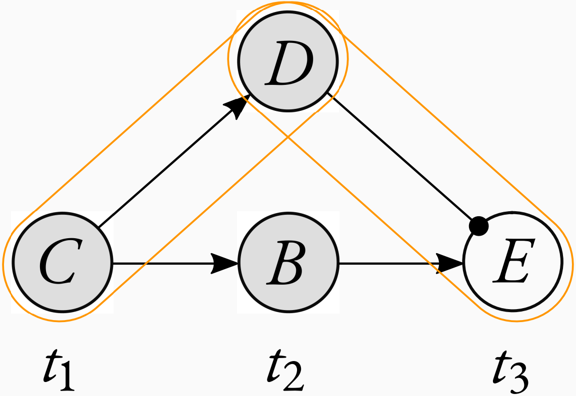
# Counterexamples to Transitivity



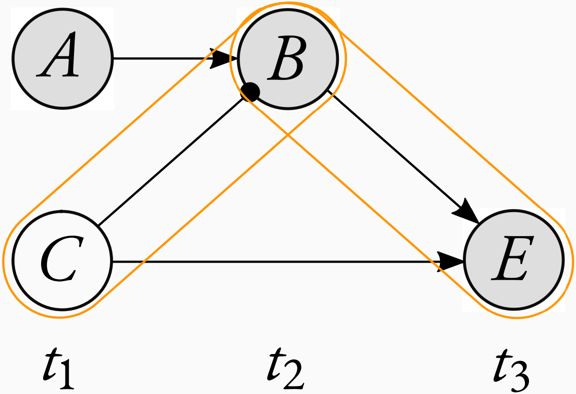
# Counterexamples to Transitivity



# Counterexamples to Transitivity



# Counterexamples to Transitivity



# Transitive Path

**Transitive Path**

In a causal model $\mathbb{M}$, a directed path $\mathbf{P}$

$$\mathbf{P} : V_1 \to V_2 \to V_3 \to \cdots \to V_N$$

is a *transitive path* iff:

(a) For each variable $V_i$ along this path, there is a pair $(v_i, v_i^*)$ of $V_i$'s actual value $v_i$ in $\mathbb{M}$, and a *contrast* value $v_i^*$,

$$(v_1, v_1^*) \to (v_2, v_2^*) \to (v_3, v_3^*) \to \cdots \to (v_N, v_N^*)$$

such that: for all $j$ between 1 and $N-1$, $V_j$'s taking on the value $v_j$, rather than $v_j^*$, caused $V_{j+1}$ to take on the value $v_{j+1}$, rather than $v_{j+1}^*$;

# Transitive Path

## Transitive Path

In a causal model $\mathbb{M}$, a directed path $\mathbf{P}$

$$\mathbf{P} : V_1 \to V_2 \to V_3 \to \cdots \to V_N$$

is a *transitive path* iff:

(a) For each variable $V_i$ along this path, there is a pair $(v_i, v_i^*)$ of $V_i$'s actual value $v_i$ in $\mathbb{M}$, and a *contrast* value $v_i^*$,

$$(v_1, v_1^*) \to (v_2, v_2^*) \to (v_3, v_3^*) \to \cdots \to (v_N, v_N^*)$$

such that: for all $j$ between 1 and $N-1$, $V_j$'s taking on the value $v_j$, rather than $v_j^*$, caused $V_{j+1}$ to take on the value $v_{j+1}$, rather than $v_{j+1}^*$;

## Transitive Path

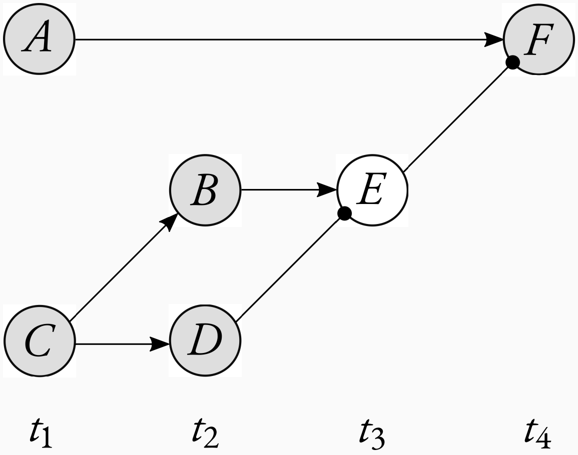
In a causal model $\mathbb{M}$, a directed path $\mathbf{P}$

$$\mathbf{P} : V_1 \to V_2 \to V_3 \to \cdots \to V_N$$

is a *transitive path* iff:

(b) Both $V_1$'s and $V_N$'s actual values are *deviant*, their contrast values *default*;

# Short Circuit, again



# Transitive Path

## Transitive Path

In a causal model $\mathbb{M}$, a directed path **P**

$$\mathbf{P} : V_1 \to V_2 \to V_3 \to \cdots \to V_N$$

is a *transitive path* iff:

(a) For each variable $V_i$ along this path, there is a pair $(v_i, v_i^*)$ of $V_i$'s actual value $v_i$ in $\mathbb{M}$, and a *contrast* value $v_i^*$,

$$(v_1, v_1^*) \to (v_2, v_2^*) \to (v_3, v_3^*) \to \cdots \to (v_N, v_N^*)$$

such that: for all $j$ between 1 and $N-1$, $V_j$'s taking on the value $v_j$, rather than $v_j^*$, caused $V_{j+1}$ to take on the value $v_{j+1}$, rather than $v_{j+1}^*$;

# Transitive Path

## Transitive Path

In a causal model $\mathbb{M}$, a directed path $\mathbf{P}$

$$\mathbf{P} : V_1 \rightarrow V_2 \rightarrow V_3 \rightarrow \cdots \rightarrow V_N$$

is a *transitive path* iff:

(a) For each variable $V_i$ along this path, there is a pair $(v_i, v_i^*)$ of $V_i$'s actual value $v_i$ in $\mathbb{M}$, and a *contrast* value $v_i^*$,

$$(v_1, v_1^*) \rightarrow (v_2, v_2^*) \rightarrow (v_3, v_3^*) \rightarrow \cdots \rightarrow (v_N, v_N^*)$$

such that: for all $j$ between 1 and $N-1$, $V_j$'s taking on the value $v_j$, rather than $v_j^*$, caused $V_{j+1}$ to take on the value $v_{j+1}$, rather than $v_{j+1}^*$;

## Transitive Path

In a causal model $\mathbb{M}$, a directed path $\mathbf{P}$

$$\mathbf{P} : V_1 \to V_2 \to V_3 \to \cdots \to V_N$$

is a *transitive path* iff:

(b) Both $V_1$'s and $V_N$'s actual values are *deviant*, their contrast values *default*;

(c) Every *departure* variable along $\mathbf{P}$ causes each of its *return* variables along $\mathbf{P}$.

# Transitive Path

## Transitive Path

In a causal model $\mathbb{M}$, a directed path $\mathbf{P}$

$$\mathbf{P} : V_1 \to V_2 \to V_3 \to \cdots \to V_N$$

is a *transitive path* iff:

(b) Both $V_1$'s and $V_N$'s actual values are *deviant*, their contrast values *default*;

(c) Every *departure* variable along $\mathbf{P}$ causes each of its *return* variables along $\mathbf{P}$.

# A Model Invariant Theory

## Preemption and Omission without Dependence?

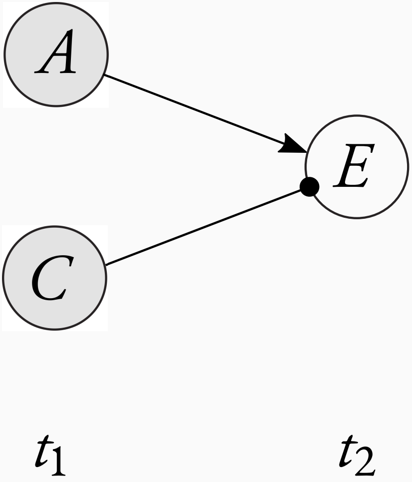
# Default Causes and Effects

- So far, we've only looked at causes and effects whose values are *deviant*
- Default variable values can also be causes and effects.

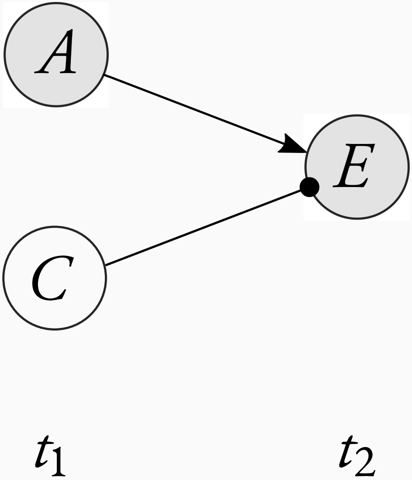
# Default Causes and Effects

- So far, we've only looked at causes and effects whose values are *deviant*
- Default variable values can also be causes and effects.

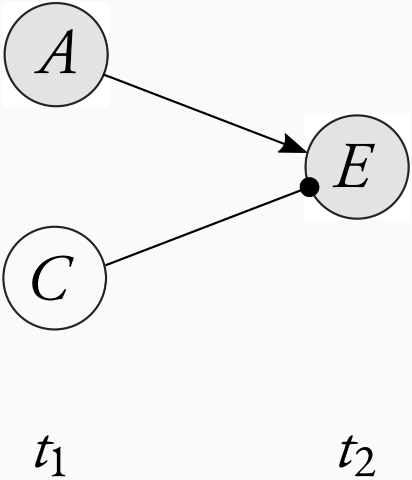
# Prevention



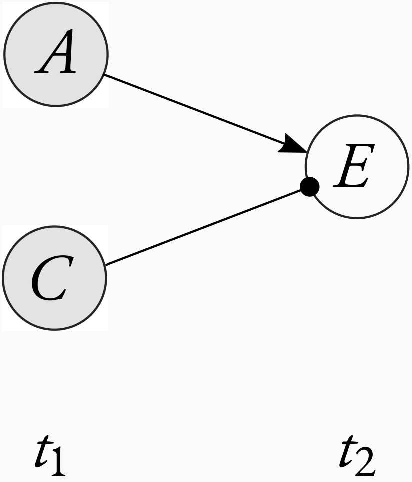
# Prevention



# Omission



# Omission



# Prevention without Dependence?

- When $C = c$ and $E = e$ are *deviant*, local counterfactual dependence suffices for causation.
- Does local counterfactual dependence suffice for causation when $C = c$ or $E = e$ are *default*, too?

# Prevention without Dependence?

- When $C = c$ and $E = e$ are *deviant*, local counterfactual dependence suffices for causation.
- Does local counterfactual dependence suffice for causation when $C = c$ or $E = e$ are *default*, too?

# Prevention without Dependence?

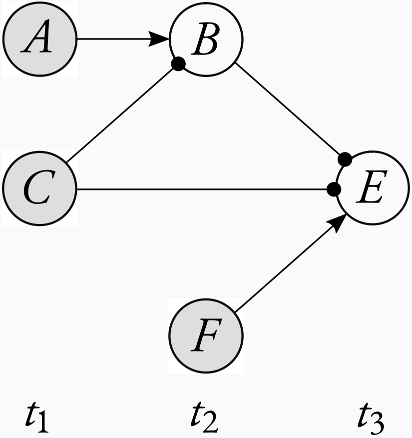



**Figure 8:** *Prevention without Dependence?*

# Prevention without Dependence?

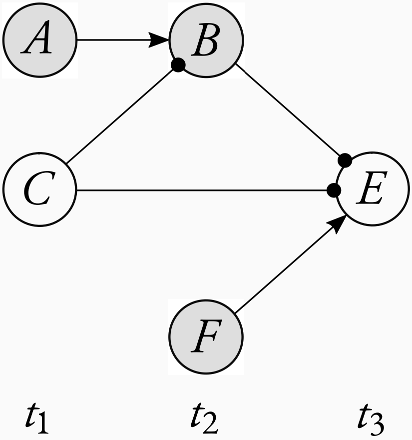
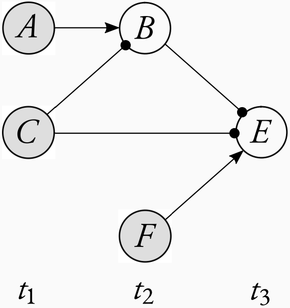



**Figure 8:** *Prevention without Dependence?*

# Prevention without Dependence?



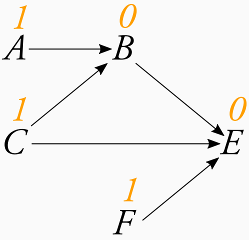

The model $\mathbb{M}_7$:

$\mathbb{U} : (A, C, F)$

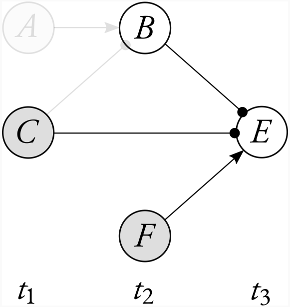
$\vec{u} : (1, 1, 1)$

$\mathbb{V} : (B, E)$

$$\mathbb{U} : \begin{pmatrix} E & := F \wedge \neg B \wedge \neg D \\ B & := A \wedge \neg C \end{pmatrix}$$



# Prevention without Dependence?



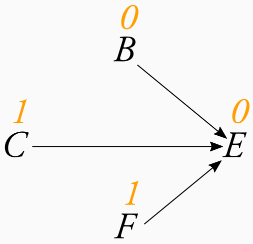

The *local model* $\mathbb{M}_7((E))$:

$\mathbb{U} : (B, C, F)$

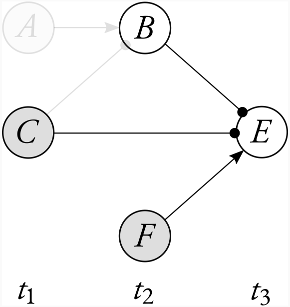
$\vec{u} : (0, 1, 1)$

$\mathbb{V} : (E)$

$\mathbb{U} : (E := F \wedge \neg B \wedge \neg D)$



# Prevention without Dependence?



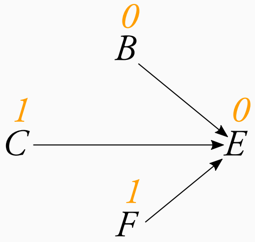

The model $\mathbb{M}_7(\langle E \rangle)$:

$\mathbb{U} : (B, C, F)$

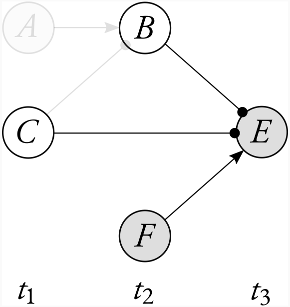
$\vec{u} : (0, 1, 1)$

$\mathbb{V} : (E)$

$\mathbb{U} : (E := F \wedge \neg B \wedge \neg D)$



# Prevention without Dependence?



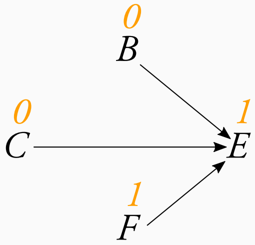

The model $\mathbb{M}_7((E))[C \to 0]$:

$\mathbb{U} : (B, C, F)$

$\vec{u} : (0, 0, 1)$

$\mathbb{V} : (E)$

$\mathbb{U} : (E := F \wedge \neg B \wedge \neg D)$



# Prevention without Dependence?

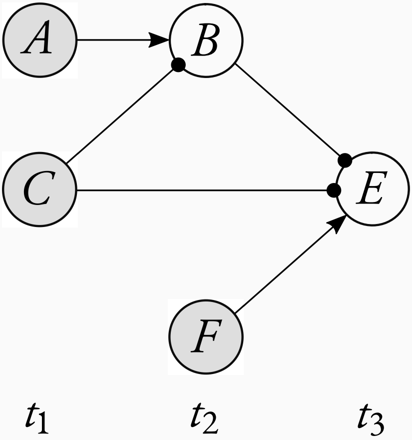



**Figure 7:** *Prevention without Dependence?*

# Prevention without Dependence?

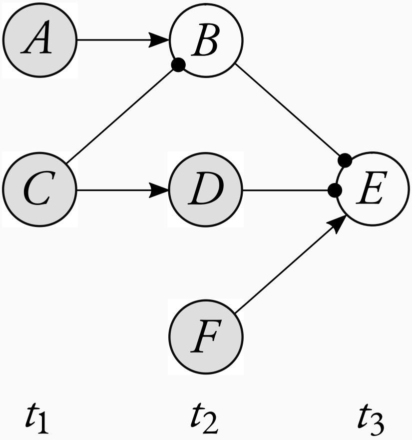



**Figure 7:** *Prevention without Dependence?*

# Prevention without Dependence?

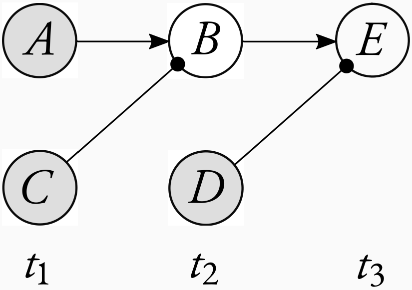



**Figure 8:** *Prevention without Dependence?*

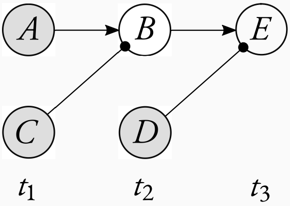
# Prevention without Dependence?

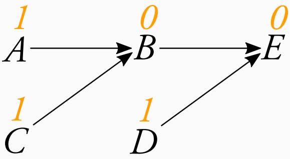
$\mathbb{U} : (A, C, D)$

$\vec{u} : (1, 1, 1)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \left( \begin{array}{ll} E & := B \wedge \neg D \\ B & := A \wedge \neg C \end{array} \right)$$



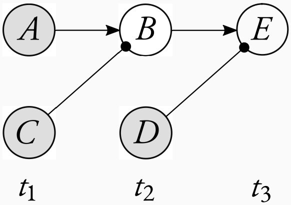
# Prevention without Dependence?

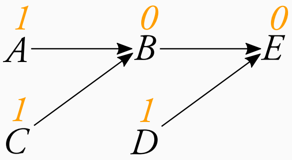
$\mathbb{U} : (A, C, D)$

$\vec{u} : (1, 1, 1)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \wedge \neg D \\ B & := A \wedge \neg C \end{pmatrix}$$



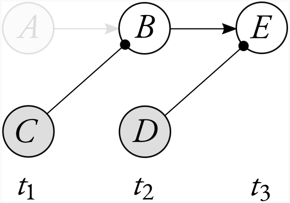
# Prevention without Dependence?

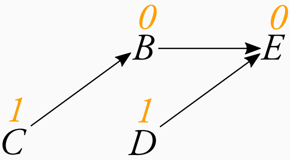
$\mathbb{U} : (C, D)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \wedge \neg D \\ B & := 1 \wedge \neg C \end{pmatrix}$$



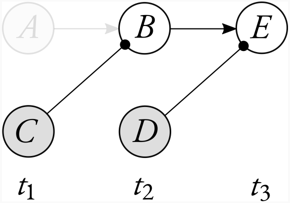
# Prevention without Dependence?

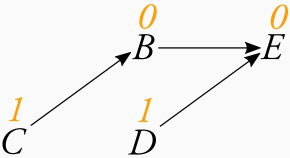
$\mathbb{U} : (C, D)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \land \neg D \\ B & := \neg C \end{pmatrix}$$



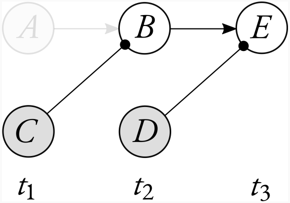
# Prevention without Dependence?

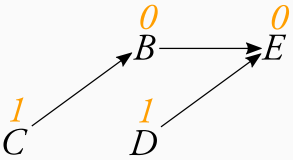
$$\mathbb{U} : (C, D)$$

$$\vec{u} : (1, 1)$$

$$\mathbb{V} : (B, E)$$

$$\mathbb{E} : \begin{pmatrix} E & := B \land \neg D \\ B & := \neg C \end{pmatrix}$$



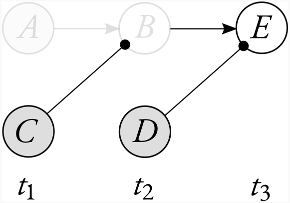
# Prevention without Dependence?

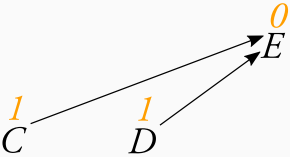
$\mathbb{U} : (C, D)$

$\vec{u} : (1, 1)$

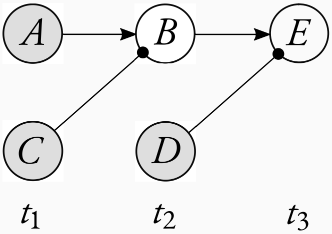
$\mathbb{V} : (E)$

$\mathbb{E} : \left( \; E := \neg C \wedge \neg D \; \right)$



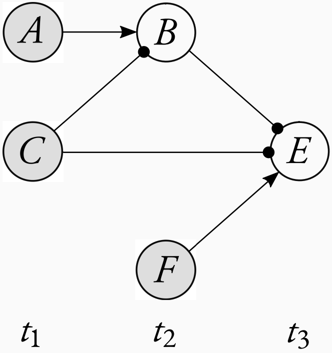



# Prevention without Dependence?




- Any model-invariant account must say that *C*'s firing kept *E* from firing iff *D*'s firing kept *E* from firing.

# Prevention without Dependence?




- Any model-invariant account must say that *C*'s firing kept *E* from firing iff *D*'s firing kept *E* from firing.

# Omission without Dependence?

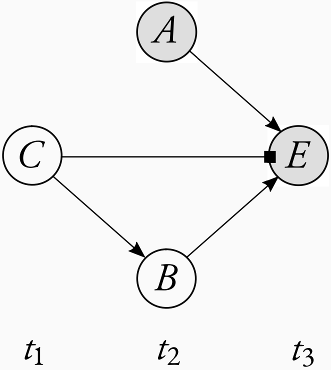



**Figure 9:** *Omission without Dependence?*

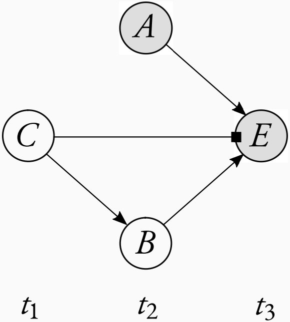
# Omission without Dependence?

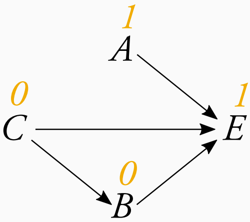
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 0)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := (A + B) > C \\ B & := C \end{pmatrix}$$



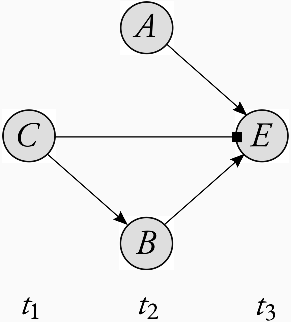
# Omission without Dependence?

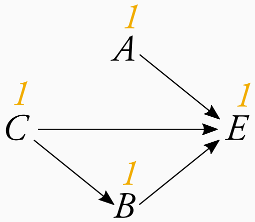
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := (A + B) > C \\ B & := C \end{pmatrix}$$



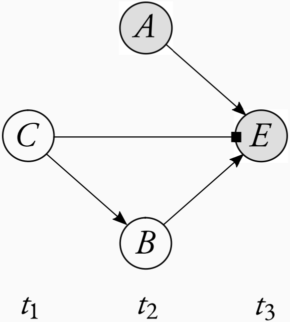
# Omission without Dependence?

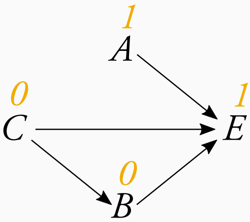
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 0)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := (A + B) > C \\ B & := C \end{pmatrix}$$



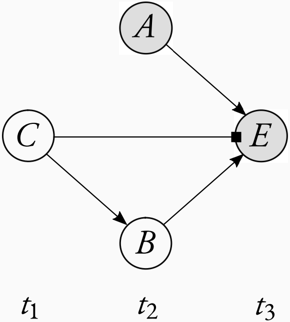
# Omission without Dependence?

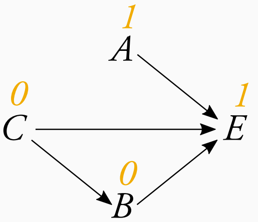
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 0)$

$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := (A + B) > C \\ B & := C \end{pmatrix}$$



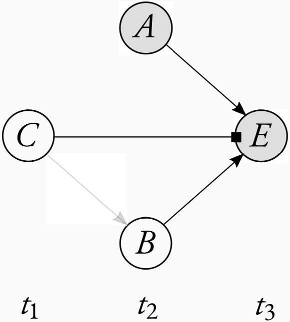
# Omission without Dependence?

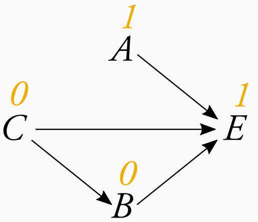
$\mathbb{U} : (A, C, B)$

$\vec{u} : (1, 0)$

$\mathbb{V} : (E)$

$$\mathbb{E} : \begin{pmatrix} E & := (A + B) > C \\ B & := C \end{pmatrix}$$



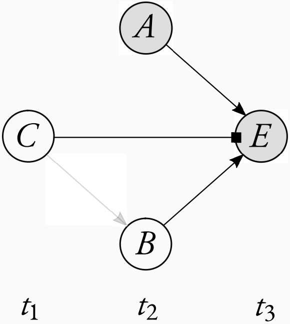
# Omission without Dependence?

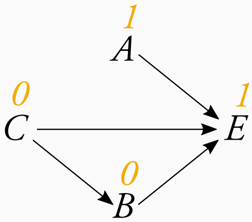
$\mathbb{U} : (A, C, B)$

$\vec{u} : (1, 0, 0)$

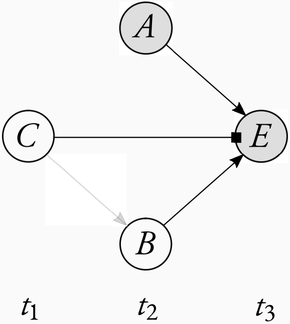
$\mathbb{V} : (E)$

$$\mathbb{E} : \begin{pmatrix} E & := (A + B) > C \\ B & := C \end{pmatrix}$$



# Omission without Dependence?

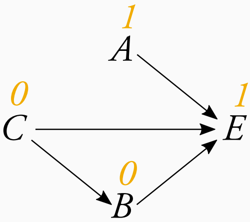



$\mathbb{U} : (A, C, B)$

$\vec{u} : (1, 0, 0)$

$\mathbb{V} : (E)$

$\mathbb{E} : \left( \; E := (A + B) > C \; \right)$



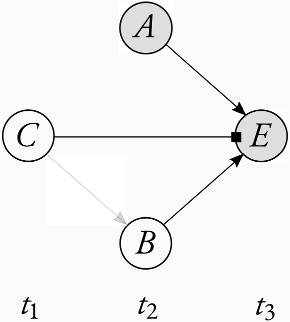
# Omission without Dependence?

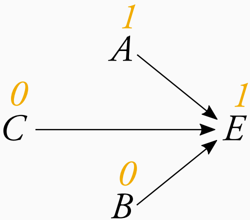
$\mathbb{U} : (A, C, B)$

$\vec{u} : (1, 0, 0)$

$\mathbb{V} : (E)$

$\mathbb{E} : \left(\ E := (A + B) > C\ \right)$



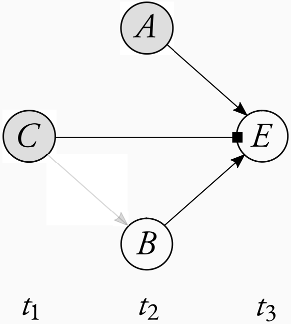
# Omission without Dependence?

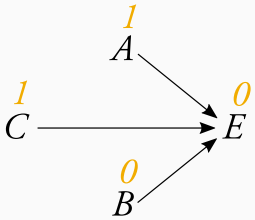
$\mathbb{U} : (A, C, B)$

$\vec{u} : (1, 1, 0)$

$\mathbb{V} : (E)$

$\mathbb{E} : \left( \; E := (A + B) > C \; \right)$



# Omission without Dependence?

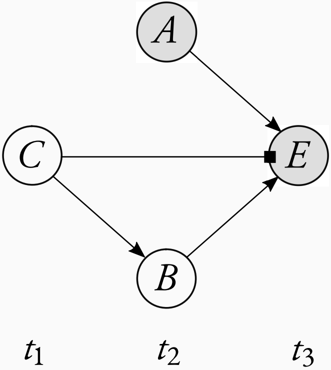



**Figure 9:** *Omission without Dependence?*

# Omission without Dependence?

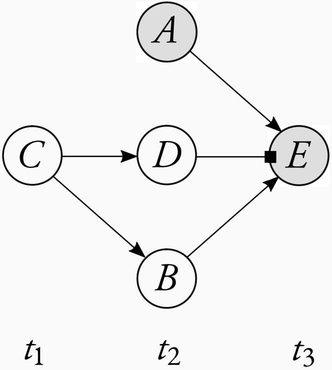



**Figure 9:** *Omission without Dependence?*

# Local Dependence

- If our account is to be model-invariant, and we accept TRANSITIVE PATH, then there can be no prevention or omission without global counterfactual dependence.
- Local counterfactual dependence only suffices for causation when
    - Both $c$ and $e$ are *deviant* values (and both $c^*$ and $e^*$ *default*)

# Local Dependence

- If our account is to be model-invariant, and we accept Transitive Path, then there can be no prevention or omission without global counterfactual dependence.
- Local counterfactual dependence only suffices for causation when
    - Both $c$ and $e$ are *deviant* values (and both $c^*$ and $e^*$ *default*)

# Local Dependence

- If our account is to be model-invariant, and we accept TRANSITIVE PATH, then there can be no prevention or omission without global counterfactual dependence.
- Local counterfactual dependence only suffices for causation when
    - Both $c$ and $e$ are *deviant* values (and both $c^*$ and $e^*$ *default*)

# A Model Invariant Theory

## Causation as Production and Dependence

### Causation as Production and Dependence

In a causal model $\mathbb{M}$, $C$'s taking on the value $c$, rather than $c^*$, caused $E$ to take on the value $e$, rather than $e^*$, iff either (Prod) or (Dep).

(Prod) Both $c$ and $e$ are *deviant* variable values, the contrasts $c^*$ and $e^*$ *defaults*, and either:

i. In the local model at $E$, $\mathbb{M}((E))$, had $C$ taken on the value $c^*$, $E$ would have taken on the value $e^*$,

$$\mathbb{M}((E)) \models C = c^* \;\Box\!\!\rightarrow E = e^*$$

or

ii. In $\mathbb{M}$, there is a transitive path leading from $C$ to $E$.

# Causation as Production and Dependence

## Causation as Production and Dependence

In a causal model $\mathbb{M}$, $C$'s taking on the value $c$, rather than $c^*$, caused $E$ to take on the value $e$, rather than $e^*$, iff either (Prod) or (Dep).

(Prod) Both $c$ and $e$ are *deviant* variable values, the contrasts $c^*$ and $e^*$ *defaults*, and either:

i. In the local model at $E$, $\mathbb{M}(\!(E)\!)$, had $C$ taken on the value $c^*$, $E$ would have taken on the value $e^*$,

$$\mathbb{M}(\!(E)\!) \models C = c^* \;\Box\!\!\rightarrow E = e^*$$

or

ii. In $\mathbb{M}$, there is a transitive path leading from $C$ to $E$.

## Causation as Production and Dependence

In a causal model $\mathbb{M}$, $C$'s taking on the value $c$, rather than $c^*$, caused $E$ to take on the value $e$, rather than $e^*$, iff either (Prod) or (Dep).

(Prod) Both $c$ and $e$ are *deviant* variable values, the contrasts $c^*$ and $e^*$ *defaults*, and either:

i. In the local model at $E$, $\mathbb{M}(\!(E)\!)$, had $C$ taken on the value $c^*$, $E$ would have taken on the value $e^*$,

$$\mathbb{M}(\!(E)\!) \models C = c^* \mathrel{\Box\!\!\rightarrow} E = e^*$$

or

ii. In $\mathbb{M}$, there is a transitive path leading from $C$ to $E$.

# Causation as Production and Dependence

## Causation as Production and Dependence

In a causal model $\mathbb{M}$, $C$'s taking on the value $c$, rather than $c^*$, caused $E$ to take on the value $e$, rather than $e^*$, iff either (Prod) or (Dep).

(Prod) Both $c$ and $e$ are *deviant* variable values, the contrasts $c^*$ and $e^*$ *defaults*, and either:

i. In the local model at $E$, $\mathbb{M}(\!(E)\!)$, had $C$ taken on the value $c^*$, $E$ would have taken on the value $e^*$,

$$\mathbb{M}(\!(E)\!) \models C = c^* \;\Box\!\!\rightarrow E = e^*$$

or

ii. In $\mathbb{M}$, there is a transitive path leading from $C$ to $E$.

### Causation as Production and Dependence

In a causal model $\mathbb{M}$, $C$'s taking on the value $c$, rather than $c^*$, caused $E$ to take on the value $e$, rather than $e^*$, iff either (Prod) or (Dep).

(Dep) In $\mathbb{M}$, had $C$ taken on the value $c^*$, $E$ would have taken on the value $e^*$,

$$\mathbb{M} \models C = c^* \;\Box\!\!\rightarrow E = e^*$$

## Causation as Production and Dependence

In a causal model $\mathbb{M}$, $C$'s taking on the value $c$, rather than $c^*$, caused $E$ to take on the value $e$, rather than $e^*$, iff either (Prod) or (Dep).

(Dep) In $\mathbb{M}$, had $C$ taken on the value $c^*$, $E$ would have taken on the value $e^*$,

$$\mathbb{M} \models C = c^* \mathrel{\Box\!\!\rightarrow} E = e^*$$

# Model-Invariance

- Suppose that we have a correct model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $U \in \mathbb{U}$ and $V \in \mathbb{V}$.
  - Neither $U$ nor $V$ are $C$ or $E$
  - $U$ and $V$ are both inessential
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-V}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-V}$

# Model-Invariance

- Suppose that we have a correct model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $U \in \mathbb{U}$ and $V \in \mathbb{V}$.
    - Neither $U$ nor $V$ are $C$ or $E$
    - $U$ and $V$ are both inessential
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-V}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-V}$

## Model-Invariance

- Suppose that we have a correct model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $U \in \mathbb{U}$ and $V \in \mathbb{V}$.
    - Neither $U$ nor $V$ are $C$ or $E$
    - $U$ and $V$ are both inessential
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-V}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-V}$

## Model-Invariance

- Suppose that we have a correct model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $U \in \mathbb{U}$ and $V \in \mathbb{V}$.
  - Neither $U$ nor $V$ are $C$ or $E$
  - $U$ and $V$ are both inessential
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-V}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-V}$

## Model-Invariance

- Suppose that we have a correct model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $U \in \mathbb{U}$ and $V \in \mathbb{V}$.
    - Neither $U$ nor $V$ are $C$ or $E$
    - $U$ and $V$ are both inessential
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-V}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-V}$

# Model-Invariance

- Suppose that we have a correct model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $U \in \mathbb{U}$ and $V \in \mathbb{V}$.
  - Neither $U$ nor $V$ are $C$ or $E$
  - $U$ and $V$ are both inessential
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-V}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-V}$

# Model-Invariance

- Suppose that we have a correct model $\mathbb{M} = (\mathbb{U}, \vec{u}, \mathbb{V}, \mathbb{E}, \mathscr{D})$, with $U \in \mathbb{U}$ and $V \in \mathbb{V}$.
    - Neither $U$ nor $V$ are $C$ or $E$
    - $U$ and $V$ are both inessential
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ caused $E = e$ in $\mathbb{M}$, then $C = c$ caused $E = e$ in $\mathbb{M}^{-V}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-U}$
- If $C = c$ *didn't* cause $E = e$ in $\mathbb{M}$, then $C = c$ *didn't* cause $E = e$ in $\mathbb{M}^{-V}$

# Model Invariance

- Thus, this account is consistent with **Model Invariance**, **Exogenous Reduction**, and **Endogenous Reduction**.

# Causation as Production and Dependence

- PROD does a reasonably good job at capturing a notion of *causal production*.
    - Production involves uninterupted, local propogation of deviant, non-interial states of affairs (rather than default, inertial states of affairs)

# Causation as Production and Dependence

- Prod does a reasonably good job at capturing a notion of *causal production*.
    - Production involves uninterupted, local propogation of deviant, non-interial states of affairs (rather than default, inertial states of affairs)

## Causation as Production and Dependence

In a causal model $\mathbb{M}$, $C$'s taking on the value $c$, rather than $c^*$, caused $E$ to take on the value $e$, rather than $e^*$, iff either (Prod) or (Dep).

(Prod) Both $c$ and $e$ are *deviant* variable values, the contrasts $c^*$ and $e^*$ *defaults*, and either:

i. In the local model at $E$, $\mathbb{M}((E))$, had $C$ taken on the value $c^*$, $E$ would have taken on the value $e^*$,

$$\mathbb{M}((E)) \models C = c^* \;\Box\!\!\rightarrow E = e^*$$

or

ii. In $\mathbb{M}$, there is a transitive path leading from $C$ to $E$.

**Causation as Production and Dependence**

In a causal model $\mathbb{M}$, $C$'s taking on the value $c$, rather than $c^*$, caused $E$ to take on the value $e$, rather than $e^*$, iff either (Prod) or (Dep).

(Dep) In $\mathbb{M}$, had $C$ taken on the value $c^*$, $E$ would have taken on the value $e^*$,

$$\mathbb{M} \models C = c^* \;\Box\!\!\rightarrow E = e^*$$

# Causation as Production and Dependence

- Hypothesis: the notion of causal production encapsulated in PROD represents the core of our concept of causation. The causal judgments licensed by the PROD clause alone are far more intuitive, natural, and widespread that those which are only licensed with the addition of the DEP clause.
    - E.g., *preemptive overdetermination*, as opposed to
    - *prevention*, *omission*, *omissive prevention*, and *double prevention*.
- However, if you accept the PROD clause, then the full strength of DEP is required in order for the account to be model-invariant.

# Causation as Production and Dependence

- Hypothesis: the notion of causal production encapsulated in PROD represents the core of our concept of causation. The causal judgments licensed by the PROD clause alone are far more intuitive, natural, and widespread that those which are only licensed with the addition of the DEP clause.
    - E.g., *preemptive overdetermination*, as opposed to
    - *prevention*, *omission*, *omissive prevention*, and *double prevention*.
- However, if you accept the PROD clause, then the full strength of DEP is required in order for the account to be model-invariant.

# Causation as Production and Dependence

- Hypothesis: the notion of causal production encapsulated in PROD represents the core of our concept of causation. The causal judgments licensed by the PROD clause alone are far more intuitive, natural, and widespread that those which are only licensed with the addition of the DEP clause.
    - E.g., *preemptive overdetermination*, as opposed to
    - *prevention*, *omission*, *omissive prevention*, and *double prevention*.
- However, if you accept the PROD clause, then the full strength of DEP is required in order for the account to be model-invariant.

# Causation as Production and Dependence

- Hypothesis: the notion of causal production encapsulated in PROD represents the core of our concept of causation. The causal judgments licensed by the PROD clause alone are far more intuitive, natural, and widespread that those which are only licensed with the addition of the DEP clause.
    - E.g., *preemptive overdetermination*, as opposed to
    - *prevention*, *omission*, *omissive prevention*, and *double prevention*.
- However, if you accept the PROD clause, then the full strength of DEP is required in order for the account to be model-invariant.

# Double Prevention without Dependence

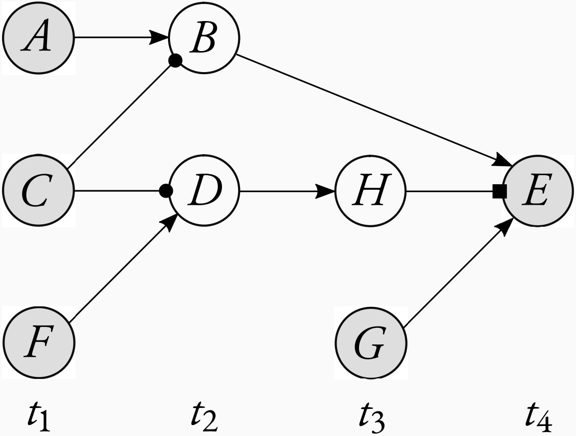



**Figure 10:** *Double Prevention without Dependence*

# Double Prevention without Dependence

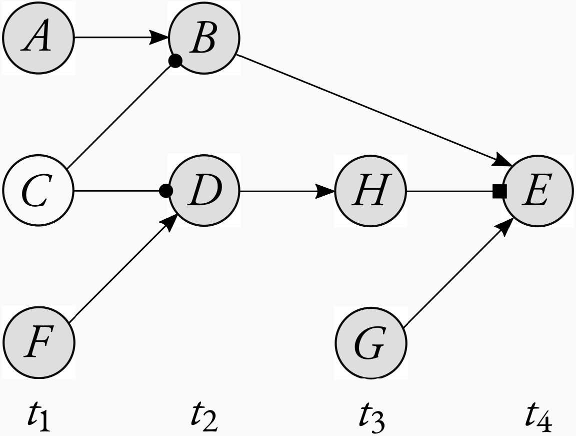



**Figure 10:** *Double Prevention without Dependence*

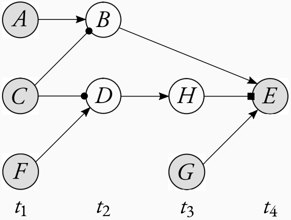
# Double Prevention without Dependence

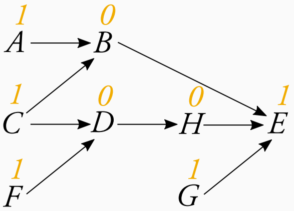
$\mathbb{U} : (A, C, F, G)$

$\vec{u} : (1, 1, 1, 1)$

$\mathbb{V} : (B, D, H, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B + G > H \\ H & := D \\ D & := F \wedge \neg C \\ B & := A \wedge \neg C \end{pmatrix}$$



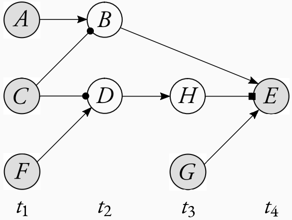
# Double Prevention without Dependence

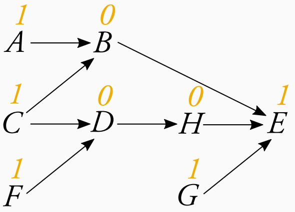
$\mathbb{U} : (A, C, F, G)$

$\vec{u} : (1, 1, 1, 1)$

$\mathbb{V} : (B, D, H, E)$

$$\mathbb{E} : \left( \begin{array}{rl} E & := B + G > H \\ H & := D \\ D & := F \wedge \neg C \\ B & := A \wedge \neg C \end{array} \right)$$



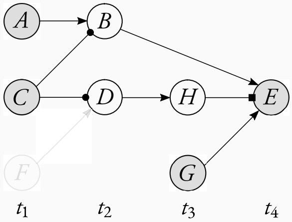
# Double Prevention without Dependence

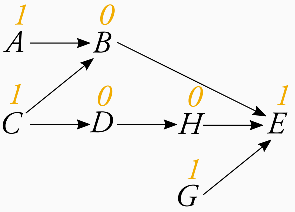
$\mathbb{U} : (A, C, G)$

$\vec{u} : (1, 1, 1)$

$\mathbb{V} : (B, D, H, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B + G > H \\ H & := D \\ D & := \neg C \\ B & := A \wedge \neg C \end{pmatrix}$$



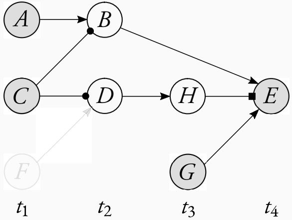
# Double Prevention without Dependence

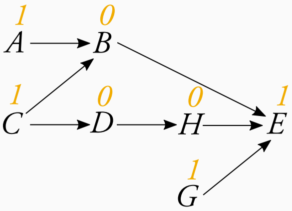
$\mathbb{U} : (A, C, G)$

$\vec{u} : (1, 1, 1)$

$\mathbb{V} : (B, D, H, E)$

$$\mathbb{E} : \left( \begin{array}{l} E := B + G > H \\ H := D \\ D := \neg C \\ B := A \wedge \neg C \end{array} \right)$$



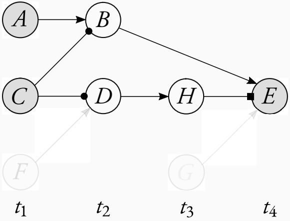
# Double Prevention without Dependence

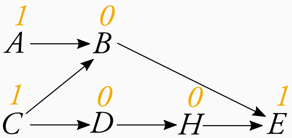
$$\mathbb{U} : (A, C)$$

$$\vec{u} : (1, 1)$$

$$\mathbb{V} : (B, D, H, E)$$

$$\mathbb{E} : \begin{pmatrix} E & := B > H \\ H & := D \\ D & := \neg C \\ B & := A \wedge \neg C \end{pmatrix}$$



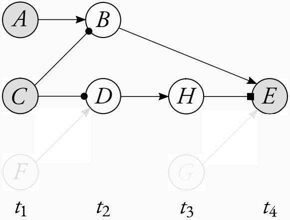
# Double Prevention without Dependence

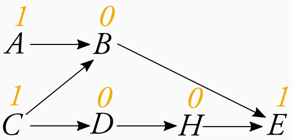
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, D, H, E)$

$$\mathbb{E} : \left( \begin{array}{rl} E & := B > H \\ H & := D \\ D & := \neg C \\ B & := A \wedge \neg C \end{array} \right)$$



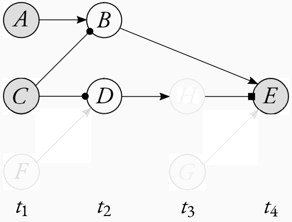
# Double Prevention without Dependence

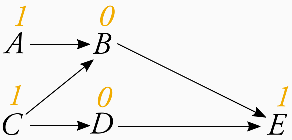
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E := B > D \\ D := \neg C \\ B := A \wedge \neg C \end{pmatrix}$$



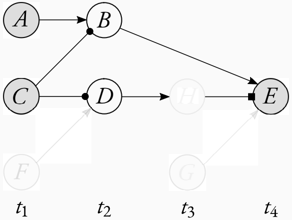
# Double Prevention without Dependence

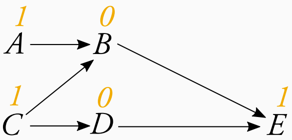
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B > D \\ D & := \neg C \\ B & := A \wedge \neg C \end{pmatrix}$$



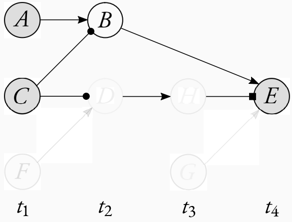
# Double Prevention without Dependence

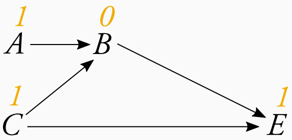
$\mathbb{U} : (A, C)$

$\vec{u} : (1, 1)$

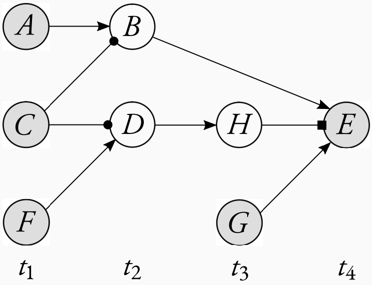
$\mathbb{V} : (B, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \lor C \\ B & := A \land \neg C \end{pmatrix}$$



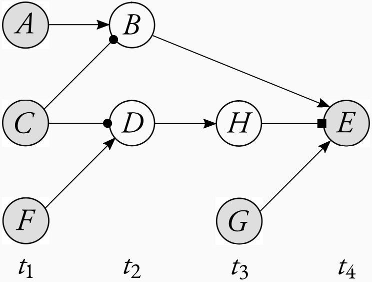



# Double Prevention without Dependence



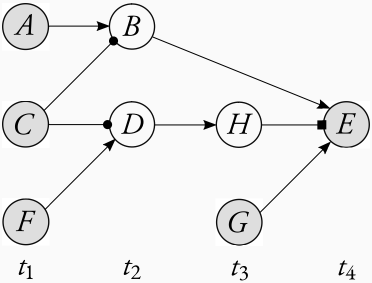

- If we wish our account to be model-invariant, and we wish to secure the intuitive verdict in *Preemptive Overdetermination*, then we must say that *C*'s firing caused *E*'s firing in figure 10.

# Double Prevention without Dependence




- So, we must have a transitive path running from $C$ to $E$.

# Double Prevention without Dependence



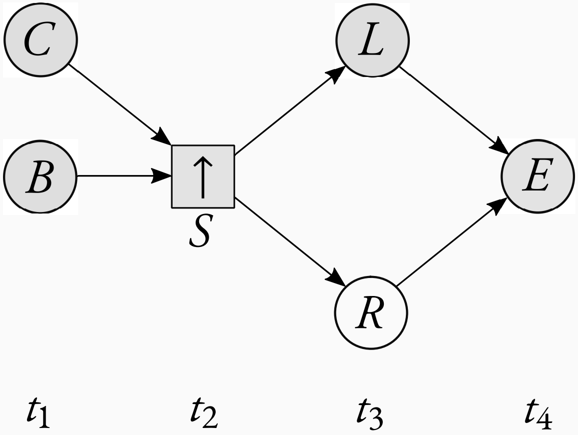

- So, we must make use of the full strength of the DEP clause.

Thank you!

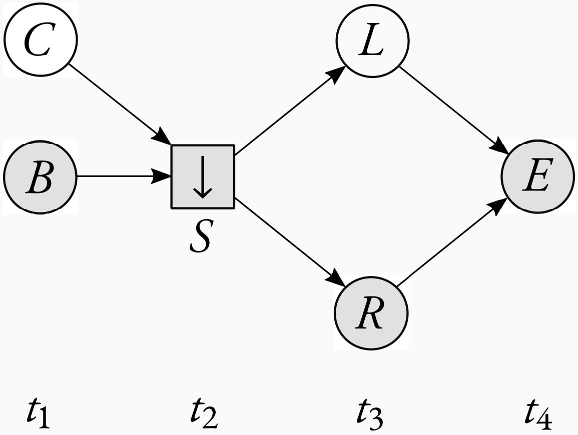
Questions?

# Extras

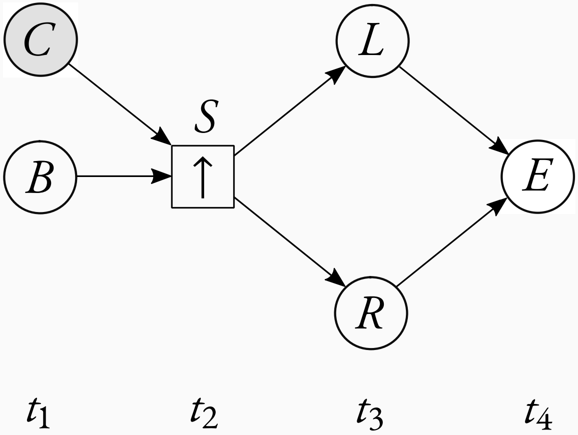
# Switching



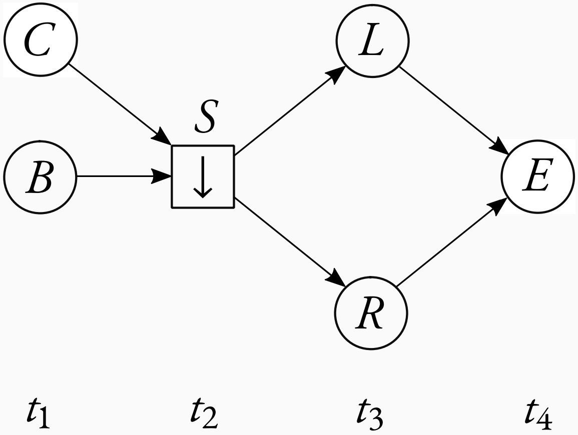
# Switching



# Switching



# Switching



# Counterexamples to Transitivity

'*Shock Edward*' is a game for three players. Carol and David each has a switch with two positions, Left and Right; to start, both are in the Left position. Carol has first turn: she can either move her switch to Right, or do nothing. David then has a turn: he can either move his switch to Right, or do nothing. The power is then turned on: if both switches are in the Left position, or both in the Right position, Edward gets an electric shock.

On this occasion the play goes as follows. Carol moves her switch to Right. David observes Carol's move; he wants Edward to get a shock, so he responds by moving his switch to Right also. Edward duly gets a shock. (McDermott, 1995)

# Counterexamples to Transitivity

- A natural reading of *Shock Edward*:
    - Carol is trying to prevent Edward from getting a shock.
    - She is one of the good guys
    - What she did is *default*.

# Counterexamples to Transitivity

- A natural reading of *Shock Edward*:
    - Carol is trying to prevent Edward from getting a shock.
    - She is one of the good guys
    - What she did is *default*.

# Counterexamples to Transitivity

- A natural reading of *Shock Edward*:
    - Carol is trying to prevent Edward from getting a shock.
    - She is one of the good guys
    - What she did is *default*.

# Counterexamples to Transitivity

- A natural reading of *Shock Edward*:
    - Carol is trying to prevent Edward from getting a shock.
    - She is one of the good guys
    - What she did is *default*.

# Counterexamples to Transitivity

$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (D, E)$

$$\mathbb{E} : \begin{pmatrix} E & \coloneqq C = D \\ D & \coloneqq C \end{pmatrix}$$

$\mathscr{D} : C = 1$ is default

$$\begin{tikzcd} & \overset{1}{D} \arrow[dr] & \\ \overset{1}{C} \arrow[ur] \arrow[rr] & & \overset{1}{E} \end{tikzcd}$$

# Counterexamples to Transitivity

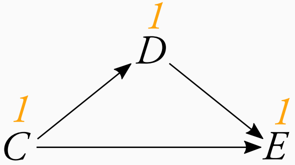
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := C = D \\ D & := C \end{pmatrix}$$

$\mathscr{D}$ : $C = 1$ is default



# Counterexamples to Transitivity

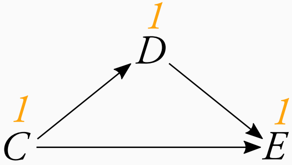
$\mathbb{U} : (C)$

$\vec{u} : (1)$

$\mathbb{V} : (D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := C = D \\ D & := C \end{pmatrix}$$

$\mathscr{D}$ : $C = 1$ is deviant



# Counterexamples to Transitivity

- If Carol's flipping the switch is deviant, then any model invariant theory of causation which gets the case of preemptive overdetermination correct will say that Carol's flipping the switch caused Edward to get a shock.

# Counterexamples to Transitivity

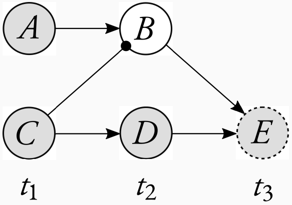
$\mathbb{U} : (A, C)$

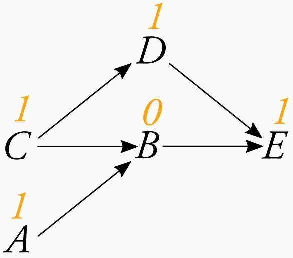
$\vec{u} : (1, 1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \veebar D \\ D & := C \\ B & := A \wedge \neg C \end{pmatrix}$$

$\mathscr{D}$ : $C = 1$ is deviant



# Counterexamples to Transitivity

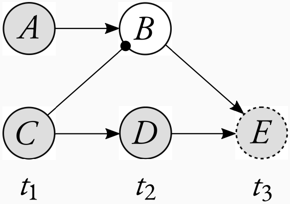
$\mathbb{U} : (A, C)$

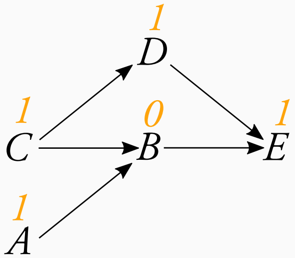
$\vec{u} : (1, 1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E := B \veebar D \\ D := C \\ B := A \wedge \neg C \end{pmatrix}$$

$\mathscr{D} : C = 1$ is deviant



# Counterexamples to Transitivity

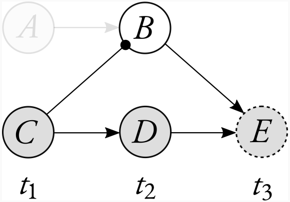
$\mathbb{U}:(C)$

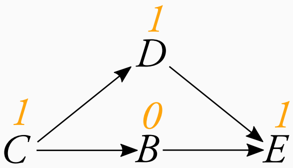
$\vec{u}:(1)$

$\mathbb{V}:(B,D,E)$

$$\mathbb{E}:\begin{pmatrix} E & := B \veebar D \\ D & := C \\ B & := \neg C \end{pmatrix}$$

$\mathscr{D}: C=1$ is deviant



# Counterexamples to Transitivity

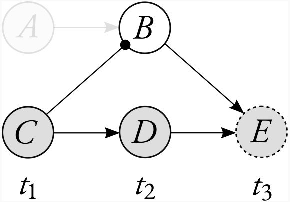
$\mathbb{U} : (C)$

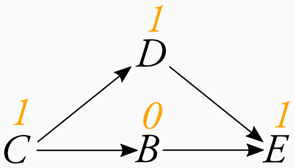
$\vec{u} : (1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E := B \veebar D \\ D := C \\ B := \neg C \end{pmatrix}$$

$\mathscr{D} : C = 1$ is deviant



# Counterexamples to Transitivity

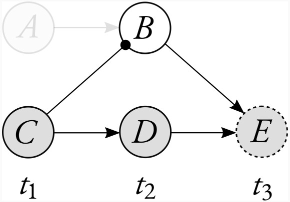
$\mathbb{U} : (C)$

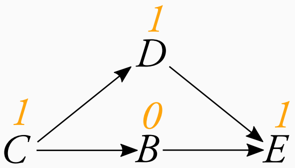
$\vec{u} : (1)$

$\mathbb{V} : (B, D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := B \veebar D \\ D & := C \\ B & := \neg C \end{pmatrix}$$

$\mathscr{D} : C = 1$ is deviant



# Counterexamples to Transitivity

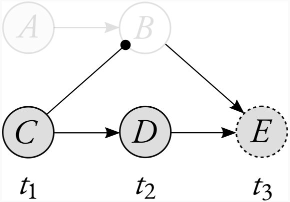
$\mathbb{U} : (C)$

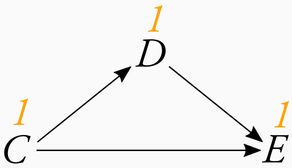
$\vec{u} : (1)$

$\mathbb{V} : (D, E)$

$$\mathbb{E} : \left( \begin{array}{rl} E & := \neg C \veebar D \\ D & := C \end{array} \right)$$

$\mathscr{D} : C = 1$ is deviant



# Counterexamples to Transitivity

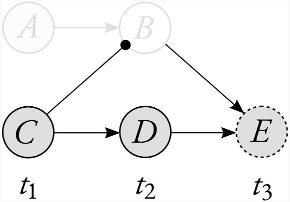
$\mathbb{U} : (C)$

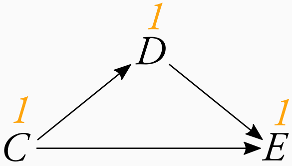
$\vec{u} : (1)$

$\mathbb{V} : (D, E)$

$$\mathbb{E} : \begin{pmatrix} E & := C = D \\ D & := C \end{pmatrix}$$

$\mathscr{D} : C = 1$ is deviant



# Counterexamples to Transitivity

- If you want
  - to give a model-invariant theory of causation and
  - to say that Carol's flipping the switch didn't cause Edward to get a shock,
- then you'll have to say that one of the variable values in *Shock Edward* is default.

# Counterexamples to Transitivity

- If you want
  - to give a model-invariant theory of causation and
  - to say that Carol's flipping the switch didn't cause Edward to get a shock,
- then you'll have to say that one of the variable values in *Shock Edward* is default.

# Counterexamples to Transitivity

- If you want
    - to give a model-invariant theory of causation and
    - to say that Carol's flipping the switch didn't cause Edward to get a shock,
- then you'll have to say that one of the variable values in *Shock Edward* is default.

# Counterexamples to Transitivity

- If you want
  - to give a model-invariant theory of causation and
  - to say that Carol's flipping the switch didn't cause Edward to get a shock,
- then you'll have to say that one of the variable values in *Shock Edward* is default.