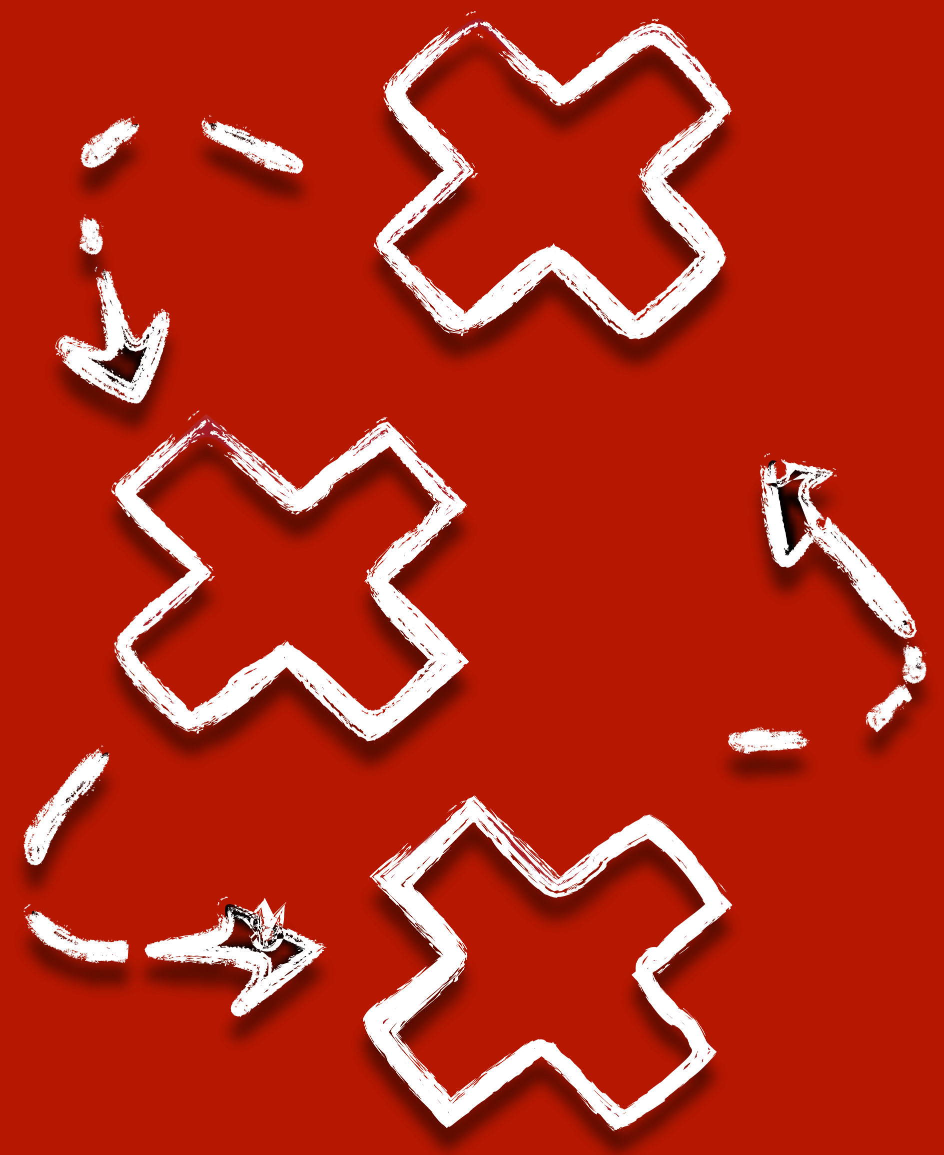


Causal Data Science

Lecture 11.1: Do-calculus and transportability

Lecturer: Sara Magliacane

UvA - Spring 2024

Let’s step back to Class 5 and 6

1	Introduction
2	Probability recap
3	Graphical models, d-separation
4	Causal graphs, Interventions, SCMs
5	Covariate adjustment: backdoor criterion
6	Covariate Frontdoor criterion, Instrumental variables
7	Counterfactuals, potential outcomes, estimating causal effects 1
8	Estimating causal effects 2 (matching, IPW)
9	Constraint based structure learning
10	Score based structure learning, restricted models
11	Do-calculus and transportability
12	Invariant Causal Prediction, Joint Causal Inference, Causality-inspired ML

We know the causal graph, how do we estimate causal effects?

(Some) identification strategies for causal effects

- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones
- **Backdoor criterion (class 5), Adjustment criterion (class 6)**

$$p(x_j \mid \mathrm{do}(x_i)) = \int_{x_{\mathbf{Z}}} p(x_j \mid x_i, x_{\mathbf{Z}}) p(x_{\mathbf{Z}}) dx_{\mathbf{Z}}$$

- **Frontdoor criterion (class 6)**

$$p(x_j \mid \mathrm{do}(x_i')) = \int_{x_{\mathbf{M}}} p(x_{\mathbf{M}} \mid x_i') \int_{x_i} p(x_j \mid x_{\mathbf{M}}, x_i') p(x_i) dx_i$$

- **Instrumental variables (class 6)**

(Some) identification strategies for causal effects

- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones
- **Backdoor criterion (class 5), Adjustment criterion (class 6)**

$$p(x_j \mid \mathrm{do}(x_i)) = \int_{x_\mathbf{Z}} p(x_j \mid x_i, x_\mathbf{Z}) p(x_\mathbf{Z}) dx_\mathbf{Z}$$

- **Frontdoor criterion (class 6)**

$$p(x_j \mid \mathrm{do}(x_i')) = \int_{x_\mathbf{M}} p(x_\mathbf{M} \mid x_i') \int_{x_i} p(x_j \mid x_\mathbf{M}, x_i') p(x_i) dx_i$$

- **Instrumental variables (class 6)**

None of these are **complete**

(they find all possible formulas to compute interventional distributions from observational ones)

(Some) identification strategies for causal effects

- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones
- **Backdoor criterion (class 5), Adjustment criterion (class 6)**

$$p(x_j \mid \mathrm{do}(x_i)) = \int_{x_{\mathbf{Z}}} p(x_j \mid x_i, x_{\mathbf{Z}}) p(x_{\mathbf{Z}}) dx_{\mathbf{Z}}$$

- **Frontdoor criterion (class 6)**

$$p(x_j \mid \mathrm{do}(x_i')) = \int_{x_{\mathbf{M}}} p(x_{\mathbf{M}} \mid x_i') \int_{x_i} p(x_j \mid x_{\mathbf{M}}, x_i') p(x_i) dx_i$$

- **Instrumental variables (class 6)**

None of these are **complete**

Do-calculus is complete

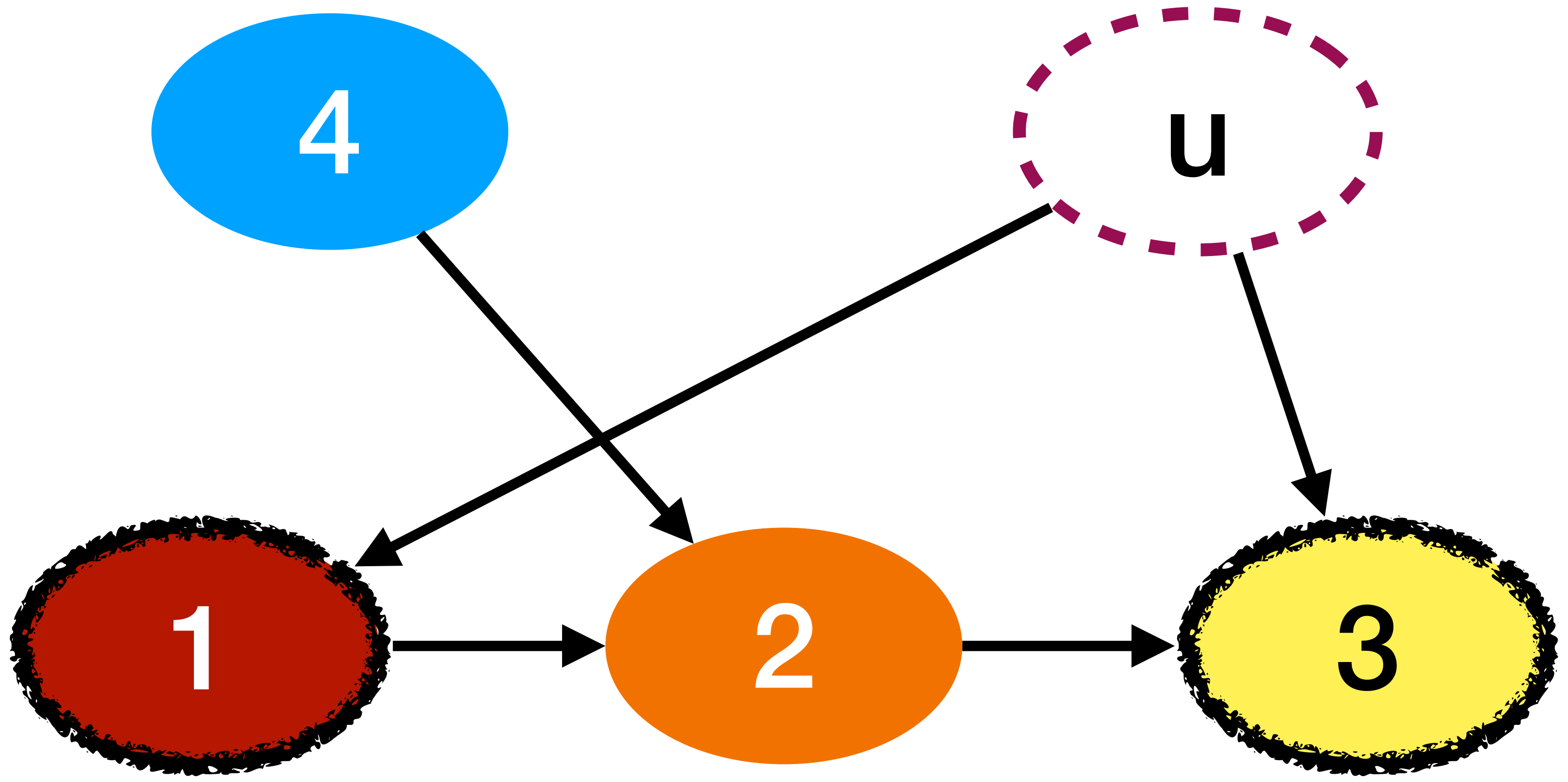
Advances topics: Do-calculus

1	Introduction
2	Probability recap
3	Graphical models, d-separation
4	Causal graphs, Interventions, SCMs
5	Covariate adjustment: backdoor criterion
6	Covariate Frontdoor criterion, Instrumental variables
7	Counterfactuals, potential outcomes, estimating causal effects 1
8	Estimating causal effects 2 (matching, IPW)
9	Constraint based structure learning
10	Score based structure learning, restricted models
11	Do-calculus and transportability
12	Invariant Causal Prediction, Joint Causal Inference, Causality-inspired ML

For more details see: https://canvas.uva.nl/courses/36213/pages/course-schedule-tentative?module_item_id=1585152

Advanced topics: Do-calculus

- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones

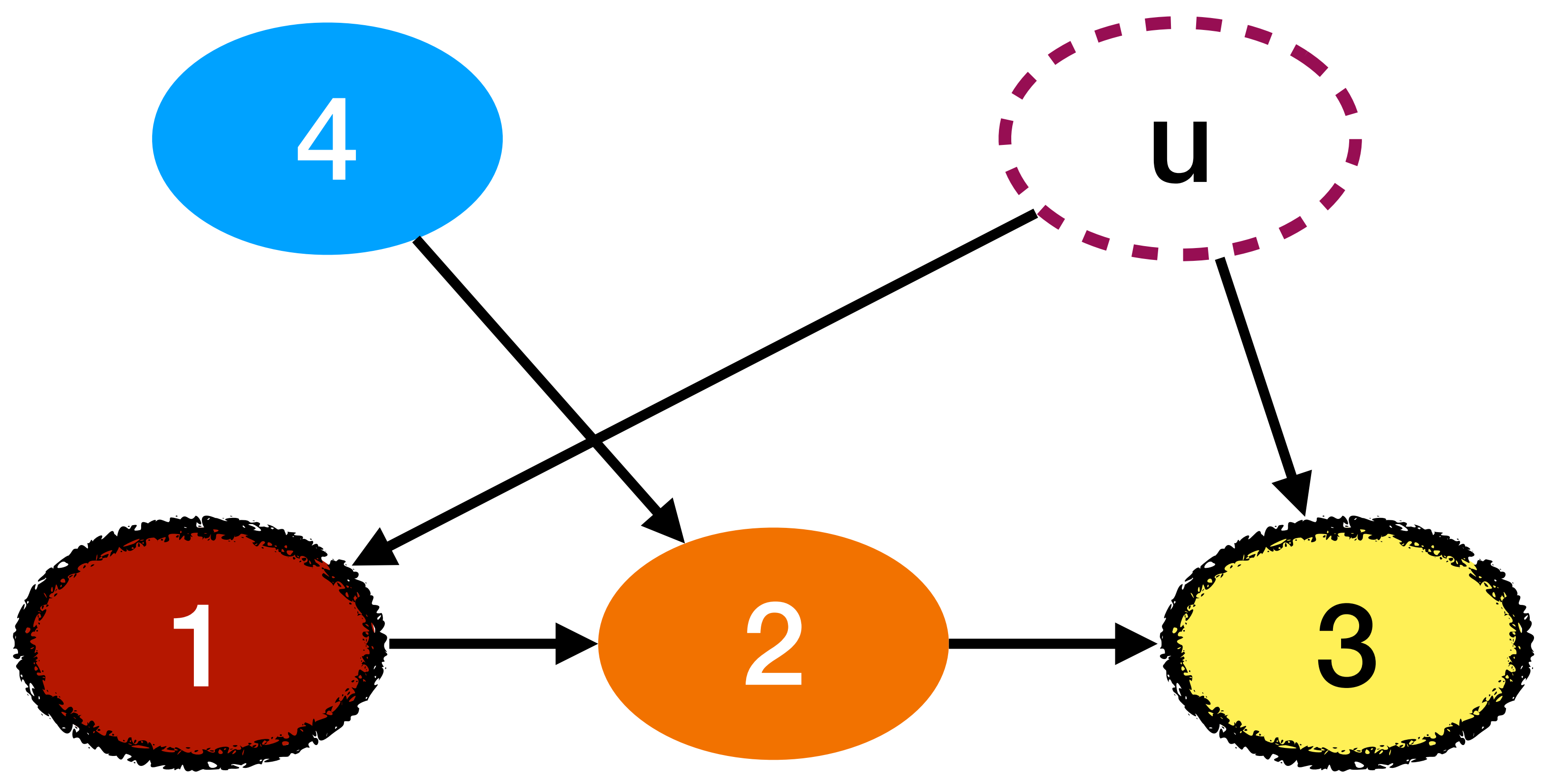


We want to estimate $P(X_3 \mid \mathrm{do}(X_1 = 1))$

Does this graph satisfy any graphical criterion?

Advanced topics: Do-calculus

- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones



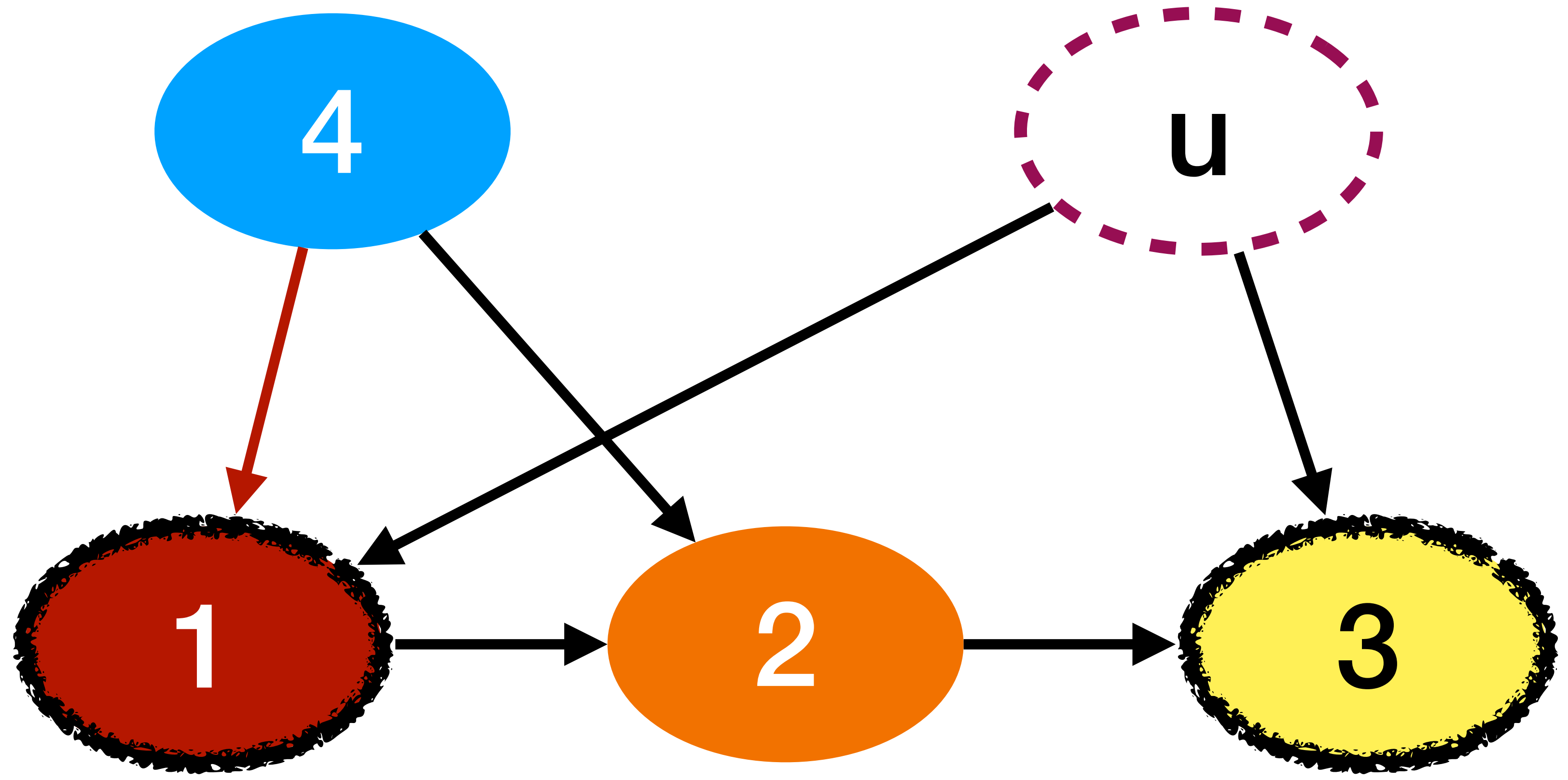
We want to estimate $P(X_3 \mid \mathrm{do}(X_1 = 1))$

Does this graph satisfy any graphical criterion?

1. $\mathbf{M}$ blocks all directed paths from i to j, **and**
2. No unblocked backdoor paths from $i \leftarrow \ldots$ to $\mathbf{M}$ with $Z = \emptyset$, **and**
3. $Z = \{i\}$ blocks all backdoor paths from $\mathbf{M} \leftarrow \ldots$ to j

Advanced topics: Do-calculus

- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones

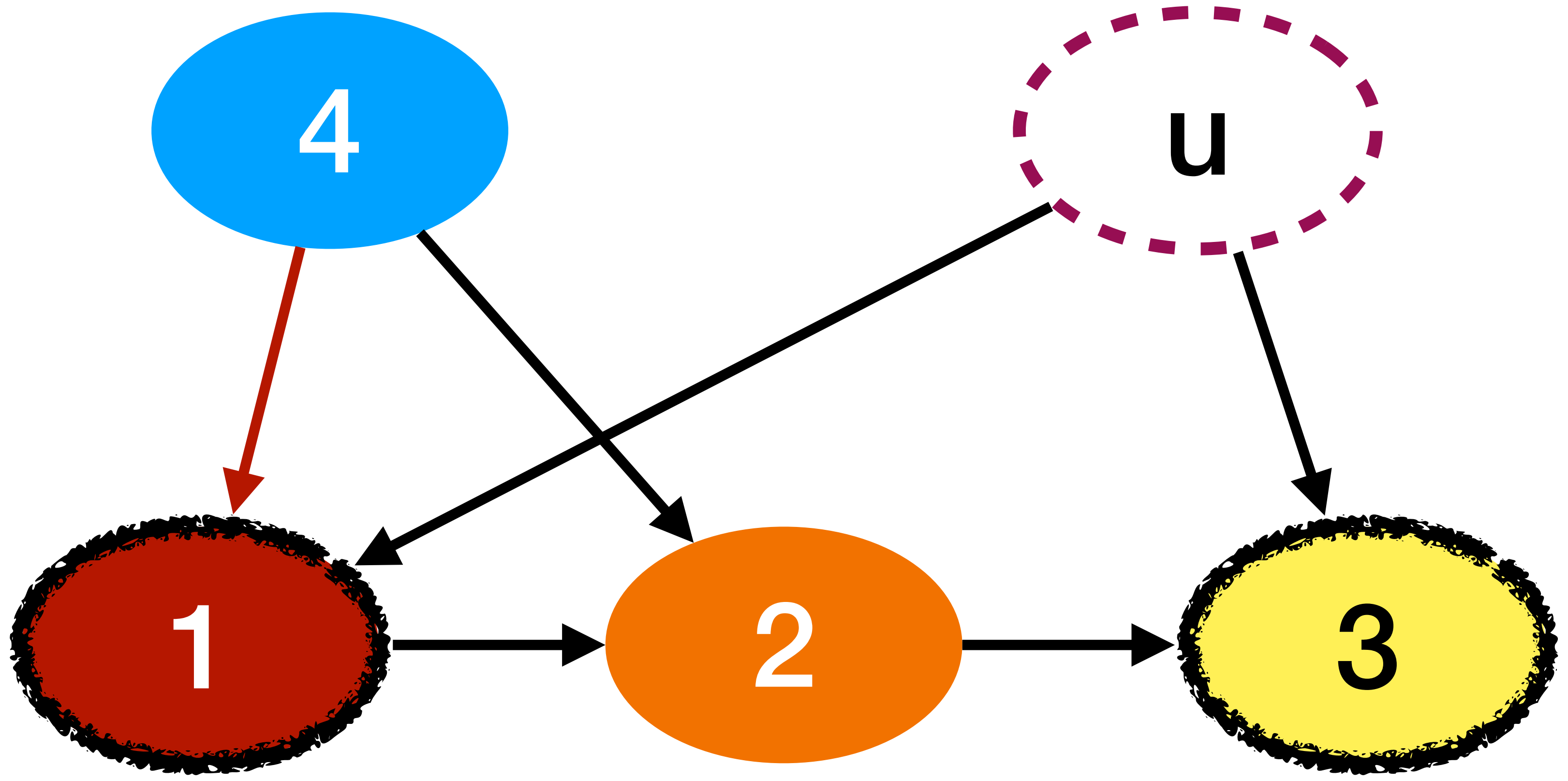


We want to estimate $P(X_3 \mid \mathrm{do}(X_1 = 1))$

And now, does this graph satisfy any graphical criterion?

Advanced topics: Do-calculus

- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones



Or maybe we want to estimate multiple treatments
$P(X_3 \mid \text{do}(X_1 = 1, X_4 = 1))$

Or maybe we want to estimate multiple outcomes
$P(X_3, X_2 \mid \text{do}(X_1 = 1, X_4 = 1))$

Advanced topics: Do-calculus

- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones
- **Do-calculus is complete:** if a formula (**an estimand**) exists for estimating an interventional distribution from observational data, it will find it

Advanced topics: Do-calculus

- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones
- **Do-calculus is complete:** if a formula (**an estimand**) exists for estimating an interventional distribution from observational data, it will find it
- **Calculus**: a set of rules that you can use to rewrite interventional distributions in equivalent expressions using interventional and observational distributions

Advanced topics: Do-calculus

- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones
- **Do-calculus is complete:** if a formula (**an estimand**) exists for estimating an interventional distribution from observational data, it will find it
- **Calculus**: a set of rules that you can use to rewrite interventional distributions in equivalent expressions using interventional and observational distributions
    - Keep applying rules until there are no interventional distributions left

Advanced topics: Do-calculus

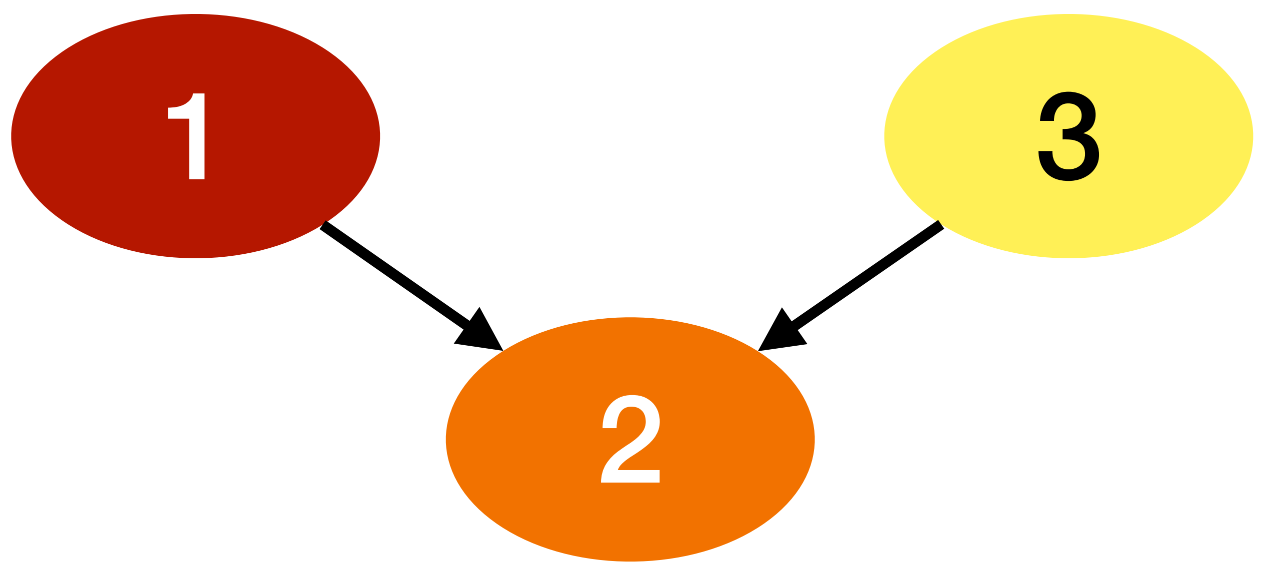
- Given a causal graph G, an **identification strategy** is a formula to estimate an interventional distribution from a combination of observational ones
- **Do-calcul** … nating an interventio …

If we cannot apply the rules anymore, but there are still interventional distributions, then the causal effect is not identifiable

- **Calculus**: a set of rules that you can use to rewrite interventional distributions in equivalent expressions using interventional and observational distributions
  - Keep applying rules until there are no interventional distributions left

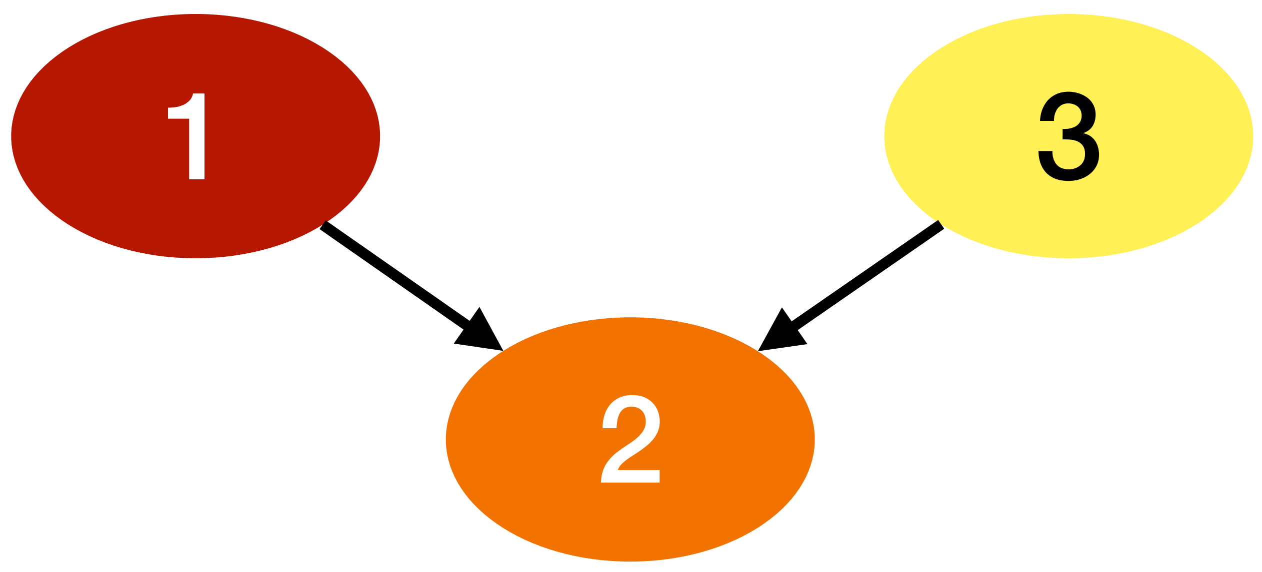
Adding auxiliary intervention variables

- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**



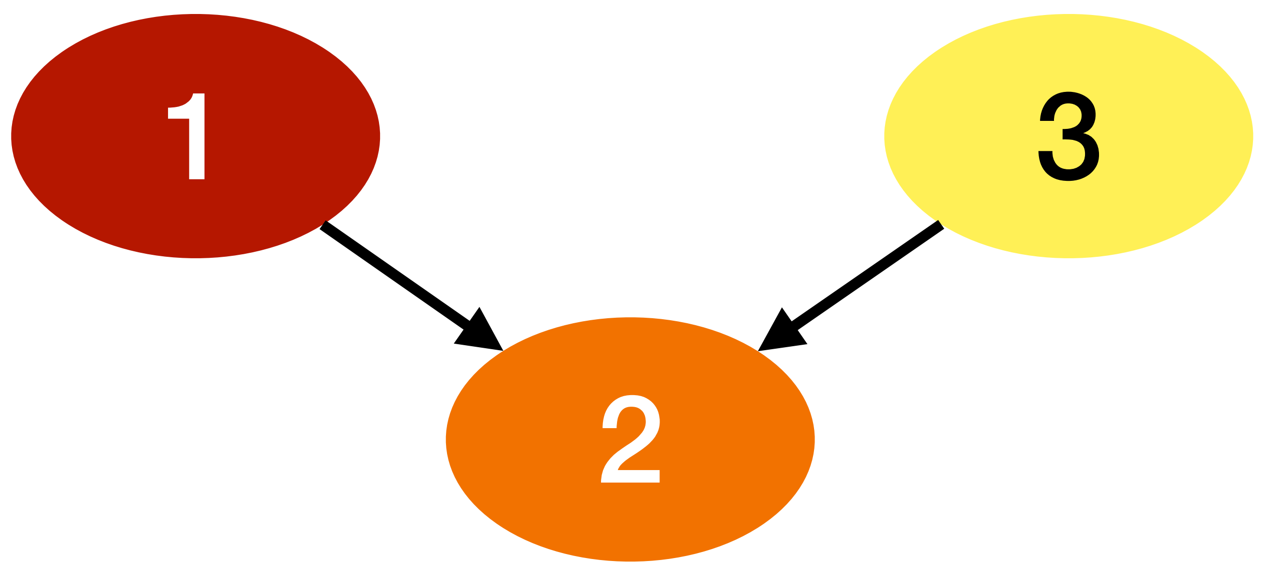
Adding auxiliary intervention variables

- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**
  - For each $i \in \mathbf{V}$ we create a new variable I_i and a new edge $I_i \to i$

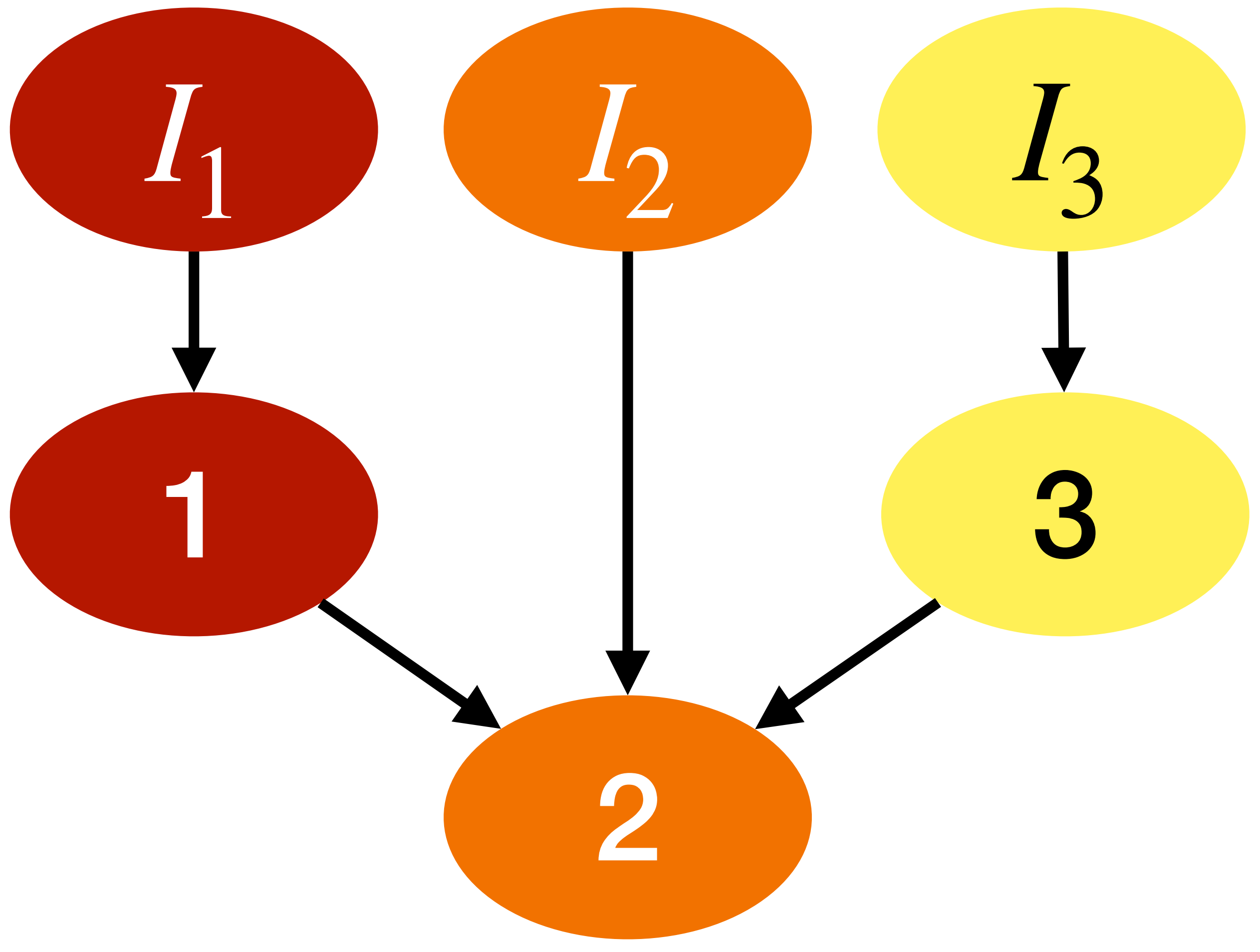


Adding auxiliary intervention variables

- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**
  - For each $i \in \mathbf{V}$ we create a new variable I_i and a new edge $I_i \to i$



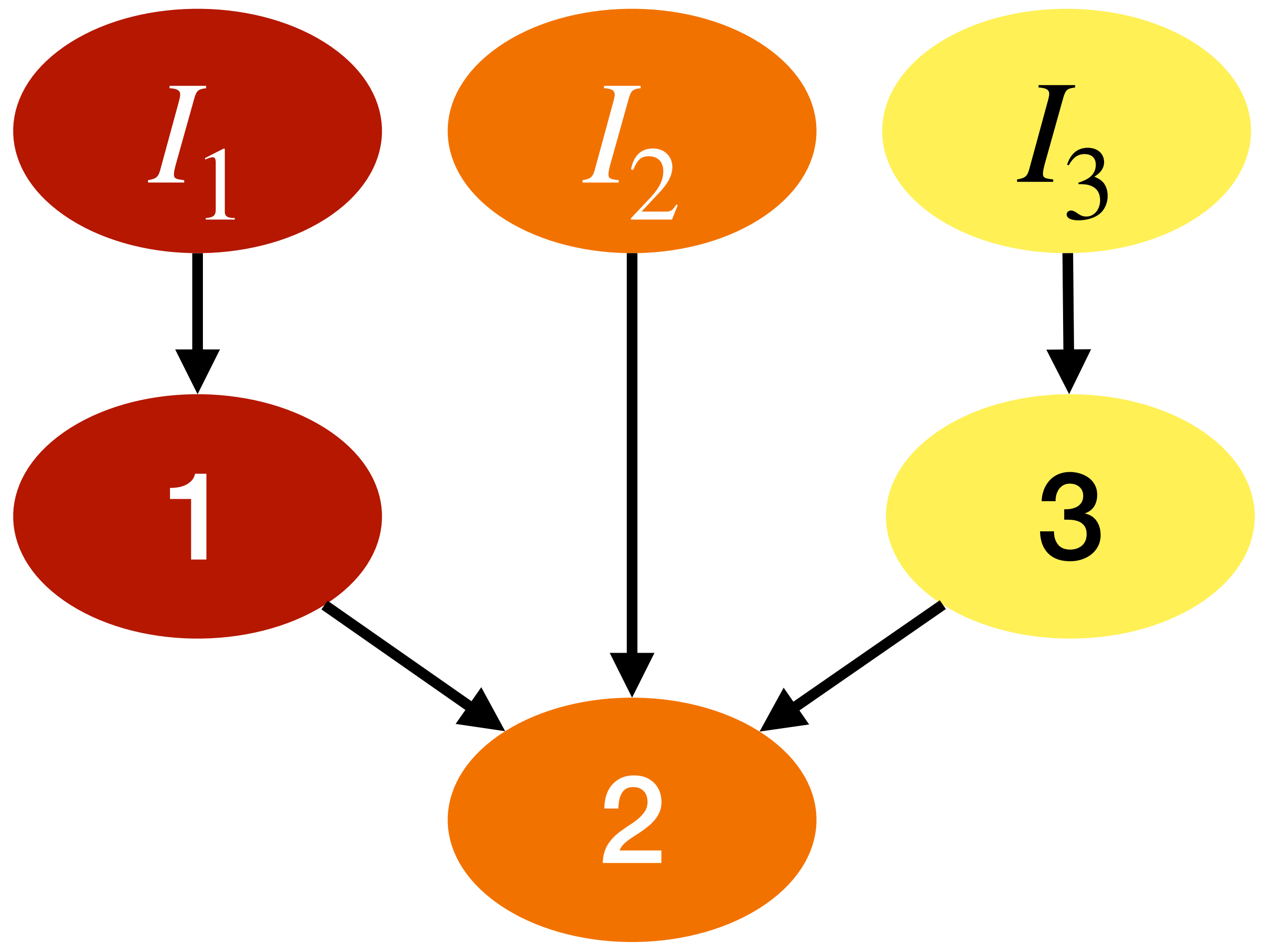
$G = (\mathbf{V}, \mathbf{E})$



$G' = (\mathbf{V} \cup \{I_i\}_{i \in \mathbf{V}}, \mathbf{E} \cup \{I_i \to i\}_{i \in \mathbf{V}})$

Adding auxiliary intervention variables

- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**
  - For each $i \in \mathbf{V}$ we create a new variable I_i and a new edge $I_i \to i$



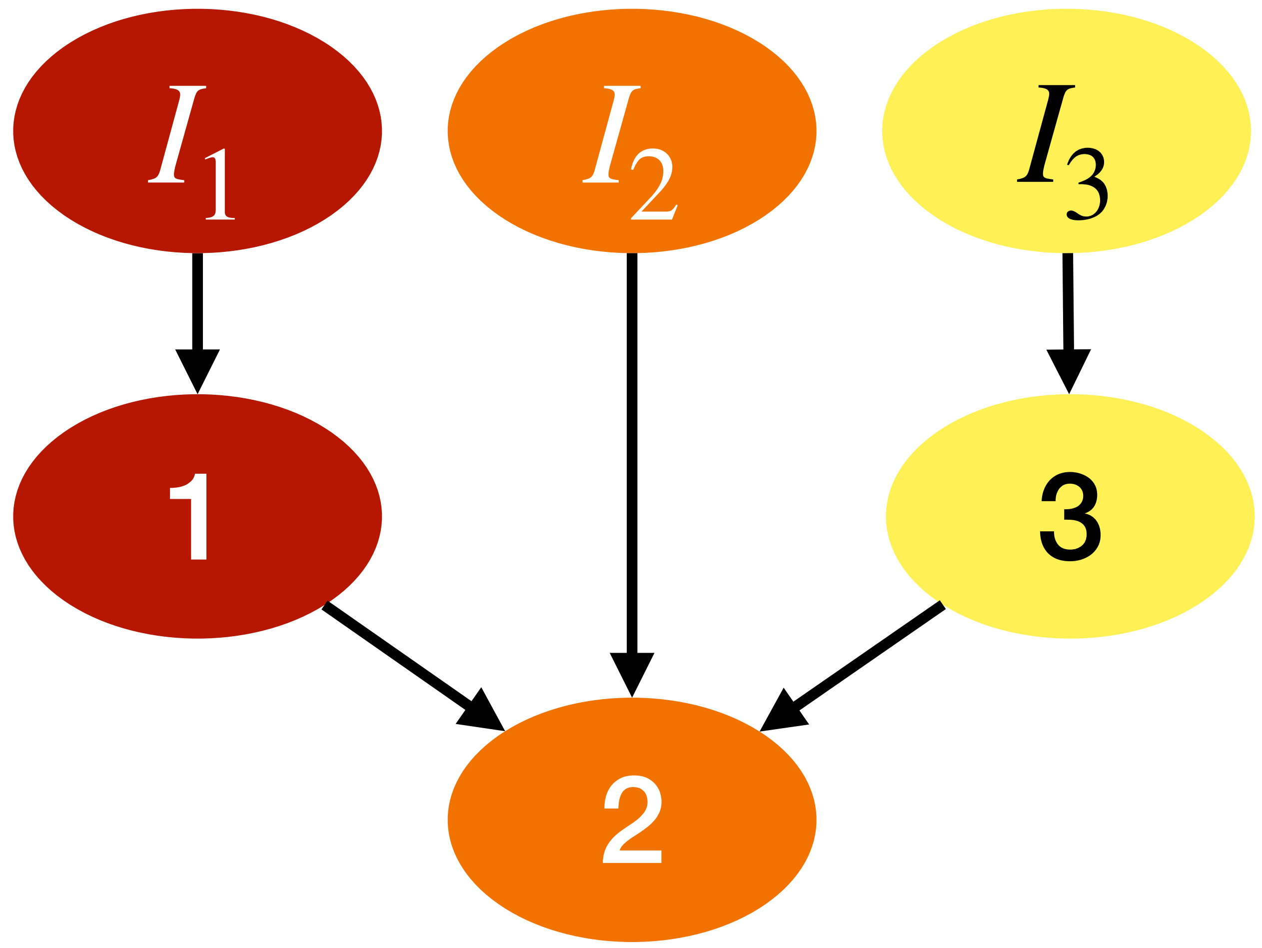
Each intervention variable I_i is:

- 0 in the observational distribution or in the interventional distribution $\mathrm{do}(X_j = x_j)$ for $i \neq j$

In other words when i is not intervened upon

Adding auxiliary intervention variables

- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**
  - For each $i \in \mathbf{V}$ we create a new variable I_i and a new edge $I_i \to i$

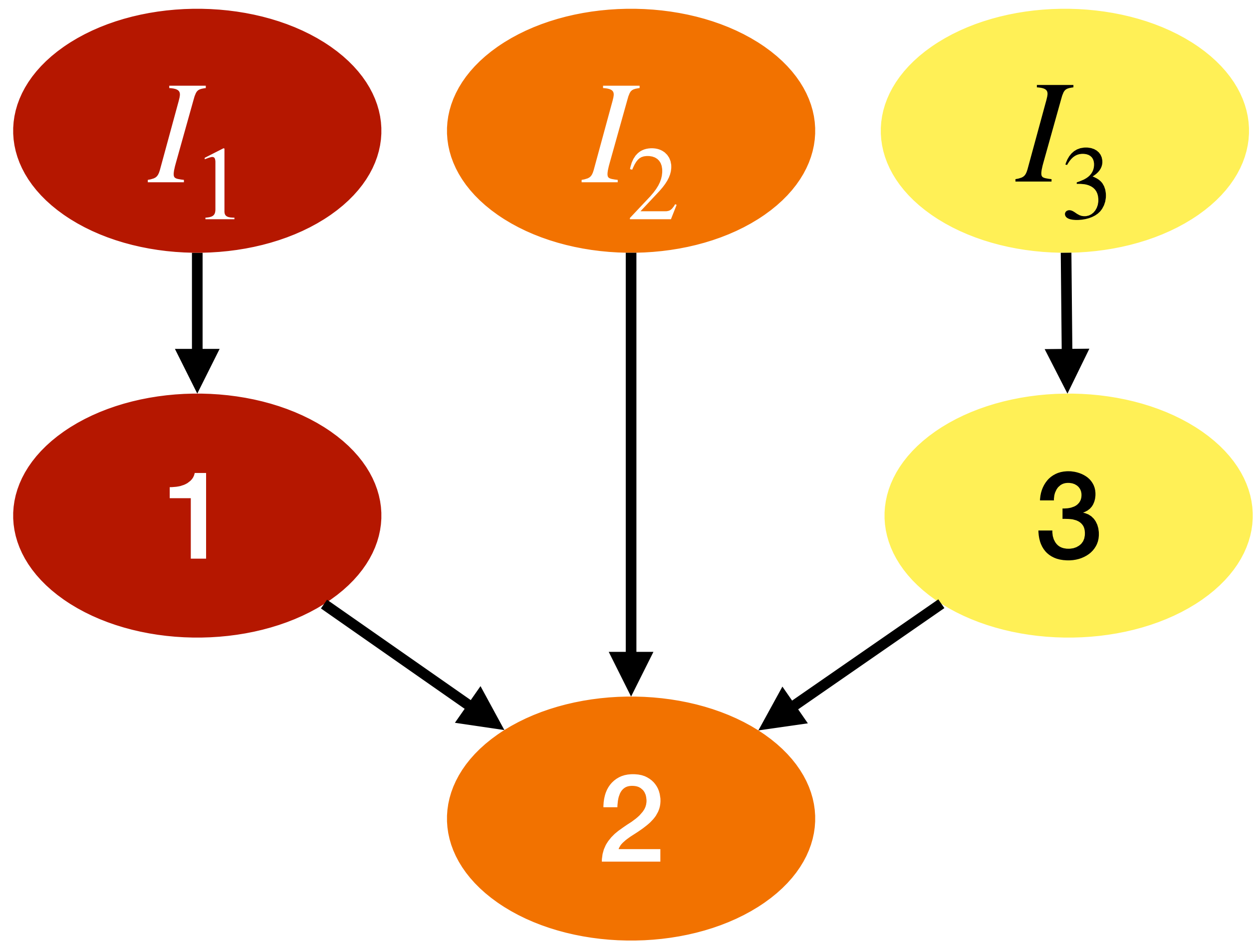


Each intervention variable I_i is:

- 0 in the observational distribution or in the interventional distribution $\mathrm{do}(X_j = x_j)$ for $i \neq j$
- 1 in the interventional distribution $\mathrm{do}(X_i = x_i)$

Adding auxiliary intervention variables

- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**
  - For each $i \in \mathbf{V}$ we create a new variable I_i and a new edge $I_i \to i$

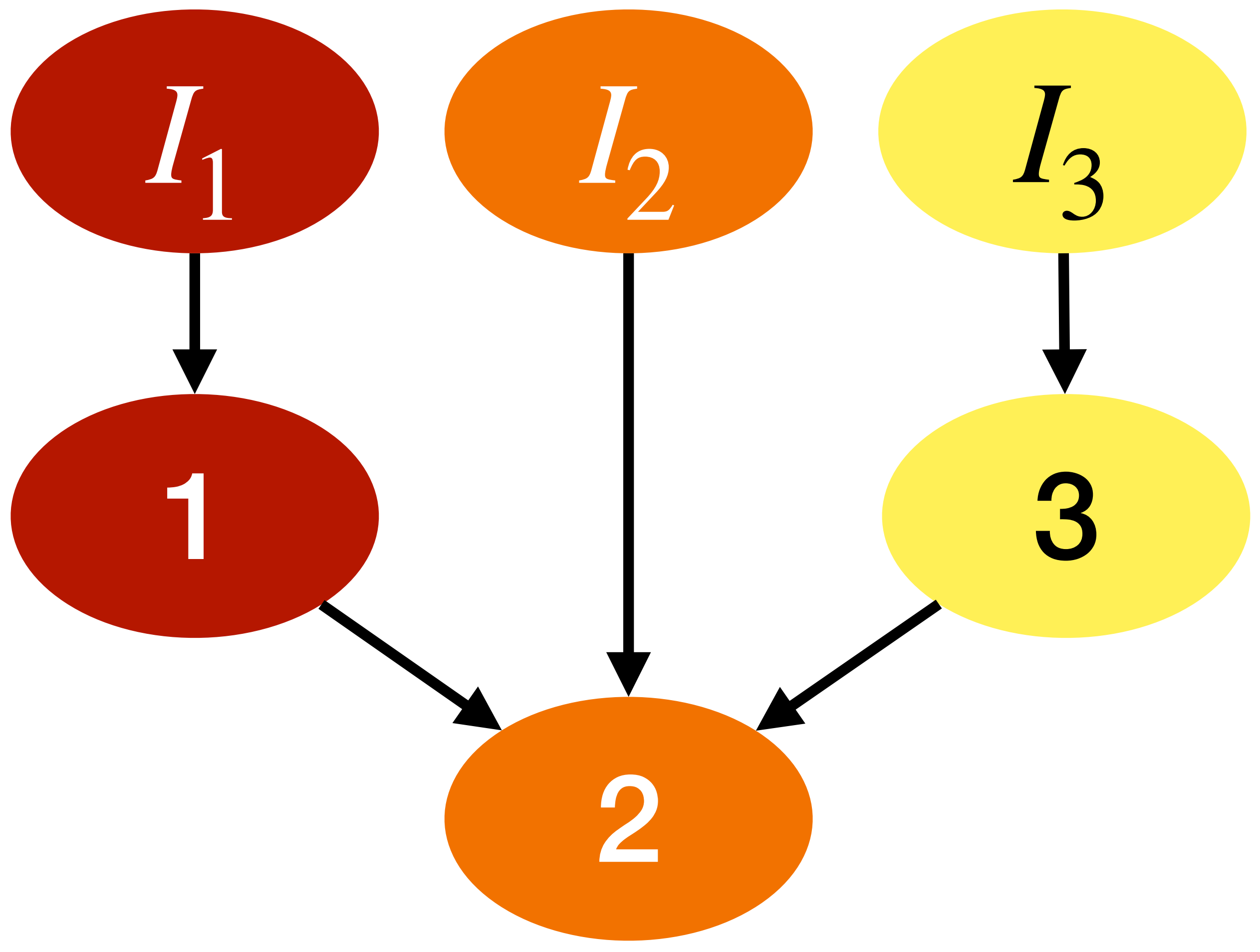


Each intervention variable I_i is:

- 0 in the observational distribution or in the interventional distribution $\mathrm{do}(X_j = x_j)$ for $i \neq j$
- 1 in the interventional distribution $\mathrm{do}(X_i = x_i)$

This graph represents jointly the observational and all possible interventional distributions

Using d-separation on augmented graph G'



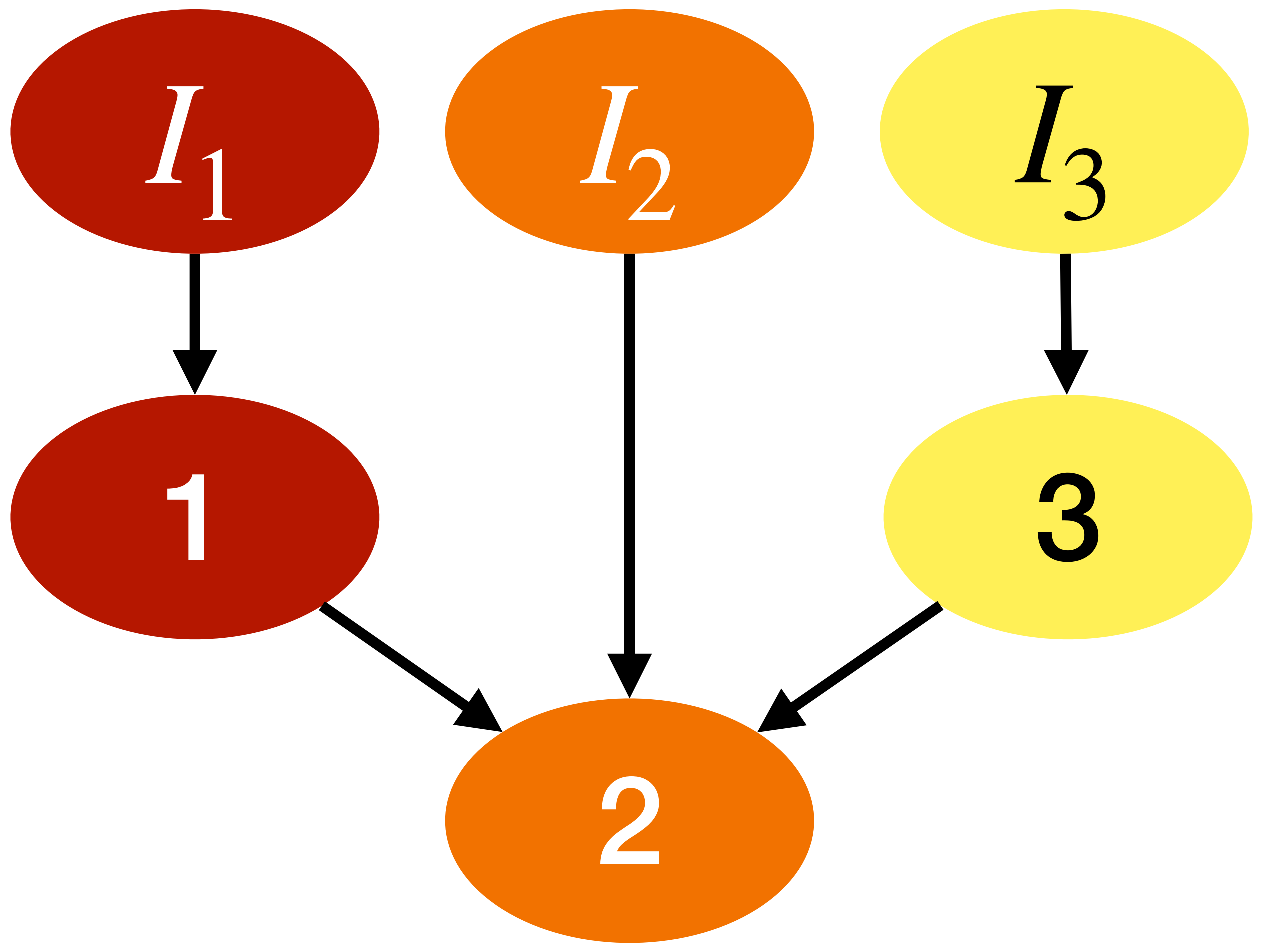
$$1 \perp_{G'} 3$$

$$I_1 \perp_{G'} 3$$

$$I_1 \not\perp_{G'} 2$$

$$I_1 \perp_{G'} 2 \mid 1$$

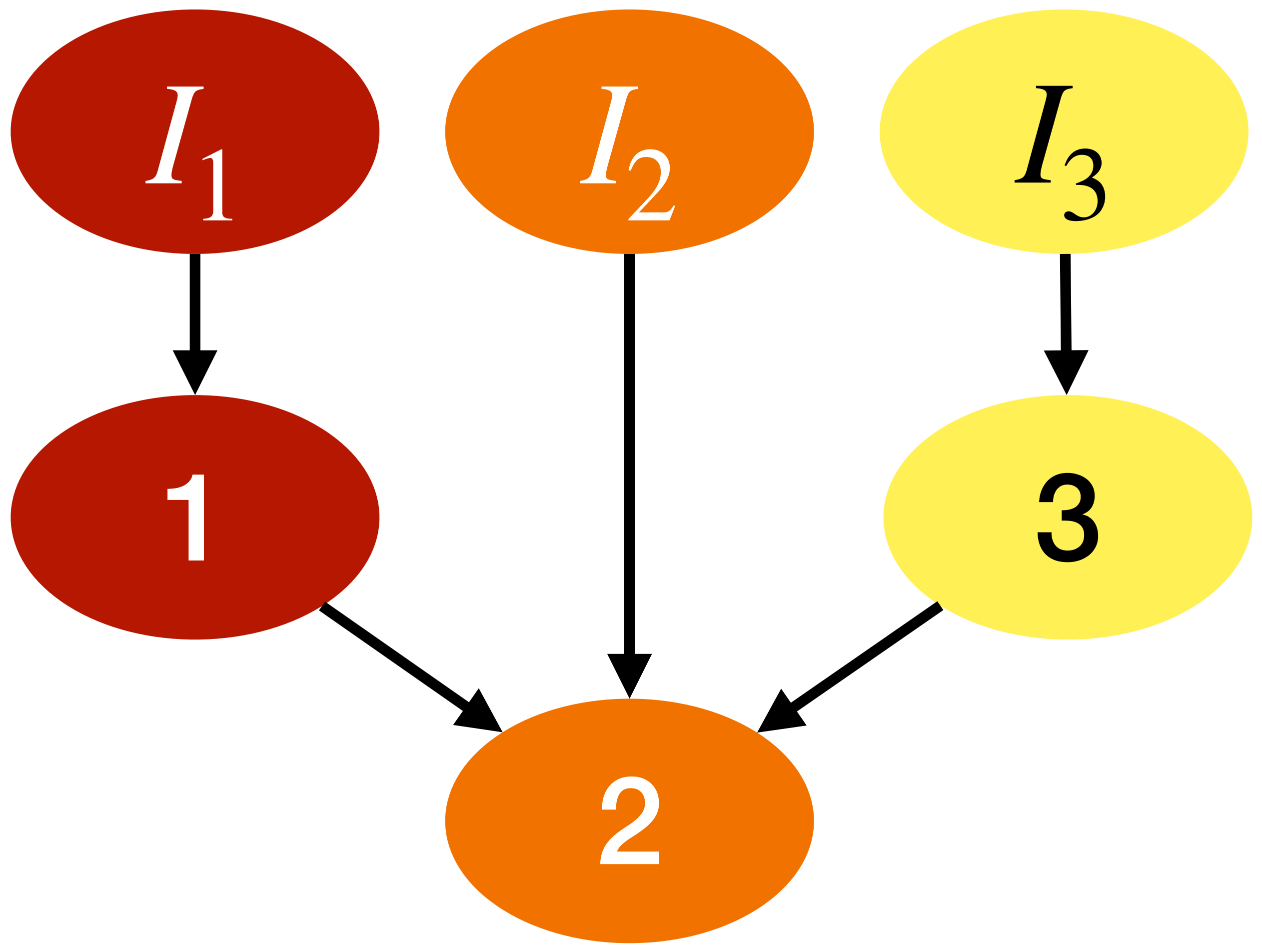
Using d-separation on augmented graph G'



$$I_1 \perp_{G'} 3 \iff_{\text{Markov+faith}} I_1 \perp\!\!\!\perp X_3$$

$$P(X_3 \mid I_1 = 0) = P(X_3 \mid I_1 = 1) = P(X_3)$$

Using d-separation on augmented graph G'

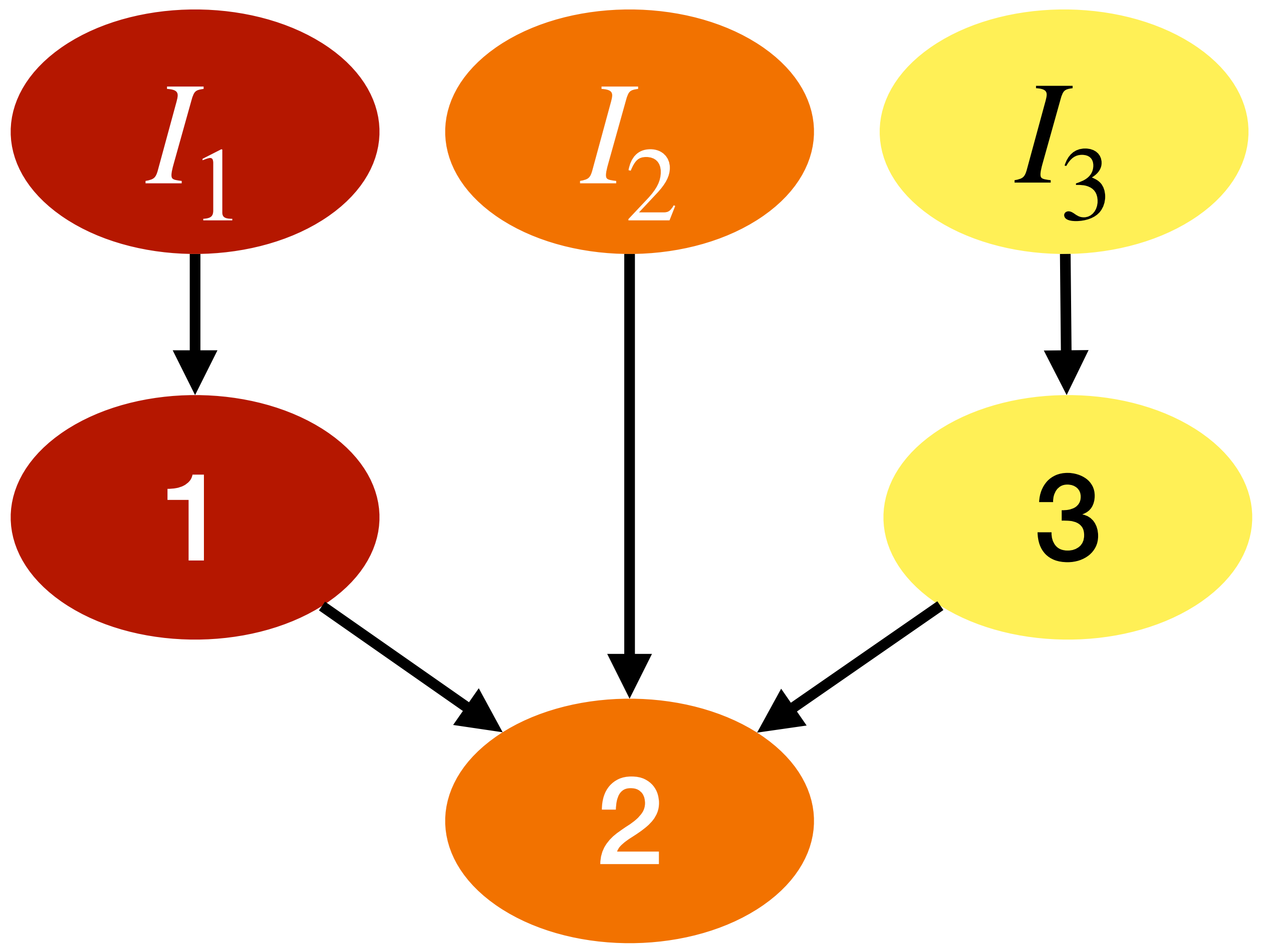


$$I_1 \perp_{G'} 3 \iff_{\text{Markov+faith}} I_1 \perp\!\!\!\perp X_3$$

$$P(X_3 \mid I_1 = 0) = P(X_3 \mid I_1 = 1) = P(X_3)$$

$$P_{\text{obs}}(X_3) = P_{\text{do}(X_1)}(X_3)$$

Using d-separation on augmented graph G'



$$I_1 \perp_{G'} 3 \iff_{\text{Markov+faith}} I_1 \perp\!\!\!\perp X_3$$

$$P(X_3 \mid I_1 = 0) = P(X_3 \mid I_1 = 1) = P(X_3)$$

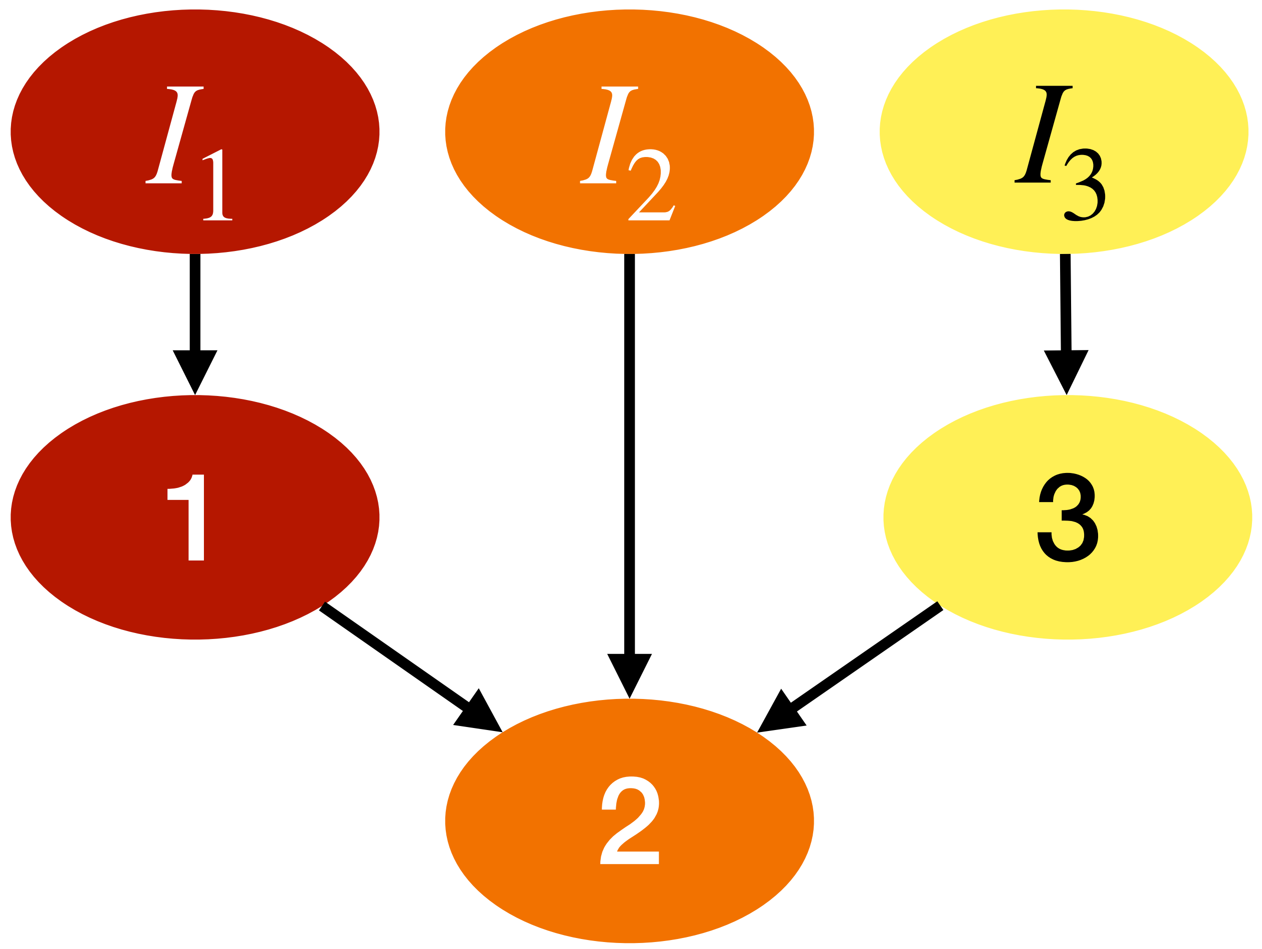
$$P_{\text{obs}}(X_3) = P_{\text{do}(X_1)}(X_3)$$

Classic do-notation in Pearl:

$$P(X_3) = P(X_3 \mid \text{do}(X_1))$$

Remember the do is **not** really **conditioning**

Using d-separation on augmented graph G'



$1 \perp_{G'} 3$ $\quad P(X_3 | X_1) = P(X_3)$

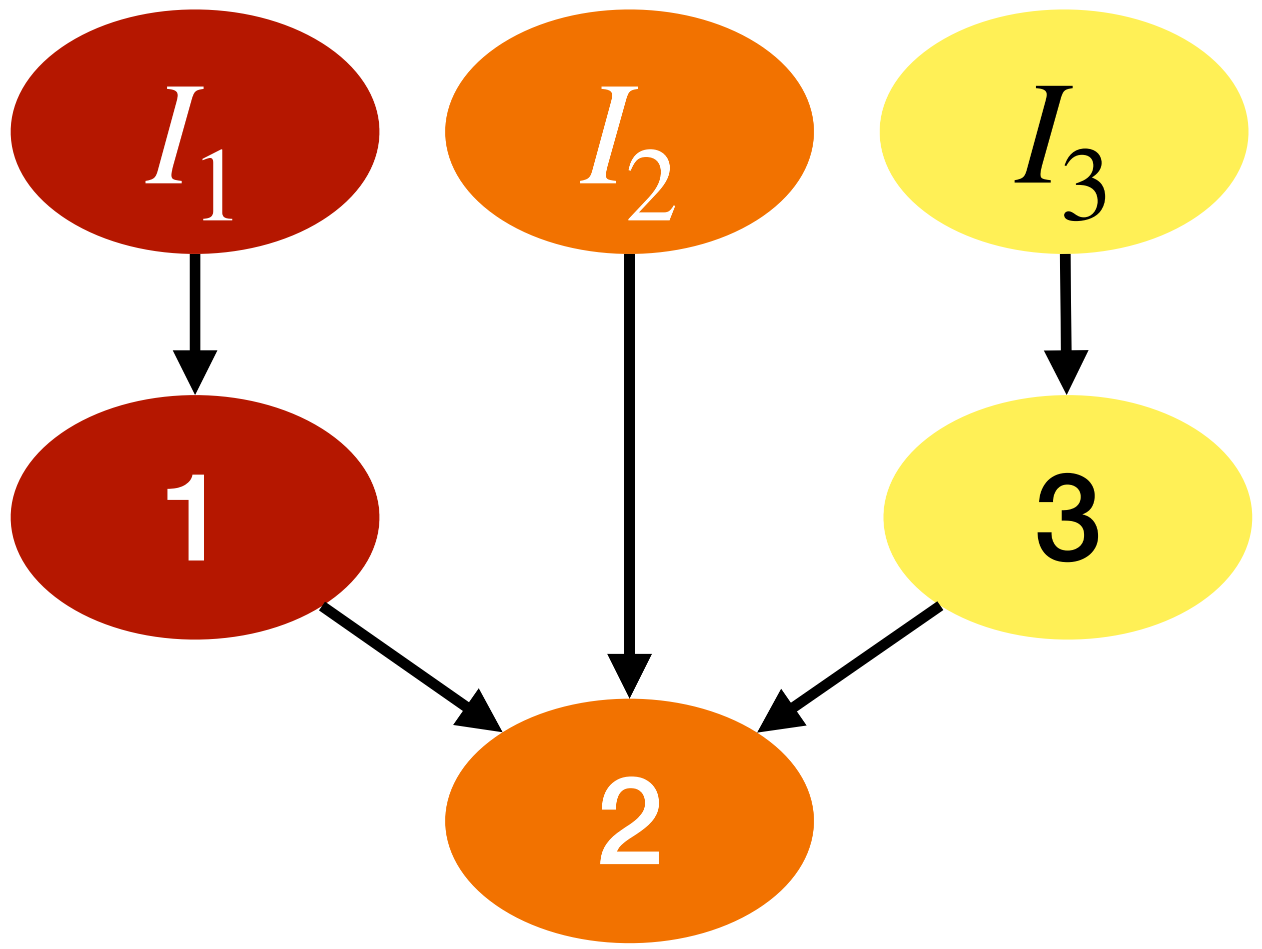
$I_1 \perp_{G'} 3$ $\quad P(X_3 | \mathrm{do}(X_1)) = P(X_3)$

$I_1 \not\perp_{G'} 2$ $\quad P(X_2 | \mathrm{do}(X_1)) \neq P(X_2)$

$I_1 \perp_{G'} 2 | 1$ $\quad P(X_2 | \mathrm{do}(X_1)) = P(X_2 | X_1)$

$$P(X_2 | I_1 = 1, X_1) = P(X_2 | I_1, X_1) = P(X_2 | X_1)$$

Using d-separation on augmented graph G'



$1 \perp_{G'} 3 \quad P(X_3 | X_1) = P(X_3)$

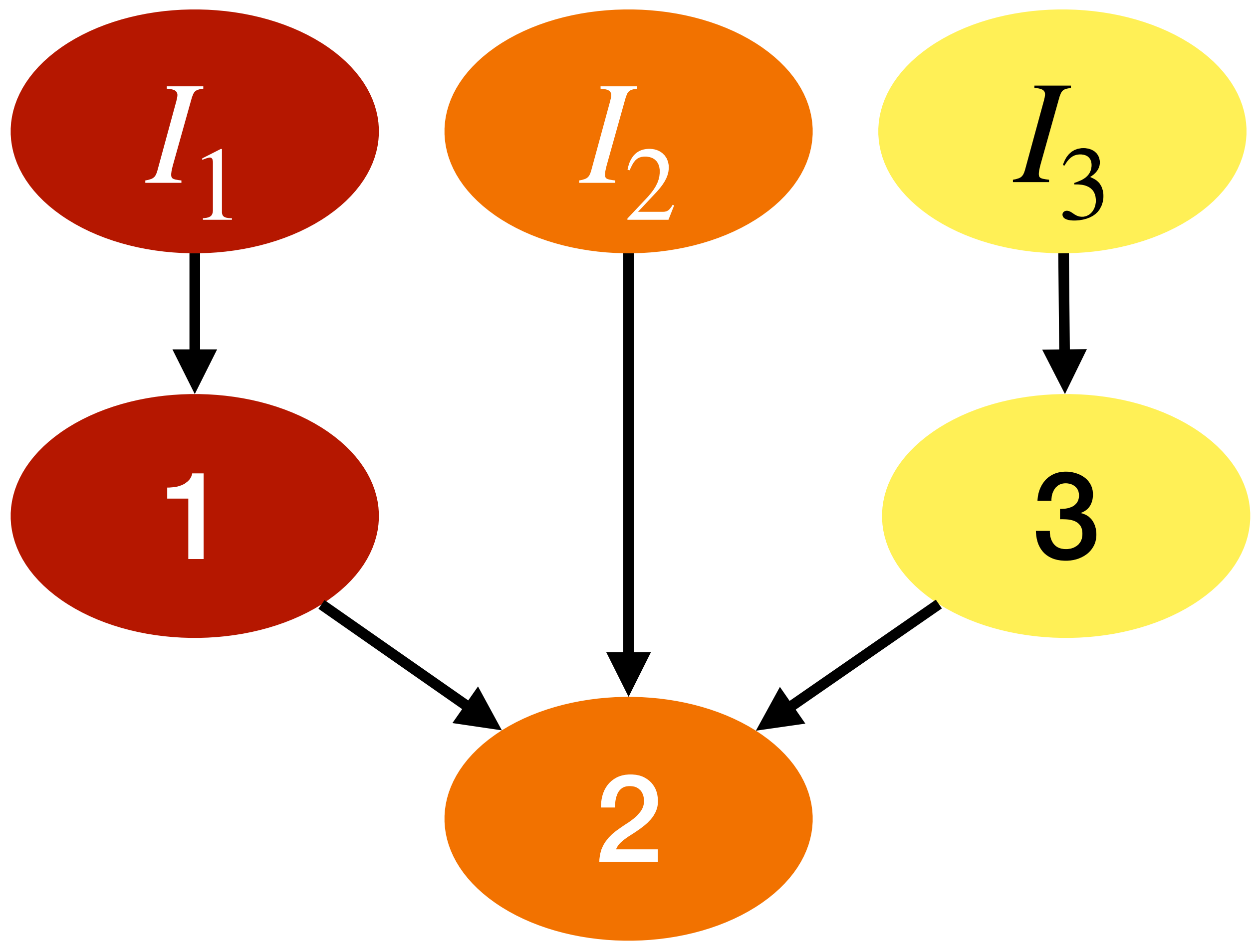
$I_1 \perp_{G'} 3 \quad P(X_3 | \mathrm{do}(X_1)) = P(X_3)$

$I_1 \not\perp_{G'} 2 \quad P(X_2 | \mathrm{do}(X_1)) \neq P(X_2)$

$I_1 \perp_{G'} 2 | 1 \quad P(X_2 | \mathrm{do}(X_1)) = P(X_2 | X_1)$

What about $P(X_3 | \mathrm{do}(X_1), X_2)$? Is it the same as $P(X_3 | X_2)$?

Using d-separation on augmented graph G'



$1 \perp_{G'} 3 \quad P(X_3 \mid X_1) = P(X_3)$

$I_1 \perp_{G'} 3 \quad P(X_3 \mid \mathrm{do}(X_1)) = P(X_3)$

$I_1 \not\perp_{G'} 2 \quad P(X_2 \mid \mathrm{do}(X_1)) \neq P(X_2)$

$I_1 \perp_{G'} 2 \mid 1 \quad P(X_2 \mid \mathrm{do}(X_1)) = P(X_2 \mid X_1)$

$\boxed{\text{What about } P(X_3 \mid \mathrm{do}(X_1), X_2)\text{? Is it the same as } P(X_3 \mid X_2)\text{?}}$

$3 \not\perp_{G'} I_1 \mid 2 \quad P(X_3 \mid \mathrm{do}(X_1), X_2) \neq P(X_3 \mid X_2)$

Do-calculus simplified version (only one intervention)

- For disjoint sets $\mathbf{A}, \mathbf{B}, \mathbf{C} \subseteq \mathbf{V}$:

1. Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_{\mathbf{A}} \mid X_{\mathbf{B}}, X_{\mathbf{C}}) = P(X_{\mathbf{A}} \mid X_{\mathbf{C}})$$

Does this remind you of anything?

Do-calculus simplified version (only one intervention)

- For disjoint sets $\mathbf{A}, \mathbf{B}, \mathbf{C} \subseteq \mathbf{V}$:

1. Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$$

2. Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid do(X_\mathbf{B} = x_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{B} = x_\mathbf{B}, X_\mathbf{C})$$

$$\forall x_\mathbf{B} : \quad P(X_\mathbf{A} \mid X_\mathbf{B} = x_\mathbf{B}, X_\mathbf{C}, I_\mathbf{B} = 1) = P(X_\mathbf{A} \mid X_\mathbf{B} = x_\mathbf{B}, X_\mathbf{C}, I_\mathbf{B} = 0)$$

Do-calculus simplified version (only one intervention)

- For disjoint sets $\mathbf{A}, \mathbf{B}, \mathbf{C} \subseteq \mathbf{V}$:

1. Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$$

2. Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid do(X_\mathbf{B} = x_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{B} = x_\mathbf{B}, X_\mathbf{C})$$

3. Rule 3: insertion/deletion of actions

$$\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid do(X_\mathbf{B}), X_\mathbf{C},) = P(X_\mathbf{A} \mid X_\mathbf{C})$$

$$P(X_\mathbf{A} \mid X_\mathbf{C}, I_\mathbf{B} = 1) = P(X_\mathbf{A} \mid X_\mathbf{C}, I_\mathbf{B} = 0)$$

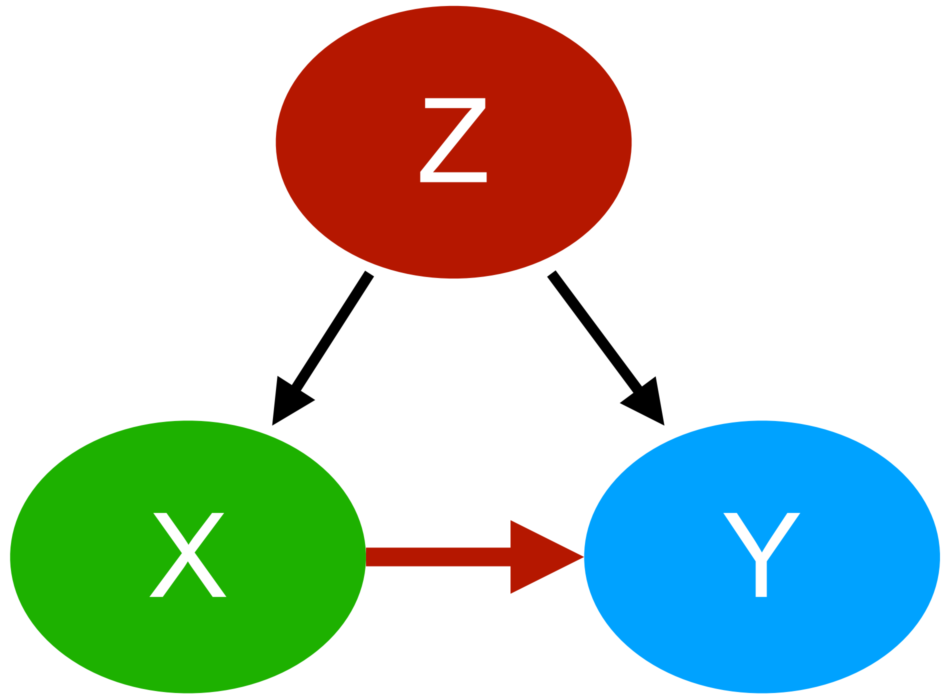
Do-calculus simplified version - example

- Rule 1: insertion/deletion of observations

  $$\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_{\mathbf{A}} \mid X_{\mathbf{B}}, X_{\mathbf{C}}) = P(X_{\mathbf{A}} \mid X_{\mathbf{C}})$$
- Rule 2: action/observation exchange

  $$\mathbf{A} \perp_{G'} I_{\mathbf{B}} \mid \mathbf{B}, \mathbf{C} \implies P(X_{\mathbf{A}} \mid \mathrm{do}(X_{\mathbf{B}}), X_{\mathbf{C}}) = P(X_{\mathbf{A}} \mid X_{\mathbf{B}}, X_{\mathbf{C}})$$
- Rule 3: insertion/deletion of actions

  $$\mathbf{A} \perp_{G'} I_{\mathbf{B}} \mid \mathbf{C} \implies P(X_{\mathbf{A}} \mid \mathrm{do}(X_{\mathbf{B}}), X_{\mathbf{C}}) = P(X_{\mathbf{A}} \mid X_{\mathbf{C}})$$



$p(y \mid do(x)) = ?$

Do-calculus simplified version - example

- Rule 1: insertion/deletion of observations

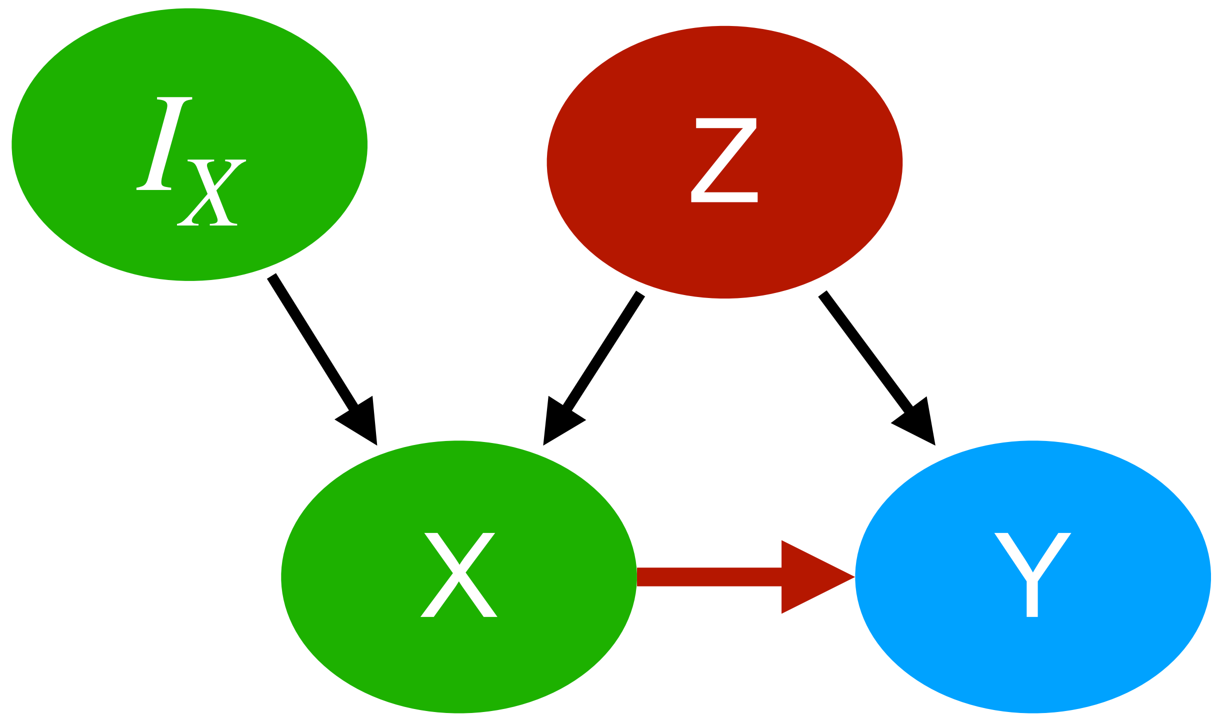
$$\mathbf{A} \perp_{G'} \mathbf{B} \,|\, \mathbf{C} \implies P(X_{\mathbf{A}} \,|\, X_{\mathbf{B}}, X_{\mathbf{C}}) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{C}})$$

- Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'} I_{\mathbf{B}} \,|\, \mathbf{B}, \mathbf{C} \implies P(X_{\mathbf{A}} \,|\, \mathrm{do}(X_{\mathbf{B}}), X_{\mathbf{C}}) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{B}}, X_{\mathbf{C}})$$

- Rule 3: insertion/deletion of actions

$$\mathbf{A} \perp_{G'} I_{\mathbf{B}} \,|\, \mathbf{C} \implies P(X_{\mathbf{A}} \,|\, \mathrm{do}(X_{\mathbf{B}}), X_{\mathbf{C}}) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{C}})$$



$p(y \,|\, do(x)) = ?$

$\boxed{\text{Step 1: add relevant intervention variable } I_X}$

$\boxed{\text{Step 2: check d-separations with } I_X \text{ and } Y}$

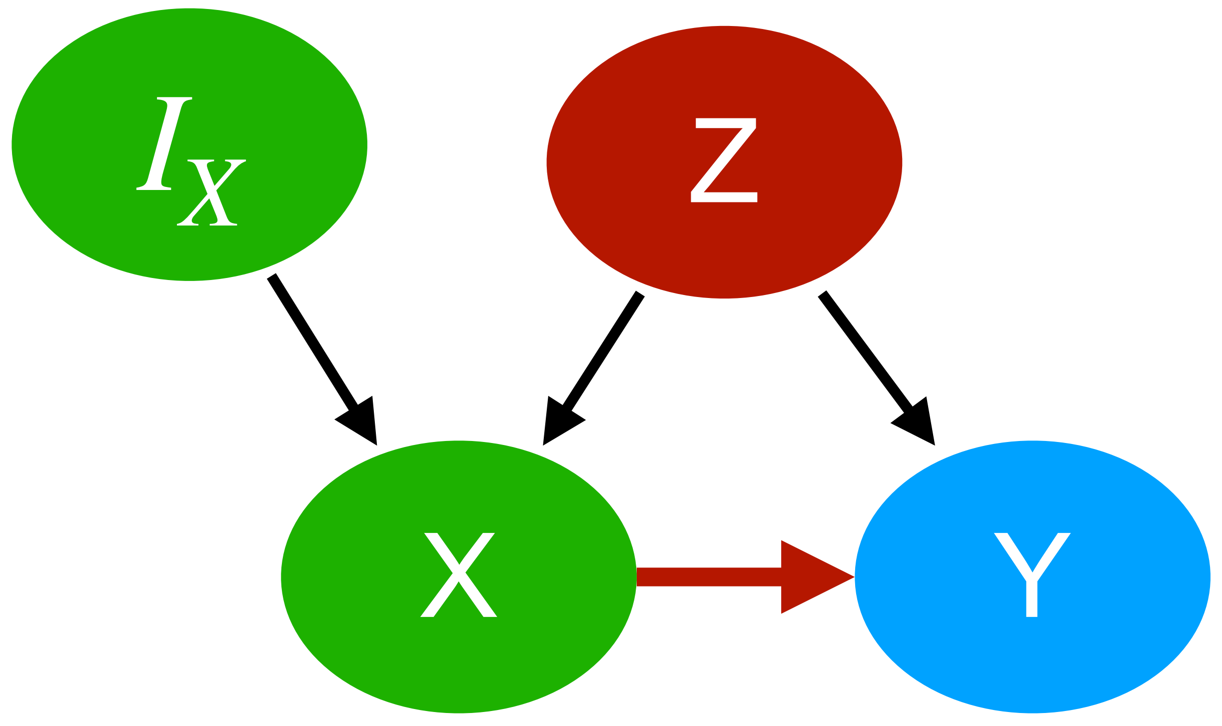
Do-calculus simplified version - example

- Rule 1: insertion/deletion of observations
  $\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$
- Rule 2: action/observation exchange
  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C})$
- Rule 3: insertion/deletion of actions
  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$

Step 2: check d-separations with I_X and Y

$Y \not\perp_d I_X$

$Y \not\perp_d I_X \mid X$



$P(Y \mid do(X)) = ?$

What do these two d-connections mean for $P(Y \mid \mathrm{do}(X))$? (check rules 2 and 3)

Do-calculus simplified version - example

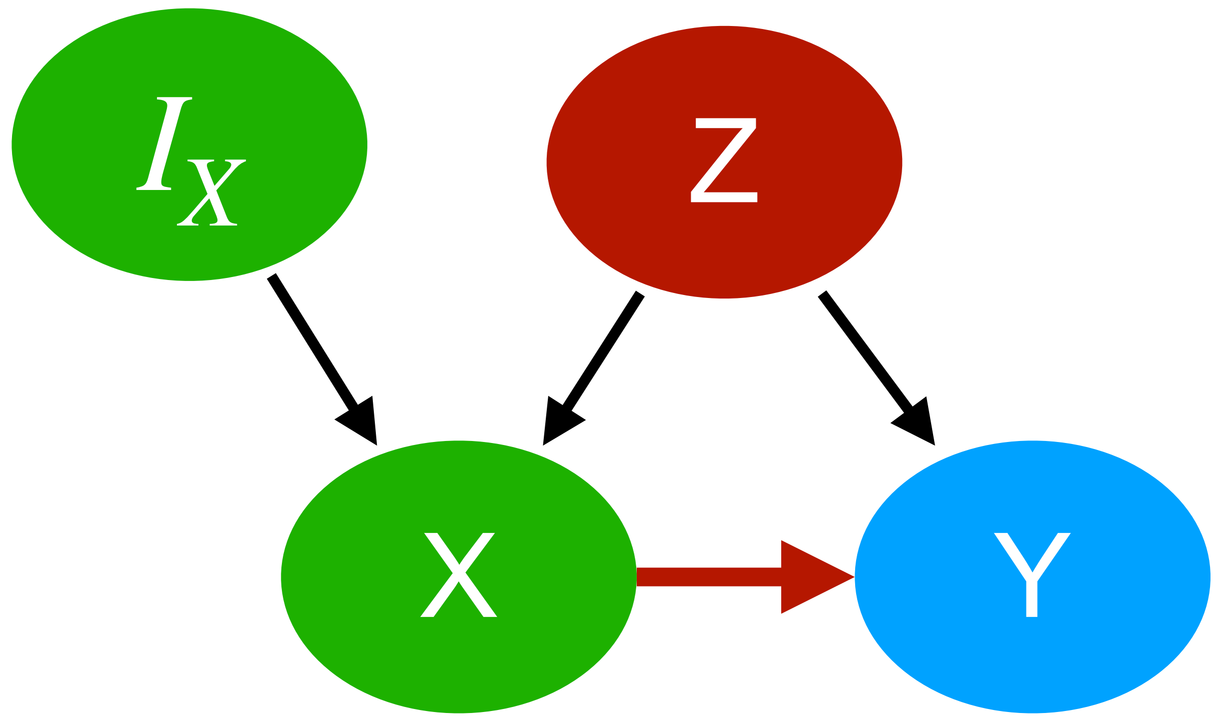
- Rule 1: insertion/deletion of observations
  $\mathbf{A} \perp_{G'} \mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{C})$
- Rule 2: action/observation exchange
  $\mathbf{A} \perp_{G'} I_\mathbf{B} \,|\, \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \,|\, \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C})$
- Rule 3: insertion/deletion of actions
  $\mathbf{A} \perp_{G'} I_\mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{C})$

Step 2: check d-separations with I_X and Y

$Y \not\perp_d I_X$

$Y \not\perp_d I_X \mid X$

$Y \perp_d I_X \mid X, Z$



$p(Y \mid do(X)) = ?$

What do this new d-separation mean for $P(Y \,|\, \mathrm{do}(X))$? (check rules 2 and 3)

Do-calculus simplified version - example

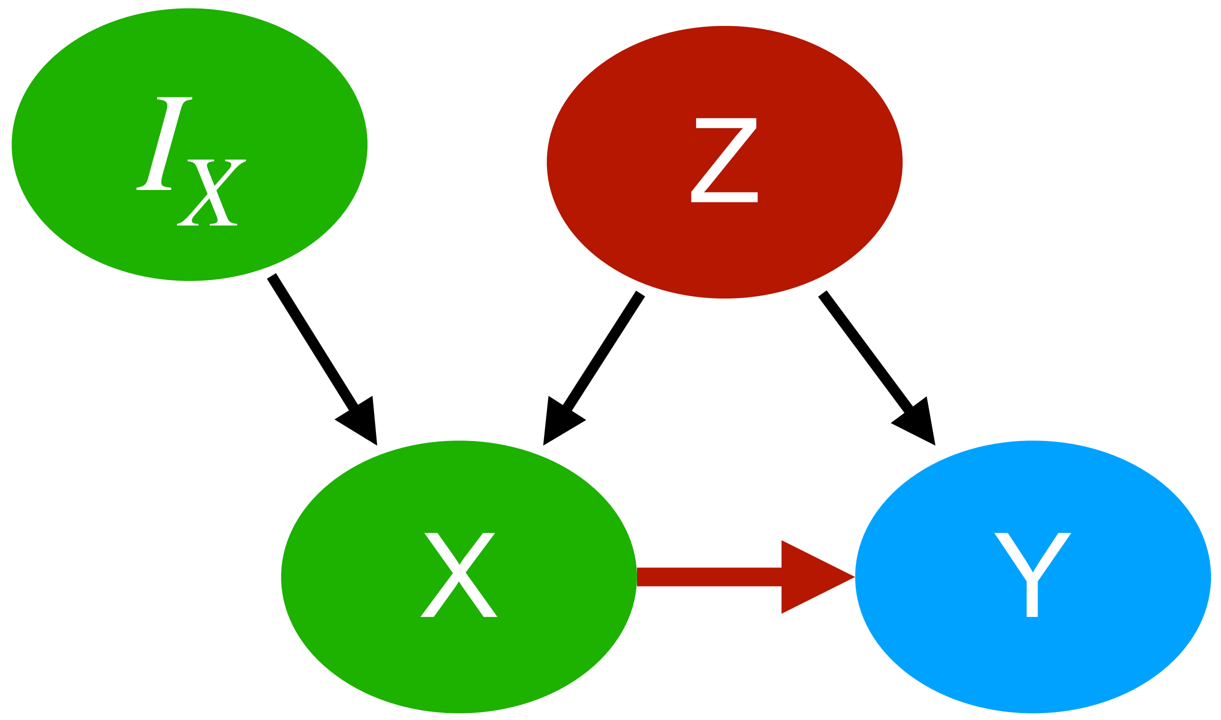
- Rule 1: insertion/deletion of observations
  $\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$
- Rule 2: action/observation exchange
  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C})$
- Rule 3: insertion/deletion of actions
  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$

Step 2: check d-separations with I_X and Y

$Y \not\perp_d I_X$

$Y \not\perp_d I_X \mid X$

$Y \perp_d I_X \mid X, Z$



$p(Y \mid do(X)) = ?$

$p(Y \mid do(X), Z) = p(Y \mid X, Z)$

(RULE 2)

Not yet there, we have $p(Y \mid \mathrm{do}(X), Z)$, but not $p(Y \mid \mathrm{do}(X))$

Do-calculus simplified version - example

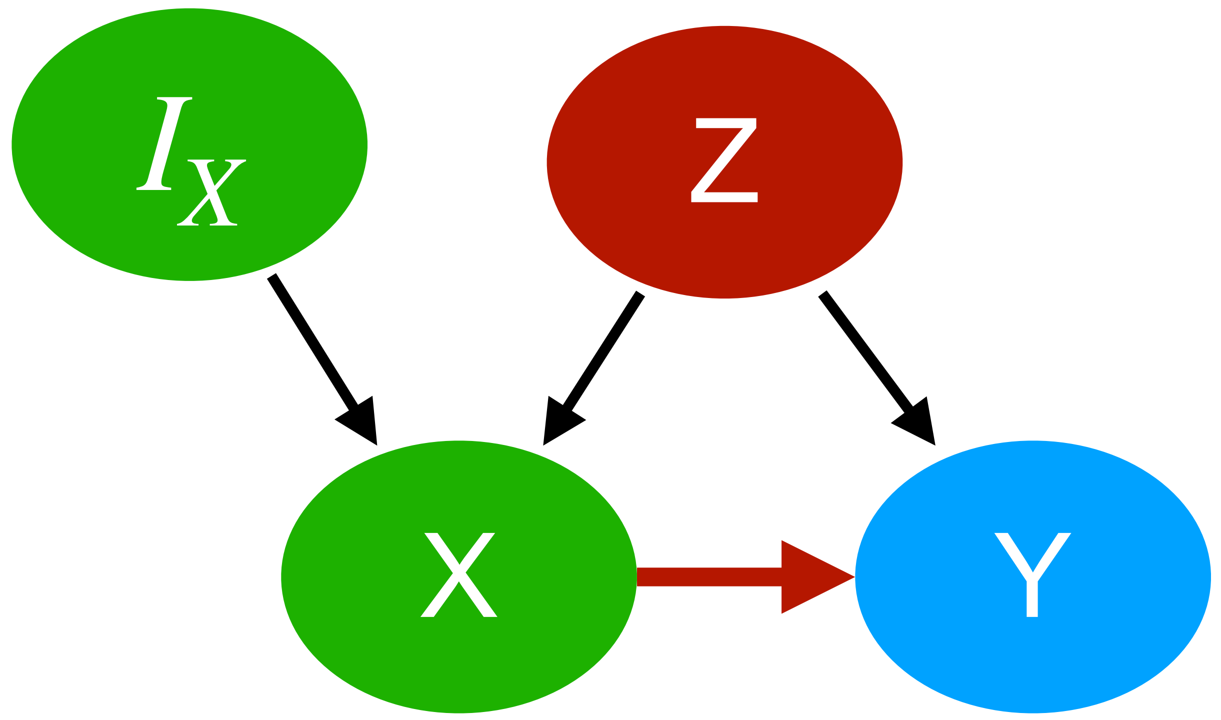
- Rule 1: insertion/deletion of observations

  $\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$
- Rule 2: action/observation exchange

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C})$
- Rule 3: insertion/deletion of actions

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$

$p(y \mid do(x), z) = p(y \mid x, z)$
(RULE 2)



$p(y \mid do(x)) = ?$

$z \perp I_x \Rightarrow p(z \mid do(x)) = p(z)$
(RULE 3)

What about other d-separations with I_X?

Do-calculus simplified version - example

- Rule 1: insertion/deletion of observations

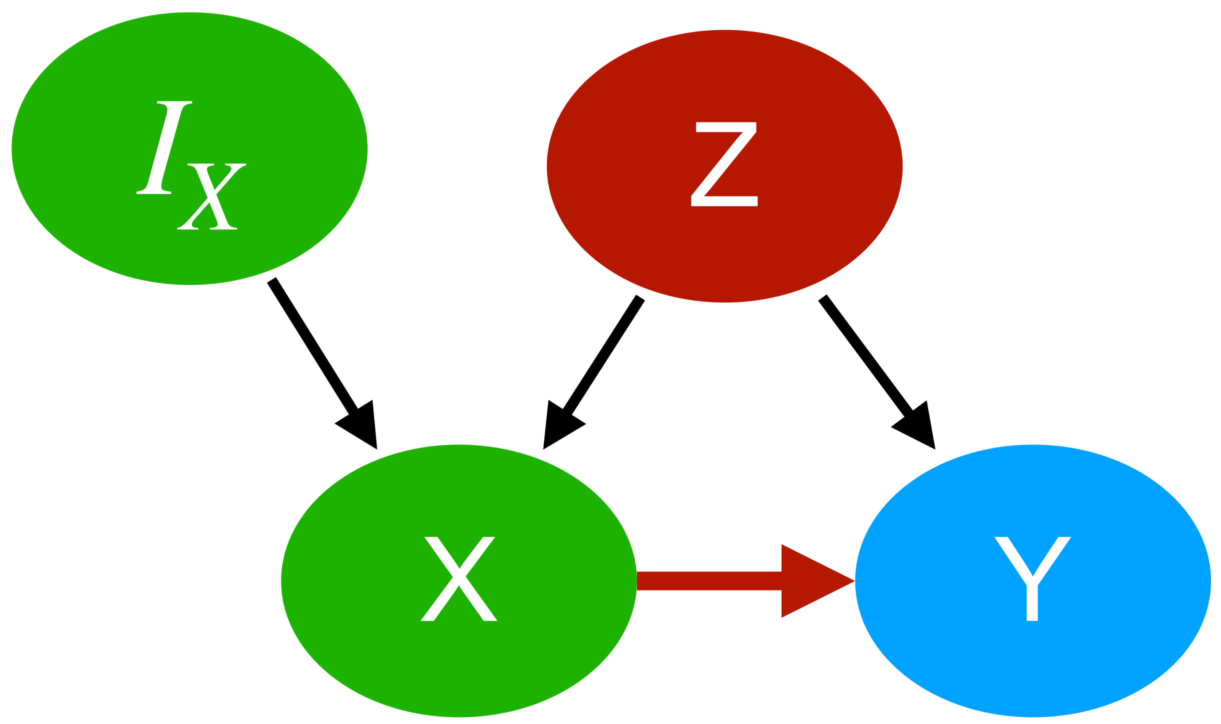
  $\mathbf{A} \perp_{G'} \mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{C})$
- Rule 2: action/observation exchange

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \,|\, \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \,|\, \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C})$
- Rule 3: insertion/deletion of actions

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{C})$

$p(y \,|\, do(x), z) = p(y \,|\, x, z)$
(RULE 2)

$p(z \,|\, do(x)) = p(z)$ (RULE 3)



$$p(X, Y, Z \,|\, \mathrm{do}(X = \tilde{x})) = p(Y \,|\, X = x, Z = z, \mathrm{do}(X = \tilde{x})) \cdot p(Z = z \,|\, \mathrm{do}(X = \tilde{x})) \cdot 1(X = \tilde{x})$$

Truncated factorisation formula for $do(X)$

$$p(X, Y, Z \,|\, \mathrm{do}(X = \tilde{x})) = p(Y \,|\, Z = z, \mathrm{do}(X = \tilde{x})) \cdot p(Z = z \,|\, \mathrm{do}(X = \tilde{x})) \cdot 1(X = \tilde{x})$$

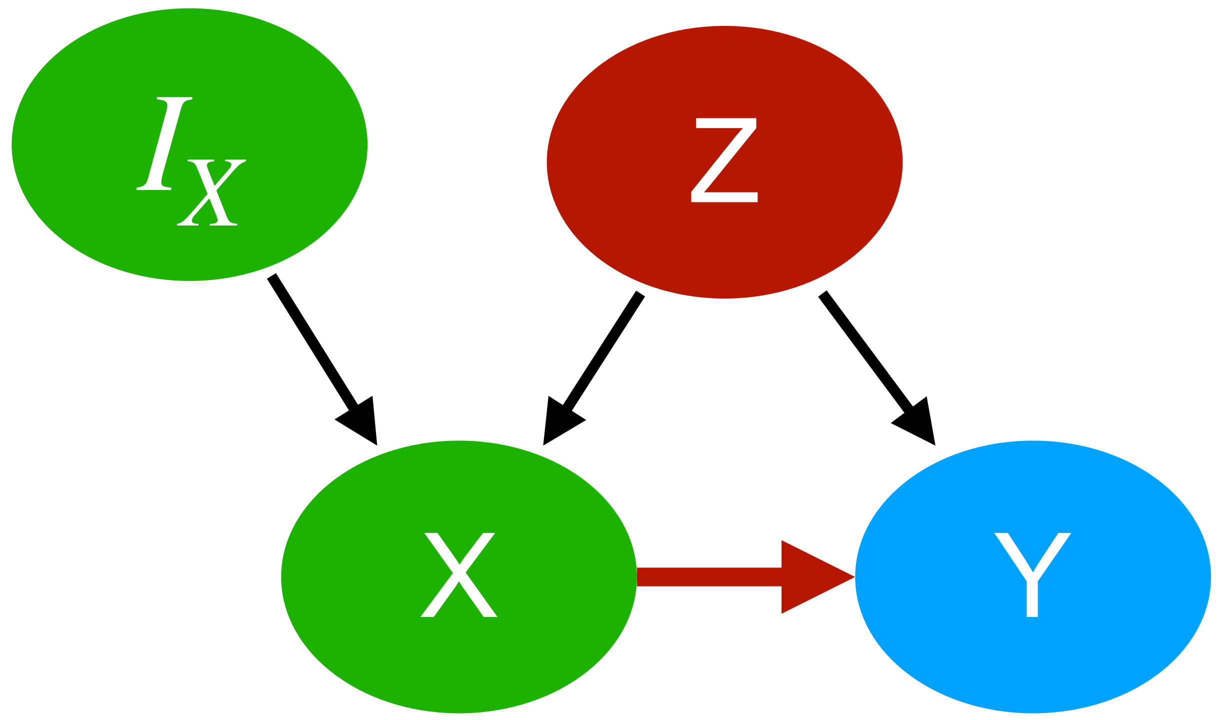
Do-calculus simplified version - example

- Rule 1: insertion/deletion of observations

  $\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$
- Rule 2: action/observation exchange

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C})$
- Rule 3: insertion/deletion of actions

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$



$$p(X, Y, Z \mid \mathrm{do}(X = \tilde{x})) = p(Y \mid Z = z, \mathrm{do}(X = \tilde{x})) \cdot p(Z = z \mid \mathrm{do}(X = \tilde{x})) \cdot 1(X = \tilde{x})$$

$p(Y \mid do(X), Z) = p(Y \mid X, Z)$ $p(Z \mid do(X)) = p(Z)$

$$p(X, Y, Z \mid \mathrm{do}(X = \tilde{x})) = p(Y \mid X = x, Z = z) \cdot p(Z = z) \cdot 1(X = \tilde{x})$$

Do-calculus simplified version - example

- Rule 1: insertion/deletion of observations

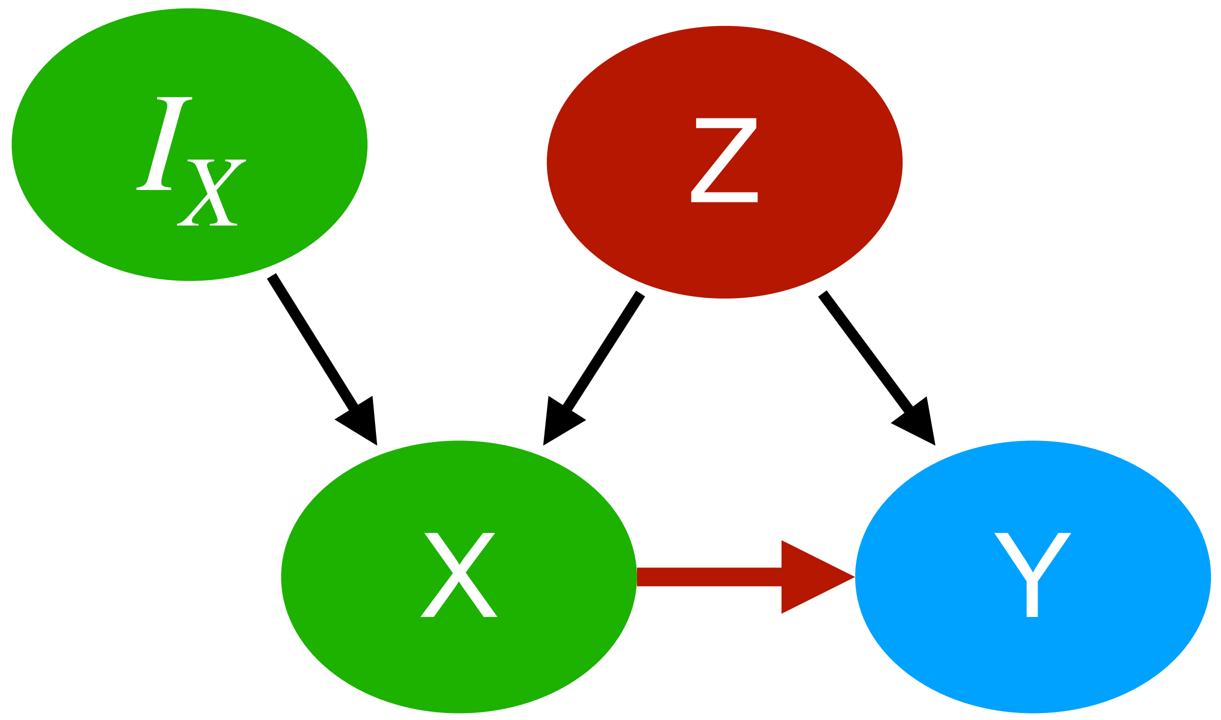
  $\mathbf{A} \perp_{G'} \mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{C})$
- Rule 2: action/observation exchange

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \,|\, \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \,|\, \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C})$
- Rule 3: insertion/deletion of actions

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{C})$

$p(y \,|\, do(x), z) = p(y \,|\, x, z)$
(RULE 2)

$p(z \,|\, do(x)) = p(z)$ (RULE 3)



$$p(X, Y, Z \,|\, \mathrm{do}(X = \tilde{x})) = p(Y \,|\, X = x, Z = z) \cdot p(Z = z) \cdot 1(X = \tilde{x})$$

$$p(Y \,|\, \mathrm{do}(X = \tilde{x})) = \int_x \int_z p(Y \,|\, X = x, Z = z) \cdot p(Z = z) \cdot 1(X = \tilde{x}) dz dx$$

Do-calculus simplified version - example

- Rule 1: insertion/deletion of observations

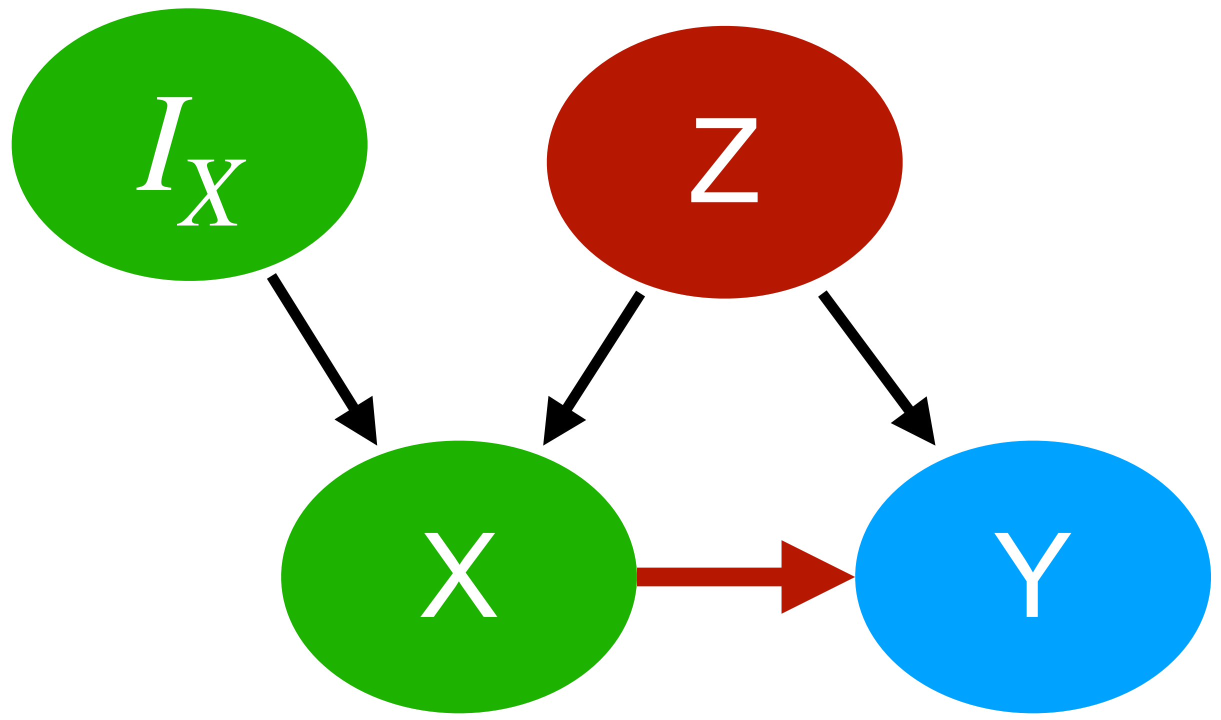
  $\mathbf{A} \perp_{G'} \mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{C})$
- Rule 2: action/observation exchange

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \,|\, \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \,|\, \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C})$
- Rule 3: insertion/deletion of actions

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{C})$

$p(y \mid do(x), z) = p(y \mid x, z)$
(RULE 2)

$p(z \mid do(x)) = p(z)$ (RULE 3)



$$p(X, Y, Z \,|\, \mathrm{do}(X = \tilde{x})) = p(Y \,|\, X = x, Z = z) \cdot p(Z = z) \cdot 1(X = \tilde{x})$$

$$p(Y \,|\, \mathrm{do}(X = \tilde{x})) = \int_x \int_z p(Y \,|\, X = x, Z = z) \cdot p(Z = z) \cdot 1(X = \tilde{x}) dz dx$$

$$p(Y \,|\, \mathrm{do}(X = \tilde{x})) = \int_z p(Y \,|\, X = \tilde{x}, Z = z) \cdot p(Z = z) dz$$

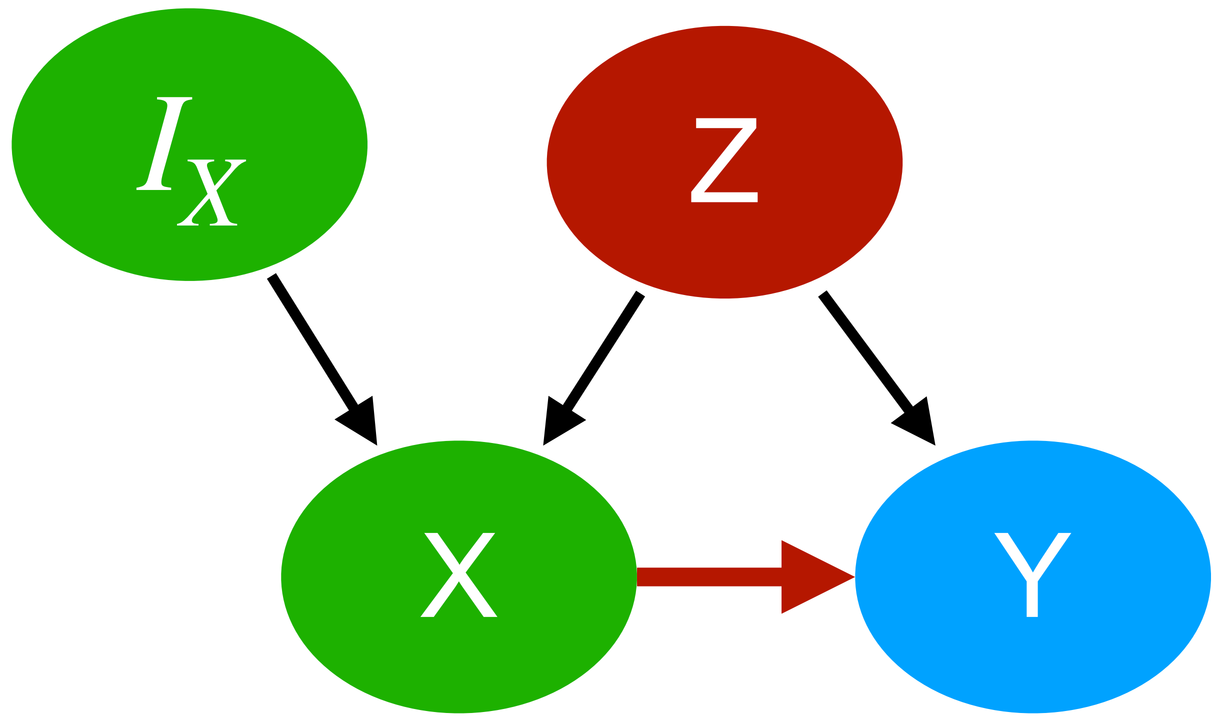
Do-calculus simplified version - example

- Rule 1: insertion/deletion of observations

  $\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$
- Rule 2: action/observation exchange

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C})$
- Rule 3: insertion/deletion of actions

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$



$$p(Y \mid \mathrm{do}(X = \tilde{x})) = \int_z p(Y \mid X = \tilde{x}, Z = z) \cdot p(Z = z) dz$$

Does this remind you of anything?

Do-calculus simplified version - example

- Rule 1: insertion/deletion of observations

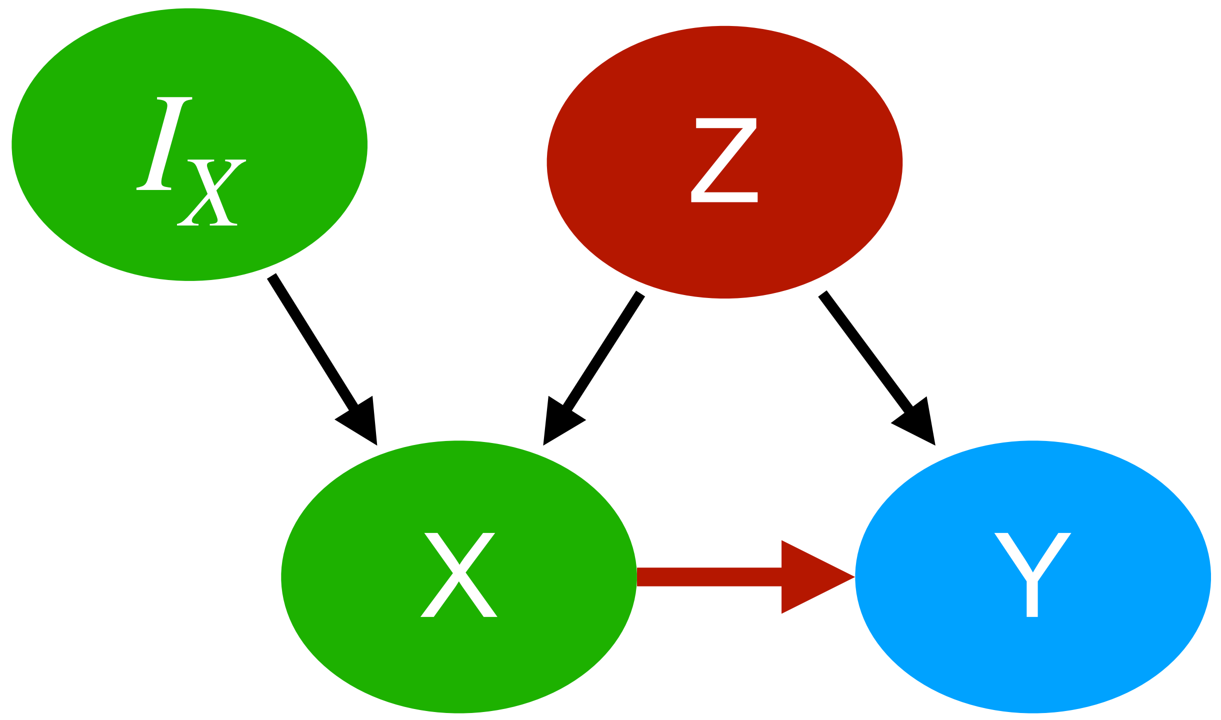
  $\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$
- Rule 2: action/observation exchange

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C})$
- Rule 3: insertion/deletion of actions

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$

$Z \perp_d I_X$ ($Z \notin Desc(X)$)

$Y \perp_d I_X \mid X, Z$ (Z blocks backdoors)

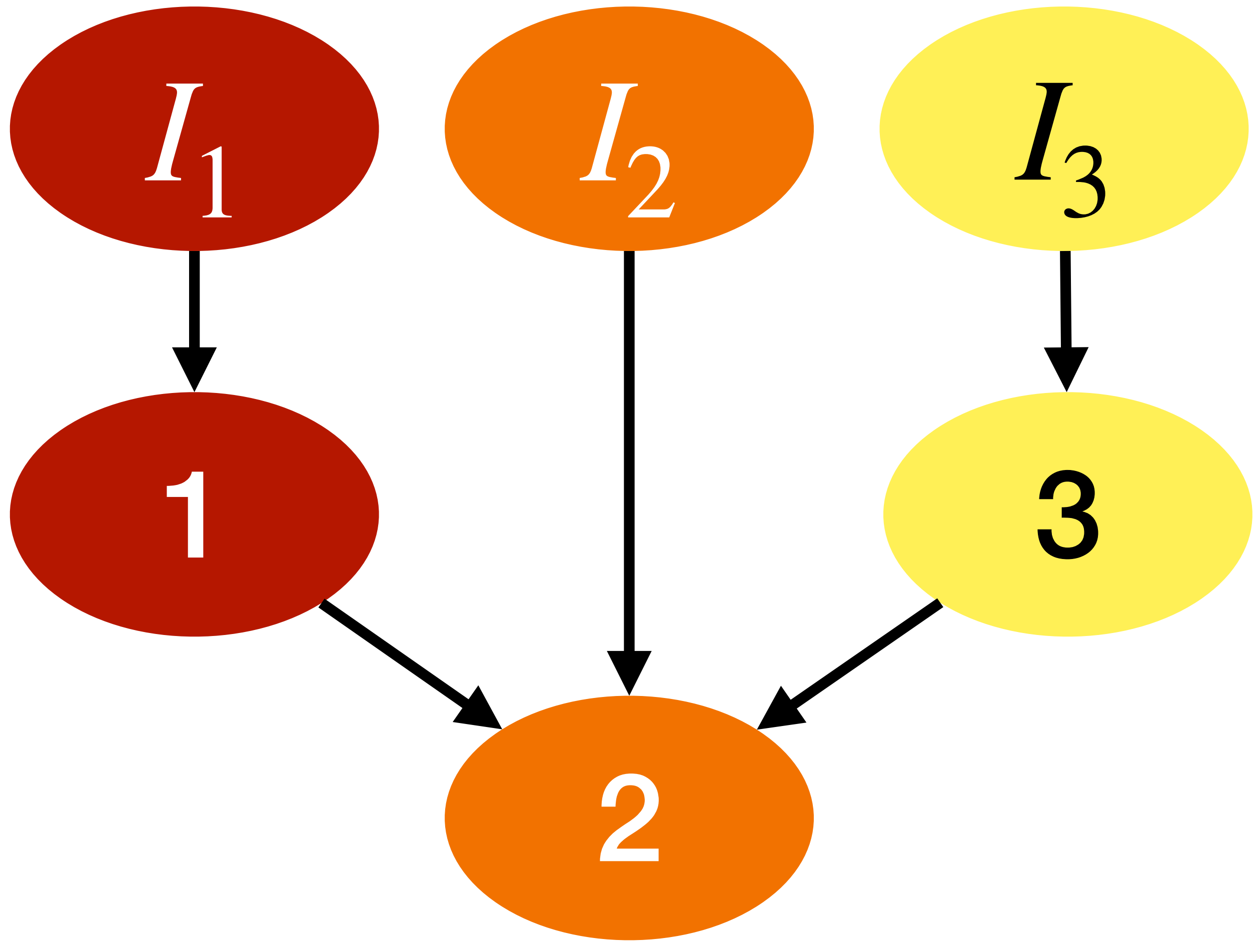


$$p(Y \mid \mathrm{do}(X = \tilde{x})) = \int_z p(Y \mid X = \tilde{x}, Z = z) \cdot p(Z = z) dz$$

BACKDOOR ADJ.

Multiple consequent interventions

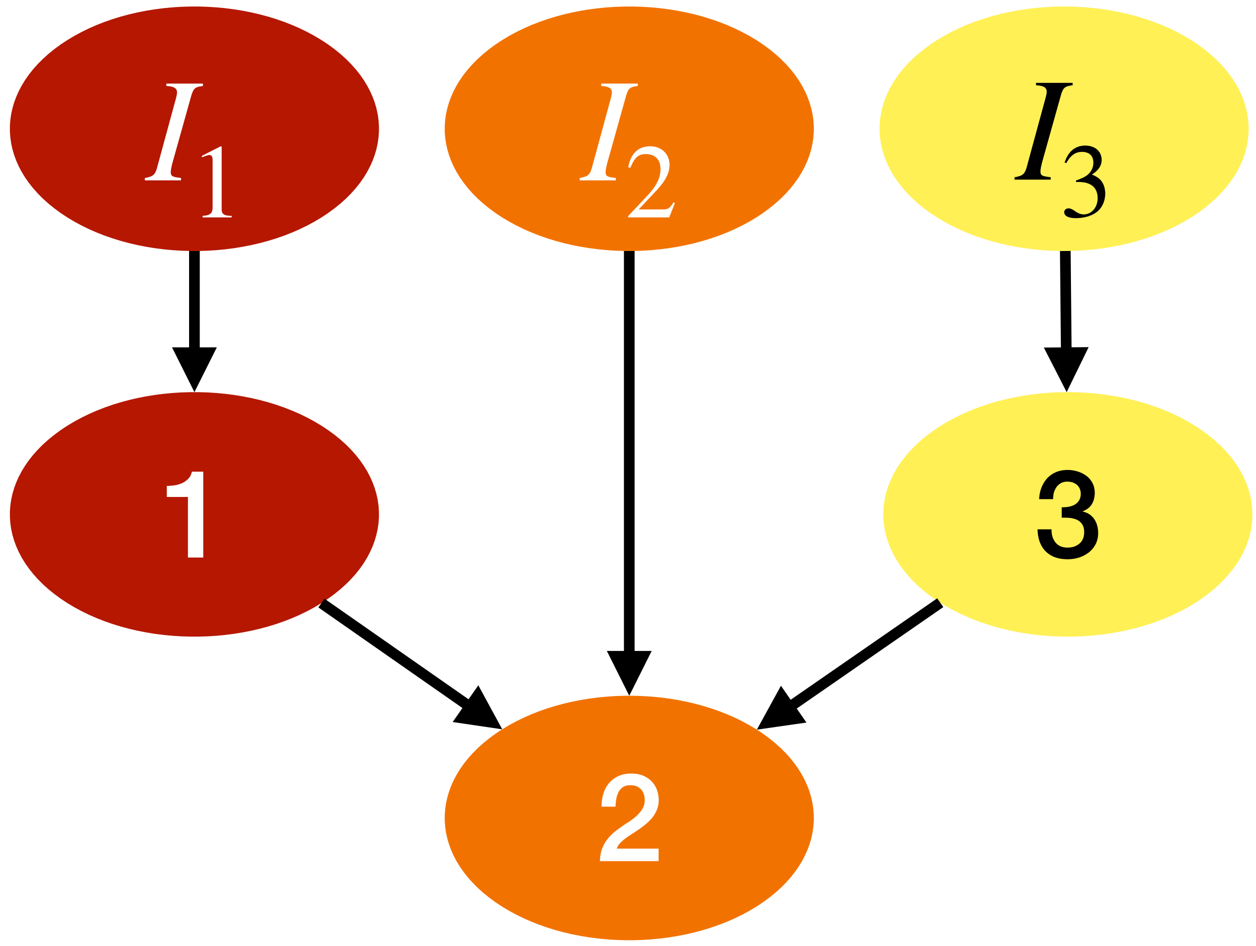
- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**
- We can create a **mutilated version** of G' after $\mathrm{do}(X_W)$ for $\mathbf{W} \subset \mathbf{V}$, $G'_{\mathrm{do}(X_{\mathbf{W}})}$



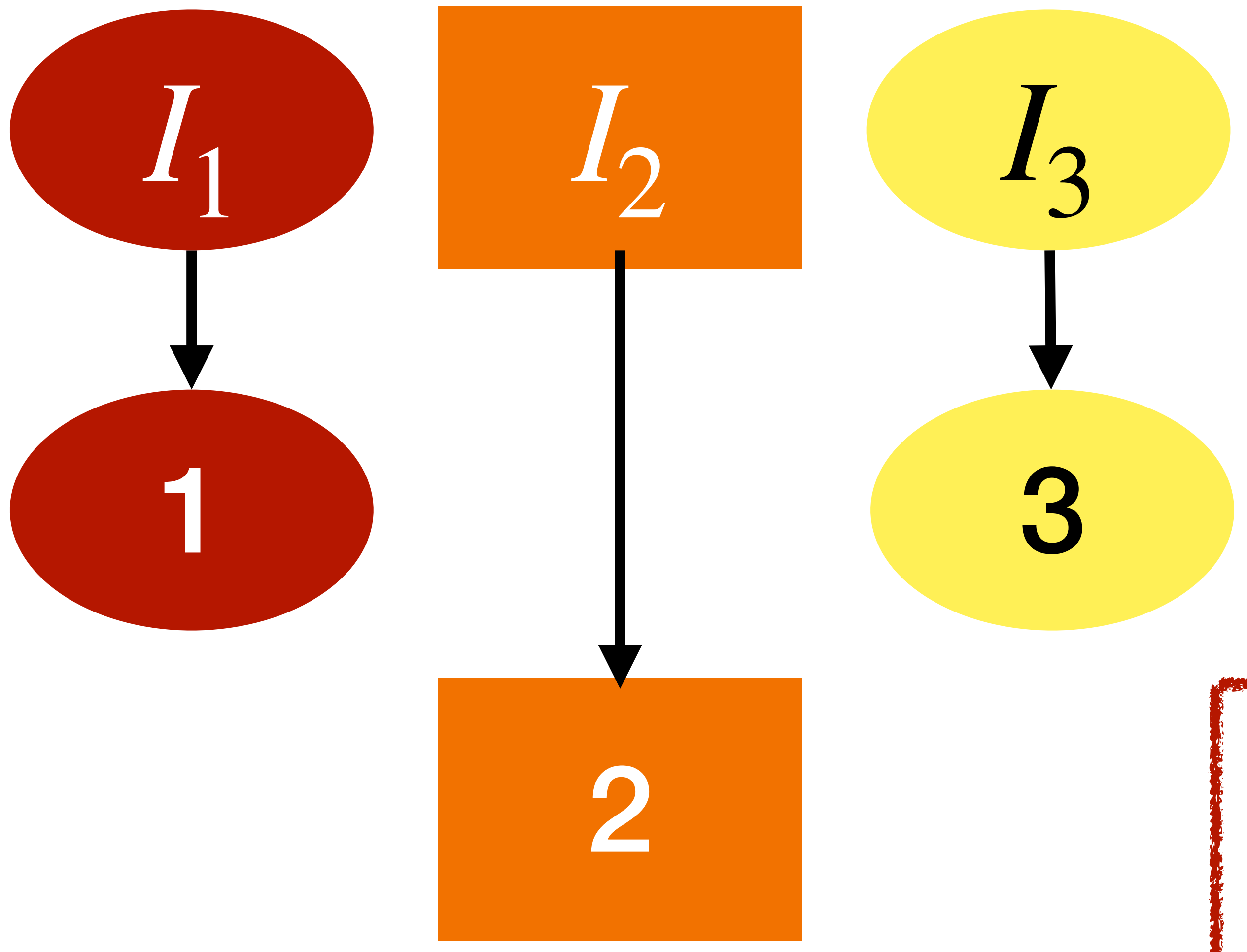
$$G' = (\mathbf{V} \cup \{I_i\}_{i \in \mathbf{V}}, \mathbf{E} \cup \{I_i \to i\}_{i \in \mathbf{V}})$$

Multiple consequent interventions

- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**
- We can create a **mutilated version** of G' after $\mathrm{do}(X_W)$ for $\mathbf{W} \subset \mathbf{V}$, $G'_{\mathrm{do}(X_\mathbf{W})}$



$G' = (\mathbf{V} \cup \{I_i\}_{i \in \mathbf{V}}, \mathbf{E} \cup \{I_i \to i\}_{i \in \mathbf{V}})$

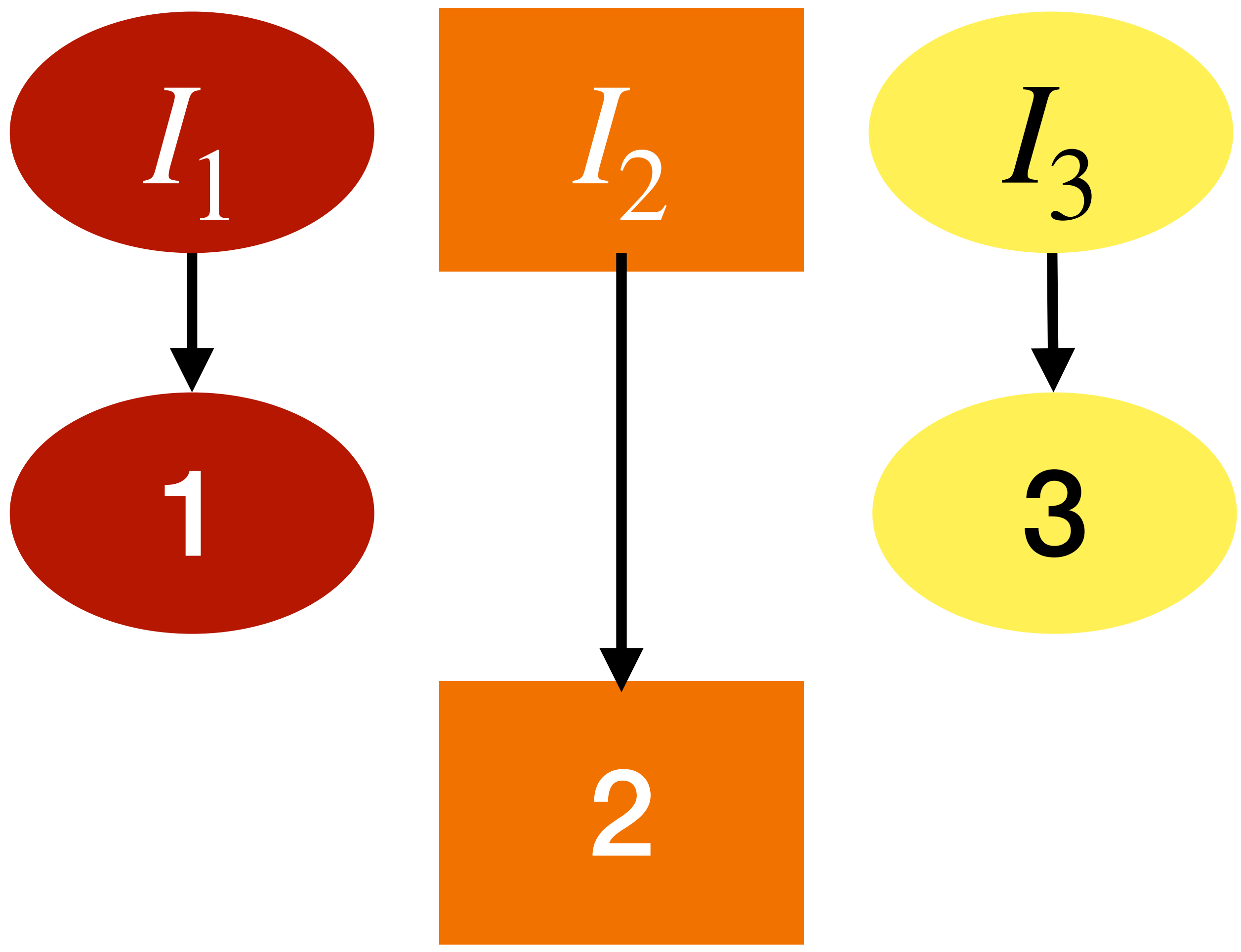


$G'_{\mathrm{do}(X_2)}$

Rectangle = fixed value

Multiple consequent interventions

- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**
- We can create a **mutilated version** of G' after $\mathrm{do}(X_W)$ for $\mathbf{W} \subset \mathbf{V}$, $G'_{\mathrm{do}(X_\mathbf{W})}$



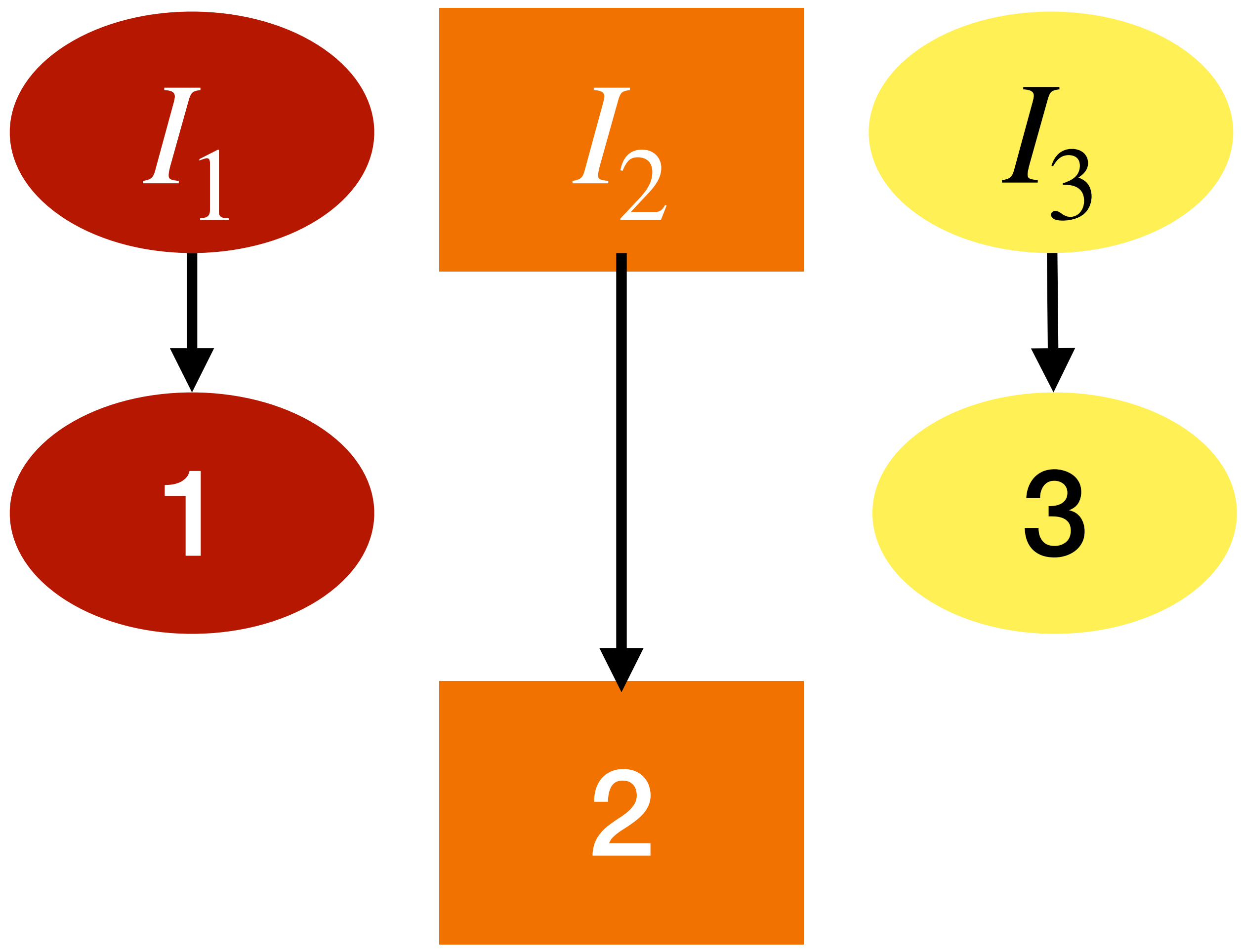
$G'_{\mathrm{do}(X_2)}$

$$I_1 \not\perp_{G'} X_3 \mid X_2$$

$$I_1 \perp_{G'_{\mathrm{do}(X_2)}} X_3 \mid X_2$$

Multiple consequent interventions

- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**
- We can create a **mutilated version** of G' after $\mathrm{do}(X_W)$ for $\mathbf{W} \subset \mathbf{V}$, $G'_{\mathrm{do}(X_\mathbf{W})}$



$G'_{\mathrm{do}(X_2)}$

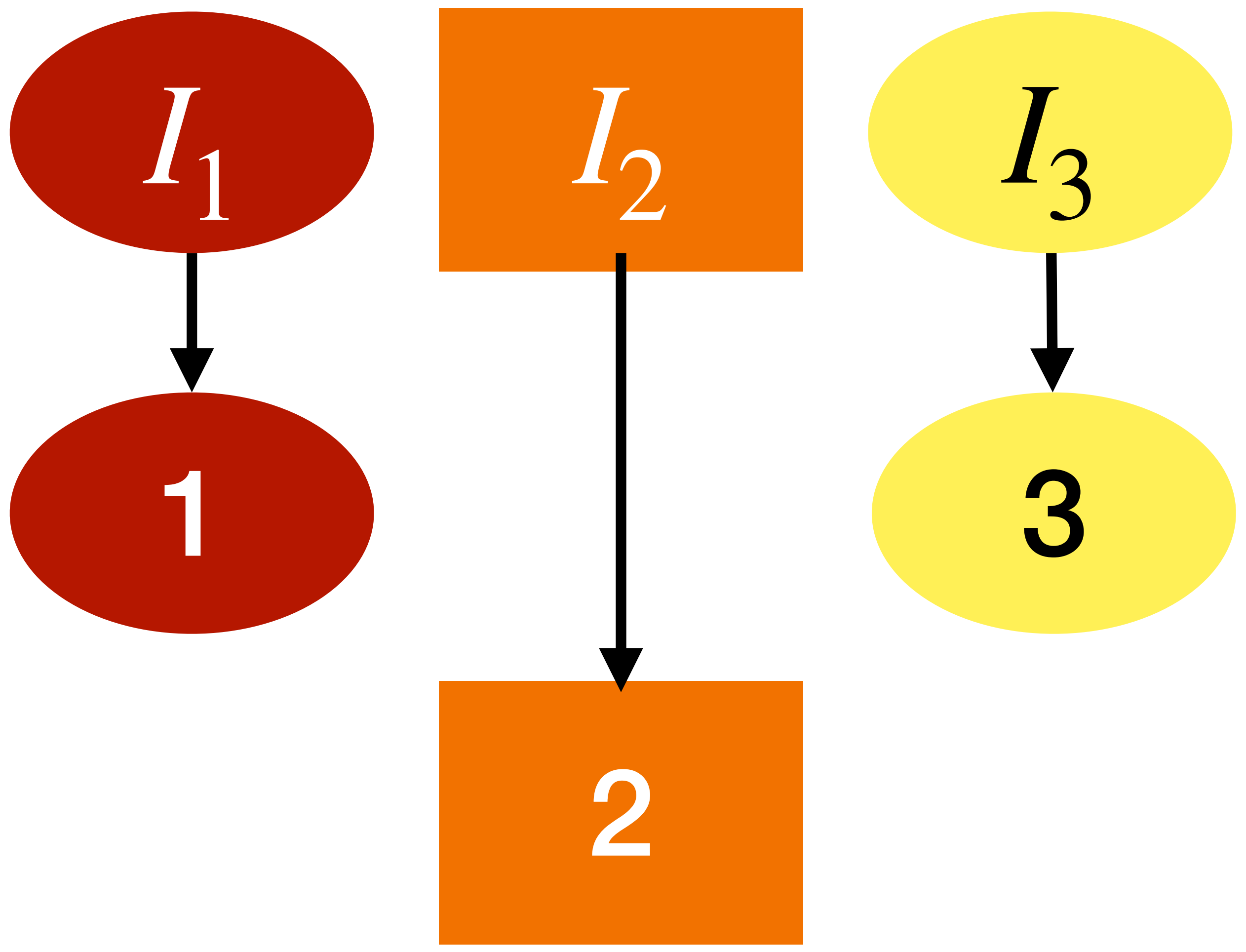
$$I_1 \perp_{G'_{\mathrm{do}(X_2)}} X_3 \mid X_2$$

$$P(X_\mathbf{Y} \mid X_\mathbf{W}, \mathrm{do}(X_\mathbf{T} = x_\mathbf{T})) := \frac{P(X_\mathbf{Y}, X_\mathbf{W} \mid \mathrm{do}(X_\mathbf{T} = x_\mathbf{T}))}{P(X_\mathbf{W} \mid \mathrm{do}(X_\mathbf{T} = x_\mathbf{T}))}$$

CONDITIONING ALWAYS AFTER INTERVENTION

Multiple consequent interventions

- Given a DAG $G = (\mathbf{V}, \mathbf{E})$ we can create an augmented graph G' with additional **intervention variables**
- We can create a **mutilated version** of G' after $\text{do}(X_W)$ for $\mathbf{W} \subset \mathbf{V}$, $G'_{\text{do}(X_\mathbf{W})}$



$G'_{\text{do}(X_2)}$

$$I_1 \perp_{G'_{\text{do}(X_2)}} X_3 \mid X_2$$

$$P(X_3 \mid X_2, I_1 = 1, \text{do}(X_2)) = P(X_3 \mid X_2, I_1 = 0, \text{do}(X_2))$$

$$P(X_3 \mid X_2, \text{do}(X_1), \text{do}(X_2)) = P(X_3 \mid X_2, \text{do}(X_2))$$

Do-calculus (intervention variable version)

- For disjoint sets $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{W} \subseteq \mathbf{V}$:

1. Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} \mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \,|\, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

$$\mathbf{A} \perp_{G'} \mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \,|\, X_\mathbf{C})$$

Do-calculus (intervention variable version)

- For disjoint sets $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{W} \subseteq \mathbf{V}$:

1. Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

2. Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

$$\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B} = x_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{B} = x_\mathbf{B}, X_\mathbf{C})$$

Do-calculus (intervention variable version)

- For disjoint sets $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{W} \subseteq \mathbf{V}$:

1. Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

2. Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

3. Rule 3: insertion/deletion of actions

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

Do-calculus example 2

- Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

- Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

- Rule 3: insertion/deletion of actions

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

$$p(Y \mid \mathrm{do}(X)) = ?$$

Do-calculus example 2

- Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

- Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

- Rule 3: insertion/deletion of actions

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

$$p(X, Y, M \mid \mathrm{do}(X = \tilde{x})) = \int_h p(M \mid X, do(X = \tilde{x})) \cdot p(Y \mid M, H = h, do(X = \tilde{x})) \cdot 1(X = \tilde{x}) dh$$

$M \perp_{G'} I_X \mid X$

$$= p(M \mid X = \tilde{x}) \cdot \int_h p(Y \mid M, H = h, do(X = \tilde{x})) dh$$

$$= p(M \mid X = \tilde{x}) \cdot p(Y \mid M, do(X = \tilde{x}))$$

Basic rules of probability

Do-calculus example 2

- Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'_{do(X_{\mathbf{W}})}} \mathbf{B} \,|\, \mathbf{C} \implies P(X_{\mathbf{A}} \,|\, X_{\mathbf{B}}, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}})) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}}))$$

- Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'_{do(X_{\mathbf{W}})}} I_{\mathbf{B}} \,|\, \mathbf{B}, \mathbf{C} \implies P(X_{\mathbf{A}} \,|\, \mathrm{do}(X_{\mathbf{B}}), X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}})) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{B}}, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}}))$$

- Rule 3: insertion/deletion of actions

$$\mathbf{A} \perp_{G'_{do(X_{\mathbf{W}})}} I_{\mathbf{B}} \,|\, \mathbf{C} \implies P(X_{\mathbf{A}} \,|\, \mathrm{do}(X_{\mathbf{B}}), X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}})) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}}))$$

$$p(X, Y, M \,|\, \mathrm{do}(X = \tilde{x})) = p(M \,|\, X = \tilde{x}) \cdot p(Y \,|\, M, do(X = \tilde{x}))$$

$$\int_x \int_m \dots$$

$$p(Y \,|\, \mathrm{do}(X)) = \int_m p(M = m \,|\, X = \tilde{x}) \cdot p(Y \,|\, M = m, do(X = \tilde{x}))\; dm$$

Do-calculus example 2

- Rule 1: insertion/deletion of observations

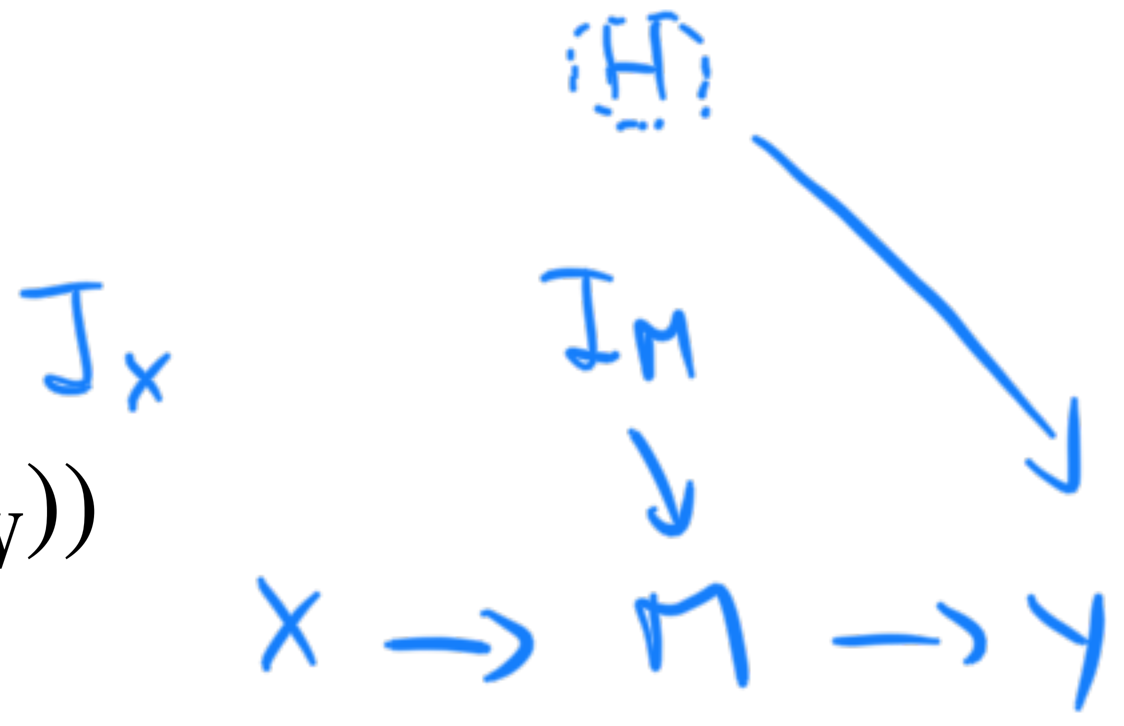
$$\mathbf{A} \perp_{G'_{do(X_{\mathbf{W}})}} \mathbf{B} \,|\, \mathbf{C} \implies P(X_{\mathbf{A}} \,|\, X_{\mathbf{B}}, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}})) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}}))$$

- Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'_{do(X_{\mathbf{W}})}} I_{\mathbf{B}} \,|\, \mathbf{B}, \mathbf{C} \implies P(X_{\mathbf{A}} \,|\, \mathrm{do}(X_{\mathbf{B}}), X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}})) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{B}}, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}}))$$

- Rule 3: insertion/deletion of actions

$$\mathbf{A} \perp_{G'_{do(X_{\mathbf{W}})}} I_{\mathbf{B}} \,|\, \mathbf{C} \implies P(X_{\mathbf{A}} \,|\, \mathrm{do}(X_{\mathbf{B}}), X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}})) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}}))$$



$$p(Y \,|\, \mathrm{do}(X = \tilde{x})) = \int_m p(M = m \,|\, X = \tilde{x}) \cdot p(Y \,|\, M = m, do(X = \tilde{x}))\, dm$$

$$Y \perp_{G'_{\mathrm{do(X)}}} I_M \,|\, M, \mathrm{do(X)},$$

$$= \int_m p(M = m \,|\, X = \tilde{x}) \cdot p(Y \,|\, do(M = m), do(X = \tilde{x}))\, dm$$

Do-calculus example 2

- Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} \mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \,|\, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

- Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \,|\, \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \,|\, \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \,|\, X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

- Rule 3: insertion/deletion of actions

$$\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \,|\, \mathbf{C} \implies P(X_\mathbf{A} \,|\, \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \,|\, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$$

$$p(Y \,|\, \mathrm{do}(X = \tilde{x})) = \int_m p(M = m \,|\, X = \tilde{x}) \cdot \boxed{p(Y \,|\, do(M = m), do(X = \tilde{x}))} \; dm$$

$$Y \perp_{G'_{\mathrm{do(M)}}} I_X \,|\, \mathrm{do(M)}$$

$$= \int_m p(M = m \,|\, X = \tilde{x}) \cdot \boxed{p(Y \,|\, do(M = m))} \; dm$$

Do-calculus example 2

- Rule 1: insertion/deletion of observations

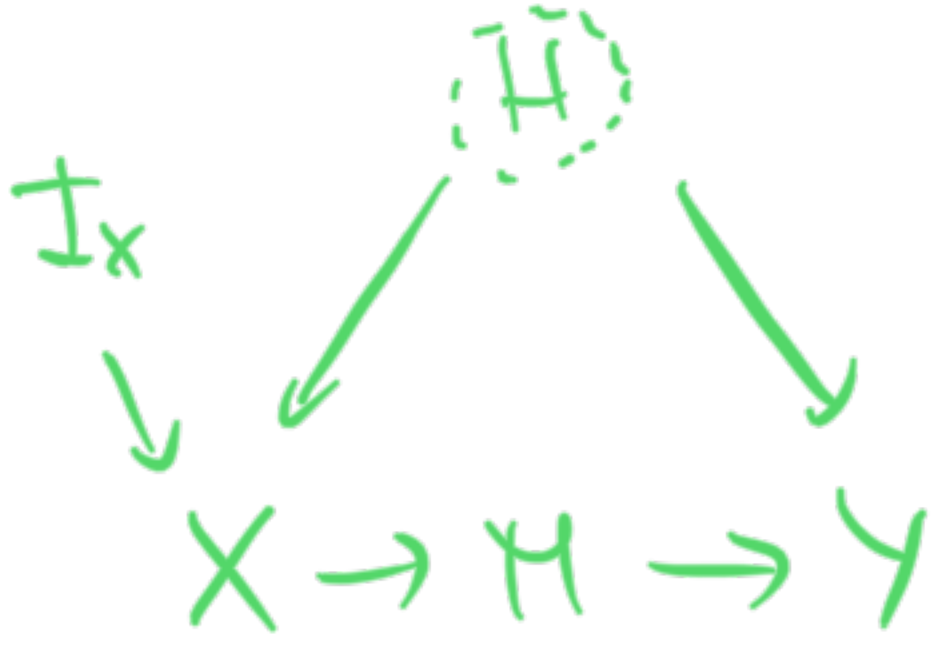
  $\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$

- Rule 2: action/observation exchange

  $\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$

- Rule 3: insertion/deletion of actions

  $\mathbf{A} \perp_{G'_{do(X_\mathbf{W})}} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}, \mathrm{do}(X_\mathbf{W})) = P(X_\mathbf{A} \mid X_\mathbf{C}, \mathrm{do}(X_\mathbf{W}))$



$$p(Y \mid \mathrm{do}(X = \tilde{x})) = \int_m p(M = m \mid X = \tilde{x}) \cdot p(Y \mid do(M = m))\, dm$$

$Y \perp_{G'} I_M \mid X, M \quad X \perp_{G'} I_M$

$$= \int_m p(M = m \mid X = \tilde{x}) \cdot \int_{x'} p(Y \mid X = x', M = m) p(X = x') dx'\, dm$$

FRONTDOOR ADJ.

Do-calculus summary

- Rule 1: insertion/deletion of observations

$$\mathbf{A} \perp_{G'_{do(X_{\mathbf{W}})}} \mathbf{B} \,|\, \mathbf{C}, \mathrm{do}(\mathbf{W}) \implies P(X_{\mathbf{A}} \,|\, X_{\mathbf{B}}, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}})) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}}))$$

- Rule 2: action/observation exchange

$$\mathbf{A} \perp_{G'_{do(X_{\mathbf{W}})}} I_{\mathbf{B}} \,|\, \mathbf{B}, \mathbf{C}, \mathrm{do}(\mathbf{W}) \implies P(X_{\mathbf{A}} \,|\, \mathrm{do}(X_{\mathbf{B}}), X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}})) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{B}}, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}}))$$

- Rule 3: insertion/deletion of actions

$$\mathbf{A} \perp_{G'_{do(X_{\mathbf{W}})}} I_{\mathbf{B}} \,|\, \mathbf{C}, \mathrm{do}(\mathbf{W}) \implies P(X_{\mathbf{A}} \,|\, \mathrm{do}(X_{\mathbf{B}}), X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}})) = P(X_{\mathbf{A}} \,|\, X_{\mathbf{C}}, \mathrm{do}(X_{\mathbf{W}}))$$

- One can show that these rules + probability axioms are **complete**
- There is a polytime algorithm for do-calculus (**ID algorithm** [Shpitser and Pearl 2006], https://cran.r-project.org/web/packages/causaleffect/index.html

Canvas quiz - do calculus (simplified version)

- Rule 1: insertion/deletion of observations

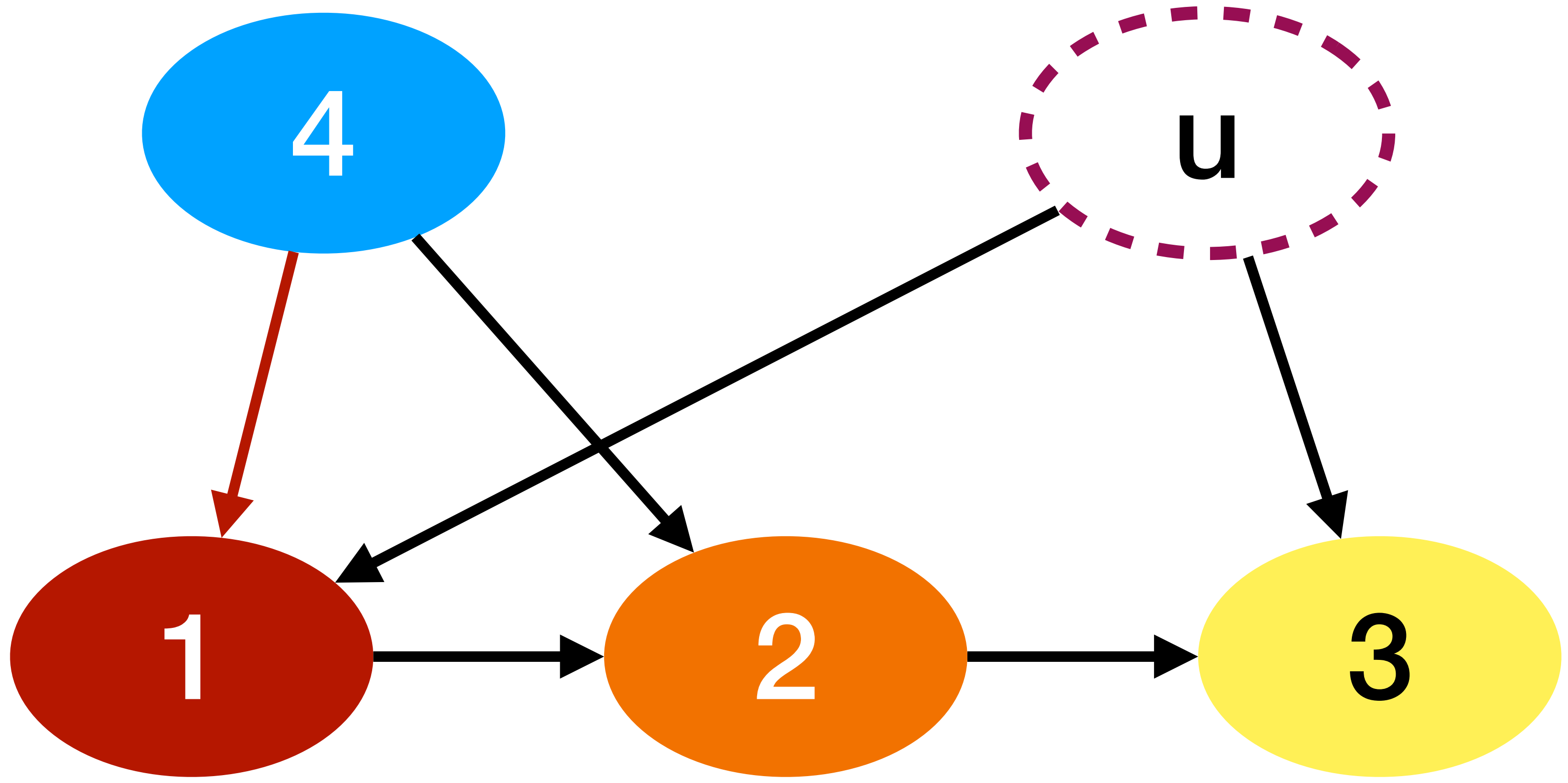
  $\mathbf{A} \perp_{G'} \mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$

- Rule 2: action/observation exchange

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{B}, \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{B}, X_\mathbf{C})$

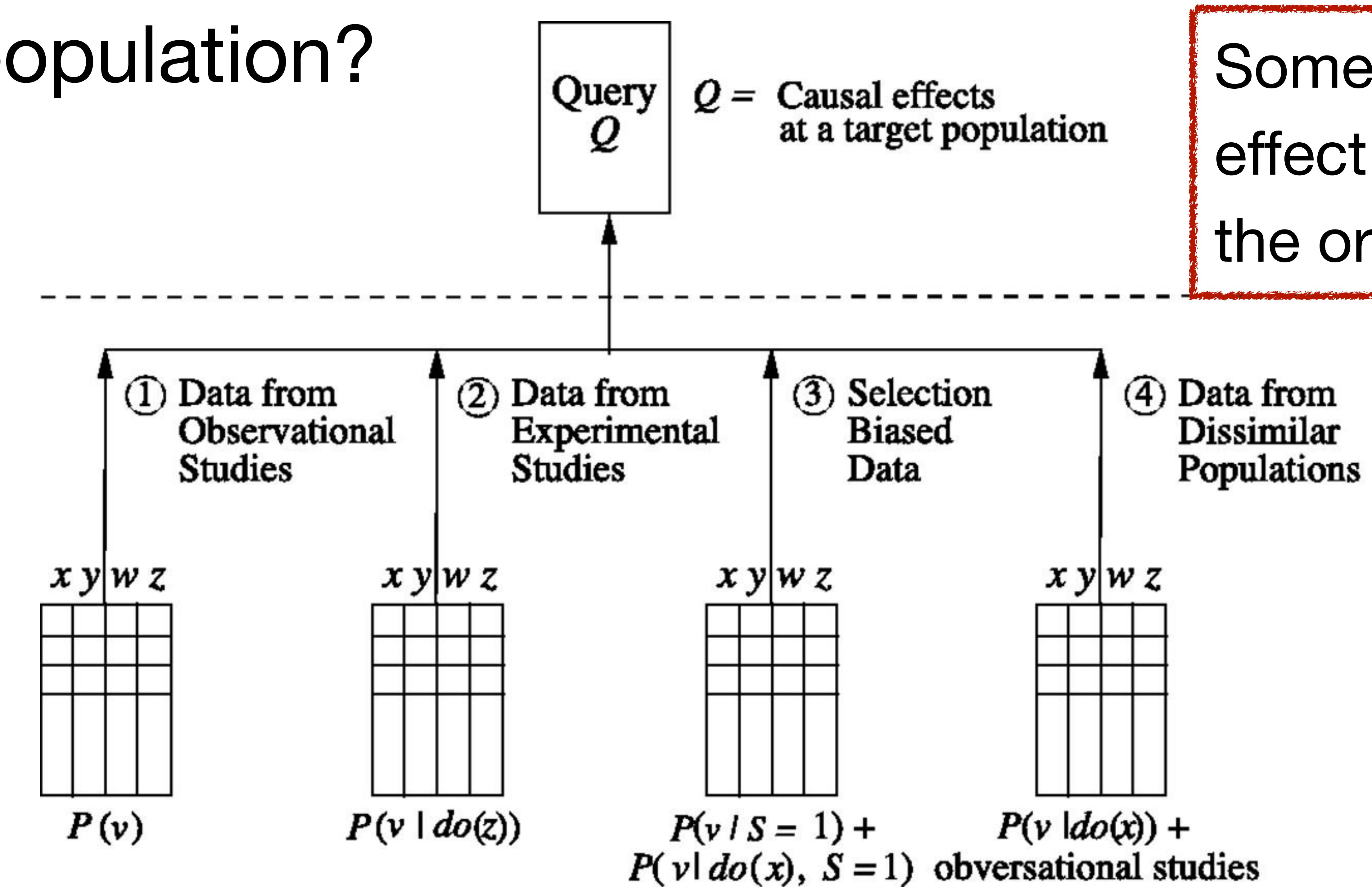
- Rule 3: insertion/deletion of actions

  $\mathbf{A} \perp_{G'} I_\mathbf{B} \mid \mathbf{C} \implies P(X_\mathbf{A} \mid \mathrm{do}(X_\mathbf{B}), X_\mathbf{C}) = P(X_\mathbf{A} \mid X_\mathbf{C})$



Transportability [Bareinboim and Pearl 2016]

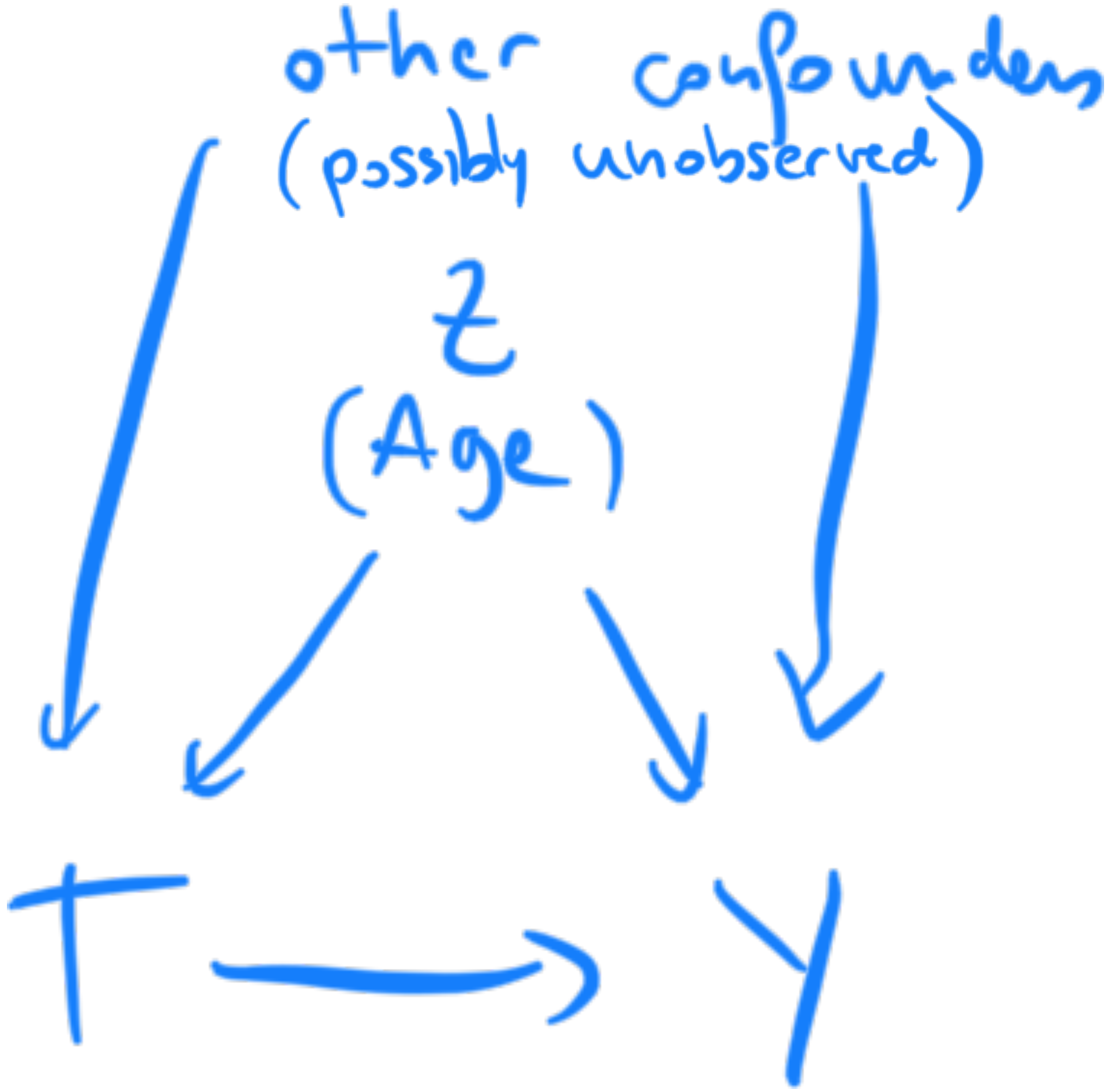
- How to combine the data from different **observational** and **experimental** conditions, each conducted on a different population, to estimate a causal effect on a target population?



Sometimes even if causal effect was not identifiable in the original population!

https://www.pnas.org/doi/10.1073/pnas.1510507113

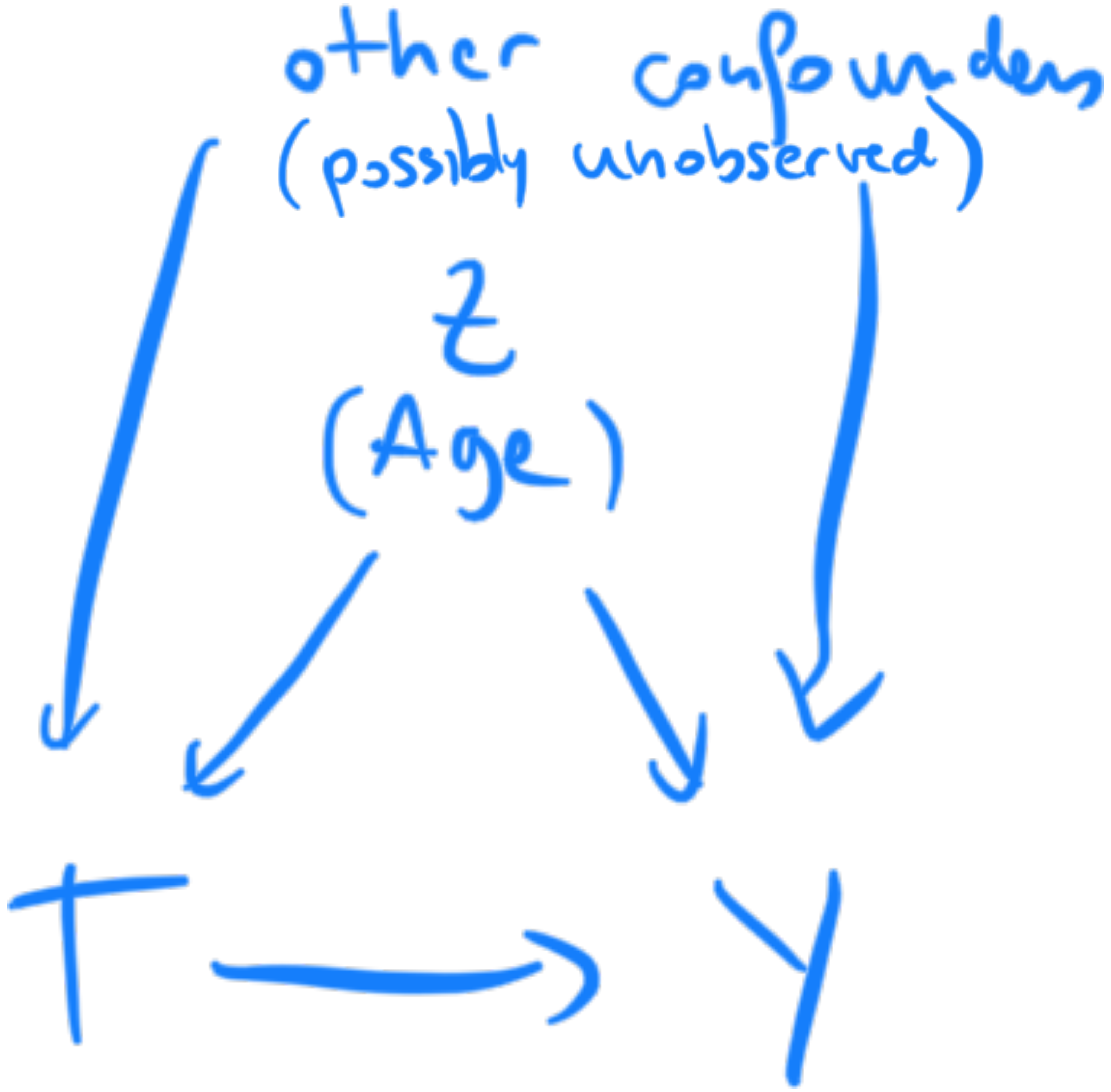
Transportability example



1. LA data
   - We can do a randomised controlled trial (RCT) and estimate $P^{LA}(Y \mid \mathrm{do}(T))$ and $P^{LA}(Y \mid \mathrm{do}(T), Z)$

2. NY data
   - Older population $P^{LA}(Z) \neq P^{NY}(Z)$
   - Other confounders are the same
   - We have $P^{NY}(Z)$, what is $P^{NY}(Y \mid \mathrm{do}(T))$?

Transportability example



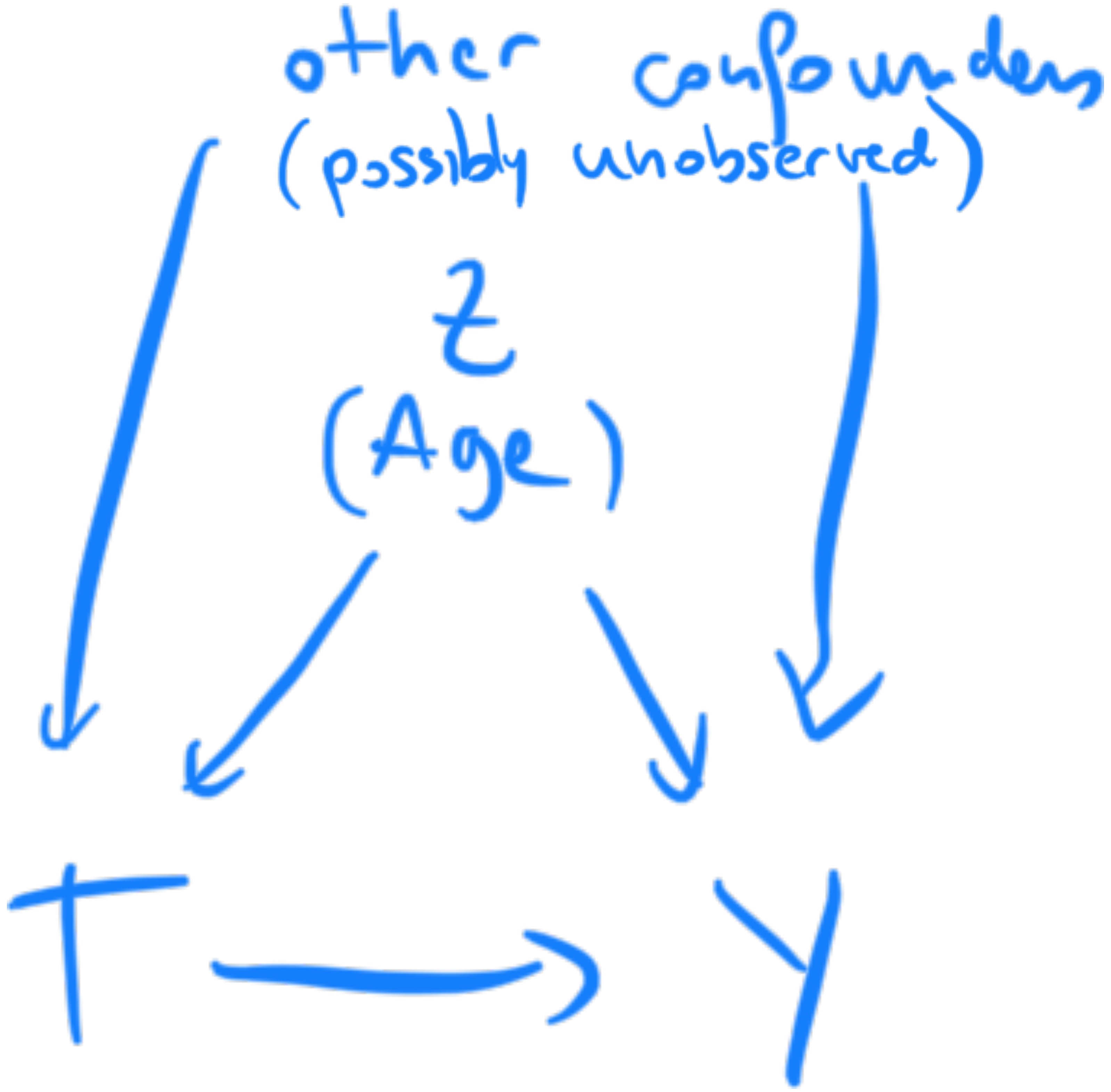
1. LA data
   - $P^{LA}(Y \mid \mathrm{do}(T), Z)$
2. NY data
   - $P^{NY}(Y \mid \mathrm{do}(T))$?

We assume Z-specific (age) effects are the same:

$$P^{LA}(Y \mid \mathrm{do}(T), Z) = P^{NY}(Y \mid \mathrm{do}(T), Z)$$

$$P^{NY}(Y \mid \mathrm{do}(T)) = \sum_Z P^{NY}(Y \mid do(T), Z) P^{NY}(Z)$$

Transportability example



1. LA data
   - $P^{LA}(Y \mid \mathrm{do}(T), Z)$
2. NY data
   - $P^{NY}(Y \mid \mathrm{do}(T))$?

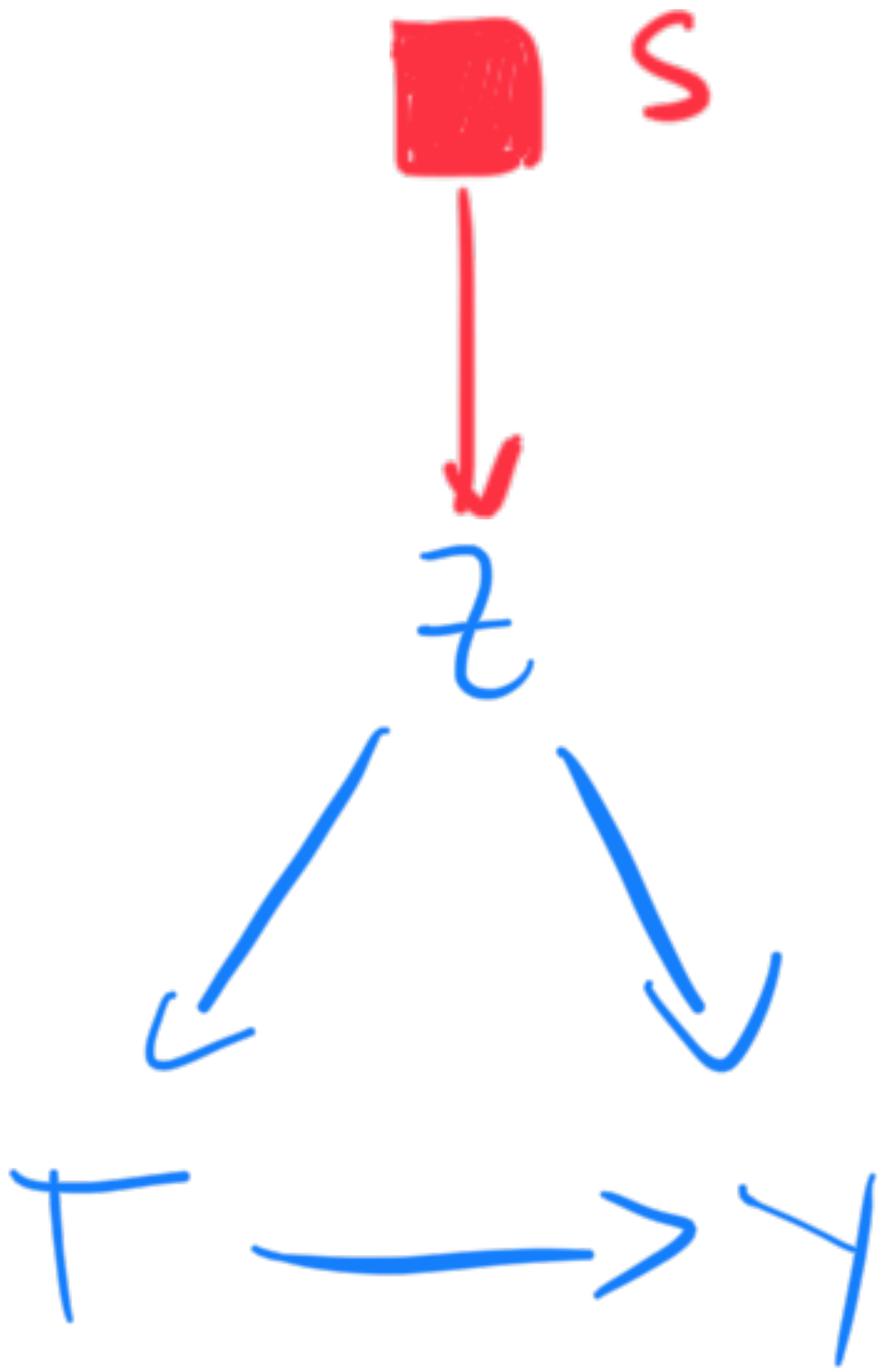
Transport formula:

Assuming Z-specific effects are invariant

$$P^{NY}(Y \mid \mathrm{do}(T)) = \sum_Z P^{LA}(Y \mid do(T), Z) P^{NY}(Z)$$

$$P^{LA}(Y \mid \mathrm{do}(T), Z) = P^{NY}(Y \mid \mathrm{do}(T), Z)$$

Selection diagrams (generalised assumptions)

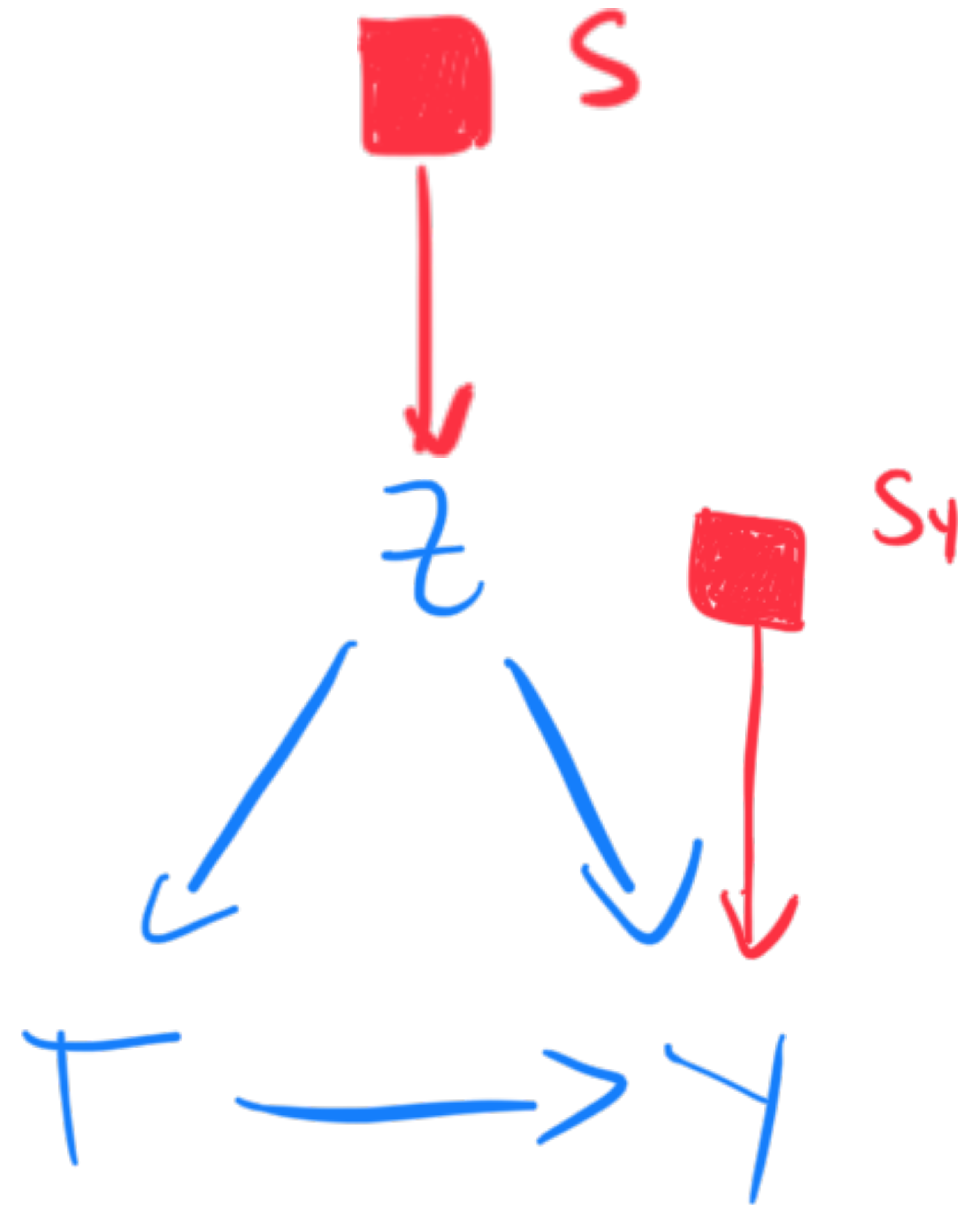


1. LA data
   - $P^{LA}(Y \mid \mathrm{do}(T), Z)$
2. NY data
   - Older population $P^{LA}(Z) \neq P^{NY}(Z)$

S encodes changes

$P^{LA}(Z) \neq P^{NY}(Z)$ $\quad$ $P(Z, S=0) \neq P(Z, S=1)$

Selection diagrams (generalised assumptions)



1. LA data
   - $P^{LA}(Y \mid \text{do}(T), Z)$
2. NY data
   - Older population $P^{LA}(Z) \neq P^{NY}(Z)$

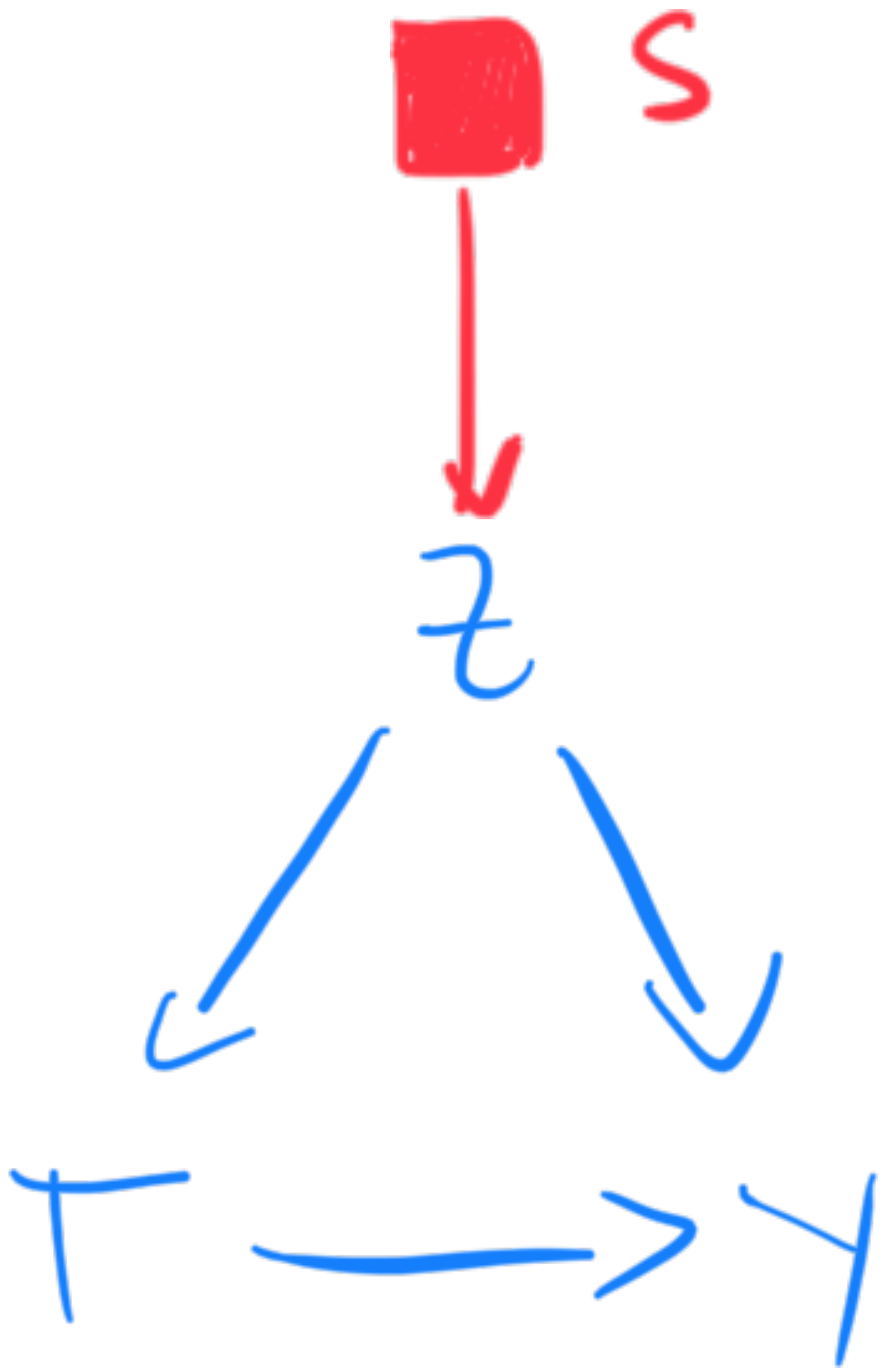
S encodes changes

$P^{LA}(Z) \neq P^{NY}(Z)$

$P^{LA}(Y \mid Z, T) \neq P^{NY}(Y \mid Z, T)$

$P^{LA}(T \mid Z) = P^{NY}(T \mid Z)$ (because no S_T)

Selection diagrams - direct transportability



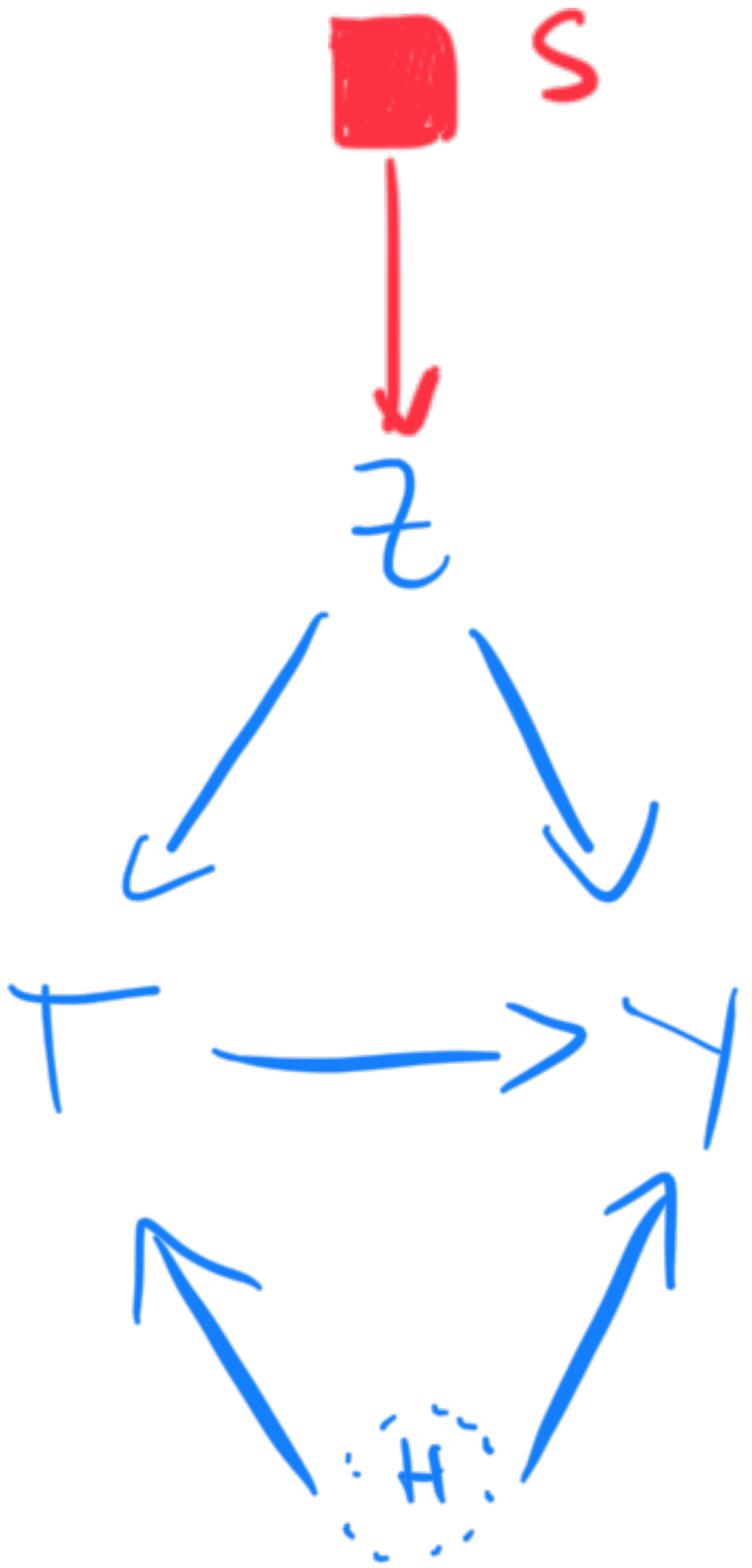
$$Y \perp_d S \mid \mathbf{Z}, \mathrm{do}(T) \implies \underbrace{P^{NY}(Y \mid \mathbf{Z}, \mathrm{do}(T))}_{\text{target}} = P^{LA}(Y \mid \mathbf{Z}, \mathrm{do}(T))$$

Before we assumed it, now it's a consequence of the graph:

$$P^{LA}(Y \mid \mathrm{do}(T), Z) = P^{NY}(Y \mid \mathrm{do}(T), Z)$$

$$P^{NY}(Y \mid \mathrm{do}(T)) = \sum_Z P^{LA}(Y \mid do(T), Z) P^{NY}(Z)$$

Selection diagrams - direct transportability



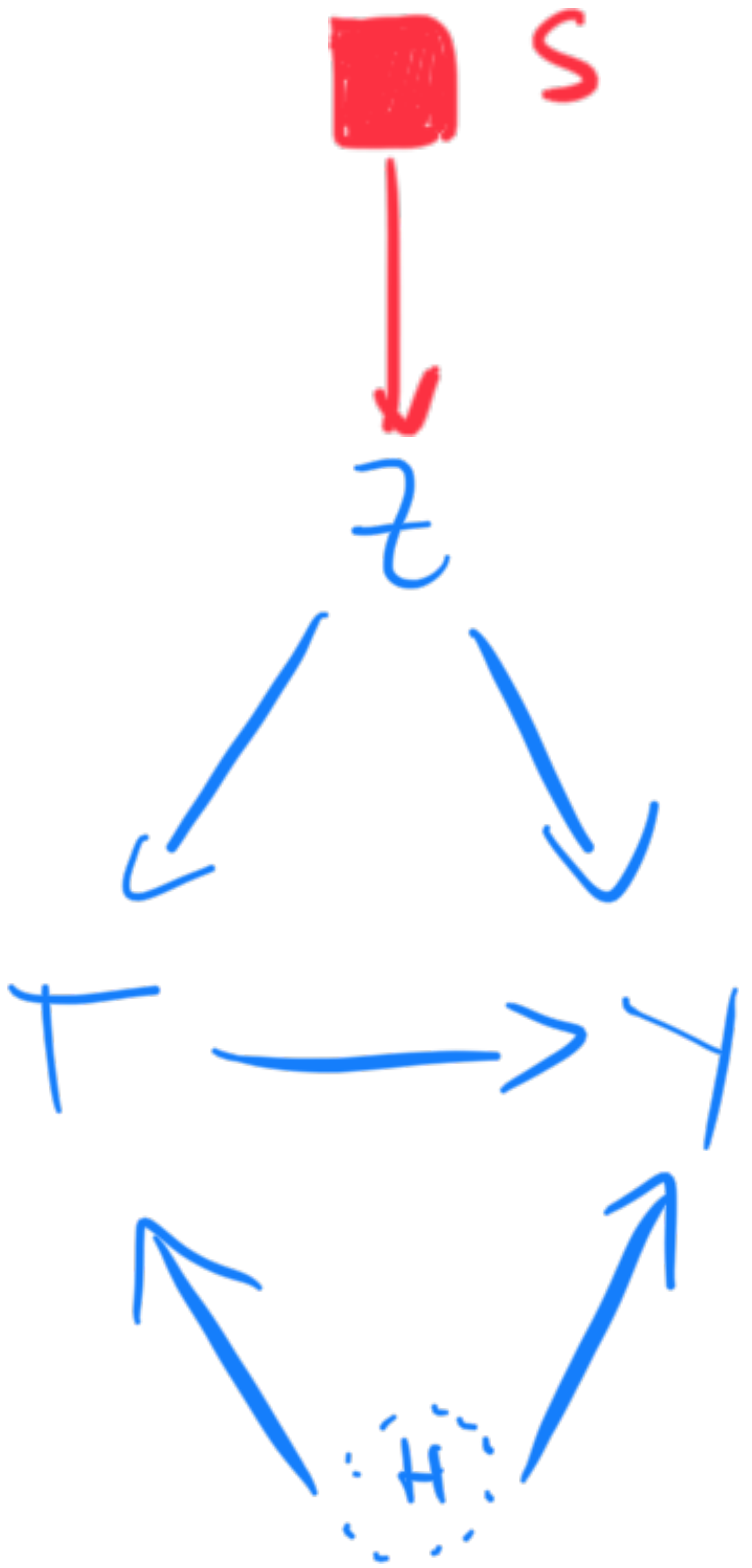
$$Y \perp_d S \mid \mathbf{Z}, \mathrm{do}(T) \implies \underbrace{P^{NY}(Y \mid \mathbf{Z}, \mathrm{do}(T))}_{\text{target}} = P^{LA}(Y \mid \mathbf{Z}, \mathrm{do}(T))$$

$$Y \perp_d S \mid \mathbf{Z}, \mathrm{do}(T) \implies P^*(Y \mid \mathbf{Z}, \mathrm{do}(T)) = P(Y \mid \mathbf{Z}, \mathrm{do}(T))$$

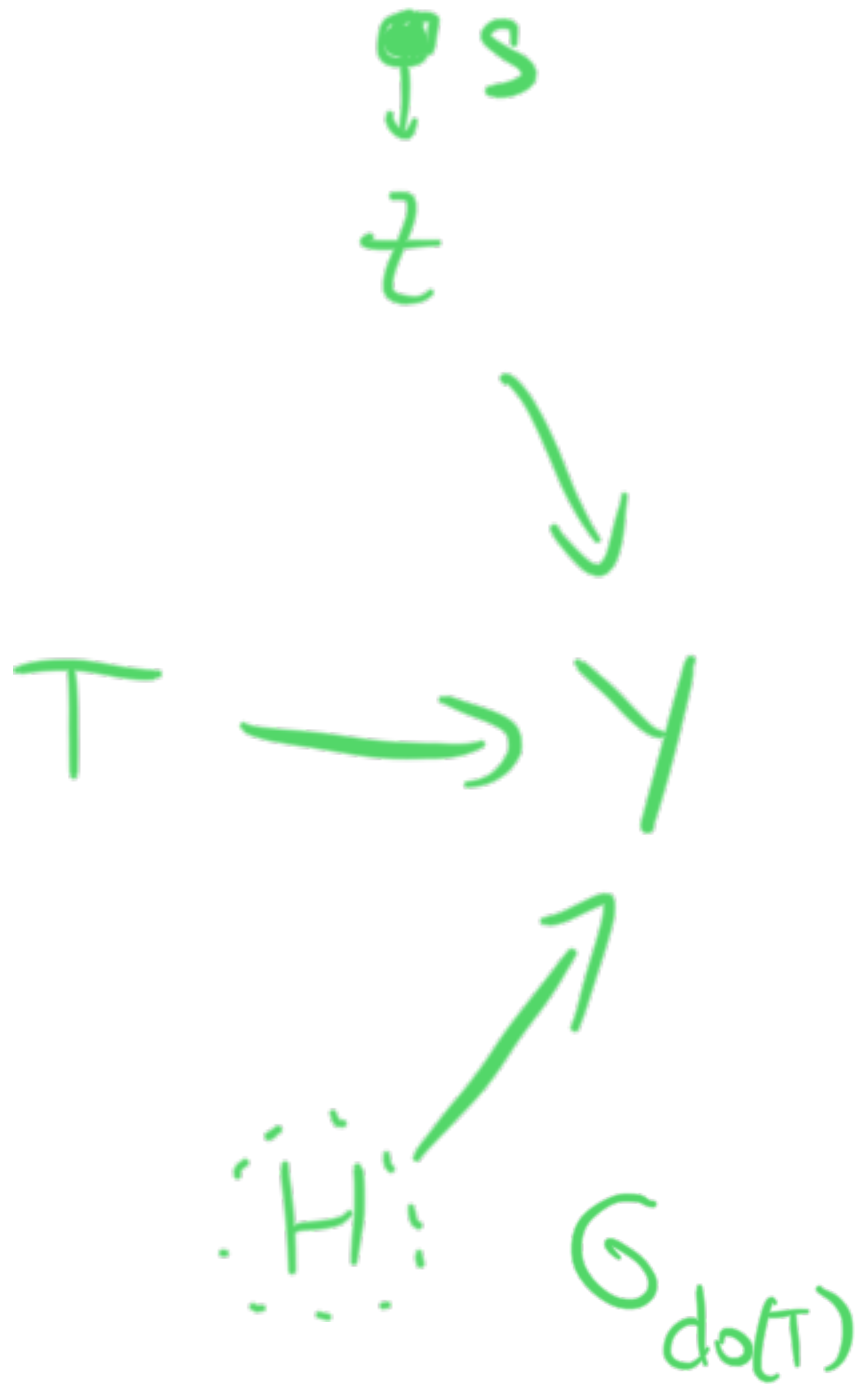
$$P^*(Y \mid \mathrm{do}(T), \mathbf{X}) = P(Y \mid \mathrm{do}(T), \mathbf{X}, S = 0)$$

We can apply do-calculus

Selection diagrams - direct transportability



$$Y \perp_d S \mid \mathbf{X}, \mathrm{do}(T) \implies \underbrace{P^*(Y \mid \mathbf{X}, \mathrm{do}(T))}_{\text{target}} = P(Y \mid \mathbf{X}, \mathrm{do}(T))$$

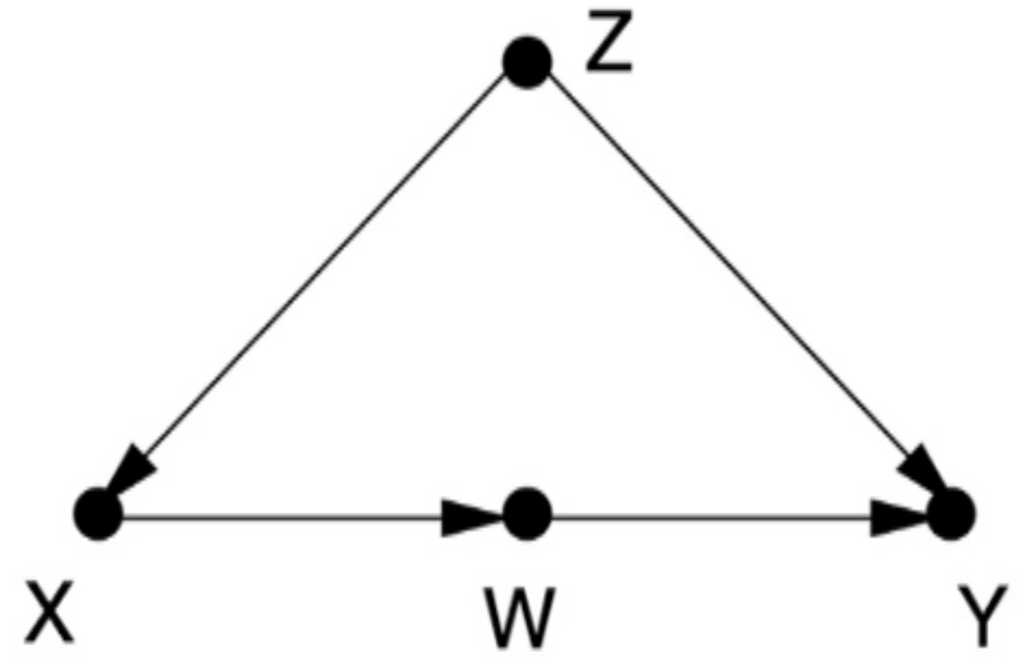


$$P^{LA}(Y \mid \mathrm{do}(T), Z) = P^{NY}(Y \mid \mathrm{do}(T), Z)$$

Selection diagrams (Book of Why Ch.10)

Causal effect of advertising surfboards X in Arkansas on purchases Y: $P(Y \mid \mathrm{do}(X))$?

(a) Target population



X = Advertisement, Y = Purchase Decision, Z = Age, W = Click-through Rate,

backdoor with Z

Z is not measured here

Selection diagrams (Book of Why Ch.10)

Causal effect of advertising surfboards X in Arkansas on purchases Y: $P(Y \mid \mathrm{do}(X))$?

(a) Target population

Arkansas a)

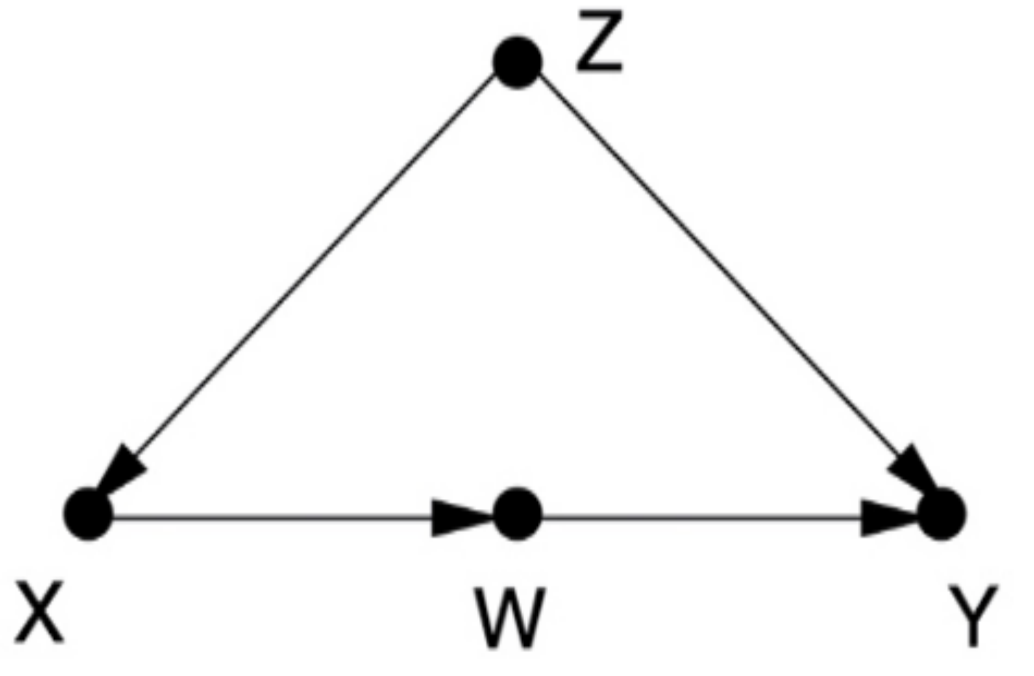
LA b)
Younger (Z)

Boston c)
lower car ownership (V)

SF d)
higher rate of (V) click through

Toronto e)
RCT
higher click through (W)

Honolulu f)
RCT
more surfboards sold (Y)



X = Advertisement, Y = Purchase Decision, Z = Age, W = Click-through Rate,

Z is not measured here

Selection diagrams (Book of Why Ch.10)

Causal effect of advertising surfboards X in Arkansas on purchases Y: $P(Y \mid \mathrm{do}(X))$?

Arkansas a)

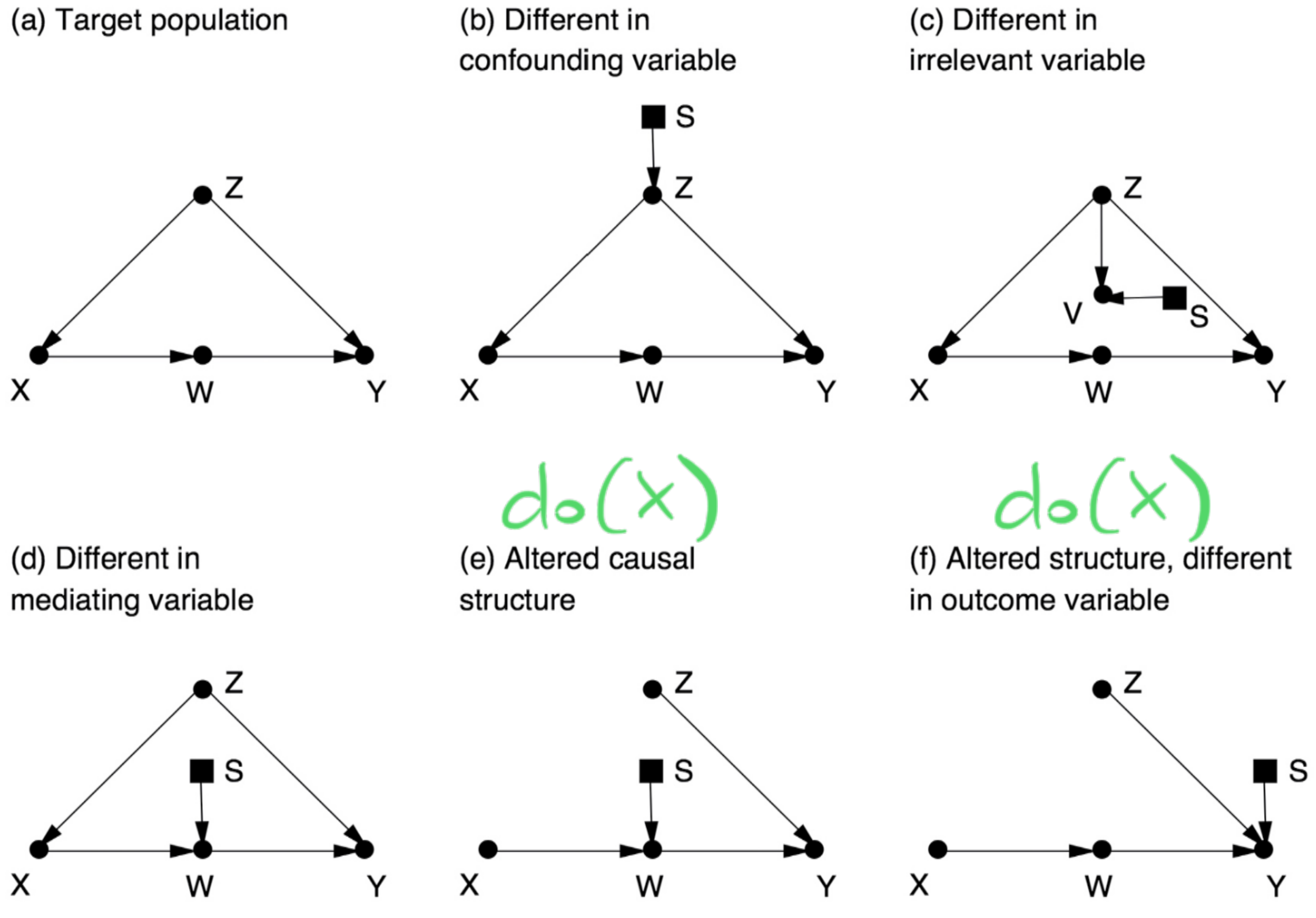
LA b) Younger (Z)

Boston c) lower car ownership (V)

SF d) higher rate of (W) click through

Toronto e) RCT higher click through (W)

Honolulu f) RCT more surfboards sold (Y)

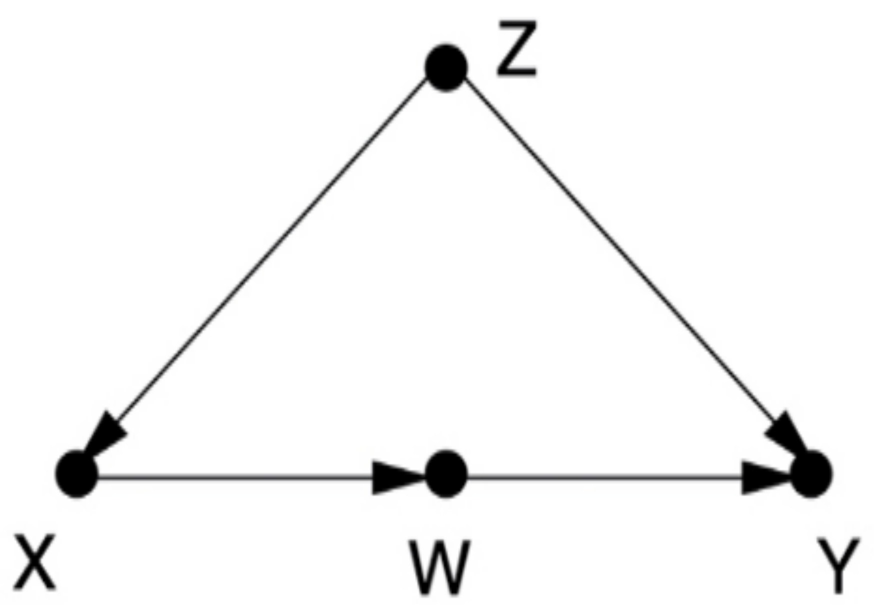


X = Advertisement, Y = Purchase Decision, Z = Age, W = Click-through Rate, V = Car Ownership, S = Indicator Variable

Selection diagrams (Book of Why Ch.10)

Causal effect of advertising surfboards X in Arkansas on purchases Y: $P(Y \mid \mathrm{do}(X))$?

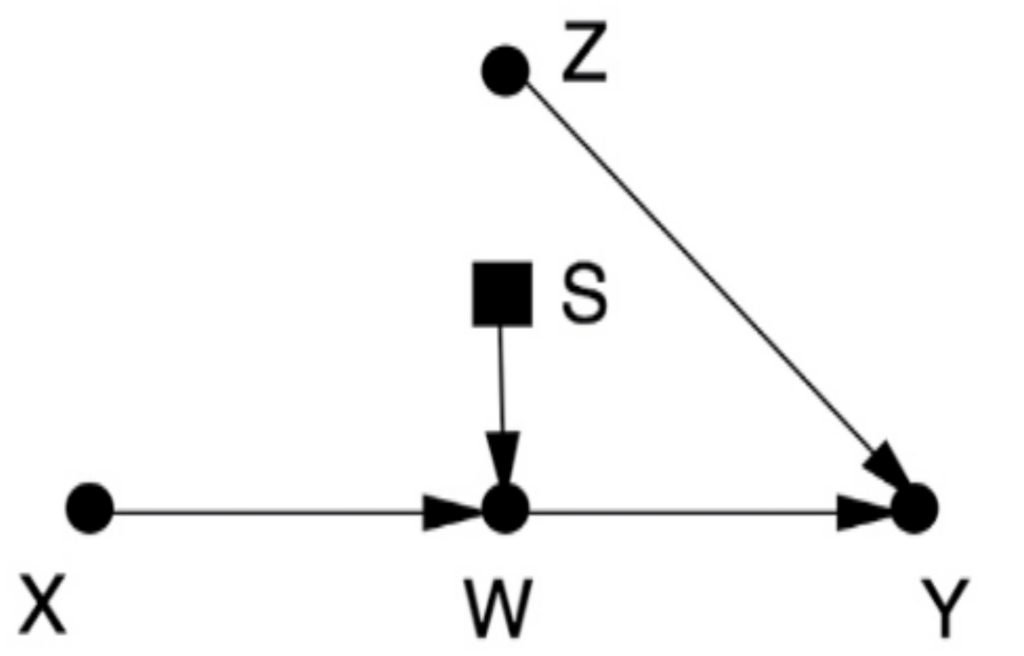
(a) Target population



$I_X \perp_d W \mid X$

$P^*(W \mid X)$

(e) Altered causal structure



$Y \perp_d S \mid W, do(X)$

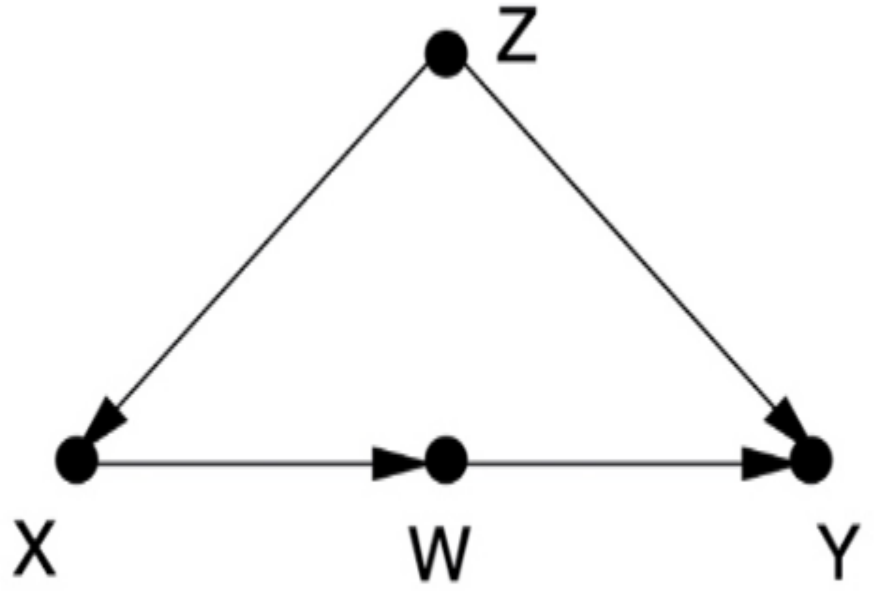
$\Rightarrow P^*(Y \mid do(X), W) = P(Y \mid do(X), W)$

$$P^*(Y \mid do(X)) = \sum_W P^*(Y \mid do(X), W) \cdot P^*(W \mid do(X))$$

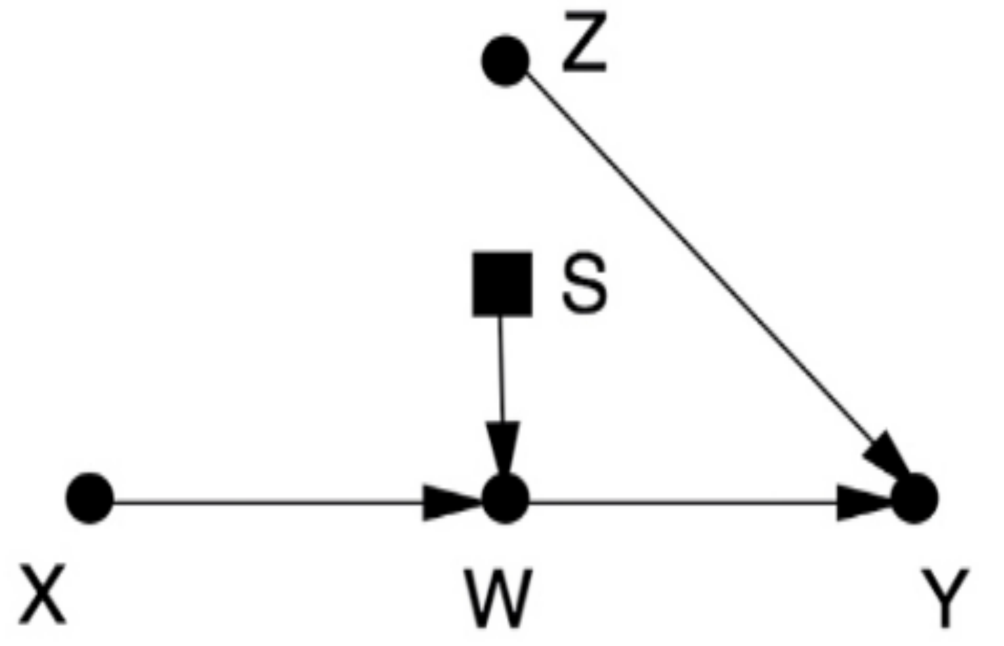
Selection diagrams (Book of Why Ch.10)

Causal effect of advertising surfboards X in Arkansas on purchases Y: $P(Y \mid \mathrm{do}(X))$?

(a) Target population



(e) Altered causal structure



$$P^*(Y \mid do(x)) = \sum_w P^*(Y \mid do(x), w) \cdot P^*(W \mid do(x)) = \sum_w P(Y \mid do(x), w) \cdot P^*(W \mid X)$$

Transportability summary

- How to combine the data from different **observational** and **experimental** conditions, each conducted on a different population, to estimate a causal effect on a target population?
- Given not only the true causal graph in the target setting, but also **the selection diagrams showing the differences in the other settings**, one can find an estimated by applying **do-calculus**
- There is an algorithm for transportability [Bareinboim and Pearl 2014] https://cran.r-project.org/web/packages/causaleffect/index.html