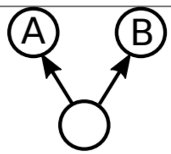
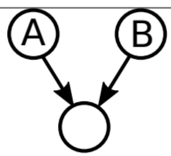
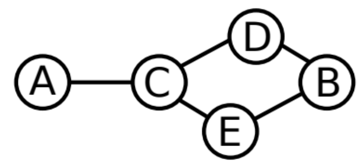
Causality, the do-calculus, confounders, etc

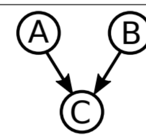
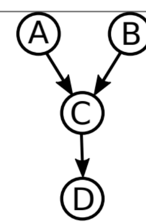
UNIVERSITY of WASHINGTON

What is independence in a causal graph?

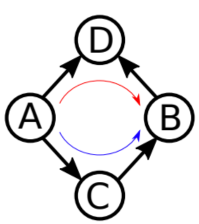
	$A \not\perp\!\!\!\perp B$
	$A \not\perp\!\!\!\perp B$
	$A \not\perp\!\!\!\perp B$
	$A \perp\!\!\!\perp B$

Conditional independence

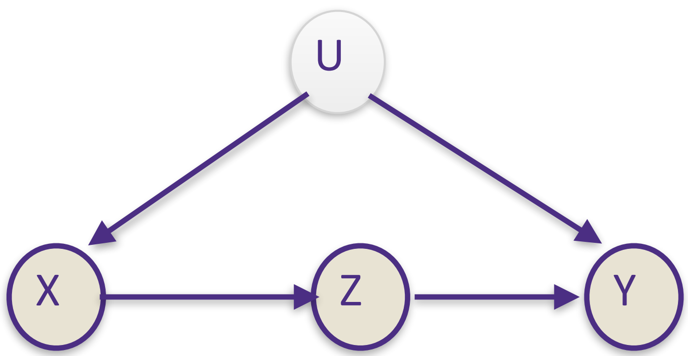
	$A \perp\!\!\!\perp B \mid C$ $A \perp\!\!\!\perp B \mid \{D, E\}$ $A \perp\!\!\!\perp B \mid \{C, D\}$ $A \not\perp\!\!\!\perp B \mid D$

	$A \perp\!\!\!\perp B$ $A \not\perp\!\!\!\perp B \mid C$
	$A \perp\!\!\!\perp B$ $A \not\perp\!\!\!\perp B \mid C$ $A \not\perp\!\!\!\perp B \mid D$

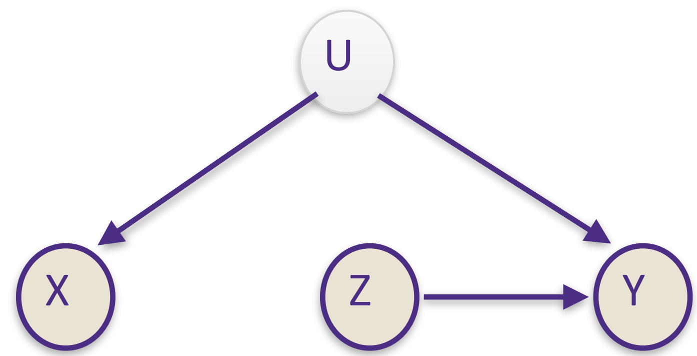
colliders

	$A \not\perp\!\!\!\perp B$ because blue path exists $A \perp\!\!\!\perp B \mid C$ because blue path is blocked by C, and red path is blocked by collider $A \not\perp\!\!\!\perp B \mid D$ because blue path is not blocked, *and* D 'explains away' influence. $A \not\perp\!\!\!\perp B \mid \{C, D\}$ because observed D 'explains away' influence.

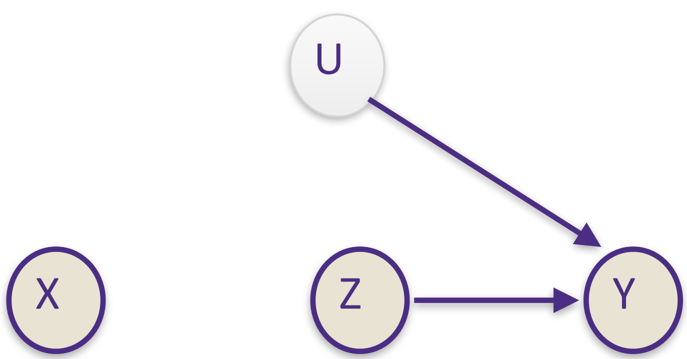
Causal graphs and interventions



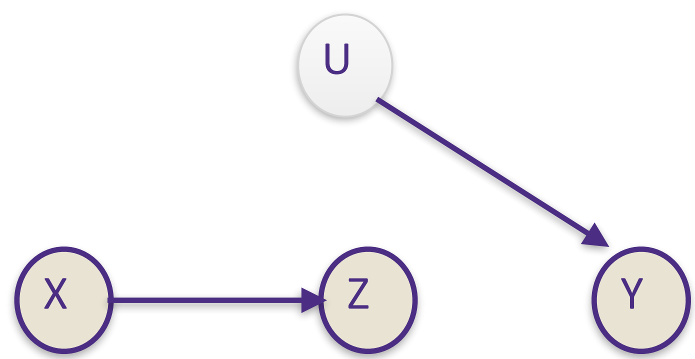
Graph G



Graph $G_{\underline{X}} = G_{\overline{Z}}$

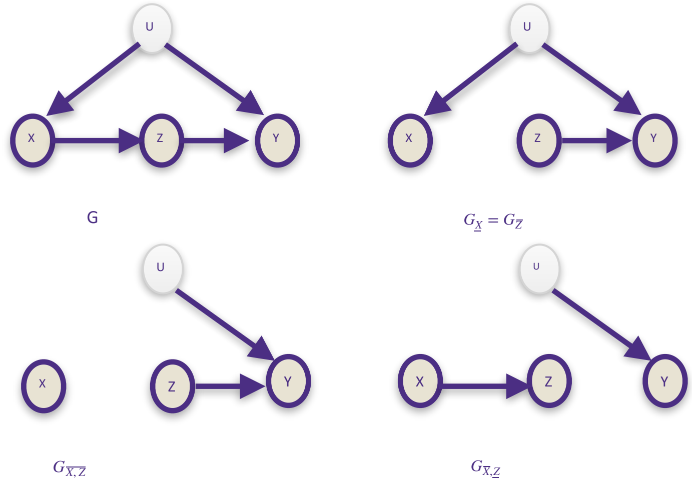


Graph $G_{\overline{X,Z}}$



Graph $G_{\overline{X},\underline{Z}}$

The do-caculus

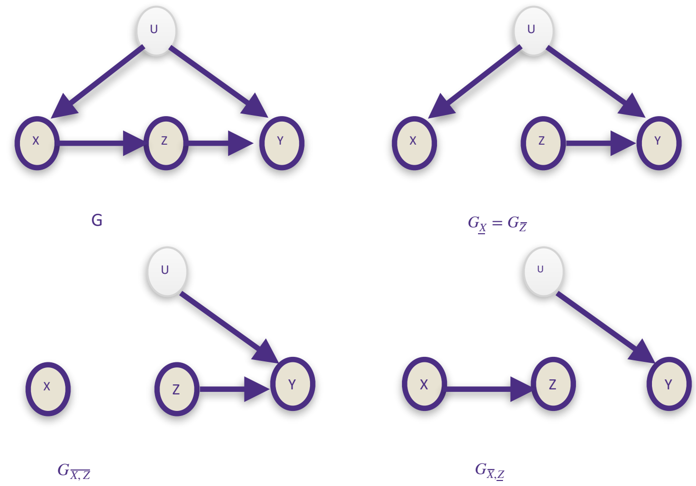


Ignorability of observations

$$\mathbb{P}[y \,|\, do(X=x), z, w] = \mathbb{P}[y \,|\, do(X=x), w] \quad \text{if } (Y \perp\!\!\!\perp Z \,|\, X, W)_{G_{\overline{X}}}$$

> **If any path in $G_{\overline{X}}$ between Y and Z are blocked by X, W and X, W aren't colliders between Y and Z**

The do-caculus

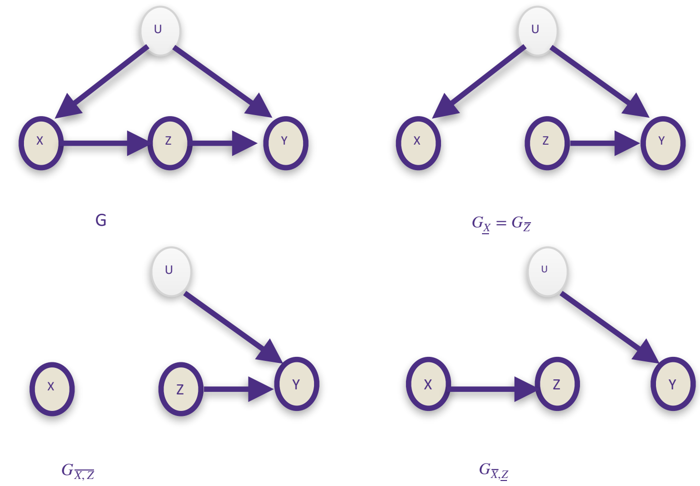


Exchangeability between action/observation (back-door criterion)

$$\mathbb{P}[y \,|\, do(X = x), do(Z = z), w] = \mathbb{P}[y \,|\, do(X = x), z, w] \text{ if } \quad (Y \perp\!\!\!\perp Z \,|\, X, W)_{G_{\overline{X},\underline{Z}}}$$

> **If every path in $G_{\overline{X},\underline{Z}}$ between Y, Z are blocked by X, W and there are no colliders between Y, Z in X, W**

The do-caculus

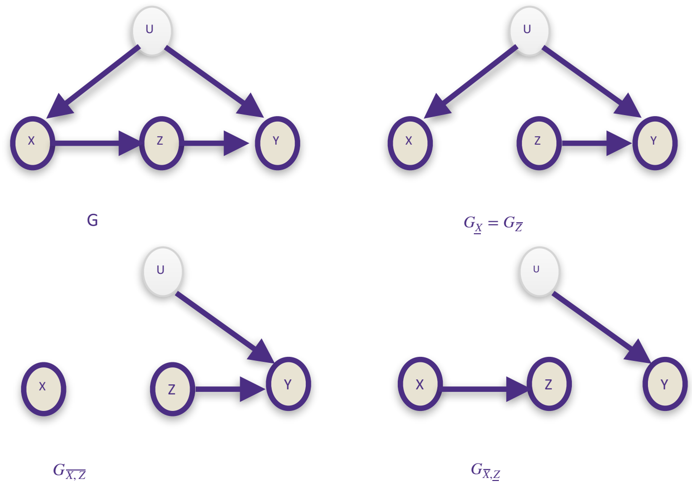


Ignorability of interventions

$$\mathbb{P}[y \,|\, do(X=x), do(Z=z), w] = \mathbb{P}[y \,|\, do(X=x), w] \text{ if } \quad (Y \perp\!\!\!\perp Z \,|\, X, W)_{G_{\overline{X},\overline{Z(W)}}}$$

> **Z(W): set of nodes in Z that aren’t ancestors of W**

The do-caculus



Ignorability of observations

$$\mathbb{P}[y \,|\, do(X=x), z, w] = \mathbb{P}[y \,|\, do(X=x), w] \quad \text{if } (Y \perp\!\!\!\perp Z \,|\, X, W)_{G_{\overline{X}}}$$

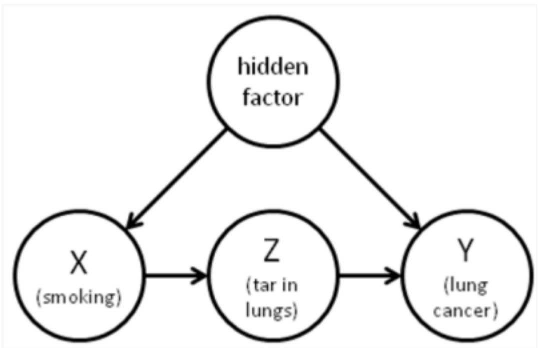
Exchangeability between action/observation (back-door criterion)

$$\mathbb{P}[y \,|\, do(X=x), do(Z=z), w] = \mathbb{P}[y \,|\, do(X=x), z, w] \text{ if } \quad (Y \perp\!\!\!\perp Z \,|\, X, W)_{G_{\overline{X},\underline{Z}}}$$

Ignorability of interventions

$$\mathbb{P}[y \,|\, do(X=x), do(Z=z), w] = \mathbb{P}[y \,|\, do(X=x), w] \text{ if } \quad (Y \perp\!\!\!\perp Z \,|\, X, W)_{G_{\overline{X},\overline{Z(W)}}}$$

The do-caculus



Ignorability of observations

$$\mathbb{P}[y \mid do(X=x), z, w] = \mathbb{P}[y \mid do(X=x), w] \quad \text{if } (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}}}$$

Exchangeability of interventions and actions (back-door)

$$\mathbb{P}[y \mid do(X=x), do(Z=z), w] = \mathbb{P}[y \mid do(X=x), z, w] \text{ if } (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X},\underline{Z}}}$$

Ignorability of interventions

$$\mathbb{P}[y \mid do(X=x), do(Z=z), w] = \mathbb{P}[y \mid do(X=x), w] \text{ if } (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X},\overline{Z(W)}}}$$

Calculate P[Y | do(X)] from observational data?

Rule 1 doesn't apply

Rule 2 doesn't apply

Rule 3 doesn't apply

$$Pr[y \mid do(x)] = \sum_z P[y \mid do(x), z] p[z \mid do(x)]$$

$$= \sum_z P[y \mid do(x), z] p[z \mid x] \qquad \text{Rule 2:} (X \perp Z)_{G_{\underline{X}}}$$

$$= \sum_z P[y \mid do(x), do(z)] p[z \mid x] \qquad \text{Rule 2:} (Y \perp Z \mid X)_{G_{\overline{X},\underline{Z}}}$$

$$= \sum_z P[y \mid do(z)] p[z \mid x] \qquad \text{Rule 3:} (Y \perp X \mid Z)_{G_{\overline{Z},\overline{X}}}$$

$$= \sum_z \left(\sum_{x'} P[y \mid x', do(z)] p[x' \mid do(z)] \right) p[z \mid x]$$

$$= \sum_z \left(\sum_{x'} P[y \mid x', z] p[x' z] \right) p[z \mid x] \qquad \text{Rule 2+3}$$

A Theory of Measurement

Features… what are features?

Let’s start with an example.

Height.

A property of a person (that is observable)

A variety of tools one can use to measure

A *construct,* operationalized as the result of using a tool to measure it

Need to make decisions about whether height includes hair, sitting v standing

Every tool will be imperfect in measuring height.

and these measurement errors may not be mean 0, may be correlated with demographics.

Features… what are features?

Another example.

Socioeconomic Status (SES)

Encompassing social + economical positioning in the world

A property of a person (that is **un**observable)

One instead **needs to infer it from other properties/constructs.**

For example, using 1 or more proxies (income, education, …)

and then defining a mapping between these and SES.

These involve making assumptions and introducing errors.

A *construct,* operationalized based on how one approximates it ^

Many examples

Constructs	Operationalizations
Teacher Effectiveness	f(test scores, last year's test scores,…)
Risk of Recidivism	f(criminal history, employment, drug use, …)
Patient Benefit	f(health records, age, weight, , …)

Constructs

A construct and its operationalization

are abstract

may or may not be reliable (are they repeatable?)
may or may not be valid (are they right?)

- Contestedness
  - does it have multiple (conflictling) theoretical understandings?
- Substantive validity
  - is the operationalization + measurement good enough?
- Structural validity
  - what operations can be done + are meaningful on these operationalizations?
- Convergent validity
  - Does the operationalization correlate strongly with other operationalizations?
- Discriminant validity
  - Does the measurement correlate with another construct w/o the constructs having similar correlations?
- Predictive validity
  - Is the operationalization/measurement predictive of other constructs or observable properties?
- Consequential validity
  - what are the consequences of the measurement + operationalization?

Fairness as a construct

Fairness itself is some abstract notion

in different contexts, interpreted differently
A ***contested*** construct!

Many many many operationalizations
None of which is appropriate all the time.

Equal opportunity is a state of fairness in which individuals are treated similarly, unhampered by artificial barriers, prejudices, or preferences, except when particular distinctions can be explicitly justified.
— One interpretation of fairness, which itself may be operationalized in many different ways

Individual versus group fairness aren't just different mathematical operationalizations; they're different theoretical understandings of fairness.

Choosing amongst these makes choices about values, not simply about mathematics.