CAUSAL DISCOVERY
Beware of the DAG!

Philip Dawid
University of Cambridge

Seeing and Doing

- Causality is about the effects of *interventions*
- To discover these we really should *experiment*
- If we can’t, is there anything sensible we can conclude from observational data?
- No amount of clever analysis of observational data can replace experimentation

Seeing

- Association
  - Describe stochastic dependence and independence
- Conditional Independence
  - We have a formal algebraic theory
    - Semi-graphoid
    - Separoid

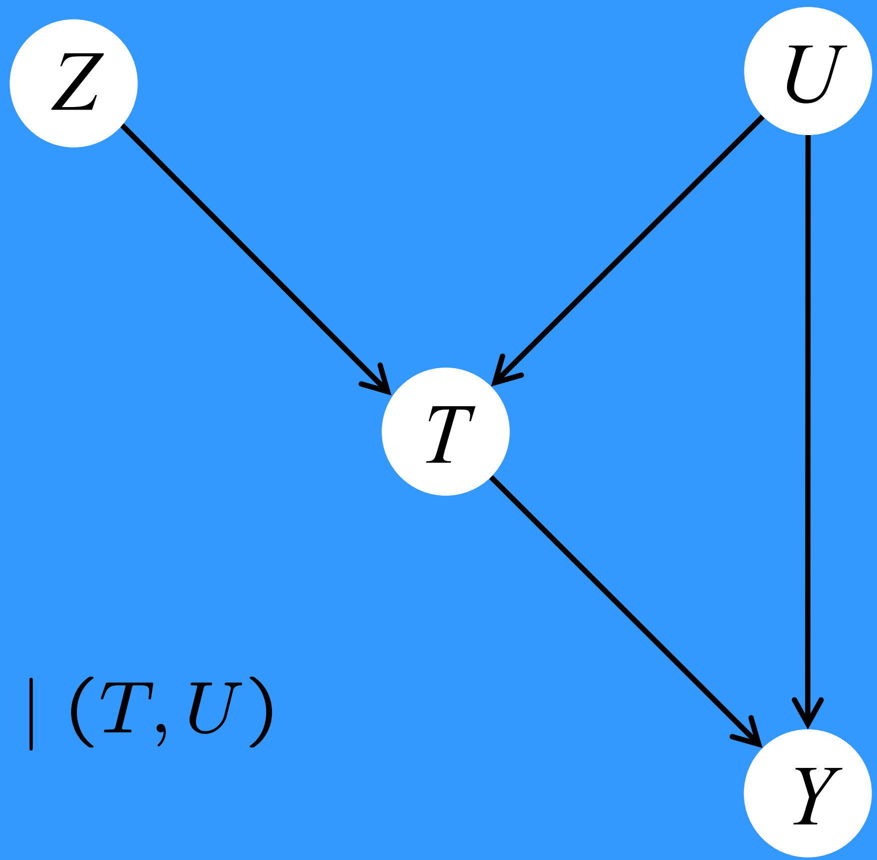
Properties of CI

$$
\begin{array}{lcl}
X \perp\!\!\!\perp Y \mid Z & \Rightarrow & Y \perp\!\!\!\perp X \mid Z \\
X \perp\!\!\!\perp Y \mid X & & \\
X \perp\!\!\!\perp Y \mid Z, \quad W \leq Y & \Rightarrow & X \perp\!\!\!\perp W \mid Z \\
X \perp\!\!\!\perp Y \mid Z, \quad W \leq Y & \Rightarrow & X \perp\!\!\!\perp Y \mid (W, Z) \\
\left.\begin{array}{l} X \perp\!\!\!\perp Y \mid Z \\ \text{and} \\ X \perp\!\!\!\perp W \mid (Y, Z) \end{array}\right\} & \Rightarrow & X \perp\!\!\!\perp (Y, W) \mid Z.
\end{array}
$$

Graphical Representation

- Certain collections of CI properties can be described and manipulated using a DAG
- A probabilistic CI property corresponds to a graphical separation property
  - d-separation
  - moralization
- That's it!

Example



$$U \perp\!\!\!\perp Z$$

$$Y \perp\!\!\!\perp Z \mid (T, U)$$

Points to Remember

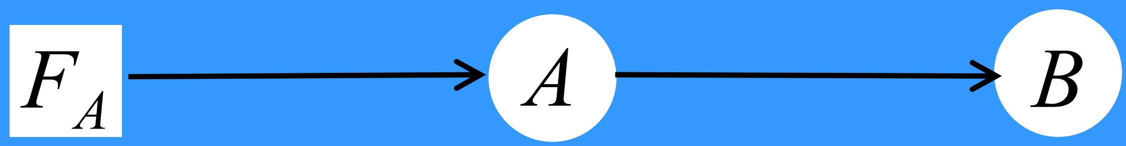
- The graph is *nothing but* an indirect way of describing the CI relationships
  - *cf.* regression
- Clear semantics of this description
- May be several alternative representations (or none)
- Arrows have no intrinsic meaning
  - CI is non-directional!
- Represented relationships unaffected by others unmentioned

Doing

Augmented DAG

with intervention indicators

Explicit causal semantics



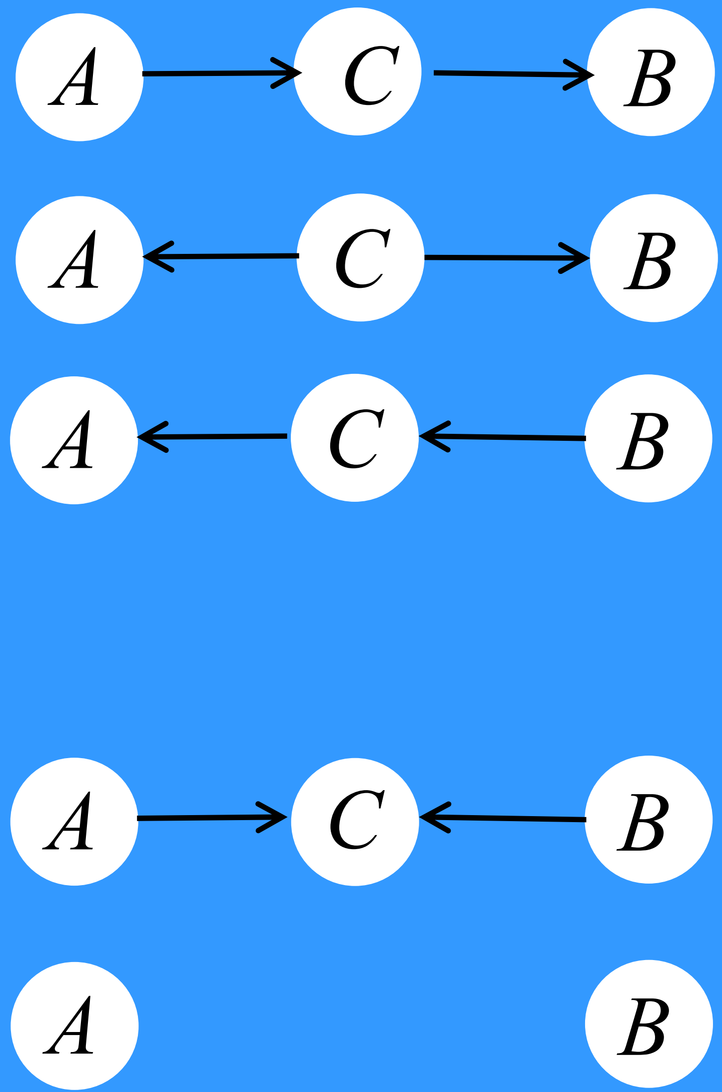
$$B \perp\!\!\!\perp F_A \mid A$$

“Reification”

In an *associational* DAG:

- (Some) arrows represent direction of influence, direct cause,…
- (Some) directed paths represent causal pathways”
- If these exist in all equivalent DAG representations,
  - or if they can be described in terms of additive noise

  they are truly causal



$$A \perp\!\!\!\perp B \mid C$$

$$A \perp\!\!\!\perp B \mid C$$

$$A \perp\!\!\!\perp B \mid C$$

$$A \perp\!\!\!\perp B$$

$$A \perp\!\!\!\perp B$$

With intervention indicators

$$F_A \rightarrow A \rightarrow C \rightarrow B \qquad \left\{ \begin{array}{llll} C & \perp\!\!\!\perp & F_A & \mid \; A \\ B & \perp\!\!\!\perp & (A, F_A) & \mid \; C \end{array} \right.$$

$$F_A \rightarrow A \leftarrow C \rightarrow B \qquad \left\{ \begin{array}{llll} C & \perp\!\!\!\perp & F_A & \\ B & \perp\!\!\!\perp & (A, F_A) & \mid \; C \end{array} \right.$$

$$F_A \rightarrow A \leftarrow C \leftarrow B \qquad \left\{ \begin{array}{llll} (B, C) & \perp\!\!\!\perp & F_A & \\ A & \perp\!\!\!\perp & B & \mid \; (F_A, C) \end{array} \right.$$

$$A \perp\!\!\!\perp B \mid (F_A, C)$$

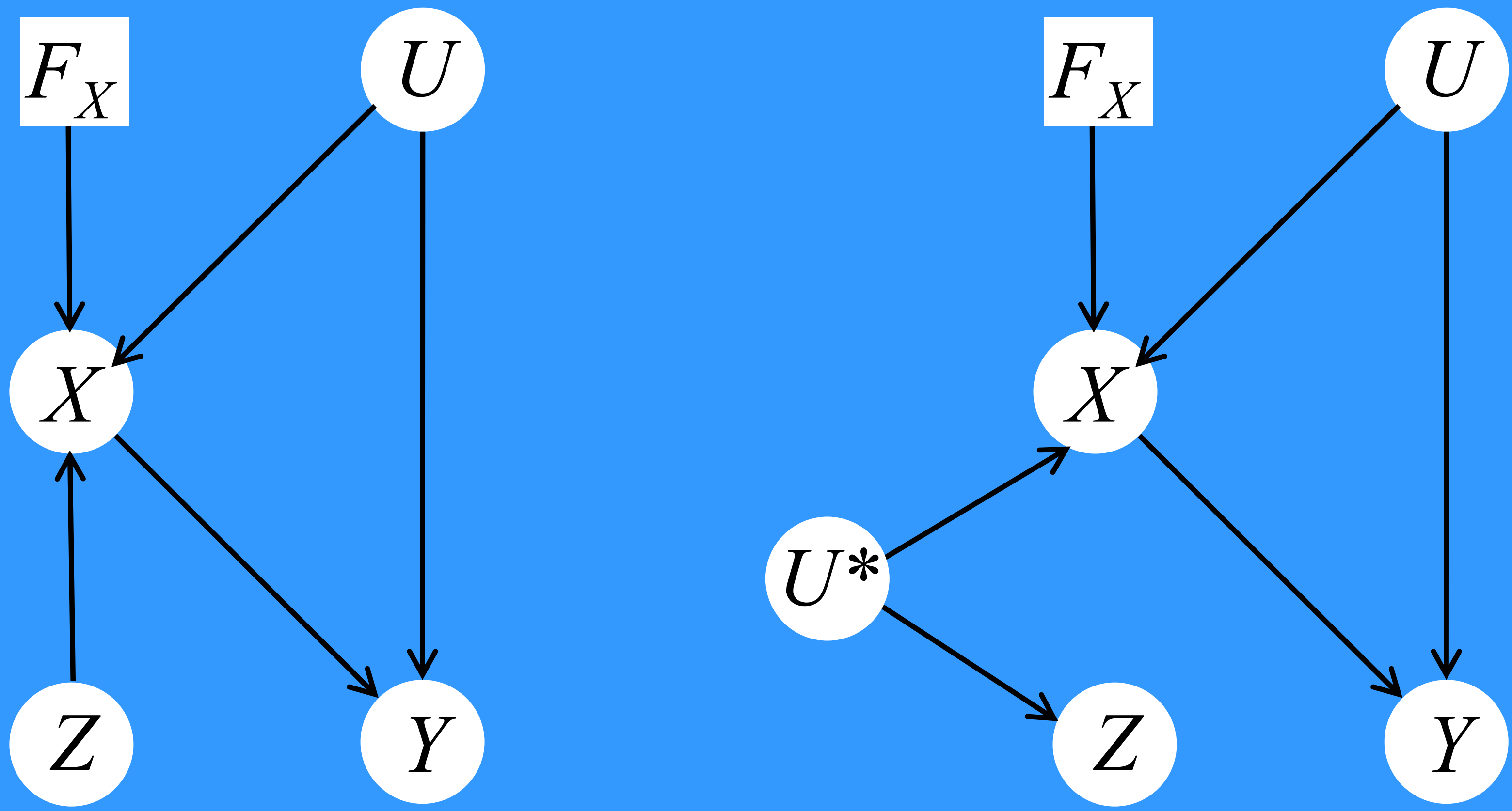
$$F_A \rightarrow A \rightarrow C \leftarrow B \qquad \left\{ \begin{array}{llll} C & \perp\!\!\!\perp & F_A & \mid \; (A, B) \\ B & \perp\!\!\!\perp & (A, F_A) & \end{array} \right.$$

Intuition and Formality

Hernan and Robins (2006):

A *causal DAG* is a DAG in which:

1) the lack of an arrow from V_j to V_m can be interpreted as the absence of a direct causal effect of V_j on V_m (relative to the other variables on the graph)
2) all common causes, even if unmeasured, of any pair of variables on the graph are themselves on the graph. In Figure 2 the inclusion of the measured variables **(*Z*, *X*, *Y*)** implies that the causal DAG must also include their unmeasured common causes **(*U*, *U**)**.



$$\perp\!\!\!\perp \{U, Z, F_X\}$$

$$Y \perp\!\!\!\perp (Z, F_X) | (U, X)$$

When can we just add intervention variables?

- Behaviour of system when kicked need not bear any relationship to its behaviour when observed
- If $\quad A \perp\!\!\!\perp B \quad (A \perp\!\!\!\perp B \mid \text{ancestors}),$

  on adding interventions, neither of A nor B can cause the other (weak causal Markov property??)
  - why need this be?

A way ahead?

- Obtain interventional as well as observational data
- Seek conditional independences involving interventions as well as observations
- Use to build augmented DAG
- Genuine causal interpretation