Identifying Causal Effects via Context-specific Independence Relations

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Santtu Tikka
Department of Mathematics and Statistics
University of Jyväskylä, Finland
santtu.tikka@jyu.fi

Antti Hyttinen
HIIT, Department of Computer Science
University of Helsinki, Finland
antti.hyttinen@helsinki.fi

Juha Karvanen
Department of Mathematics and Statistics
University of Jyväskylä, Finland
juha.t.karvanen@jyu.fi

Abstract

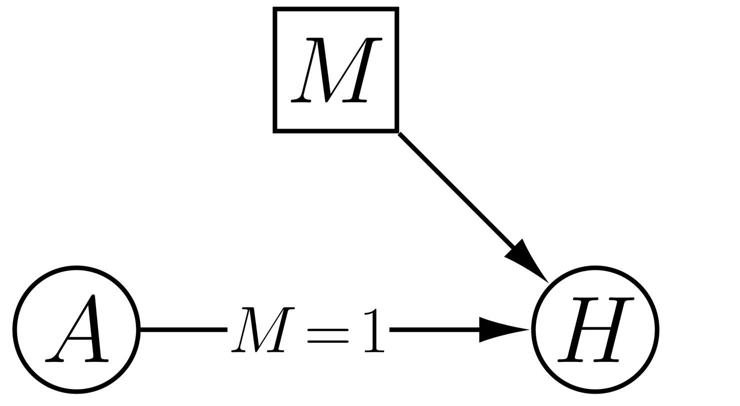
- Causal effect identification considers whether an interventional probability distribution can be uniquely determined from a passively observed distribution in a given causal structure.
- If the generating system induces context-specific independence (CSI) relations, the existing identification procedures and criteria based on do-calculus are inherently incomplete.
- We design a calculus and an automated search procedure for identifying causal effects in the presence of CSIs.
- We demonstrate that a small number of CSI-relations may be sufficient to turn a previously non-identifiable instance to identifiable.

Context-specific Independence and Labeled DAGs

CSIs arise from local causal mechanisms. For a single example consider an antibiotic (A) that normally has a dose–response effect on the number of bacteria (H). A genetic mutation (M) makes the bacteria resistant to the antibiotic meaning that in the context of this mutation the dose and the number of bacteria are independent:

$$P(H \mid A, M=1) = P(H \mid M=1) \quad \text{denoted as} \quad H \perp\!\!\!\perp A \mid M=1$$
$$P(H \mid A, M=0) \neq P(H \mid M=0) \quad \text{(possibly)}$$

Labeled Directed Acyclic Graphs (LDAG) offer a simple and intuitive way to represent the causal structure and local CSIs. Interventional distribution $P(Y \mid \mathrm{do}(X))$ can be modeled as $P(Y \mid X, I_X = 1)$, where I_X is an intervention node, hence do-operation is not needed explicitly.



Notation

- Label ℓ on $X \rightarrow Y$ is an assignment to $L = pa(Y) \setminus X$ (other parents of Y) encoding a local CSI $X \perp\!\!\!\perp Y \mid L = \ell$
- Symbol * denotes any assignment
- Square nodes are unobserved

In the example case, we mark label $M = 1$ on the edge $A \rightarrow H$ to show that antibiotic A can have a dose-response effect on the number of bacteria H only if a genetic mutation M has not taken place.

Problem Formulation

Input: An LDAG G over a set of nodes $\boldsymbol{V}$, a probability distribution $P(\boldsymbol{W})$ for $\boldsymbol{W} \subset \boldsymbol{V}$ and a query $P(\boldsymbol{Y} \mid \mathrm{do}(\boldsymbol{X}), \boldsymbol{Z})$ such that $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \subset \boldsymbol{V}$.

Task: Output a formula for $P(\boldsymbol{Y} \mid \mathrm{do}(\boldsymbol{X}), \boldsymbol{Z})$ over $P(\boldsymbol{W})$ or decide that it is non-identifiable.

Computational Complexity

Theorem 1. Deciding non-identifiability of a causal effect in an LDAG over $\boldsymbol{V}$ and a passively observed distribution over $\boldsymbol{W} \subset \boldsymbol{V}$ is NP-hard.

CSI-calculus for Determining Identifiability

Rule 1 (Insertion/Deletion of observations):

$$P(\boldsymbol{Y}_1, \boldsymbol{y}_2 \mid \boldsymbol{Z}_1, \boldsymbol{z}_2, \boldsymbol{X}_1, \boldsymbol{x}_2) = P(\boldsymbol{Y}_1, \boldsymbol{y}_2 \mid \boldsymbol{X}_1, \boldsymbol{x}_2) \text{ if } \boldsymbol{Y}_1, \boldsymbol{Y}_2 \perp\!\!\!\perp \boldsymbol{Z}_1, \boldsymbol{Z}_2 \mid \boldsymbol{X}_1, \boldsymbol{x}_2$$

Rule 2 (Marginalization/Sum-rule):

$$P(\boldsymbol{Y}_1, \boldsymbol{y}_2 \mid \boldsymbol{X}_1, \boldsymbol{x}_2) = \textstyle\sum_{\boldsymbol{Z}} P(\boldsymbol{Y}_1, \boldsymbol{y}_2, \boldsymbol{Z} \mid \boldsymbol{X}_1, \boldsymbol{x}_2)$$

Rule 3 (Conditioning):

$$P(\boldsymbol{Y}_1 \mid \boldsymbol{Z}_1, \boldsymbol{z}_2, \boldsymbol{X}_1, \boldsymbol{x}_2) = \frac{P(\boldsymbol{Y}_1, \boldsymbol{Z}_1, \boldsymbol{z}_2 \mid \boldsymbol{X}_1, \boldsymbol{x}_2)}{\sum_{\boldsymbol{Y}_1} P(\boldsymbol{Y}_1, \boldsymbol{Z}_1, \boldsymbol{z}_2 \mid \boldsymbol{X}_1, \boldsymbol{x}_2)}$$

Rule 4 (Product-rule):

$$P(\boldsymbol{Y}_1, \boldsymbol{y}_2, \boldsymbol{Z}_1, \boldsymbol{z}_2 \mid \boldsymbol{X}_1, \boldsymbol{x}_2) = P(\boldsymbol{Y}_1, \boldsymbol{y}_2 \mid \boldsymbol{Z}_1, \boldsymbol{z}_2, \boldsymbol{X}_1, \boldsymbol{x}_2) P(\boldsymbol{Z}_1, \boldsymbol{z}_2 \mid \boldsymbol{X}_1, \boldsymbol{x}_2)$$

Rule 5 (General-by-case reasoning):

$$P(\boldsymbol{Y}_1, \boldsymbol{y}_2, 1-z \mid \boldsymbol{X}_1, \boldsymbol{x}_2) = P(\boldsymbol{Y}_1, \boldsymbol{y}_2 \mid \boldsymbol{X}_1, \boldsymbol{x}_2) - P(\boldsymbol{Y}_1, \boldsymbol{y}_2, z \mid \boldsymbol{X}_1, \boldsymbol{x}_2)$$

Rule 6 (Case-by-case reasoning):

$$P(\boldsymbol{Y}_1, \boldsymbol{y}_2, Z \mid \boldsymbol{X}_1, \boldsymbol{x}_2) = \begin{cases} P(\boldsymbol{Y}_1, \boldsymbol{y}_2, Z = 0 \mid \boldsymbol{X}_1, \boldsymbol{x}_2) & \text{if } Z = 0 \\ P(\boldsymbol{Y}_1, \boldsymbol{y}_2, Z = 1 \mid \boldsymbol{X}_1, \boldsymbol{x}_2) & \text{if } Z = 1 \end{cases}$$

Rule 7 (Case-by-general reasoning (a)):

$$P(\boldsymbol{Y}_1, \boldsymbol{y}_2, z \mid \boldsymbol{X}_1, \boldsymbol{x}_2) = P(\boldsymbol{Y}_1, \boldsymbol{y}_2, Z \mid \boldsymbol{X}_1, \boldsymbol{x}_2)\big|_{Z=z}$$

Rule 8 (Case-by-general reasoning (b)):

$$P(\boldsymbol{Y}_1, \boldsymbol{y}_2 \mid \boldsymbol{X}_1, \boldsymbol{x}_2, z) = P(\boldsymbol{Y}_1, \boldsymbol{y}_2 \mid \boldsymbol{X}_1, \boldsymbol{x}_2, Z)\big|_{Z=z}$$

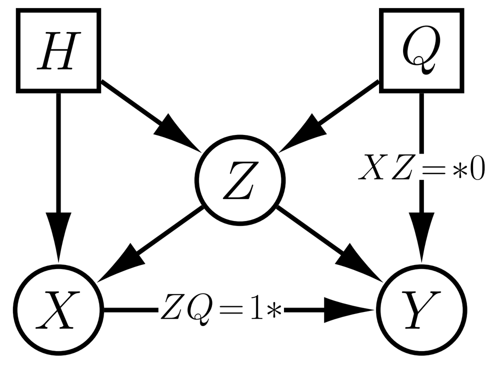
Theorem 2. Do-calculus is a special case of CSI-calculus.

A Search over the Rules of CSI-calculus

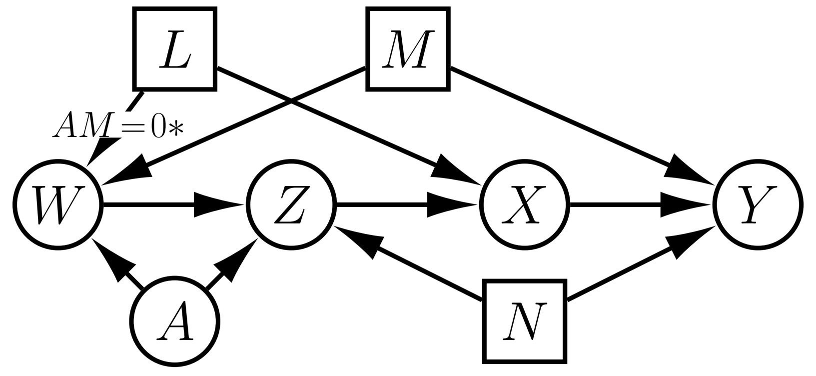
A forward search algorithm over the rules of CSI-calculus is used to obtain identifying formulas and derivations. Starting from the input $P(\boldsymbol{W})$, only identifiable terms are derived. During the search, a novel separation criterion is applied to verify non-local CSIs. Contexts are combined and redundant contexts are eliminated to avoid unnecessary computation.

New Identifiability Results Unobtainable via do-calculus

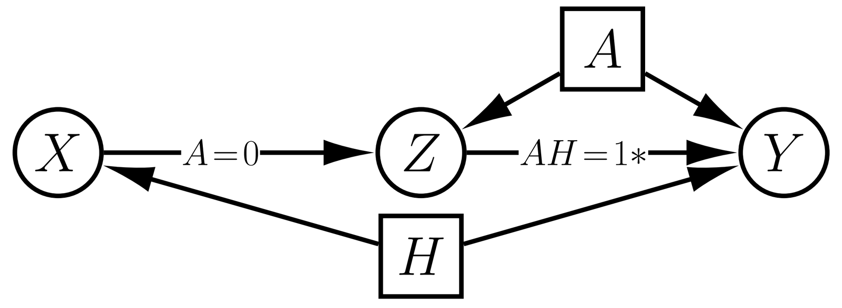
With the search we are able to identify $P(Y \mid \mathrm{do}(X))$ in the following LDAGs.



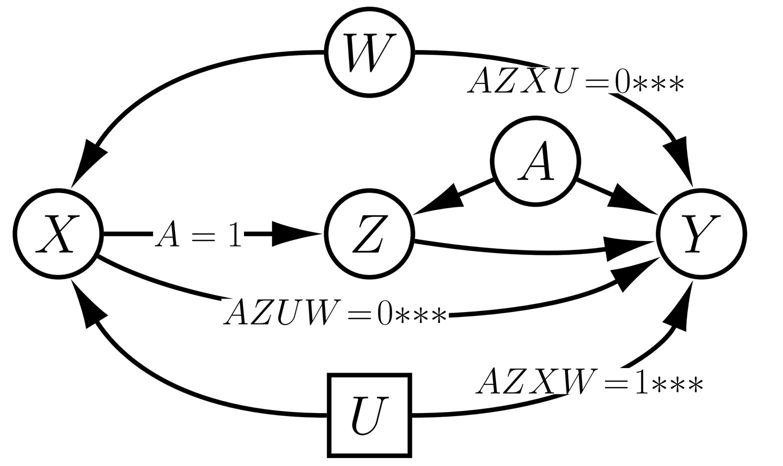
$$P(Y \mid \mathrm{do}(X)) =$$
$$P(Y \mid Z=0, X) P(Z=0) + P(Y \mid Z=1) P(Z=1)$$



$$P(Y \mid \mathrm{do}(X)) =$$
$$\textstyle\sum_Z P(Z \mid A=0) P(Y \mid X, Z, A)$$



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



$$P(Y \mid \mathrm{do}(X)) =$$
$$P(A=1) \textstyle\sum_W P(Y \mid X, W, A=1) P(W \mid A=1)$$
$$+ P(A=0) \left(\textstyle\sum_Z P(Z \mid X, A=0) \right.$$
$$\left. \textstyle\sum_{X'} P(Y \mid X', Z, A=0) P(X' \mid A=0) \right)$$