Seoul National University
Graduate School of Data Science

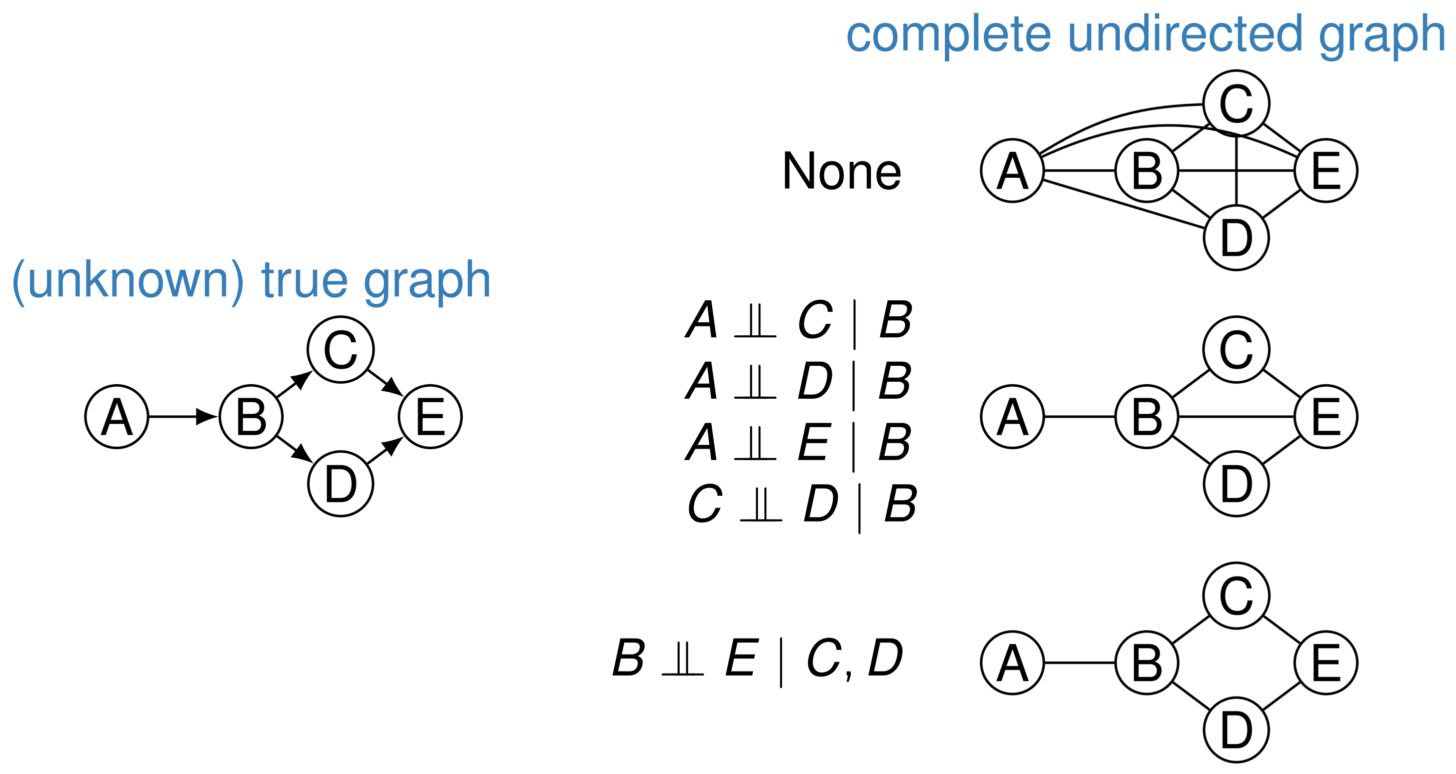
Causal Discovery with Deductive Reasoning: One Less Problem

Jonghwan Kim, Inwoo Hwang, Sanghack Lee

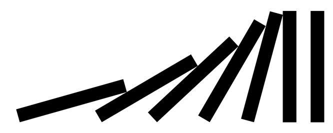
Causality Lab.

Motivation & Background

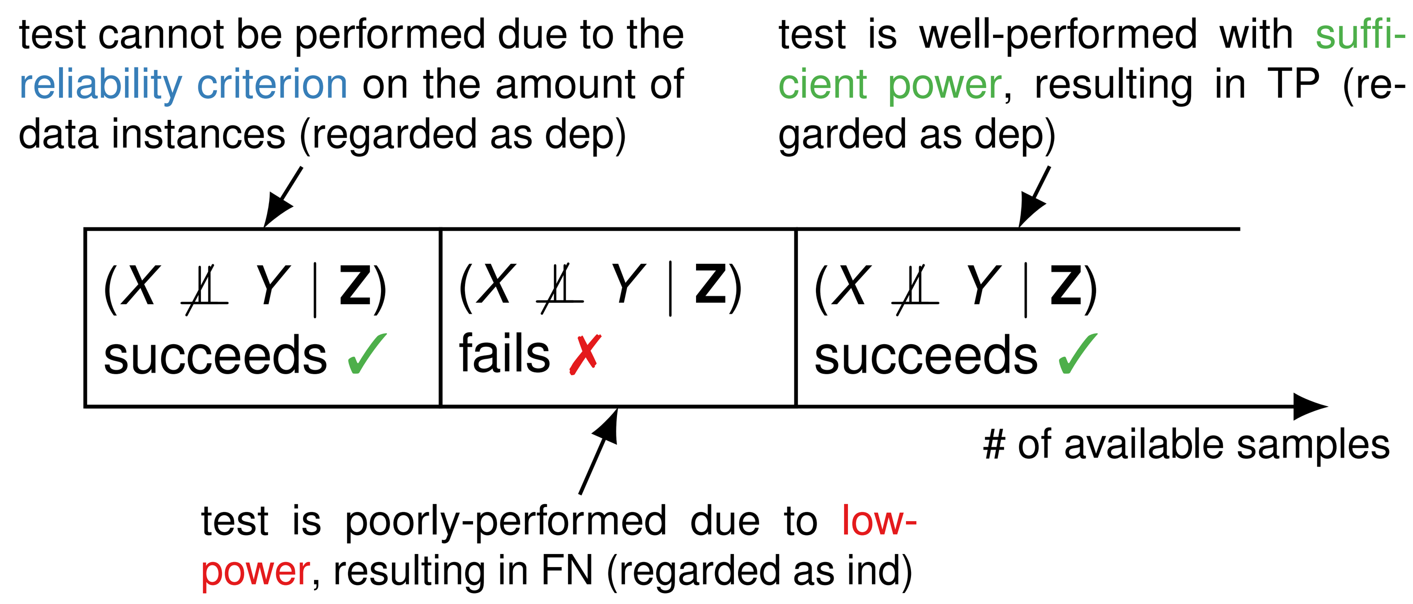
- Constraint-based causal discovery utilizes multiple **conditional independence tests** (CITs) to induce underlying causal structure of data.
- Starting from a complete undirected graph, PC algorithm removes edges based on CIT results accordingly.



- However, high-order CITs with low power often yield false negatives, propagating errors throughout the structure learning process.



- Prior works have tackled this issue with either **simple heuristics** or **complicated routines** with heavy computational burden.



Deductive Reasoning for Causal Discovery

Q. How can we properly correct unreliable CIT results for robust structure learning?

A. Utilize relationships with other CIT results for correcting unreliable CIT result!

Ingredients: Graphoid Axioms

- Graphoid axioms (Pearl and Paz, 1987) can be used to constrain CI statements by other CI statements.
- Under the faithfulness, we have more relaxed rules derived from graphoid axioms as follows: (*selected*)

Symmetry: $(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \iff (\mathbf{Y} \perp\!\!\!\perp \mathbf{X} \mid \mathbf{Z})$

Decomposition: $(\mathbf{X} \perp\!\!\!\perp \mathbf{Y}, \mathbf{W} \mid \mathbf{Z}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \land (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z})$

Contraction: $(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \land (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z}, \mathbf{Y}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y}, \mathbf{W} \mid \mathbf{Z})$

Weak Transitivity: $(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \land (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}, W) \implies (\mathbf{X} \perp\!\!\!\perp W \mid \mathbf{Z}) \lor (W \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z})$

Deducing Dependence

We provide a condition for deducing higher-order dependence statement with lower-order CI statements.

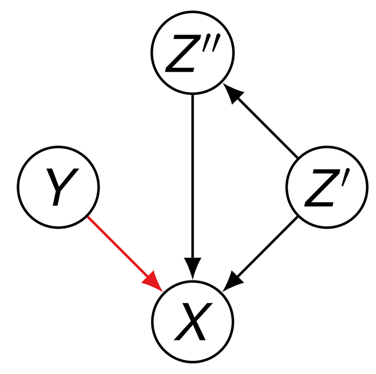
Theorem Deduce-Dep

Under the faithful Bayesian network $(\mathcal{G}, P)$, let $\{X\} \sqcup \{Y\} \sqcup \mathbf{Z} \subseteq \mathbf{V}$, $Z \in \mathbf{Z}$, and $\mathbf{Z}' = \mathbf{Z} \setminus \{Z\}$.
If

$$(X \not\perp\!\!\!\perp Y \mid \mathbf{Z}') \oplus \left((X \not\perp\!\!\!\perp Z \mid \mathbf{Z}') \land (Y \not\perp\!\!\!\perp Z \mid \mathbf{Z}')\right),$$

then $(X \not\perp\!\!\!\perp Y \mid \mathbf{Z})$ **holds**.

Illustrative Example



CD algorithm tries to discover the left. It tries to examine $X - Y$ where the following are accurately obtained:

$(X \not\perp\!\!\!\perp Y \mid Z')$ and $(X \not\perp\!\!\!\perp Y \mid Z'')$.

Unfortunately, the relationship between X and Y is relatively weak, and we wrongly obtained:

$(X \perp\!\!\!\perp Y \mid Z', Z'')$,

We may have a doubt about the CIT result, worrying about its power being low. Then, we may examine the following CI between Y and Z'' given Z', where the CIT correctly yields $(Y \perp\!\!\!\perp Z'' \mid Z')$. In such case, we can indeed induce $(X \not\perp\!\!\!\perp Y \mid Z', Z'')$ from the two existing CIT results and $(Y \perp\!\!\!\perp Z'' \mid Z')$ via applying rules derived from graphoid axioms.

Algorithm: Deduce-Dep

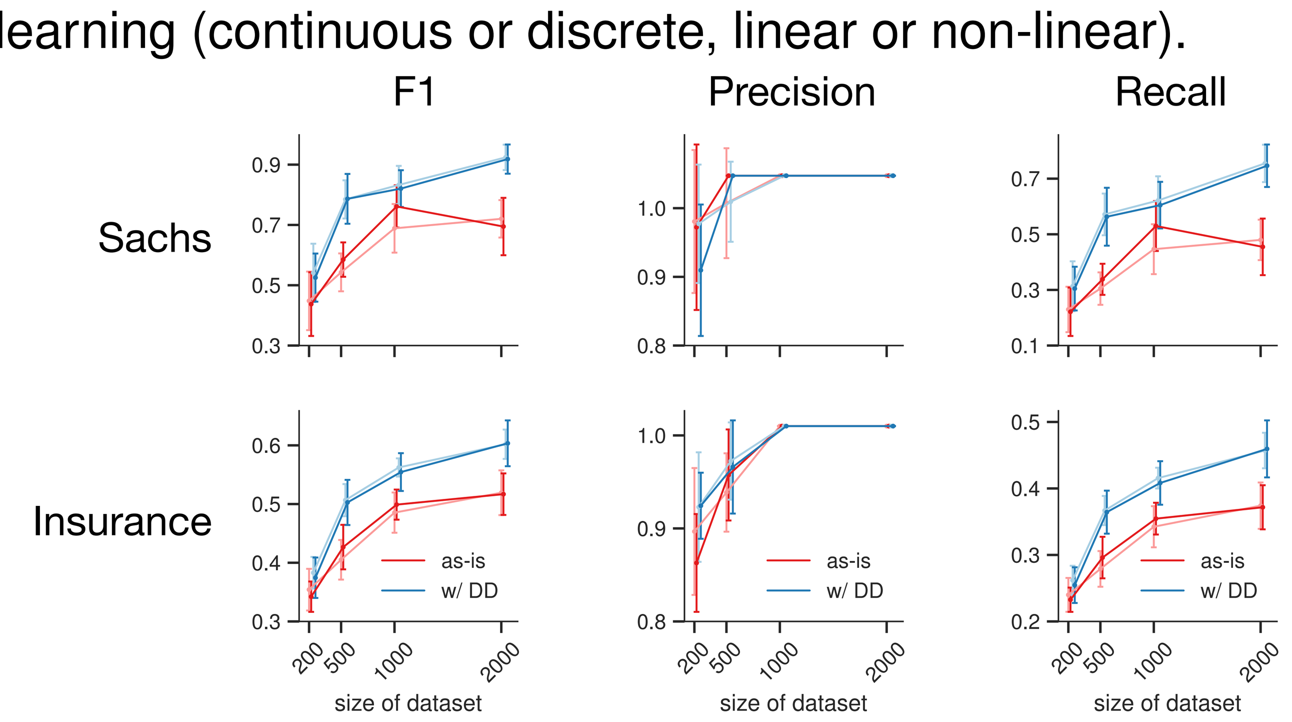
- Our method replaces unreliable CIT results with deductively reasoned results from lower-order CITs, which are deemed more *reliable*.
- Our method can be effortlessly plugged into any constraint-based structure learning algorithm.

Input: $\{X\}, \{Y\}, \mathbf{Z}$ disjoint subsets of $\mathbf{V}$, reliability threshold K (default 1)
Output: Whether $(X \not\perp\!\!\!\perp Y \mid \mathbf{Z})$ is deducible or not.
if $|\mathbf{Z}| \leq K$ **return** FALSE
for $Z \in \mathbf{Z}$
  $\mathbf{Z}' \leftarrow \mathbf{Z} \setminus \{Z\}$
  for $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ in $\{(X, Y, \mathbf{Z}'), (X, Z, \mathbf{Z}'), (Y, Z, \mathbf{Z}')\}$
    if $(\mathbf{A} \perp\!\!\!\perp \mathbf{B} \mid \mathbf{C})$ **and not** DEDUCE-DEP$(\mathbf{A}, \mathbf{B}, \mathbf{C})$
      mark $(\mathbf{A}; \mathbf{B} \mid \mathbf{C})$ as $\perp\!\!\!\perp$
    else mark $(\mathbf{A}; \mathbf{B} \mid \mathbf{C})$ as $\not\perp\!\!\!\perp$
  if $(X \not\perp\!\!\!\perp Y \mid \mathbf{Z}') \oplus \left((X \not\perp\!\!\!\perp Z \mid \mathbf{Z}') \land (Y \not\perp\!\!\!\perp Z \mid \mathbf{Z}')\right)$
    return TRUE
return FALSE

Input: a set of variables $\mathbf{V}$, and CI tester
Output: a CPDAG
Initialize $\mathcal{G}$ with a complete undirected graph
for $k \in 1, 2, \ldots,$
  for an ordered pair of adjacent vertices $(X, Y) \in \mathcal{G}$ s.t. $|Ne(\{X\})_{\mathcal{G}} \setminus \{Y\}| \geq k$
    for $\mathbf{S} \subseteq Ne(\{X\})_{\mathcal{G}} \setminus \{Y\}$ s.t. $|\mathbf{S}| = k$
      if $(X \perp\!\!\!\perp Y \mid \mathbf{S})$
        if not DEDUCE-DEP$(\{X\}, \{Y\}, \mathbf{S})$
          Remove X-Y from $\mathcal{G}$
  else break
Orient $\mathcal{G}$ for unshielded colliders
Complete orientation of $\mathcal{G}$ with Meek's rules
return $\mathcal{G}$

Experimental Results

Under data-scarce scenarios, our method improves the performance of structure learning (continuous or discrete, linear or non-linear).



Conclusion

Contribution

- We proposed a **practical correction method** for unreliable CITs by leveraging rules derived from graphoid axioms.
- Our method can be effortlessly adapted to any constraint-based structure learning algorithm.
- Empirical evaluation reveals that our method properly corrects the unreliable CITs, improving the performance of structure learning.

Future Work

- Combining our method with false positive control methods might ensure a more robust causal structure learning.