Local Identifying Causal Relations in the Presence of Latent Variables

Zheng Li[*1], Zeyu Liu[*1], Feng Xie[1], Hao Zhang[2], Chunchen Liu[3], Zhi Geng[1]

*Equal contribution

1. Beijing Technology and Business University

2. SIAT, Chinese Academy of Sciences

3. LingYang Co.Ltd, Alibaba Group

Correspondence to: Zheng Li <zhengli0060@gmail.com>, Feng Xie <xiefeng009@gmail.com>

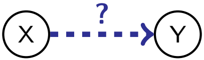
June 26, 2025

Contents

Motivation

Table: Observed Dataset

	V_1	...	X	...	Y	...	V_{m-1}	V_m
1	$v_{1,1}$	...	x_1	...	y_1	...	$v_{1,m-1}$	$v_{1,m}$
2	$v_{2,1}$	...	x_2	...	y_2	...	$v_{2,m-1}$	$v_{2,m}$
...	...	...	...	...	...	...	...	...
n	$v_{n,1}$	...	x_n	...	y_n	...	$v_{n,m-1}$	$v_{n,m}$



Is X an ancestor of Y?

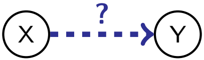
Problem

How to identify the causal relationship between X and Y in the presence of latent variables?

Motivation

Table: Observed Dataset

	V_1	$\dots$	X	$\dots$	Y	$\dots$	V_{m-1}	V_m
1	$v_{1,1}$	...	x_1	...	y_1	...	$v_{1,m-1}$	$v_{1,m}$
2	$v_{2,1}$	...	x_2	...	y_2	...	$v_{2,m-1}$	$v_{2,m}$
...	...	...	...	...	...	...	...	...
n	$v_{n,1}$	...	x_n	...	y_n	...	$v_{n,m-1}$	$v_{n,m}$



Is X an ancestor of Y?

Problem

How to identify the causal relationship between X and Y in the presence of latent variables?

Learning global causal structure? (FCI, RFCI, etc.)

The complexity of these methods grows exponentially

with respect to the number of all observed variables.

Goal

Table: Observed Dataset

	V_1	...	X	...	Y	...	V_{m-1}	V_m
1	$v_{1,1}$	...	x_1	...	y_1	...	$v_{1,m-1}$	$v_{1,m}$
2	$v_{2,1}$	...	x_2	...	y_2	...	$v_{2,m-1}$	$v_{2,m}$
...	...	...	...	...	...	...	...	...
n	$v_{n,1}$	...	x_n	...	y_n	...	$v_{n,m-1}$	$v_{n,m}$

Is X an ancestor of Y?

Task: Under the standard assumptions of the causal Markov condition, the causal Faithfulness condition, and no selection bias, our objective is to

- **Characterize** the local graphical features of different types of causal relationships between a pair of variables X and Y.
- **Develop** a fully local algorithm to determine the causal relationship between X and Y.

Preliminaries

Given a Markov equivalent class of MAGs, there are three possible types of causal relationships (Zhang 2006):

- A variable X is an **invariant ancestor** of a variable Y if there is a directed path from X to Y in every equivalent MAG.
- A variable X is an **invariant non-ancestor** of a variable Y if there is no directed path from X to Y in any equivalent MAG.
- A variable X is a **possible ancestor** of variable Y if X is neither an invariant ancestor nor an invariant non-ancestor of Y.

Markov Blanket

- The *Markov blanket* of a variable X is the smallest set conditioned on which all other variables are statistically independent of X.
- Graphically, in a DAG $\mathcal{D}$, this is the set of parents, children, and children's parents of vertex X, denoted by $MB(X, \mathcal{D})$.
- In a MAG, the *Markov blanket* of a vertex X, noted as $MB(X, \mathcal{M})$, comprises: (1) vertices adjacent to X, and (2) vertices reachable from X via collider paths despite being non-adjacent.

Markov Blanket

In a MAG, the *Markov blanket* of a vertex X, noted as $MB(X, \mathcal{M})$, comprises:

- (1) vertices adjacent to X, and
- (2) vertices reachable from X via collider paths despite being non-adjacent.

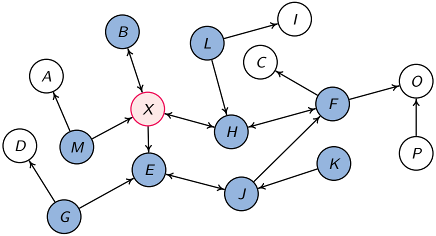


Figure: The illustrative example for MB, where X is the target of interest and the blue vertices belong to $MB(X, \mathcal{M})$.

How can we characterise the local features of causal relationships?

The well-known **local Markov property for DAGs** states that, given a target variable X, it is independent of its non-descendants conditioned on *a particular set* of variables, i.e. , its *parents*. Formally,

$$X \perp\!\!\!\perp \{\mathbf{V} \setminus De(X)\} \mid Pa(X).$$

However, the parents of the target variable may not be observed, in a system where latent variables may exist.

Define a local Markov property for MAGs

Definition 1 (Arrow-Collider Path)

In a PAG or a MAG, a path $\pi = \langle V_0, \ldots, V_n \rangle$ is called an arrow-collider path from V_0 to V_n if every non-endpoint vertex is a collider on π, and the edge between V_0 and V_1 points into V_0, i.e. , $V_0 \leftrightarrow V_1 \cdots \leftarrow\!\!* V_n$. If $n = 1$, π simplifies to $V_0 \leftarrow\!\!* V_1$.

Building on Definition 1, we define the *particular set* graphically in Definition 2.

Definition 2 (Augmented Parent Set)

Let $\mathcal{G}$ be a PAG or a MAG. The augmented parent set of a vertex X, denoted as $Pa^*(X, \mathcal{G})$, is defined as follows: for any vertex $V \in \mathbf{O}$, $V \in Pa^*(X, \mathcal{G})$ if and only if there exists an arrow-collider path π from X to V such that:

(1) *in a MAG, X is a non-ancestor of every vertex on π, including V.*

(2) *in a PAG, X is an invariant non-ancestor of all vertices on π, including V.*

Local Markov Property for MAGs

Definition 1 (Local Markov Property for MAGs)

Let $\mathcal{M}$ be the MAG over $\mathbf{O}$, and let $Pre(X, \mathcal{M})$ denote the pre-treatment vertices of X in $\mathcal{M}$, i.e. , the vertices for which X is not an ancestor. The local Markov property for the MAG states that for every variable $X \in \mathbf{O}$, the following property holds:

$$X \perp\!\!\!\perp Pre(X, \mathcal{M}) \setminus Pa^*(X, \mathcal{M}) \mid Pa^*(X, \mathcal{M}) \quad (1)$$

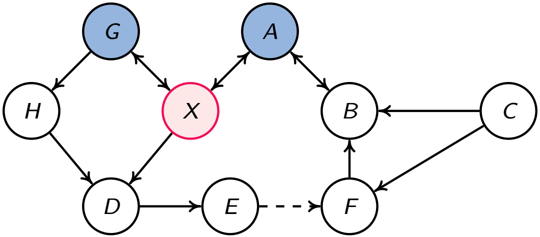


Figure: A MAG

$De(X, \mathcal{M}) = \{D, E, F, B\}$,

$Pre(X, \mathcal{M}) = \{A, C, G, H\}$,

$Pa^*(X, \mathcal{M}) = \{A, G\}$

The local Markov property for X is:

$$X \perp\!\!\!\perp \{C, H\} \mid \{A, G\}$$

Local Characterization: Invariant non-Ancestor

Theorem 1

*Let $\mathcal{P}$ be the PAG over $\mathbf{O}$. For any pair of vertices (X, Y) in $\mathcal{P}$, X is an **invariant non-ancestor** of Y if and only if $X \perp\!\!\!\perp Y \mid Pa^*(X, \mathcal{P})$.*

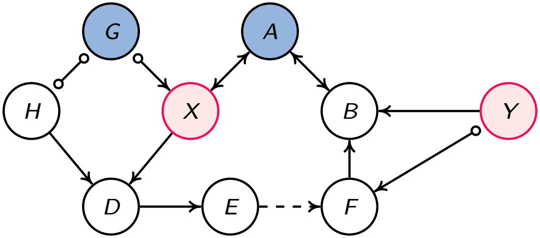


Figure: A PAG

$Pa^*(X, \mathcal{P}) = \{A, G\}$,

$X \perp\!\!\!\perp Y \mid \{A, G\}$

Consequently, X is an invariant non-ancestor of Y.

Intuitively, given a PAG $\mathcal{P}$, $Pa^*(X, \mathcal{P})$ blocks all non-causal paths from X to Y in $\mathcal{P}$.

Local Characterization: Invariant Ancestor

Definition 3 (Explicit Invariant Ancestor)

Given a PAG $\mathcal{P}$, a vertex X is an explicit invariant ancestor of another vertex Y if a common directed path exists from X to Y in every MAG within $[\mathcal{P}]$.

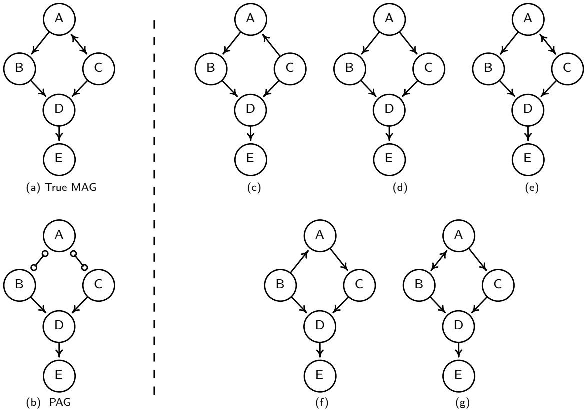
Definition 4 (Implicit Invariant Ancestor)

Given a PAG $\mathcal{P}$, a vertex X is an implicit invariant ancestor of another vertex Y if X is an invariant ancestor of Y, but there is ***no*** *directed path from X to Y common to every MAG within $[\mathcal{P}]$.*

Remark 1

We define X as an invariant ancestor of Y if there exists a directed path from X to Y in every MAG in $[\mathcal{P}]$. If these directed paths are identical across all MAGs in $[\mathcal{P}]$, the invariant ancestor relation is explicit; otherwise, it is implicit.

Explicit/Implicit Invariant Ancestor Examples



(a) True MAG (b) PAG (c) (d) (e) (f) (g)

- **Explicit invariant ancestor**: Let $X = D$ and $Y = E$. A common directed path $D \to E$ is present in every graph within Figures (c-g), thereby establishing D as an explicit invariant ancestor of E.

- **Implicit invariant ancestor**: Let $X = A$ and $Y = D$. There is a directed path from A to D that can be observed in every graph within Figures (c-g), but there is no common directed path from A to D in these graphs. Therefore, A is considered an implicit invariant ancestor of D.

Local Characterization: Explicit Invariant Ancestor

Definition 5 (**Circle-Collider Path**)

In a PAG, a path $\pi = \langle V_0, \ldots, V_n \rangle$ *is called a circle-collider path from* V_0 *to* V_n *if every non-endpoint vertex is a collider on* π*, and the edge between* V_0 *and* V_1 *is undirected relative to* V_0*, i.e. ,* $V_0 \circ\!\!\rightarrow V_1 \cdots \leftarrow\!\!* V_n$*. If* $n = 1$*,* π *simplifies to* $V_0 \circ\!\!-\!\!* V_1$*.*

Definition 6 (**Augmented Non-directed Neighbor Set**)

Let $\mathcal{P}$ *be a PAG, the augmented non-directed neighbor set of a vertex* X*, denoted as* $Ne^*(X, \mathcal{P})$*, is defined as follows: For any vertex* $V \in \mathbf{O}$*,* $V \in Ne^*(X, \mathcal{P})$ *if and only if there exists a circle-collider path* $\pi = \langle X = V_0, V_1, \ldots, V_n = V \rangle$ *from* X *to* V *such that for every* $2 \leq i \leq n$*,* X *is an invariant non-ancestor of* V_i [a]*.*

[a] When $n = 1$, the condition becomes redundant and may be omitted.

Local Characterization: Explicit Invariant Ancestor

Theorem 2

*Let $\mathcal{P}$ be the PAG over $\mathbf{O}$. For any pair of vertices (X, Y) in $\mathcal{P}$, X is an **explicit invariant ancestor** of Y if and only if $X \not\perp\!\!\!\perp Y \mid Pa^*(X,\mathcal{P}) \cup Ne^*(X,\mathcal{P})$.*

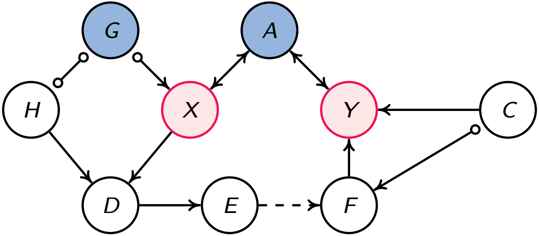


Figure: A PAG

$Pa^*(X,\mathcal{P}) = \{A, G\}$,

$Ne^*(X,\mathcal{P}) = \emptyset$,

$X \not\perp\!\!\!\perp Y \mid \{A, G\}$

X is an explicit invariant ancestor of Y.

Intuitively, given a PAG $\mathcal{P}$, $Ne^*(X,\mathcal{P})$ blocks all partially directed paths(except directed paths) from X to Y in $\mathcal{P}$.

Local Characterization: Implicit Invariant Ancestor

Definition 7

Let $\mathcal{P}$ be a PAG and let $\mathbb{M}$ represent the set of maximal cliques of the induced subgraph of $\mathcal{P}$ over $PossCh(X, \mathcal{P}) \cup NondNe(X, \mathcal{P})$. The set of augmented non-directed neighbor of a vertex X relative to a maximal clique $\mathbf{M} \in \mathbb{M}$, denoted as $Ne^*(X_{\mathbf{M}}, \mathcal{P})$. For any vertex $V \in \mathbf{O}$, $V \in Ne^*(X_{\mathbf{M}}, \mathcal{P})$ if and only if there exists a circle-collider path $\pi = \langle X = V_0, V_1, \ldots, V_n = V \rangle$ from X to V such that (1) for every $2 \leq i \leq n$, X is an invariant non-ancestor of V_i and (2) $V_1 \in \mathbf{M}$.

Local Characterization: Implicit Invariant Ancestor

Theorem 3

*Let $\mathcal{P}$ be the PAG over $\mathbf{O}$. For any pair of vertices (X, Y) in $\mathcal{P}$, X is an **implicit invariant ancestor** of Y if and only if (1) $X \perp\!\!\!\perp Y \mid Pa^*(X, \mathcal{P}) \cup Ne^*(X, \mathcal{P})$, but (2) $X \not\perp\!\!\!\perp Y \mid Pa^*(X, \mathcal{P}) \cup Ne^*(X_{\mathbf{M}}, \mathcal{P})$ for every $\mathbf{M} \in \mathbb{M}$.*

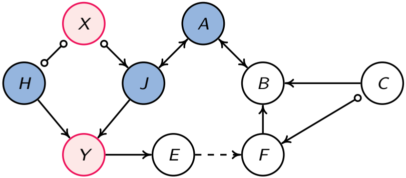


Figure: A PAG

$$Pa^*(X, \mathcal{P}) = \emptyset,$$

$$Ne^*(X, \mathcal{P}) = \{A, J, H\},$$

$$(1) X \perp\!\!\!\perp Y \mid \{A, J, H\}$$

Local Characterization: Implicit Invariant Ancestor

Theorem 3

*Let $\mathcal{P}$ be the PAG over $\mathbf{O}$. For any pair of vertices (X, Y) in $\mathcal{P}$, X is an **implicit invariant ancestor** of Y if and only if (1) $X \perp\!\!\!\perp Y \mid Pa^*(X, \mathcal{P}) \cup Ne^*(X, \mathcal{P})$, but (2) $X \not\perp\!\!\!\perp Y \mid Pa^*(X, \mathcal{P}) \cup Ne^*(X_{\mathbf{M}}, \mathcal{P})$ for every $\mathbf{M} \in \mathbb{M}$.*

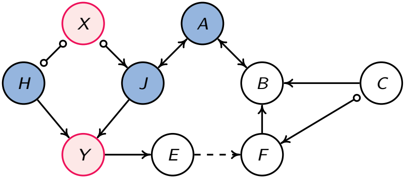


Figure: A PAG

$PossCh(X, \mathcal{P}) \cup NondNe(X, \mathcal{P}) = \{H, J\}$,

$\mathbb{M} = \{\mathbf{M}_1 = \{H\}, \mathbf{M}_2 = \{J\}\}$

$Ne^*(X_{\mathbf{M}_1}, \mathcal{P}) = \{H\}$,

$Ne^*(X_{\mathbf{M}_2}, \mathcal{P}) = \{A, J\}$,

$(2) X \not\perp\!\!\!\perp Y \mid \{H\}, X \not\perp\!\!\!\perp Y \mid \{A, J\}$

Local Characterization: Implicit Invariant Ancestor

Theorem 3

*Let $\mathcal{P}$ be the PAG over $\mathbf{O}$. For any pair of vertices (X, Y) in $\mathcal{P}$, X is an **implicit invariant ancestor** of Y if and only if (1) $X \perp\!\!\!\perp Y \mid Pa^*(X, \mathcal{P}) \cup Ne^*(X, \mathcal{P})$, but (2) $X \not\perp\!\!\!\perp Y \mid Pa^*(X, \mathcal{P}) \cup Ne^*(X_{\mathbf{M}}, \mathcal{P})$ for every $\mathbf{M} \in \mathbb{M}$.*

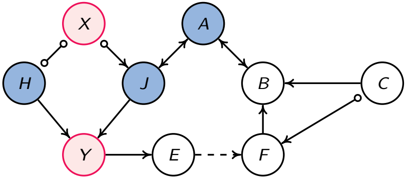


Figure: A PAG

X is an implicit invariant ancestor of Y.

Intuitively, similar to the role of $Ne^*(X, \mathcal{P})$ in $\mathcal{P}$, $Ne^*(X_{\mathbf{M}}, \mathcal{P})$ can block all partially directed paths (excluding directed paths) from X to Y that pass through $\mathbf{M}$ in $\mathcal{P}$.

Local Characterization: Invariant Ancestor and Possible Ancestor

Corollary 1

Let $\mathcal{P}$ be the PAG over $\mathbf{O}$, and let $\mathbb{M}$ denote the set of maximal cliques of the induced subgraph of $\mathcal{P}$ over $PossCh(X,\mathcal{P}) \cup NondNe(X,\mathcal{P})$. For any pair of vertices (X, Y) in $\mathcal{P}$, X is an ***invariant ancestor*** *of Y if and only if $X \not\perp\!\!\!\perp Y \mid Pa^*(X,\mathcal{P}) \cup Ne^*(X,\mathcal{P})$ or $X \not\perp\!\!\!\perp Y \mid Pa^*(X,\mathcal{P}) \cup Ne^*(X_{\mathbf{M}},\mathcal{P})$ for every $\mathbf{M} \in \mathbb{M}$.*

Theorem 4

Let $\mathcal{P}$ be the PAG over $\mathbf{O}$. For any pair of vertices (X, Y) in $\mathcal{P}$, X is a ***possible ancestor*** *of Y if and only if neither Theorem 1 nor Corollary 1 applies.*

Locally Learning Conditional Sets

To find the conditional sets $Pa^{*}(X, \mathcal{P})$, $Ne^{*}(X, \mathcal{P})$, and $Ne^{*}(X_{\mathbf{M}}, \mathcal{P})$ for every maximal clique $\mathbf{M} \in \mathbb{M}$, we need to answer the following two questions:

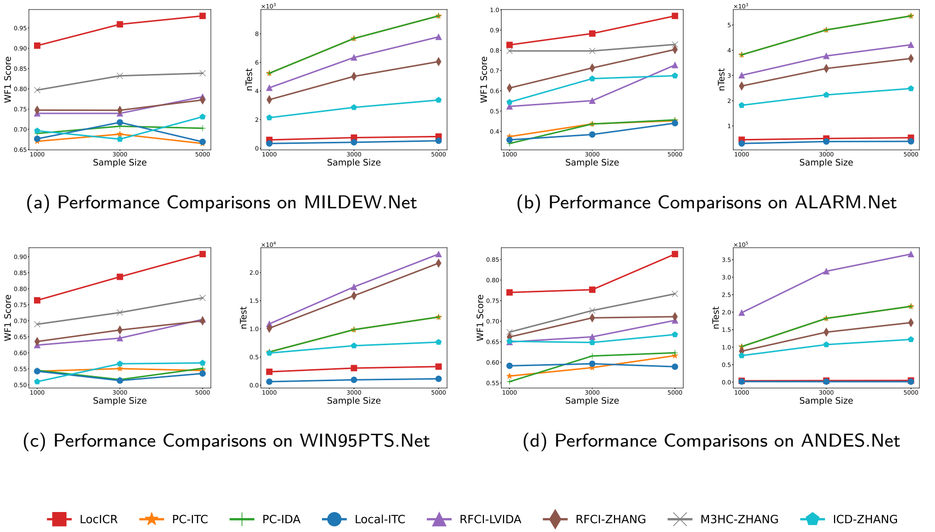
- How to discover which vertices in $MB(X)$ are connected to X via arrow-collider paths or circle-collider paths?

  This involves locally learning the induced subgraph of the global PAG over $MB^{+}(X)$, i.e. , $\mathcal{P}_{MB+(X)}$.

- How to determine whether X is an invariant non-ancestor of the vertices on these paths?

  This requires identifying all vertices in $MB(X)$ for which X is an invariant non-ancestor, denoted as $Pre_{MB}(X)$.

Experimental Results on Benchmark Network Structures



(a) Performance Comparisons on MILDEW.Net

(b) Performance Comparisons on ALARM.Net

(c) Performance Comparisons on WIN95PTS.Net

(d) Performance Comparisons on ANDES.Net

Figure: Performance of eight algorithms on four benchmark networks

Application to General Social Survey Data

The dataset we used contains six observed variables, and 1380 samples (O. D. Duncan, Featherman, and B. Duncan 1972).

- We first selected the father's occupation (X) and the son's education (Y) as the target variable pair, connected by a direct edge. Our method identifies the father's occupation as an invariant ancestor of the son's education.
- Next, we selected father's occupation (X) and son's income (Y), as well as father's education (X) and son's income (Y) as the target variable pairs. In both cases, X and Y are connected by directed paths. Our method identifies both father's occupation and father's education as invariant ancestors of the son's income.
- Finally, for son's income (X) and number of siblings (Y), which are neither connected by a direct edge nor a directed path from X to Y. Our method finds the son's income as an invariant non-ancestor of the number of siblings.

Application to Gene Expression Data

The dataset we used contains 33 genes, and 118 samples (Wille et al. 2004).

- We first selected *DXR* (X) and *MCT* (Y), as well as *HMGS* (X) and *HMGR1* (Y), as the target pairs, both of which are connected by a direct edge. Our method identifies *DXR* as an invariant ancestor of *MCT*, and similarly, *HMGS* as an invariant ancestor of *HMGR1*.
- Next, we selected *AACT1* (X) and *FPPS1* (Y), as well as *DXPS3* (X) and *CMK* (Y), where each pair is connected by a directed path from X to Y. Our method identifies *AACT1* as an invariant ancestor of *FPPS1*, and likewise, *DXPS3* as an invariant ancestor of *CMK*.
- Finally, we considered *PPDS1* (X) and *DXPS1* (Y), as well as *DXPS1* (X) and *DXPS3* (Y), neither of which is connected by a direct edge or a directed path from X to Y. Our method finds that *PPDS1* is an invariant non-ancestor of *DXPS1*, and similarly, *DXPS1* is an invariant non-ancestor of *DXPS3*.

Conclusion and Future Work

Conclusion: We addressed the problem of locally learning causal relations from observational data without assuming causal sufficiency.

- We provided sufficient and necessary local characterizations for identifying invariant ancestors, invariant non-ancestors, and possible ancestors.
- We introduced LocICR, a novel algorithm for local causal discovery. We proved that LocICR is complete, matching the accuracy of existing global methods.

Future Work:

- How can background knowledge be utilized to further aid in locally identifying the causal relationships in the presence of latent variables?
- How does local causal discovery from multi-environment data work?

Thank You!

Questions and Discussion

References I

Duncan, Otis Dudley, David L Featherman, and Beverly Duncan (1972). "Socioeconomic background and achievement". In: Seminar Press.

Wille, Anja et al. (2004). "Sparse graphical Gaussian modeling of the isoprenoid gene network in Arabidopsis thaliana". In: *Genome biology* 5, pp. 1–13.

Zhang, Jiji (2006). "Causal inference and reasoning in causally insufficient systems". PhD thesis. Citeseer.