Introduction to Causal Discovery

More on (T)PC in practice: Tiers and finite sample problems

Christine Bang
Leibniz Institute for Prevention Research & Epidemiology – BIPS

April 16, 2024
EuroCIM 2024, Copenhagen

Tiered background knowledge

Constraint-based causal discovery:

separation in graph $\Leftrightarrow$ (conditional) independence

Tiered background knowledge

Constraint-based causal discovery:

separation in graph $\Leftrightarrow$ (conditional) independence

Theoretical issue: Can at best estimate an equivalence class.

Practical issue: Algorithm sensitive to statistical errors.

Tiered background knowledge

Constraint-based causal discovery:

separation in graph $\Leftrightarrow$ (conditional) independence

Theoretical issue: Can at best estimate an equivalence class.

Practical issue: Algorithm sensitive to statistical errors.

With background knowledge: Estimate restricted equivalence class represented by an MPDAG [Perković et al., 2017].

$\Rightarrow$ Contains information additional to independence.

$\Rightarrow$ Estimate more robust to statistical errors [Petersen et al., 2021, Bang et al., 2024].

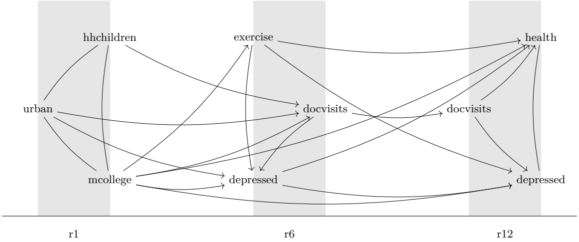
Informativeness

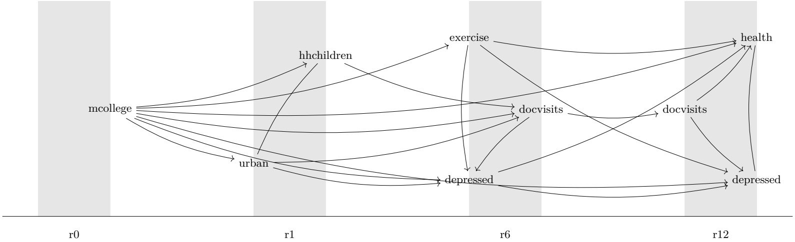
CPDAGs: Encode (conditional) independencies

MPDAGs: Encode (conditional) independencies and additional causal information

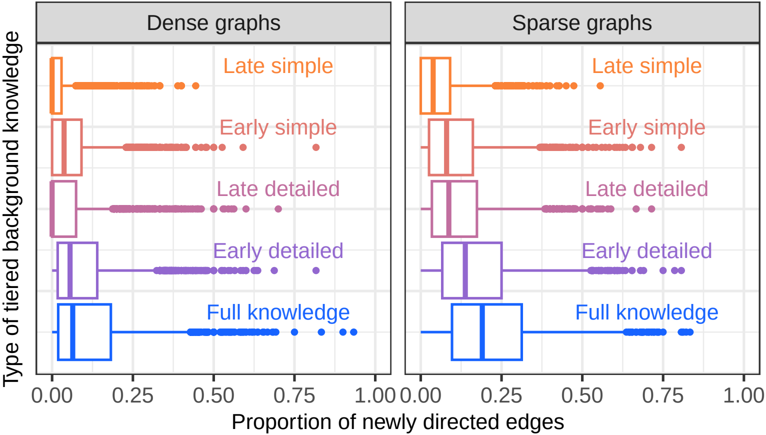
DAGs: Encode (conditional) independencies and all causal information

Informativeness



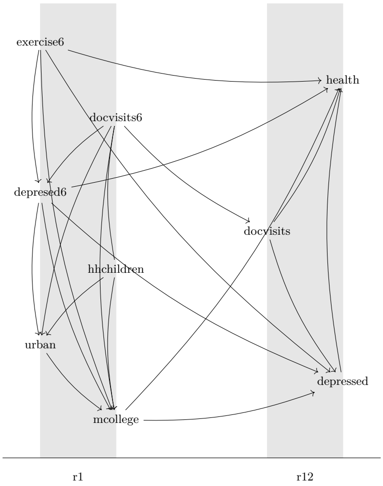


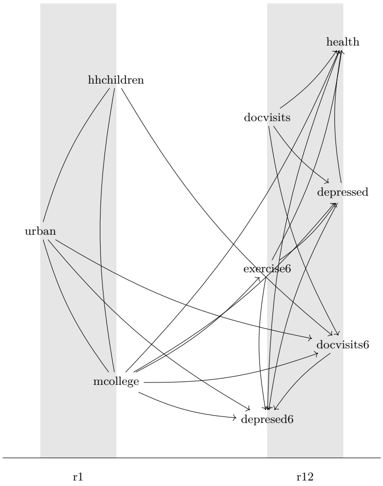
Informativeness



For details: Bang and Didelez [2023]

Finite sample data





Finite sample data

Statistical tests might yield incorrect independence results (e.g. directed edges contradicting the flow of time)

⇒ With background knowledge we do not necessarily get more informative graphs – but we expect fewer errors.

Finite sample data

Statistical tests might yield incorrect independence results (e.g. directed edges contradicting the flow of time)

⇒ With background knowledge we do not necessarily get more informative graphs – but we expect fewer errors.

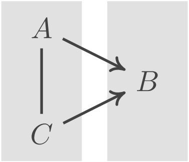
Moreover: Inconsistent independencies might result in conflicting edges – can be resolved by background knowledge.

Robustness

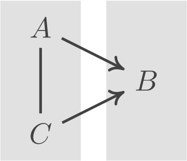
Fewer edges tests:

If $A \not\perp\!\!\!\perp C \mid \emptyset$ and $\mathrm{tier}(A) \leq \mathrm{tier}(C) < \mathrm{tier}(B)$,
then $A \not\perp\!\!\!\perp C \mid \{B\}$.

$\Rightarrow$ fewer type II errors (higher edge recall).



Robustness



Fewer edges tests:

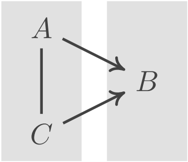
If $A \not\perp\!\!\!\perp C \mid \emptyset$ and $\mathrm{tier}(A) \leq \mathrm{tier}(C) < \mathrm{tier}(B)$, then $A \not\perp\!\!\!\perp C \mid \{B\}$.

$\Rightarrow$ fewer type II errors (higher edge recall).

Resolving conflicts:

Assume that both $A \to B \leftarrow C$ and $B \to C \leftarrow D$ then this might be resolved by background knowledge $\Rightarrow$ fewer conflicts.

Robustness



Fewer edges tests:

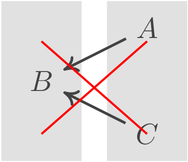
If $A \not\perp\!\!\!\perp C \mid \emptyset$ and $\mathrm{tier}(A) \leq \mathrm{tier}(C) < \mathrm{tier}(B)$,
then $A \not\perp\!\!\!\perp C \mid \{B\}$.
$\Rightarrow$ fewer type II errors (higher edge recall).

Resolving conflicts:

Assume that both $A \to B \leftarrow C$ and $B \to C \leftarrow D$ then this might be resolved by background knowledge $\Rightarrow$ fewer conflicts.

Fewer incorrect directed edges:

Suppose we incorrectly got $A \perp\!\!\!\perp C \mid \emptyset$
and $A \not\perp\!\!\!\perp C \mid \{B\}$,
but $\mathrm{tier}(B) < \mathrm{tier}(A) \leq \mathrm{tier}(C)$
$\Rightarrow$ here v-structure ruled out by the tiers.



Unobserved confounding

We now relax the assumption of no unobserved confounding and consider ‘latent DAGs’.

Unobserved confounding

We now relax the assumption of no unobserved confounding and consider ‘latent DAGs’.

Represent (pure) latent confounding using bidirected edges:

$$\begin{tikzcd} & L \arrow[dl, bend right] \arrow[dr, bend left] & \\ A & & B \end{tikzcd} \quad \Longrightarrow \quad \begin{tikzcd} A \arrow[rr, leftrightarrow, bend left] & & B \end{tikzcd}$$

Assume that all directed edges might be confounded:

$$\begin{tikzcd} A \arrow[r] \arrow[r, leftrightarrow, bend left] & B \end{tikzcd} \quad \Longrightarrow \quad \begin{tikzcd} A \arrow[r] & B \end{tikzcd}$$

Unobserved confounding

We now relax the assumption of no unobserved confounding and consider ‘latent DAGs’.

Represent (pure) latent confounding using bidirected edges:

$$\begin{tikzcd} & L \arrow[dl] \arrow[dr] & \\ A & & B \end{tikzcd} \quad \Longrightarrow \quad \begin{tikzcd} A \arrow[rr, bend left, leftrightarrow] & & B \end{tikzcd}$$

Assume that all directed edges might be confounded:

$$\begin{tikzcd} A \arrow[r] \arrow[r, bend left, leftrightarrow] & B \end{tikzcd} \quad \Longrightarrow \quad \begin{tikzcd} A \arrow[r] & B \end{tikzcd}$$

Task: Adapt the expert graph!

References

Christine W Bang and Vanessa Didelez. Do we become wiser with time? On causal equivalence with tiered background knowledge. Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence, pages 119–129. PMLR, 2023.

Christine W Bang, Janine Witte, Ronja Foraita, and Vanessa Didelez. Efficient use of tiered background knowledge for causal discovery with cohort data. Working paper, 2024.

Emilija Perković, Markus Kalisch, and Marloes H Maathuis. Interpreting and using cpdags with background knowledge. In Proceedings of the 2017 Conference on Uncertainty in Artificial Intelligence (UAI2017), pages ID–120. AUAI Press, 2017.

Anne H Petersen, Merete Osler, and Claus T Ekstrøm. Data-driven model building for life-course epidemiology. American Journal of Epidemiology, 190(9):1898–1907, 2021.