Imperial College London

Argumentative Causal Discovery

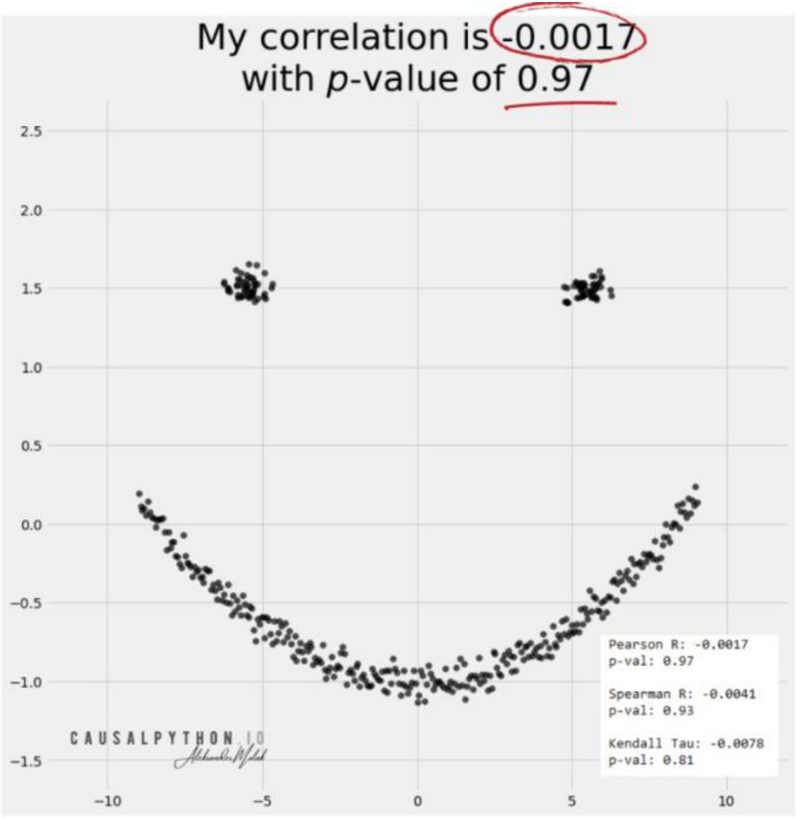
Improving causal discovery through logic and argumentation

Fabrizio Russo

Imperial College London

Correlation is not Causation

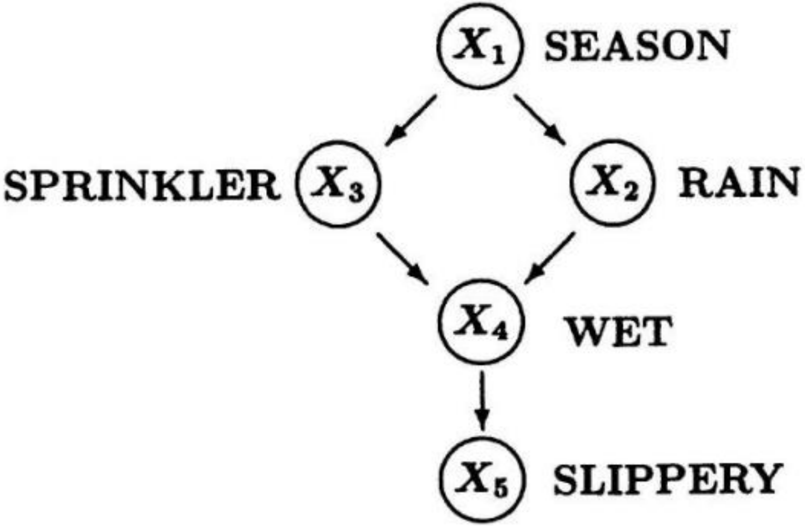
Data does not always play nice



Imperial College London

Graphical Models

Why?

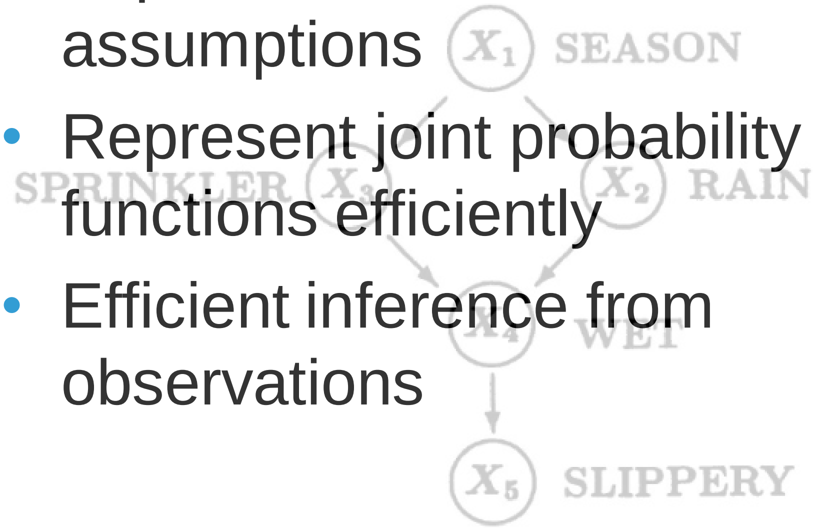


Pearl J. (2009)

Graphical Models

Why?

- Express substantive assumptions
- Represent joint probability functions efficiently
- Efficient inference from observations

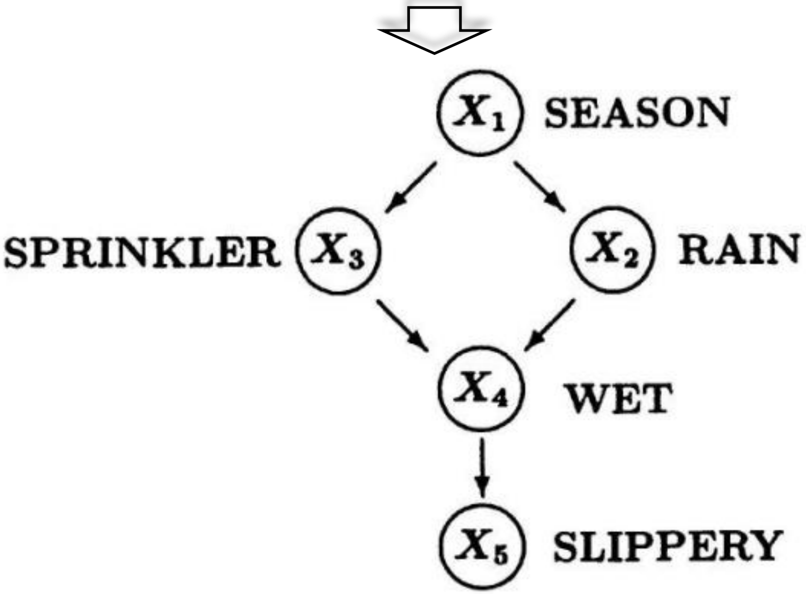


Pearl J. (2009)

Imperial College London

Bayesian Structure Learning aka Causal Discovery

Season	Sprinkler	Rain	Wet	Slippery
Dry	ON	YES	YES	YES
Wet	OFF	NO	NO	NO
...	...	...	...	...



Given some data, retrieve the causal graph underlying the data generating process

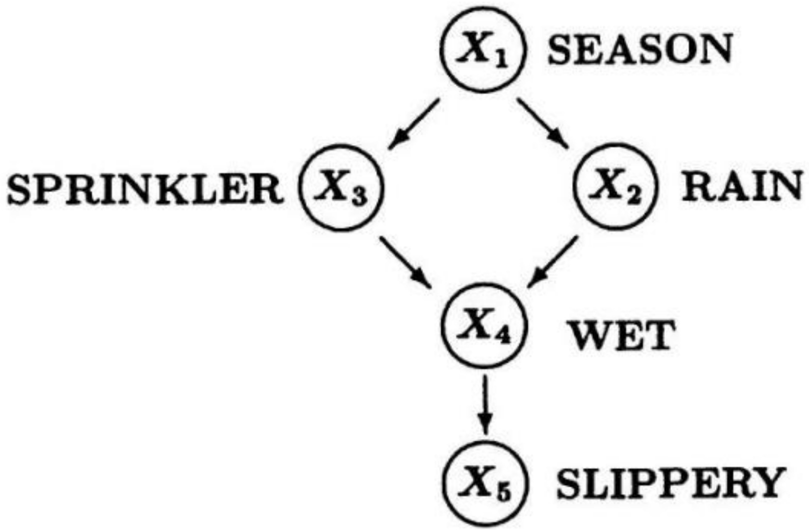
Bayesian Networks for a simple conversation*

Q1: If the season is dry, and the pavement is slippery, did it rain?

A1: Unlikely, it is more likely the sprinkler was ON.

Q2: But what if we SEE that the sprinkler is OFF?

A2: Then it is more likely that it rained



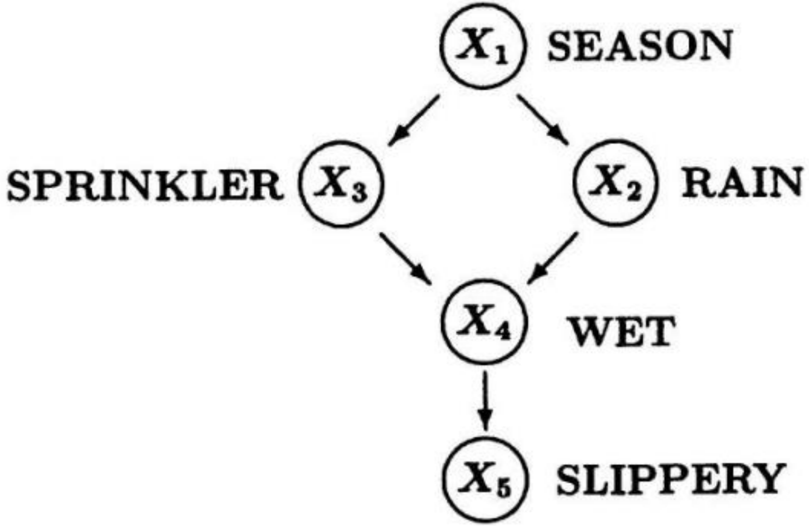
*From Rudolf Kruse and Alexander Dockhorn, Causal Networks, open6.pptx (ovgu.de)

Imperial College London

Bayesian Networks for a simple conversation*

Q3: Do you mean that if we actually turn the sprinkler OFF, the rain will be more likely?

A3: No, the likelihood of rain would remain the same

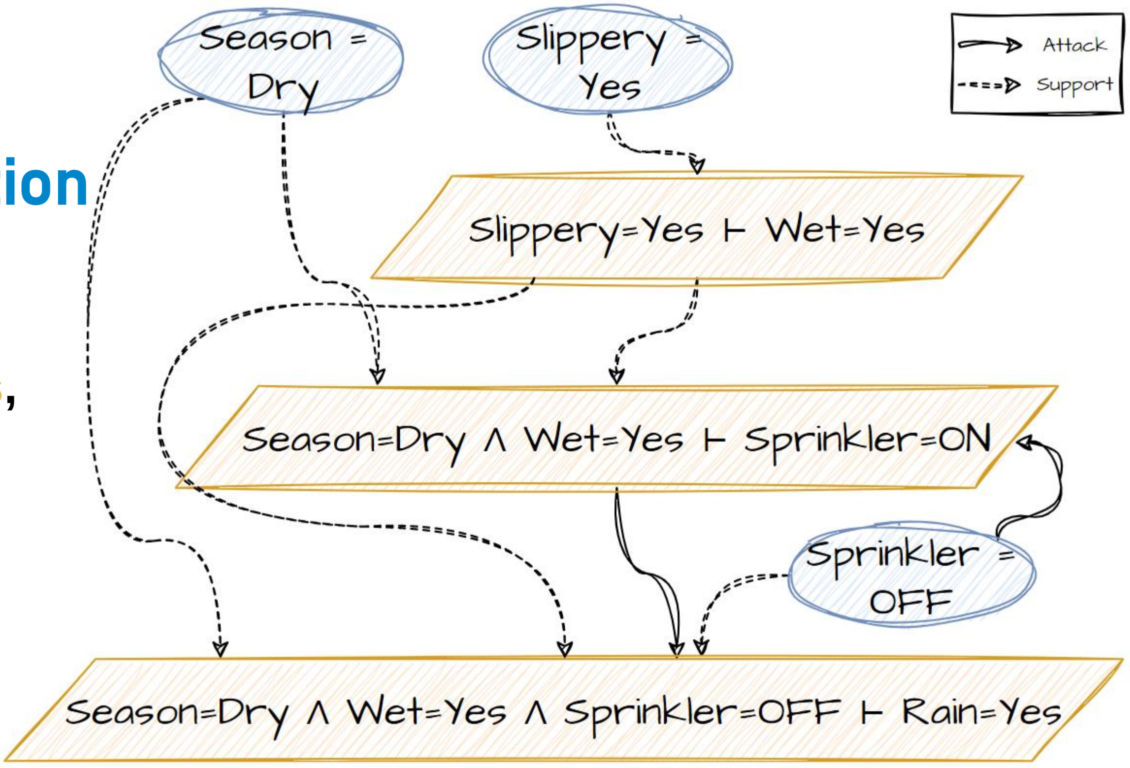


*From Rudolf Kruse and Alexander Dockhorn, Causal Networks, open6.pptx (ovgu.de)

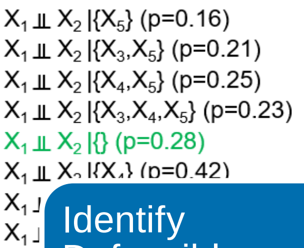
Imperial College London

Structured Argumentation

Given assumptions and rules, reason about defeasible and conflicting knowledge

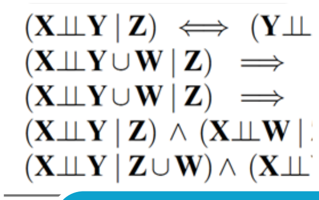


Argumentative Causal Discovery – Overview



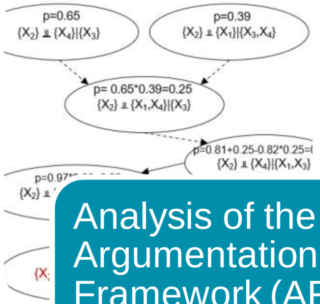
Identify Defeasible Facts

- Run constrained-based Causal Discovery Algorithms



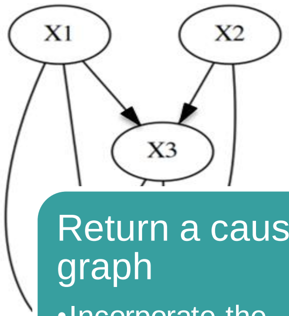
Rule-based Reasoning

- Augment tests from data with Pearlian axiomatisation and *d-separation* rules



Analysis of the Argumentation Framework (AF)

- Identify "strongest" independence test set that complies with rules



Return a causal graph

- Incorporate the results of the argumentative analysis into the causal graph

Imperial College London

PC algorithm (Spirtes et al, 1991)

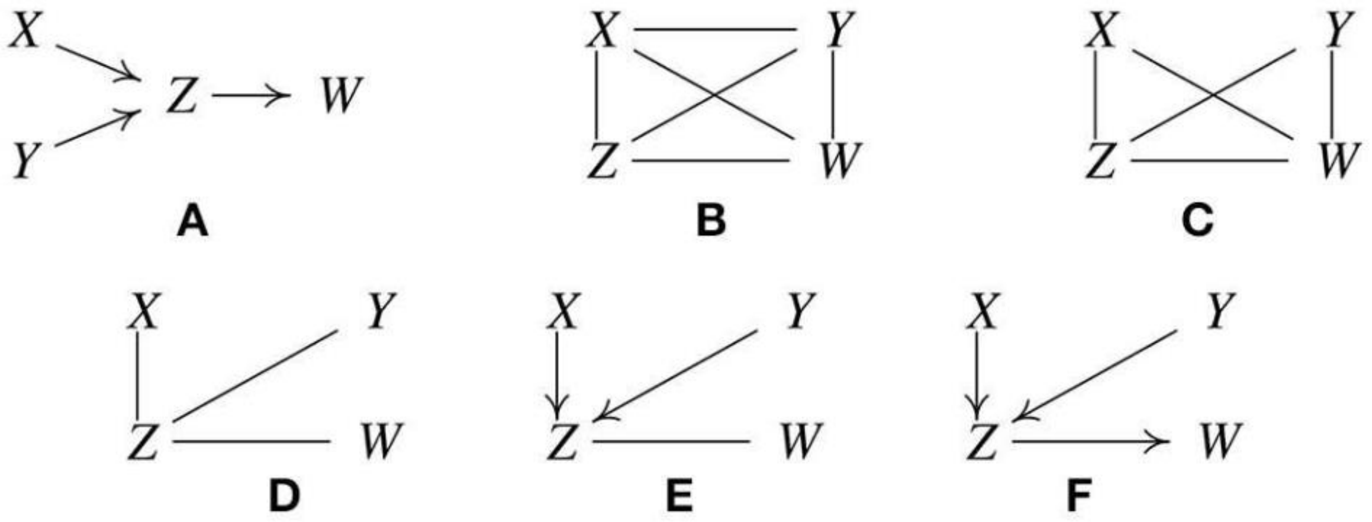


FIGURE 1 | Illustration of how the PC algorithm works. **(A)** Original true causal graph. **(B)** PC starts with a fully-connected undirected graph. **(C)** The $X - Y$ edge is removed because $X \perp\!\!\!\perp Y$. **(D)** The $X - W$ and $Y - W$ edges are removed because $X \perp\!\!\!\perp W \mid Z$ and $Y \perp\!\!\!\perp W \mid Z$. **(E)** After finding v-structures. **(F)** After orientation propagation.

Figure from Glymour et al (2019)

Imperial College London

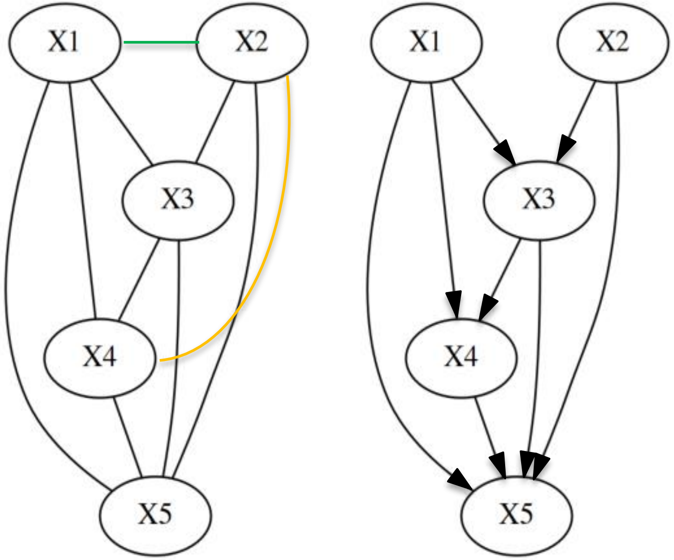
Identifying Defeasible Facts

A continuation of the skeleton example in Colombo and Maathius (2014)

Imperial College London

Step 1 – Skeleton Estimation with Max-PC (Ramsey, 2016)

- $X_1 \perp\!\!\!\perp X_2$ (p-value= 0.28)
- $X_2 \perp\!\!\!\perp X_4$ (p-value= 0.18)
  - Edge is correctly removed, but for the wrong reason
  - It should be $X_2 \perp\!\!\!\perp X_4 \mid \{X_1, X_3\}$



Estimated Skeleton – Green edges removed

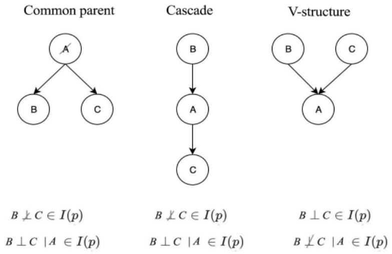
True DAG

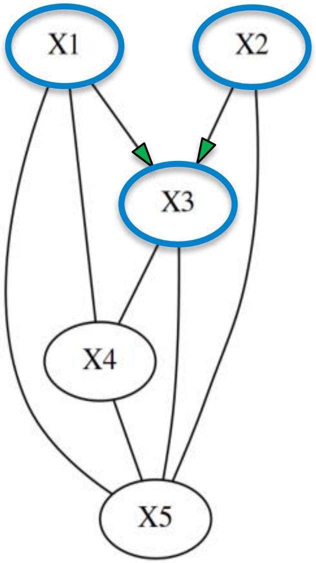
Imperial College London

Step 2 – V-structures Estimation

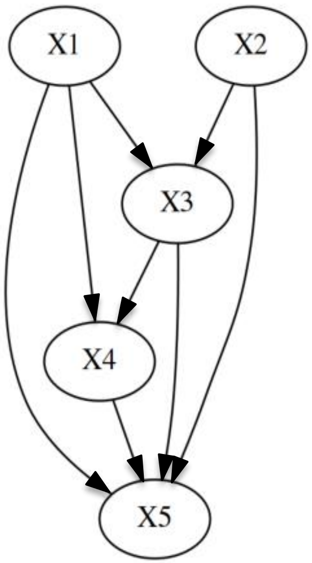
- Consider the triple $\{X_1, X_3, X_2\}$
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_3\}$ (p=0.35)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_3, X_4\}$ (p=0.39)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_3, X_5\}$ (p=0.21)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_3, X_4, X_5\}$ (p=0.23)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{\}$ (p=0.28)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_4\}$ (p=0.43)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_5\}$ (p=0.16)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_4, X_5\}$ (p=0.24)
- 0.43>0.39=> $X_1 \not\perp\!\!\!\perp X_2 \mid \{X_3\}$

d-separation Rules - Rebane and Pearl, 1987.





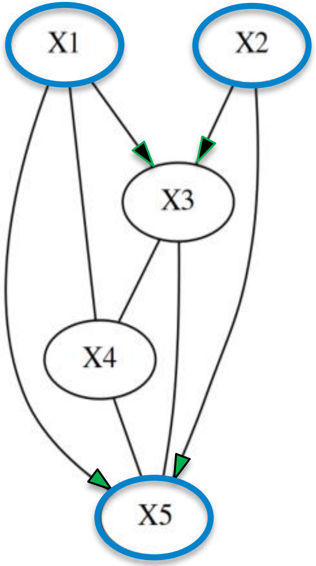
Estimated V-structure – colour arrowheads



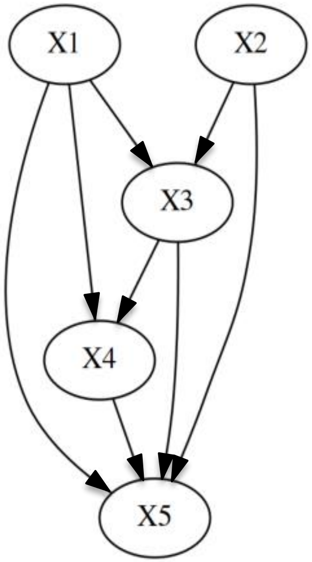
True DAG

Step 2 – V-structures Estimation

- Consider the triple $\{X_1,X_5,X_2\}$
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_5\}$ (p=0.16)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_3,X_5\}$ (p=0.21)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_4,X_5\}$ (p=0.25)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_3,X_4,X_5\}$ (p=0.23)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{\}$ (p=0.28)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_4\}$ (p=0.43)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_5\}$ (p=0.16)
  - $X_1 \perp\!\!\!\perp X_2 \mid \{X_3,X_4\}$ (p=0.39)
- 0.42>0.25=> $X_1 \not\perp\!\!\!\perp X_2 \mid \{X_5\}$



Estimated V-structure – colour arrowheads

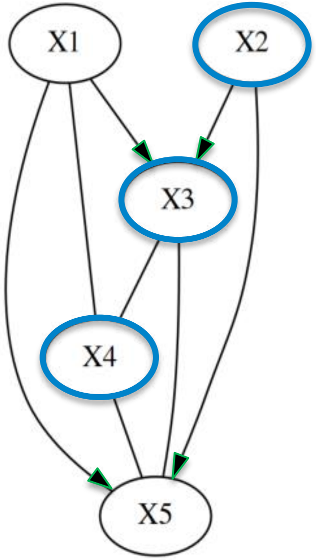


True DAG

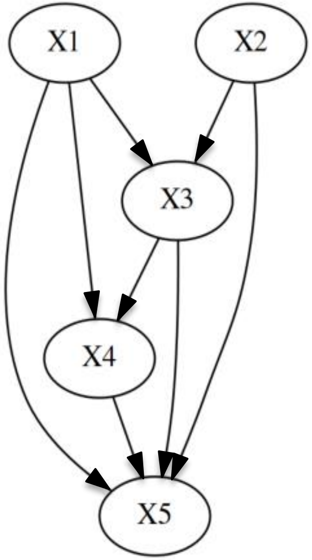
Imperial College London

Step 2 – V-structures Estimation

- Consider the triple $\{X_2,X_3,X_4\}$
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_3\}$ (p=0.65)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1,X_3\}$ (p=0.81)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_5,X_3\}$ (p=0.75)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1,X_5,X_3\}$ (p=0.97)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{\}$ (p=0.18)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1\}$ (p=0.27)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_5\}$ (p=0.24)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1,X_5\}$ (p=0.39)
- 0.39<0.97=> $X_2 \perp\!\!\!\perp X_4 \mid \{X_3\}$
- No V-structure



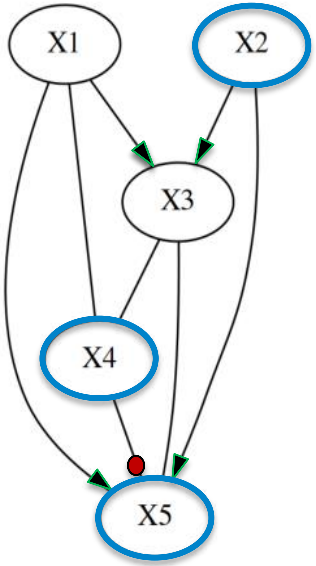
Estimated V-structure
– colour arrowheads



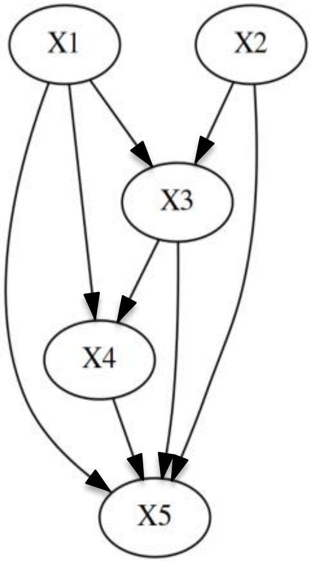
True DAG

Step 2 – V-structures Estimation

- Consider the triple $\{X_2, X_5, X_4\}$
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_5\}$ ($p=0.24$)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1, X_5\}$ ($p=0.39$)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_3, X_5\}$ ($p=0.75$)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1, X_3, X_5\}$ ($p=0.97$)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{\}$ ($p=0.18$)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1\}$ ($p=0.27$)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_3\}$ ($p=0.66$)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1, X_3\}$ ($p=0.81$)
- $0.81 < 0.97 \Rightarrow X_2 \perp\!\!\!\perp X_4 \mid \{X_5\}$
- No V-structure ← Sample issue



Estimated V-structure – colour arrowheads



True DAG

Imperial College London

Step 3 – Orientation Rules

- $X_2 \rightarrow X_3$ and $X_3 - X_4$
  - X_4 is not ancestor of X_3
  - Orient: $X_3 \rightarrow X_4$
- $X_1 \rightarrow X_3$ and $X_3 \rightarrow X_4$ and $X_1 - X_4$
  - Orient: $X_1 \rightarrow X_4$
- Cascade error: $X_3 \rightarrow X_5$ if $X_4 \rightarrow X_5$ (by Rule 2)

From Meek, 1995. Dashed line: either direction. Solid line: undirected

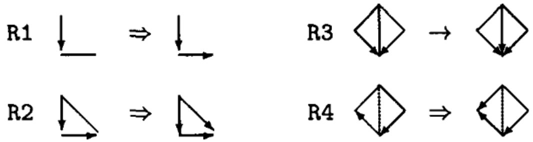
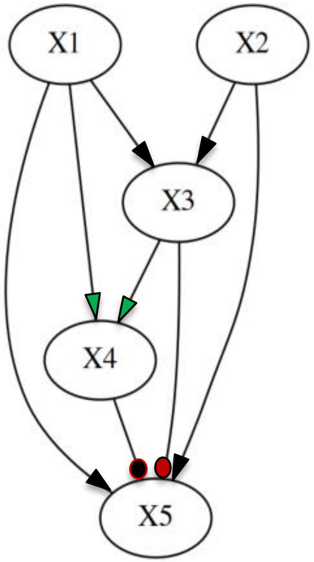
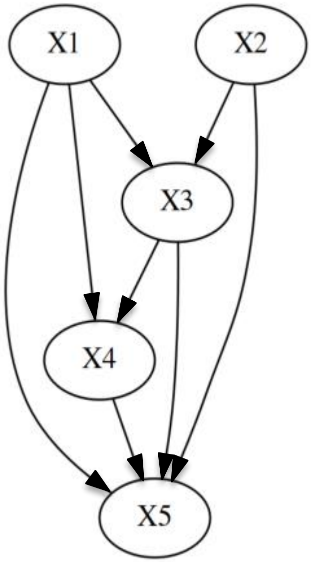


Figure 1: Orientation rules for patterns



Estimated V-structure – colour arrowheads



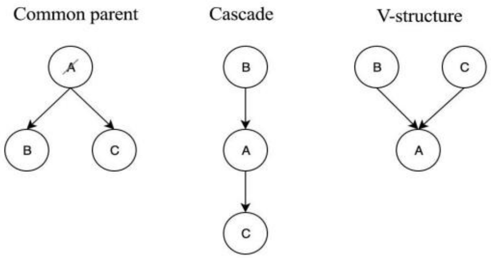
True DAG

Reasoning about Independence tests

Integrate Pearl's axiomatic representation with V-structures and Meek's rules

Imperial College London

Graphical and Axiomatic Rules



Common parent	Cascade	V-structure
$B \not\perp C \in I(p)$	$B \not\perp C \in I(p)$	$B \perp C \in I(p)$
$B \perp C \mid A \in I(p)$	$B \perp C \mid A \in I(p)$	$B \not\perp C \mid A \in I(p)$

d-separation Rules - Rebane and Pearl (1987)

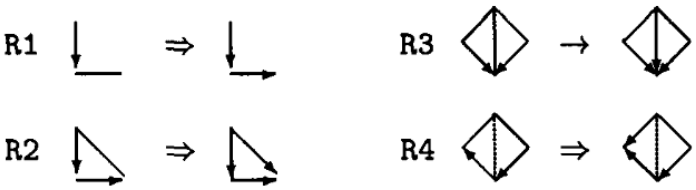


Figure 1: Orientation rules for patterns

Orientation Rules - Meek (1995)

$$
\begin{aligned}
&\textbf{(Symmetry)} && (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \iff (\mathbf{Y} \perp\!\!\!\perp \mathbf{X} \mid \mathbf{Z}) \\
&\textbf{(Decomposition)} && (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z}) \\
&\textbf{(Weak Union)} && (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \mathbf{W}) \\
&\textbf{(Contraction)} && (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z} \cup \mathbf{Y}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) \\
&\textbf{(Intersection)} && (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \mathbf{W}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z} \cup \mathbf{Y}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z})
\end{aligned} \qquad (5)
$$

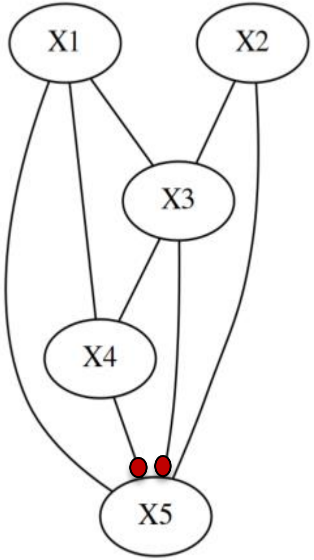
Graphoid Rules - Pearl and Paz (1987)

(5) From Bromberg and Maathius (2009). Adapted from Pearl and Paz (1987)

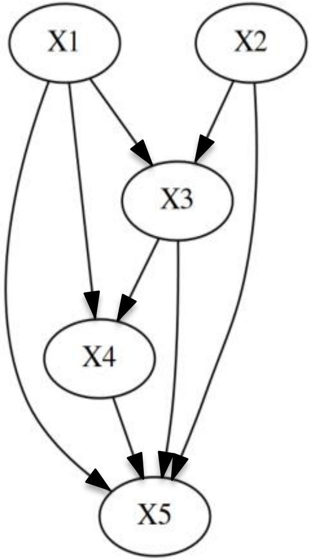
Imperial College London

Step 2 – V-structures Estimation

- Consider the triple $\{X_2, X_5, X_4\}$
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_5\}$ (p=0.24)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1, X_5\}$ (p=0.39)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_3, X_5\}$ (p=0.75)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1, X_3, X_5\}$ (p=0.97)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{\}$ (p=0.18)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1\}$ (p=0.27)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_3\}$ (p=0.66)
  - $X_2 \perp\!\!\!\perp X_4 \mid \{X_1, X_3\}$ (p=0.81)
- Sample issue: 0.81<0.97=> $X_2 \perp\!\!\!\perp X_4 \mid \{X_5\}$
- Cascade error: $X_3 \rightarrow X_5$ if $X_4 \rightarrow X_5$ (by Rule 2)



Estimated V-structure – colour arrowheads



True DAG

Imperial College London

Apply "Graphoid Axioms"

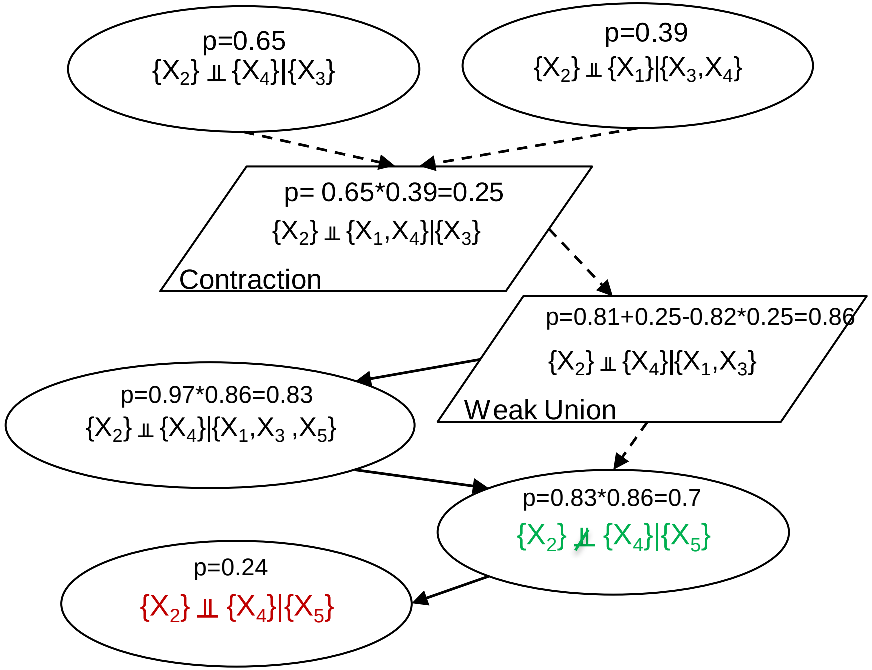
$$
\begin{array}{lll}
\text{(Symmetry)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \iff (\mathbf{Y} \perp\!\!\!\perp \mathbf{X} \mid \mathbf{Z}) & \\
\text{(Decomposition)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z}) & \\
\text{(Weak Union)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \mathbf{W}) & (5) \\
\text{(Contraction)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z} \cup \mathbf{Y}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) & \\
\text{(Intersection)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \mathbf{W}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z} \cup \mathbf{Y}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) &
\end{array}
$$

- $X_2 \perp\!\!\!\perp X_4 \mid \{X_3\}$ (p=0.65)
- $X_2 \perp\!\!\!\perp X_1 \mid \{X_3, X_4\}$ (p=0.39)
- Apply Contraction:
  - $\{X_2 \perp\!\!\!\perp X_4 \mid \{X_3\}, X_2 \perp\!\!\!\perp X_1 \mid \{X_3, X_4\}\} \implies X_2 \perp\!\!\!\perp \{X_1, X_4\} \mid \{X_3\}$
- Apply Weak Union:
  - $\{X_2 \perp\!\!\!\perp \{X_1, X_4\} \mid \{X_3\}\} \implies X_2 \perp\!\!\!\perp X_4 \mid \{X_1, X_3\}$

Imperial College London

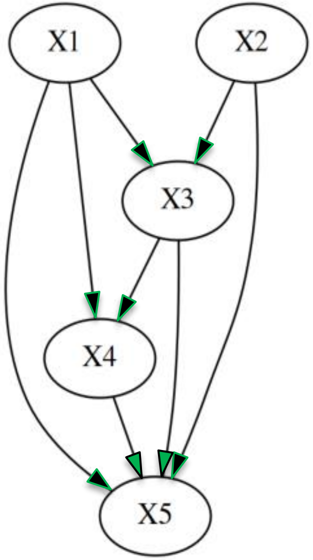
Argumentation Framework

- Objective: **decide which set of independencies is the strongest and adjust the causal graph accordingly**
- Use gradual semantics for bipolar graphs e.g.
  - DF-quad (Rago et al, 2016)
  - T-(co)norms (Jedwabny et al, 2020)

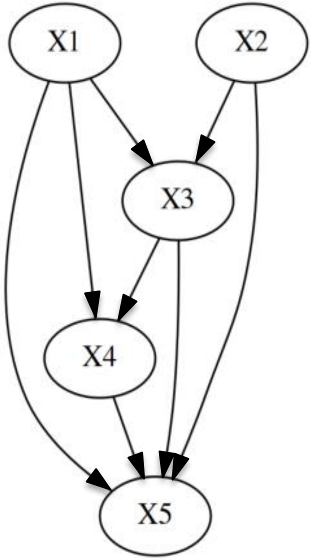


Imperial College London

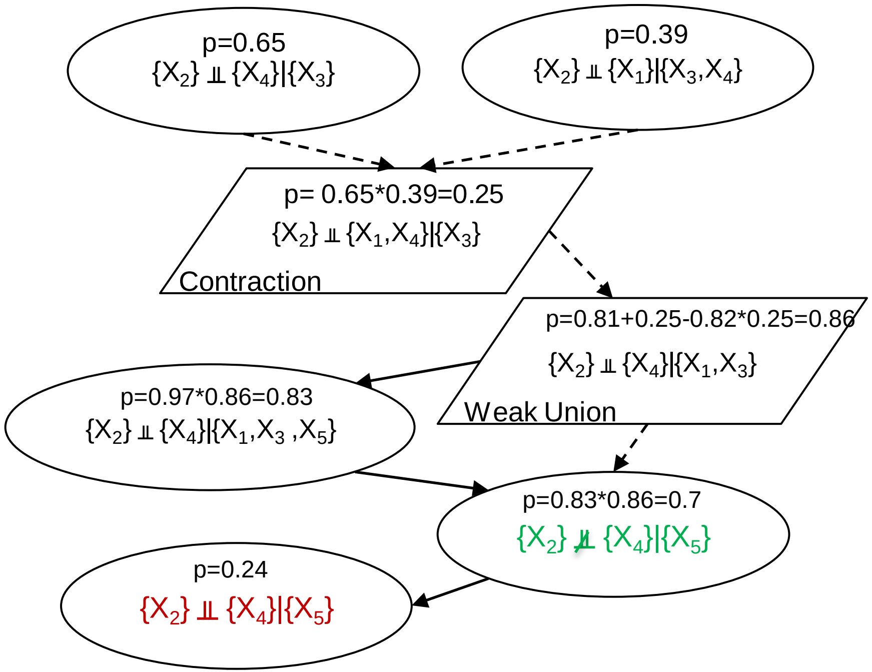
Get back to the graph



Estimated V-structure – colour arrowheads



True DAG



Imperial College London

Conclusions

Add rigour and constraints to data-driven results

Imperial College London

Conclusions

- Preliminary work
- Debating about the orientation phases as well as skeleton should significantly improve on the current benchmarks for causal discovery
- Experimentation with ABA+ and T-Norms semantics is the current focus
- Extensions to different independence tests and causal discovery algorithms

Imperial College London

Questions?

Imperial College London

Appendix

Imperial College London

Previous work – Argumentative Independence Tests (AIT)*

Use

Preference-based argumentation

To incorporate information from

An *Axiomatic Representation*

to resolve inconsistencies from

Multiple conditional independence tests

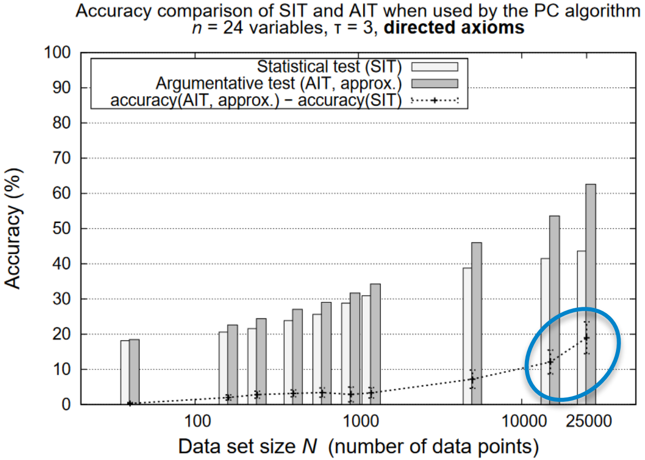
underlying

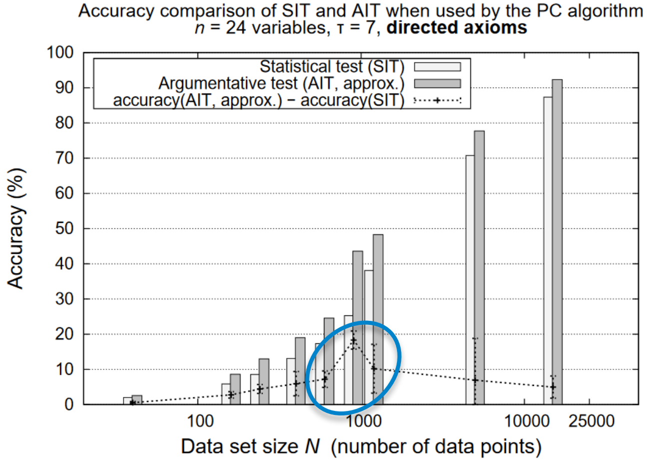
The PC Algorithm

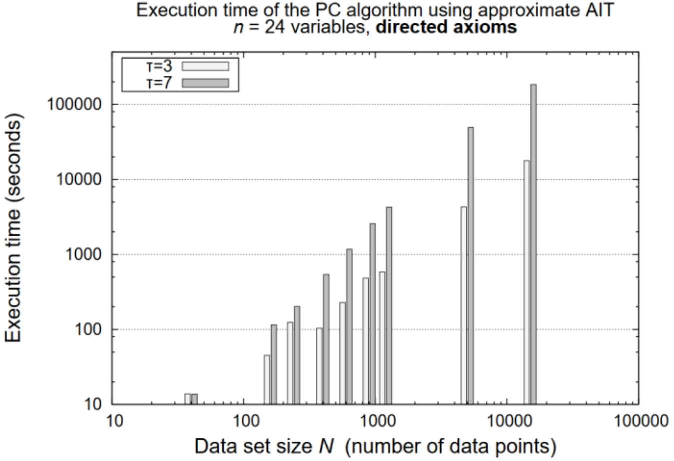
*Bromberg, F and Margaritis, D, Improving the Reliability of Causal Discovery from Small Data Sets Using Argumentation, 2009

Imperial College London

- AIT used in combination with PC algorithm
- Accuracy: Improvements of up to 20%
- Cost:
  - significantly lower for sparse domains (τ=3)
  - Scales linearly with *N*







Imperial College London

Argumentative Causal Discovery (ACD)

Use

Structured Argumentation and Gradual Semantics

To incorporate information from

An *Axiomatic Representation* and *Graphical Rules*

to resolve inconsistencies from

Multiple conditional independence tests

and incorporate

Various CD techniques in the debate

Imperial College London

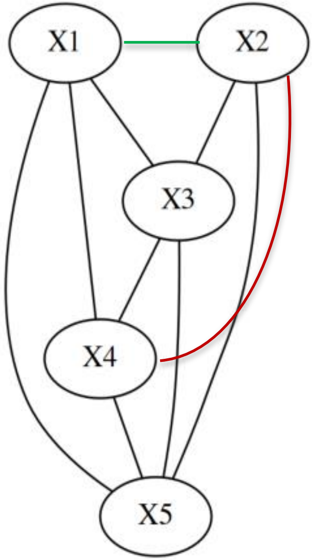
Step 1 – Skeleton Estimation with PC-Stable

The following tests result in independencies with depth (conditioning size) equal to 0

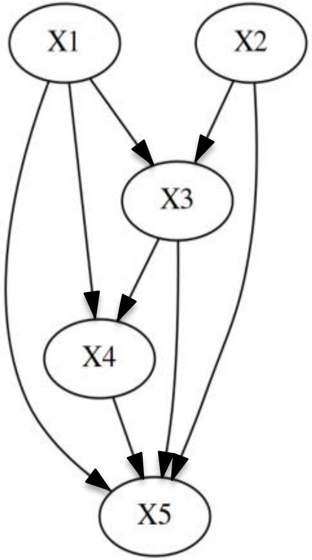
- $X_1 \perp\!\!\!\perp X_2 | \{\}$ (p-value= 0.28)
- $X_2 \perp\!\!\!\perp X_4 | \{\}$ (p-value= 0.18) edge is removed for the wrong reason

Once all depth=0 are calculated, depth 1, 2, 3 will follow, depending on the results from the shallower depths

No other independencies found



Estimated Skeleton – Green edges removed

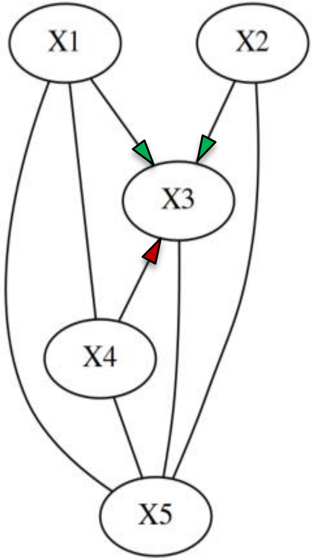


True DAG

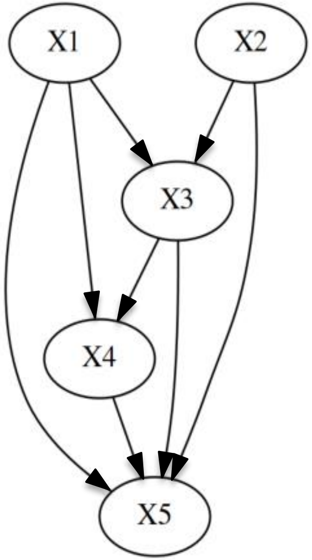
Imperial College London

Step 2 – V-structures Estimation

- Consider the triple $\{X_1, X_3, X_2\}$
  - X_3 is not in the separating set of $\{X_1, X_2\}$ i.e. $X_1 \not\perp\!\!\!\perp X_2 \mid X_3$
    - Problem: $X_1 \perp\!\!\!\perp X_2 \mid X_3$ is never calculated, only assumed given that $X_1 \perp\!\!\!\perp X_2 \mid \{\}$
- Consider the triple $\{X_2, X_3, X_4\}$
  - X_3 is not in the separating set of $\{X_2, X_4\}$ i.e. $X_2 \not\perp\!\!\!\perp X_4 \mid X_3$
    - Problem: $X_1 \perp\!\!\!\perp X_2 \mid X_3$ is never calculated, only assumed given that $X_2 \perp\!\!\!\perp X_4 \mid \{\}$
    - Cascade error: wrong direction between X_4 and X_3



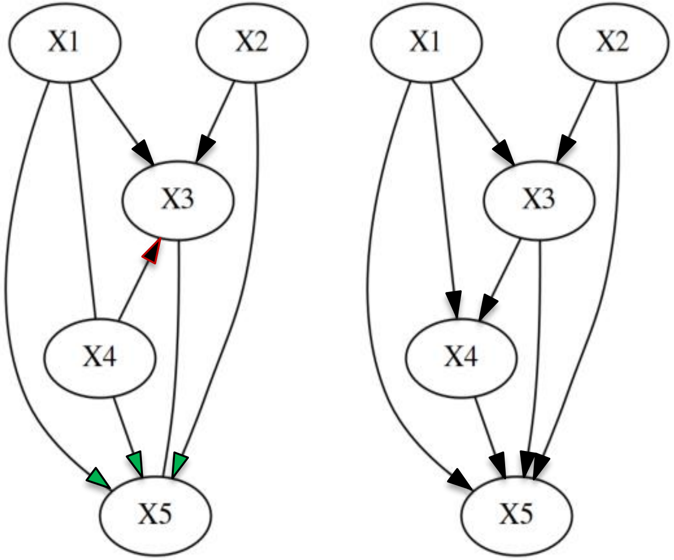
Estimated V-structure – colour arrowheads



True DAG

Step 2 – V-structures Estimation

- Consider the triple $\{X_1,X_5,X_2\}$
  - X_5 is not in the separating set of $\{X_1,X_2\}$ i.e. $X_1 \not\perp\!\!\!\perp X_2 \mid X_5$
    - Problem: $X_1 \perp\!\!\!\perp X_2 \mid X_5$ is never calculated, only assumed given that $X_1 \perp\!\!\!\perp X_2 \mid \{\}$
- Consider the triple $\{X_2,X_5,X_4\}$
  - X_5 is not in the separating set of $\{X_2,X_4\}$ i.e. $X_2 \not\perp\!\!\!\perp X_4 \mid X_5$
    - Problem: $X_2 \perp\!\!\!\perp X_4 \mid X_5$ is never calculated, only assumed given that $X_2 \perp\!\!\!\perp X_4 \mid \{\}$



Estimated V-structure – colour arrowheads

True DAG

Imperial College London

Step 3 – Orientation Rules

- None of Meek's rules apply to the given state of the graph
  - Cascade error: $X_1 \rightarrow X_4$ and $X_3 \rightarrow X_5$ if $X_3 \rightarrow X_4$ was correctly identified (Rule 2)

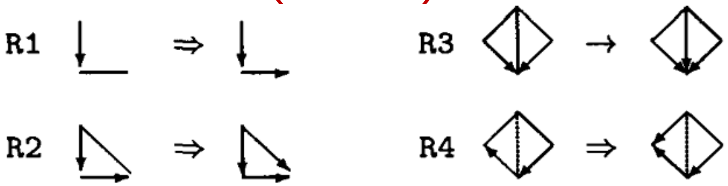
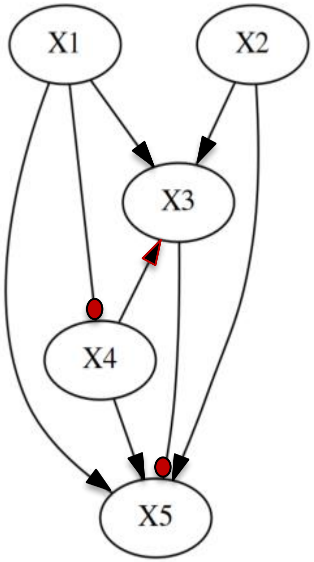
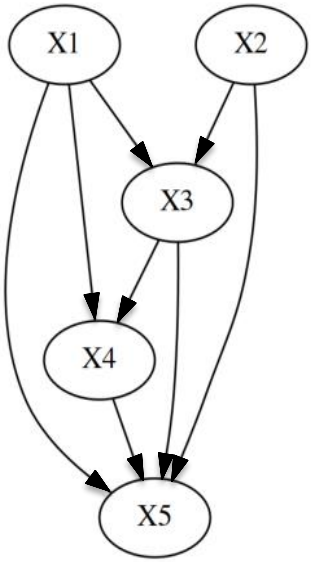


Figure 1: Orientation rules for patterns

From Meek, 1995. Dashed line: either direction. Solid line: undirected



Estimated V-structure – colour arrowheads



True DAG

Imperial College London

Independence Knowledge Base (IKB)

In certain situations where the experimenter knows that the underlying distribution belongs to the class of Bayesian Networks (acyclic and with independent error terms), it is appropriate to use the specific axioms of Eq. (6) instead of the general axioms of Eq. (5).

$$
\begin{array}{rl}
\textbf{(Symmetry)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \iff (\mathbf{Y} \perp\!\!\!\perp \mathbf{X} \mid \mathbf{Z}) \\
\textbf{(Composition)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) \\
\textbf{(Decomposition)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z}) \\
\textbf{(Intersection)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \mathbf{W}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z} \cup \mathbf{Y}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) \\
\textbf{(Weak Union)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \mathbf{W}) \\
\textbf{(Contraction)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z} \cup \mathbf{Y}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}) \\
\textbf{(Weak Transitivity)} & (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \gamma) \implies (\mathbf{X} \perp\!\!\!\perp \gamma \mid \mathbf{Z}) \vee (\gamma \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \\
\textbf{(Chordality)} & (\alpha \perp\!\!\!\perp \beta \mid \gamma \cup \delta) \wedge (\gamma \perp\!\!\!\perp \delta \mid \alpha \cup \beta) \implies (\alpha \perp\!\!\!\perp \beta \mid \gamma) \vee (\alpha \perp\!\!\!\perp \beta \mid \delta)
\end{array}
\qquad (6)
$$

Imperial College London

Propositionalization of Inference Rules

To construct the set of single-headed rules we consider, for each axiom, all possible contrapositive versions of it that have a single head. To illustrate, consider the Weak Transitivity axiom

$$(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \gamma) \implies (\mathbf{X} \perp\!\!\!\perp \gamma \mid \mathbf{Z}) \vee (\gamma \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z})$$

from which we obtain the following set of single-headed rules:

$$\begin{aligned}
(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \gamma) \wedge (\mathbf{X} \not\perp\!\!\!\perp \gamma \mid \mathbf{Z}) &\implies (\gamma \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \\
(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \gamma) \wedge (\gamma \not\perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) &\implies (\mathbf{X} \perp\!\!\!\perp \gamma \mid \mathbf{Z}) \\
(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \gamma) \wedge (\gamma \not\perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \not\perp\!\!\!\perp \gamma \mid \mathbf{Z}) &\implies (\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \\
(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\gamma \not\perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \not\perp\!\!\!\perp \gamma \mid \mathbf{Z}) &\implies (\mathbf{X} \not\perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \gamma).
\end{aligned}$$

Imperial College London

Propositionalization of Inference Rules

To obtain decomposed rules we apply the Decomposition axiom to every single-headed rule to produce only propositions over singletons. To illustrate, consider the Intersection axiom:

$$(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \mathbf{W}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z} \cup \mathbf{Y}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \cup \mathbf{W} \mid \mathbf{Z}).$$

In the above the consequent coincides with the antecedent of the Decomposition axiom, and we thus replace the Intersection axiom with a decomposed version:

$$(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z} \cup \mathbf{W}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z} \cup \mathbf{Y}) \implies (\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}) \wedge (\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z}).$$

Finally, note that it is easy to show that this rule is equivalent to two single-headed rules, one implying $(\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z})$ and the other implying $(\mathbf{X} \perp\!\!\!\perp \mathbf{W} \mid \mathbf{Z})$.