Computing lower and upper bounds on the probability of causal statements

Elena Sokolova, Martine Hoogman, Perry Groot, Tom Claassen, and Tom Heskes
Radboud University,
Nijmegen, The Netherlands

Reliability of a causal statement

How to estimate?

- Translate p-values to probabilities (Triantafilou, 2014)
- Based on Bayesian scores of DAGs (Claassen and Heskes, 2012)

Bayesian Constraint-based Causal Discovery:

Combination of Constraint base and Bayesian score base approach

Improves Constraint based approach (FCI) by using Bayesian approach to estimate the reliability of causal statements, avoiding propagation of unreliable decisions

T. Claassen, T. Heskes. **A Bayesian approach to constraint based causal inference**. In *UAI 2012*

BCCD

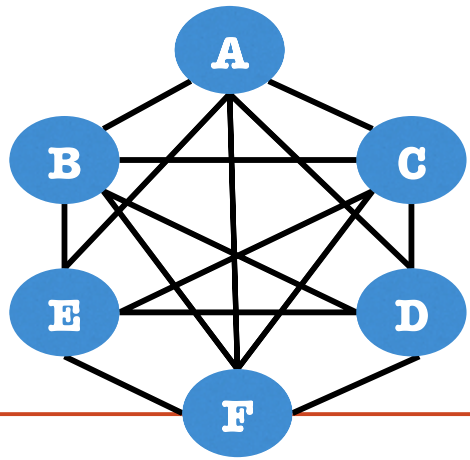
Basic idea:

- **Step 0** Start with a fully connected graph.
- **Step 1** Estimate the reliability of a causal statement ($X \Rightarrow Y$) using Bayesian score.
- **Step 2** If a causal statement declares a variable conditionally independent, delete an edge.
- **Step 3** Combine causal statements to infer new statements
- **Step 4** Rank all causal statements and orient edges in the graph.

BCCD

Estimate the reliability of a causal statement

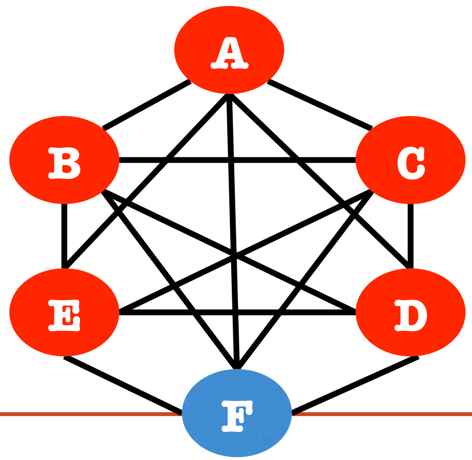
- Get possible subsets of K variables in the graph



BCCD

Estimate the reliability of a causal statement

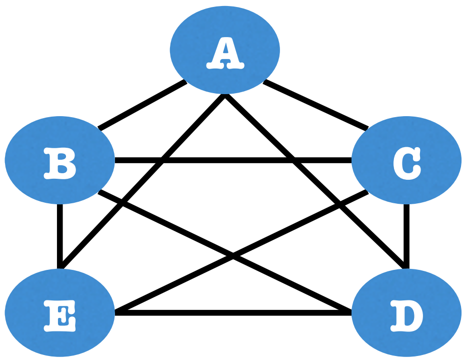
- Get possible subsets of K variables in the graph



BCCD

Estimate the reliability of a causal statement

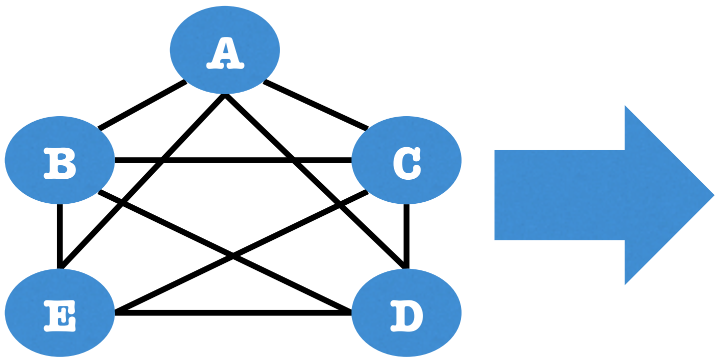
- Get possible subsets of K variables in the graph



BCCD

Estimate the reliability of a causal statement

- Get possible subsets of K variables in the graph
- Infer causal statement from this subgraph
- Estimate the reliability of the causal statement



$P(A \Rightarrow B) = 60\%$

$P(D \nRightarrow C) = 76\%$

$P(E \perp A \mid B) = 55\%$

BCCD

- Combine causal statements to infer new statements

List of statements:

$P(D \nRightarrow C)=76\%$

$P(A \Rightarrow B)=60\%$

$P(B \Rightarrow C)=60\%$

.....

'$A \Rightarrow B$'+'$B \Rightarrow C$' $\vdash$ '$A \Rightarrow C$'

What is the probability of new statement?

Estimation of the probability for new statement

- Logical statement- combination of OR and AND

- Fréchet inequalities

  Probability of an **intersection (AND)** of events
  - $\max(0, P(A) + P(B) - 1) \leq P(A \cap B) \leq \min(P(A), P(B))$

  Probability of a **union (OR)** of events
  - $\max(P(A), P(B)) \leq P(A \cup B) \leq \min(1, P(A) + P(B))$

Dependency between statement form and probability bound

Typical Statement form: $(L_1 \cup L_2) \cap (L_3 \cup L_1 \cap L_2)$

Possible standard forms to use:

- Disjunctive normal form (DNF): (… ∪ …) ∩(… ∪ …)- **best upper bound**

- Conjunctive normal form (CNF): (…∩…) ∪ (…∩…)- **best lower bound**

Method advantages and disadvantages

- Advantages:
  - More accurate estimate of probability
- Disadvantages:
  - Computationally expensive
  - Can explode (use approximation)

Approximation

‘A ⇒ B’+’B ⇒ C’ ⊢ ‘A ⇒ C’

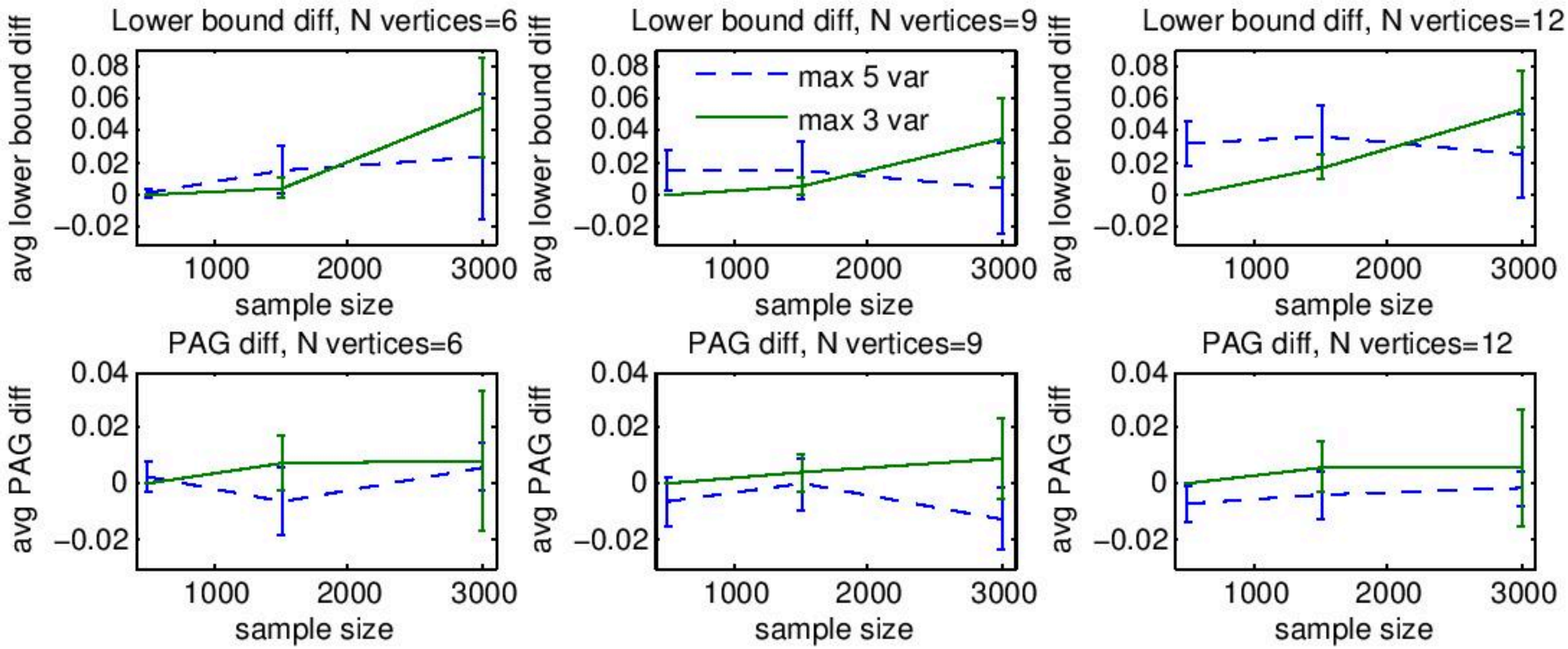
‘C ⇒ D’+’D ⇒ E’ ⊢ ‘C ⇒ E’

‘A ⇒ E’

Delayed: (‘A ⇒ B’+’B ⇒ C’) + (‘C ⇒ D’+’D ⇒ E)

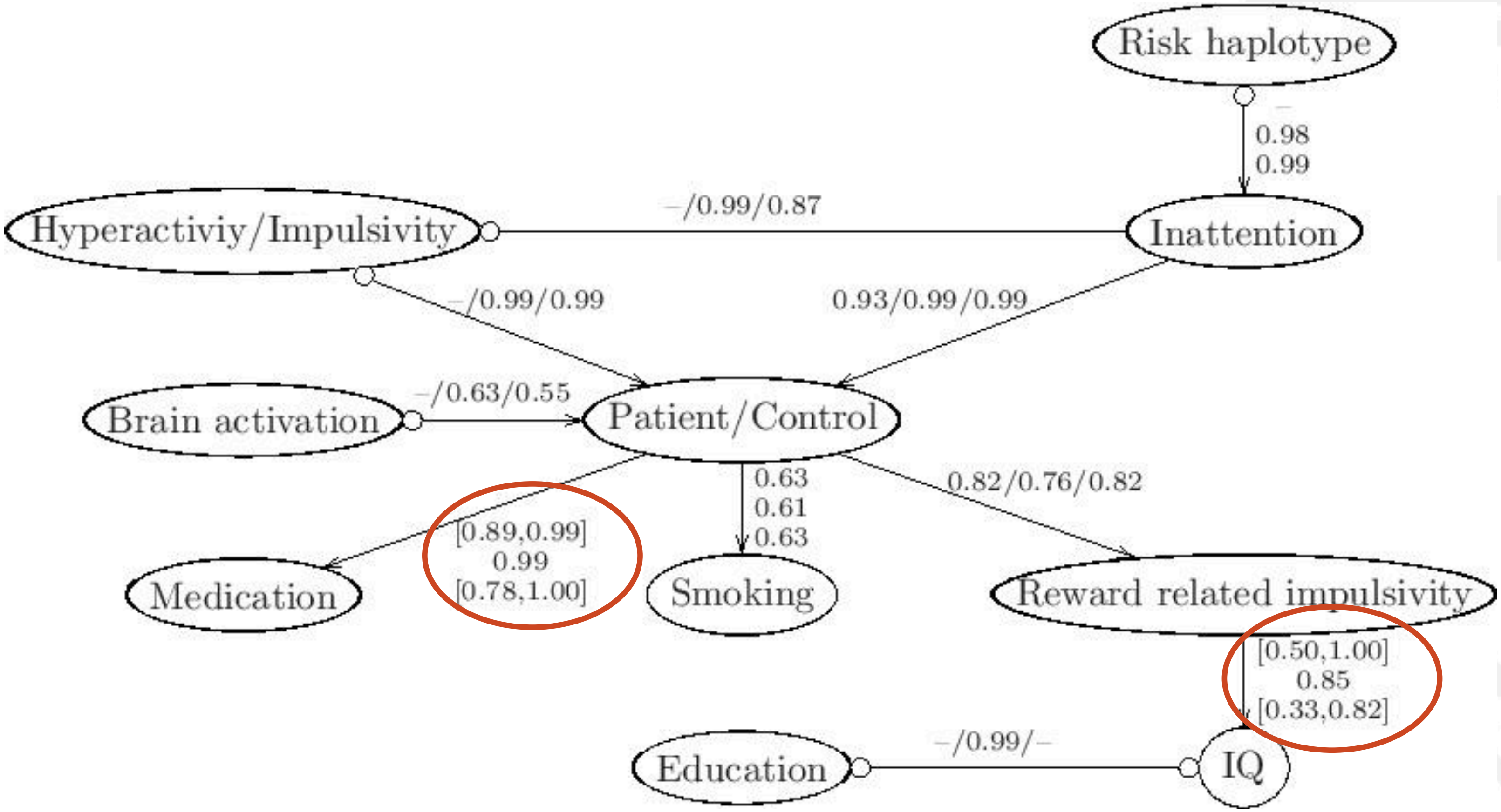
Greedy: ‘A ⇒ C’ + ‘C ⇒ E’

Simulation data: Impact of approximation



Causal Graph

A → **B**: A causes B
A ↔ **B**: latent common cause
A — **B**: selection bias
o: cannot distinguish between arrow and tail



Conclusions and Future work

- Provides guidelines for setting up new (intervention) experiments
- Use linear programming (SAT solvers) for more accurate bounds

Thank you for your attention!

Types of causal statements

- $A \to B$: causal effect from A to B
- $A \leftrightarrow B$: *latent common cause*
- $A - B$: *Selection bias*

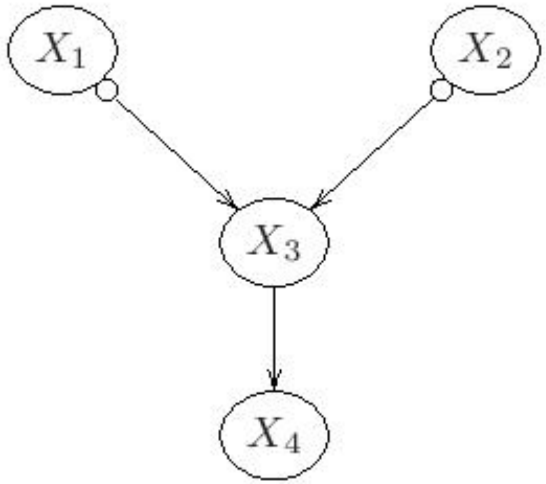
To decide $A \to B$ we should infer:

- "$A \Rightarrow B$" (*A causes B, tail)*
- "$B \not\Rightarrow A$" (*B does not cause A, arrow*)

When partially unclear use ●:

- *$A \bullet\!\!\to B$ is either $A \to B$ or $A \leftrightarrow B$*

Example Y-structure



- K=3

0^{st} *level:* $\Theta_1 : [X_4 \nRightarrow X_3]$

1^{st} *level:* $\Gamma_1 : [(X_4 \nRightarrow X_2) \wedge (X_4 \nRightarrow S)] \quad \wedge \quad \Gamma_2 : [(X_3 \Rightarrow X_4)]$

2^{nd} *level:* $\Psi_1 : [(X_3 \Rightarrow X_2) \vee (X_3 \Rightarrow X_4) \vee (X_3 \Rightarrow S)] \quad \wedge \Psi_2 : [(X_3 \nRightarrow X_2) \wedge (X_3 \nRightarrow S)]$