When are Graphical Causal Models not Good Models?

CAPITS 2008

Jan Lemeire
September 12th 2008

Vrije Universiteit Brussel

Overview

- The Kolmogorov Minimal Sufficient Statistic (KMSS).
- Bayesian Networks as Minimal Descriptions of Probability Distributions.
- When the minimal Bayesian network is the KMSS.
- When the minimal Bayesian network is NOT the KMSS.

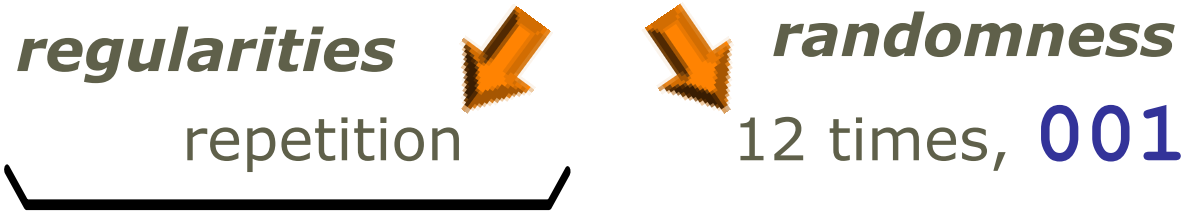
Overview

- **The Kolmogorov Minimal Sufficient Statistic (KMSS).**
- Bayesian Networks as Minimal Descriptions of Probability Distributions.
- When the minimal Bayesian network is the KMSS.
- When the minimal Bayesian network is NOT the KMSS.

Randomness versus Regularity

- **0010010010010010010010010010010010010010**

  Regular string: compressible, low complexity



Model that minimally describes regularities (qualitative props)

= *Kolmogorov Minimal Sufficient Statistic (KMSS)*

- **0110001101011010101110010011010001011 10**

  Random string: incompressible, maximal complexity

  But: it is no *meaningful* information, only *accidental* information

Overview

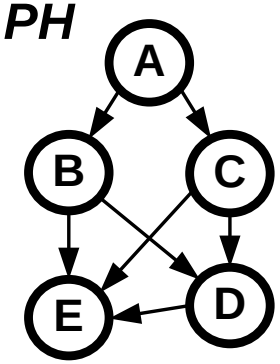
- The Kolmogorov Minimal Sufficient Statistic (KMSS).
- **Bayesian Networks as Minimal Descriptions of Probability Distributions.**
- When the minimal Bayesian network is the KMSS.
- When the minimal Bayesian network is NOT the KMSS.

Description of Probability Distributions with Bayesian networks

Joint Probability Distribution
P(A, B, C, D, E)

=

GRAPH

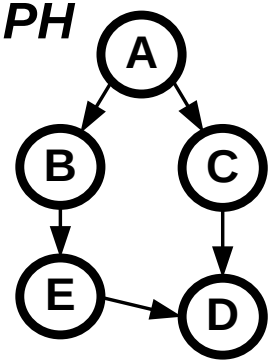


CPDs

P(A)
P(B|A)
P(C|A)
P(D|B, C)
P(E|B, C, D)

=

GRAPH



CPDs

P(A)
P(B|A)
P(C|A)
P(E|B)
P(D|C, E)

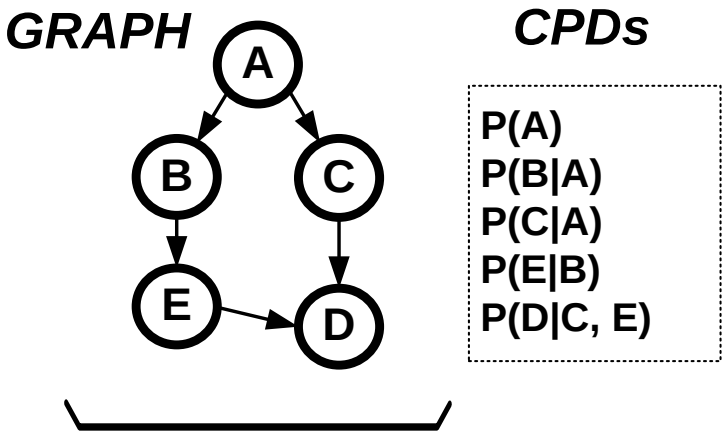
Multiple Bayesian networks

➡ select minimal one

Meaningful Information of Probability Distributions

Joint Probability Distribution P(A, B, C, D, E)

=



meaningful information

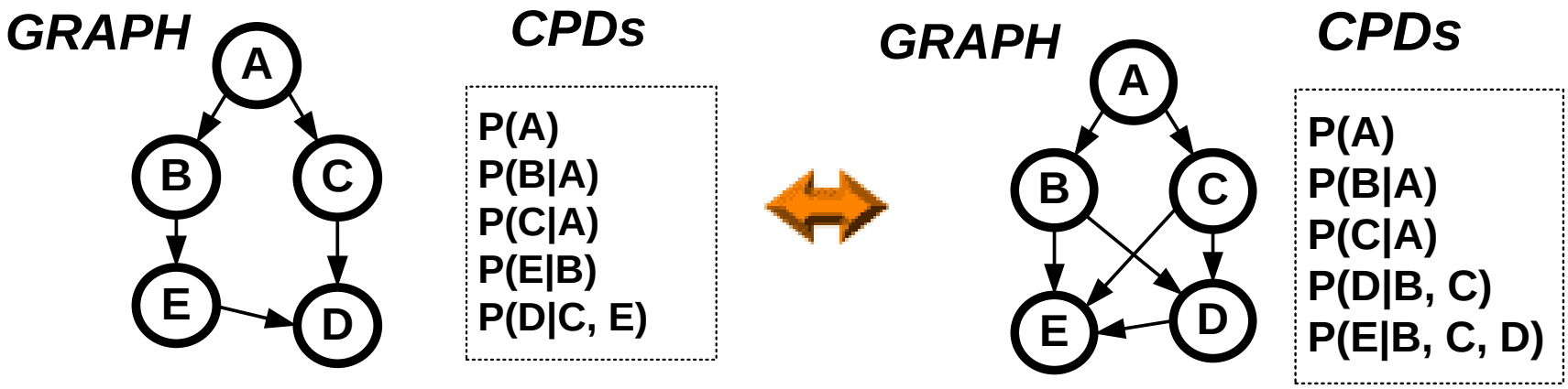
- Regularities: Conditional Independencies
- **Kolmogorov Minimal Sufficient Statistic** if graph and CPDs are incompressible

Representation of Independencies

- Graph (DAG) of a Bayesian network is a description of conditional independencies
- Faithfulness: *All* conditional independencies of the distribution are described by the graph.
- Theorem: If a faithful Bayesian network exists for a distribution, it is the minimal Bayesian network.

Regularities cause unfaithfulness

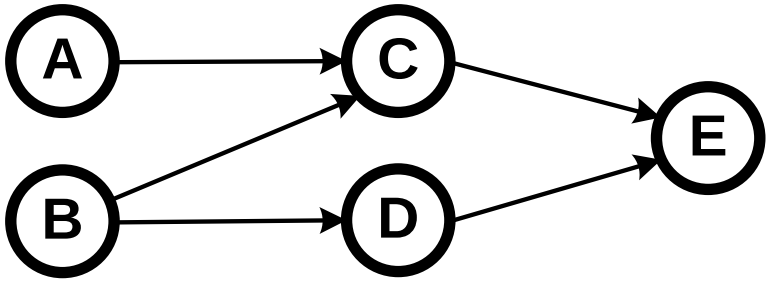
- Theorem: A Bayesian network for which the concatenation of the CPDs is incompressible, is faithful.



Overview

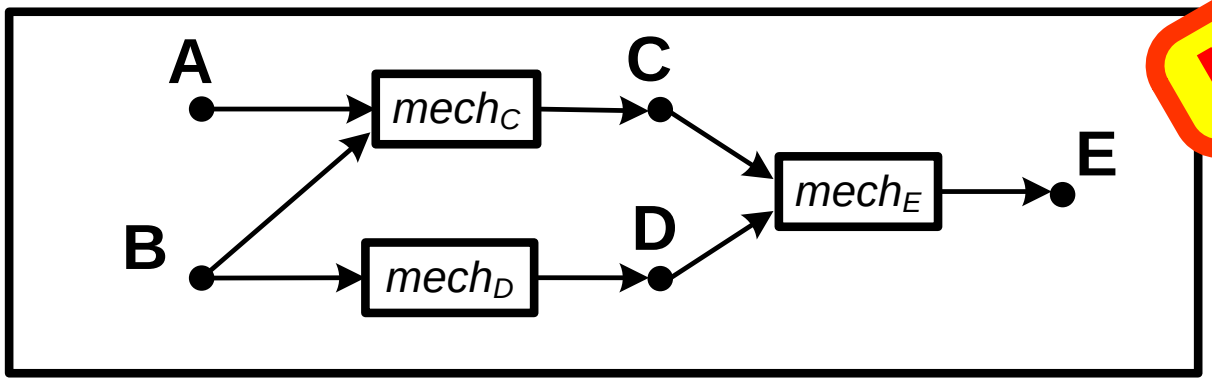
- The Kolmogorov Minimal Sufficient Statistic (KMSS).
- Bayesian Networks as Minimal Descriptions of Probability Distributions.
- **When the minimal Bayesian network is the KMSS.**
- When the minimal Bayesian network is NOT the KMSS.

When the minimal Bayesian network is the KMSS.



- No other regularities than the independencies present in graph
- Model is faithful
- KMSS = DAG
- ➡ quasi-unique and minimal decomposition of the system
- CPDs are independent

The Top-Ranked Hypothesis



Number 1!!

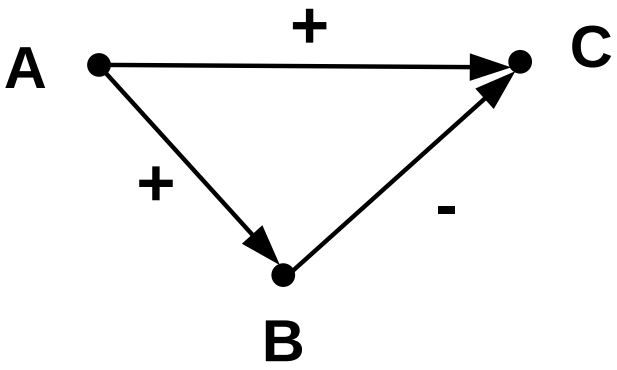
Comes with certificate

Highly recommended

- **Each CPD corresponds to an independent part of reality, a mechanism**
  - = Modularity and autonomy
  - ➡ possibility to predict the effect of changes to the system (interventions)

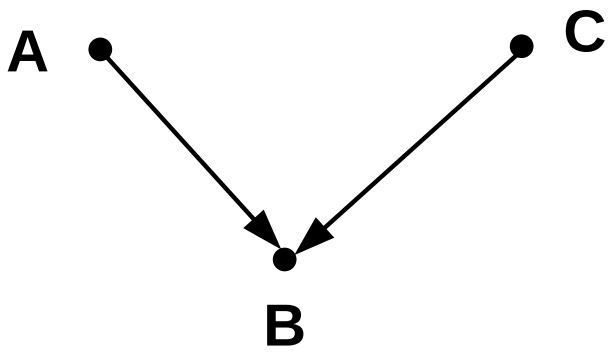
- Causal component = Reductionism

Except… World can be More Complex



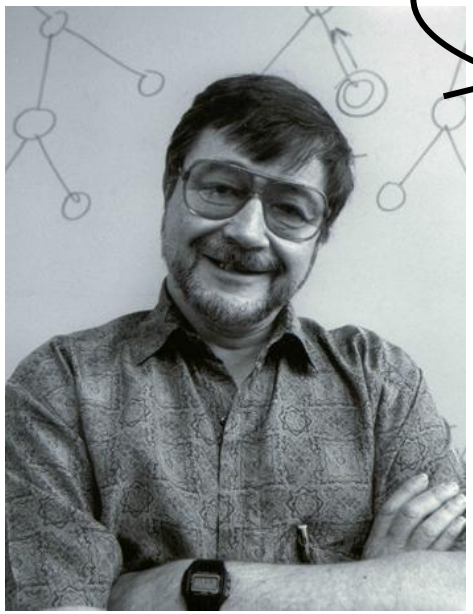
$A \perp\!\!\!\perp C$

True Model



Minimal Model

There is no absolute guarantee, the KMSS might be a bit too simplistic



Great job, Jan!

Overview

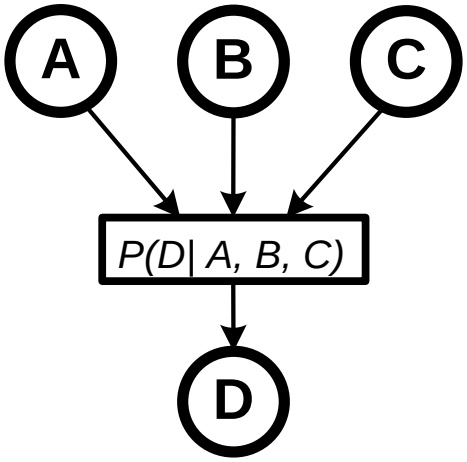
- The Kolmogorov Minimal Sufficient Statistic (KMSS).
- Bayesian Networks as Minimal Descriptions of Probability Distributions.
- When the minimal Bayesian network is the KMSS.
- **When the minimal Bayesian network is NOT the KMSS.**

When the minimal Bayesian network is NOT the KMSS…

- There are non-modeled regularities, DAG+CPDs is compressible
  A. Compressibility of an individual CPD
  B. Compressibility of a set of CPDs

- *Case studies*:
  - True Causal Model in set of minimal Bayesian networks?
  - Faithfulness?
  - Modularity?

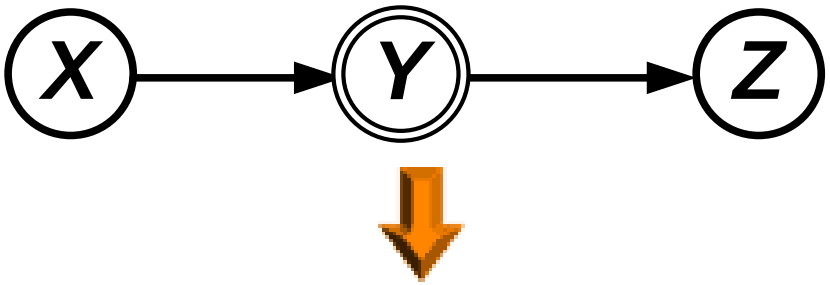
(A1) Local Structure



A	B	C	P(D|A, B, C)
0	0	0	**0.4**
0	0	1	**0.6**
0	1	0	**0.7**
0	1	1	**0.7**
1	0	0	**0.3**
1	0	1	**0.3**
1	1	0	**0.3**
1	1	1	**0.3**

- A single CPD can be described shorter, without affecting the rest of the model
  - Local structure [Friedman and Goldszmidt, 1996]
  - Context-specific independencies [Boutilier, 1996]
- Everything OK.

(A2) Deterministic Relations

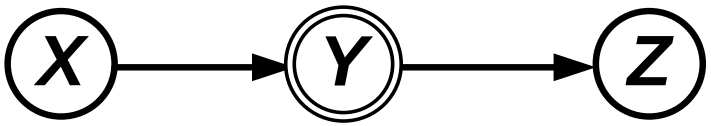


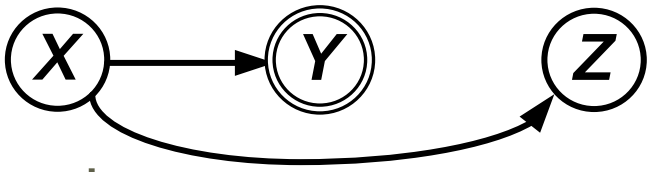
$Y=f(X)$

$X \perp\!\!\!\perp Z \mid Y$

$Y \perp\!\!\!\perp Z \mid X$

Violation of the intersection condition





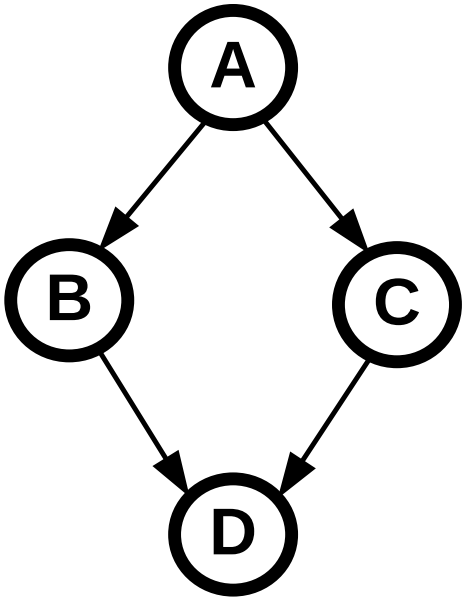
Two minimal Bayesian networks.

Both unfaithful

Conclusions for Compressibility of an individual CPD

- CPDs are independent
- ➡ modularity is still plausible 
- True Causal Model in set of minimal Bayesian networks! 
- But faithfulness may become invalid 
  - Constraint-based algorithms may fail 

(B1) Meta-mechanism



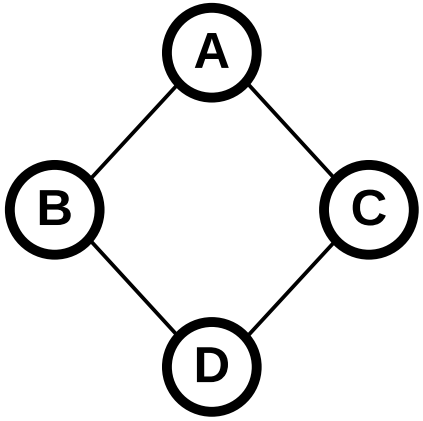
$D \perp\!\!\!\perp A$

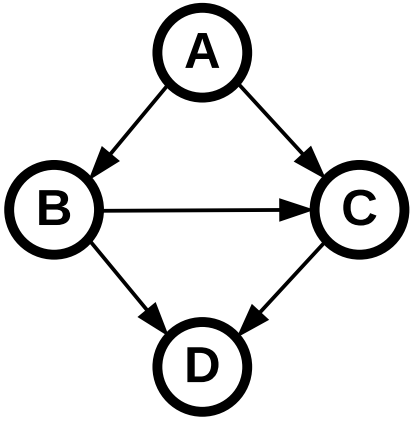
- Influence A->B->D and A->C->D exactly balance
- ➡ Unfaithfulness
- Learned correctly

- We may assume a global mechanism that controls mechanisms such that they neutralize
  - E.g. evolution

(B2) Markov Networks

- Different model class!!





One of the minimal Bayesian Networks.
Unfaithful & non-modular

Conclusions

- Faithfulness cannot be guaranteed
- Modularity cannot be guaranteed when dependent CPDs
- Regularities/Model class under consideration must be properly chosen
  - Augmentation of Bayesian networks with other qualitative properties
- *Faithfulness* = ability of a model to explicitly explain all regularities of the data