Learning Causal Graphical Models with Latent Variables

Sam Maes, Stijn Meganck, Philippe Leray
sammaes@vub.ac.be

Laboratoire d'Informatique, Traitement de l'Information, Systèmes.

LEARNING'06
EPSEVG

02/10/2006 Vilanova i la Geltrú

Overview

Overview

Introduction

Subtasks of causal modeling with latent variables:

- Structure learning from:
  - observational data
  - experimental data
- Learning parameters
- Probabilistic inference
- Causal inference

Problem

No integral approach for all these subtasks in the presence of latent variables.

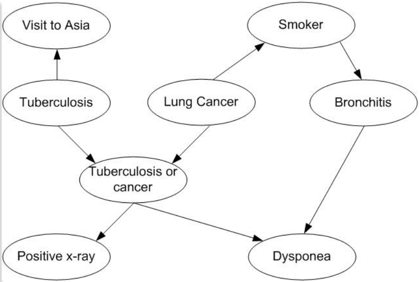
- Causal inference: **semi-Markovian causal models**
- Structure learning from observational data: **ancestral graphs**

Overview

Bayesian Networks (BN)

Probabilistic graphical model

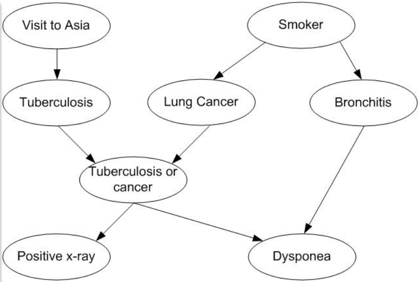
- Independence model
- Probability distribution
  - $P(V) = \prod_{X \in V} P(X|Pa(X))$
- Probabilistic inference
  - Predict consequence of observation
  - $P(\text{cancer}|\text{smoker,x-ray})$



Causal Bayesian networks (CBN)

Causal probabilistic graphical model

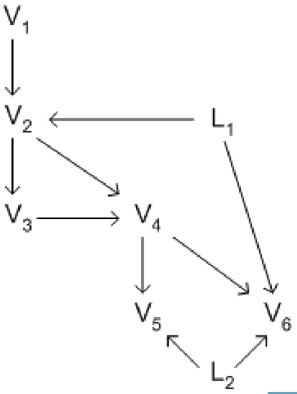
- BN where every $\rightarrow$ corresponds to an immediate causal relation:
  - $X \rightarrow Y$: manipulating variable X changes the distribution of variable Y
- Causal inference
  - Predict consequence of manipulation
  - $P(\text{cancer}|\text{do}(\text{TBC}=\text{true}))$



Modeling Latent Variables

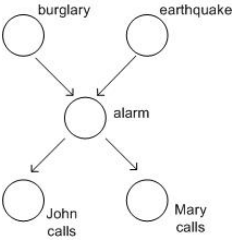
Some variables unobserved

- Model variables implicitly,
  i.e. no estimating of cardinality and distributions.
- Two main approaches:
  - Semi-Markovian Causal Models
  - Maximal Ancestral Graphs



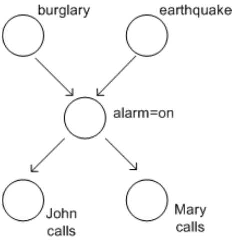
Probabilistic vs Causal Inference

original model



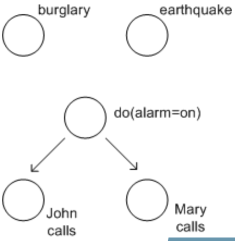
after observation

- instantiate the observed variables
- propagate



after manipulation

- replace the former causes
- instantiate
- propagate



With latent variables

Causal inference becomes more complicated:

- **replace the former causes**
- instantiate
- propagate

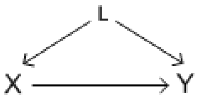
$$\begin{tikzcd} & L \arrow[dl] \arrow[dr] & \\ X \arrow[r] & Z \arrow[r] & Y \end{tikzcd}$$

$P(Y = y|do(X = x))$: manipulate variable X and study the effect on Y.

Overview

Our assumptions

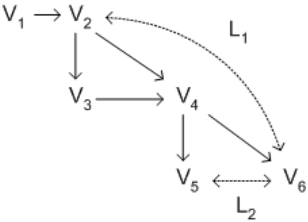
- stability: independences are structural
- max. 1 immediate connection between any 2 variables in the underlying DAG
- no selection bias
- discrete variables



Representation for causal inference

semi-Markovian causal models (SMCM)

- directed edge represents an immediate causal relation
- bi-directed edge represents a latent common cause
- importance: every model with arbitrary latent variables can be transformed into a SMCM
- a joint probability distribution: e.g. $P(V_1, \ldots, V_6)$



Inference in SMCMs

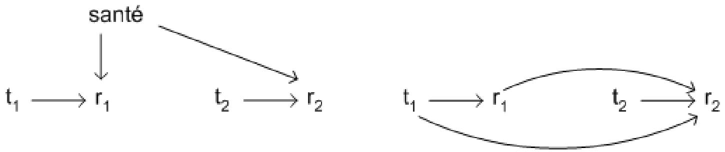
causal inference algorithm exists (Tian & Pearl), but:

- no efficient parametrisation
- no probabilistic inference algorithm
- no learning algorithm

Representation for learning

The classs of DAGs is not complete under marginalisation of latent variables.

I.e., a DAG of the observable variables can not represent exactly all the independences present between the variables.

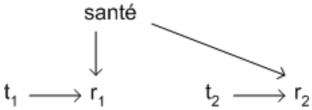


Maximal Ancestral Graphs (MAG)

Maximal ancestral graph without conditioning

Graph with:

- directed edges: have an ancestral meaning $\neq$ causal
- bi-directed edges: represent latent common cause
- maximum 1 edge between 2 variables:
  ancestral relation absorbs latent common cause
- every absent edge represents an independence relation



Learning from Observational Data

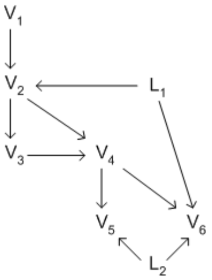
Constraint-based algorithm

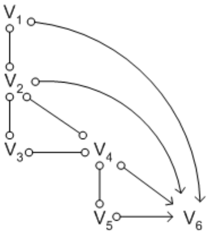
- Learns from independencies in the data
- Fast Causal Inference (FCI) algorithm
- Rules for orienting edges
- **Problem:** only learns upto Markov equivalence

Markov Equivalence Class

Complete partial ancestral graph (CPAG):

- 3 possible edge marks: o, -, >
- how to fill in the edges with o if we want the causal model ?

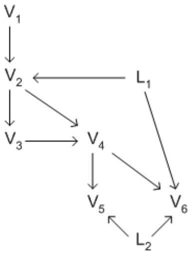


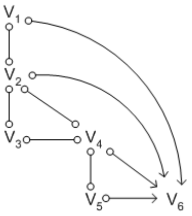


Uncertainty in CPAGs

3 possible underlying explanations for each edge:

- immediate causal relation $V_1 \longrightarrow V_2$
- latent variable $V_2 \leftrightarrow V_6$
- "inducing path" between V_1 et V_6
  - V_1 can not be separated from V_6 by conditioning on observable variables
  - observationally it seems that there is an immediate connection





Inference in MAGs

- only learning upto Markov equivalence class
- limited causal inference algorithm (only those causal expressions that are equal for the complete equivalence class)
- no probabilistic inference
- no parametrisation for discrete variables

Therefore:

Use observational learning to learn a CPAG, then use experiments to transform into a SMCM to allow inference.

Overview

CPAG $\longrightarrow$ SMCM

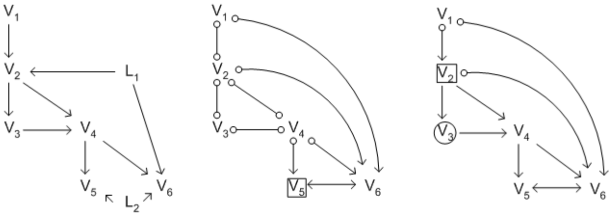
Perform experiments to differentiate between the different cases:

- Type 1: resolve $o\!\!\longrightarrow$
- Type 2: resolve o—o
- Remove edges due to inducing paths

CPAG $\rightarrow$ SMCM (Type 1)

Type 1: resolve $\circ\!\!\rightarrow$

- $exp(A) \not\rightsquigarrow B$: $A \leftrightarrow B$
- $exp(A) \rightsquigarrow B$:
  - $\nexists$pot. dir. path $A \dashrightarrow B$ of length ≥ 2: $A \rightarrow B$
  - $\exists$pot. dir. path $A \dashrightarrow B$ of length ≥ 2:
    block each pot. dir. path by conditioning on a set D
    - $exp(A)|D \rightsquigarrow B$: $A \rightarrow B$
    - $exp(A)|D \not\rightsquigarrow B$: $A \leftrightarrow B$



CPAG $\longrightarrow$ SMCM (Type 2)

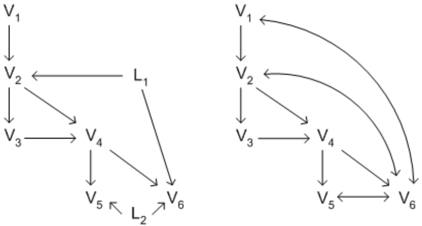
Type 2: resolve $o\!-\!o$

- easily transformed into Type 1

CPAG $\longrightarrow$ SMCM (ctd.)

Remove edges due to inducing paths

- recognize the edges $A \leftrightarrow B$ or $A \rightarrow B$ possibly created due to an inducing path
- block each inducing path between A, B with experiments E
- block each other path between A, B by conditioning on D
- $exp(E)|D$:
  - no dependence in the exper. data: remove the arc
  - still dependence in the exper. data: leave the arc



Conclusion

Conclusion

- an approach to learn SMCMs from a combination of observational and experimental data

Other and future work

- parametrisation for SMCMs
- optimise the order of the experiments wrt several decision criteria
- infer edges after each experiment