

L'inférence Causale dans des Modèles Graphiques avec des Variables Latentes

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

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

JFRB '06
LAMIH

18/05/2006 Valenciennes



- 1 Introduction
- 2 Inférence Causale
- 3 Apprentissage
- 4 Paramétrisation

- 1 Introduction
- 2 Inférence Causale
- 3 Apprentissage
- 4 Paramétrisation

Composants de la modélisation causale avec variables latentes:

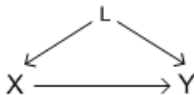
- Apprentissage de structure à base de données:
 - observables
 - expérimentales
- Apprentissage de paramètres
- Inférence probabiliste
- Inférence causale

Pas d'approche intégrale pour tout ces composants dans la présence de variables latentes.

- Inférence causale:
modèles causaux semi-Markoviens
- Apprentissage de structure à base de données observables:
graphs ancestraux

Hypothèses

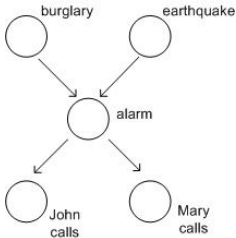
- stabilité: les indépendences sont structurelles
- max. une connection immédiate par 2 variables
- pas de biais de sélection



- 1 Introduction
- 2 Inférence Causale**
- 3 Apprentissage
- 4 Paramétrisation

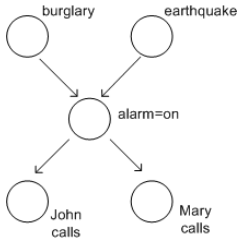
Inférence Probabiliste vs Causale

modèle original



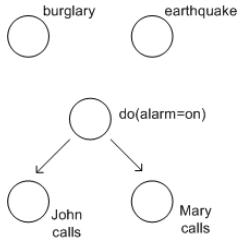
après observation

- instantier les variables observées
- propager



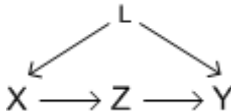
après manipulation

- remplacer les anciennes causes
- instantier
- propager



L'inférence causale devient plus compliquée:

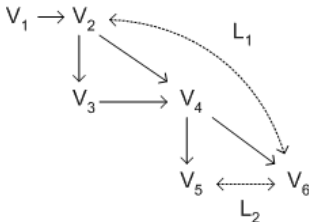
- remplacer les anciennes causes
- instantier
- propager



$P(Y = y | do(X = x))$: manipuler variable X et étudier l'effet sur Y.

modèles causaux semi-Markoviens (SMCM)

- arc dirigé représente une relation causale autonome
- arc bi-dirigé représente une cause commune latente
- importance: chaque modèle avec variables latentes arbitraires peut être transformé en un SMCM
- une loi jointe: e.g. $P(V_1, \dots, V_6)$



algorithme d'inférence causale existe (Tian & Pearl), mais:

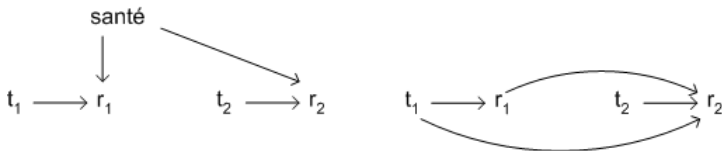
- pas de paramétrisation efficace
- pas d'algorithme d'inférence probabiliste
- pas d'algorithme d'apprentissage

Plan

- 1 Introduction
- 2 Inférence Causale
- 3 Apprentissage**
- 4 Paramétrisation

La classe des DAGs n'est pas complète sous marginalisation.

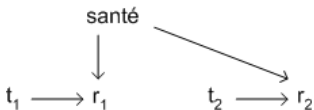
I.e., un DAG des variables observables ne peut pas représenter exactement toutes les indépendances présentes entre les variables.



Graphe ancestral maximal sans conditionnement

Graphe avec des arcs:

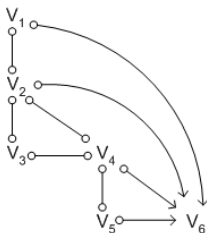
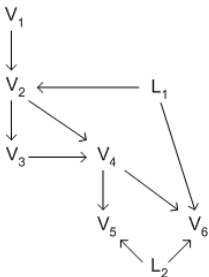
- dirigés qui ont une signification ancestrale \neq causale
- bi-dirigés
- maximum 1 arc entre 2 variables
- chaque arc absent représente une relation d'indépendance



Classe d'équivalence de Markov

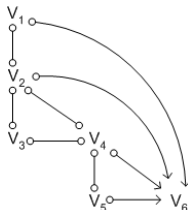
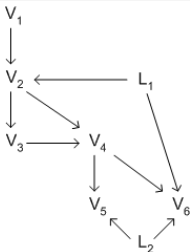
Graph complet partiellement ancestral (CPAG):

- Fast Causal Inference (FCI)
- Règles pour orienter des arcs
- 3 extrémités d'arcs possibles: \circ , $-$, $>$



3 explications possibles pour chaque arc:

- relation causale $V_1 \rightarrow V_2$
- variable latente $V_2 \leftrightarrow V_6$
- "inducing path" entre V_1 et V_6
 - V_1 ne peut être séparé de V_6 en utilisant des variables observées
 - à cause de la maximalité des modèles FCI trouve un arc



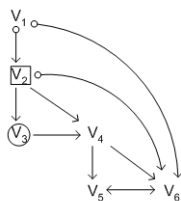
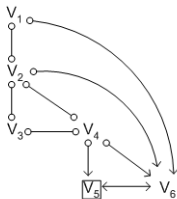
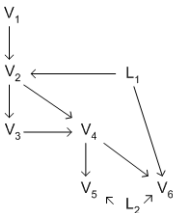
Exécuter des expériences pour différentier entre les cas:

- Type 1: résoudre $o \rightarrow$
- Type 2: résoudre $o - o$
- Enlever les arcs en trop

CPAG \rightarrow SMCM (Type 1)

Type 1: résoudre $o \rightarrow$

- $exp(A) \not\sim B: A \leftrightarrow B$
- $exp(A) \rightsquigarrow B:$
 - \nexists chemin pot.dir. $A \dashrightarrow B$ de longueur $\geq 2: A \rightarrow B$
 - \exists chemin pot.dir. $A \dashrightarrow B$ de longueur $\geq 2:$
bloque chaque chemin part.dir. en conditionnant sur un ensemble D
 - $exp(A)|D \rightsquigarrow B: A \rightarrow B$
 - $exp(A)|D \not\sim B: A \leftrightarrow B$



CPAG → SMCM (Type 2)

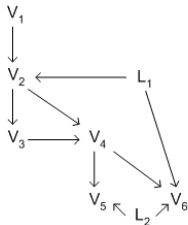
Type 2: résoudre $o-o$

- Facilement transformé en Type 1

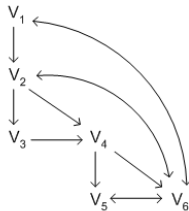
CPAG → SMCM (ctd.)

Enlever les arcs en trop

- reconnaître les arcs $A \leftrightarrow B$ où $A \rightarrow B$ possiblement créé par un "inducing path"
- bloquer chaque "inducing path" entre A, B avec des expér. E
- bloquer chaque autre chemin entre A, B en condit. sur un D
- $exp(E)|D$:
 - $cov(A, B) \neq 0$: laisser l'arc
 - $cov(A, B) = 0$: enlever l'arc



Sam Maes



L'inférence Causale

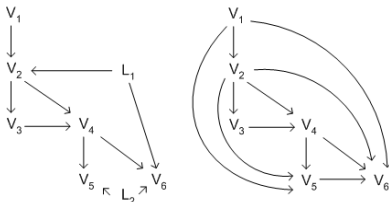
Plan

- 1 Introduction
- 2 Inférence Causale
- 3 Apprentissage
- 4 Paramétrisation**

Paramétrisation discrète

Représenter le SMCM avec un DAG

- ce DAG est un I-map du SMCM
- i.e. toutes les indép. du DAG sont aussi dans le SMCM
- utiliser ce DAG pour:
 - apprendre des paramètres
 - l'inférence probabiliste
- pour l'inférence causale:
 - utiliser le SMCM pour la structure causale
 - utiliser le DAG pour les paramètres



- optimiser l'ordre des expériences par rapport à plusieurs critères de décisions
- étudier l'optimalité de la paramétrisation