A Characterization of Markov Equivalence Classes of Relational Causal Models under Path Semantics

Sanghack Lee
joint work with Vasant Honavar

Penn State University

June 27, UAI 2016
Overview

- Relational Causal Model (RCM, Maier et al. 2010) is
  - a generalization of Causal Bayesian Network (CBN, causal DAG)
  - one of relational models (between PRM & DAPER).

- Generalized
  - (causal) Markov condition, (causal) faithfulness
  - d-separation

- Characterization of Markov equivalence of RCM
  - When do two RCMs yield the same independence relations?
  - Generalized existing ideas for Markov equivalence of DAG.

- Basis for a sound and complete causal discovery algorithm
Overview

- **Relational Causal Model** (RCM, Maier et al. 2010) is
  - a generalization of **Causal Bayesian Network** (CBN, causal DAG)
  - one of relational models (between PRM & DAPER).

- **Generalized**
  - (causal) Markov condition, (causal) faithfulness
  - d-separation

- Characterization of Markov equivalence of RCM
  - When do two RCMs yield the same independence relations?
  - Generalized existing ideas for Markov equivalence of DAG.

- Basis for a sound and complete causal discovery algorithm
Overview

- **Relational Causal Model** (RCM, Maier et al. 2010) is
  - a generalization of **Causal Bayesian Network** (CBN, causal DAG)
  - one of relational models (between PRM & DAPER).

- **Generalized**
  - (causal) Markov condition, (causal) faithfulness
  - d-separation

- **Characterization of Markov equivalence of RCM**
  - When do two RCMs yield the same independence relations?
  - Generalized existing ideas for Markov equivalence of DAG.

- Basis for a sound and complete causal discovery algorithm
Overview

- Relational Causal Model (RCM, Maier et al. 2010) is
  - a generalization of Causal Bayesian Network (CBN, causal DAG)
  - one of relational models (between PRM & DAPER).

- Generalized
  - (causal) Markov condition, (causal) faithfulness
  - d-separation

- Characterization of Markov equivalence of RCM
  - When do two RCMs yield the same independence relations?
  - Generalized existing ideas for Markov equivalence of DAG.

- Basis for a sound and complete causal discovery algorithm
BACKGROUND

- Relational Schema $\mathcal{S}$
- Relational Skeleton $\sigma$
- Relational Causal Model $\mathcal{M}$
- Ground Graph $\mathcal{G}_\mathcal{M}$
Relational Schema $\mathcal{S}$

$\mathcal{S} = (\mathcal{E}, \mathcal{R}, \mathcal{A}, \text{card})$

Entity classes $\mathcal{E}$, Relationship classes $\mathcal{R}$, Attribute classes $\mathcal{A}$

Cardinality constraints, $\mathcal{R} \times \mathcal{E} \rightarrow \{\text{one}, \text{many}\}$

Maier [2014]
Relational Skeleton $\sigma \in \Sigma_S$

- an instance of the given relational schema $S$
  - $\Sigma_S$, all possible instantiations

- an undirected bipartite graph
  - node = item (i.e., entity or relationship, $i$, $j$)
  - edge = the participation of an entity in a relationship

modified from Maier [2014]
Relational Skeleton $\sigma \in \Sigma_S$

- an instance of the given relational schema $S$
  - $\Sigma_S$, all possible instantiations

- an undirected bipartite graph
  - node = item (i.e., entity or relationship, $i$, $j$)
  - edge = the participation of an entity in a relationship

modified from Maier [2014]
Relational Causal Model

\[ \mathcal{M} = (S, D, \Theta) \]
with a set of relational dependencies \( D \),
and relevant functions or parameters \( \Theta \)
Relational Causal Model

\[ M = (S, D, \Theta) \]

with a set of relational dependencies \( D \),
and relevant functions or parameters \( \Theta \)

**relational dependency**

*Success* of a product depends on the *Competence* of its developer(s).

\[ \text{[Product, Develops, Employee]} \cdot \text{Competence} \rightarrow \text{[Product].Success} \]

*relational path*
Relational Causal Model

\[ \mathcal{M} = (S, D, \Theta) \]
with a set of relational dependencies \( D \),
and relevant functions or parameters \( \Theta \)

relational dependency

Success of a product depends on the Competence of its developer(s).

\[ [Product, Develops, Employee] \cdot Competence \to \forall Success \]

relational path
Relational Causal Model

\[ M = (S, D, \Theta) \]

with a set of relational dependencies \( D \),
and relevant functions or parameters \( \Theta \)

Maier [2014]
Relational Causal Model: Class Dependency Graph

\[ M = (S, D, \Theta) \]

with a set of relational dependencies \( D \),
and relevant functions or parameters \( \Theta \)

Class Dependency Graph \( G^M_A \)

acyclicity of an RCM
\[ = \text{acyclicity of its CDG} \]
\[ = A \text{ is partially-ordered.} \]
Ground Graph $G^M_\sigma$

- is an instance of an RCM $M$ given a relational skeleton $\sigma$
- is a CBN of item-attributes (e.g., $i.X$, $paul.Salary$)

**instantiating relational dependencies**

$$j.Y \rightarrow i.X \in G^M_\sigma \quad \text{if} \; \exists P. Y \rightarrow \forall X \in D \; \text{and} \; j \in P|_{i}^\sigma$$
Ground Graph $G^M_\sigma$

- is an instance of an RCM $\mathcal{M}$ given a relational skeleton $\sigma$
- is a CBN of item-attributes (e.g., $i.X$, $paul.Salary$)

instantiating $[E, D, P, F, B].\text{Budget} \rightarrow \forall \text{Salary} @ paul$
Ground Graph $\mathcal{G}_\sigma^\mathcal{M}$

- is an instance of an RCM $\mathcal{M}$ given a relational skeleton $\sigma$
- is a CBN of item-attributes (e.g., $i.X$, $paul.Salary$)

instantiating $[E, D, P, F, B].\text{Budget} \rightarrow \forall \text{Salary} @ paul$

\[
\{\text{accessories, devices}\} = [E, D, P, F, B]|_{paul}^\sigma
\]
Ground Graph $G^M_\sigma$: Path Semantics

- is an instance of an RCM $M$ given a relational skeleton $\sigma$
- is a CBN of item-attributes (e.g., $i.X$, $paul\cdot Salary$)

instantiating $[E, D, P, F, B]. Budget \rightarrow \forall Salary @ paul$

$accessories \in [E, D, P, F, B] |_{paul}$
Ground Graph $\mathcal{G}_\sigma^M$ : Path Semantics

- is an instance of an RCM $\mathcal{M}$ given a relational skeleton $\sigma$
- is a CBN of item-attributes (e.g., $i.X$, $paul.Salary$)

instantiating $[E, D, P, F, B].Budget \rightarrow \forall Salary @ paul$

$devices \in [E, D, P, F, B]_\sigma |_{paul}$
Ground Graph $\mathcal{G}_\sigma^M$

- is an instance of an RCM $\mathcal{M}$ given a relational skeleton $\sigma$
- is a CBN of item-attributes (e.g., $i.X$, $paul.Salary$)

**instantiating relational dependencies**

\[ j.Y \rightarrow i.X \in \mathcal{G}_\sigma^M \quad \text{if} \quad \exists P. Y \rightarrow \forall X \in D \text{ and } j \in P_{\mid \sigma} \]
Relational Schema \(\rightarrow\) Relational Causal Model

Relational Skeleton(s) \(\rightarrow\) Ground Graph(s)

given

instantiated
MARKOV EQUIVALENCE of RCMs
Two DAGs $\mathcal{G}$ and $\mathcal{G}'$ are equivalent under Markov condition, $[\mathcal{G}] = [\mathcal{G}']$, if they entail the same independence relations (= d-separation).
Markov Equivalence of DAG: Review

Two DAGs $\mathcal{G}$ and $\mathcal{G}'$ are equivalent under Markov condition, $[\mathcal{G}] = [\mathcal{G}']$, if they entail the same independence relations (= d-separation).

\[ \text{DAG } \mathcal{G} \]

\[ \begin{array}{c}
W \\
\rightarrow
\\
\rightarrow
\\
\rightarrow
\\
Y \\
\rightarrow
\\
Z
\end{array} \]

\[ \text{Pattern of } \mathcal{G} \]

\[ \begin{array}{c}
W \\
\rightarrow
\\
\rightarrow
\\
\rightarrow
\\
V
\end{array} \]

\begin{align*}
\text{unshielded colliders (e.g., } &\{W \rightarrow X \leftarrow Z\}\text{)} \\
[\mathcal{G}] = [\mathcal{G}'] &\iff \text{pattern}(\mathcal{G}) = \text{pattern}(\mathcal{G}') \quad \text{[Verma and Pearl, 1990]}
\end{align*}
Two DAGs $\mathcal{G}$ and $\mathcal{G}'$ are equivalent under Markov condition, $[\mathcal{G}] = [\mathcal{G}']$, if they entail the same independence relations (= d-separation).

unshielded non-colliders & acyclicity

Meek’s rules [Meek, 1995], &
PDAG extensibility [Dor and Tarsi, 1992]
Markov Equivalence of DAG: Review

- DAG $\mathcal{G}$
- $\text{pattern}(\mathcal{G})$
- Unshielded Colliders
- Unshielded Non-colliders
- CPDAG
- Acyclicity
Markov Equivalence of RCM: Plan

RCM $\mathcal{M}$

? ? ?

? ? ?

? ? ?

? ? ?

? ? ?

? ? ?
Markov Equivalence of RCM

Two RCMs $\mathcal{M}$ and $\mathcal{M}'$ are equivalent under Markov condition, $[\mathcal{M}] = [\mathcal{M}']$, if they entail the same set of relational d-separation.
Markov Equivalence of RCM

Two RCMs $\mathcal{M}$ and $\mathcal{M}'$ are equivalent under Markov condition, $[\mathcal{M}] = [\mathcal{M}']$, if they entail the same set of relational d-separation.

Relational d-separation [Maier et al., 2013] generalizes d-separation among variables to among relational variables.

Example

$$[E].Salary \perp \perp [E, D, P, D, E].Competence \mid \{[E].Competence, [E, D, P, F, B].Budget\}$$
Markov Equivalence of RCM

Two RCMs $\mathcal{M}$ and $\mathcal{M}'$ are equivalent under Markov condition, $[\mathcal{M}] = [\mathcal{M}']$, if they entail the same set of relational d-separation.

Relational d-separation [Maier et al., 2013] generalizes d-separation among variables to among relational variables.

Example - base item class

$[E].\text{Salary} \indep [E, D, P, D, E].\text{Competence} |$

\{[E].\text{Competence}, [E, D, P, F, B].\text{Budget}\}
Markov Equivalence of RCM

Two RCMs $\mathcal{M}$ and $\mathcal{M}'$ are equivalent under Markov condition, $[\mathcal{M}] = [\mathcal{M}']$, if they entail the same set of relational d-separation.

Relational d-separation generalizes d-separation among variables (i.e., attributes) to among relational variables.

**Relational d-separation** = $\forall$ d-separation

Let $U$, $V$, $W$ be relational variables starting with $B \in \mathcal{E} \cup \mathcal{R}$,

$$(U \perp \!\!\!\perp V \mid W)_{\mathcal{M}} \triangleq \forall_{\sigma \in \Sigma_S} \forall_{i \in \sigma(B)} (U|_{i}^{\sigma} \perp \!\!\!\perp V|_{i}^{\sigma} \mid W|_{i}^{\sigma})_{G_{\sigma}^{\mathcal{M}}}$$

for every relational skeleton
for every base item
Markov Equivalence of RCM

Two RCMs $\mathcal{M}$ and $\mathcal{M}'$ are equivalent under Markov condition, $[\mathcal{M}] = [\mathcal{M}']$, if they entail the same set of relational d-separation.
A Necessary and Sufficient Condition

**Theorem**

\[ [\mathcal{M}] = [\mathcal{M}'] \iff \forall \sigma \in \Sigma_s [G^\mathcal{M}_\sigma] = [G^\mathcal{M}'_\sigma] \]

- **Sufficiency:**
  from the definition of relational d-separation

- **Necessity:**
  1. Different **adjacencies**:
     \[ \exists i. X \rightarrow j. Y \implies \exists P. Y \rightarrow \forall X \implies \exists S \forall X \perp \perp P. Y | S \]
  2. Different **unshielded colliders**:
     \[ \exists (i.X, j.Y, k.Z) \implies \exists (\forall X, P.Y, R.Z) \implies \exists S \forall X \perp \perp R.Z | S \]
Pattern of RCM

<table>
<thead>
<tr>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>adjacencies of $\mathcal{M}$ + orientations from canonical unshielded colliders of $\mathcal{M}$.</td>
</tr>
</tbody>
</table>

- **Problem**: infinite # of canonical unshielded (non-)colliders.
  \[
  \{(\forall X, P.Y, R.Z)\} \text{ of } \mathcal{M} \leftrightarrow \{(i.X, j.Y, k.Z)\} \text{ of } \forall \sigma \in \Sigma \mathcal{S} \mathcal{G} \mathcal{M}_{\sigma}.
  \]

- **Solution**: enumerate a sufficient subset of canonical unshielded triples to retrieve a pattern.
Pattern of RCM

Definition

adjacencies of $\mathcal{M}$ +
orientations from canonical unshielded colliders of $\mathcal{M}$.

**Problem:** infinite # of canonical unshielded (non-)colliders.

$\{(\forall x, P \cdot Y, R \cdot Z)\}$ of $\mathcal{M}$ $\leftrightarrow$ $\{(i \cdot X, j \cdot Y, k \cdot Z)\}$ of $\forall \sigma \in \Sigma S^M_\sigma$.

**Solution:** enumerate a sufficient subset of canonical unshielded triples to retrieve a pattern.
Pattern of RCM

Definition

adjacencies of $\mathcal{M}$ +
orientations from canonical unshielded colliders of $\mathcal{M}$.

Problem: infinite # of canonical unshielded (non-)colliders.

$\{(\forall X, P.Y, R.Z)\}$ of $\mathcal{M}$ $\leftrightarrow$ $\{(i.X, j.Y, k.Z)\}$ of $\forall \sigma \in \Sigma G^M_{\sigma}$.

Solution: enumerate a sufficient subset of canonical unshielded triples to retrieve a pattern.
RCM \( \mathcal{M} \) \rightarrow \text{pattern}(\mathcal{M}) \rightarrow \text{a sufficient subset of} \rightarrow \text{Canonical Unshielded Colliders} \rightarrow \ldots \
canonical unshielded collider
([E]. Competence, [[E, D, P]. Success], [E, D, P, D, E]. Competence) canonical unshielded collider
canonical unshielded collider


Pattern of RCM
Completed Partially-directed RCM: CPRCM

- acyclicity: $\mathcal{A}$ is a partially-ordered set. CDG $\mathcal{G}_\mathcal{A}$
- canonical unshielded non-colliders
  - e.g., ([B].Budget, {[B].Revenue}, [B, F, P].Success)
Completed Partially-directed RCM: CPRCM

- acyclicity: $\mathcal{A}$ is a partially-ordered set. CDG $\mathcal{G}_A$^

- canonical unshielded non-colliders = $\mathcal{A}$-level non-colliders
  e.g., ($Budget$, $Revenue$, $Success$)
Completed Partially-directed RCM: CPRCM

- **acyclicity**: $\mathcal{A}$ is a partially-ordered set. CDG $\mathcal{G}^M_{\mathcal{A}}$
- canonical unshielded non-colliders = $\mathcal{A}$-level non-colliders
e.g., (Budget, Revenue, Success)
- generalized PDAG extensibility with (un)shielded non-colliders.
Completed Partially-directed RCM: CPRCM

- acyclicity: $\mathcal{A}$ is a partially-ordered set. CDG $\mathcal{G}_\mathcal{A}^\mathcal{M}$
- canonical unshielded non-colliders = $\mathcal{A}$-level non-colliders
e.g., (Budget, Revenue, Success)
- generalized PDAG extensibility with (un)shielded non-colliders.

Pattern-CDG

$\text{pattern}(\mathcal{M})$
Completed Partially-directed RCM: CPRCM

- **acyclicity**: $\mathcal{A}$ is a partially-ordered set. CDG $\mathcal{G}_A^M$
- **canonical unshielded non-colliders** = $\mathcal{A}$-level non-colliders
  e.g., ($\text{Budget}$, $\text{Revenue}$, $\text{Success}$)
- **generalized PDAG extensibility** with (un)shielded non-colliders.
RCM $\mathcal{M}$

- Canonical Unshielded Colliders
  - a sufficient subset of $\text{pattern}(\mathcal{M})$
- Canonical Unshielded Non-colliders
  - a sufficient subset of $\text{pattern}(\mathcal{M})$

1. $g^\text{pattern}(\mathcal{M})$
2. completed $g^\text{pattern}(\mathcal{M})$
3. CPRCM

Acyclicity
Summary & Future work

- RCM generalizes CBN
- Markov equivalence of RCM generalizes that of CBN.
  - adjacencies and unshielded (non-)colliders.
  - generalized PDAG extensibility with non-colliders.
- a sound mechanism for relational d-separation
- relax assumptions (e.g., acyclicity)
- accurate, non-parametric, CI tests for relational data (non-iid)
- robust causal discovery algorithm
Summary & Future work

- RCM generalizes CBN
- Markov equivalence of RCM generalizes that of CBN.
  - adjacencies and unshielded (non-)colliders.
  - generalized PDAG extensibility with non-colliders.
- A sound mechanism for relational d-separation
- Relax assumptions (e.g., acyclicity)
  - Accurate, non-parametric, CI tests for relational data (non-iid)
  - Robust causal discovery algorithm
Summary & Future work

- RCM generalizes CBN
- Markov equivalence of RCM generalizes that of CBN.
  - adjacencies and unshielded (non-)colliders.
  - generalized PDAG extensibility with non-colliders.

- a sound mechanism for relational d-separation
- relax assumptions (e.g., acyclicity)
- accurate, non-parametric, CI tests for relational data (non-iid)
- robust causal discovery algorithm
thank you

meet me @ poster session


