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

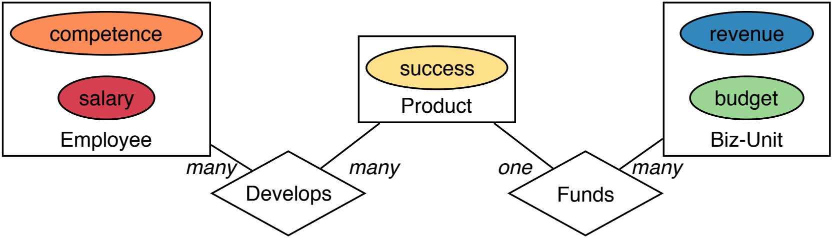
- Relational Schema $\mathcal{S}$
- Relational Skeleton σ
- Relational Causal Model $\mathcal{M}$
- Ground Graph $\mathcal{G}_{\sigma}^{\mathcal{M}}$

Relational Schema $\mathcal{S}$

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

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

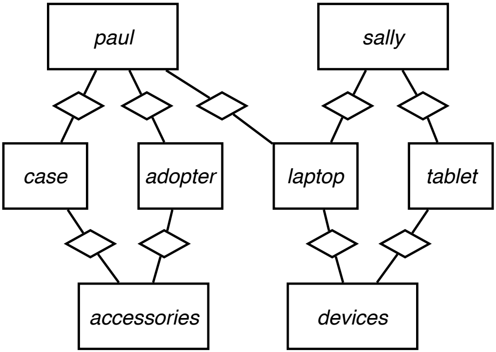
  Cardinality constraints, $\mathcal{R} \times \mathcal{E} \to \{\mathit{one}, \mathit{many}\}$



Maier [2014]

Relational Skeleton $\sigma \in \Sigma_{\mathcal{S}}$

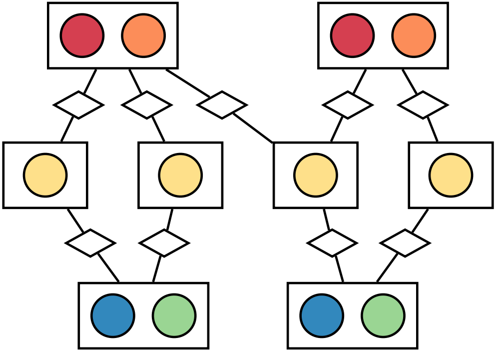
- an instance of the given relational schema $\mathcal{S}$
  - $\Sigma_{\mathcal{S}}$, all possible instantiations
- an undirected bipartite graph
  - node = item (i.e., entity or relationship, *i*, *j*)
  - edge = the participation of an entity in a relationship



modified from Maier [2014]

Relational Skeleton $\sigma \in \Sigma_{\mathcal{S}}$

- an instance of the given relational schema $\mathcal{S}$
  - $\Sigma_{\mathcal{S}}$, 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

- $\mathcal{M} = (\mathcal{S}, \mathbf{D}, \mathbf{\Theta})$
  with a set of relational dependencies **D**,
  and relevant functions or parameters **Θ**

Relational Causal Model

- $\mathcal{M} = (\mathcal{S}, \mathbf{D}, \mathbf{\Theta})$
  with a set of relational dependencies **D**,
  and relevant functions or parameters **Θ**

relational dependency

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

$$\underbrace{[\textit{Product}, \textit{Develops}, \textit{Employee}]}_{\textit{relational path}}.\textit{Competence} \rightarrow [\textit{Product}].\textit{Success}$$

Relational Causal Model

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

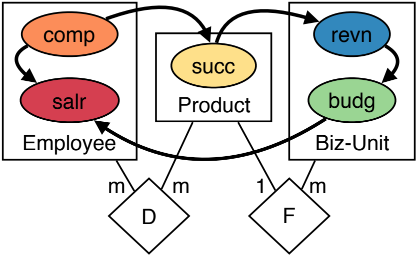
relational dependency

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

$$\underbrace{[\textit{Product}, \textit{Develops}, \textit{Employee}]}_{\text{relational path}}.\textit{Competence} \rightarrow \mathcal{V}_{\textit{Success}}$$

Relational Causal Model

- $\mathcal{M} = (\mathcal{S}, \mathbf{D}, \mathbf{\Theta})$
  with a set of relational dependencies **D**,
  and relevant functions or parameters **Θ**

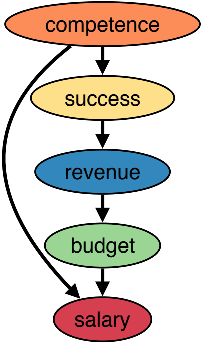


Maier [2014]

$$[E].Competence \rightarrow \mathcal{V}_{Salary},$$
$$[P, D, E].Competence \rightarrow \mathcal{V}_{Success},$$
$$[B, F, P].Success \rightarrow \mathcal{V}_{Revenue},$$
$$[B].Revenue \rightarrow \mathcal{V}_{Budget},$$
$$[E, D, P, F, B].Budget \rightarrow \mathcal{V}_{Salary}$$

Relational Causal Model: Class Dependency Graph

- $\mathcal{M} = (\mathcal{S}, \mathbf{D}, \boldsymbol{\Theta})$
  with a set of relational dependencies $\mathbf{D}$,
  and relevant functions or parameters $\boldsymbol{\Theta}$



Class Dependency Graph $\mathcal{G}_{\mathcal{A}}^{\mathcal{M}}$

acyclicity of an RCM

= acyclicity of its CDG

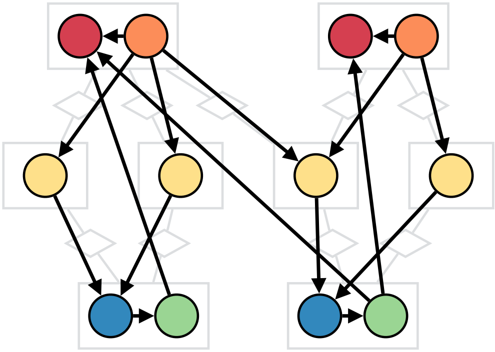
= $\mathcal{A}$ is partially-ordered.

Ground Graph $\mathcal{G}_\sigma^{\mathcal{M}}$

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

instantiating relational dependencies

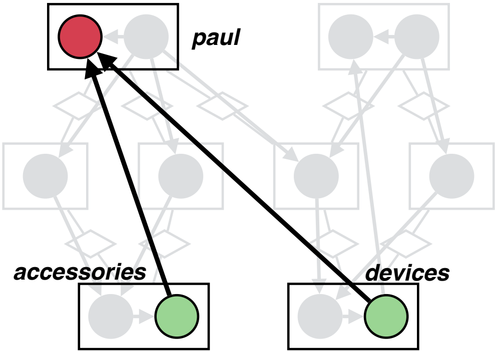
$$j.Y \to i.X \in \mathcal{G}_\sigma^{\mathcal{M}} \qquad \text{if } \exists P.Y \to \mathcal{V}_X \in \mathbf{D} \text{ and } j \in P|_i^\sigma$$



Ground Graph $\mathcal{G}_\sigma^\mathcal{M}$

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

instantiating $[E, D, P, F, B].Budget \rightarrow \mathcal{V}_{Salary}$ @ paul

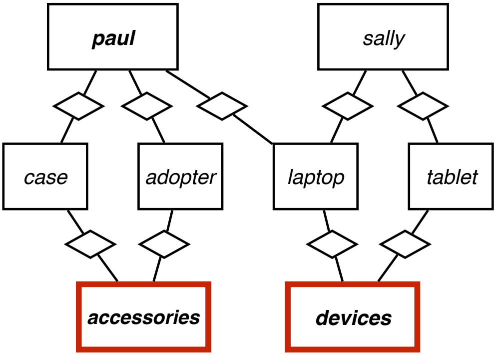


Ground Graph $\mathcal{G}_\sigma^{\mathcal{M}}$

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

instantiating $[E, D, P, F, B].Budget \to \mathcal{V}_{Salary}$ @ paul

$$\{accessories, devices\} = [E, D, P, F, B]|_{paul}^{\sigma}$$

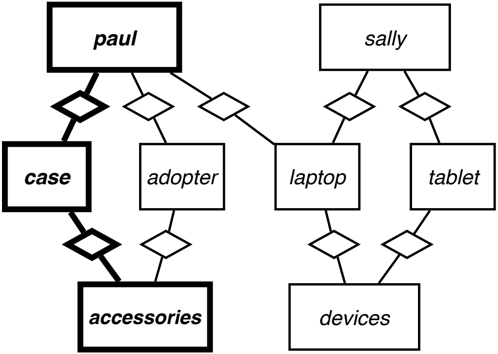


Ground Graph $\mathcal{G}_{\sigma}^{\mathcal{M}}$: Path Semantics

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

instantiating $[E, D, P, F, B].Budget \rightarrow \mathcal{V}_{Salary}$ @ paul

$$accessories \in [\mathbf{E}, \mathbf{D}, \mathbf{P}, \mathbf{F}, \mathbf{B}]|^{\sigma}_{paul}$$

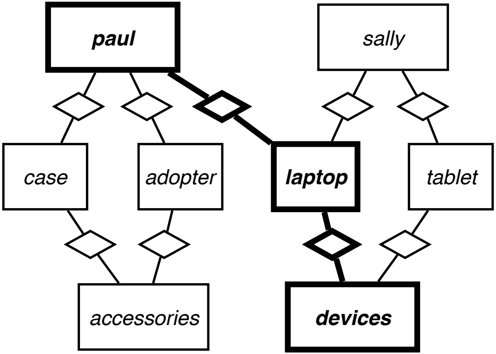


Ground Graph $\mathcal{G}_\sigma^{\mathcal{M}}$: Path Semantics

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

instantiating $[E, D, P, F, B].Budget \rightarrow \mathcal{V}_{Salary}$ @ paul

$$devices \in [\mathbf{E}, \mathbf{D}, \mathbf{P}, \mathbf{F}, \mathbf{B}]|^{\sigma}_{paul}$$

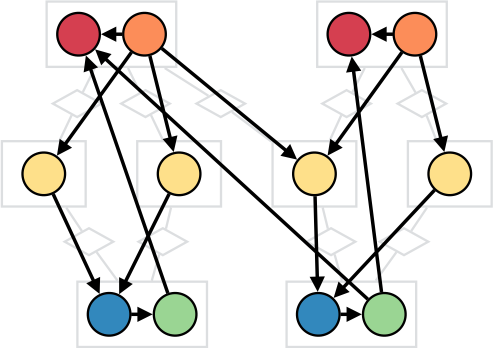


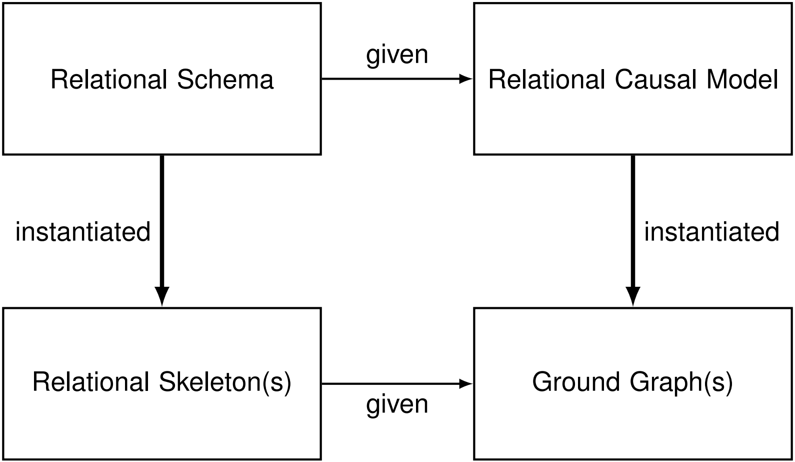
Ground Graph $\mathcal{G}_\sigma^{\mathcal{M}}$

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

instantiating relational dependencies

$$j.Y \rightarrow i.X \in \mathcal{G}_\sigma^{\mathcal{M}} \qquad \text{if } \exists P.Y \rightarrow \mathcal{V}_X \in \mathbf{D} \text{ and } j \in P|_i^\sigma$$

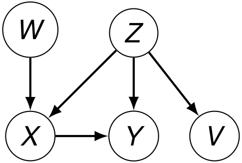



Relational Schema
given
Relational Causal Model
instantiated
instantiated
Relational Skeleton(s)
given
Ground Graph(s)

MARKOV EQUIVALENCE of RCMs

Markov Equivalence of DAG: Review

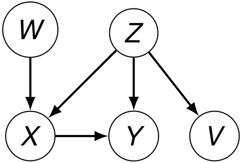
Two DAGs $\mathcal{G}$ and $\mathcal{G}'$ are equivalent under Markov condition, $[\mathcal{G}] = [\mathcal{G}']$, if they entail the same independence relations (= d-separation).



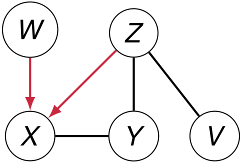
DAG $\mathcal{G}$

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).



DAG $\mathcal{G}$



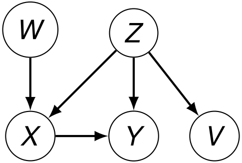
Pattern of $\mathcal{G}$

unshielded colliders (e.g., $\{W \rightarrow X \leftarrow Z\}$)

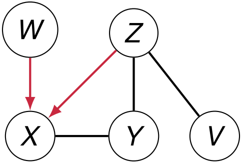
$[\mathcal{G}] = [\mathcal{G}'] \Leftrightarrow \textit{pattern}(\mathcal{G}) = \textit{pattern}(\mathcal{G}')$ [Verma and Pearl, 1990]

Markov Equivalence of DAG: Review

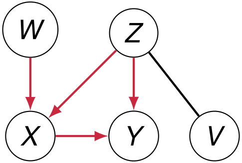
Two DAGs $\mathcal{G}$ and $\mathcal{G}'$ are equivalent under Markov condition, $[\mathcal{G}] = [\mathcal{G}']$, if they entail the same independence relations (= d-separation).



DAG $\mathcal{G}$



Pattern of $\mathcal{G}$

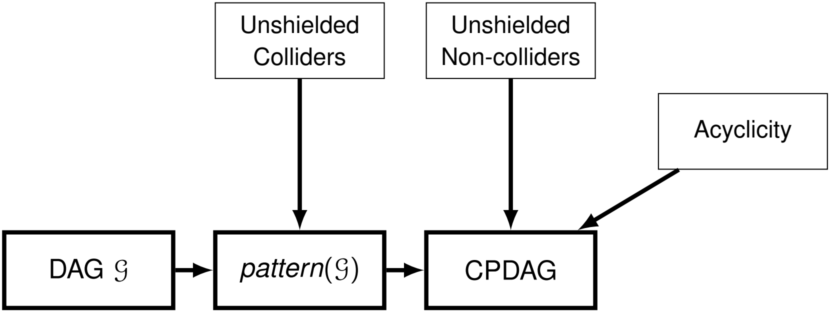


CPDAG of $\mathcal{G}$

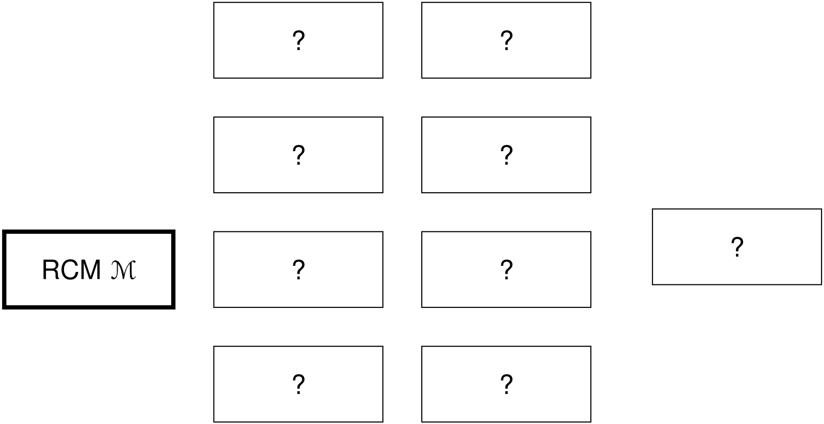
unshielded non-colliders & acyclicity

Meek's rules [Meek, 1995], &
PDAG extensibility [Dor and Tarsi, 1992]

Markov Equivalence of DAG: Review



Markov Equivalence of RCM: Plan



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

$$[E].Salary \perp\!\!\!\perp [E, D, P, D, E].Competence \mid \{[E].Competence, [E, D, P, F, B].Budget\}$$

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

relational d-separation = ∀d-separation

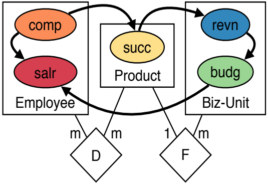
Let U, V, $\mathbf{W}$ be relational variables starting with $B \in \mathcal{E} \cup \mathcal{R}$,

$$(U \perp\!\!\!\perp V \mid \mathbf{W})_{\mathcal{M}} \triangleq \forall_{\sigma \in \Sigma_{\mathcal{S}}} \forall_{i \in \sigma(B)} (U|_i^\sigma \perp\!\!\!\perp V|_i^\sigma \mid \mathbf{W}|_i^\sigma)_{\mathcal{G}_\sigma^{\mathcal{M}}}$$

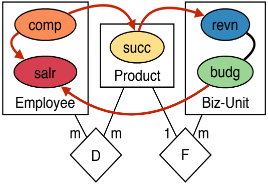
for every relational skeleton → $\forall_{\sigma \in \Sigma_{\mathcal{S}}}$
for every base item → $\forall_{i \in \sigma(B)}$

Markov Equivalence of RCM

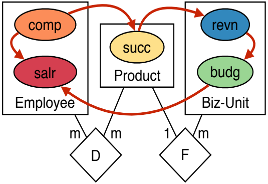
Two RCMs $\mathcal{M}$ and $\mathcal{M}'$ are equivalent under Markov condition, $[\mathcal{M}] = [\mathcal{M}']$, if they entail the same set of relational d-separation.



$\mathcal{M}$



Pattern of $\mathcal{M}$



CPRCM of $\mathcal{M}$

A Necessary and Sufficient Condition

Theorem

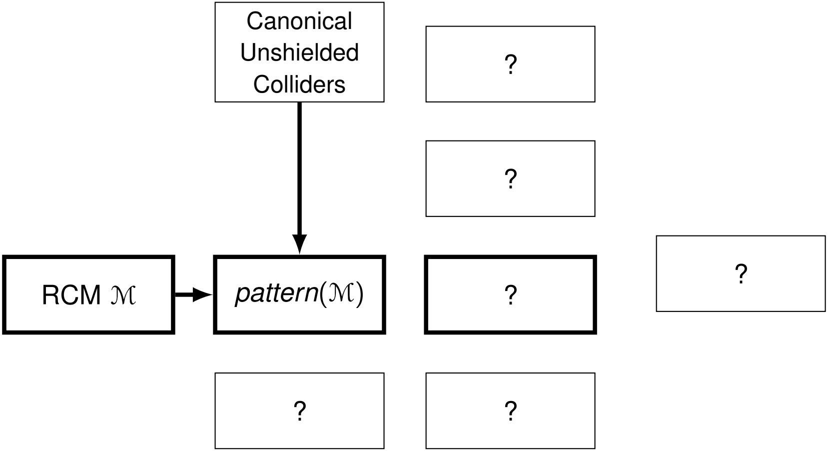
$$[\mathcal{M}] = [\mathcal{M}'] \quad \Leftrightarrow \quad \forall_{\sigma \in \boldsymbol{\Sigma}_{\mathcal{S}}} [\mathcal{G}_{\sigma}^{\mathcal{M}}] = [\mathcal{G}_{\sigma}^{\mathcal{M}'}]$$

- Sufficiency:
  from the definition of relational d-separation
- Necessity:
  1. Different adjacencies:

     $$\exists i.X - j.Y \quad \Rightarrow \quad \exists P.Y - \mathcal{V}_X \quad \Rightarrow \quad \exists_{\mathbf{S}} \mathcal{V}_X \perp\!\!\!\perp P.Y \mid \mathbf{S}$$

  2. Different unshielded colliders:

     $$\exists (i.X, j.Y, k.Z) \;\Rightarrow\; \exists (\mathcal{V}_X, \mathbf{P}.Y, R.Z) \;\Rightarrow\; \exists_{\mathbf{S}} \mathcal{V}_X \perp\!\!\!\perp R.Z \mid \mathbf{S}$$


Canonical Unshielded Colliders
?
?
RCM $\mathcal{M}$
pattern($\mathcal{M}$)
?
?
?
?

Pattern of RCM

Definition

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

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

  $\{(\mathcal{V}_X, \mathbf{P}.Y, R.Z)\}$ of $\mathcal{M}$ $\leftrightarrow$ $\{(i.X, j.Y, k.Z)\}$ of $\forall_{\sigma \in \Sigma_{\mathcal{S}}} \mathcal{G}_{\sigma}^{\mathcal{M}}$.

- **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.

  $\{(\mathcal{V}_X, \mathbf{P}.Y, R.Z)\}$ of $\mathcal{M} \quad \leftrightarrow \quad \{(i.X, j.Y, k.Z)\}$ of $\forall_{\sigma \in \Sigma_{\mathcal{S}}} \mathcal{G}_{\sigma}^{\mathcal{M}}$.

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

Pattern of RCM

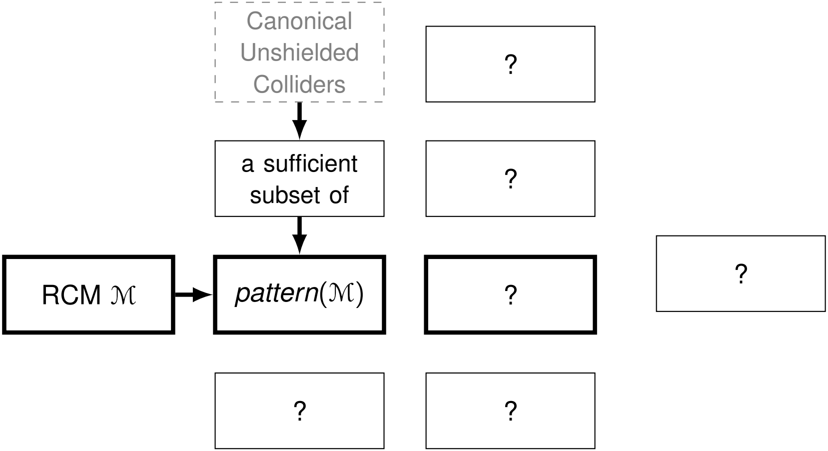
Definition

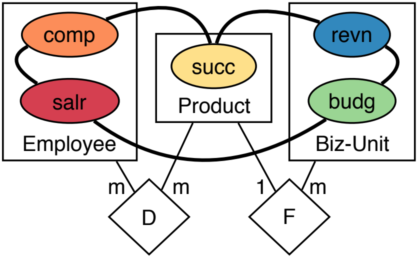
adjacencies of $\mathcal{M}$ +
orientations from canonical unshielded colliders of $\mathcal{M}$.

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

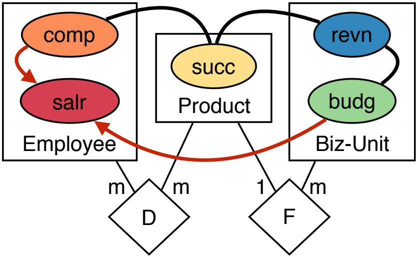
  $$\{(\mathcal{V}_X, \mathbf{P}.Y, R.Z)\} \text{ of } \mathcal{M} \quad \leftrightarrow \quad \{(i.X, j.Y, k.Z)\} \text{ of } \forall_{\sigma \in \Sigma_{\mathcal{S}}} \mathcal{G}_\sigma^{\mathcal{M}}.$$

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

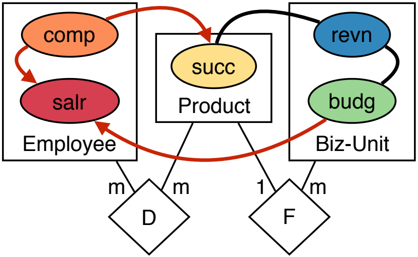

Canonical Unshielded Colliders
a sufficient subset of
RCM $\mathcal{M}$
pattern($\mathcal{M}$)
?
?
?
?
?
?
?



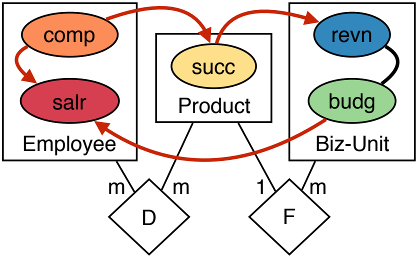
undirected



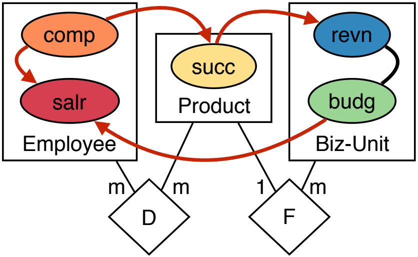
$([B].Budget, \{[B, F, P, D, E].Salary\}, [B, F, P, D, E].Competence)$
canonical unshielded collider



([*E*].*Competence*, {[*E*, *D*, *P*].*Success*}, [*E*, *D*, *P*, *D*, *E*].*Competence*)
canonical unshielded collider



([*P*].*Success*, {[*P*, *F*, *B*].*Revenue*}, [*P*, *F*, *B*, *F*, *P*].*Success*)
canonical unshielded collider



Pattern of RCM

Completed Partially-directed RCM: CPRCM

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

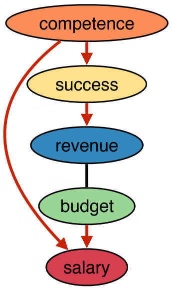
Completed Partially-directed RCM: CPRCM

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

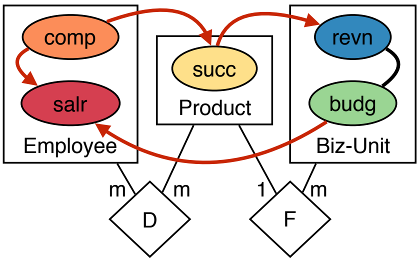
Completed Partially-directed RCM: CPRCM

- 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



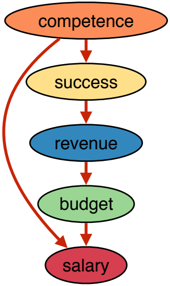
pattern($\mathcal{M}$)



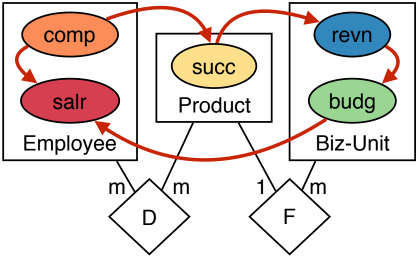
Completed Partially-directed RCM: CPRCM

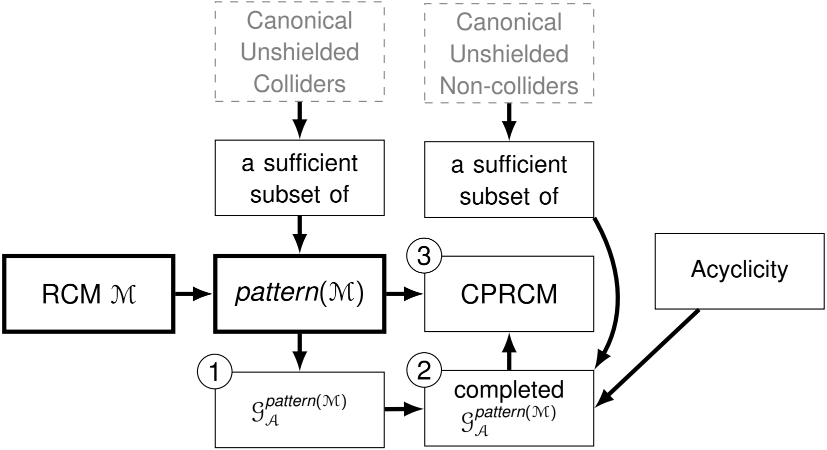
- acyclicity: $\mathcal{A}$ is a partially-ordered set. CDG $\mathcal{G}_{\mathcal{A}}^{\mathcal{M}}$
- canonical unshielded non-colliders = $\mathcal{A}$-level non-colliders
  e.g., (*Budget*, *Revenue*, *Success*)
- generalized PDAG extensibility with (un)shielded non-colliders.

CP-CDG



CPRCM





Canonical Unshielded Colliders
Canonical Unshielded Non-colliders
a sufficient subset of
a sufficient subset of
RCM $\mathcal{M}$
pattern($\mathcal{M}$)
3
CPRCM
Acyclicity
1
$\mathcal{G}_{\mathcal{A}}^{pattern(\mathcal{M})}$
2
completed $\mathcal{G}_{\mathcal{A}}^{pattern(\mathcal{M})}$

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

References I

Dorit Dor and Michael Tarsi. A simple algorithm to construct a consistent extension of a partially oriented graph. Technical Report R-185, Cognitive Systems Laboratory, UCLA, 1992.

Marc Maier, Brian Taylor, and David Jensen. Learning causal models of relational domains. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, pages 531–538. AAAI Press, 2010.

Marc Maier, Katerina Marazopoulou, David Arbour, and David Jensen. A sound and complete algorithm for learning causal models from relational data. In *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence*, pages 371–380. AUAI Press, 2013.

Marc E Maier. *Causal Discovery for Relational Domains: Representation, Reasoning, and Learning*. PhD thesis, University of Massachusetts Amherst, 2014.

Christopher Meek. Causal inference and causal explanation with background knowledge. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 403–410. Morgan Kaufmann, 1995.

Tom Verma and Judea Pearl. Equivalence and synthesis of causal models. In *Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence*, pages 220–227. AUAI Press, 1990.