Learning Relational Causal Models with Cycles through Relational Acyclification

Ragib Ahsan[1], David Arbour[2], Elena Zheleva[1]

[1]University of Illinois at Chicago, Chicago, IL
[2]Adobe Research
[1]{rahsan3, ezheleva}@uic.edu, [2]arbour@adobe.com

Abstract

In real-world phenomena which involve mutual influence or causal effects between interconnected units, equilibrium states are typically represented with cycles in graphical models. An expressive class of graphical models, *relational causal models*, can represent and reason about complex dynamic systems exhibiting such cycles or feedback loops. Existing cyclic causal discovery algorithms for learning causal models from observational data assume that the data instances are independent and identically distributed which makes them unsuitable for relational causal models. At the same time, causal discovery algorithms for relational causal models assume acyclicity. In this work, we examine the necessary and sufficient conditions under which a constraint-based relational causal discovery algorithm is sound and complete for *cyclic relational causal models*. We introduce *relational acyclification*, an operation specifically designed for relational models that enables reasoning about the identifiability of cyclic relational causal models. We show that under the assumptions of relational acyclification and σ-faithfulness, the relational causal discovery algorithm RCD is sound and complete for cyclic models. We present experimental results to support our claim.

Introduction

Most of the tools and methods developed for causal discovery rely on a graphical representation based on Bayesian networks which assume independent and identically distributed (i.i.d) instances. Probabilistic relational models (Getoor et al. 2007) have been developed that relax this assumption. The key advantage of these *relational* models is that they can represent systems involving multiple types of entities interacting with each other with some probabilistic dependence. Causal reasoning over such relational systems is key to understanding many real-world phenomena, such as social influence.

Influence in complex dynamic systems is often mutual and represented by a feedback loop or cycle in the relational model. Identifying mutual influence in relational models is of great interest in the research community. For example, social scientists and marketing experts are interested to study the social dynamics between people and products in social networks (Bakshy et al. 2011; Bakshy, Messing, and Adamic 2015; Ogburn, Shpitser, and Lee 2020). However, there is a

lack of available methods for discovering mutual influence or cycles in complex relational systems.

Sound and complete algorithms have been proposed for learning relational causal models from observational data (Maier et al. 2013; Lee and Honavar 2016a,b). However, they assume acyclicity and thus cannot reason about mutual influence or cycles. In a recent work, Ahsan et al. (2022) develop σ-abstract ground graph (σ-AGG), a sound and complete representation for cyclic relational causal models. Even though σ-AGG is shown to be sound and complete for cyclic relational causal models, to the best of our knowledge no work on discovering σ-AGG or identifying relational cycles from observational data exists in the literature.

The closest works on cyclic causal discovery are mostly from the domain of Bayesian networks. Richardson (1996) develop a cyclic causal discovery (CCD) algorithm which is shown to be sound but not complete. In recent work, Mooij and Claassen (2020) provide necessary conditions for constraint-based causal discovery algorithms developed for acyclic causal models, such as PC (Pearl et al. 2000) and FCI (Spirtes et al. 2000), to be sound and complete for cyclic causal models under sigma-separation criteria. There are several other algorithms for cyclic causal discovery from i.i.d samples (Rothenhäusler et al. 2015; Strobl 2019a) but no such algorithm exists for cyclic relational causal models.

In this work, we examine the necessary and sufficient conditions for which constraint-based relational causal discovery can be shown to be sound and complete for cyclic relational causal models under σ-separation. We introduce *relational acyclification*, an operation that helps to reason over the scope of cyclic relational causal models which are identifiable with constraint-based causal discovery algorithms. Following this criterion, we show that RCD (Maier et al. 2013), a pioneering relational causal discovery algorithm for acyclic relational models, is sound and complete for cyclic relational models under σ-separation and causal sufficiency assumption. We provide experimental results on synthetic relational models in support of our claims. We also demonstrate the effectiveness of the algorithm on a real-world dataset.

Related Work

Richardson (1996) develop cyclic causal discovery (CCD) algorithm for directed cyclic graphs under the causal sufficiency assumption. They provide a characterization of the

equivalence class of cyclic causal models. They show that a class of graphs called Partial Ancestral Graphs (PAG) is sufficient to represent the equivalence class of cyclic causal models. Two models are equivalent if they entail the same set of d-separation relationships. There are a few caveats about CCD and its choice of representation. Markov property of d-separation holds for directed cyclic graphs only with linear structural equation models (SEM). So, it is not sure how to extend CCD to more general models beyond linear SEMs. Moreover, CCD is not complete in the sense that it does not guarantee to produce the maximally oriented PAG.

Rothenhäusler et al. (2015) develop a general discovery algorithm (BackShift) that allows latent confounders in addition to cycles. The method relies on equilibrium data of the model recorded under a specific kind of intervention called *shift interventions*. (Strobl 2019b) develop the CCI algorithm which allows both latent confounders and selection bias apart from cycles. Similar to CCD, both Backshift and CCI are restricted to only linear SEM models. CCI considers a different representation for equivalence class, called maximal almost ancestral graph (MAAG) which is carefully chosen to allow for both latent confounders and selection bias.

The FCI algorithm is a constraint-based causal discovery algorithm designed specifically for acyclic causal models with latent confounders (Spirtes et al. 2000). Mooij and Claassen (2020) show that FCI is sound and complete for cyclic models under σ-separation criteria which is different than d-separation and not restricted to only linear models. They also show that any constraint-based causal discovery algorithm (PC, FCI (Spirtes et al. 2000) etc.) which is sound and complete for acyclic causal models, can be shown to be sound and complete for cyclic causal models under some background knowledge (i.e. sufficiency) and assumptions.

Maier et al. (2013) develop the first sound and complete algorithm that can discover the dependencies of a relational causal model under the assumption of d-faithfulness, sufficiency, acyclicity. It is designed based on the PC algorithm with some additional steps introduced specifically to handle relational aspects of the representation. They utilize the *abstract ground graph*, an abstract representation that allows answering relational queries based on d-separation criterion (Maier, Marazopoulou, and Jensen 2013). RCD introduces Relational Bi-variate Orientation (RBO)- an orientation rule specifically designed for relational models. Lee and Honavar (2016b) develop an efficient version of RCD named RCD-Light which requires polynomial time to run. They also develop an alternative algorithm RpCD based on *path semantic* which describes a unique way of defining relational paths (Lee and Honavar 2016a).

Ahsan et al. (2022) develop σ-abstract ground graph, a sound and complete abstract representation for cyclic relational causal models under σ-separation. They introduce relational σ-separation and show that this criterion can consistently answer relational queries on cyclic relational models.

Preliminaries

We present only a high-level overview of the important notions used in this paper due to space constraint [1]. For further details, we refer readers to the relevant literature (Pearl et al. 2000; Maier et al. 2013; Lee and Honavar 2015; Forré and Mooij 2017; Mooij and Claassen 2020).

Cyclic Graphical Causal Models

The most common graphical representation for causal models is *directed acyclic graphs (DAGs)*. DAGs provide ways for natural causal interpretation and satisfy the Markov property under d-separation. A more general class of graphs are *directed cyclic graphs (DCGs)* which drop the assumption of acyclicity (and allow feedback loops). These graphs are appropriate for (possibly cyclic) structural causal models (SCMs) where the corresponding Markov properties and causal interpretation are more subtle (Bongers et al. 2021). Cyclic SCMs are useful to represent causal semantics of equilibrium states in dynamical systems (Bongers et al. 2021).

Directed cyclic graphs offer certain properties that help model cyclic causal models. Given a directed cyclic graph $G = (V, E)$, all nodes on directed cycles passing through node $i \in V$ together form the strongly connected component $SC_G(i) = AN_G(i) \bigcap DE_G(i)$ of i where $AN_G(i)$ and $DE_G(i)$ refers to the ancestors and descendants of node $i \in V$. The set of conditional independence entailed in DCG, G is refereed to as independence model $IM(G)$.

Unlike DAGs, DCGs are not guaranteed to satisfy the Markov property in a general case under d-separation. Instead, a more general notion of separation, called σ-separation satisfies the Markov property of DCGs (Forré and Mooij 2017). σ-separation states that a non-collider blocks a path only if it points to another node in the path which belongs to a different strongly connected component (Mooij and Claassen 2020). σ-faithfulness refers to the property which states that all statistical dependencies found in the distribution generated by a given causal model are entailed by the σ-separation relationships.

Richardson (1996) show that a class of graphs called Partial Ancestral Graphs (PAG) is a sufficient representation for the equivalence class of cyclic causal models represented by DCGs. PAGs have also been shown to be a sufficient representation for causal discovery with cycles and unobserved confounders (Mooij and Claassen 2020). Since we are assuming no selection bias for simplicity, we will only discuss directed PAGs (DPAG) in this study.

Forré and Mooij (2017) introduced an operation called *acyclification* for directed cyclic graphs that generates DAGs with equivalent independence models as the given DCG. It allows a single DPAG to represent the ancestral relationship of a DCG G and all its acyclifications G'.

Definition 1 (Acyclification (Forré and Mooij 2017)). *Given a DCG $G = (\mathcal{V}, \mathcal{E})$, an acyclification of G is a DAG $G' = (\mathcal{V}, \mathcal{E}')$ with*
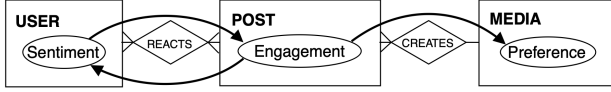

Figure 1: Example of a cyclic relational model. There are three entity types (USER, POST, MEDIA) and two relationship types (REACTS, CREATES) among them. Attributes are shown in oval shapes. The bold arrows refer to the relational dependencies.

 i the same nodes $\mathcal{V}$;

 ii for any pair of nodes i, j such that $i \notin SC_G(j)$: $i \to j \in \mathcal{E}'$ iff there exists a node k such that $k \in SC_G(j)$ and $i \to k \in \mathcal{E}$;

 iii for any pair of distinct nodes i, j such that $i \in SC_G(j)$: $i \to j \in \mathcal{E}'$ or $i \leftarrow j \in \mathcal{E}'$;

Proposition 1 ((Mooij and Claassen 2020)). *For any DCG G and any acyclification G' of G, $IM_\sigma(G) = IM_\sigma(G') = IM_d(G')$ where $IM_\sigma(G)$ and $IM_d(G)$ refers to the independence model of the given DCG G under σ-separation and d-separation respectively.*

Mooij and Claassen (2020) provide the necessary conditions under which constraint-based causal discovery algorithms for acyclic causal models, such as PC (Pearl et al. 2000) and FCI (Spirtes et al. 2000), are sound and complete in the presence of cycles under σ-separation. Their result depends on the following assumptions:

A 1. *The underlying causal model is σ-faithful.*

A 2. *There exists one or more valid acyclifications of the given causal model which contains the same set of ancestral relationships as the given model.*

Corollary 1 ((Mooij and Claassen 2020)). *The PC algorithm with Meek's orientation rules is sound, arrowhead-complete, tail-complete, and Markov complete (in the σ-separation setting without selection bias) for directed cyclic graphs. (Mooij and Claassen 2020)*

Relational Causal Models (RCMs)

Relational causal models are an expressive class of graphical models that can represent probabilistic dependence among instances of different entity types interacting with each other following a specific schema. We use a simplified Entity-Relationship model to describe relational models following previous work (Heckerman, Meek, and Koller 2007; Maier et al. 2013; Lee and Honavar 2020). A relational schema $\mathcal{S} = \langle \mathcal{E}, \mathcal{R}, \mathcal{A}, card \rangle$ represents a relational domain where $\mathcal{E}$, $\mathcal{R}$ and $\mathcal{A}$ refer to the set of entity, relationship and attribute classes respectively. Figure 1 shows an example relational model that describes a simplified user-media engagement system. The cardinality constraints are shown with crow's feet notation— a user can react to multiple posts, multiple users can react to a post, and a post can be created by only a single media entity.

A relational causal model $\mathcal{M} = \langle \mathcal{S}, \mathcal{D} \rangle$, is a collection of relational dependencies defined over schema $\mathcal{S}$. *Relational dependencies* consist of two relational variables,

cause, and effect. As an example, consider the following relational dependency [Post, Reacts, User].Sentiment $\to$ [Post].Engagement which states that the engagement of a post is affected by the actions of users who react to that post. Note that, all causal dependencies are defined with respect to a specific *perspective* (entity type). A relational model $\mathcal{M} = (\mathcal{S}, \mathcal{D})$ is said to be *cyclic* if the set of relational dependencies $\mathcal{D}$ construct one or more directed cycles of arbitrary length. There is a direct feedback loop in the relational model of Figure 1 making it a cyclic relational causal model.

A realization of a relational model $\mathcal{M}$ with a relational skeleton is referred to as the *ground graph $GG_\mathcal{M}$*. It is a directed graph consisting of attributes of entities in the skeleton as nodes and relational dependencies among them as edges. Figure 2a shows the ground graph for the relational model from Figure 1. A *σ-abstract ground graph* (σ-AGG) is an abstract representation that captures the dependencies consistent in all possible ground graphs and represents them in a directed graph. σ-AGGs are defined for a specific perspective and *hop threshold*, h. Hop threshold refers to the maximum length of the relational paths. Figure 2b presents the σ-AGG from the perspective of USER with $h = 6$.

Conditional independence facts are only useful when they hold across all ground graphs that are consistent with the model. Maier, Marazopoulou, and Jensen (2013) show that relational d-separation is sufficient to achieve that for acyclic models. In recent work, (Ahsan et al. 2022) introduced relational σ-separation criteria specifically for cyclic relational models which directly follows from the definition of relational d-separation except it uses σ-separation criterion instead of d-separation.

Relational Causal Discovery (RCD)

The RCD algorithm developed by Maier et al. (2013) is the first sound and complete algorithm that can discover the dependencies of a relational causal model (RCM) under the assumption of d-faithfulness, sufficiency, acyclicity, and a maximum hop threshold h. It is designed based on the PC algorithm with some additional steps introduced specifically to handle relational aspects of the representation. Maier et al. (2013) provides theoretical guarantees for soundness and completeness of RCD.

Relational Causal Discovery with Cycles

Cyclic relational causal models (CRCM) are relational causal models where dependencies form one or more directed cycles (Ahsan et al. 2022). The cycles or feedback loops can represent equilibrium states in dynamic systems. Consider the example from Figure 1 where sentiments of users and engagements in a media post may reach an equilibrium. Identifying such cycles or feedback loops from observational samples requires proper representation and a learning algorithm. Ahsan et al. (2022) introduce an abstract representation, σ-AGG that entails all the conditional independence relations consistent across all ground graphs of the model and shows that it is sound and complete under σ-separation. Given σ-AGG representation, discovering CRCM transforms into the problem of learning the σ-AGG from observational

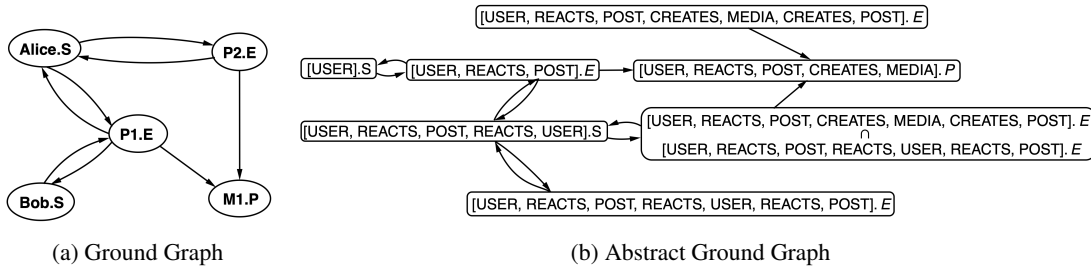(a) Ground Graph   (b) Abstract Ground Graph

Figure 2: Fragments of a relational ground graph, and σ-abstract ground graph corresponding to the relational causal model from Figure 1. Here, S, E, P refers to the attributes Sentiment, Engagement, Preference correspondingly. The arrows represent relational dependencies.

samples of a relational model. Since σ-AGG is a DCG, we can consider DPAGs to represent the equivalence class of σ-AGG following the previous work of Richardson (1996).

Problem 1 (Cyclic Relational Causal Discovery). *Given observational samples from a σ-faithful cyclic relational causal model $\mathcal{M} = \langle \mathcal{S}, \mathcal{D} \rangle$ with hop threshold h, learn the maximally oriented DPAG that contains the corresponding σ-AGGs of $\mathcal{M}$.*

RCD for Cyclic Relational Causal Models

The RCD algorithm developed by Maier et al. (2013) is the first sound and complete constraint-based algorithm that can learn relational dependencies of a relational causal model (RCM) under the assumption of d-faithfulness, sufficiency, acyclicity, and a maximum hop threshold h. It is designed based on the PC algorithm with additional steps introduced specifically to handle relational aspects of the representation.

Following the recent development by Mooij and Claassen (2020) (Corollary 1), and considering that RCD is developed based on the PC algorithm, a natural question arises: *Is RCD sound and complete for cyclic relational causal models?* To the best of our knowledge, no prior work addresses this question. More generally, the effectiveness and theoretical guarantees of existing relational causal structure learning algorithms for cyclic RCMs under σ-separation are not studied in the current literature.

A Counterexample We present a counterexample that shows that RCD is not sound and complete for discovering cyclic relational causal models in general. Figure 3a shows a CRCM with three entity types A,B,C, and two relationship types AB, BC and maximum hop threshold $h = 2$. The attribute types X1, Y1, and Z1 refer to the attributes of entity types A, B, and C respectively. There are three relational dependencies: 1) [A, AB, B].Y1 $\rightarrow$ [A].X1, 2) [B, AB, A].X1 $\rightarrow$ [B].Y1, and 3) [B, BC, C].Z1 $\rightarrow$ [B].Y1. The first two dependencies form a feedback loop. Figure 3b shows the true σ-AGG built from perspective A with maximum hop threshold $h = 4$ [2]. Figure 3c shows the output of RCD with

---

[2] A sound and complete AGG contains nodes with higher hop threshold than the model(h) (Maier, Marazopoulou, and Jensen 2013). However, the hop threshold of relational dependencies cannot exceed h. The same argument holds for σ-AGGs as well.

a σ-separation oracle. We see that RCD orients arrows [A, AB, B].Y1 $\rightarrow$ [A].X1 and [A, AB, B].Y1 $\rightarrow$ [A, AB, B, AB, A].X1 which refers to the relational dependency [A, AB, B].Y1 $\rightarrow$ [A].X1. However, the true model contains a feedback loop between [A, AB, B].Y1 and [A].X1. This example shows that RCD, even with σ-separation oracle produces incorrect edge orientations.

Relational Acyclification for Cyclic Relational Causal Models

In this section, we present relational acyclification which enables the discovery of relational causal models with cycles. We also discuss how to read off features of the true model from the output of the discovery algorithm.

Relational Acyclification

The counterexample in the previous subsection shows that the RCD algorithm is not sound and complete for general cyclic RCMs under σ-separation. For the given counterexample, RCD orients edges that contradict the given relational model. In order to understand what causes this error and to find a solution, we focus on the acyclification operation introduced by Forré and Mooij (2017) which is a key tool for the generalization results by Mooij and Claassen (2020).

Figure 4 shows an acyclification of the σ-AGG presented in Figure 3b following definition 1. Here we see the edges [A, AB, B, BC, C].Z1 $\rightarrow$ [A].X1 and [A, AB, B, BC, C].Z1 $\rightarrow$ [A, AB, B, AB, A].X1 which does not follow the relational model since the hop threshold of such dependencies ($h = 4$) exceed the hop threshold of the given model ($h = 2$). The definition of acyclification, as given by (Forré and Mooij 2017) essentially considers all the nodes or entities to be of the same entity type. As a result, applying it directly to relational models creates erroneous results. We propose a new definition of acyclification for relational models which specifically mentions that the maximum hop threshold of an acyclification can be different than the hop threshold of the original model.

Definition 2 (Relational Acyclification). *Given a relational schema $\mathcal{S} = (\mathcal{E}, \mathcal{R}, \mathcal{A}, card)$, σ-AGG $G = (V, E)$, and a hop threshold h, a relational acyclification of G is a σ-AGG $G' = (V, E')$ with hop threshold $h' \geq h$ containing*
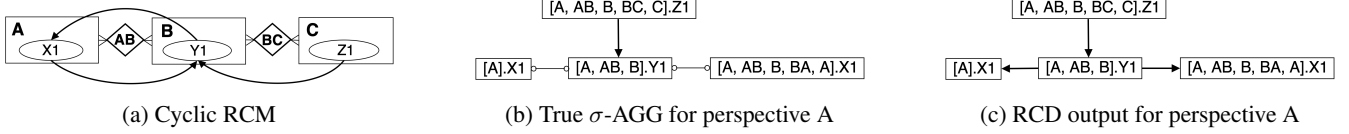
    i the same nodes V;

(a) Cyclic RCM
(b) True σ-AGG for perspective A
(c) RCD output for perspective A

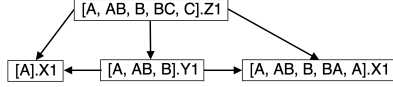Figure 3: Counterexample showing RCD produces incorrect output for cyclic RCM under σ-separation.

Figure 4: Invalid acyclification of σ-AGG from Figure 3b

ii for any pair of nodes $P.X, Q.Y$ such that $P.X \notin SC_G(Q.Y)$: $P.X \to Q.Y \in E'$ iff there exists a node $R.Z$ such that $R.Z \in SC_G(Q.Y)$ and $P.X \to R.Z \in E$ and $P.X \to Q.Y$ is a valid relational dependency with maximum hop threshold h';

iii for any pair of distinct nodes $P.X, Q.Y$ such that $P.X \in SC_G(Q.Y)$: $P.X \to Q.Y \in E'$ or $P.X \leftarrow Q.Y \in E'$;

The definition of relational acyclification follows from Definition 1 where the main distinction is that it allows a new bound on the maximum hop threshold which is different than the bound of the original model. The implication of this is that the potential dependencies RCD considers in building the skeleton, may not be sufficient for soundness and completeness.

Maximum Hop Threshold for Relational Acyclification

Definition 2 suggests that the maximum hop threshold used in a relational acyclification of a σ-AGG may be higher than the hop threshold of the given model. It is important to characterize the maximum bound of relational acyclifications for allowing practical implementation of the RCD algorithm for cyclic models. The following proposition provides the maximum bound on the hop threshold of relational acyclifications.

Proposition 2. *Given a relational model $\mathcal{M} = (\mathcal{S}, \mathcal{D})$ with hop threshold h and corresponding σ-AGG $G = (V, E)$ with a given perspective, the hop threshold h' of any relational acyclification G' of G can be at most $\lfloor \frac{2+l^c}{2} \rfloor h$ where l^c refers to the length of the longest cycle of dependencies in the relational model $\mathcal{M}$.*

The need for higher hop thresholds arises for the additional edges drawn for any incoming edges to a strongly connected component (Definition 1). Any such incoming edge has a maximum hop threshold h of the given model. In order to reach the farthest node in the cycle where each dependency can be of at most h hop threshold, we need at most $\lfloor \frac{l^c}{2} \rfloor h$ hop threshold where l^c refers to the length of the cycle. So, in total it can be at most $h + \lfloor \frac{l^c}{2} \rfloor h = \lfloor \frac{2+l^c}{2} \rfloor h$. Note that in order to calculate an upper bound on the hop threshold of relational acyclification we need to assume the maximum length of any cycle, l^c in the given relational model.

Soundness and Completeness of RCD for Cyclic Relational Causal Models

We consider RCD as a mapping $\mathcal{P}_{RCD}$ from independence models (on variables V) to DPAGs (with vertex set V), which maps the independence model of a σ-AGG G to the DPAG $\mathcal{P}_{RCD}(IM_\sigma(G))$. We assume the following:

A 3. *There exists one or more valid relational acyclifications with hop threshold not exceeding the hop threshold of the given relational causal model ($h' = h$).*

A 4. *The degree of any entity in the relational skeleton is greater than one.*

Assumption 3 follows from Assumption 2 and also limits the set of relational causal models for which RCD can be shown to be sound and complete. Assumption 4 satisfies the soundness and completeness of σ-AGG (Ahsan et al. 2022).

Theorem 1. *Considering Assumption 1, 3, 4 and causal sufficiency holds, RCD is*

(i) sound: for all σ-AGGs G, $\mathcal{P}_{RCD}(IM_\sigma(G))$ contains G;

(ii) arrowhead complete: for all σ-AGGs G: if $i \notin AN_{\tilde{G}}(j)$ for any DCG $\tilde{G}$ that is σ-Markov equivalent to G, then there is an arrowhead $j \ast\!\!\to i$ in $\mathcal{P}_{RCD}(IM_\sigma(G))$

(iii) tail complete: for all σ-AGGs G, if $i \in AN_{\tilde{G}}(j)$ in any DCG $\tilde{G}$ that is σ-Markov equivalent to G, then there is a tail $i \to j$ in $\mathcal{P}_{RCD}(IM_\sigma(G))$;

(iv) Markov complete: for all σ-AGGs G_1 and G_2, G_1 is σ-Markov equivalent to G_2 iff $\mathcal{P}_{RCD}(IM_\sigma(G_1)) = \mathcal{P}_{RCD}(IM_\sigma(G_2))$

in the σ-separation setting given sufficient hop threshold.

Proof. The main idea of the proof is very similar to the proof of Theorem 1 from Mooij and Claassen (2020) where they prove the soundness and completeness of FCI for cyclic models under σ-separation.

To prove soundness, let G be a σ-AGG and $\mathcal{P} = \mathcal{P}_{RCD}(IM_\sigma(G))$. The acyclic soundness of RCD means that for all AGGs G', $\mathcal{P}_{RCD}(IM_\sigma(G'))$ contains G'. Hence, by Definition 2 and Assumption 3, $\mathcal{P}$ contains G' for all acyclifications G'. But then $\mathcal{P}$ must contain G which can be easily shown using Proposition 3 of Mooij and Claassen (2020).

To prove arrowhead completeness, let G be a σ-AGG and suppose that $i \notin AN_{\tilde{G}}(j)$ in any DCG $\tilde{G}$ that is σ-Markov equivalent to G. Since G' is σ-Markov equivalent to G, this implies in particular that for all AGGs $\tilde{G}$ that are d-Markov equivalent to G', $i \notin AN_{\tilde{G}}(j)$. Because of the acyclic arrowhead completeness of RCD, there must be an arrowhead $j \ast\!\!\to i$ in $\mathcal{P}_{RCD}(IM_\sigma(G')) = \mathcal{P}_{RCD}(IM_\sigma(G))$. Tail completeness is proved similarly.
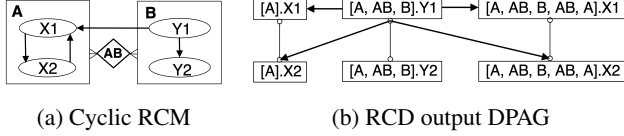
(a) Cyclic RCM   (b) RCD output DPAG

Figure 5: An example cyclic relational model and its corresponding DPAG output by RCD under σ-separation.

To prove Markov completeness: Definition 2 and Proposition 1 imply both $IM_\sigma(G_1) = IM_d(G_1')$ and $IM_\sigma(G_2) = IM_d(G_2')$. From the acyclic Markov completeness of RCD[3], it then follows that G_1' must be d-Markov equivalent to G_2', and hence G_1 must be σ-Markov equivalent to G_2. □

The statement of this theorem can be seen as a special case of the generalization claim (Theorem 2) by Mooij and Claassen (2020). There is an important point to discuss about Assumption 3. Even though Assumption 3 limits the scope of possible relational causal models, it is possible to modify RCD in a way so that it can work for models with relational acyclification having hop threshold higher than the hop threshold of the given model ($h' > h$). The intuition here is that the skeleton building process should consider this new hop threshold h' (which is upper bounded by $\lfloor \frac{2+l^c}{2} \rfloor h$) rather than the true hop threshold h. However, it requires further proof of soundness and completeness with this modified skeleton. We leave this for future work.

Identification of Relational (Non-)Cycles

Mooij and Claassen (2020) show that the patterns in strongly connected components in DCGs can be used as a sufficient condition for identifying the absence of certain cyclic causal relations in a complete DPAG. Given Definition 2, the same condition holds for relational models and σ-AGGs as well. We present the necessary and sufficient conditions for identifying non-cycles in the output of RCD following Proposition 10 by (Mooij and Claassen 2020):

Proposition 3. *Let G be a σ-AGG and denote by $\mathcal{P} = \mathcal{P}_{RCD}(IM_\sigma(G))$ the corresponding complete DPAG output by RCD. Let $i \neq j$ be two nodes in $\mathcal{P}$. If there is an edge $i \circ\!\!-\!\!\circ j$ in $\mathcal{P}$, and all nodes k for which $k *\!\!\rightarrow i$ is in $\mathcal{P}$ also have an edge of the same type $k *\!\!\rightarrow j$ (i.e., the two edge marks at k are the same) in $\mathcal{P}$, then there exists a DCG $\tilde{G}$ with $j \in SC_{\tilde{G}}(i)$ that is σ-Markov equivalent to G, but also a DCG H with $j \notin SC_H(i)$ that is σ-Markov equivalent equivalent to G.*

In other words, under the conditions of this proposition, it is not identifiable from $\mathcal{P}$ alone whether j and i are part of a causal cycle, but they are candidates of being part of a cycle. Figure 5 shows an example of this identifiability criteria. Figure 5b shows the output DPAG of an example cyclic RCM from Figure 5a. The edges between nodes [A].X1, [A].X2 and [A, AB, B, AB, A].X1, [A, AB, B, AB, A].X2 satisfies

---

[3]Since relational d-separation is sound and complete for AGG (Maier, Marazopoulou, and Jensen 2013)

the conditions given in Proposition 3. It means they could be part of a cycle but it is not possible to confirm that based on the output alone.

Experiments

In this section, we examine the effectiveness of RCD for cyclic RCMs using both synthetically generated cyclic RCMs satisfying relational acyclification criteria and a demonstration with a real-world dataset. Since there is no other algorithm designed to discover cyclic RCMs, we compare against the vanilla RCD with d-separation oracle.

Experimental Setup

We follow the procedure introduced by Maier et al. (2013) for synthetic data generation except we allow feedback loops. We generate 100 random cyclic relational causal models over randomly generated schema for each of the following combinations: entities (1–3); relationships (one less than the number of entities) with cardinalities selected uniformly at random; attributes per item drawn from Pois($\lambda = 1$) + 1; and the number of relational dependencies (4, 6, 8, 10, 12) limited by a hop threshold of 2 and at most 3 parents per variable. Moreover, for a consistent evaluation, we select models with maximum density of σ-AGGs within a specific range (0.05 to 0.1 for multiple entity types and 0.3 to 0.5 for single entity type). We enforce a feedback loop among the dependencies. Note that a single feedback loop can introduce arbitrary length cycles based on the structure of the model. This procedure yields a total of 15,000 synthetic models. We refer to the version of RCD with d-separation and σ-separation oracles as d-RCD and σ-RCD respectively.[4]

Evaluation

The goal of the evaluation is to compare the learned causal models with the true causal models. However, the output object for cyclic RCMs is PAGs instead of CPDAGs. Moreover, it is expected that the skeleton of the output PAG might be different from the true causal model. For this reason, we evaluate the algorithms based on the ancestral relationships. We identify the ancestral relationships entailed by the output and the σ-AGG of the true model and report F1-score for comparison. For a sound and complete algorithm, we expect to see perfect F1-scores. Moreover, we consider the identification criterion given in Proposition 3 and evaluate the algorithms based on their ability to correctly identify edges as possible cycle candidates. We report F1-score for this evaluation as well.

Results

Figure 6 shows the comparison of d-RCD and σ-RCD based on *isPossibleAncestor* (top row) and *isPossibleCycle* (bottom row) queries on synthetically generated relational models. The columns represent the increased number of entity types (left to right). The x-axis shows the number of dependencies and y-axis shows F1-scores. In the leftmost column, we see the results for single entity models. The top left and bottom
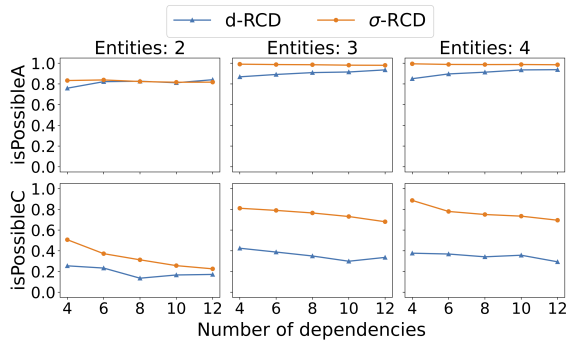
---

[4]Code available at https://github.com/edgeslab/sRCD

Figure 6: Comparison of F1-score for d-RCD and σ-RCD based on *isPossibleAncestor* (top row) and *isPossibleCycle* (bottom row) queries. The number of entity types increases from left to right.
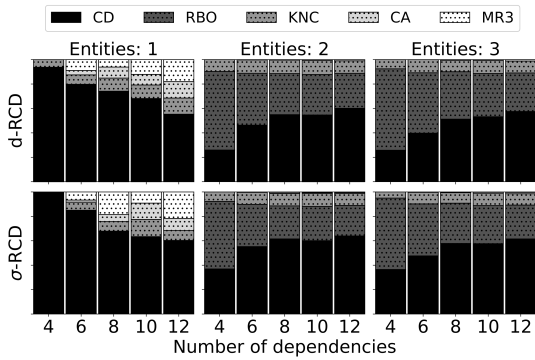

Figure 7: Frequency of edge orientation rules for d-RCD (top) and σ-RCD for different numbers of entity types and dependencies.

left figures are equivalent to running the PC algorithm with d- and σ-separation oracles respectively. The rest of the figures represent proper relational models. In the top row, we see perfect F1-score for σ-RCD for relational models. However, for single entity models, F1-score is slightly lower than d-RCD. However, we recorded 100% recall for σ-RCD. We see a general upward trend from left to right which is intuitive since higher number of dependencies make the models increasingly denser. On the other hand, for isPossibleCycle query the trend is generally downward since denser models make the identification harder.

Figure 7 shows the percentage of orientation rules used for d-RCD (top row) and σ-RCD (bottom row). The leftmost column refers to the single entity case where no RBO is in effect. We can see some subtle differences in the distribution of rules for d-RCD and σ-RCD. For the small number of dependencies (i.e. 4) only CD (collider detection) rule activates with σ-RCD where d-RCD utilizes both CD and KNC (known non-collider). The increased number of dependencies shows the difference in the overall distribution. For the middle and right column, a significant difference is seen in the percentage of times rule MR3 (Meek rule 3) is executed for σ-RCD compared to d-RCD. These differences indicate
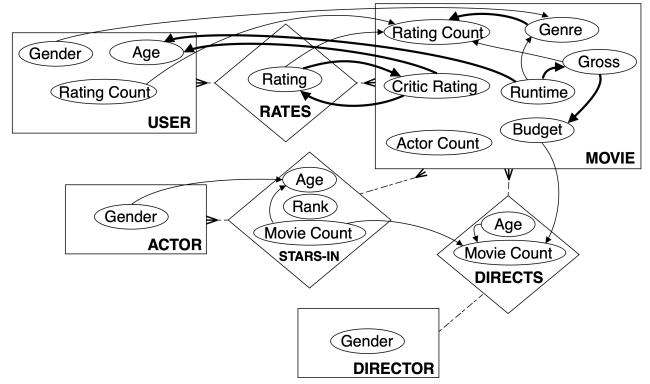

Figure 8: A possible cyclic relational model of MovieLens+ based on the output of RCD (Maier et al. 2013).

that the algorithms learn fundamentally different structures.

Demonstration on Real-world Data

Maier et al. (2013) show the output of RCD on a sample of MovieLens dataset (www.grouplens.org) based on an approximate conditional independence test using the significance of coefficients in linear regression. Their output contains undirected edges which are potential candidates for cycle edges. Figure 8 shows a possible cyclic relational model which corresponds to the original output. Following Proposition 3, we can infer that the edge between *[Movie].Rating Count* and *[Movie].Genre* cannot be part of any cycles or feedback loops. Some undirected edges can be oriented based on domain knowledge (i.e. Budget can cause gross income but not the other way around). There exist many possible orientations of dependencies that agrees with the RCD output. We show one plausible case with a feedback loop between *user rating* and *critic ratings* of a movie. It is possible that rating information is public and users and critics influence each other with their ratings. However, it is also possible that there exists one or more unobserved confounders which influence both *user ratings* and *critic ratings*.

Conclusion

Feedback loops and cycles are mostly avoided in the study of causal inference and discovery. Despite several methods developed for cyclic causal discovery from i.i.d samples, no causal discovery algorithm exists for cyclic relational causal models even though cycles are ubiquitous in real-world relational systems. In this work, we introduce relational acyclification operation which facilitates the theoretical guarantees for identifiability of such models. We prove that an existing state-of-the-art relational discovery algorithm, RCD, is sound and complete for cyclic relational causal models for which valid relational acyclification exists. To the best of our knowledge, this discovery is the first of its kind. We hope that this work will play an important role in the study of mutual influence and interference in complex relational systems.

Acknowledgments

References

Ahsan, R.; Fatemi, Z.; Arbour, D.; and Zheleva, E. 2022. Relational Causal Models with Cycles: Relational Causal Models with Cycles: Representation and Reasoning. In *1st Conference on Causal Learning and Reasoning (CLeaR 2022)*.

Bakshy, E.; Hofman, J. M.; Mason, W. A.; and Watts, D. J. 2011. Everyone's an influencer: quantifying influence on Twitter. In *WSDM*.

Bakshy, E.; Messing, S.; and Adamic, L. A. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239): 1130–1132.

Bongers, S.; Forré, P.; Peters, J.; and Mooij, J. M. 2021. Foundations of structural causal models with cycles and latent variables. *The Annals of Statistics*, 49(5).

Forré, P.; and Mooij, J. M. 2017. Markov properties for graphical models with cycles and latent variables. *arXiv preprint arXiv:1710.08775*.

Getoor, L.; Friedman, N.; Koller, D.; Pfeffer, A.; and Taskar, B. 2007. Probabilistic relational models. *Introduction to statistical relational learning*, 8.

Heckerman, D.; Meek, C.; and Koller, D. 2007. Probabilistic entity-relationship models, PRMs, and plate models. *Introduction to statistical relational learning*, 201–238.

Lee, S.; and Honavar, V. 2020. Towards robust relational causal discovery. In *Uncertainty in Artificial Intelligence*, 345–355. PMLR.

Lee, S.; and Honavar, V. G. 2015. Lifted Representation of Relational Causal Models Revisited: Implications for Reasoning and Structure Learning. In *ACI@UAI*.

Lee, S.; and Honavar, V. G. 2016a. A Characterization of Markov Equivalence Classes of Relational Causal Models under Path Semantics. In *UAI*.

Lee, S.; and Honavar, V. G. 2016b. On Learning Causal Models from Relational Data. In *AAAI*.

Maier, M.; Marazopoulou, K.; Arbour, D.; and Jensen, D. 2013. A sound and complete algorithm for learning causal models from relational data. *UAI 2013*.

Maier, M.; Marazopoulou, K.; and Jensen, D. 2013. Reasoning about independence in probabilistic models of relational data. *arXiv preprint arXiv:1302.4381*.

Mooij, J. M.; and Claassen, T. 2020. Constraint-Based Causal Discovery using Partial Ancestral Graphs in the presence of Cycles. In *Conference on Uncertainty in Artificial Intelligence*, 1159–1168. PMLR.

Ogburn, E. L.; Shpitser, I.; and Lee, Y. 2020. Causal inference, social networks and chain graphs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 183(4): 1659–1676.

Pearl, J.; et al. 2000. Models, reasoning and inference. *Cambridge, UK: CambridgeUniversityPress*, 19.

Richardson, T. 1996. A discovery algorithm for directed cyclic graphs. In *Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence*.

Rothenhäusler, D.; Heinze, C.; Peters, J.; and Meinshausen, N. 2015. BACKSHIFT: Learning causal cyclic graphs from unknown shift interventions. *Advances in Neural Information Processing Systems*, 28.

Spirtes, P.; Glymour, C. N.; Scheines, R.; Heckerman, D.; Meek, C.; Cooper, G.; and Richardson, T. 2000. *Causation, prediction, and search*. MIT press.

Strobl, E. V. 2019a. A constraint-based algorithm for causal discovery with cycles, latent variables and selection bias. *International Journal of Data Science and Analytics*, 8(1): 33–56.

Strobl, E. V. 2019b. A constraint-based algorithm for causal discovery with cycles, latent variables and selection bias. *International Journal of Data Science and Analytics*, 1–24.