σ-Maximal Ancestral Graphs[1]

Binghua Yao & Joris M. Mooij

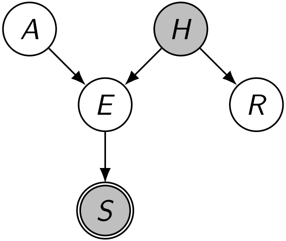
binghuayao71210@gmail.com & J.M.Mooij@uva.nl

University of Amsterdam

[1]The presentation of this paper at the conference was financially supported by the Amsterdam ELLIS Unit.

An Example from Richardson (1998)

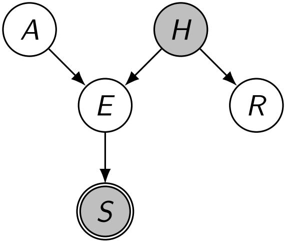
Suppose we conduct a randomized trial of an ineffective drug that causes unpleasant side-effects:



- A: Treatment assignment (randomized)
- E: Side-effects of the drug
- H: Patient's general health level
- R: Recovery speed
- S: Whether the patient stays in the study

An Example from Richardson (1998)

Suppose we conduct a randomized trial of an ineffective drug that causes unpleasant side-effects:

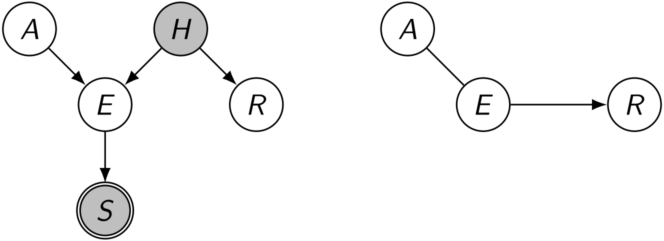


Apply the *d*-separation criterion (assuming faithfulness):

$$A \overset{d}{\not\perp} R \mid S \;\Rightarrow\; X_A \not\perp\!\!\!\perp X_R \mid X_S$$

Patients in the treatment group who remain in the study tend to be healthier than those in the control group, since unhealthy patients taking the drug are more likely to drop out!

Why Use Ancestral Graphs



Key idea: It is not possible to accurately represent a causally insufficient system using a Directed Acyclic Graph (DAG) over observed variables alone.

Ancestral graphs allow us to capture the influence of latent (selection) variables in the causal process that generates the data.

Maximal Ancestral Graphs

Maximal Ancestral Graphs (MAGs), introduced by Richardson and Spirtes (2002), provide an abstract representation of DAGs in the presence of latent (selection) variables.

Definition (MAG)

A *MAG* is a mixed graph H with edge types $\{\rightarrow, \leftarrow, \leftrightarrow, -\}$, satisfying the following conditions:

1. **Simplicity:** At most one edge exists between any two distinct nodes, and no node has a self-loop.
2. **Ancestrality:**
   - If H contains an anterior path $a \multimap \cdots \multimap b$ for $a \neq b$, then it must not contain an edge $a \leftarrow\!* \; b$.
   - If H contains an undirected edge $a - b$, then no node c satisfies $c *\!\!\rightarrow a$.
3. **Maximality:** No inducing path exists between any two non-adjacent nodes in H.

Representing DAGs by MAGs

MAGs encode both **ancestral relations** and *d*-**separation properties** of the DAGs they represent.

Definition (Representing Rules)

Let H be a MAG with nodes V, and let G be a DAG with nodes $V^+ = V \cup L \cup S$. We say that H *represents* G *given* S if all of the following conditions hold:

1. Two distinct nodes $a, b \in V$ are adjacent in H if and only if there exists an inducing path between a and b given $L \cup S$ in G.
2. If H contains the edge $a \leftarrow\!\!* \, b$, then $a \notin \mathrm{Anc}_G(\{b\} \cup S)$.
3. If H contains the edge $a \,-\!\!* \, b$, then $a \in \mathrm{Anc}_G(\{b\} \cup S)$.

Example: a MAG Representing DAGs

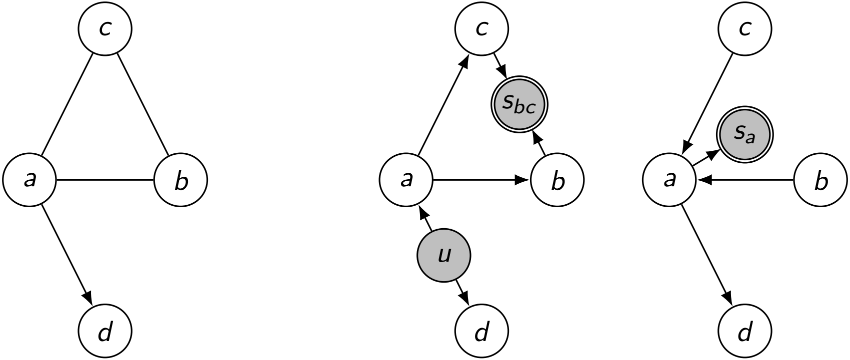


Figure: An example MAG H (left) that represents DAG G_1 (middle) given $S = \{s_{bc}\}$ and DAG G_2 (right) given $S = \{s_a\}$.

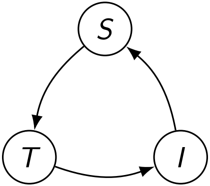
Why MAGs Matter

- MAGs abstract the causal structure of DAGs when latent (selection) variables are present.
- They encode ancestral relations and *d*-separations among observed variables.
- MAGs provide the theoretical foundation for proving the soundness and completeness of the FCI algorithm for causal discovery.
- Their structure supports formal reasoning, including a version of do-calculus tailored to partial observability.

Limitations of MAGs

A fundamental limitation of MAGs is their inability to represent **cyclic causal relationships**, making them unsuitable for systems with feedback mechanisms.

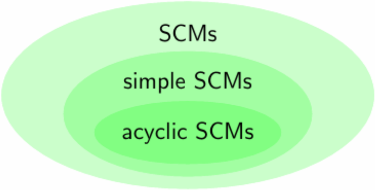
Example: Arctic Climate Feedback Loop



- T: Temperature
- I: Amount of sea ice
- S: Amount of sunlight absorbed

Modeling Cyclic Causality with Simple SCMs

Structural Causal Models (SCMs) can naturally model cycles. We consider here the subclass of **simple SCMs** (Bongers et al., 2021).



Simple SCMs extend acyclic SCMs to allow for (weak) cyclic causal relations while preserving the convenient properties of acyclic SCMs.

Definition (Simple SCM)

An SCM is called **simple** if any subset of its structural equations can be solved uniquely for its associated variables in terms of the other variables that appear in these equations.

d-Separation vs. σ-Separation

In the general cyclic case, the notion of d-separation is too strong (Spirtes, 1995). A solution is to replace d-separation with σ-separation.

Definition (σ-separation (Bongers et al., 2021))

We say that a path $q_0 \ast\!\!-\!\!\ast \cdots \ast\!\!-\!\!\ast\, q_n$ in DMG G is **σ-blocked by Z** if:

1. $q_0 \in Z$ or $q_n \in Z$, or
2. it contains a collider $q_k \notin \mathrm{Anc}_G(Z)$, or
3. it contains a non-collider $q_k \in Z$ that points to a neighboring node on the path in another strongly connected component.

If all paths in G between any node in set X and any node in set Y are σ-blocked by a set Z, we say that X is **σ-separated** from Y by Z, and we write $X \overset{\sigma}{\perp} Y \mid Z$.

σ-separation is different from d-separation in general, but reduces to d-separation in the acyclic case.

Our Contribution: σ-MAGs

We generalize MAGs to a new class of graphs called σ-MAGs, which are capable of representing richer Directed Mixed Graphs (DMGs) that include feedback loops.

For σ-MAGs, we develop a comprehensive theoretical framework:

- We define a tailored m-separation criterion for σ-MAGs, which corresponds to the σ-separation used in DMGs.
- We build on the ideas of Spirtes and Richardson (1996) to characterize Markov equivalence classes of σ-MAGs.

Formal Definition

Definition (σ-MAG)

A *σ-MAG* is a mixed graph H with edge types $\{\rightarrow, \leftarrow, \leftrightarrow, —\}$, satisfying the following conditions:

1. **Simplicity:** At most one edge exists between any two distinct nodes, and no node has a self-loop.
2. **Ancestrality:**
   - If H contains an anterior path $a \mathrel{-\!\!*} \cdots \mathrel{-\!\!*} b$ for $a \neq b$, then it must not contain an edge $a \mathrel{\leftarrow\!\!*} b$.
   - ~~If H contains an undirected edge $a — b$, then no node c satisfies $c \mathrel{*\!\!\rightarrow} a$.~~
3. **σ-maximality:**
   - **Maximality:** No inducing path exists between any two non-adjacent nodes.
   - **σ-completeness:** If H contains a triple $a \mathrel{*\!\!\rightarrow} b — c$, then a and c must be adjacent. Moreover, if $b — d$ is also present, then c and d must be adjacent.

Representing DMGs by σ-MAGs

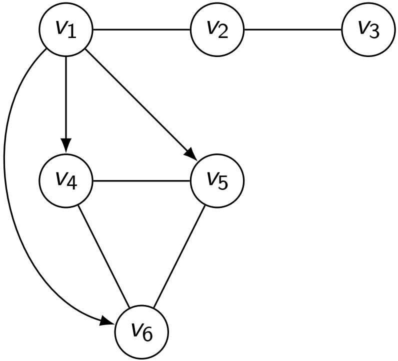
σ-MAGs provide an abstract representation of (possibly cyclic) DMGs:

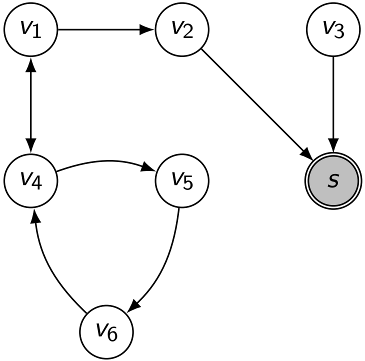
Definition (Representing Rules for σ-MAGs)

Let H be a σ-MAG with nodes V, and let G be a DMG with nodes $V^+ = V \cup S$. We say that H *represents* G *given* S if the following conditions hold:

1. Two nodes $a, b \in V$ are adjacent in H if and only if there exists ~~an inducing path~~ a σ-inducing path between a and b given S in G.
2. If H contains the edge $a \leftarrow\!\!* \; b$, then $a \notin \mathrm{Anc}_G(\{b\} \cup S)$.
3. If H contains the edge $a \rightarrow\!\!* \; b$, then $a \in \mathrm{Anc}_G(\{b\} \cup S)$.

Example: a σ-MAG Representing a DMG





A σ-MAG H (left) represents a DMG G (right) given $S = \{s\}$

m-Separation in σ-MAGs

We generalize the classical m-separation criterion (Richardson and Spirtes, 2002) to σ-MAGs:

Definition (m-separation for σ-MAGs)

We say that a path $q_0 \ast\!\!-\!\!\ast \cdots \ast\!\!-\!\!\ast\ q_n$ in σ-MAG H is m-**blocked by** Z if:

1. it contains a collider $q_k \notin \mathrm{Anc}_H(Z)$, or
2. it contains a non-collider $q_k \in Z$, or
3. it contains a subpath of the form $q_{k-1} \ast\!\!\rightarrow q_k - q_{k+1}$ or $q_{k-1} - q_k \leftarrow\!\!\ast\ q_{k+1}$.

If all paths in H between any node in set X and any node in set Y are m-blocked by a set Z, we say that X is m-**separated** from Y given Z, and we write $X \overset{m}{\perp} Y \mid Z$.

Separation Equivalence

Theorem (Spirtes and Richardson (1996))

Let G be a DAG with nodes $V^+ = V \cup L \cup S$, and let H be a MAG with nodes V that represents G given S. For all subsets $X, Y, Z \subseteq V$, the following holds:

$$X \underset{H}{\overset{m}{\perp}} Y \mid Z \qquad \Longleftrightarrow \qquad X \underset{G}{\overset{d}{\perp}} Y \mid Z \cup S.$$

Theorem (Separation Equivalence)

Let G be a DMG with nodes $V^+ = V \cup S$, and let H be a σ-MAG with nodes V that represents G given S. For all subsets $X, Y, Z \subseteq V$, the following holds:

$$X \underset{H}{\overset{m}{\perp}} Y \mid Z \qquad \Longleftrightarrow \qquad X \underset{G}{\overset{\sigma}{\perp}} Y \mid Z \cup S.$$

m-Markov Equivalence

Two MAGs H_1, H_2 with the same nodes V are *m-Markov equivalent* if for all subsets $X, Y, Z \subseteq V$, the following holds:

$$X \underset{H_1}{\overset{m}{\perp}} Y \mid Z \quad \Longleftrightarrow \quad X \underset{H_2}{\overset{m}{\perp}} Y \mid Z.$$

Theorem (m-Markov Equivalence (Spirtes and Richardson, 1996))

Two MAGs H_1, H_2 with the same nodes V are m-Markov equivalent if and only if the following conditions hold:

1. *H_1 and H_2 have the same adjacencies.*
2. *H_1 and H_2 have the same unshielded colliders.*
3. *Let π be a discriminating path for a node q in H_1, and let π' be the corresponding path in H_2. If π' is also a discriminating path for q, then q is a collider on π in H_1 if and only if it is a collider on π' in H_2.*

Generalized m-Markov Equivalence

Two σ-MAGs H_1, H_2 with the same nodes V are *m-Markov equivalent* if for all subsets $X, Y, Z \subseteq V$, the following holds:

$$X \overset{m}{\underset{H_1}{\perp}} Y \mid Z \quad \Longleftrightarrow \quad X \overset{m}{\underset{H_2}{\perp}} Y \mid Z.$$

Theorem (Generalized m-Markov Equivalence)

Two σ-MAGs H_1, H_2 with the same nodes V are m-Markov equivalent if and only if the following conditions hold:

1. *H_1 and H_2 have the same adjacencies.*
2. *H_1 and H_2 have the same unshielded colliders.*
3. *Let π be a discriminating path for a node q in H_1, and let π' be the corresponding path in H_2. If π' is also a discriminating path for q, then q is a collider on π in H_1 if and only if it is a collider on π' in H_2.*

Discussion

The σ-MAG we propose **extends MAGs** to settings with **cycles**, offering a solid foundation for future work in causal discovery.

- Developing sound and complete **extensions of the FCI algorithm** applicable to data generated by simple SCMs in the presence of selection bias and feedback.
- Serving as a key step towards a generalized **do-calculus** for FCI outputs in settings that include both latent confounding and cyclic causation.

References I

Bongers, S., Forré, P., Peters, J., and Mooij, J. M. (2021). Foundations of structural causal models with cycles and latent variables. *Annals of Statistics*, 49(5):2885–2915.

Richardson, T. and Spirtes, P. (2002). Ancestral graph Markov models. *The Annals of Statistics*, 30(4):962–1030.

Richardson, T. S. (1998). Chain graphs and symmetric associations. In *Learning in graphical models*, pages 231–259. Springer.

Spirtes, P. (1995). Directed cyclic graphical representations of feedback models. In *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence (UAI-95)*, pages 491–499.

Spirtes, P. and Richardson, T. (1996). A polynomial time algorithm for determining DAG equivalence in the presence of latent variables and selection bias. In *Proceedings of the 6th International Workshop on Artificial Intelligence and Statistics (AISTATS-1996)*, pages 489–500.

Thank you for your attention!

A Useful Lemma

Lemma (Fundamental Property of σ-MAGs)

Let H be a σ-MAG. If H contains a triple of the form $a \ast\!\!\rightarrow b — c$, then the edge between a and c is of the same type as the edge between a and b, and the neighbors of b and c are complete.