Algorithmic syntactic causal identification

Dhurim Cakiqi*, Max A. Little*

*School of Computer Science, University of Birmingham, UK

May 20, 2024

Causal Bayes nets (CBNs) are probabilistic models in which *causal influences* between *random variables* are expressed via the use of *graphs* with *nodes* in these graphs being the random variables and *directed edges* indicating the direction of causality between them [Pearl, 2009, Bareinboim et al., 2020]. Every such *directed acyclic graph* (DAG) with *latent* (*unobserved*) nodes has a corresponding *acyclic directed mixed graph* (ADMG) which is obtained from the DAG through *latent projection* which simplifies the DAG whilst preserving its *causal d-separation* properties [Richardson et al., 2012, Pearl, 2009].

For an ADMG with no bidirected edges (thus, no latent variables, equivalent to a CBN over a DAG), it is always possible to derive any interventional distribution from the joint distribution over the variables in the DAG using the *truncated factorization* [Pearl, 2009]. However, more generally, in the presence of unobserved confounding (e.g. models having bidirected edges in the ADMG) this is no longer true and only certain interventional distributions can be derived from the observed variables [Shpitser, 2008, Bareinboim et al., 2020]. Pearl's *do-calculus* [Pearl, 2009] is a set of three algebraic distribution transformations which it has been shown are necessary and sufficient for deriving the interventional distribution where this is possible [Shpitser, 2008]. These algebraic transformations can be applied ad-hoc or, more systematically, using Shpitser's *ID algorithm* to derive a desired interventional distribution [Shpitser, 2008]. More recently, the specific conditions under which any particular interventional distribution can be determined from the observed variables using do-calculus or some other systematic algorithm, has been simplified in terms of *fixing operations* and *reachable subgraphs* in causal ADMGs [Richardson et al., 2012]. Exploiting the same reasoning, Richardson et al. [2012] show how fixing operations can be combined in a simple algorithm which achieves the same result.

This algorithm, as with most algorithms for causal inference, is expressed in terms of CBNs using random variables and classical probabilities where probabilistic conditioning indicates the direction of causal inference in an ADMG. Such causal identification algorithms rely on simultaneous manipulation of the ADMG, tracking the consequence of such manipulations on the corresponding (joint) distribution over that graph. As long as the appropriate *Markov property* holds [Bareinboim et al., 2020], which guarantees the consistency of the distribution with the CBN, then this is a valid procedure for deriving the desired interventional distribution. Nonetheless, there are many practical settings where probabilistic modelling is inappropriate, such as relational databases [Patterson, 2017], hardware description languages, distributed systems modelled by Petri nets and most modern machine learning algorithms [Little, 2019]. In these settings there is no such Markov property therefore it appears that the existing causal identification algorithms are inapplicable in these wider, non-probabilistic applications.

A different and more recently explored direction which might circumvent this limitation is to change the fundamental axiomatic basis of the modelling language to use (*monoidal*) *category theory* instead. This amounts to a fundamental reformulation of CBNs that, rather than organizing causal models around sets, measure theory and graph topology which requires the additional complexity of Markov properties to bind these together, instead views CBNs from the simpler and more abstract vantage

point of *structured compositional processes*. Causal modelling and inference in terms of *string diagrams* representing such processes has shown considerable promise. Building on work by Fong [2013], Cho and Jacobs [2019] formulated the essential concepts of Bayesian reasoning as strings, following which Jacobs et al. [2021] provided an exposition of causal identification under a slightly extended form of the *front-door* causal scenario for *affine Markov categories* [Fritz, 2020]. Since then, string analogues of do-calculus and d-separation have been described [Yin and Zhang, 2022, Fritz and Klingler, 2023] and explicit description of extensions of the categorical string diagram approach to causal modelling in non-probabilistic settings such as machine learning [Cakiqi and Little, 2022].

Symmetric monoidal categories (SMCs) are algebraic structures which capture the notion of simultaneous *sequential* and (in our application) *parallel composition* of maps between types. Examples of such categories include ordinary sets and functions between these sets with the cartesian product indicating parallel composition, the category of sets and relations [Fong, 2013], (*affine*) *Markov categories* of sample spaces with conditional distributions modelled by sets and *probability monads* between them [Fritz, 2020] or other non-deterministic monads in arbitrary semifields [Cakiqi and Little, 2022].

Here we present our main result. Richardson et al. [2012, Theorem 49] is a re-formulation of the *ID algorithm* [Shpitser, 2008] for causal identification in general causal models with latent variables, in terms of fixing operations on conditional ADMGs (CADMGs). In this section we provide a purely syntactic description of the same algorithm which uses only the structural information in the ADMG.

In the ADMG $\mathcal{G}$, consider the set of *cause* $\mathbf{A} \subset \mathbf{V}^{\mathcal{G}}$ and effect variables $\mathbf{Y} \subset \mathbf{V}^{\mathcal{G}}$, where $\mathbf{A}$ and $\mathbf{Y}$ do not intersect. Now consider the set of variables $\mathbf{Y}^{\star} = \mathrm{an}_{\mathcal{G}_{\mathbf{V}^{\mathcal{G}}\setminus\mathbf{A}}}(\mathbf{Y})$ and $\mathbf{D}^{\star}$ the set of districts of the subgraph $\mathcal{G}_{\mathbf{Y}^{\star}}$. The signature of the *syntactic causal effect*, $\Sigma^{\mathcal{G}}_{\mathbf{Y}|\mathrm{do}(\mathbf{A})}$, of $\mathbf{A}$ on $\mathbf{Y}$ is *identifiable* if, for every district $\mathbf{D}' \in \mathbf{D}^{\star}$ the set of nodes $\mathbf{V}^{\mathcal{G}}\setminus\mathbf{D}'$ is a valid fixing sequence. If identifiable, this causal effect is given by the following composite signature manipulation,

$$\Sigma^{\mathcal{G}}_{\mathbf{Y}|\mathrm{do}(\mathbf{A})} = \mathrm{Hide}_{\mathbf{Y}^{\star}\setminus\mathbf{Y}}\left(\bigcup_{\mathbf{D}'\in\mathbf{D}^{\star}} \mathrm{Simple}\left(\mathrm{Fixseq}_{\mathbf{V}^{\mathcal{G}}\setminus\mathbf{D}'}\left(\Sigma^{\mathcal{F}}\right)\right)\right). \tag{1}$$

For the case of the *front-door* ADMG, used for instance in mediation analysis, applying (1) obtains the following signature,

$$\Sigma^{\mathcal{G}}_{Y|\mathrm{do}(X)} = \left(\left\{X, X', Y, Z\right\}, \left\{x, y, z\right\}, \left\{x' : 1 \to X', z : X \to Z, y : X'Z \to Y\right\}\right). \tag{2}$$

This is the purely syntactic categorical analogue of the *front-door adjustment formula*. As an example interpretation, consider the Markov category with discrete sample spaces $X' \mapsto \Omega_X$, $Z \mapsto \Omega_Y$, $Y \mapsto \Omega_Y$ and with conditional distributions $x' \mapsto p\left(X'\right)$, $z \mapsto p\left(Z|X\right)$ and $y \mapsto p\left(Y|X,Z\right)$, then (2) is the familiar discrete interventional distribution [Pearl, 2009],

$$p\left(Y = y|\mathrm{do}\left(X = x\right)\right) = \sum_{z\in\Omega_Z} p\left(Z = z|X = x\right) \sum_{x'\in\Omega_X} p\left(Y = y|X' = x', Z = z\right) p\left(X' = x'\right). \tag{3}$$

As another example interpretation, consider *potential function models* in machine learning in which the variables X, Y, Z are arbitrary sets, then we derive the following potential model

$$f\left(y|do\left(x\right)\right) = \min_{z\in Z} f\left(z|x\right) + \min_{x'\in X}\left[f\left(y|x', z\right) + f\left(x'|\right)\right]. \tag{4}$$

References

E. Bareinboim, J.D. Correa, D. Ibeling, and T. Icard. *On Pearl's Hierarchy and the Foundations of Causal Inference*. ACM Books, 2020.

D. Cakiqi and M.A. Little. Non-probabilistic Markov categories for causal modeling in machine learning. In *ACT 2022: Applied Category Theory*, 2022.

K. Cho and B. Jacobs. Disintegration and Bayesian inversion via string diagrams. *Mathematical Structures in Computer Science*, 29(7):938–971, March 2019.

B. Fong. Causal theories: A categorical perspective on Bayesian networks, 2013.

T. Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Advances in Mathematics*, 370:107239, August 2020.

T. Fritz and A. Klingler. The d-separation criterion in categorical probability. *Journal of Machine Learning Research*, 24(46):1–49, 2023.

B. Jacobs, A. Kissinger, and F. Zanasi. Causal inference via string diagram surgery: A diagrammatic approach to interventions and counterfactuals. *Mathematical Structures in Computer Science*, 31(5): 553–574, 2021.

M.A. Little. *Machine Learning for Signal Processing: Data Science, Algorithms, and Computational Statistics*. Oxford University Press, 2019.

E. Patterson. Knowledge representation in bicategories of relations. page arXiv:1706.00526, 2017.

J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2009.

T.S. Richardson, R.J. Evans, J.M. Robins, and I. Shpitser. Nested Markov properties for acyclic directed mixed graphs. In *UAI'12: Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*, page 13. ACM, 2012.

I. Shpitser. Complete identification methods for the causal hierarchy. *Journal of Machine Learning Research*, pages 1941–1979, 2008.

Y. Yin and J. Zhang. Markov categories, causal theories, and the do-calculus. page arXiv:2204.04821, 2022.