Markov Categories in Causal Inference
HPS/PI/CS 110 Paper 2

Pearl Chen

March 1, 2024

1 Introduction

Category theory is the study of mathematical structures. Subsequently, viewing mathematical objects from a categorical perspective can shed light on the nature of their properties. The process often entails investigating what set of necessary or sufficient constraints can produce a certain desired structure, or what generalizations can be made from these constraints to further the known results. These are precisely the motivation behind what this paper will discuss regarding the results of Yin and Zhang on causal theory and Markov categories [YZ22]. The primary result we focus on will be the t-separation conditions on three causal interpretations [YZ22, Corollary 4.12], which builds on top of Fong's causal theory formulation and [Fon13] induces Pearl's do-calculus.

Specifically, the result at hand capture the necessary and sufficient t-separation conditions for conditional independence, conditionality, and conditionally irrelevance. The conditions specified come from the constraints in causal screen-off morphisms in a Markov category [YZ22, Theorem 4.10], which is derived from the paper's t-separation conditions on the decomposition of causal effect morphisms [YZ22, Theorem 4.8]. The t-separation here can be thought of as a looser variant of d-separation, and the three rules in Pearl's do-calculus can be derived as a specific case of the results.

2 Preliminaries

We borrow all terminology from the paper, and will omit here the category theory definitions described in §2 (namely objects, morphisms, category, strict/symmetric monoidal category, and Markov category). However, it is worth highlighting that the categorical description of Markov Categories involves simply endowing a semigroup structure (symmetric monoidal structure) on any generic category C, so that a Markov category can be seen as a co-commutative coalgebra over the objects in C. The only additional constraint is on the discard morphism, which is very innocent. There is nothing inherently causal in this definition, which goes to show the generality of any results derived under this framework. In causal models alone, the authors note that for example SEM models can be described by the Markov category Set, while a probability law together with intervention distributions over a set of causal variables (causal Bayesian networks) can be described by the Markov category $\mathsf{FinStoch}$ of finite sets and stochastic matrix morphisms.

The causality component comes into play in the definition of *causal theory* $\mathsf{Cau}(G)$ of a DAG $G = (V, A, s, t)$, which is defined to be a free Markov category of words over V with morphisms generated from only the causal arrows in G along with duplicates and discards [Fon13]. In this sense the morphisms are purely causal. Then, a morphism $f : u \to v$ between words u, v translates to a causal relationship from the variables in u to those in v. This is denoted as $[v||u]$ or the causal effect of u on v. In the case $[v||u]$ is nontrivial, there exists a purely causal relationship. It is worth noting that the authors remedy this categorical formulation so that all words considered consists of no duplicate variables (singular) and are unordered (after total ordering).

One important definition in the result is t-separation [YZ22, Definition 4.7]. This definition is graphical in nature, making it comparable to d-separation. The authors define X to be *forward-t-separated* from Y by Z if all directed paths from any $i \in X$ to $j \in Y$ not crossing X, Y contains some

$k \in Z$. X is *backward-t-separated* from Y by Z if all paths π from any $i_0 \in Y$ to any $i_n \in X$ not crossing X, Y, satisfying

$$\pi = i_0 \leftarrow ... \leftarrow i_j \rightarrow ... \rightarrow i_n$$

for some $j \in [1, n]$, contain some $k \in Z$. X and Y are said to be *t-separated* by Z if they are both forward and backward t-separated by Z. Compared to d-separation [JP16], notice that the conditions for t-separation are identical to the chain and fork rule, but colliders are not addressed in t-separation. However, this does not matter for our result. First, any collider in the conditioning set is not a t-separation in either direction along its path. On the other hand, for colliders not in the conditioning set, equalities in do-calculus can be recovered from the observation that any collider backward t-separates its children and the empty set.

3 Result

3.1 Do-Calculus in the Markov Category

The main result of the paper culminates in the three rules as follows. If M is a Markov category induced by $\mathsf{Cau}(G)$ via strong Markov functor $F : \mathsf{Cau}(G) \to \mathsf{M}$, then

1. If v, w are t-separated by u in G, then $F(v), F(w)$ are *conditionally independent* given $F(u)$ in $F([vw||u])$ in M;
2. If v is backward-t-separated from w by u in G, then $F([w||uw])$ is a *conditional* of $F([vw||u])$ in M;
3. If v is forward-t-separated from w by u in G, then $F(v)$ is *conditionally irrelevant* to $F(w)$ given $F(u)$ over $F([w||uv])$,

where conditionally independent, conditional, and conditionally irrelevant are specifically defined in terms of morphisms in M [YZ22, Definition 4.9], which we will not repeat here. The strong Markov functor F ensures that these rules are applicable regardless of the model in question. Notice these rules only apply in the forward direction, and are not necessity conditions. This is because, interestingly, they are insensitive to faithfulness violations.

As a specific case, when $\mathsf{M} = \mathsf{FinStoch}$, the rules are analogous to Pearl's three rules in do-calculus respectively. For example, conditional independence of v, w given u in $[vw||u]$ implies $f := [vw||u] : u \to v \otimes w$ satisfies

$$\begin{aligned} f &= (1_v \otimes \epsilon_w \otimes \epsilon_v \otimes 1_w) \circ (f \otimes f) \circ \delta_u \\ &= ((1_v \otimes \epsilon_w) \circ f) \otimes ((\epsilon_v \otimes 1_w) \circ f) \circ \delta_u \\ &= [v||u] \otimes [w||u] \circ \delta_u \end{aligned} \tag{1}$$

with the last equality following from [YZ22, Lemma 4.5] since $v \cap w = \varnothing$ is implied by t-separation, for if there exists $a \in v \cap w$ then a t-connects v and w via the trivial path. Letting X, Y, Z respectively denote the sets of random variables represented by u, w, v, we can write

$$P(Y, Z|do(X)) = P(Y|do(X))P(Z|do(X)) \iff P(Y|do(X), Z) = P(Y|do(X)),$$

which is a special case of the first rule of do-calculus with empty non-intervention conditioning set. Similarly, we can derive special cases of rules 2 and 3 as $P(Y|do(X), do(Z)) = P(Y|do(X), Z)$ and $P(Y|do(X), Z) = P(Y|do(X))$ respectively.

The authors extend these special cases to show that, together, rules 2 and 3 entail the full version to be an equivalent formulation of do-calculus. This is also where the aforementioned observation regarding colliders can be used to reconcile the t- and d-separations.

This result illustrates the algebraic nature of the do-calculus with a simpler separation condition, and yields a more general formulation of do-calculus in causal Bayesian networks. From definitions, it also follows that these derived morphisms are "purely causal".

3.2 Faithfulness

The paper briefly mentions that deterministic SEM models always satisfy the decomposition property [YZ22, Theorem 4.8], regardless of backward-t-separation. This causes the converse of the rules described by our result to fail because the decomposition t-separation condition informs the screen-off morphism conditions, which constitutes conditional irrelevance in Rule 3. The decomposability property of $[vw||u]$ over v is analogous to that of probabilistic distributions $P(vw|u) = P(w|vu)P(v|u)$, stating that

$$[vw||u] = (1_v \otimes [w||vu]) \circ ((\delta_v \circ [v||u]) \otimes 1_u) \circ \delta_u.$$

As an example, the graph in Figure 1 from class admits deterministic $P(ZX|Y) = P(Z|Y)P(X|Y)$, but ZX are d-connected conditional on Y. Similarly, notice that XZ are not backward-t-separated by Y. However, in $\mathsf{Cau}(G) = \mathsf{Set}$, the morphism $[XZ||Y] : Y \to X \otimes Z$ is also deterministic, and so are $[X||YZ]$ and $[Y||Z]$, satisfying decomposability quite trivially. Although one may argue that the

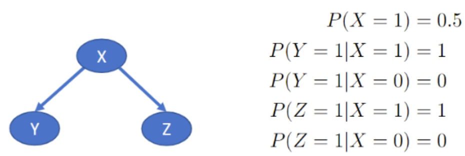

Figure 1: A causal graph G from class, with Structural Equation model $Y = X$ and $Z = X$.

variables examined are constitutions and are not valid causal variables, since the paper clearly addresses these cases somewhat, we think it's fair to say that one area of improvement can be to specify these constraints in theorem statements, which it does not at the moment. The "if and only" conditions are untrue throughout Theorems 4.8, 4.10, and 4.11, leading up to the main result in Corollary 4.12, where the authors finally address the viability of the converse. I personally found this inconsistency to be odd.

Upon closer examination, we can find that the reason the proofs fail in the converse direction is because the string diagram [YZ22, Definition 3.6] induced by causal relations is purely causal, and subsequently cannot be expected to be "faithful" in the sense that a probabilistic independence may exist along some nontrivial morphism. In this case, it makes sense that the proofs which use string diagram properties (morphism properties) would be inapplicable in the converse generally, leading to a main result insensitive to faithfulness violations.

While this may seem like an inefficiency on the part of the formulation, we find it rather illustrative of the quality of faithfulness violations as natural causal consequences. By aiming for a purely causal structure, even in the generalized categorical framework, we are not guaranteed the converse of sufficiency results. As common in abstract mathematics, the value of this result comes from not only its application in general, but the insight we gain about the specific structures studied. In this case, we find that the purely causal component of the do-calculus does not have a converse, hinting that causal discovery cannot be fully inferred starting from a purely causal assumption.

4 Future Work

Following the main result of the paper, it seems natural to conjecture there exist more categorical frameworks for other approaches in causal inference, and it may be interesting to explore what the causal implications in existing mathematical structures may be. For example, given a directed graph $G = (V, A, s, t)$, an object in the $\mathsf{DiGraph}$ category, if we can map it via some functor $F : \mathsf{DiGrph} \to \mathsf{Top}$ into the category of topological spaces while capturing its path properties, it could be illuminating to investigate the parallel between t- or d-separations and the topological separation conditions such as Hausdorff, regular, or normal separations. Perhaps this may allow for the generalization of causal variables to exist in a continuous space, on top of the traditional discrete setting. A probability measure on top of the resulting structure may connect causal inference with topological dynamics, further unifying and generalizing the methods in causal discovery.

References

[Fon13] Brendan Fong. Causal theories: A categorical perspective on bayesian networks. *arXiv preprint arXiv:1301.6201*, 2013.

[JP16] Nicholas P. Jewell Judeas Pearl, Madelyn Glymour. d-separation. In *Causal Inference in Statistics*. 2016.

[YZ22] Yimu Yin and Jiji Zhang. Markov categories, causal theories, and the do-calculus. 4 2022. Submitted on 11 April 2022.

LLM (ChatGPT4) was used to brainstorm in §4 and to generate bibliography.