
UNIFYING CAUSAL INFERENCE AND REINFORCEMENT LEARNING USING HIGHER-ORDER CATEGORY THEORY *

A PREPRINT

Sridhar Mahadevan

Adobe Research and University of Massachusetts, Amherst
smahadev@adobe.com, mahadeva@umass.edu

September 15, 2022

ABSTRACT

We present a unified formalism for structure discovery of causal models and predictive state representation (PSR) models in reinforcement learning (RL) using higher-order category theory. Specifically, we model structure discovery in both settings using simplicial objects, contravariant functors from the category of ordinal numbers into any category. Fragments of causal models that are equivalent under conditional independence – defined as causal horns – as well as subsequences of potential tests in a predictive state representation – defined as predictive horns – are both special cases of horns of a simplicial object, subsets resulting from the removal of the interior and the face opposite a particular vertex. Latent structure discovery in both settings involve the same fundamental mathematical problem of finding extensions of horns of simplicial objects through solving lifting problems in commutative diagrams, and exploiting weak homotopies that define higher-order symmetries. Solutions to the problem of filling “inner” vs “outer” horns leads to various notions of higher-order categories, from Kan complexes to quasicategories and ∞ -category theory. We define the abstract problem of structure discovery in both settings in terms of adjoint functors between the category of universal causal models or universal decision models and its simplicial object representation. In general, the left adjoint functor from a simplicial object X to a category \mathcal{C} is lossy, preserving only relationships up to a certain order defined by homotopical equivalences. In contrast, the right adjoint defining the nerve of a category constructs a lossless encoding of a category as a simplicial object.

Keywords AI · Category Theory · Causal Inference · Simplicial Objects · Machine Learning · Statistics

1 Introduction

Causal inference (Pearl, 2009a; Imbens and Rubin, 2015; Spirtes et al., 2000) and predictive state representations (PSRs) (Singh et al., 2004) in reinforcement learning (Sutton and Barto, 1998), whose roots go back to earlier work on subspace identification in linear systems (Van Overschee and De Moor, 1996) and even earlier work on algebraic theories of context-free languages Chomsky and Schützenberger (1963) and algebraic automata theory (Give’on and Arbib, 1968), both involve structure discovery of a latent variable model through interventions. The use of superficially dissimilar representations – directed acyclic graphs (DAGs) (Pearl, 1989), hybrid undirected and directed graphs (Lauritzen and Richardson, 2002) and hyperedge graphs (Forré and Mooij, 2017; Evans, 2018) in causal inference, versus Hankel matrix and Hilbert space embeddings of dynamical systems – have long obscured their deeper connections. Structure discovery in causal inference and PSRs both involve the determination of a latent structure, which is directional at lower orders, but homotopy equivalences at higher orders induce symmetries. In particular, causal inference involves determining a structure, such as a DAG that encodes direct causal effects between a pair of objects, but multiple DAG models are equivalent because of symmetries induced by conditional independences (Dawid, 2001; Studený et al., 2010a) and correlations induced by latent unobservable confounders that are only revealed over higher-order simplices (e.g., DAGs over $n \geq 3$ vertices). PSRs represent “hidden state” in dynamical systems by constructing a series of tests,

*Draft under revision. Comments welcome.

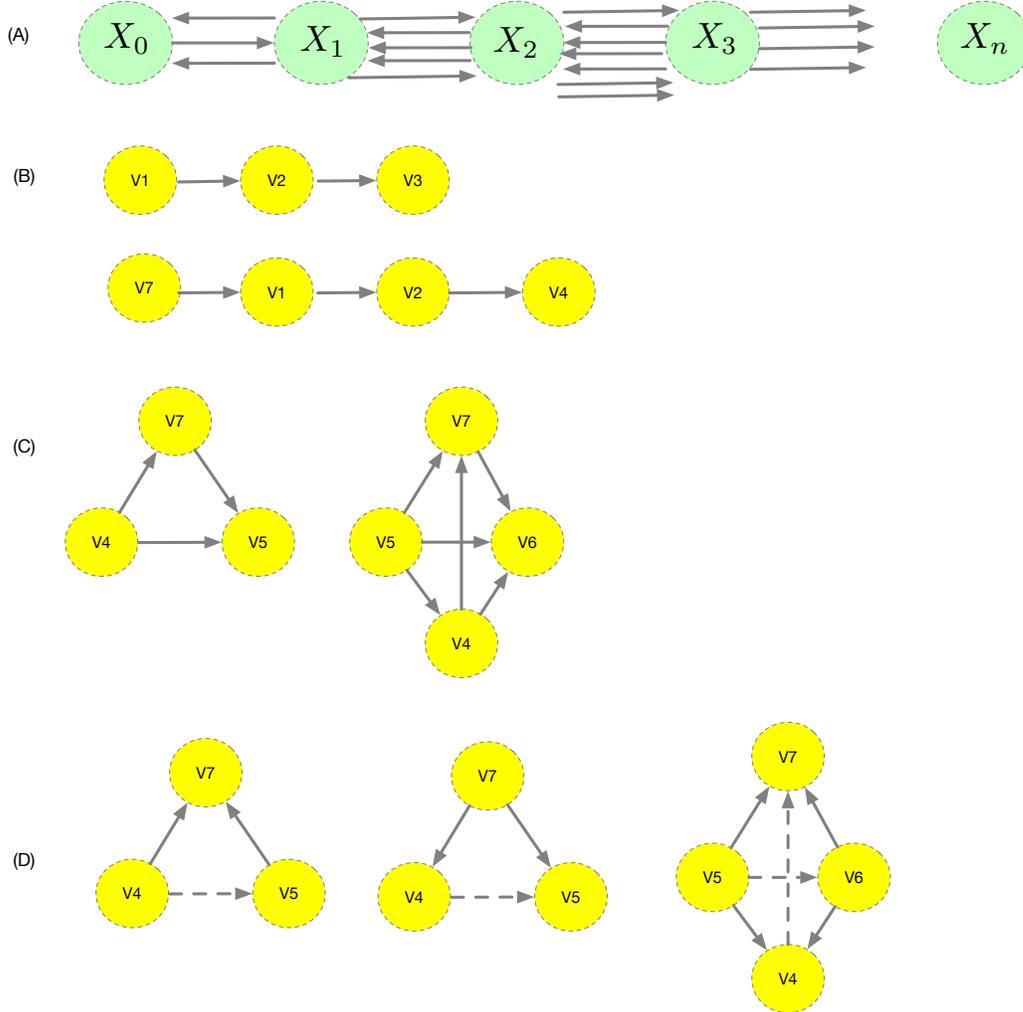


Figure 1: (A): In this paper, we unify structure prediction in causal inference and RL in terms of a simplicial object $X_n, n \geq 0$, contravariant functors from ordinal numbers into a category that originated in algebraic topology (May, 1992), but has since become the foundation for higher-order category theory (Boardman and Vogt, 1973; Joyal, 2002; Lurie, 2022). The internal structure of X_n involves composable morphisms of dimension n . 0th order objects in X_0 involve only identity morphisms from an object to itself. 1th order objects in X_1 define binary arrows between pairs of objects. Higher-order objects in $X_n, n \geq 2$, induce higher-order relationships including symmetries. Structure discovery in causal inference and RL is modeled as the extension problems of “filling horns”, where the horn Λ_k^n is a subset of a contravariant functor of an n -simplex that results from removing its interior and the face opposite the k -vertex. (B): Predictive state horns are partial fragments of potential tests, which can faithfully embed a category of dynamical systems into a simplicial object, and can be seen as a special case of the *nerve* of a category that defines a fully faithful embedding of arbitrary categories in terms of simplicial objects. (C): Causal “inner” horns, fragments of causal models, define a *quasicategory* (Joyal, 2002), a simplicial object where composition of oriented n -simplices involves solving a lifting problem in a commutative diagram under a weak homotopy. (D) “Outer” horns define more general compositional structures that are analyzed in ∞ -category theory (Lurie, 2022).

a free algebra over paired actions and observation that under certain conditions is guaranteed to faithfully embed a dynamical system without requiring the need to explicitly construct “belief states”, probability distributions over latent states, as in a partially observable Markov decision process (POMDP). The core tests define a minimal left \mathbf{k} -module over an Abelian group defined as a free algebra over all possible tests, and induce higher-order symmetries.

We formulate structure discovery of causal and PSR models in terms of *simplicial objects* (see Figure 1 and Figure 2) which were originally introduced as a combinatorial representation in algebraic topology (May, 1992), but have become the foundation for higher-order category theory (Boardman and Vogt, 1973; Joyal, 2002; Lurie, 2022). Formally, simplicial objects May (1992) are contravariant functors $X : \Delta^o \rightarrow \mathcal{C}$ from the category of ordinal numbers, whose objects are $[n] = (0, 1, \dots, n), n \geq 0$, and whose arrows are non-decreasing maps, into an underlying category, such as a universal causal model (Mahadevan, 2022a), or a universal decision model (Mahadevan, 2021a). The 0th order simplices are objects in a universal causal model, or a universal decision model. The 1-simplices define arrows encoding directional relationships between causal or decision objects. Simplicial objects of higher dimensions $n \geq 2$ encode higher-order relationships between groups of objects.

The general problem of horn filling has been solved in different ways in higher-order category theory, including weak Kan complexes Boardman and Vogt (1973), quasi-categories (Joyal, 2002), and ∞ -categories (Lurie, 2022). We will review this literature, and then show the connections to structure discovery in causal inference and RL. Structure prediction in causal inference and RL both can be formulated as filling horns of simplicial objects, subsets of functors that result from interventions that removes the interior and the face opposite a given vertex. These interventions can be defined in terms of a sequence of elementary order-preserving morphisms in the category of ordinal numbers Δ , whose objects are the ordinals $[n] = \{0, 1, \dots, n\}$ and whose arrows are defined as compositions of elementary injections $d_i : [n-1] \rightarrow [n]$ and elementary surjections $s_i : [n] \rightarrow [n-1]$ (see Figure 2). In particular, the units in a causal study or the states of an underlying dynamical system are $x, y, \dots \in X_0$, the 0-simplices of X . If $x, y \in X_0$, and $f \in X_1$, we denote a potential causal effect or action in a dynamical system by the arrow $f : x \rightarrow y$ to denote that $x = d_1 f$ and $y = d_0 f$. Here, d_i is an elementary injection $d_i : [n-1] \rightarrow [n]$ that skips element i . The set of potential causal influences or actions between x and y is denoted as $X_1(x, y)$. The degenerate $s_0 x$ denotes the unit morphism (a cycle) mapping $\mathbf{1}_x : x \rightarrow x$, where $s_i : [n] \rightarrow [n-1]$ is the elementary surjection that repeats element $i \in [n]$. Higher-order simplices $X_i, i \geq 2$ are used to capture higher-order causal effects or action transitions among groups of causal objects or dynamical system states of size $n > 2$. Note in Figure 1, there is one arrow out of X_0 , denoting the self-mapping from an object to itself. The two arrows from X_1 to X_0 represent the head and tail of each morphism in X_1 . There are in general $n+1$ arrows in and out of X_n , for $n > 0$. For example, X_2 has three arrows going back into X_1 , because those are the three arrows that are contained in the 2-simplex.

2 Simplicial Objects

Simplicial objects have long been a foundation for algebraic topology (May, 1999, 1992), and more recently higher-order category theory (Boardman and Vogt, 1973; Joyal, 2002; Lurie, 2022). The category Δ has non-empty ordinals $[n] = \{0, 1, \dots, n\}$ as objects, and order-preserving maps $[m] \rightarrow [n]$ as arrows. An important property in Δ is that any mapping is decomposable as a composition of an injective and a surjective mapping, each of which is decomposable into a sequence of elementary injections $d_i : [n] \rightarrow [n+1]$, which omits $i \in [n]$, and a sequence of elementary surjections $s_i : [n] \rightarrow [n-1]$, which repeats $i \in [n]$. The fundamental simplex $\Delta([n])$ is the presheaf of all morphisms into $[n]$, that is, the representable functor $\Delta(-, [n])$.²

We will denote simplicial objects over the category of sets by $X[n] = X_n$ to mean the functor $[n] \rightarrow X$, mapping both the objects $[n] = \{0, 1, \dots, n\}$ as well as its morphisms to a set X . Thus, X_0 is a set of objects, which are mappings from $[0] = \{0\}$ to X , and as $[0]$ contains only one element, each such mapping must pick out a single element of X . Note that as the only morphism in $[0]$ is $0 \rightarrow 0$, its image gives the identity mapping on each element in the set X . Similarly, X_1 is the functor mapping $[1] = \{0, 1\}$ and its non-identity morphism $0 \rightarrow 1$ to X , which defines arrows between elements of X_0 . Proceeding, $X[2]$ is the functor mapping $[2] = \{0, 1, 2\}$, with its non-identity morphisms $0 \rightarrow 1, 1 \rightarrow 2, 0 \rightarrow 2$ into a category, so it picks out “triangles”, or oriented simplices of order 2.

The Yoneda Lemma MacLane (1971) assures us that an n -simplex $x \in X_n$ can be identified with the corresponding map $\Delta[n] \rightarrow X$. Every morphism $f : [n] \rightarrow [m]$ in Δ is functorially mapped to the map $\Delta[m] \rightarrow \Delta[n]$ in \mathcal{S} . Figure 2) illustrates simplicial objects, in particular showing that any order-preserving morphism can be decomposed into a sequence of elementary degeneracy and face operators.³

²A representable functor is one that can be faithfully embedded in the category of **Sets**.

³The LaTeX source for this diagram is from <http://homepages.math.uic.edu/~jlv/seminars/infinitycats/inftyats.tex>.

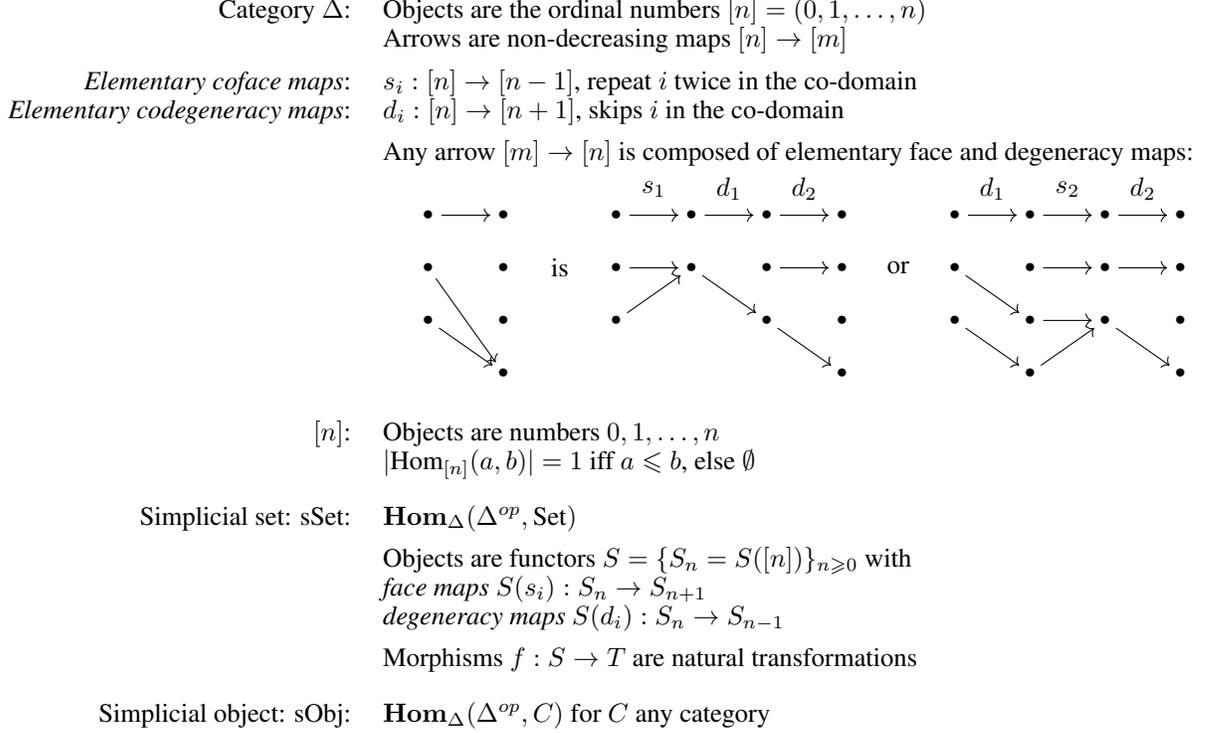


Figure 2: Simplicial objects are contravariant functors from ordinal numbers into any category.

Any morphism in the category Δ can be defined as a sequence of *degeneracy* and *face* operators, where the degeneracy operator $\delta_i : [n-1] \rightarrow [n], 0 \leq i \leq n$ is defined as:

$$\delta_i(j) = \begin{cases} j, & \text{for } 0 \leq j \leq i-1 \\ j+1 & \text{for } i \leq j \leq n-1 \end{cases}$$

Analogously, the face operator $\sigma_j : [n+1] \rightarrow [n]$ is defined as

$$\sigma_j(k) = \begin{cases} j, & \text{for } 0 \leq k \leq j \\ k-1 & \text{for } j < k \leq n+1 \end{cases}$$

The compositions of these arrows define certain well-known properties (May, 1992; Richter, 2020):

$$\begin{aligned} \delta_j \circ \delta_i &= \delta_i \circ \delta_{j-1}, \quad i < j \\ \sigma_j \circ \sigma_i &= \sigma_i \circ \sigma_{j+1}, \quad i \leq j \\ \sigma_j \circ \delta_i(j) &= \begin{cases} \sigma_i \circ \sigma_{j+1}, & \text{for } i < j \\ 1_{[n]} & \text{for } i = j, j+1 \\ \sigma_{i-1} \circ \sigma_j, & \text{for } i > j+1 \end{cases} \end{aligned}$$

Definition 1. The set of all functors $[n] \rightarrow C$ for any category C is denoted as $\mathcal{N}_n(C)$. The image of each abstract face operator $\sigma_j : [n+1] \rightarrow [n]$ under the covariant functor $N_n(C)$ is denoted as $N_n(\sigma_j) = d_j$, and the image of each abstract degeneracy operator $\delta_i : [n-1] \rightarrow [n]$ under the covariant functor $N_n(C)$ is given as $N_n(\delta_i) = s_i$.

2.1 Full and Faithful Embedding of Categories

Definition 2. A **covariant functor** $F : C \rightarrow D$ from category C to category D , and defined as the following:

- An object FX (sometimes written as $F(x)$) of the category D for each object X in category C .

- An arrow $F(f) : FX \rightarrow FY$ in category \mathcal{D} for every arrow $f : X \rightarrow Y$ in category \mathcal{C} .
- The preservation of identity and composition: $F id_X = id_{FX}$ and $(Ff)(Fg) = F(g \circ f)$ for any composable arrows $f : X \rightarrow Y, g : Y \rightarrow Z$.

Definition 3. A **contravariant functor** $F : \mathcal{C} \rightarrow \mathcal{D}$ from category \mathcal{C} to category \mathcal{D} is defined exactly like the covariant functor, except all the arrows are reversed. In the contravariant functor $F : \mathcal{C}^{\text{OP}} \rightarrow \mathcal{D}$, every morphism $f : X \rightarrow Y$ is assigned the reverse morphism $Ff : FY \rightarrow FX$ in category \mathcal{D} .

- For every object X in a category \mathcal{C} , there exists a covariant functor $\mathcal{C}(X, -) : \mathcal{C} \rightarrow \mathbf{Set}$ that assigns to each object Z in \mathcal{C} the set of morphisms $\mathcal{C}(X, Z)$, and to each morphism $f : Y \rightarrow Z$, the pushforward mapping $f_* : \mathcal{C}(X, Y) \rightarrow \mathcal{C}(X, Z)$.
- For every object X in a category \mathcal{C} , there exists a contravariant functor $\mathcal{C}(-, X) : \mathcal{C}^{\text{OP}} \rightarrow \mathbf{Set}$ that assigns to each object Z in \mathcal{C} the set of morphisms $\mathbf{Hom}_{\mathcal{C}}(X, Z)$, and to each morphism $f : Y \rightarrow Z$, the pullback mapping $f^* : \mathbf{Hom}_{\mathcal{C}}(Z, X) \rightarrow \mathcal{C}(Y, X)$. Note how ‘‘contravariance’’ implies the morphisms in the original category are reversed through the functorial mapping, whereas in covariance, the morphisms are not flipped.

Definition 4. Let $\mathcal{F} : \mathcal{C} \rightarrow \mathcal{D}$ be a functor from category \mathcal{C} to category \mathcal{D} . If for all arrows f the mapping $f \rightarrow Ff$

- injective, then the functor \mathcal{F} is defined to be **faithful**.
- surjective, then the functor \mathcal{F} is defined to be **full**.
- bijective, then the functor \mathcal{F} is defined to be **fully faithful**.

Definition 5. A pair of **adjoint functors** is defined as $F : \mathcal{C} \rightarrow \mathcal{D}$ and $G : \mathcal{D} \rightarrow \mathcal{C}$, where F is considered the right adjoint, and G is considered the left adjoint.

$$\mathcal{D} \begin{array}{c} \xrightarrow{G} \\ \xleftarrow{F} \end{array} \mathcal{C}.$$

Definition 6. The **nerve** of a category \mathcal{C} is the set of composable morphisms of length n , for $n \geq 1$. Let $N_n(\mathcal{C})$ denote the set of sequences of composable morphisms of length n .

$$\{C_o \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_n} C_n \mid C_i \text{ is an object in } \mathcal{C}, f_i \text{ is a morphism in } \mathcal{C}\}$$

The set of n -tuples of composable arrows in \mathcal{C} , denoted by $N_n(\mathcal{C})$, can be viewed as a functor from the simplicial object $[n]$ to \mathcal{C} . Note that any nondecreasing map $\alpha : [m] \rightarrow [n]$ determines a map of sets $N_m(\mathcal{C}) \rightarrow N_n(\mathcal{C})$. The nerve of a category \mathcal{C} is the simplicial set $N_{\bullet} : \Delta \rightarrow \mathbf{Set}$, which maps the ordinal number object $[n]$ to the set $N_n(\mathcal{C})$.

The importance of the nerve of a category comes from a key result Lurie (2022), showing it defines a full and faithful embedding of a category:

Theorem 1. (Lurie, 2022, Tag 002Y): The **nerve functor** $N_{\bullet} : \mathbf{Cat} \rightarrow \mathbf{Set}$ is fully faithful. More specifically, there is a bijection θ defined as:

$$\theta : \mathbf{Cat}(\mathcal{C}, \mathcal{C}') \rightarrow \mathbf{Set}_{\Delta}(N_{\bullet}(\mathcal{C}), N_{\bullet}(\mathcal{C}'))$$

In general, the functor G from a simplicial object X to a category \mathcal{C} can be lossy. For example, we can define the objects of \mathcal{C} to be the elements of X_0 , and the morphisms of \mathcal{C} as the elements $f \in X_1$, where $f : a \rightarrow b$, and $d_0 f = a$, and $d_1 f = b$, and $s_0 a, a \in X$ as defining the identity morphisms 1_a . Composition in this case can be defined as the free algebra defined over elements of X_1 , subject to the constraints given by elements of X_2 . For example, if $x \in X_2$, we can impose the requirement that $d_1 x = d_0 x \circ d_2 x$. Such a definition of the left adjoint would be quite lossy because it only preserves the structure of the simplicial object X up to the 2-simplices. The right adjoint from a category to its associated simplicial object, in contrast, constructs a full and faithful embedding of a category into a simplicial set. In particular, the nerve of a category is such a right adjoint.

Example 1. Given a category \mathcal{C} , and an n -simplex σ of the simplicial set $N_n(\mathcal{C})$, which we can identify with the sequence:

$$\sigma = C_o \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_n} C_n$$

the face operator d_0 applied to σ yields the sequence

$$d_0\sigma = C_1 \xrightarrow{f_2} C_2 \xrightarrow{f_3} \dots \xrightarrow{f_n} C_n$$

where the object C_0 is “deleted” along with the morphism f_0 leaving it.

Example 2. Given a category \mathcal{C} , and an n -simplex σ of the simplicial set $N_n(\mathcal{C})$, which we can identify with the sequence:

$$\sigma = C_o \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_n} C_n$$

the face operator d_n applied to σ yields the sequence

$$d_n\sigma = C_0 \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_{n-1}} C_{n-1}$$

where the object C_n is “deleted” along with the morphism f_n entering it.

Example 3. Given a category \mathcal{C} , and an n -simplex σ of the simplicial set $N_n(\mathcal{C})$, which we can identify with the sequence:

$$\sigma = C_o \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_n} C_n$$

the face operator $d_i, 0 < i < n$ applied to σ yields the sequence

$$d_i\sigma = C_0 \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots C_{i-1} \xrightarrow{f_{i+1} \circ f_i} C_{i+1} \dots \xrightarrow{f_n} C_n$$

where the object C_i is “deleted” and the morphisms f_i is composed with morphism f_{i+1} .

Example 4. Given a category \mathcal{C} , and an n -simplex σ of the simplicial set $N_n(\mathcal{C})$, which we can identify with the sequence:

$$\sigma = C_o \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_n} C_n$$

the degeneracy operator $s_i, 0 \leq i \leq n$ applied to σ yields the sequence

$$s_i\sigma = C_0 \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots C_i \xrightarrow{\mathbf{1}_{C_i}} C_i \xrightarrow{f_{i+1}} C_{i+1} \dots \xrightarrow{f_n} C_n$$

where the object C_i is “repeated” by inserting its identity morphism $\mathbf{1}_{C_i}$.

Definition 7. Given a category \mathcal{C} , and an n -simplex σ of the simplicial set $N_n(\mathcal{C})$, which we can identify with the sequence:

$$\sigma = C_o \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_n} C_n$$

we define σ as a **degenerate** simplex if some f_i in the above sequence is an identity morphism, in which case C_i and C_{i+1} are equal.

3 Horns of Simplicial Objects

One of the fundamental contributions of this paper is to show that structure discovery in causal inference and RL both involve solving extensions problems defined by commutative diagrams, which are defined as “horn filling” in higher-order category theory (Boardman and Vogt, 1973; Joyal, 2002; Lurie, 2022).

3.1 Simplicial Subsets and Horns

We now describe ways of modeling subsets of simplicial objects, which are simplicial objects in their own right. We introduce the crucial concept of the *boundary* and *horn* of a simplicial set.

Definition 8. The **standard simplex** Δ^n is the simplicial set defined by the construction

$$([m] \in \Delta) \mapsto \mathbf{Hom}_\Delta([m], [n])$$

By convention, $\Delta^{-1} := \emptyset$. The standard 0-simplex Δ^0 maps each $[n] \in \Delta^{op}$ to the single element set $\{\bullet\}$.

Definition 9. Let S_\bullet denote a simplicial set. If for every integer $n \geq 0$, we are given a subset $T_n \subseteq S_n$, such that the face and degeneracy maps

$$d_i : S_n \rightarrow S_{n-1} \quad s_i : S_n \rightarrow S_{n+1}$$

applied to T_n result in

$$d_i : T_n \rightarrow T_{n-1} \quad s_i : T_n \rightarrow T_{n+1}$$

then the collection $\{T_n\}_{n \geq 0}$ defines a **simplicial subset** $T_\bullet \subseteq S_\bullet$.

Definition 10. The **boundary** is a simplicial set $(\partial\Delta^n) : \Delta^{op} \rightarrow \mathbf{Set}$ defined as

$$(\partial\Delta^n)([m]) = \{\alpha \in \mathbf{Hom}_\Delta([m], [n]) : \alpha \text{ is not surjective}\}$$

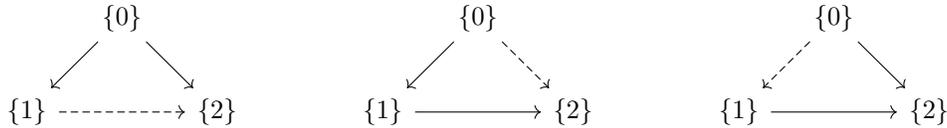
Note that the boundary $\partial\Delta^n$ is a simplicial subset of the standard n -simplex Δ^n .

Definition 11. The **Horn** $\Lambda_i^n : \Delta^{op} \rightarrow \mathbf{Set}$ is defined as

$$(\Lambda_i^n)([m]) = \{\alpha \in \mathbf{Hom}_\Delta([m], [n]) : [n] \not\subseteq \alpha([m]) \cup \{i\}\}$$

Intuitively, the Horn Λ_i^n can be viewed as the simplicial subset that results from removing the interior of the n -simplex Δ^n together with the face opposite its i th vertex. Lifting problems provide elegant ways to define basic notions in a wide variety of areas in mathematics. For example, the notion of injective and surjective functions, the notion of separation in topology, and many other basic constructs can be formulated as solutions to lifting problems.

Let us illustrate this abstract definition with the following diagrams. Consider the problem of composing 1-dimensional simplices to form a 2-dimensional simplicial object. Each simplicial subset of an n -simplex induces a *horn* Λ_k^n , where $0 \leq k \leq n$. Intuitively, a horn is a subset of a simplicial object that results from removing the interior of the n -simplex and the face opposite the i th vertex. Consider the three horns defined below. The dashed arrow \dashrightarrow indicates edges of the 2-simplex Δ^2 not contained in the horns.



The inner horn Λ_1^2 is the middle diagram above, and admits an easy solution to the “horn filling” problem of composing the simplicial subsets. The two outer horns on either end pose a more difficult challenge. A considerable elaboration of the theoretical machinery in category theory is required to describe the various solutions proposed, which led to different ways of defining higher-order category theory (Boardman and Vogt, 1973; Joyal, 2002; Lurie, 2022), which we summarize below.

Definition 12. Let \mathcal{C} be a category. A **lifting problem** in \mathcal{C} is a commutative diagram σ in \mathcal{C} .

$$\begin{array}{ccc} A & \xrightarrow{\mu} & X \\ \downarrow f & & \downarrow p \\ B & \xrightarrow{\nu} & Y \end{array}$$

Definition 13. Let \mathcal{C} be a category. A **solution to a lifting problem** in \mathcal{C} is a morphism $h : B \rightarrow X$ in \mathcal{C} satisfying $p \circ h = \nu$ and $h \circ f = \mu$ as indicated in the diagram below.

$$\begin{array}{ccc} A & \xrightarrow{\mu} & X \\ \downarrow f & \nearrow h & \downarrow p \\ B & \xrightarrow{\nu} & Y \end{array}$$

Definition 14. Let \mathcal{C} be a category. If we are given two morphisms $f : A \rightarrow B$ and $p : X \rightarrow Y$ in \mathcal{C} , we say that f has the **left lifting property** with respect to p , or that p has the **right lifting property** with respect to f if for every pair of morphisms $\mu : A \rightarrow X$ and $\nu : B \rightarrow Y$ satisfying the equations $p \circ \mu = \nu \circ f$, the associated lifting problem indicated in the diagram below.

$$\begin{array}{ccc} A & \xrightarrow{\mu} & X \\ \downarrow f & \nearrow h & \downarrow p \\ B & \xrightarrow{\nu} & Y \end{array}$$

admits a solution given by the map $h : B \rightarrow X$ satisfying $p \circ h = \nu$ and $h \circ f = \mu$.

Example 5. Given the paradigmatic non-surjective morphism $f : \emptyset \rightarrow \{\bullet\}$, any morphism p that has the right lifting property with respect to f is a **surjective mapping**.

$$\begin{array}{ccc} \emptyset & \xrightarrow{\mu} & X \\ \downarrow f & \nearrow h & \downarrow p \\ \{\bullet\} & \xrightarrow{\nu} & Y \end{array}$$

Example 6. Given the paradigmatic non-injective morphism $f : \{\bullet, \bullet\} \rightarrow \{\bullet\}$, any morphism p that has the right lifting property with respect to f is an **injective mapping**.

$$\begin{array}{ccc} \{\bullet, \bullet\} & \xrightarrow{\mu} & X \\ \downarrow f & \nearrow h & \downarrow p \\ \{\bullet\} & \xrightarrow{\nu} & Y \end{array}$$

Definition 15. If S is a collection of morphisms in category \mathcal{C} , a morphism $f : A \rightarrow B$ has the **left lifting property with respect to S** if it has the left lifting property with respect to every morphism in S . Analogously, we say a morphism $p : X \rightarrow Y$ has the **right lifting property with respect to S** if it has the right lifting property with respect to every morphism in S .

Definition 16. Let C and C' be a pair of objects in a category \mathcal{C} . We say C is a **retract** of C' if there exists maps $i : C \rightarrow C'$ and $r : C' \rightarrow C$ such that $r \circ i = \text{id}_C$.

Definition 17. Let \mathcal{C} be a category. We say a morphism $f : C \rightarrow D$ is a **retract of another morphism** $f' : C \rightarrow D$ if it is a retract of f' when viewed as an object of the functor category $\mathbf{Hom}([1], \mathcal{C})$. A collection of morphisms T of \mathcal{C} is **closed under retracts** if for every pair of morphisms f, f' of \mathcal{C} , if f is a retract of f' , and f' is in T , then f is also in T .

3.2 Fibrations and Kan Complexes

Definition 18. Let $f : X \rightarrow S$ be a morphism of simplicial sets. We say f is a **Kan fibration** if, for each $n > 0$, and each $0 \leq i \leq n$, every lifting problem

$$\begin{array}{ccc} \Lambda_i^n & \xrightarrow{\sigma_0} & X \\ \downarrow & \nearrow \sigma & \downarrow f \\ \Delta^n & \xrightarrow{\bar{\sigma}} & S \end{array}$$

admits a solution. More precisely, for every map of simplicial sets $\sigma_0 : \Lambda_i^n \rightarrow X$ and every n -simplex $\bar{\sigma} : \Delta^n \rightarrow S$ extending $f \circ \sigma_0$, we can extend σ_0 to an n -simplex $\sigma : \Delta^n \rightarrow X$ satisfying $f \circ \sigma = \bar{\sigma}$.

Example 7. Given a simplicial set X , then a projection map $X \rightarrow \Delta^0$ that is a Kan fibration is called a **Kan complex**.

Example 8. Any isomorphism between simplicial sets is a Kan fibration.

Example 9. The collection of Kan fibrations is closed under retracts.

Definition 19. Let X and Y be simplicial sets, and suppose we are given a pair of morphisms $f_0, f_1 : X \rightarrow Y$. A **homotopy** from f_0 to f_1 is a morphism $h : \Delta^1 \times X \rightarrow Y$ satisfying $f_0 = h|_{0 \times X}$ and $f_1 = h|_{1 \times X}$.

3.3 Higher-order Categories

We now formally introduce higher-order categories, building on the framework proposed in a number of formalisms (Boardman and Vogt, 1973; Joyal, 2002; Lurie, 2022).

Definition 20. (Lurie, 2022) An ∞ -category is a simplicial object S_\bullet which satisfies the following condition:

- For $0 < i < n$, every map of simplicial sets $\sigma_0 : \Lambda_i^n \rightarrow S_\bullet$ can be extended to a map $\sigma : \Delta^n \rightarrow S_\bullet$.

This definition emerges out of a common generalization of two other conditions on a simplicial set S_i :

1. **Property K:** For $n > 0$ and $0 \leq i \leq n$, every map of simplicial sets $\sigma_0 : \Lambda_i^n \rightarrow S_\bullet$ can be extended to a map $\sigma : \Delta^n \rightarrow S_\bullet$.
2. **Property C:** for $0 < i < n$, every map of simplicial sets $\sigma_0 : \Lambda_i^n \rightarrow S_i$ can be extended *uniquely* to a map $\sigma : \Delta^n \rightarrow S_\bullet$.

Simplicial objects that satisfy property K were defined above to be Kan complexes. Simplicial objects that satisfy property C above can be identified with the nerve of a category, which yields a full and faithful embedding of a category in the category of sets. Definition 20 generalizes both of these definitions, and was called a *quasicategory* in (Joyal, 2002) and *weak Kan complexes* in (Boardman and Vogt, 1973) when \mathcal{C} is a category.

4 Universal Reinforcement Learning

In this section, we briefly review the framework of universal decision models (Mahadevan, 2021a), where a decision making object is defined as a particular type of category. We will focus in particular on predictive state representations (PSRs) Singh et al. (2004), but we will include a brief review of Markov decision processes (MDPs), which are a stochastic dynamical system widely used in the literature on reinforcement learning. We show that predictive state representations define a simplicial object representation of a category, and prove the Universal PSR theorem based on the fact that the nerve of a PSR defines a full and faithful embedding of the category of PSRs in the category of sets. We show that simplicial object representations of PSRs define a quasicategory.

4.1 UDMs based on Markov Decision Processes

We now briefly describe the category of UDMs, where each object represents a (finite) Markov decision process (MDP) (Puterman, 1994). Recall that an MDP is defined by a tuple $\langle S, A, \Psi, P, R \rangle$, where S is a discrete set of states, A is the discrete set of actions, $\Psi \subset S \times A$ is the set of admissible state-action pairs, $P : \Psi \times S \rightarrow [0, 1]$ is the transition probability function specifying the one-step dynamics of the model, where $P(s, a, s')$ is the transition probability of moving from state s to state s' in one step under action a , and $R : \Psi \rightarrow \mathbb{R}$ is the expected reward function, where $R(s, a)$ is the expected reward for executing action a in state s . MDP homomorphisms can be viewed as a principled way of abstracting the state (action) set of an MDP into a “simpler” MDP that nonetheless preserves some important properties, usually referred to as the stochastic substitution property (SSP).

Definition 21. A UDM MDP homomorphism (Ravindran and Barto, 2003) from object $M = \langle S, A, \Psi, P, R \rangle$ to $M' = \langle S', A', \Psi', P', R' \rangle$, denoted $h : M \rightarrow M'$, is defined by a tuple of surjections $\langle f, \{g_s | s \in S\} \rangle$, where $f : S \rightarrow S', g_s : A_s \rightarrow A'_{f(s)}$, where $h((s, a)) = \langle f(s), g_s(a) \rangle$, for $s \in S$, such that the stochastic substitution property and reward respecting properties below are respected:

$$P'(f(s), g_s(a), f(s')) = \sum_{s'' \in [s']_f} P(s, a, s'') \quad (1)$$

$$R'(f(s), g_s(a)) = R(s, a) \quad (2)$$

Given this definition, the following result is straightforward to prove.

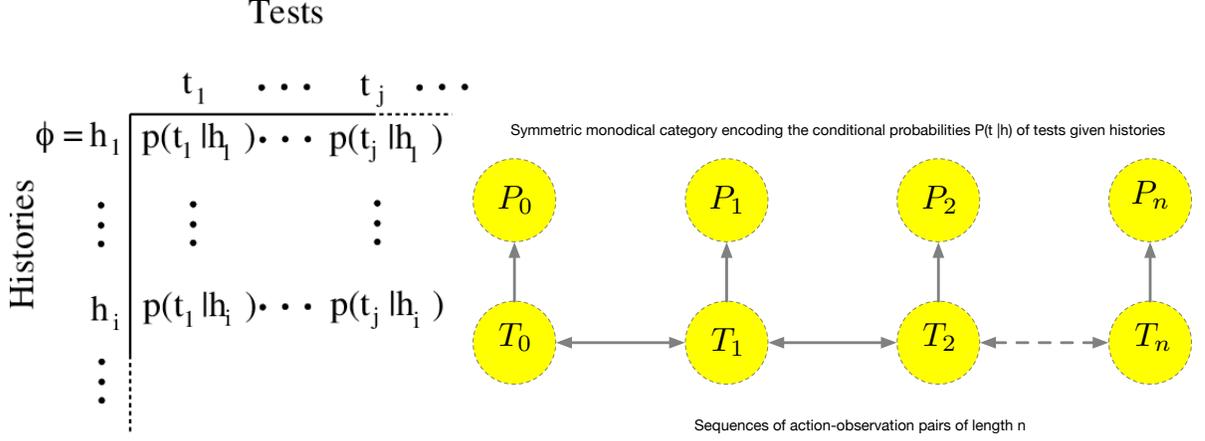


Figure 3: Left: A predictive state representation (PSR) can be defined in terms of its underlying systems dynamics matrix (Singh et al., 2004). The columns of this (infinite) matrix are tests $t = a_1 o_1 \dots a_n o_n$ and the rows are histories $h = a_1 o_1 \dots a_k o_k$. Right: the simplicial object encoded by a PSR. T_n represents the n -simplicial object of tests or histories, and P_n represents a symmetric monoidal category Fong and Spivak (2018) that can be used to encode conditional probabilities $P(t|h)$ of each test t given its history h . Note that as in Figure 2, there are n arrows leading from the n -simplex T_n back to the $n - 1$ simplex T_{n-1} as well as to the $n + 1$ -simplex T_{n+1} , which are not shown for clarity. The conditional probability for each test t as a function of history h , $P(t|h)$ can be defined using an enriched symmetric monoidal category, exactly as defined in (Bradley et al., 2022), which defines sentence completions in English by deep NLP networks using conditional probabilities as enriched categories.

Theorem 2. The UDM category \mathcal{C}_{MDP} is defined as one where each object c is defined by an MDP, and morphisms are given by MDP homomorphisms defined by Equation 1.

Proof: Note that the composition of two MDP homomorphisms $h : M_1 \rightarrow M_2$ and $h' : M_2 \rightarrow M_3$ is once again an MDP homomorphism $h' \circ h : M_1 \rightarrow M_3$. The identity homomorphism is easy to define, and MDP homomorphisms, being surjective mappings, obey associative properties. \square

4.2 UDM Category of Predictive State Representations

We now define the UDM (sub)category \mathcal{C}_{PSR} of predictive state representations (Thon and Jaeger, 2015), based on the notion of homomorphism defined for PSRs proposed in (Soni and Singh, 2007). Recall that a PSR is (in the simplest case) a discrete controlled dynamical system, characterized by a finite set of actions A , and observations O . At each clock tick t , the agent takes an action a_t and receives an observation $o_t \in O$. A *history* is defined as a sequence of actions and observations $h = a_1 o_1 \dots a_k o_k$. A *test* is a possible sequence of future actions and observations $t = a_1 o_1 \dots a_n o_n$. A test is successful if the observations $o_1 \dots o_n$ are observed in that order, upon execution of actions $a_1 \dots a_n$. The probability $P(t|h)$ is a prediction of that a test t will succeed from history h .

A state ψ in a PSR is a vector of predictions of a suite of *core tests* $\{q_1, \dots, q_k\}$. The prediction vector $\psi_h = \langle P(q_1|h) \dots P(q_k|h) \rangle$ is a sufficient statistic, in that it can be used to make predictions for any test. More precisely, for every test t , there is a $1 \times k$ projection vector m_t such that $P(t|h) = \psi_h \cdot m_t$ for all histories h . The entire predictive state of a PSR can be denoted Ψ .

Definition 22. In the UDM category \mathcal{C}_{PSR} defined by PSR objects, the morphism from object Ψ to another Ψ' is defined by a tuple of surjections $\langle f, v_\psi(a) \rangle$, where $f : \Psi \rightarrow \Psi'$ and $v_\psi : A \rightarrow A'$ for all prediction vectors $\psi \in \Psi$ such that

$$P(\psi' | f(\psi), v_\psi(a)) = P(f^{-1}(\psi') | \psi, a) \quad (3)$$

for all $\psi' \in \Psi, \psi \in \Psi, a \in A$.

Theorem 3. The UDM category \mathcal{C}_{PSR} is defined by making each object c represent a PSR, where the morphisms between two PSRs $h : c \rightarrow d$ is defined by the PSR homomorphism defined in (Soni and Singh, 2007).

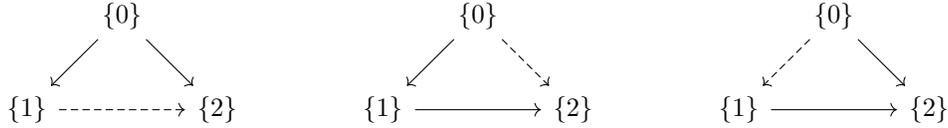
Proof: Once again, given the homomorphism definition in Definition 22, the UDM category \mathcal{P}_{PSR} is easy to define, given the surjectivity of the associated mappings f and v_ψ . \square

4.3 PSR Test Discovery as Horn Filling of a Simplicial Object

The extensive literature on PSRs has explored ways of constructing the set of core tests needed to serve as a sufficient statistic for a full and faithful embedding of any dynamical system that can be represented as a PSR. Here, we just want to illustrate that the process of elaborating a partial test $t = a_1 o_1 \dots a_i o_i$ by adding a new candidate action observation pair $a_{i+1} o_{i+1}$ can be viewed as filling the horn of a simplicial object.

Definition 23. A **predictive horn** of a PSR $\Lambda_i^n : \Delta^{op} \rightarrow \mathcal{C}_{PSR}$ is defined as the simplicial subset of the functor mapping $[n]$ to the category \mathcal{C}_{PSR} defining a PSR model. Recall the horn removes the interior of an n -simplex Δ^n together with the face opposite its i th vertex.

Let us illustrate this notion of horn filling using the same diagram as shown earlier, but this time, interpreting the morphisms shown below as tests in a PSR.



The inner horn Λ_1^2 is the middle diagram above, and admits an easy solution to the ‘‘horn filling’’ problem of composing the simplicial subsets. The two outer horns on either end pose a more difficult challenge. Stated in the language of PSRs, the inner horn filling problem represents the problem of composing two tests, defined as images of abstract tests $t_{01} : 0 \rightarrow 1$ and $t_{12} : 1 \rightarrow 2$ into the category defined by a PSR. Each image of an abstract test leads to an actual test in a PSR model, for example $F(t_{01}) = a_i o_i$ and $F(t_{12}) = a_j o_j$. These two tests can be composed to form a test $F(t_{01} \circ t_{12}) = F(t_{02}) = a_i o_i a_j o_j$. The outer horn filling problem, in contrast, involves inducing a missing test that ‘‘fills in’’ between a test of length 2 and a test of length 1.

First, let us define more precisely how to construct the associated simplicial object from a PSR. We can define this construction following the systems dynamics matrix in Figure 3. Both tests and histories are viewed as n -simplices, constructed from sequences of action-observation pairs. For example, in Figure 3, the 1-simplex encodes a unit test or history $a_i o_j$ of a single action observation pair. The conditional probabilities $P(t|h)$ of each test t given its history h is defined using an enriched symmetric monoidal category \mathcal{P} . Basically, we are treating the process of combining a test t with a history h as a morphism from h to $t \circ h$. The set of all morphisms $\mathbf{Hom}_{\mathcal{C}}(h, th)$ is defined as an enriched symmetric monoidal category that can encode conditional probabilities. (Bradley et al., 2022) gives a detailed treatment of how to model conditional probabilities for the problem of defining the compositional structure of deep NLP models, where for each fragment of an English sentence, like $h = \text{‘‘I went to the grocery store...’’}$, each possible completion, such as $t = \text{‘‘to get some milk’’}$ is defined as a conditional probability $P(t|h)$. This problem is exactly the same as defining the conditional probability of PSR tests, and we refer the reader to (Bradley et al., 2022) for additional details.

Definition 24. The **nerve** of a PSR \mathcal{C} is the set of sequences of composable tests of length n , for $n \geq 1$. Let $N_n(\mathcal{C})$ denote the set of sequences of composable tests

$$\{C_o \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_n} C_n \mid C_i \text{ is an object in } \mathcal{C}, f_i \text{ is a morphism in } \mathcal{C}\}$$

$$\mathcal{N}_{PSR}(\mathcal{C}) \xrightleftharpoons[F]{G} \mathbf{C}_{PSR}.$$

We can use this result to abstractly define a universal PSR theorem, which states that any PSR can be defined completely by its nerve.

Theorem 4. Universal PSR Theorem: The **nerve functor** defined by a PSR $N_\bullet : \mathbf{Cat}_{PSR} \rightarrow \mathbf{Set}$ is fully faithful. More specifically, there is a bijection θ defined as:

$$\theta : \mathbf{Cat}(\mathcal{C}_{PSR}, \mathcal{C}'_{PSR}) \rightarrow \mathbf{Set}_{\Delta}(N_\bullet(\mathcal{C}_{PSR}), N_\bullet(\mathcal{C}'_{PSR}))$$

Theorem 5. PSRs form a quasicategory: The **nerve functor** defined by a PSR $N_\bullet : \mathbf{Cat}_{PSR} \rightarrow \mathbf{Set}$ forms a quasicategory Joyal (2002).

Proof: The proof of both theorems follows directly from more basic results establishing that the nerve of a category is a fully faithful embedding of it. Since we showed above in Theorem 3 that PSRs form a category, the nerve embedding

is fully faithful. It also follows from (Joyal, 2002) that the nerve functor defined by the category of PSRs defines a quasicategory. •

In particular, the nerve of a PSR defines a quasicategory, and this result shows that its nerve is a full and faithful embedding of a category as a simplicial object. The significance of this theorem is it shows that a considerable amount of theoretical machinery in higher-order category theory can be brought to bear on the problem of structure discovery of a PSR. There are many elaborations of this basic result, which we will postpone to a subsequent paper. The definition of nerve leads to an important topological characterization of any category.

Definition 25. The **classifying space** of a PSR \mathcal{C} is the topological space defined by its nerve functor $|N_\bullet(\mathcal{C})|$.

The classifying space of a category gives a way to define an algebraic invariant of a PSR, which we define below as its singular homology.

4.4 Singular Homology of a PSR

We will now describe the singular homology of a PSR. First, we need to define more concretely the topological n -simplex that provides a concrete way to attach a topology to a simplicial object. Our definitions below build on those given in (Lurie, 2022). For each integer n , define the topological space $|\Delta_n|$ realized by the object Δ_n as

$$|\Delta_n| = \{t_0, t_1, \dots, t_n \in \mathbb{R}^{n+1} : t_0 + t_1 + \dots + t_n = 1\}$$

This is the familiar n -dimensional simplex over n variables. For any PSR model, its classifying space $|\mathcal{N}_\bullet(\mathcal{C})|$ defines a topological space. We can now define the *singular n -simplex* is a continuous mapping $\sigma : |\Delta_n| \rightarrow |\mathcal{N}_\bullet(\mathcal{C})|$. Every singular n -simplex σ induces a collection of $n - 1$ -dimensional simplices called *faces*, denoted as

$$d_i\sigma(t_0, \dots, t_{n-1}) = (t_0, t_1, \dots, t_{i-1}, 0, t_i, \dots, t_{n-1})$$

Define the set of all morphisms $\text{Sing}_n(X) = \mathbf{Hom}_{\text{TOP}}(\Delta_n, |\mathcal{N}_\bullet(\mathcal{C})|)$ as the set of singular n -simplices of $|\mathcal{N}_\bullet(\mathcal{C})|$.

Definition 26. For any topological space defined by a PSR $|\mathcal{N}_\bullet(\mathcal{C})|$, the **singular homology groups** $H_*(|\mathcal{N}_\bullet(\mathcal{C})|; \mathbf{Z})$ are defined as the homology groups of a chain complex

$$\dots \xrightarrow{\partial} \mathbf{Z}(\text{Sing}_2(|\mathcal{N}_\bullet(\mathcal{C})|)) \xrightarrow{\partial} \mathbf{Z}(\text{Sing}_1(|\mathcal{N}_\bullet(\mathcal{C})|)) \xrightarrow{\partial} \mathbf{Z}(\text{Sing}_0(|\mathcal{N}_\bullet(\mathcal{C})|))$$

where $\mathbf{Z}(\text{Sing}_n(|\mathcal{N}_\bullet(\mathcal{C})|))$ denotes the free Abelian group generated by the set $\text{Sing}_n(|\mathcal{N}_\bullet(\mathcal{C})|)$ and the differential ∂ is defined on the generators by the formula

$$\partial(\sigma) = \sum_{i=0}^n (-1)^i d_i\sigma$$

Intuitively, a chain complex builds a sequence of vector spaces that can be used to construct an algebraic invariant of a PSR from its classifying space by choosing the left \mathbf{k} module \mathbf{Z} to be a vector space. Each differential ∂ then becomes a linear transformation whose representation is constructed by modeling its effect on the basis elements in each $\mathbf{Z}(\text{Sing}_n(X))$.

5 Universal Causal Model

We now state more formally the Universal Causal Model (UCM) framework (Mahadevan, 2022a), which provides some background to simplicial objects formulation below.

Definition 27. A **universal causal model** (UCM) is defined as a tuple $\langle \mathcal{C}, \mathcal{X}, \mathcal{I}, \mathcal{O}, \mathcal{E} \rangle$ where each of the components is specified as follows:

1. \mathcal{C} is a category of causal objects that interact with each other, and their patterns of interactions are captured by a set of morphisms $\mathbf{Hom}_{\mathcal{C}}(X, Y)$ between object X and Y . Categories where the **Hom** morphisms can be defined as a set are referred to as *locally small* categories, which are the ones of primary interest to us in this paper. In particular, a \mathbf{V} -enriched category \mathcal{C} is one where the **Hom** morphisms are defined over the category \mathbf{V} , which allows capturing additional structure that might exist in the set of morphisms.

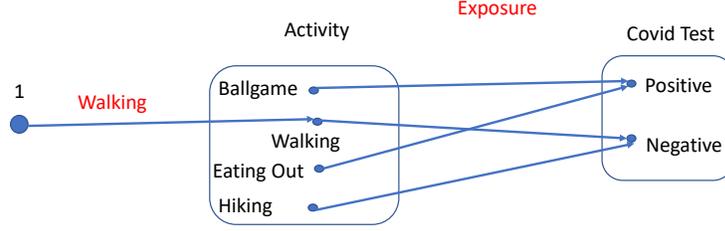


Figure 4: A simple category \mathcal{C} for studying causal effects of exposure to getting Covid-19 infections. The morphism **Exposure** maps a finite set object **Activity** to the (Boolean) object **Covid Test**, where the morphism could represent any arbitrary measure-theoretic Witsenhausen (1975); Heymann et al. (2021), probabilistic Pearl (2009b), or topological relationship Mahadevan (2021c). In category theory, an “element” of a set (such as **Walking**) is formally denoted by a morphism from the unit object **1** (which has one object and one identity morphism) to the set that is defined by the label it maps to. Causal interventions are viewed as morphisms into an object, such as determining the causal effect of **Activity** on **Covid Test** by intervening on the **Activity** variable by setting it to **Walking**. In this paper, we model interventions as elementary face and degeneracy mappings on simplicial objects.

2. \mathcal{X} is a set of *construction* objects that allow constructing complex causal models from elementary parts. Examples include monoidal tensor product of two categories $\mathcal{M} \otimes \mathcal{M}$, Galois extensions and profunctors between two partially ordered sets to represent resource models, co-limits and decorated cospans to represent electric circuits, and so on. In this paper, we use the construction tools from the theory of simplicial objects.
3. \mathcal{I} is an intervention category of intervention objects, which map from some category \mathcal{E} of experimental designs into the causal category \mathcal{C} to implement an experimental design. Interventions can be simple, such as setting the value of a variable to a specific value (see Figure 4), in which case \mathcal{I} is a comma category, or they can be more complex, such as setting the value of a variable to some arbitrary marginal distribution, such as the edge-intervention model proposed by Janzing et al. (2013). There is a large literature on treatment planning in causal inference, and many of the state of the art applications of causal inference, such as bipartite experiments Zigler and Papadogeorgou (2018) use sophisticated experimental designs. Any of these experimental designs can be accommodated in our functorial intervention framework. In this paper, we use the elementary face and degeneracy operators defined earlier on simplicial objects.
4. \mathcal{O} is a functor category of observation objects, which provide a (partial, perhaps noisy) view of causal objects in \mathcal{C} . For example, an object in the covariant functor category of presheafs $\text{Hom}_{\mathcal{C}}(X, -)$ of morphisms out of X can be viewed as a measurement of the “state” of X . More general examples include *bisimulation morphisms* used in the literature in reinforcement learning and software systems Mahadevan (2021a). Adam and Dahleh (2019) propose modeling observation as an order-preserving function Φ from a system \mathcal{S} modeled by a lattice to an observation structure \mathcal{O} , which is also a lattice. They define the composition of subsystems into a larger system as a *join* $S_1 \vee S_2$, where a system exhibits a generative effect if $\Phi(S_1 \vee S_2) \neq \Phi(S_1) \vee \Phi(S_2)$.
5. \mathcal{E} is a functor category of evaluation objects, which map a given causal model into an evaluation category for the purposes of evaluating the effects of an intervention. For an example, for a network economy Nagurney (1999), the evaluation functor E is defined as a mapping from the causal network economy model to the category of real-valued vector spaces defined by the vector field F of a variational inequality (VI) $\langle F(x^*), (x - x^*) \rangle \geq 0, \forall x \in \mathcal{C}$. Here, the vector field $F : \mathcal{C} \rightarrow \mathbb{R}^n$ is a functor from the causal category \mathcal{C} of network economy models (see Mahadevan (2021a)) to the category of n -dimensional Euclidean space. Solving a VI means finding equilibrium flows x^* on the network that represent stable patterns of trade. For example, due to the recent war in Ukraine, global supply chains from grain to natural gas and oil have been disrupted, and this intervention requires the global economic system to find a new equilibrium. The problem of studying causal interventions in network economies was recently studied by us in a previous paper Mahadevan (2021b), and we refer the reader to that paper for more details. In a causal DAG model, the evaluation functor maps the causal category into the category of probability spaces (i.e., the rules of Pearl’s do-calculus Pearl (2009b) give conditions under which $P(Y|do(X)) = P(Y|X)$).

5.1 Causal Inference over Simplicial Objects

We now turn to formulating causal discovery in terms of adjoint functors between a category, defining a universal causal model (Mahadevan, 2022a), and a simplicial object. We define causal interventions as images of the elementary injective maps $d_i : [n] \rightarrow [n + 1]$, which skips element i in its image, and surjective maps, $s_j : [n] \rightarrow [n - 1]$, which repeats element i , in the category of ordinal numbers. Defining causal inference over simplicial objects requires choosing a compositional structure over higher-order simplicial objects $n \geq 2$, which involves requires a significant amount of higher-order category theory, including weak Kan complexes Boardman and Vogt (1973), quasicategories (Joyal, 2002) and ∞ -categories (Lurie, 2022), which were reviewed above. We pose the problem of causal discovery in simplicial objects in terms of adjoints between categories and simplicial objects. The nerve of a category is a fully faithful right adjoint from a category \mathcal{C} to an associated simplicial object $\mathcal{N}_{\mathcal{C}}$. Its left adjoint is a destructive functor that only preserves information up to the 2-simplices. We relate these ideas to existing literature on causal discovery from conditional independence oracles on datasets.

Many previous studies of causal discovery from interventions, including the *conservative* family of intervention targets (Hauser and Bühlmann, 2012), path queries (Bello and Honorio, 2018), and *separating systems* of finite sets or graphs. (Eberhardt, 2008; Hauser and Bühlmann, 2012; Kocaoglu et al., 2017; Katona, 1966; Mao-cheng, 1984) can all be viewed as imposing a topology on the sets of variables being intervened.

5.2 Causal Horns and Conditional Independence

We now introduce the concept of causal horns, subsets of simplicial objects, which like the case of PSRs above, represent fragments of the complete simplicial object that is constructed during the process of structure discovery. In the literature on causal discovery, it is common to use conditional independence as a guide to causal discovery. Conditional independence is a *symmetric* property, as we will see below from examining a well-known axiom system for conditional independence in statistics.

Definition 28. A **Causal Horn** $\Lambda_i^n : \Delta^{op} \rightarrow \mathcal{C}$ is defined as

$$(\Lambda_i^n)([m]) = \{\alpha \in \mathbf{Hom}_{\Delta}([m], [n]) : [n] \not\subseteq \alpha([m]) \cup \{i\}\}$$

Intuitively, a causal horn Λ_i^n is a fragment of a causal model (see Figure 6) that can be viewed as the simplicial subset that results from removing the interior of the n -simplex Δ^n together with the face opposite its i th vertex. Note that \mathcal{C}_n defines the n -simplex associated with the horn of a causal model defined by the category \mathcal{C} (recall that a simplicial object is a contravariant functor from the category of ordinal numbers Δ to the category \mathcal{C} defining causal models.).

Definition 29. A **separoid** Dawid (2001) defines a category over preordered set (\mathcal{S}, \leq) , namely \leq is reflexive and transitive, equipped with a *ternary* relation $\perp\!\!\!\perp$ on triples (x, y, z) , where $x, y, z \in \mathcal{S}$ satisfy the following properties:

- **S1:** (\mathcal{S}, \leq) is a join semi-lattice.
- **P1:** $x \perp\!\!\!\perp y \mid x$
- **P2:** $x \perp\!\!\!\perp y \mid z \Rightarrow y \perp\!\!\!\perp x \mid z$
- **P3:** $x \perp\!\!\!\perp y \mid z$ and $w \leq y \Rightarrow x \perp\!\!\!\perp w \mid z$
- **P4:** $x \perp\!\!\!\perp y \mid z$ and $w \leq y \Rightarrow x \perp\!\!\!\perp y \mid (z \vee w)$
- **P5:** $x \perp\!\!\!\perp y \mid z$ and $x \perp\!\!\!\perp w \mid (y \vee z) \Rightarrow x \perp\!\!\!\perp (y \vee w) \mid z$

A **strong separoid** also defines a category. A strong separoid is defined over a lattice \mathcal{S} has in addition to a join \vee , a meet \wedge operation, and satisfies an additional axiom:

- **P6:** If $z \leq y$ and $w \leq y$, then $x \perp\!\!\!\perp y \mid z$ and $x \perp\!\!\!\perp y \mid w \Rightarrow x \perp\!\!\!\perp y \mid z \wedge w$

It is well known that causal DAG structures can only be recovered up to an equivalence class under observation. For example, the three DAG models $A \rightarrow B \rightarrow C$, $A \leftarrow B \leftarrow C$, and the diverger $A \leftarrow B \rightarrow C$ cannot be discriminated from observations alone, because they define the same conditional independence property $A \perp\!\!\!\perp C \mid B$. Parameterizing these models as probability distributions implies that Bayes rule allows inferring $P(B|A)$ from $P(A|B)$, thus these models are equivalent from purely observational data. To discover the exact structure requires making interventions (e.g., intervening on a variable is usually done by clamping its value $X = x$, and thereby deleting any arrows entering the variable from other objects). The same problem applies as well to non-graphical representations

of conditional independence, such as integer-valued multisets (Studený, 2010), defined as an integer-valued multiset function $u : \mathbb{Z}^{\mathcal{P}(\mathbb{Z})} \rightarrow \mathbb{Z}$ from the power set of integers, $\mathcal{P}(\mathbb{Z})$ to integers \mathbb{Z} . An imset is defined over partially ordered set (poset), defined as a distributive lattice of disjoint (or non-disjoint) subsets of variables. The bottom element is denoted \emptyset , and top element represents the complete set of variables N . A full discussion of the probabilistic representations induced by imsets is given (Studený, 2010). We will only focus on the aspects of imsets that relate to its conditional independence structure, and its topological structure as defined by the poset. A *combinatorial* imset is defined as:

$$u = \sum_{A \subset N} c_A \delta_A$$

where c_A is an integer, δ_A is the characteristic function for subset A , and A potentially ranges over all subsets of N . An *elementary* imset is defined over $(a, b \perp\!\!\!\perp A)$, where a, b are singletons, and $A \subset N \setminus \{a, b\}$. A *structural* imset is defined as one where the coefficients can be rational numbers. For a general DAG model $G = (V, E)$, an imset in standard form (Studený, 2010) is defined as

$$u_G = \delta_V - \delta_{\emptyset} + \sum_{i \in V} (\delta_{\mathbf{Pa}_i} - \delta_{i \cup \mathbf{Pa}_i})$$

Figure 5 shows an example imset for DAG models over three variables, defined by an integer valued function over the lattice of subsets. Each of the three DAG models shown defines exactly the same imset function. Studený (2010) gives a detailed analysis of imsets as a non-graphical representation of conditional independence.

Figure 5: An illustration of an integer-valued multiset (imset) consisting of a lattice of subsets over three elements for representing conditional independences in DAG models. All three DAG models are represented by the same imset. We can view an imset as a functor from a simplicial object (shown as the dark blue filled triangle) to a left k module. Note that a causal intervention on a DAG generates a horn of the simplicial object shown in dark blue.

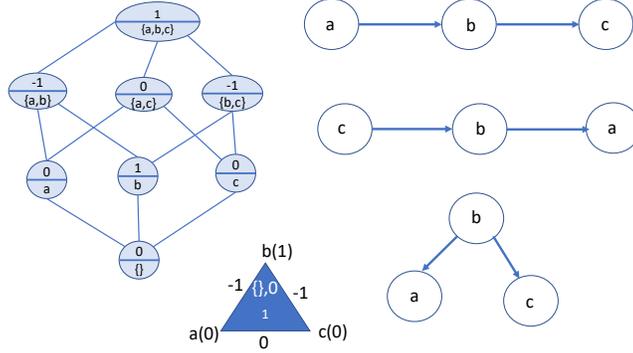


Figure 6 illustrates the space of possible causal DAG models over 3 variables, with their associated imset representations (see (Studený et al., 2010b) for details). Each vertex in this figure shows a potential DAG model. To explain this representation in more detail, the undirected edges represent the essential graph, namely for each edge of the form $a - b$, there are two possible structures that it represents $a \rightarrow b$ and $a \leftarrow b$. These two structures are homotopically equivalent, that is, they define an equivalence class on the space of all models.

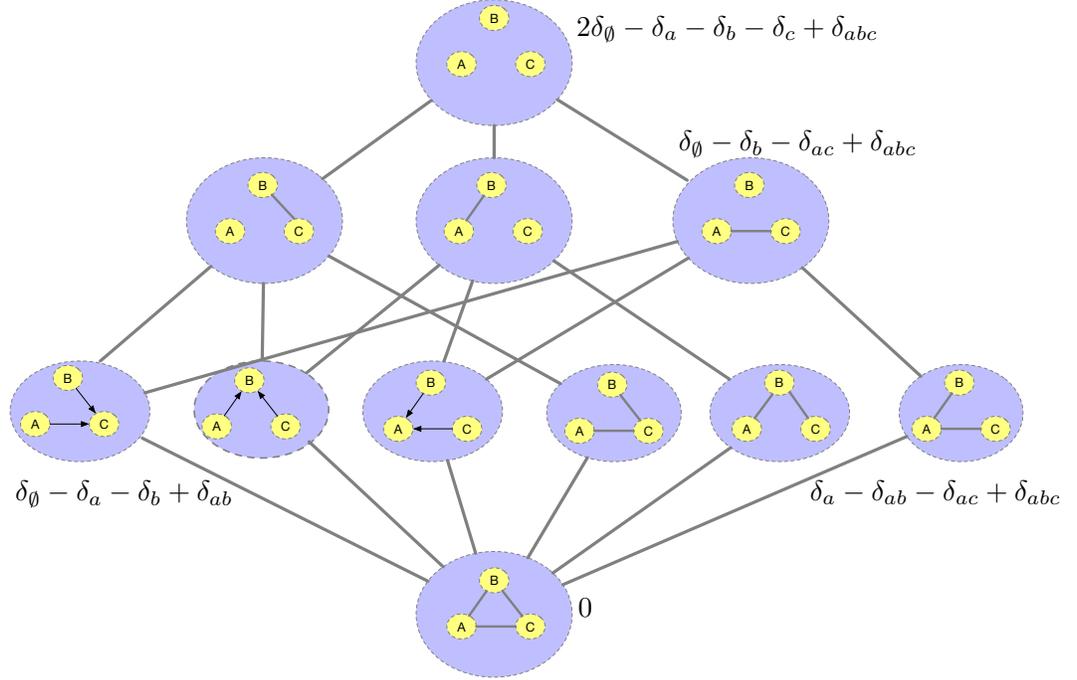
5.3 Causal Discovery as Adjoint Functors

Following the case for PSR structure discovery, we can also formulate causal discovery in terms of a pair of adjoint functors between a category \mathbf{C} and a simplicial object X .

$$X_n \begin{array}{c} \xrightarrow{G} \\ \xleftarrow{F} \end{array} \mathbf{C}.$$

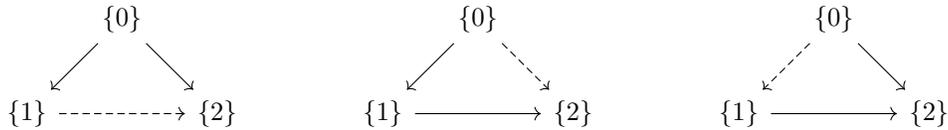
In general, the functor G from a simplicial object X to a category \mathbf{C} can be lossy. For example, we can define the objects of \mathbf{C} to be the elements of X_0 , and the morphisms of \mathbf{C} as the elements of $f \in X_1$, where $f : a \rightarrow b$, and

Figure 6: Causal discovery through the space of all models over 3 variables shown with their associated inset representations (Studený et al., 2010b). Each candidate DAG defines a *causal horn*, a simplicial subobject of the complete simplex on Δ^2 , and the process of causal structure discovery can be viewed in terms of the abstract horn filling problem defined above for higher-order categories.



$d_0 f = a$, and $d_1 f = b$, and $s_0 a, a \in X$ as defining the identity morphisms 1_a . Composition in this case can be defined as the free algebra defined over elements of X_1 , subject to the constraints given by elements of X_2 . For example, if $x \in X_2$, we can impose the requirement that $d_1 x = d_0 x \circ d_2 x$. Such a definition of the left adjoint would be quite lossy because it only preserves the structure of the simplicial object X up to the 2-simplices. The right adjoint from a category to its associated simplicial object, in contrast, constructs a full and faithful embedding of a category into a simplicial set. In particular, the nerve of a category is such a right adjoint.

Consider the three horns defined below, once again, and let us relate them to the partial DAG structures shown in Figure 6. Notice all three horns are illustrated by candidate DAG structures, for example, the rightmost horn below corresponds to the partial DAG $A \rightarrow C \leftarrow B$, under the mapping $0 \rightarrow A, 1 \rightarrow B, 2 \rightarrow C$. The edge that represents moving from this partial DAG to the complete DAG at the bottom of Figure 6 corresponds exactly to the problem of filling the outer horn shown below on the right.



Let us consider a more complex case of horn filling over a 3-simplex $f : \Lambda_2^3 \rightarrow \mathcal{C}$ for any category \mathcal{C} that represents a causal model. The “data” that is used to solve the “horn filling” problem is as follows:

- There are 4 objects, represented as 0-simplices x_0, x_1, x_2 and x_3 .
- There are 6 1-simplices, which we can define as potential causal arrows $f_{0,1}, f_{1,2}, f_{2,3}, f_{0,2}, f_{0,3}$ and $f_{1,3}$. Each of these arrows $f_{i,j}$ defines two 0-faces x_i and x_j , the basic units in a causal study.
- There are only 3 2-simplices $f_{0,2,3}, f_{0,1,2}$ and $f_{1,2,3}$. As we are dealing with the horn Λ_2^3 , the 2-simplex $f_{0,1,3}$ is missing. Each of the given 2-simplices defines a homotopy equivalent composition, for example $f_{0,2,3} : f_{2,3} \circ f_{0,2} \sim f_{0,3}$.

The problem of horn filling is to deduce the missing 2-simplex $f_{0,1,3}$ and the missing 3-simplex $f_{0,1,2,3}$ by defining a homotopy of compositions of homotopies:

$$f_{0,1,2,3} : f_{0,2,3} \circ f_{0,1,2} \equiv f_{0,1,3} \circ f_{1,2,3}$$

Definition 30. The **nerve** of a causal model \mathcal{C} is the set of sequences of composable arrows of length n , for $n \geq 1$. Let $N_n(\mathcal{C})$ denote the set of sequences of composable arrows

$$\{C_0 \xrightarrow{f_1} C_1 \xrightarrow{f_2} \dots \xrightarrow{f_n} C_n \mid C_i \text{ is an object in } \mathcal{C}, f_i \text{ is a morphism in } \mathcal{C}\}$$

We can define the causal discovery problem in terms of the right adjoint from the nerve of a causal model to the category. In general, the functor mapping a simplicial object to a category is lossy, because it only preserves properties of the simplicial object up to a certain level. In the case of causal discovery, we can exploit symmetry properties from conditional independence to design this functorial mapping, as was discussed above.

$$\mathcal{N}_\bullet(\mathcal{C}) \begin{array}{c} \xrightarrow{G} \\ \xleftarrow{F} \end{array} \mathbf{C}.$$

The definition of nerve leads to an important topological characterization of any causal category.

Definition 31. The **classifying space** of a causal model \mathcal{C} is the topological space $|N_\bullet(\mathcal{C})|$.

The classifying space of a causal model gives a way to define an algebraic invariant.

5.4 The Singular Homology of a Causal Model

We will now describe the singular homology of a causal model, using the imset representation as an example shown earlier in Figure 6.

Definition 32. For any topological space X defined by a causal model, the **singular homology groups** $H_*(X; \mathbf{Z})$ are defined as the homology groups of a chain complex

$$\dots \xrightarrow{\partial} \mathbf{Z}(\text{Sing}_2(X)) \xrightarrow{\partial} \mathbf{Z}(\text{Sing}_1(X)) \xrightarrow{\partial} \mathbf{Z}(\text{Sing}_0(X))$$

where $\mathbf{Z}(\text{Sing}_n(X))$ denotes the free Abelian group generated by the set $\text{Sing}_n(X)$ and the differential ∂ is defined on the generators by the formula

$$\partial(\sigma) = \sum_{i=0}^n (-1)^i d_i \sigma$$

Intuitively, a chain complex builds a sequence of vector spaces that can be used to construct an algebraic invariant of a topological space by choosing the left \mathbf{k} module \mathbf{Z} to be a vector space. Each differential ∂ then becomes a linear transformation whose representation is constructed by modeling its effect on the basis elements in each $\mathbf{Z}(\text{Sing}_n(X))$.

Example 10. Let us illustrate the singular homology groups defined by an integer-valued multiset Studeny (2010) used to model conditional independence. Imsets over a DAG of three variables $N = \{a, b, c\}$ shown previously as Figure 5 can be viewed as a finite discrete topological space. For this topological space X , the singular homology groups $H_*(X; \mathbf{Z})$ are defined as the homology groups of a chain complex

$$\mathbf{Z}(\text{Sing}_3(X)) \xrightarrow{\partial} \mathbf{Z}(\text{Sing}_2(X)) \xrightarrow{\partial} \mathbf{Z}(\text{Sing}_2(X)) \xrightarrow{\partial} \mathbf{Z}(\text{Sing}_1(X)) \xrightarrow{\partial} \mathbf{Z}(\text{Sing}_0(X))$$

where $\mathbf{Z}(\text{Sing}_i(X))$ denotes the free Abelian group generated by the set $\text{Sing}_i(X)$ and the differential ∂ is defined on the generators by the formula

$$\partial(\sigma) = \sum_{i=0}^4 (-1)^i d_i \sigma$$

The set $\text{Sing}_n(X)$ is the set of all morphisms $\mathbf{Hom}_{\text{Top}}(|\Delta_n|, X)$. For an imset over the three variables $N = \{a, b, c\}$, we can define the singular n -simplex σ as:

$$\sigma : |\Delta^4| \rightarrow X \text{ where } |\Delta^n| = \{t_0, t_1, t_2, t_3 \in [0, 1]^4 : t_0 + t_1 + t_2 + t_3 = 1\}$$

The n -simplex σ has a collection of faces denoted as $d_0\sigma, d_1\sigma, d_2\sigma$ and $d_3\sigma$. If we pick the k -left module \mathbf{Z} as the vector space over real numbers \mathbb{R} , then the above chain complex represents a sequence of vector spaces that can be used to construct an algebraic invariant of a topological space defined by the integer-valued multiset. Each differential ∂ then becomes a linear transformation whose representation is constructed by modeling its effect on the basis elements in each $\mathbf{Z}(\text{Sing}_n(X))$. An alternate approach to constructing a chain homology for an integer-valued multiset is to use Möbius inversion to define the chain complex in terms of the nerve of a category (see our recent work on categoroids (Mahadevan, 2022b) for details).

6 Summary

In this paper we presented a unified formalism for structure discovery in causal inference and reinforcement learning using the framework of simplicial objects, contravariant functors from the category of ordinal numbers into any category, which form the basis for higher-order category theory. We showed that structure discovery in both causal inference and RL can be defined as filling the horns of simplicial objects, and showed that a number of constructions in higher-order category theory, such as extensions and lifting problems, weak Kan complexes, nerve of a category, and chain complexes could be brought to bear on the design of new algorithms. Specifically, we show that fragments of causal models that are equivalent under conditional independence – defined as causal – as well as fragments of potential tests in a predictive state representation – defined as predictive horns – are both special cases of horns of a simplicial object of order n , resulting from the removal of the interior and the face opposite a particular vertex in an n -simplex. Causal and predictive-state horns both result from interventions on simplicial objects, defined as images of a sequence of elementary order-preserving morphisms from the category of ordinal numbers Δ into an underlying category encoding a universal causal or decision model. Latent structure discovery in both settings involve the same fundamental mathematical problem of finding extensions of horns of simplicial objects through solving lifting problems in commutative diagrams, and exploiting weak homotopies that define higher-order symmetries. The problem of filling “inner horns” is solved using quasicategories and weak Kan extensions, whereas filling “outer” horns requires ideas from ∞ -category theory. We define the common abstract problem of structure discovery in terms of adjoint functors between a universal causal or decision model category and its simplicial object representation. In general, the left adjoint functor from a simplicial object X to a category \mathcal{C} is lossy, preserving only relationships up to a certain order defined by homotopical equivalences. In contrast, the right adjoint defining the nerve of a category constructs a lossless encoding of a category as a simplicial object.

References

- J.P. May. *Simplicial Objects in Algebraic Topology*. Chicago Lectures in Mathematics. University of Chicago Press, 1992. ISBN 9780226511818. URL <https://books.google.com/books?id=QGjwV0gyQnIC>.
- M. Boardman and Rainer Vogt. *Homotopy invariant algebraic structures on topological spaces*. Springer, Berlin, 1973.
- A. Joyal. Quasi-categories and kan complexes. *Journal of Pure and Applied Algebra*, 175(1):207–222, 2002. ISSN 0022-4049. doi:[https://doi.org/10.1016/S0022-4049\(02\)00135-4](https://doi.org/10.1016/S0022-4049(02)00135-4). URL <https://www.sciencedirect.com/science/article/pii/S0022404902001354>. Special Volume celebrating the 70th birthday of Professor Max Kelly.
- Jacob Lurie. Kerodon. <https://kerodon.net>, 2022.
- Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, USA, 2nd edition, 2009a. ISBN 052189560X.
- Guido W. Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, USA, 2015. ISBN 0521885884.
- Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search, Second Edition*. Adaptive computation and machine learning. MIT Press, 2000. ISBN 978-0-262-19440-2.
- Satinder P. Singh, Michael R. James, and Matthew R. Rudary. Predictive state representations: A new theory for modeling dynamical systems. In David Maxwell Chickering and Joseph Y. Halpern, editors, *UAI '04, Proceedings of the 20th Conference in Uncertainty in Artificial Intelligence, Banff, Canada, July 7-11, 2004*, pages 512–518. AUAI Press, 2004.

- Richard S. Sutton and Andrew G. Barto. *Reinforcement learning - an introduction*. Adaptive computation and machine learning. MIT Press, 1998. ISBN 978-0-262-19398-6. URL <https://www.worldcat.org/oclc/37293240>.
- Peter Van Overschee and Bart De Moor. *Subspace identification for linear systems. Theory, implementation, applications. Incl. 1 disk*, volume xiv, pages xiv + 254. Springer, 01 1996. ISBN 0-7923-9717-7. doi:10.1007/978-1-4613-0465-4.
- N. Chomsky and M.P. Schützenberger. The algebraic theory of context-free languages*. In P. Braffort and D. Hirschberg, editors, *Computer Programming and Formal Systems*, volume 35 of *Studies in Logic and the Foundations of Mathematics*, pages 118–161. Elsevier, 1963. doi:[https://doi.org/10.1016/S0049-237X\(08\)72023-8](https://doi.org/10.1016/S0049-237X(08)72023-8). URL <https://www.sciencedirect.com/science/article/pii/S0049237X08720238>.
- Y. Give'on and M.A. Arbib. Algebra automata ii: The categorical framework for dynamic analysis. *Information and Control*, 12(4):346–370, 1968. ISSN 0019-9958. doi:[https://doi.org/10.1016/S0019-9958\(68\)90381-1](https://doi.org/10.1016/S0019-9958(68)90381-1). URL <https://www.sciencedirect.com/science/article/pii/S0019995868903811>.
- Judea Pearl. *Probabilistic reasoning in intelligent systems - networks of plausible inference*. Morgan Kaufmann series in representation and reasoning. Morgan Kaufmann, 1989.
- Steffen L. Lauritzen and Thomas S. Richardson. Chain graph models and their causal interpretations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(3):321–348, 2002. doi:<https://doi.org/10.1111/1467-9868.00340>. URL <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00340>.
- Patrick Forré and Joris M. Mooij. Markov properties for graphical models with cycles and latent variables, 2017.
- Robin J. Evans. Margins of discrete Bayesian networks. *The Annals of Statistics*, 46(6A):2623 – 2656, 2018. doi:10.1214/17-AOS1631. URL <https://doi.org/10.1214/17-AOS1631>.
- A. Philip Dawid. Separoids: A mathematical framework for conditional independence and irrelevance. *Ann. Math. Artif. Intell.*, 32(1-4):335–372, 2001. doi:10.1023/A:1016734104787. URL <https://doi.org/10.1023/A:1016734104787>.
- Milan Studený, Jiří Vomlel, and Raymond Hemmecke. A geometric view on learning bayesian network structures. *International Journal of Approximate Reasoning*, 51(5):573–586, 2010a. ISSN 0888-613X. doi:<https://doi.org/10.1016/j.ijar.2010.01.014>. URL <https://www.sciencedirect.com/science/article/pii/S0888613X10000216>. PGM-2008.
- Sridhar Mahadevan. On the universality of diagrams for causal inference and the causal reproducing property, 2022a. URL <https://arxiv.org/abs/2207.02917>.
- Sridhar Mahadevan. Universal decision models. *CoRR*, abs/2110.15431, 2021a. URL <https://arxiv.org/abs/2110.15431>.
- J.P. May. *A Concise Course in Algebraic Topology*. Chicago Lectures in Mathematics. University of Chicago Press, 1999. ISBN 9780226511832. URL <https://books.google.com/books?id=g8SG03R1bpgC>.
- Saunders MacLane. *Categories for the Working Mathematician*. Springer-Verlag, New York, 1971. Graduate Texts in Mathematics, Vol. 5.
- B. Richter. *From Categories to Homotopy Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2020. ISBN 9781108479622. URL <https://books.google.com/books?id=pnzUDwAAQBAJ>.
- Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Statistics. Wiley, 1994. ISBN 978-0-47161977-2. doi:10.1002/9780470316887. URL <https://doi.org/10.1002/9780470316887>.
- Balaraman Ravindran and Andrew G. Barto. SMDP homomorphisms: An algebraic approach to abstraction in semi-markov decision processes. In Georg Gottlob and Toby Walsh, editors, *IJCAI-03, Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, Acapulco, Mexico, August 9-15, 2003*, pages 1011–1018. Morgan Kaufmann, 2003. URL <http://ijcai.org/Proceedings/03/Papers/145.pdf>.
- Brendan Fong and David I Spivak. Seven sketches in compositionality: An invitation to applied category theory, 2018. URL <http://arxiv.org/abs/1803.05316>. cite arxiv:1803.05316Comment: 341+xi pages.
- Tai-Danae Bradley, John Terilla, and Yiannis Vlassopoulos. An enriched category theory of language: From syntax to semantics. *La Matematica*, mar 2022. doi:10.1007/s44007-022-00021-2. URL <https://doi.org/10.1007/2Fs44007-022-00021-2>.
- Michael R. Thon and Herbert Jaeger. Links between multiplicity automata, observable operator models and predictive state representations: a unified learning framework. *J. Mach. Learn. Res.*, 16:103–147, 2015. URL <http://dl.acm.org/citation.cfm?id=2789276>.

- Vishal Soni and Satinder P. Singh. Abstraction in predictive state representations. In *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence, July 22-26, 2007, Vancouver, British Columbia, Canada*, pages 639–644. AAAI Press, 2007. URL <http://www.aaai.org/Library/AAAI/2007/aaai07-101.php>.
- Dominik Janzing, David Balduzzi, Moritz Grosse-Wentrup, and Bernhard Schölkopf. Quantifying causal influences. *The Annals of Statistics*, 41(5):2324 – 2358, 2013. doi:10.1214/13-AOS1145. URL <https://doi.org/10.1214/13-AOS1145>.
- Corwin M. Zigler and Georgia Papadogeorgou. Bipartite causal inference with interference, 2018.
- Elie M. Adam and Munther A. Dahleh. Generativity and interactional effects: an overview, 2019. URL <https://arxiv.org/abs/1911.10406>.
- A. Nagurney. *Network Economics: A Variational Inequality Approach*. Kluwer Academic Press, 1999.
- Sridhar Mahadevan. Causal inference in network economics. *CoRR*, abs/2109.11344, 2021b. URL <https://arxiv.org/abs/2109.11344>.
- Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, USA, 2nd edition, 2009b. ISBN 052189560X.
- H. S. Witsenhausen. The intrinsic model for discrete stochastic control: Some open problems. In A. Bensoussan and J. L. Lions, editors, *Control Theory, Numerical Methods and Computer Systems Modelling*, pages 322–335, Berlin, Heidelberg, 1975. Springer Berlin Heidelberg. ISBN 978-3-642-46317-4.
- Benjamin Heymann, Michel de Lara, and Jean-Philippe Chancelier. Causal inference theory with information dependency models, 2021. URL <https://arxiv.org/abs/2108.03099>.
- Sridhar Mahadevan. Causal homotopy, 2021c. URL <https://arxiv.org/abs/2112.01847>.
- Alain Hauser and Peter Bühlmann. Characterization and greedy learning of interventional markov equivalence classes of directed acyclic graphs. *J. Mach. Learn. Res.*, 13:2409–2464, 2012. URL <http://dl.acm.org/citation.cfm?id=2503320>.
- Kevin Bello and Jean Honorio. Computationally and statistically efficient learning of causal bayes nets using path queries. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pages 10954–10964, 2018. URL <https://proceedings.neurips.cc/paper/2018/hash/a0b45d1bb84fe1bedbb8449764c4d5d5-Abstract.html>.
- Frederick Eberhardt. Almost optimal intervention sets for causal discovery. In David A. McAllester and Petri Myllymäki, editors, *UAI 2008, Proceedings of the 24th Conference in Uncertainty in Artificial Intelligence, Helsinki, Finland, July 9-12, 2008*, pages 161–168. AUAI Press, 2008. URL https://dslpitt.org/uai/displayArticleDetails.jsp?mmnu=1&smnu=2&article_id=1948&proceeding_id=24.
- Murat Kocaoglu, Karthikeyan Shanmugam, and Elias Bareinboim. Experimental design for learning causal graphs with latent variables. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 7018–7028, 2017. URL <https://proceedings.neurips.cc/paper/2017/hash/291d43c696d8c3704cdbe0a72ade5f6c-Abstract.html>.
- Gyula Katona. On separating systems of a finite set. *Journal of Combinatorial Theory*, 2(1):174–194, 1966.
- CAI Mao-cheng. On separating systems of graphs. *Discrete Mathematics*, 49(1):15–20, 1984. ISSN 0012-365X. doi:[https://doi.org/10.1016/0012-365X\(84\)90146-8](https://doi.org/10.1016/0012-365X(84)90146-8). URL <https://www.sciencedirect.com/science/article/pii/0012365X84901468>.
- M. Studeny. *Probabilistic Conditional Independence Structures*. Information Science and Statistics. Springer London, 2010. ISBN 9781849969482. URL <https://books.google.com.gi/books?id=bGFRcgAACAAJ>.
- Milan Studený, Jiří Vomlel, and Raymond Hemmecke. A geometric view on learning bayesian network structures. *International Journal of Approximate Reasoning*, 51:573–586, 06 2010b. doi:10.1016/j.ijar.2010.01.014.
- Sridhar Mahadevan. Categoroids: Universal conditional independence, 2022b. URL <https://arxiv.org/abs/2208.11077>.