Universal Representations for AI

Sridhar Mahadevan

Adobe Research, U.Mass, Amherst
www.cics.umass.edu/∼mahadeva

2022 年 9 月 29 日

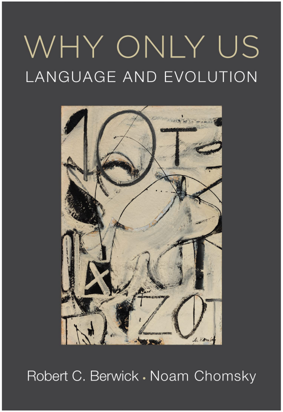
Adobe Research

UMassAmherst
College of Information
& Computer Sciences

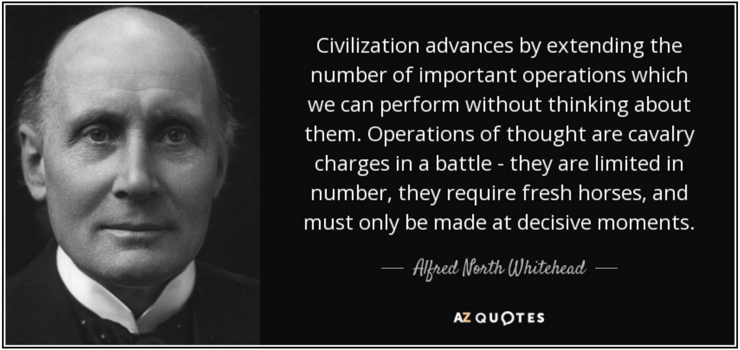
Recent Papers on Universal AI

- Universal Decision Models
- On The Universality of Diagrams for Causal Inference and the Causal Reproducing Property
- Categoroids: Universal Conditional Independence
- **Unifying Causal Inference and Reinforcement Learning using Higher-Order Category Theory**
- On Arxiv or my UMass web page: `www.cics.umass.edu/~mahadeva`

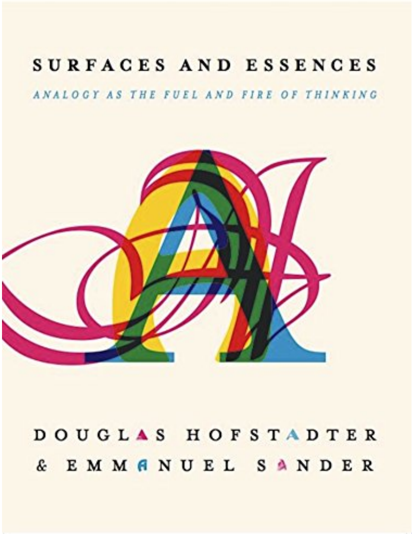
Language and Thought



The Importance of Abstraction



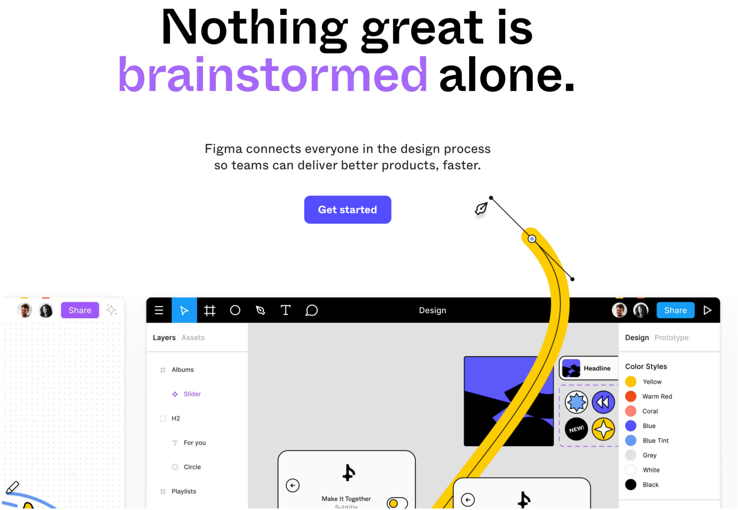
Analogy and Metaphor in Language



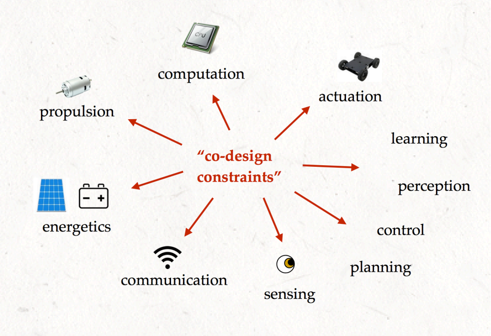
Emerging Directions in Computing

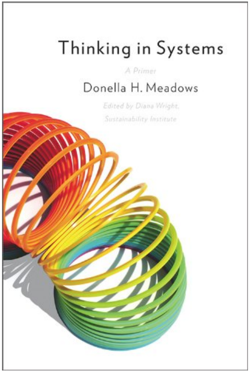
- Technological advances have greatly increased our ability our ability to perform tasks without "thinking" about them
- Computing in the cloud has decentralized collaborative activities and made remote work a reality
- Adobe just announced it will acquire Figma for $20 billion! (https://www.figma.com)
- Blockchain, the "Internet of Money", will redefine computing and economics in the 21st century

Collaborative Design using Figma



The Science of Collaborative Design





https://co-design.science/papers/

Category Theory: The Power of Abstraction

- Objects are defined by their interactions
- Functors define analogies across categories
- Natural transformations between functors give rise to universal representations
- A huge library of design tools!
  - Pullbacks, pushforwards, (co)equalizers
  - Limits and colimits
  - Kan extensions
  - Braided monoidal categories
  - Cospans, operads, props, ...

Category Theory in AI/CS/ML

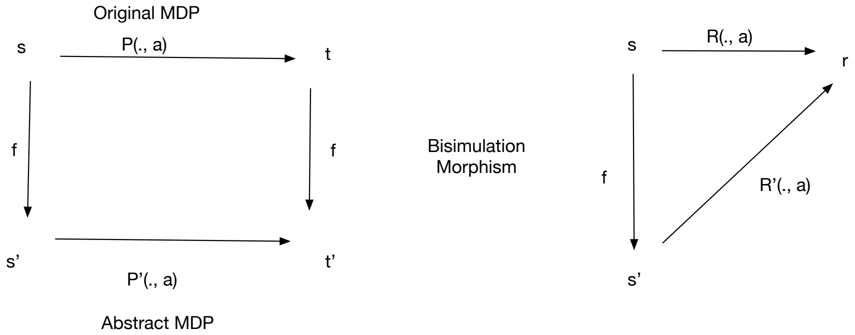
- Long been a foundation for programming languages (e.g., Haskell is widely used to implement blockchains such as Cardano)
- Uniform Manifold Approximation and Projection (https://umap-learn.readthedocs.io/en/latest/): widely used for visualizing high-dimensional data
- Clustering as a Functor (Gunnar Carlsson): overcomes Kleinberg's impossibility theorem.
- Seven Sketches in Compositionality (MIT Text on applied category theory): https://arxiv.org/abs/1803.05316
- **Offers unparalleled power to abstract complex objects and their interactions**

Reinforcement Learning

Markov Decision Processes: $\langle S, A, \Psi, P, R \rangle$

- 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
- $P : \Psi \times S \to [0, 1]$ is the transition probability function specifying the one-step dynamics of the model
- $R : \Psi \to \mathbb{R}$ is the expected reward function

RL as a Category: MDP Homomorphisms



MDP Homomorphism

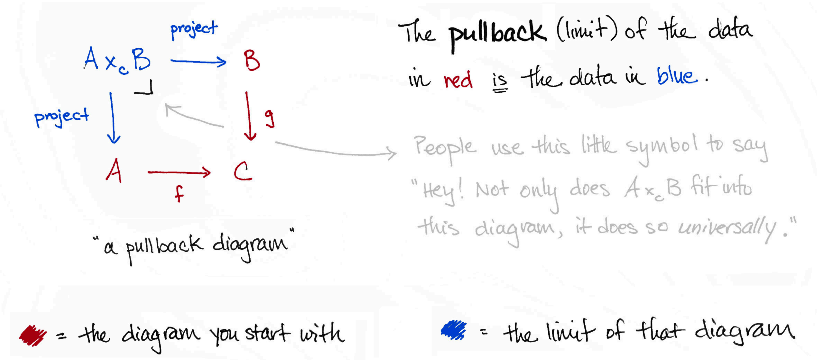
An MDP homomorphism from MDP $M = \langle S, A, \Psi, P, R \rangle$ to $M' = \langle S', A', \Psi', P', R' \rangle$, denoted $h : M \twoheadrightarrow M'$, is defined by

- A tuple of surjections $\langle f, \{g_s | s \in S\} \rangle$
- where $f : S \twoheadrightarrow S', g_s : A_s \twoheadrightarrow A'_{f(s)}$
- $h((s, a)) = \langle f(s), g_s(a) \rangle$, for $s \in S$
- Stochastic substitution property and reward respecting properties below are respected:

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

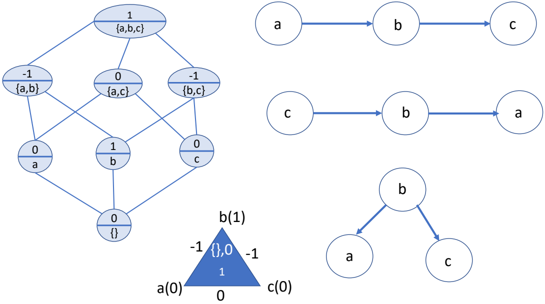
$$R'(f(s), g_s(a)) = R(s, a) \qquad (2)$$

Pullback: Reasoning with Diagrams



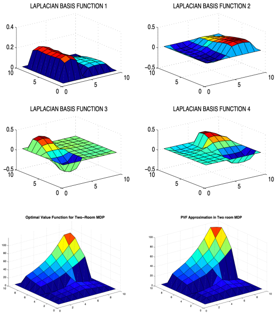
www.math3ma.com

Category of Causal Models



- Pullback, pushforward: joins and meets
- Initial and terminal objects: $\emptyset$, V

Two Functors from MDPs to Graphs



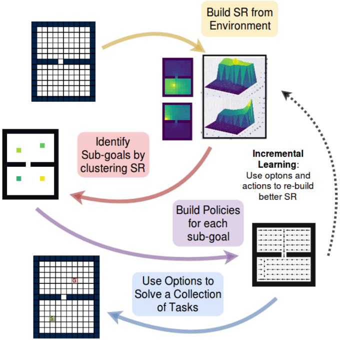
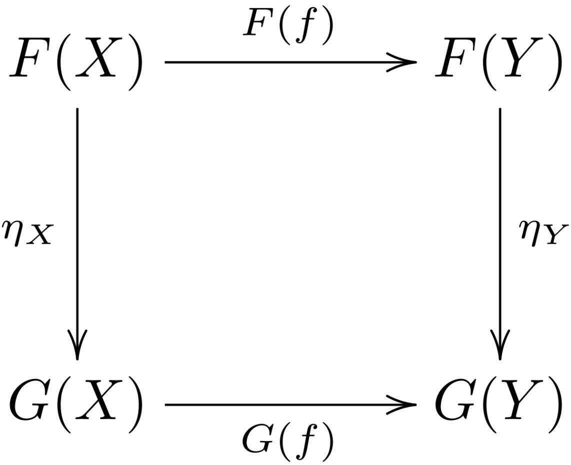


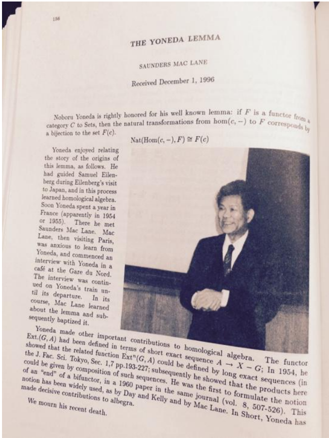
Figure 1: An overview of Successor Options Framework

Natural Transformations between two Functors



Le Lemme de la Gare du Nord: Yoneda Lemma

$$\mathbf{Hom}_{\mathcal{C}}(X, Y) \simeq \mathbf{Nat}(\mathbf{Hom}_{\mathcal{C}}(-, X), \mathbf{Hom}_{\mathcal{C}}(-, Y))$$



156

THE YONEDA LEMMA

SAUNDERS MAC LANE

Received December 1, 1996

Noboru Yoneda is rightly honored for his well known lemma: if F is a functor from a category C to Sets, then the natural transformations from $\hom(c,-)$ to F corresponds by a bijection to the set $F(c)$.

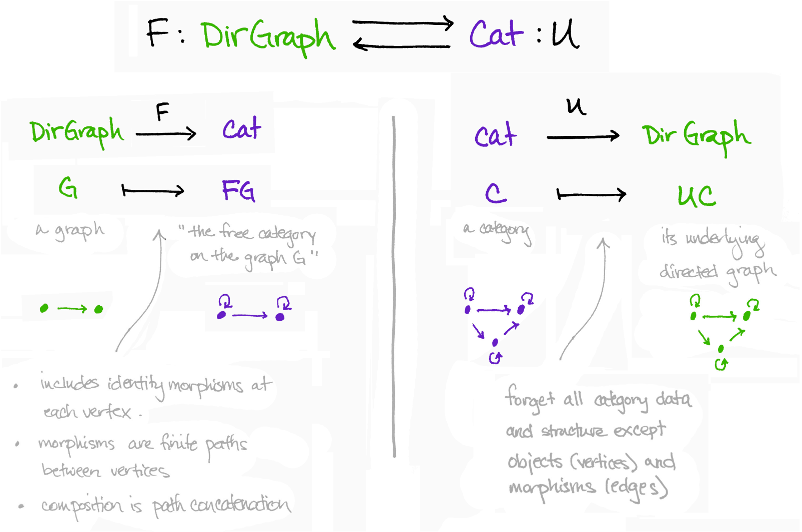
$$\mathrm{Nat}(\mathrm{Hom}(c,-), F) \cong F(c)$$

Yoneda enjoyed relating the story of the origins of this lemma, as follows. He had guided Samuel Eilenberg during Eilenberg's visit to Japan, and in this process learned homological algebra. Soon Yoneda spent a year in France (apparently in 1954 or 1955). There he met Saunders Mac Lane. Mac Lane, then visiting Paris, was anxious to learn from Yoneda, and commenced an interview with Yoneda in a café at the Gare du Nord. The interview was continued on Yoneda's train until its departure. In its course, Mac Lane learned about the lemma and subsequently baptized it.

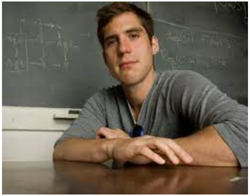
Yoneda made other important contributions to homological algebra. The functor Ext.(G, A) had been defined in terms of short exact sequence $A \to X - G$; In 1954, he showed that the related function $\mathrm{Ext}^n(G, A)$ could be defined by long exact sequences (in the J. Fac. Sci. Tokyo, Sec. 1,7 pp.193-227; subsequently he showed that the products here could be given by composition of such sequences. He was the first to formulate the notion of an "end" of a bifunctor, in a 1960 paper in the same journal (vol. 8, 507-526). This notion has been widely used, as by Day and Kelly and by Mac Lane. In Short, Yoneda has made decisive contributions to albegra.

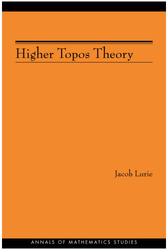
We mourn his recent death.

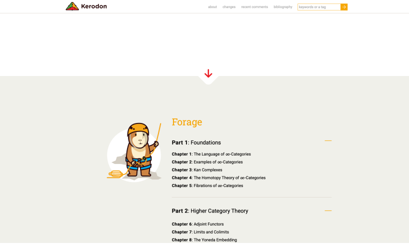
Adjunctions



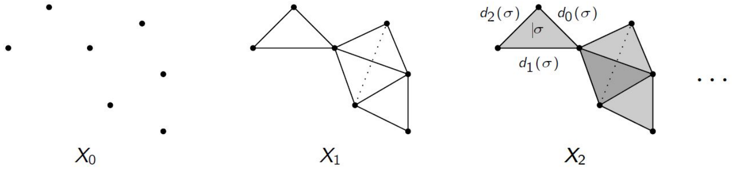
Higher-Order Category Theory

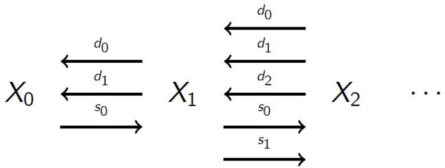






Simplicial Objects: Higher-Order Category Theory





Constructing Simplicial Objects from a Category

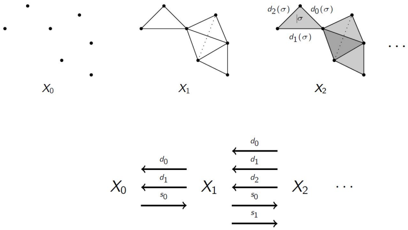
- The **nerve** of a category $\mathcal{C}$ is the set of composable morphisms of length n, for $n \geq 1$. Let $N_n(\mathcal{C})$ denote the set of sequences of composable morphisms of length n.

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

- **Theorem:** The **nerve functor** $N_\bullet : \mathbf{Cat} \to \mathbf{Set}$ is fully faithful. More specifically, there is a bijection θ defined as:

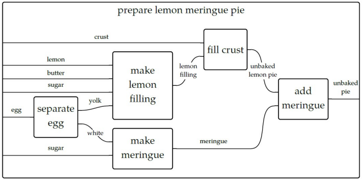
$$\theta : \mathbf{Cat}(\mathcal{C}, \mathcal{C}') \to \mathbf{Set}_\Delta(N_\bullet(\mathcal{C}), N_\bullet(\mathcal{C}'))$$

Adjunctions between Categories and Simplicial Objects

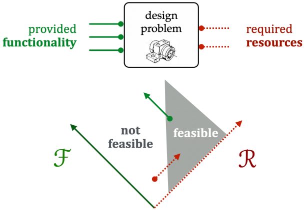


- Category to Simplicial Object: Nerve functor is a full and faithful embedding
- Simplicial object to Category: Lossy functor

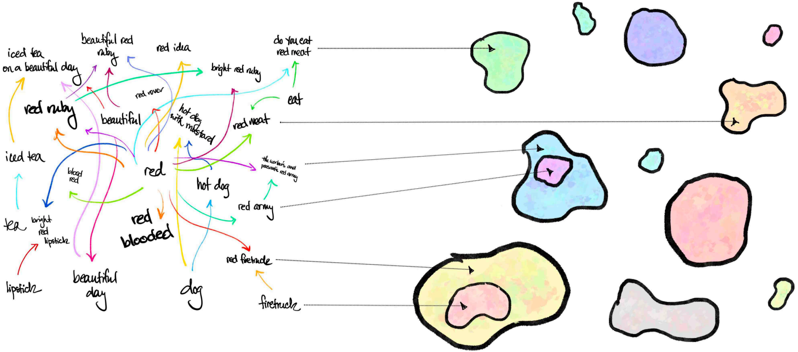
Braided Monoidal Categories (Censi, Fong, Spivak)



$\mathrm{dp}\colon \mathcal{F}^{\mathrm{op}} \times \mathcal{R} \to_{\mathbf{Pos}} \mathrm{Bool}$

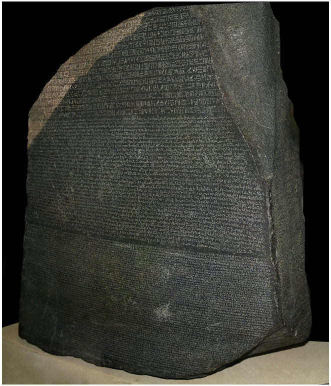


Category Theory and Language: Words and Meaning

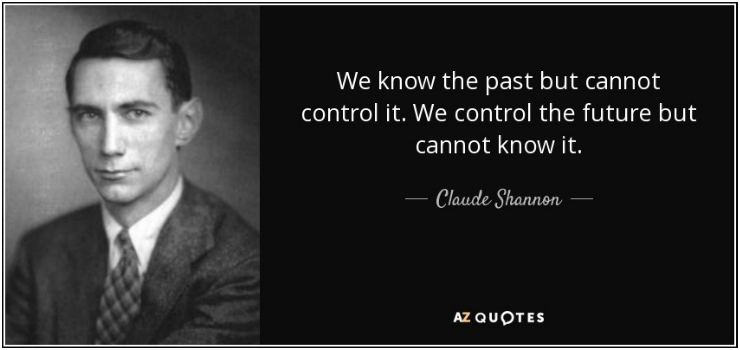


https://www.math3ma.com/categories/category-theory
https://arxiv.org/abs/2106.07890

Deciphering Ancient Language: Indus Script



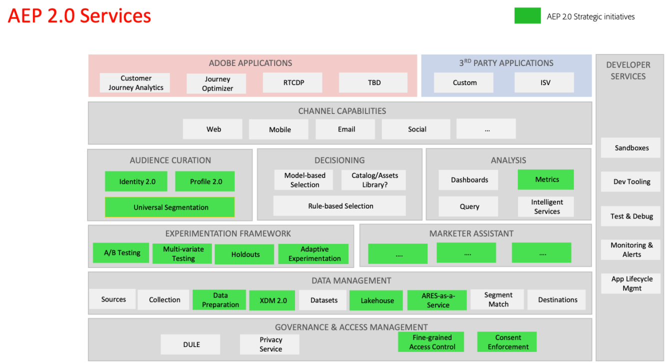
Past vs. Future



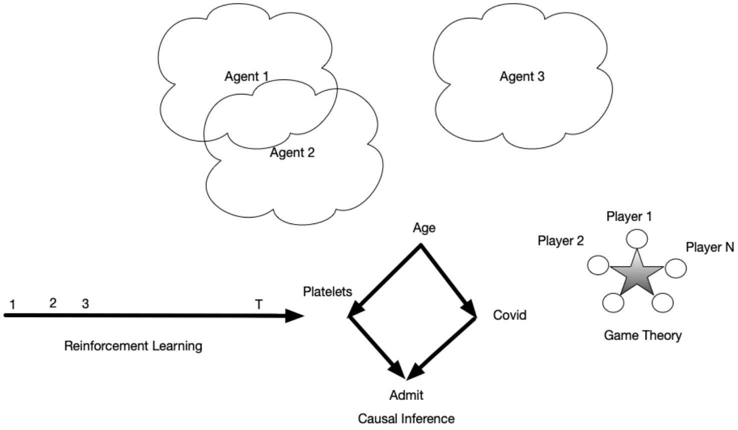
Challenge: Universal Decision Making

- We reason causally to understand our world
- We compete to gain advantage in commerce, sports and war
- We often act without complete information, but try to make better decisions over time

Adobe Experience Platform: Datamining the Past



Universal Decision Models: Controlling the Future



Category of Decision Problems

- Define categories of decision objects: causal models, RL models, Nash games, etc.
- Construct functors from one decision category to another
- Build universal representations for structure discovery

Universal Decision Models

Decision Objects: $\langle A, (\Omega, \mathcal{B}, P), (U_\alpha, \mathcal{F}_\alpha, \mathcal{I}_\alpha)_{\alpha \in A} \rangle$:

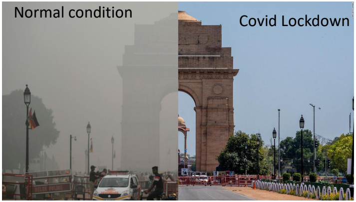
- A describes a finite universe of decision points (e.g., in RL, causal inference, or Nash games)
- $(\Omega, \mathcal{B}, P)$ is a probability space
- $(U_\alpha, \mathcal{F}_\alpha)$ is a measurable decision space from which a decision $u \in U_\alpha$ is chosen by α.
- Policy of agent α: any measurable function $\pi_\alpha : \prod_\beta U_\beta \rightarrow U_\alpha$.
- Each agent's policy is measurable from its **information field** $\mathcal{I}_\alpha$, a subfield of the overall product space $(\prod_\beta U_\beta, \prod_\beta \mathcal{F}_\beta)$, to the σ-algebra $\mathcal{F}_\alpha$.

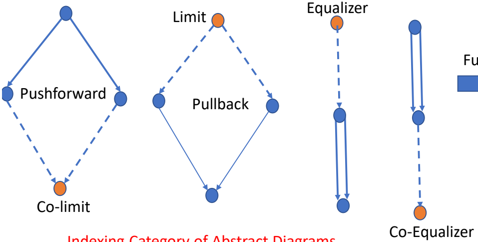
Scaling by Exploiting Conditional Independence

- Markov property: the past is conditionally independent of the future, given the current "state"
- Sufficient statistic: $(\theta \perp\!\!\!\perp X \mid T(X))$
- Graphoids, Separoids, Imsets: See my Categoroids paper
  - **Separoid** $(\mathcal{S}, \leq, \perp\!\!\!\perp)$ is a join semi-lattice.
  - **P1:** $x \perp\!\!\!\perp y \mid x$
  - **P2:** $x \perp\!\!\!\perp y \mid z \quad \Rightarrow \quad y \perp\!\!\!\perp x \mid z$
  - **P3:** $x \perp\!\!\!\perp y \mid z$ and $w \leq y \quad \Rightarrow \quad x \perp\!\!\!\perp w \mid z$
  - **P4:** $x \perp\!\!\!\perp y \mid z$ and $w \leq y \quad \Rightarrow \quad x \perp\!\!\!\perp y \mid (z \vee w)$
  - **P5:** $x \perp\!\!\!\perp y \mid z$ and $x \perp\!\!\!\perp w \mid (y \vee z) \quad \Rightarrow \quad x \perp\!\!\!\perp (y \vee w) \mid z$

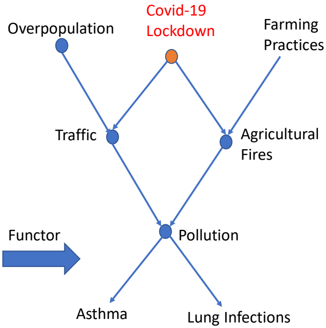
Universal Causality

Pollution in New Delhi, India





Indexing Category of Abstract Diagrams



Actual Causal Model

Yoneda Lemma Defines Universal Causality

Causal Reproducing Property: All causal influences between any two objects X and Y in a Universal Causal Model $\mathcal{M} = \langle \mathcal{C}, \mathcal{X}, \mathcal{I}, \mathcal{O}, \mathcal{E} \rangle$ is reproducible from presheaf functor objects

$$\mathbf{Hom}_{\mathcal{C}}(X, Y) \simeq \mathbf{Nat}(\mathbf{Hom}_{\mathcal{C}}(-, X), \mathbf{Hom}_{\mathcal{C}}(-, Y))$$

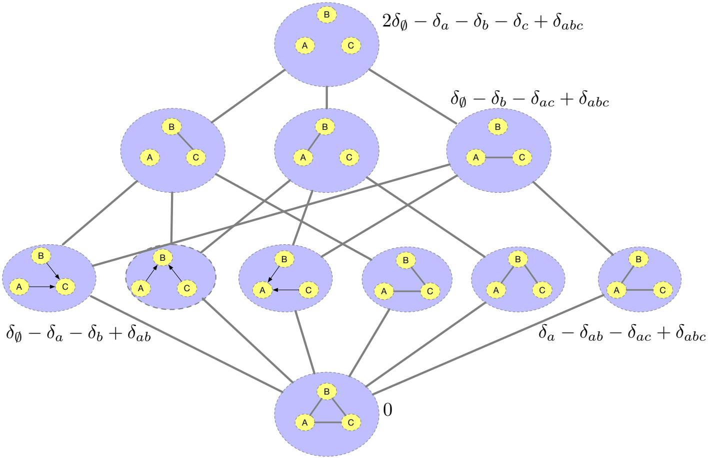
The Density Theorem in the Theory of Sheaves

Universal Causal Theorem: Given any universal causal model defined as a tuple $\mathcal{M} = \langle \mathcal{C}, \mathcal{X}, \mathcal{I}, \mathcal{O}, \mathcal{E} \rangle$, any causal inference in $\mathcal{M}$ can be represented as a co-limit of a diagram of representable objects in a unique way.

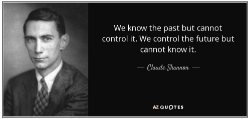
Structure Discovery by solving Lifting Problems

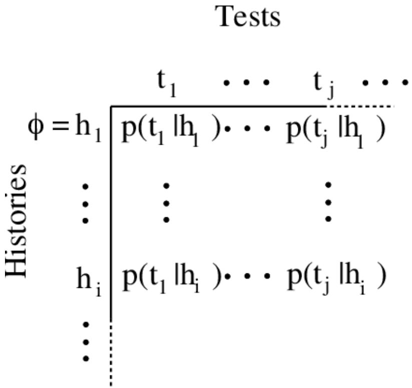
$$
\begin{array}{ccc}
A & \xrightarrow{\mu} & X \\
{\scriptstyle f}\big\downarrow & & \big\downarrow{\scriptstyle p} \\
B & \xrightarrow{\nu} & Y
\end{array}
\qquad\qquad
\begin{array}{ccc}
A & \xrightarrow{\mu} & X \\
{\scriptstyle f}\big\downarrow & \overset{h}{\nearrow} & \big\downarrow{\scriptstyle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

Lifting Problem in Causal Discovery



Predictive State Representations





Predictive State Representations

- Finite set of actions A and observations O.
- A *history*: sequence of actions and observations $h = a_1 o_1 \ldots a_k o_k$.
- A *test*: possible sequence of future actions and observations $t = a_1 o_1 \ldots a_n o_n$.
- $P(t|h)$ is a prediction test t will succeed from history h.
- State ψ: a vector of predictions of *core tests* $\{q_1, \ldots, q_k\}$.
- The prediction vector $\psi_h = \langle P(q_1|h) \ldots P(q_k|h)\rangle$ is a sufficient statistic. The predictive state of a PSR is denoted Ψ.

Category of PSRs

A **PSR homomorphism** from a PSR Ψ to another PSR Ψ' is defined as:

- A tuple of surjections $\langle f, v_\psi(a) \rangle$
- where $f : \Psi \to \Psi'$ and $v_\psi : A \to A'$ for all prediction vectors $\psi \in \Psi$
- such that

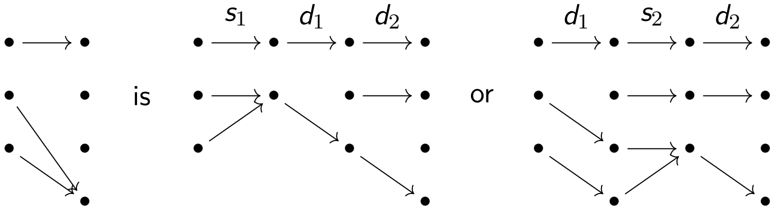
$$P(\psi' | f(\psi), v_\psi(a)) = P(f^{-1}(\psi') | \psi, a) \tag{3}$$

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

Category of Ordinal Numbers

- Category Δ: objects are non-empty ordinals $[n] = \{0, 1, \ldots, n\}$
- Arrows: non-decreasing maps $f\colon [m] \to [n]$.
- Elementary injections $d_i : [n] \to [n+1]$, which omits $i \in [n]$
- Elementary surjections $s_i : [n] \to [n-1]$, which repeats $i \in [n]$.
- Fundamental simplex $\Delta([n])$: Yoneda functor $\Delta(-, [n])$.

Elementary Surjections and Injections



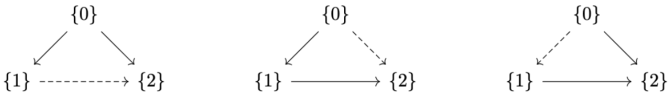
Simplicial Objects are Functors!

- Simplicial objects over sets: $X[n] = X_n : [n] \to X$
- X_0 is a set of objects: $f : [0] = \{0\}$ to X
- Arrows are in X_1: functors mapping $[1] = \{0, 1\}$ and its non-identity morphism $0 \to 1$ to X.
- 2-simplices $X[2]$: functor mapping $[2] = \{0, 1, 2\}$, with its non-identity morphisms $0 \to 1, 1 \to 2, 0 \to 2$ into a category.
- Yoneda Lemma: An n-simplex $x \in X_n$ is defined by the presheaf functor $\Delta[n] \to X$.

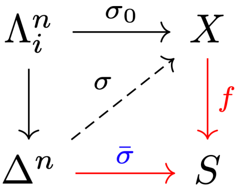
Structure Discovery by Filling Horns of Simplicial Objects

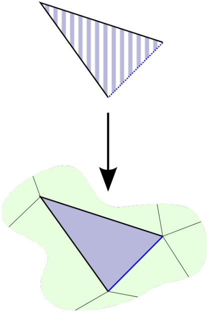
The **Horn** $\Lambda^n_i : \Delta^{op} \rightarrow \mathbf{Set}$ is defined as

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



Solving Lifting Problems with Kan Fibrations





Higher-Order Category Theory

- Weak Kan complexes
- Quasicategories
- ∞-categories

Universal PSR Theorems

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:

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

Theorem: The **nerve functor** defined by a PSR $N_\bullet : \mathbf{Cat}_{PSR} \rightarrow \mathbf{Set}$ forms a quasicategory.

Singular Homology Group Defined by a Topological Space

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

$$\ldots \xrightarrow{\partial} \mathbf{Z}(\mathrm{Sing}_2(X)) \xrightarrow{\partial} \mathbf{Z}(\mathrm{Sing}_1(X)) \xrightarrow{\partial} \mathbf{Z}(\mathrm{Sing}_0(X))$$

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

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

Classifying Space of a PSR

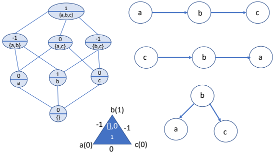
- The **classifying space** of a PSR $\mathcal{C}$ is the topological space defined by its nerve functor $|N_\bullet(\mathcal{C})|$.
- For any topological space defined by a PSR $|\mathcal{N}_\bullet(\mathcal{C})|$, the **singular homology groups** $H_*(|\mathcal{N}_\bullet(\mathcal{C})|; \mathbf{Z})$ are defined as the homology groups of a chain complex

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

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

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

Singular Homology Defined by a Causal Model



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

$$\partial(\sigma) = \sum_{i=0}^{3} (-1)^i d_i \sigma$$

$$\sigma_n : |\Delta^n| \rightarrow X \text{ where } |\Delta^n| = \{t_0, \ldots, t_n\} \in [0,1]^n : \sum_i t_i = 1\}$$

Summary: Universal AI

- How to unify different models of decision making?
- How to build universal representations?
- How to solve the universal structure discovery problem?
- (Higher-order) category theory provides answers