Universal Decision Making

A Categorial Framework using Information Fields

Sridhar Mahadevan

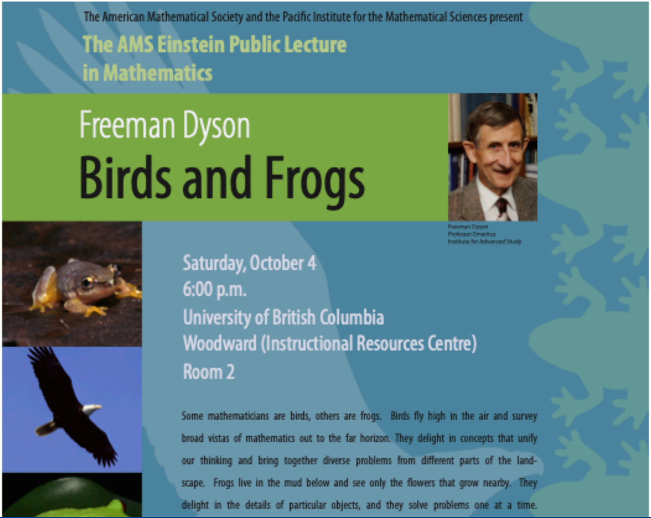
Adobe Research, U.Mass, Amherst

2026 年 4 月 22 日

Adobe Research

UMassAmherst
College of Information & Computer Sciences

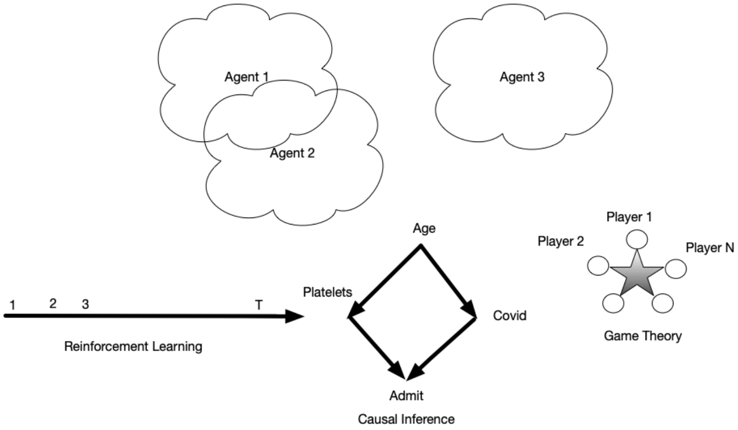
Birds and Frogs (Freeman Dyson)



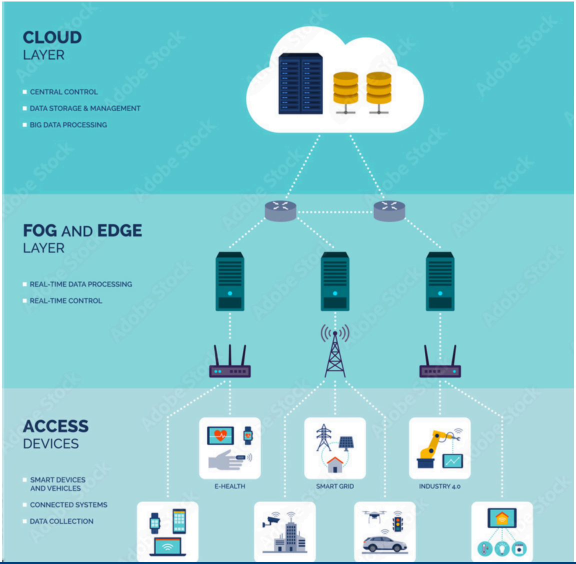
Challenges

- What are the *universal properties* of decision making?
  - Causal inference, Game theory, Reinforcement Learning, . . .
- How do we formulate the *information structures* underlying decision making?
- Agents have a priori no ordering – no notion of time!
- Organization of agents may depend on random state of nature or task at hand

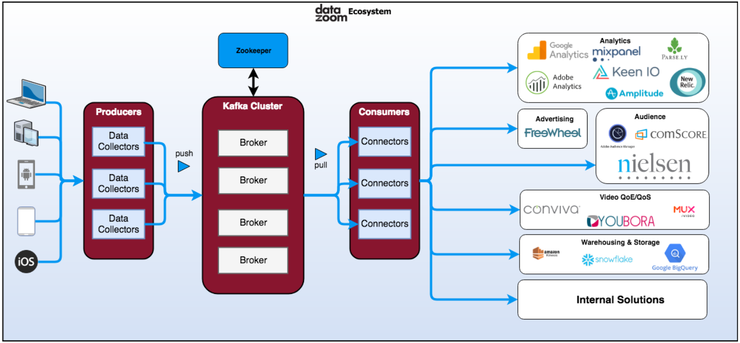
Universal Decision Making via Information Fields



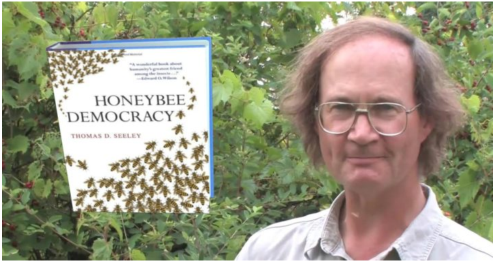
Decentralized Decision Making in Cloud Computing



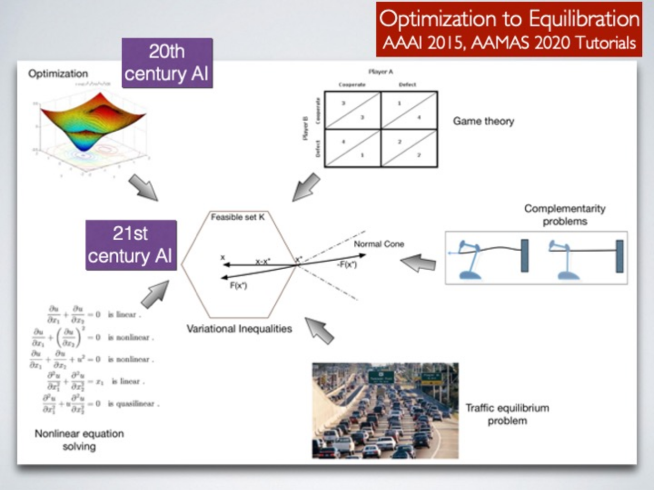
Decision Making in Apache Data Centers



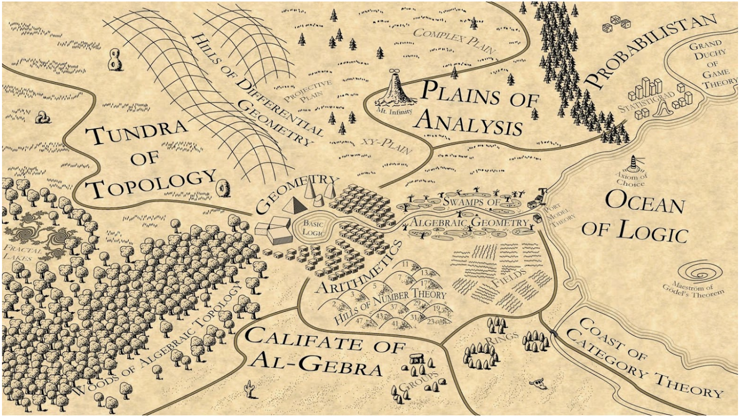
Group Decision Making in Honeybees



Decision Making in Network Economics



Mathematistan (Martin Kuppe)



Principles of Category Theory

- Unifying framework that revolutionized math over the past 50-60 years.
- Instead of describing objects (e.g., sets), characterize their interactions.
- Functors map from one category to another (e.g., $f : \mathbf{Top} \rightarrow \mathbf{Grp}$).
- Universal properties characterize an object uniquely up to isomorphism
- Natural transformations map between two functors
- Yoneda lemma: fully faithful embedding of categorial objects

Categories: Objects and Morphisms

A category $\mathcal{C}$ is

- A *collection* of *objects* $X, Y, \ldots$
- A collection of *morphisms* $f, g, \ldots$, where $f : X \to Y$ is the morphism whose domain is X and co-domain is Y.
- For each pair of morphisms f, g, such that the co-domain of f is the same as the domain of g, there is a composite morphism $g \circ f$, simply defined as the composition of g and f (where f is applied first, followed by g), defined as $gf : X \to Z$.
- Each object X has associated with it an *identity* morphism $1_X : X \to X$, whose composition with any other morphism $f : X \to Y$ is defined as $1_Y f = f = f 1_X = f$.
- *Associativity*, whereby given morphisms $f : X \to Y, g : Y \to Z, h : Z \to W$, the composite morphism $hgf : X \to W$ is associative.

Examples of Categories

- **Set**: Objects are sets, morphisms are mappings on sets.
- **Top:** Topological spaces are objects, and continuous functions as its morphisms.
- **Group:** Groups are its objects, and group homomorphisms as its morphisms.
- **Graph:** Graphs are objects, and graph morphisms (mapping vertices to vertices, preserving adjacency properties) as its morphisms.
- **Poset:** Partially ordered sets as its objects and order-preserving functions as its morphisms.
- **Meas:** Measurable spaces are its objects and measurable functions as its morphisms.

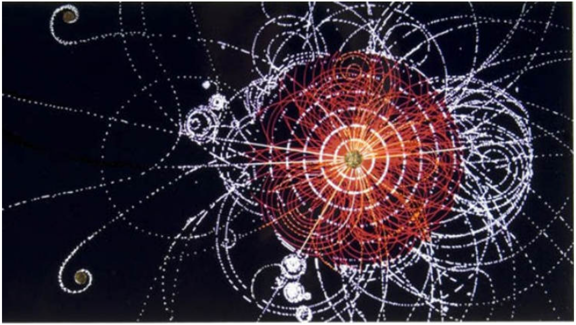
Categories vs. Sets

Set theory	Category theory
set	object
subset	subobject
truth values $\{0, 1\}$	subobject classifier Ω
power set $P(A) = 2^A$	power object $P(A) = \Omega^A$
bijection	isomorphims
injection	monic arrow
surjection	epic arrow
singleton set $\{*\}$	terminal object $\mathbf{1}$
empty set $\emptyset$	initial object $\mathbf{0}$
elements of a set X	morphism $f : \mathbf{1} \to X$
	non-global element $Y \to X$
	functors, nat. transformations, limits, colimits

What is an MDP?

- Parts-based view: states, actions, rewards, stochastic dynamics.
- Categorial view: interaction of an MDP with other objects.
- Ravi Vakil's particle accelerator analogy:" You work at a particle accelerator. You want to understand some particle. All you can do is throw other particles at it and see what happens. If you understand how your mystery particle responds to all possible test particles at all possible test energies, then you know everything there is to know about your mystery particle."

Building a "particle accelerator" for MDPs?



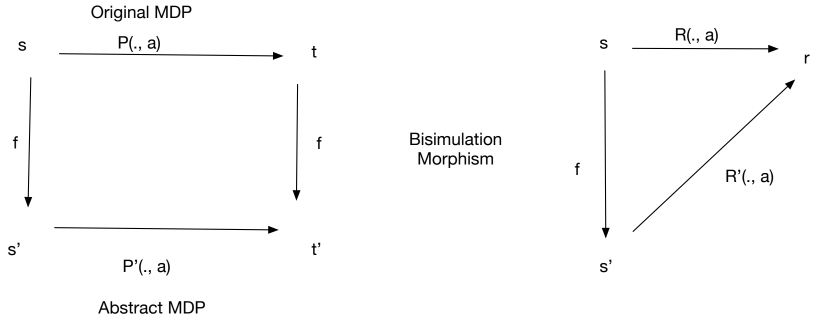
- Intuition: treat an MDP like a "black box"
- Smash things into it, or use it to smash other things!

MDPs form a Category

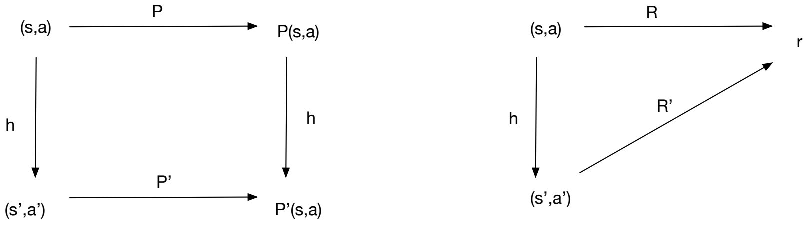
Objects are MDPs: $\langle S, A, \Psi, P, R\rangle$

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

MDP Homomorphisms



MDP Homomorphisms



MDP Homomorphisms [Ravindran and Barto]

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:

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

$$R'(f(s), g_s(a)) = R(s, a) \tag{2}$$

Predictive State Representations

PSR (and earlier models, like multiplicity automata, observer operator models etc.) form categories:

- 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 entire predictive state of a PSR can be denoted Ψ.

PSR Homomorphisms [Soni et al., AAAI]

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 $f : \Psi \rightarrow \Psi'$ and $v_\psi : A \rightarrow A'$ for all prediction vectors $\psi \in \Psi$
- such that

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

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

SIAM J. CONTROL
Vol. 9, No. 2, May 1971

ON INFORMATION STRUCTURES, FEEDBACK AND CAUSALITY*

H. S. WITSENHAUSEN†

Abstract. A finite number of decisions, indexed by $\alpha \in A$, are to be taken. Each decision amounts to selecting a point in a measurable space $(U_\alpha, \mathscr{F}_\alpha)$. Each decision is based on some information fed back from the system and characterized by a subfield $\mathscr{I}_\alpha$ of the product space $(\prod_\alpha U_\alpha, \prod_\alpha \mathscr{F}_\alpha)$. The decision function for each α can be any function γ_α measurable from $\mathscr{I}_\alpha$ to $\mathscr{F}_\alpha$.

A property of the $\{\mathscr{I}_\alpha\}_{\alpha \in A}$ is defined which assures that the setup has a causal interpretation. This property implies that for any combination of choices of the γ_α, the closed loop equations have a unique solution.

The converse implication is false, when card $A > 2$.

1. Introduction. In control-oriented works on dynamic games (in particular, stochastic control problems) one usually finds a "dynamic equation" describing the evolution of a "state" in response to decision (control) variables of the players and to random variables. One also finds "output equations" which define output variables for a player as functions of the state, decision and random variables. Then the information structure is defined by allowing each decision variable to be any desired (measurable) function of the output variables generated for that player up to that time.

Such a setup assumes that the time order in which the various decision variables are selected is fixed in advance. It assumes that each player acts as if he had responsibility only for one station. It assumes that this station has perfect memory.

For large complex systems these tacit assumptions are unlikely to hold.

Two-Player Game

A partial information game $\mathcal{G} = \langle A, (\Omega, \mathcal{B}, P), (U_\alpha, \mathcal{P}_\alpha)_{\alpha \in A} \rangle$:

- Set of players A, with probability space $(\Omega, \mathcal{B}, P)$
- Decision space: $(U_\alpha, \mathcal{P}_\alpha)_{\alpha \in A}$, where $\mathcal{P}_\alpha$ is a partition of Ω.
- Simple two-player game: $A = \alpha, \beta$.
  - State of nature: $\Omega = \{1, 2, \ldots, 9\}$, $\mathcal{B} = 2^\Omega$, $P\{i : i \in \Omega\} = \frac{1}{9}$.
  - Information partition: $\mathcal{P}_\alpha = \{\{1, 2, 3\}, \{4, 5, 6\}, \{7, 8, 9\}\}$.
  - Information partition: $\mathcal{P}_\beta = \{\{1, 2, 3, 4\}, \{5, 6, 7, 8\}, \{9\}\}$.
  - Suppose true state of nature is $\omega \in \Omega = 1$
  - Knowledge of α: $\mathcal{P}^1_\alpha = \{1, 2, 3\}$
  - Knowledge of β: $\mathcal{P}^1_\beta = \{1, 2, 3, 4\}$

Join and Meet Operations

Partition fields form a lattice:

- $\mathcal{P}_\alpha = \{\{1,2,3\},\{4,5,6\},\{7,8,9\}\}$.
- $\mathcal{P}_\beta = \{\{1,2,3,4\},\{5,6,7,8\},\{9\}\}$.
- The **join** $\mathcal{P}_\alpha \vee \mathcal{P}_\beta$: Coarsest common refinement
- $\mathcal{P}_\alpha \vee \mathcal{P}_\beta = \{\{1,2,3\},\{4\},\{5,6\},\{7,8\},\{9\}\}$.
- The **meet** $\mathcal{P}_\alpha \wedge \mathcal{P}_\beta$: Finest partition refined by $\mathcal{P}_\alpha$ and $\mathcal{P}_\beta$.
- $\mathcal{P}_\alpha \wedge \mathcal{P}_\beta = \{\{1,2,3,4,5,6,7,8,9\}$.

Updating priors

Suppose α and β observe the event $A = \{3, 4\}$

- $\mathcal{P}_\alpha^A = \{\{1, 2, 3\}, \{4, 5, 6\}\ \}$, leading to its assessment of the probability of the state of nature being $P_\alpha(\omega) = \frac{P(A \cap P_\alpha(\omega))}{P_\alpha(\omega)} = \frac{1}{3}$.
- $\mathcal{P}_\alpha^B = \{\{1, 2, 3, 4\}\}$, leading to its assessment of the probability of the state of nature being $P_\beta(\omega) = \frac{P(A \cap P_\alpha(\omega))}{P_\alpha(\omega)} = \frac{1}{2}$.

Sigma algebras

We generalize from partition fields to sigma fields (or algebras)

- Decisions are taken in a **measurable space** $(U, \mathcal{F})$ is defined as a set U along with a σ-algebra $\mathcal{F}$ of subsets of U.
- Policies: A **measurable function** $f : U \to V$ is defined to be any function defined over measurable spaces in its domain and range, namely if $(U, \mathcal{F}_U)$ is the measurable space over its domain, and $(V, \mathcal{F}_V)$ is the measurable space over the range.
- For any finite measurable space $(U, \mathcal{F})$, its σ-algebra $\mathcal{F}$ can be generated purely from a partition of $U = \{P_1, \ldots, P_k\}$, by forming the union of all possible subsets in the partition.

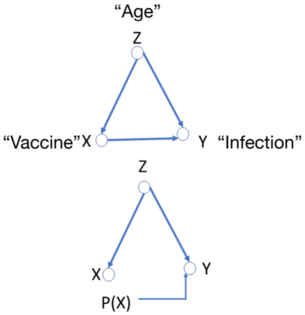
Witsenhausen's Intrinsic Model

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

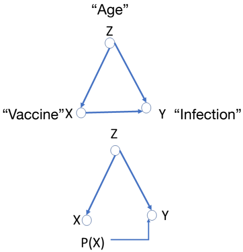
- A describes a finite universe of elements (e.g., agents, causal variables, discrete states) representing decision points
- $(\Omega, \mathcal{B}, P)$ is a probability space representing the inherent stochastic state of nature due to randomness
- $(U_\alpha, \mathcal{F}_\alpha)$ is a measurable space from which a decision $u \in U_\alpha$ is chosen by α.
- The policy of agent α can be any function $\pi_\alpha : \prod_\beta U_\beta \rightarrow U_\alpha$.
- Each agent's policy is any function $\pi_\alpha : \prod_\beta U_\beta \rightarrow U_\alpha$ 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 $\mathcal{F}_\alpha$.

Causal Inference

- Causal influence $\mathcal{C}_S^{\phi} = D_{\phi}(P \parallel P_S)$, for some set of edges S, is defined as the ϕ-divergence between pre-intervention distribution P and post-intervention distribution P_S [Mahadevan, Arxiv, 2021]
- Example DAG: $P(x, y, z) = P(z)P(x|z)P(y|x, z)$ $P_{X \rightarrow Y}(x, y, z) = P(z)P(x|z) \sum_{x'} P(y|x', z)P(x')$.



Causal Information Fields [Heymann et al., Arxiv, 2021]



- $A = \{X, Y, Z\}, U_X = U_Y = U_Z = \{0, 1\}$.
- σ-algebras: $\mathcal{F}_X = \mathcal{F}_Y = \mathcal{F}_Z = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$.
- States of nature: $\Omega = \{0, 1\}^3$, Borel topology $\mathcal{B} = 2^\Omega$.
- Information fields:
  - $\mathcal{I}_Z \subset \mathcal{B}_Z \times \{\emptyset, \Omega_X\} \times \{\emptyset, \Omega_Y\} \times \{\emptyset, U_Z\}$.
  - $\mathcal{I}_Y \subset \{\emptyset, \Omega_Z\} \times \{\emptyset, \Omega_X\} \times \mathcal{F}_X \times \mathcal{B}_Z \times \mathcal{F}_Y \times \{\emptyset, U_Y\}$.

Solvable Intrinsic Model

An intrinsic system $\langle A, (\Omega, \mathcal{B}, P), (U_\alpha, \mathcal{F}_\alpha, \mathcal{I}_\alpha)_{\alpha \in A} \rangle$ is said to be **solvable** if for every state of nature $\omega \in \Omega$, and every control law $\pi \in \Pi_A$, the set of simultaneous equations given below has one and only one solution $u \in U$.

$$u_\alpha = \pi_\alpha(h) \equiv \pi_\alpha(\omega, u) \tag{4}$$

Here, π_α can be viewed as a projection from the joint decision h taken by the entire ensemble of decision makers in the intrinsic model.

Causality in Intrinsic Models

An intrinsic model $\langle A, (\Omega, \mathcal{B}, P), (U_\alpha, \mathcal{F}_\alpha, \mathcal{I}_\alpha)$ is said to be **causal** if

- There exists $\phi : H \rightarrow S$, where S is the set of total orderings of decision makers in A,
- such that for $1 \leq k \leq n$, and any ordered set $(\alpha_1, \ldots, \alpha_k)$ of distinct elements from A, the set $E \subset H$ on which $\phi(h)$ begins with the same ordering $(\alpha_1, \ldots, \alpha_k)$ satisfies the following causality condition:

$$\forall F \in \mathcal{F}_{\alpha_k}, \quad E \cap F \in \mathcal{F}(\{\alpha_1, \ldots, \alpha_{k-1}\}) \tag{5}$$

Classes of Intrinsic Models

1. **Monic:** $A = \{\alpha\}$, $\mathcal{I}_\alpha \subset \mathcal{F}(\emptyset)$.
2. **Team:** $\mathcal{I}_\alpha \subset \mathcal{F}(\emptyset)$.
3. **Sequential:** There exists a fixed ordering $\{\alpha_1, \ldots, \alpha_n\}$ of decision makers from A such that for any $1 \leq k \leq n$, it holds that $\mathcal{I}_{\alpha_k} \subset \mathcal{F}(\{\alpha_1, \ldots, \alpha_{k-1}\}$.
4. **Classical:** An intrinsic model is called classical if it is sequential, and $\mathcal{I}_0 \in \mathcal{F}(\emptyset)$, $\mathcal{I}_{k-1} \subset \mathcal{I}_k$, for all $k = 2 \ldots, n$.
5. **Without self-information:** An intrinsic system has no self-information if for all computing elements $\alpha \in A$, it holds that $\mathcal{I}_\alpha \subset \mathcal{F}(A - \{\alpha\})$.

Sequential Intrinsic Models

1. Probability space: $(\Omega, \mathcal{B}, P)$
2. Measurable decision spaces $(U_t, \mathcal{F}_t), t = 1, \ldots, T$ at each time point.
3. Information fields $\mathcal{I}_t \subset \mathcal{B} \times \mathcal{F}_1 \ldots \mathcal{F}_T$
4. Exists permutation $p : \{1, \ldots, T\} \rightarrow \{1, \ldots, T\}$ such that for $t = 1, \ldots T$, the information field $\mathcal{I}_t \subset \mathcal{B} \times \mathcal{F}_{p(1)} \times \mathcal{F}_{p(2)}, \ldots, \mathcal{F}_{p(t-1)} \times \{\emptyset, \mathcal{F}_{p(t)}\} \times \ldots \times \{\emptyset, \mathcal{F}_{p(T)}\}$.
5. Cost function $c : (\Omega \times U_{1:T} \times, \mathcal{B} \times \mathcal{F}_{1:T}) \rightarrow (\mathbb{R}, \mathbb{B})$
6. Objective: minimize cost function $\inf_\pi E[c(\omega, U_1, \ldots, U_T)]$ exactly, or to within ϵ.

Common Knowledge in Sequential Intrinsic Models [Nayyar, 2018]

- The **common knowledge** for the t^{th} decision maker in a sequential intrinsic model is defined as

$$\mathcal{C}_t = \bigcap_{s=t}^{T} \mathcal{I}_s \tag{6}$$

- Coarsening property: $\mathcal{C}_t \subset \mathcal{I}_t$: immediate from definition.
- Nestedness property: $\mathcal{C}_t \subset \mathcal{C}_{t+1}$: immediate from definition.
- Common observations: There exist observations $Z_1, \ldots, Z_T$ with Z_t taking values in a finite measurable space $(Z_t, 2^{Z_t})$, and $Z_t = \eta_t(\omega, U_1, \ldots, U_{t-1})$ such that $\sigma(Z_{1:t}) = \mathcal{C}_t$.

Category of Intrinsic Models

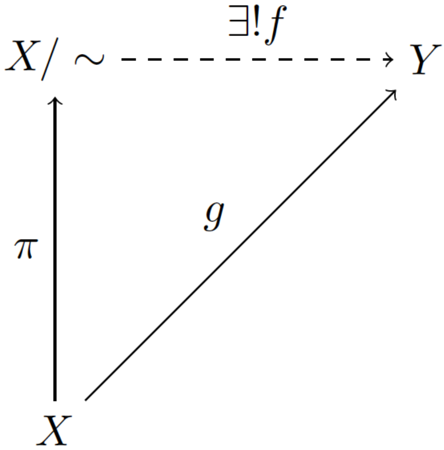
- Objects are intrinsic models $\langle A, (\Omega, \mathcal{B}, P), (U_\alpha, \mathcal{F}_\alpha, \mathcal{I}_\alpha)_{\alpha \in A}$.
- Morphisms are **bisimulation** relationship between two intrinsic models $M = \langle A, (\Omega, \mathcal{B}, (U_\alpha, \mathcal{F}_\alpha, \mathcal{I}_\alpha)_{\alpha \in A}$ and $M' = \langle A', (\Omega', \mathcal{B}', (U'_\alpha, \mathcal{F}'_\alpha, \mathcal{I}'_\alpha)_{\alpha \in A'}$, denoted as $M \twoheadrightarrow M'$, is defined as is defined by a tuple of surjections as follows:
    - A surjection $f : A \twoheadrightarrow A'$ that maps decision points in A to corresponding points in A'.
    - As f is surjective, it induces an equivalence class in A such that $x \sim y, x, y \in A$ if and only if $f(x) = f(y)$.
    - A surjection $g : H \twoheadrightarrow H'$, where $H = \Omega \times \prod_{\alpha \in A} U_\alpha$, with the product σ-algebra $\mathcal{B} \times \prod_{\alpha \in A} \mathcal{F}_\alpha$, and $H' = \Omega' \times \prod_{\alpha \in A'} U'_\alpha$, with the corresponding σ-algebra $\mathcal{B}' \times \prod_{\alpha \in A'} \mathcal{F}'_\alpha$.
    - The **quotient information field** of a collection of agents $[\alpha]_f$ is defined as the join of the information fields of each agent:

$$\mathcal{I}_{[\alpha]} = \bigvee_{\beta \in [\alpha]_f} \mathcal{I}_\alpha \tag{7}$$

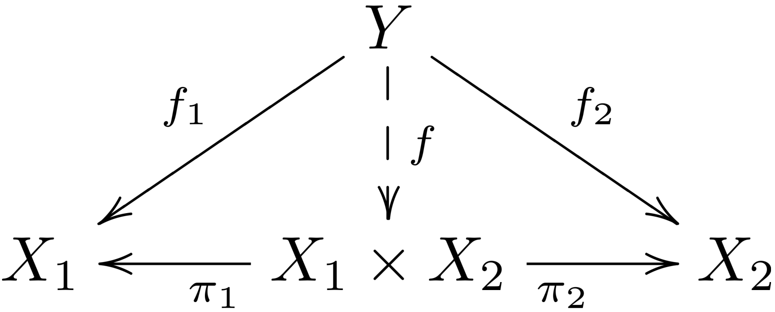
Subsystem Topology of Intrinsic Models [Witsenhausen]

- A subset of decision makers $B \subset A$ form a **subsystem** if for all $\alpha \in B$, $\mathcal{I}_\alpha \subset \mathcal{F}(B)$.
- The subsystem $\langle B, (\Omega, \mathcal{B}, P), (U_\alpha, \mathcal{F}_\alpha, \mathcal{I}_{\alpha B})_{\alpha \in B} \rangle$ has an induced information subfield $\mathcal{I}_{\alpha B}$, which is the canonical projection of $\mathcal{I}_B$ upon H_B.
- The **closure** of a decision maker $\alpha \in A$ in an intrinsic model $\langle A, (\Omega, \mathcal{B}, P), (U_\alpha, \mathcal{F}_\alpha, \mathcal{I}_\alpha)_{\alpha \in A} \rangle$ is the smallest subsystem containing α, denoted by $\overline{\{\alpha\}}$.
- The **preorder** relationship between decision makers, denoted $\alpha \leftarrow \beta$ is defined by the containment between the closure sets, namely $\alpha \leftarrow \beta$ if and only if $\overline{\{\alpha\}} \subset \overline{\{\beta\}}$.

Quotients



Products



Meets in a Poset are Products

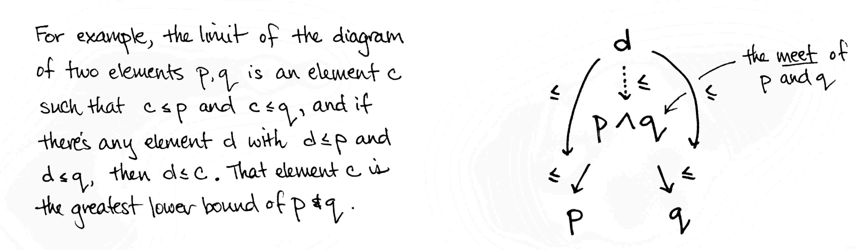
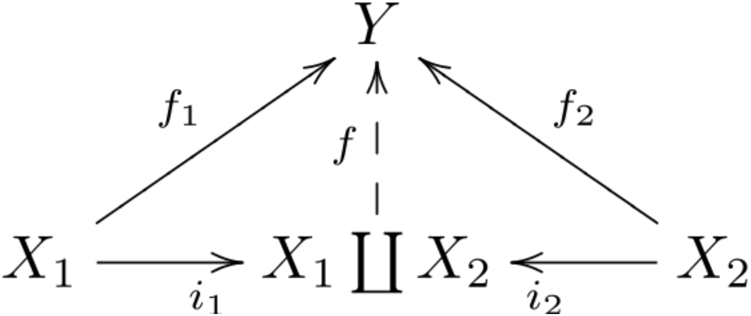


Figure source:
https://www.math3ma.com/blog/limits-and-colimits-part-3

Co-Products



Covariant Functors

A **covariant functor** $F : \mathcal{C} \rightarrow \mathcal{D}$ from category $\mathcal{C}$ to category $\mathcal{D}$ is defined as the following:

- An object $\mathcal{F}X$ of the category $\mathcal{D}$ for each object X in category $\mathcal{C}$.
- A morphism $\mathcal{F}f : \mathcal{F}X \rightarrow \mathcal{F}Y$ in category $\mathcal{D}$ for every morphism $f : X \rightarrow Y$ in category $\mathcal{C}$.
- The preservation of identity and composition: $\mathcal{F}\ id_X = id_{\mathcal{F}X}$ and $(\mathcal{F}g)(\mathcal{F}f) = \mathcal{F}(gf)$ for any composable morphisms $f : X \rightarrow Y, g : Y \rightarrow Z$.

Examples of Covariant Functors

- The "forgetful" functor $F : \mathcal{C}_{\text{MDP}} \rightarrow \textbf{Set}$ that maps an MDP into its set of states S.
- The "PVF" functor $F : \mathcal{C}_{\text{MDP}} \rightarrow \textbf{Graph}$ that maps an MDP into an undirected graph over states S, with an undirected edge between actual transitions.
- The "Top" functor $F : \mathcal{C}_{\text{MDP}} \rightarrow \textbf{Top}$ that maps an MDP into the category of topological spaces.
- The "Grp" functor $F : \mathcal{C}_{\text{MDP}} \rightarrow \textbf{Group}$ that maps an MDP into the category of groups.
- The "Hom" functor $F : \mathcal{C}_{\text{MDP}} \rightarrow \textbf{Modules}$ that maps an MDP into the category of modules.

Contravariant Functors

- A **contravariant functor** $F : \mathcal{C} \to \mathcal{D}$ from category $\mathcal{C}$ to category $\mathcal{D}$ is defined exactly like the covariant functor, except all the mappings are reversed.
- Contravariant functor $\mathcal{F} : C^{\text{op}} \to D$, every morphism $f : X \to Y$ is assigned the reverse morphism $\mathcal{F}f : \mathcal{F}Y \to \mathcal{F}X$ in category $\mathcal{D}$.

Functorial Representations

- 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 $\mathcal{C}(X, Z)$, and to each morphism $f : Y \rightarrow Z$, the pullback mapping $f^* : \mathcal{C}(Z, X) \rightarrow \mathcal{C}(Y, X)$.

Fully Faithful Representations

Let $\mathcal{F} : \mathcal{C} \rightarrow \mathcal{D}$ be a functor from category $\mathcal{C}$ to category $\mathcal{D}$. If for all objects X and Y in $\mathcal{C}$, the map $\mathcal{C}(X, Y) \rightarrow \mathcal{D}(\mathcal{F}X, \mathcal{F}Y)$, denoted as $f \mapsto \mathcal{F}f$ is

- 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**.

Natural Transformations

- Given two functors $\mathcal{F}, \mathcal{G} : \mathcal{C} \rightarrow \mathcal{D}$ that map from category $\mathcal{C}$ to category $\mathcal{D}$, a *natural transformation* $\eta : \mathcal{F} \implies \mathcal{G}$ consists of a morphism $\eta_X : \mathcal{F}X \rightarrow \mathcal{G}X$ for each object X in $\mathcal{C}$.
- For any two functors $\mathcal{F}, \mathcal{G} : \mathcal{C} \rightarrow \mathcal{D}$, let $\text{Nat}(\mathcal{F}, \mathcal{G})$ denote the natural transformations from $\mathcal{F}$ to $\mathcal{G}$. If $\eta_X : \mathcal{F}X \rightarrow \mathcal{G}X$ is an isomorphism for each X in category $\mathcal{C}$, then the natural transformation η is called a **natural isomorphism** and $\mathcal{F}$ and $\mathcal{G}$ are naturally isomorphic, denoted as $\mathcal{F} \cong \mathcal{G}$.

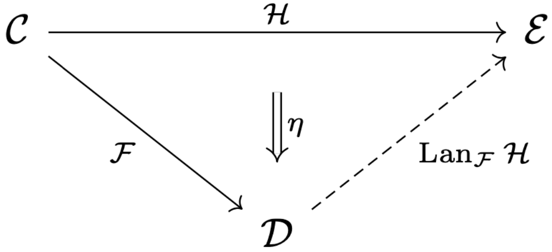
Yoneda Lemma and Pre-Sheaf Representations

Yoneda Lemma: For every object X in category $\mathcal{C}$, and every contravariant functor $\mathcal{F} : \mathcal{C}^{\text{Op}} \rightarrow \mathbf{Set}$, the set of natural transformations from $\mathcal{C}(-, X)$ to $\mathcal{F}$ is isomorphic to $\mathcal{F}X$.

- One of the deepest results in category theory.
- A pair of objects are isomorphic $X \cong Y$ if and only if the corresponding contravariant functors are isomorphic, namely $\mathcal{C}(-, X) \cong \mathcal{C}(-, Y)$.
- Given any two categories $\mathcal{C}, \mathcal{D}$, we can always define the new category $\mathcal{D}^{\mathcal{C}}$, whose objects are functors $\mathcal{C} \rightarrow \mathcal{D}$, and whose morphisms are natural transformations.
- If we take $\mathcal{D} = \mathbf{Set}$, and consider the contravariant version $\mathbf{Set}^{\mathcal{C}^{\text{Op}}}$, we obtain a category whose objects are presheafs.
- Presheafs have some very nice properties, which makes them a *topos*.

Kan Extensions: Approximating Functors

A **left Kan extension** of a functor $H : \mathcal{C} \rightarrow \mathcal{E}$ along F, another functor $F : \mathcal{C} \rightarrow \mathcal{D}$, is a functor $\text{Lan}_F H : \mathcal{D} \rightarrow \mathcal{E}$ with a natural transformation $\eta : H \implies \text{Lan}_F H \circ F$ such that for any other such pair $(G : \mathcal{D} \rightarrow \mathcal{E}, \gamma : H \implies G \circ F)$, γ factors uniquely through η.



Summary

- We proposed a categorial formulation of decision making
  - Decision making objects: MDPs, PSRs, intrinsic models
  - Morphisms: bisimulations
  - Functors: Probe a category by mapping it into a different space
  - Yoneda lemma shows how to construct (pre-sheaf) fully faithful representations
- Birds-eye view of decision making
- Plenty of "frog"-like work to do in terms of applications!