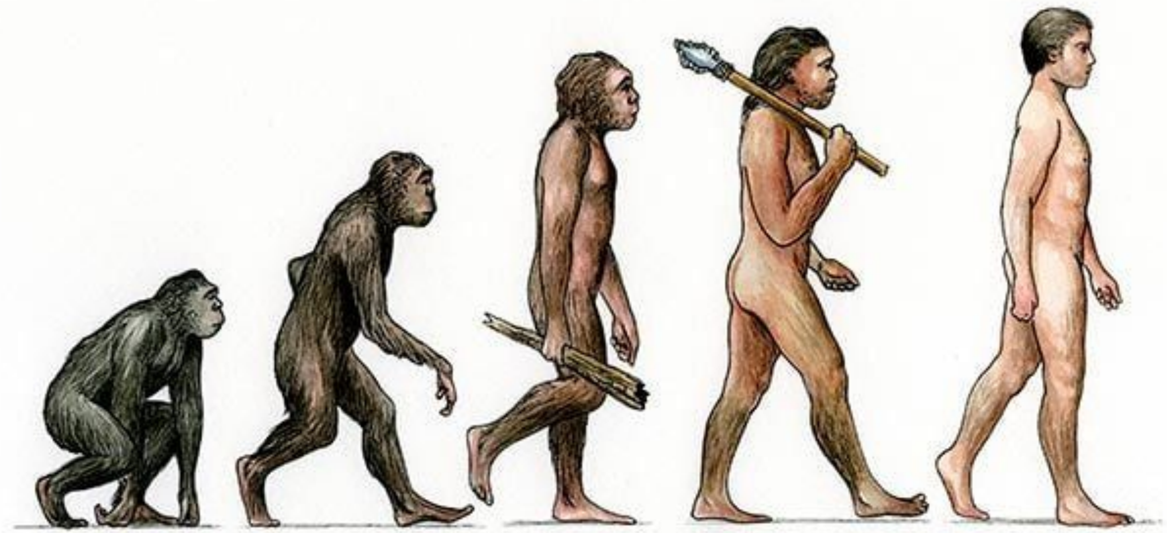
Physics of Learning, and Structure

Siyuan Guo

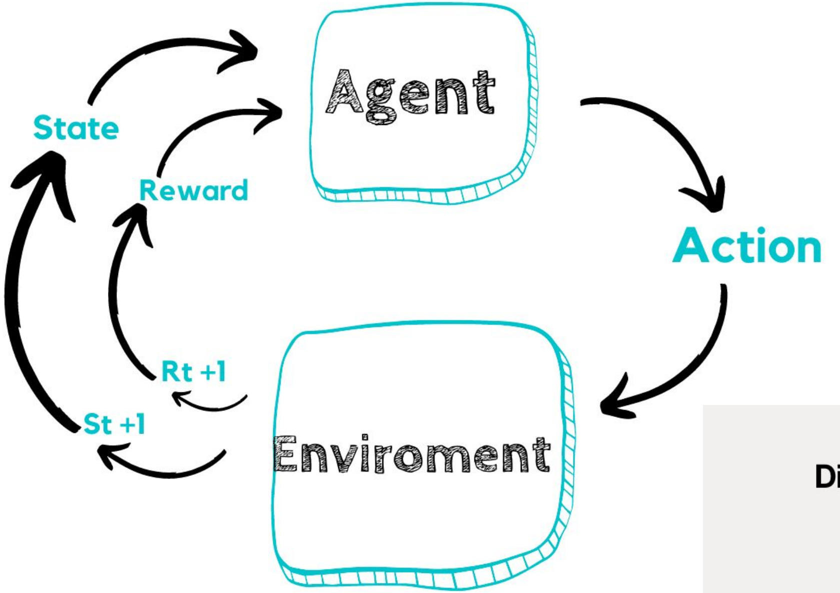
The Learning Brain



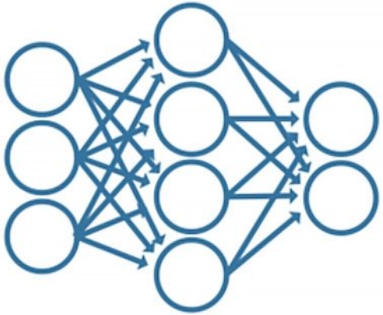
Understand Intelligence

- **Intelligence develops under evolution and the need to survive.**
- **Efficient Learning is key to intelligence arising.**

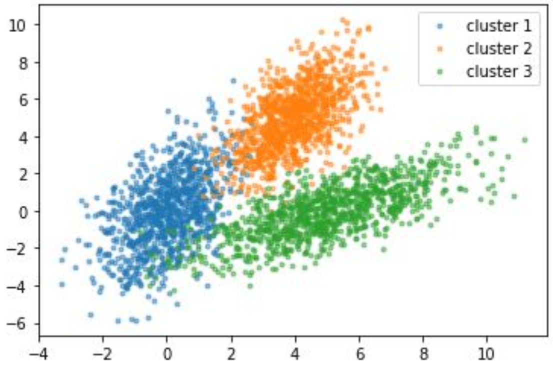
[1] Francois Chollet. ARC Challenges. *On the measure of intelligence.* Preprint. 2019.

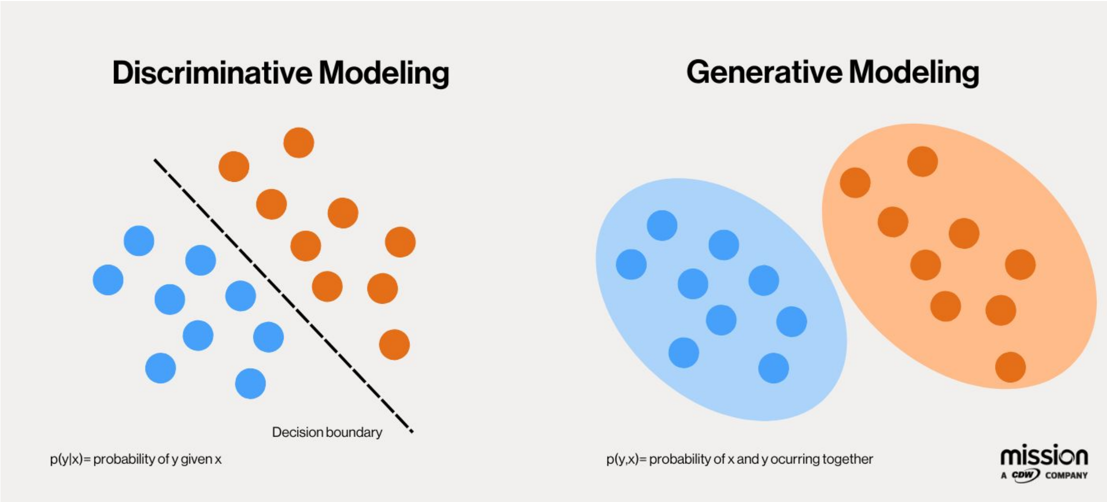


Reinforcement Learning



Unsupervised Learning



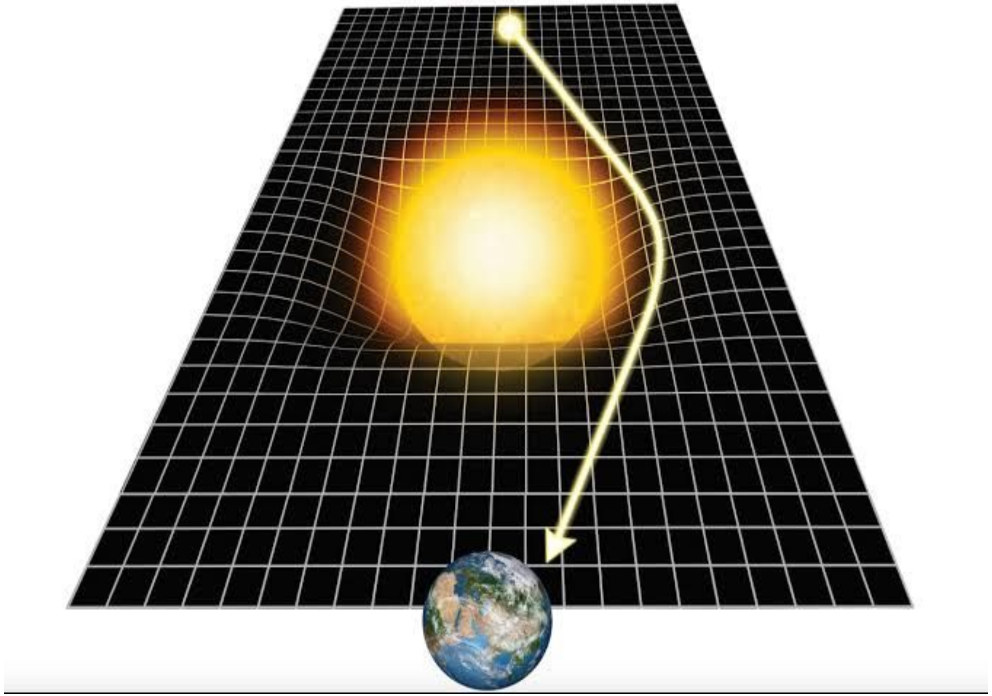


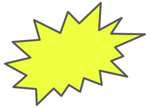
Parameter Estimation

Principle of Least Time [1]

A ray of light travels from point A to point B chooses a path along which the time is the least or minimum.

$$\mathrm{T} = \int_A^B ds = \int_A^B \frac{1}{v} ds$$

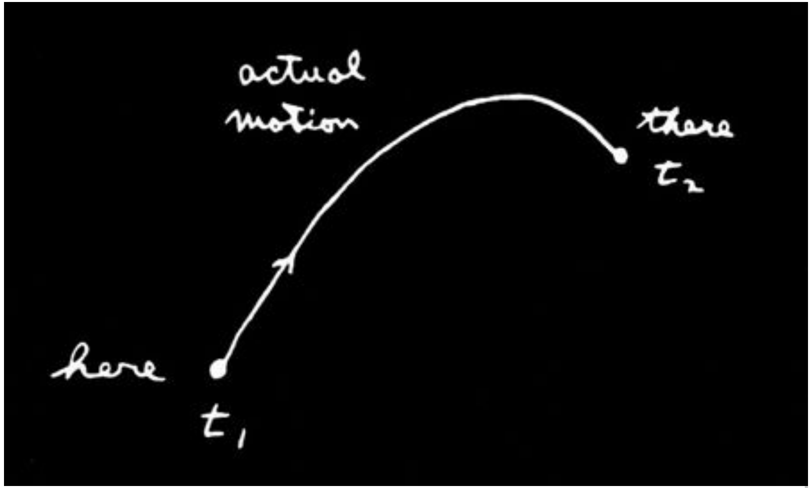




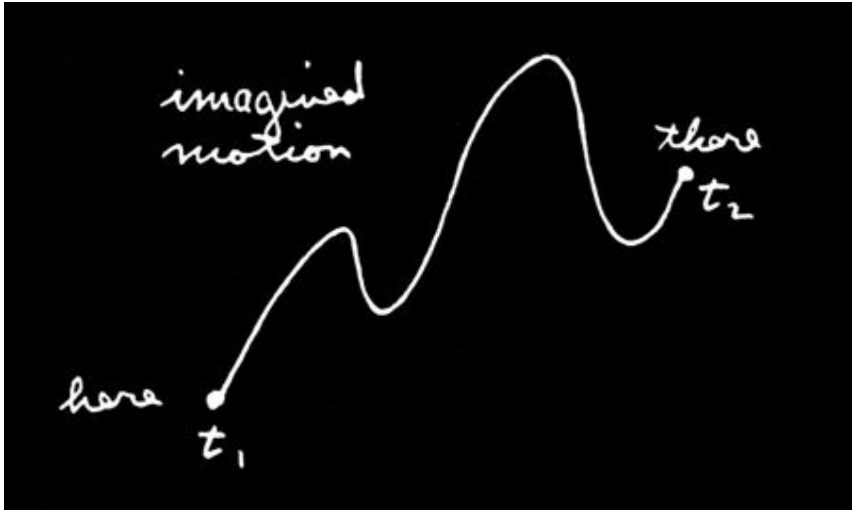
Designing an efficient learning system is as if walking along an information space such that we can find a minimum path that reaches the generalization error in the shortest time.

[1] Max Born and Emil Wolf. Principles of Optics: 60th Anniversary Edition. Cambridge University Press, 7 edition, 2019

Hamilton's Law: Principle of Least Action [2]



$$\int_{t_1}^{t_2} \left[\frac{1}{2}m\left(\frac{dx}{dt}\right)^2 - mxg\right] dt$$



The true path is the one for which that integral is least.

[2] Feynman, Richard P. (Richard Phillips), 1918-1988. The Feynman Lectures on Physics.

Learning, too, follows the laws of physics – principle of least action.

Physics of Learning: A Lagrangian perspective to different learning paradigms

Siyuan Guo *
Department of Computer Science, University of Cambridge
Max Planck Institute for Intelligent Systems
United Kingdom & Germany

Bernhard Schölkopf
Max Planck Institute for Intelligent Systems
ELLIS Institute Tübingen
Germany

* *Correspondence:* `siyuan.guo@tuebingen.mpg.de`

Abstract

We study the problem of building an efficient learning system. Efficient learning processes information in the least time, i.e., building a system that reaches a desired error threshold with the least number of observations. Building upon least action principles from physics, we derive classic learning algorithms, Bellman's optimality equation in reinforcement learning, and the Adam optimizer in generative models from first principles, i.e., the Learning *Lagrangian*. We postulate that learning searches for stationary paths in the Lagrangian, and learning algorithms are derivable by seeking the stationary trajectories.

Understand Kinematics Movement in Information Space

Table 1: Overview of Kinematics in Information Space.

	In-context Learning	**Pre-train Learning**
Data source	$s_1, \ldots, s_t$	$B_1, B_2, \ldots, B_t$
Position at time t: $P(t)$	$-\log p(s_1, s_2, \ldots, s_t)$	$\epsilon_t = \epsilon(B_1, \ldots, B_t)$
Velocity at time t $v(t) = \lim_{\delta \to 0} \frac{P(t+\delta)-P(t)}{\delta}$	$-\log p(s_{t+1} \mid s_{\leq t})$	$\epsilon_{t+1} - \epsilon_t$
Acceleration at time t $a(t) = \lim_{\delta \to 0} \frac{v(t+\delta)-v(t)}{\delta}$	$-\log p(s_{t+2} \mid s_{\leq t+1}) + \log p(s_{t+1} \mid s_{\leq t})$	$\ldots$
Applications (velocity curve)	In-context learning loss	Test loss curve

Learning as a deceleration process.

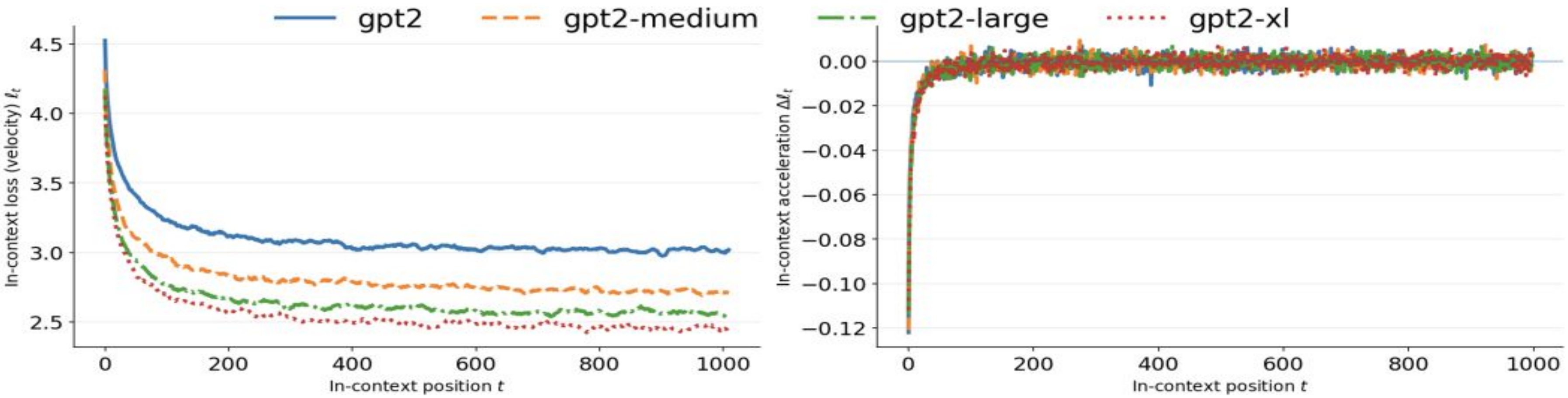


Figure 1: Expected test-time in-context learning velocity and acceleration: (Left) In-context per-token loss $\ell_t = v(t) = \mathbb{E}[-\log p_\theta(x_t \mid x_{<t})]$; (Right) In-context per-token difference in loss $\Delta\ell_t = a(t) = \mathbb{E}[\ell_{t+1} - \ell_t]$. In-context learning (as shown in the right) is a deceleration process, meaning loss goes down but less quickly as time progresses. A similar phenomenon is expected in training and test loss. Here, in-context loss is evaluated on OpenWebText.

Optimizing Physics Inspired Lagrangians.

	Physics
Fermat's principle	$T = \int_A^B dt$
Hamiltonian	$H(\mathbf{x}, \mathbf{p}) = \mathbf{p} \cdot \dot{\mathbf{x}} - L(\mathbf{x}, \dot{\mathbf{x}})$
the *Lagrangian*	$S[\xi] = \int L\, dt$, where $L = T - V$

Efficient learning is as if designing a process of walking along the information space such that it takes the least time (1) to reach the desired error threshold.

$$\mathrm{T}(\delta) = \min_{\mathbf{s}} \int_0^\infty \Theta(\epsilon[\mathbf{s}] - \delta)\, dt$$

(1) Least time = least sample size if information content of each sample is approximately similar. It is open investigation on how to quantify time when varying information content.

Learning Objective cf. Fermat

$$\mathrm{T}(\delta) = \min_{\mathbf{s}} \int_0^{\infty} \Theta(\epsilon[\mathbf{s}] - \delta)\, dt$$

T_{sample}	⬇	Intrinsic Intelligence	⬆
$T_{computation}$	⬇	Real-life processing time	⬇

Unknown generalization error a priori for optimization.

Optimize

$$\mathrm{T}(\delta) = \min_{\mathbf{s}} \int_0^\infty \Theta(\epsilon[\mathbf{s}] - \delta)\, dt$$

→

Choose the observational path such that the generalization error is minimized with the least number of observations.

Demo on Parametric Models: Linear Regression

$$y = x^T \beta + \epsilon, x \in \mathbb{R}^p, \mathbb{E}[\epsilon] = 0, \mathrm{Var}(\epsilon) = \sigma^2$$

$$\textbf{generalization error} : \epsilon(\mathbf{x})$$

$$\epsilon(\mathbf{x}) = \sigma^2 + \sigma^2 \mathrm{tr}((X^T X)^{-1} \mathbb{E}[x x^T]),$$

Where x is the test data point and $\mathbf{X}$ is the sequence of observations as rows of X.

Assume inputs satisfy unit norm $||x_i||_2 = 1$ and uniformly drawn from the unit sphere $\mathbb{S}^{p-1}$

Active Learning

$$\min_{\mathbf{x}} \epsilon(\mathbf{x}) = \sigma^2 + \sigma^2 \frac{p}{n}$$

Key: Find $\mathbf{x}$ such that $S = X^T X = \frac{n}{p} I_p$

Does optimal data path walks along a continuous path?

$$S = X^T X = \sum x_i x_i^T$$

$x_i x_i^T \neq \frac{1}{p} I_p$ due to rank difference

Blocks of p data points to always remain on the most efficient learning path.

Planning is needed to learn continuously in the most efficient way.

Reinforcement Learning

$$\mathrm{T}(\delta) = \min_{\mathbf{s}} \int_0^\infty \Theta(\epsilon[\mathbf{s}] - \delta)\, dt$$

Reinforcement Learning is optimizing the above function. Unknown error = reward

1. Optimize action subject to constraints is optimizing data / state path in RL, cf. min_s.
2. Unknown step-wise generalization is replaced by reward r(s, a).(speed can be incorporated, e.g., r = -1)

Find states $s_0, s_1, \ldots, s_n$
and actions $a_0, a_1, \ldots, a_{n-1}$
to maximize the objective function J,

$$J = h(s_n) + \int_0^{t_f} r(s_t, a_t, t) dt,$$

subject to constraint $s_{t+1} = f(s_t, a_t)$,
and t_f is the final time.

Technical Details

How Bellman Optimality Equation is a solution for seeking stationary path in the Lagrangian, and its connection to Hamiltonian system (row 2 below).

	Physics
Fermat's principle	$T = \int_A^B dt$
Hamiltonian	$H(\mathbf{x}, \mathbf{p}) = \mathbf{p} \cdot \dot{\mathbf{x}} - L(\mathbf{x}, \dot{\mathbf{x}})$
the *Lagrangian*	$S[\xi] = \int L\, dt, \text{where } L = T - V$

Constrained Lagrangian:

$$\mathcal{L}(\{\mathbf{s}\},\{\mathbf{a}\},\lambda) = h(\mathbf{s}_n) + \sum_{k=0}^{n-1}(r(\mathbf{s}_k,\mathbf{a}_k,k) + (f(\mathbf{s}_k,\mathbf{a}_k) - s_{k+1})^T\lambda_{k+1}$$

Seeking stationary solutions: $\frac{\partial \mathcal{L}}{\partial s_k} = \frac{\partial \mathcal{L}}{\partial a_k} = \frac{\partial \mathcal{L}}{\partial \lambda} = 0, \forall k$

Define discrete-time Hamiltonian: $H^{(k)}(\mathbf{s},\mathbf{a},\lambda) = r(\mathbf{s},\mathbf{a},k) + f(\mathbf{s},\mathbf{a})^T\lambda$

Re-write the Lagrangian:

$$\mathcal{L} = h(\mathbf{s}_n) - \mathbf{s}_n^T\lambda_n + \mathbf{s}_0^T\lambda_0 + \sum_{k=1}^{n-1}\left(H^{(k)}(\mathbf{s}_k,\mathbf{a}_k,\lambda_{k+1}) - \mathbf{s}_k^T\lambda_k\right)$$

$$d\mathcal{L} = (\nabla_{\mathbf{s}} h(\mathbf{s}_n) - \lambda_n)^T d\mathbf{s}_n + \lambda_0^T d\mathbf{s}_0 + \sum_{k=1}^{n-1}\left(\frac{\partial H^{(k)}}{\partial \mathbf{s}_k} - \lambda_k\right)^T d\mathbf{s}_k + \left(\frac{\partial H^{(k)}}{\partial \mathbf{a}_k}\right) d\mathbf{a}_k$$

[1] Derivation is classical control theory textbook material.

$$d\mathcal{L} = (\nabla_{\mathbf{s}} h(\mathbf{s}_n) - \lambda_n)^T d\mathbf{s}_n + \lambda_0^T d\mathbf{s}_0 + \sum_{k=1}^{n-1} \left(\frac{\partial H^{(k)}}{\partial \mathbf{s}_k} - \lambda_k\right)^T d\mathbf{s}_k + \left(\frac{\partial H^{(k)}}{\partial \mathbf{a}_k}\right) d\mathbf{a}_k$$

Stationary solutions needs to satisfy below constraints:

$$\lambda_n = \nabla_{\mathbf{s}} h(\mathbf{s}_n)$$

Terminal Reward

$$\lambda_k = \frac{\partial r(\mathbf{s}_k, \mathbf{a}_k, k)}{\partial \mathbf{s}_k} + \frac{\partial f(\mathbf{s}_k, \mathbf{a}_k)}{\partial \mathbf{s}_k}^T \lambda_{k+1}$$

$$\mathbf{a}_k = \arg\max_u H^{(k)}(\mathbf{s}_k, u, \lambda_{k+1}) \text{ implies } \frac{\partial H^{(k)}}{\partial \mathbf{a}_k} = 0$$

Let $\lambda_k = \nabla_{\mathbf{s}} V(\mathbf{s}_k)$

$$\nabla_{\mathbf{s}} V(\mathbf{s}_k) = \frac{\partial r(\mathbf{s}_k, \mathbf{a}_k, k)}{\partial \mathbf{s}_k} + \frac{\partial f(\mathbf{s}_k, \mathbf{a}_k)}{\partial \mathbf{s}_k}^T \nabla_{\mathbf{s}} V(\mathbf{s}_{k+1})$$

$$V(\mathbf{s}_k) = r(\mathbf{s}_k, \mathbf{a}_k, k) + V(f(\mathbf{s}_k, \mathbf{a}_k))$$

Bellman Optimality Equation [1]

$$V(\mathbf{s}_k) = \max_u \{r(\mathbf{s}_k, u, k) + V(f(\mathbf{s}_k, u))\}$$

[1] Richard Bellman. Dynamic programming and stochastic control processes. 1958.

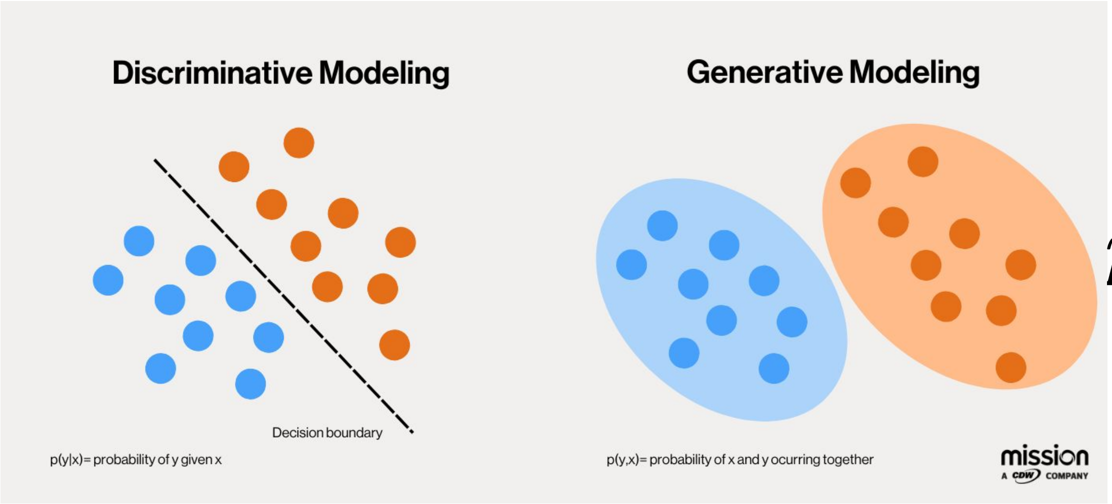
The solution for stationary path in the Lagrangian, written in terms of rewards, satisfies Bellman's optimality equation. Thus, optimizing Bellman's equation is searching for the stationary path.

$$H^{(k)}(\mathbf{s}, \mathbf{a}, \lambda) = r(\mathbf{s}, \mathbf{a}, k) + f(\mathbf{s}, \mathbf{a})^T \lambda$$

	Physics
Fermat's principle	$T = \int_A^B dt$
Hamiltonian	$H(\mathbf{x}, \mathbf{p}) = \mathbf{p} \cdot \dot{\mathbf{x}} - L(\mathbf{x}, \dot{\mathbf{x}})$
the *Lagrangian*	$S[\xi] = \int L\, dt$, where $L = T - V$

Efficient Learning <> Least Time <> Seek Stationary Path

$p_\theta(y \mid x)$



$p_\theta(x)$

$$L = T - V$$
$$= \text{Kinetic} - \text{Potential}$$

$$\text{Newtonian: } \frac{1}{2}mv^2 - mxg$$

The Spark for A Postulation

Potential energy is something intrinsic about the task: $\ell(\theta; x)$

Kinetic energy? And efficiency?

Cramer Rao Lower Bound:

Let $\hat{\theta}$ be an unbiased estimator. Then under regularity conditions,

$$\text{Var}(\theta) - I(\theta)^{-1}$$

is positive semi-definite. An unbiased estimator attains the lower bound is an efficient estimator.

$$I(\theta) := \mathbb{E}[(\nabla_\theta \ell(\theta; x))(\nabla_\theta \ell(\theta; x))^T]$$
$$= -\mathbb{E}[\frac{\partial^2}{\partial\theta\partial\theta^T}\ell(\theta; x)]$$

Newtonian: $\frac{1}{2}mv^2 - mxg$

A Postulation

$$\mathcal{L}(\ell, \nabla_\theta \ell) = T - V = \frac{1}{2P}(\nabla_\theta \ell)^T F(\theta)^{-1}(\nabla_\theta \ell) - \ell(\theta; x)$$

where P is the number of model parameters.

Stationary solution satisfies the Euler-Lagrange equation for scalar field theory

$$\mathbb{E}[\frac{\partial \mathcal{L}}{\partial \ell}] = \mathbb{E}\Big[\frac{\partial}{\partial t}(\frac{\partial \mathcal{L}}{\partial \dot{\ell}}) + \sum_i \frac{\partial}{\partial \theta_i}\frac{\partial \mathcal{L}}{\partial(\partial \ell / \partial \theta_i)}\Big]$$

$$-1 = \mathbb{E}[\nabla_\theta \cdot \frac{\partial \mathcal{L}}{\partial \nabla_\theta l}]$$

$$-1 = \frac{1}{P}\mathbb{E}[\nabla_\theta \cdot (F^{-1}\nabla_\theta l)] \quad \text{due to } \frac{\partial \mathcal{L}}{\partial \nabla_\theta \ell} = F^{-1}\nabla_\theta \ell$$

$$-1 = \frac{1}{P}\text{tr}(\nabla_\theta (F(\theta)^{-1}) \underbrace{\mathbb{E}[\nabla_\theta l]}_{=0 \text{ at stationary points}} + F^{-1}\mathbb{E}[\nabla_\theta^2 l]]) = \frac{1}{P}\text{tr}\big(F^{-1}\underbrace{\mathbb{E}[\nabla_\theta^2 l]}_{=-F}\big) = -1$$

The solution at stationary points needs to be an MLE.

A Postulation cont.

Learning dynamics of the loss fields operationalizes through changes in particle dynamics (i.e., changes in parameter):

$$\mathcal{L}(\ell, \nabla_\theta \ell) = T - V = \frac{1}{2P}(\nabla_\theta \ell)^T F(\theta)^{-1}(\nabla_\theta \ell) - \ell(\theta; x)$$

$$\frac{1}{2} m \dot{\theta}^T \dot{\theta} - \ell(\theta), m = \frac{1}{P}, \dot{\theta} = F^{-1/2} \nabla_\theta \ell$$

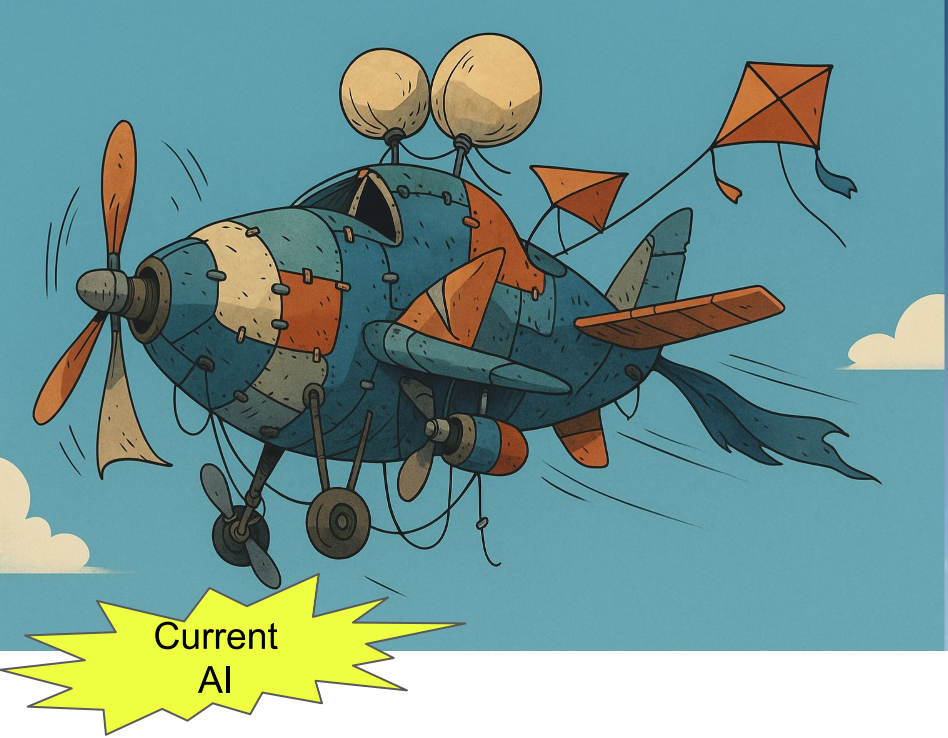
$$\textit{RMSprop: } \theta_{t+1} \leftarrow \theta_t - \alpha \frac{g_t}{\sqrt{v_t} + \epsilon}, \quad (31)$$

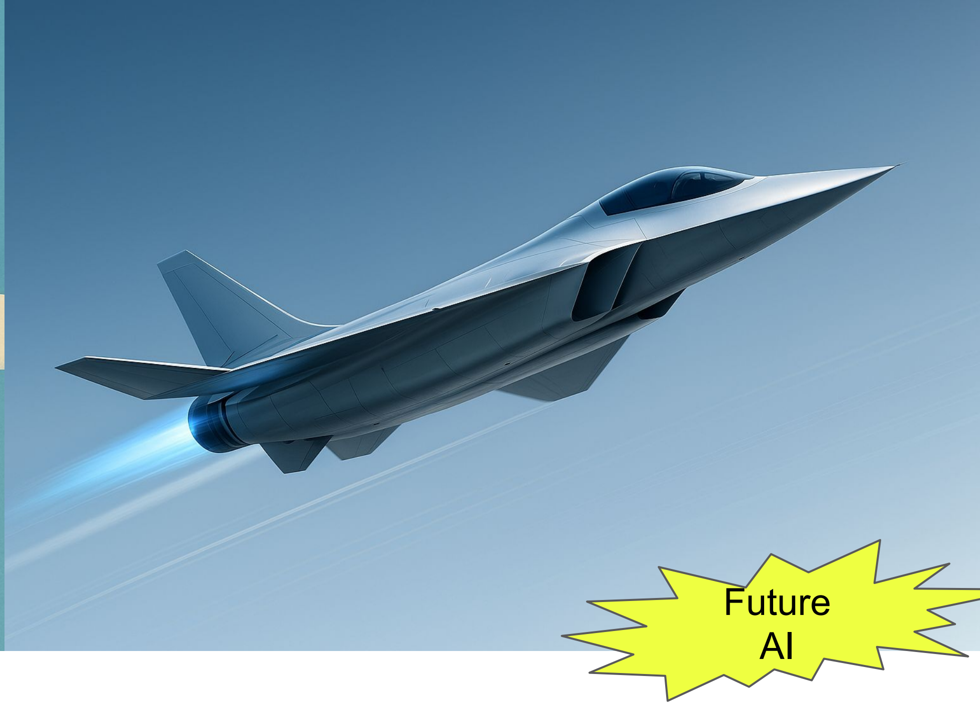
$$\textit{Adam: } \theta_{t+1} \leftarrow \theta_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}, \quad (32)$$

where $g_t = \nabla_{\theta_t} \ell$, $v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t \odot g_t$, and $m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$, $\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$, $\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$, and ϵ are added for numerical stability. From the Lagrangian, one can also predict the inefficiency of SGD, as it does not satisfy the Euler-Lagrange equation. Combining with Section 3.2

	Physics	Learning
Fermat's principle	$T = \int_A^B dt$	$T = \int_{\epsilon[\emptyset]}^{\epsilon[\mathbf{s}]} dt$ [*]
Hamiltonian	$H(\mathbf{x}, \mathbf{p}) = \mathbf{p} \cdot \dot{\mathbf{x}} - L(\mathbf{x}, \dot{\mathbf{x}})$	$H(\mathbf{s}, \mathbf{a}, \lambda) = r(\mathbf{s}, \mathbf{a}) + f(\mathbf{s}, \mathbf{a})^T \lambda$ [†]
the *Lagrangian*	$L = T - V$	$L(\ell, \nabla_\theta \ell) = \frac{1}{2}(\nabla_\theta \ell)^T F^{-1} \nabla_\theta \ell - \ell(\theta)$ [*]
	Applications	**Algorithms**
Fermat's principle	Parametric Models	A-optimality (Atkinson et al., 2007)
Hamiltonian	Reinforcement Learning	Bellman's Equation (Bellman, 1958)
the *Lagrangian*	Generative Models / Supervised Learning	Adam (Kingma, 2014) / RMSprop (Tieleman, 2012)

Notes: T in Fermat's principle denotes time taken to travel from point A to point B; $\epsilon[\emptyset], \epsilon[\mathbf{s}]$ is the generalization error after observing zero data to data sequence $\mathbf{s} := s_1, s_2, \ldots$; H is the (physical) Hamiltonian system with position $\mathbf{x}$ and momentum $\mathbf{p}$ and Lagrangian L; $H(\mathbf{s}, \mathbf{a}, \lambda)$ is the reinforcement learning correspondent with state $\mathbf{s}$, action $\mathbf{a}$, reward $r(\mathbf{s}, \mathbf{a})$, transition dynamics $f(\mathbf{s}, \mathbf{a})$ and momentum equivalent λ; $L = T - V$ represents kinetic energy minus potential energy; ℓ denotes some log-likelihood function; $\nabla_\theta \ell$ is gradient with respect to model parameters $\theta \in \mathbb{R}^P$; F^{-1} denotes the inverse Fisher information. Bold symbols are vectors; $(\cdot)^\top$ is transpose; $\dot{x}$ is derivative with respect to time. The learning Lagrangian indicated via [†] means it is classic textbook material in control theory (see Todorov (2006)). Learning Lagrangians indicated by [*] are proposed in this work; to the best of our knowledge, no prior published work exists as of September 2025.



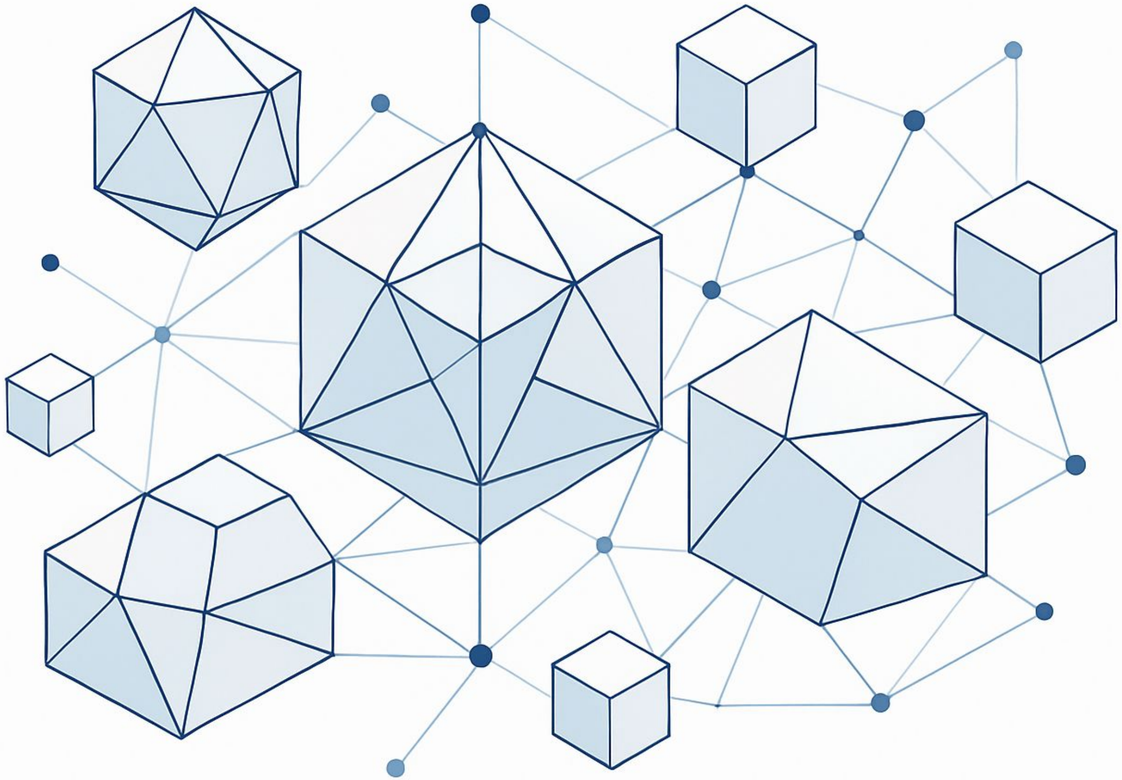


1. Postulate: Learning too adheres to the laws of physics.
2. Demo: the principle of least action.

→ Testable predictions via Community Efforts.

(looking for discussions / collaborations on experiments to verify)

Structures



Structures in Causality.

A way to describe data in terms of structure.

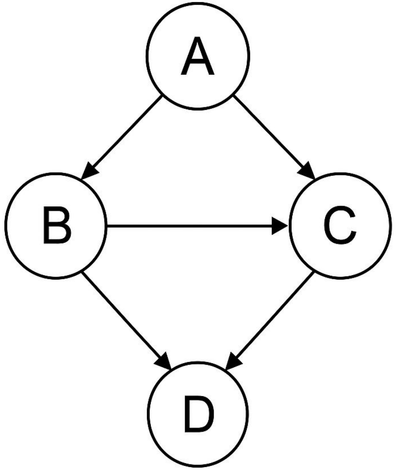
A set of observables $X_1, X_2, \ldots, X_n$

A set of structural assignments:

$$\boldsymbol{X_i := f_i(X_{pa}, U_i), \forall i}$$

The model is Markovian if U_i are all jointly independent.

– Structural Causal Model [1]



[1] Pearl J. Causality. 2nd ed. Cambridge University Press; 2009.

Empirically Infer Structure from Data

Object and Perception process are independent.

$$P(X_i \mid PA_i)$$

Is independent of

$$P(X_j \mid PA_j)$$

In the sense of:
No information exchange
No intervention propagation.

Causal de Finetti

Causal de Finetti: On the Identification of Invariant Causal Structure in Exchangeable Data

Siyuan Guo[1,2*] Viktor Tóth[1*] Bernhard Schölkopf[2] Ferenc Huszár[1]
[1]University of Cambridge [2]Max Planck Institute for Intelligent Systems
{syg26,fh277}@cam.ac.uk toth.viktor7400@gmail.com
bs@tuebingen.mpg.de

NeurIPS 2023

Theorem.
Let $\{(X_i, Y_i)\}$ be an infinite sequence of binary random variables.

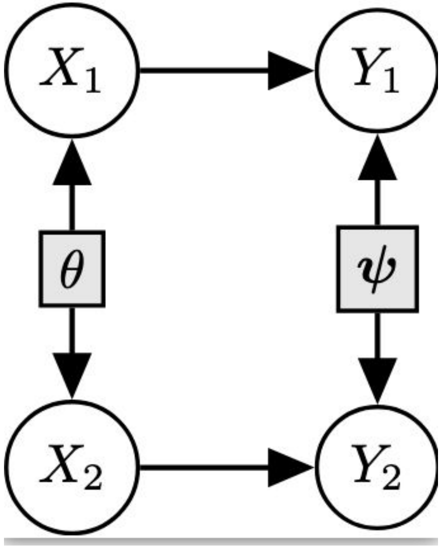
Suppose:

1. The sequence is infinitely exchangeable,
2. $\forall n \in \mathbb{N} : Y_{[n]} \perp\!\!\!\perp X_{n+1} \mid X_{[n]}$, where $[n] = \{1, \ldots, n\}$

Then there exists two latent variables θ, ψ with suitable probability measure μ, ν such that

$$P(X_1 = x_1, Y_1 = y_1, \ldots, X_n = x_n, Y_n = y_n)$$
$$= \int \prod_{i=1}^{n} p(y_i | x_i, \psi) p(x_i \mid \theta) d\mu(\theta) d\nu(\psi)$$

Causal de Finetti



Theorem.
Let $\{(X_i, Y_i)\}$ be an infinite sequence of binary random variables.

Suppose:

1. The sequence is infinitely exchangeable,
2. $\forall n \in \mathbb{N} : Y_{[n]} \perp\!\!\!\perp X_{n+1} \mid X_{[n]}$, where $[n] = \{1, \ldots, n\}$

Not Inform and Not Influence:
$\exists \theta, \psi$ and $\theta \perp\!\!\!\perp \psi$

Then there exists two latent variables θ, ψ with suitable probability measure μ, ν such that

$$P(X_1 = x_1, Y_1 = y_1, \ldots, X_n = x_n, Y_n = y_n)$$
$$= \int \prod_{i=1}^{n} p(y_i | x_i, \psi) p(x_i \mid \theta) d\mu(\theta) d\nu(\psi)$$

Causal de Finetti

Exchangeable (de Finett):

$$P(X_1, \ldots, X_n) = P(X_{\pi(1)}, \ldots, X_{\pi(n)})$$

$$P(X_1, \ldots, X_n) = \int \prod_{i=1}^n p(x_i \mid \theta) d\mu(\theta)$$

$$Y_{[n]} \perp\!\!\!\perp X_{[n]} \to P(Y \mid X)$$

$$X_{n+1} \sim P(X)$$

$$(X_n, Y_n) \not\perp (X_m, Y_m)$$

Theorem.
Let $\{(X_i, Y_i)\}$ be an infinite sequence of binary random variables.

Suppose:

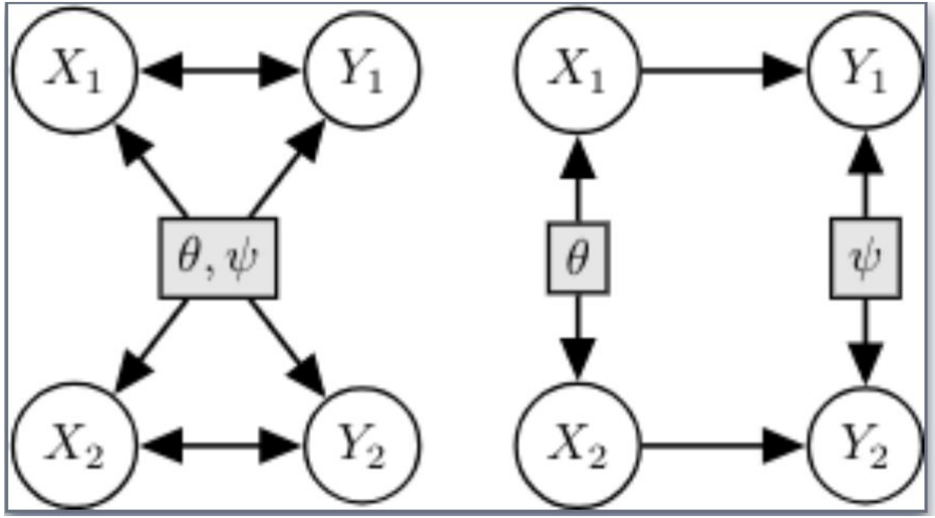
1. The sequence is infinitely exchangeable,
2. $\forall n \in \mathbb{N} : Y_{[n]} \perp\!\!\!\perp X_{n+1} \mid X_{[n]}$, where $[n] = \{1, \ldots, n\}$

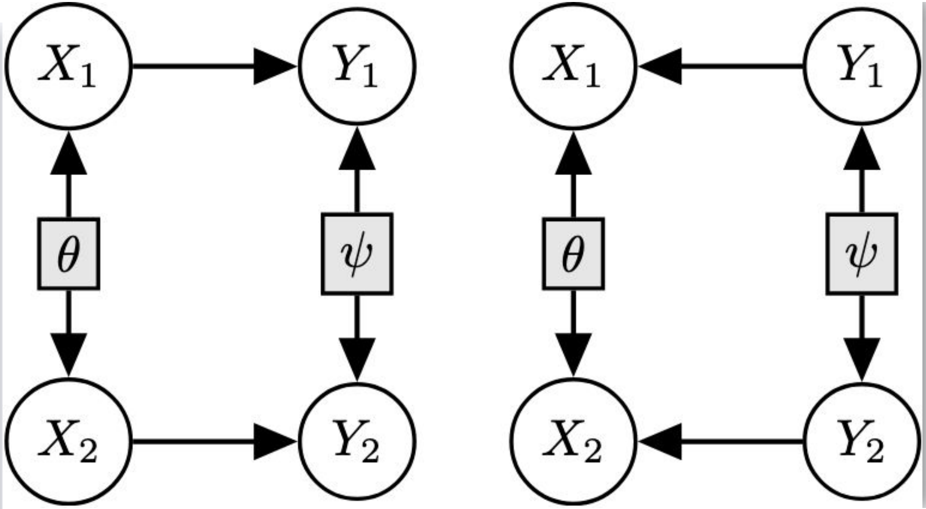
Then there exists two latent variables θ, ψ with suitable probability measure μ, ν such that

$$P(X_1 = x_1, Y_1 = y_1, \ldots, X_n = x_n, Y_n = y_n)$$

$$= \int \prod_{i=1}^{n} p(y_i | x_i, \psi) p(x_i \mid \theta) d\mu(\theta) d\nu(\psi)$$

Impact: Understanding and Structure Discovery





Independent mechanisms under mixture data means independent latent variables controlling each mechanisms.

$$X_1 \perp\!\!\!\perp Y_2 \mid X_2 \qquad X_1 \perp\!\!\!\perp Y_2 \mid Y_1$$

Heterogenous data sources enable structure identification.

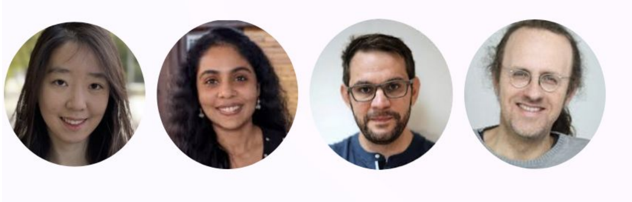
Effect Estimation

Do Finetti: on Causal Effects for Exchangeable Data

Siyuan Guo[14]* **Chi Zhang**[2] **Karthika Mohan**[3] **Ferenc Huszár**[4†] **Bernhard Schölkopf**[1†]
[1]Max Planck Institute for Intelligent Systems [2]Toyota Research Institute
[3] Oregon State University [4] University of Cambridge
† Equal supervision

Abstract

We study causal effect estimation in a setting where the data are not i.i.d. (independent and identically distributed). We focus on exchangeable data satisfying an assumption of independent causal mechanisms. Traditional causal effect estimation frameworks, e.g., relying on structural causal models and do-calculus, are typically limited to i.i.d. data and do not extend to more general exchangeable generative processes, which naturally arise in multi-environment data. To address this gap, we develop a generalized framework for exchangeable data and introduce a truncated factorization formula that facilitates both the identification and estimation of causal effects in our setting. To illustrate potential applications, we introduce a causal Pólya urn model and demonstrate how intervention propagates effects in exchangeable data settings. Finally, we develop an algorithm that performs simultaneous causal discovery and effect estimation given multi-environment data.



NeurIPS 2024 **oral**

Representation Learning

IDENTIFIABLE EXCHANGEABLE MECHANISMS FOR CAUSAL STRUCTURE AND REPRESENTATION LEARNING

Patrik Reizinger[* 1,3], **Siyuan Guo**[* 1,2], **Ferenc Huszár**[2], **Bernhard Schölkopf** [† 1,3], **and Wieland Brendel** [† 1,3,4]

[1]Max Planck Institute for Intelligent Systems, Tübingen, Germany
[2]University of Cambridge, Cambridge, United Kingdom
[3]ELLIS Institute Tübingen, Tübingen, Germany
[4]Tübingen AI Center, Tübingen, Germany

ABSTRACT

Identifying latent representations or causal structures is important for good generalization and downstream task performance. However, both fields developed rather independently. We observe that several structure and representation identifiability methods, particularly those that require multiple environments, rely on exchangeable non–i.i.d. (independent and identically distributed) data. To formalize this connection, we propose the Identifiable Exchangeable Mechanisms (IEM) framework to unify key representation and causal structure learning methods. IEM provides a unified probabilistic graphical model encompassing causal discovery, Independent Component Analysis, and Causal Representation Learning. With the help of the IEM model, we generalize the Causal de Finetti theorem of Guo et al. (2024a) by relaxing the necessary conditions for causal structure identification in exchangeable data. We term these conditions cause and mechanism variability, and show how they imply a duality condition in identifiable representation learning, leading to new identifiability results.

ICLR 2025 **spotlight**

Counterfactual Reasoning

Counterfactual reasoning: an analysis of in-context emergence

Moritz Miller[12]* **Bernhard Schölkopf**[12] **Siyuan Guo**[13]
[1]Max Planck Institute for Intelligent Systems [2]ETH Zurich [3] University of Cambridge

Abstract

Large-scale neural language models (LMs) exhibit remarkable performance in in-context learning: the ability to learn and reason the input context on the fly without parameter update. This work studies in-context counterfactual reasoning in language models, that is, to predict the consequences of changes under hypothetical scenarios. We focus on studying a well-defined synthetic setup: a linear regression task that requires noise abduction, where accurate prediction is based on inferring and copying the contextual noise from factual observations. We show that language models are capable of counterfactual reasoning in this controlled setup and provide insights that counterfactual reasoning for a broad class of functions can be reduced to a transformation on in-context observations; we find self-attention, model depth, and data diversity in pre-training drive performance in Transformers. More interestingly, our findings extend beyond regression tasks and show that Transformers can perform noise abduction on sequential data, providing preliminary evidence on the potential for counterfactual story generation. Our code is available here.

NeurIPS 2025

Reinforcement Learning

Skill Learning via Policy Diversity Yields Identifiable Representations for Reinforcement Learning

Patrik Reizinger*[1,2], Bálint Mucsányi*[3,4], Siyuan Guo*[1,5], Benjamin Eysenbach[6], Bernhard Schölkopf†[1,2], and Wieland Brendel†[1,2]

[1]Max Planck Institute for Intelligent Systems, Tübingen, Germany
[2]ELLIS Institute Tübingen, Tübingen, Germany
[3]University of Tübingen, Tübingen, Germany
[4]Tübingen AI Center, Tübingen, Germany
[5]University of Cambridge, Cambridge, United Kingdom
[6]Princeton University, Department of Computer Science, Princeton, United States

Abstract

Self-supervised feature learning and pretraining methods in reinforcement learning (RL) often rely on information-theoretic principles, termed mutual information skill learning (MISL). These methods aim to learn a representation of the environment while also incentivizing exploration thereof. However, the role of the representation and mutual information parametrization in MISL is not yet well understood theoretically. Our work investigates MISL through the lens of identifiable representation learning by focusing on the Contrastive Successor Features (CSF) method. We prove that CSF can provably recover the environment's ground-truth features up to a linear transformation due to the inner product parametrization of the features and skill diversity in a discriminative sense. This first identifiability guarantee for representation learning in RL also helps explain the implications of different mutual information objectives and the downsides of entropy regularizers. We empirically validate our claims in MuJoCo and DeepMind Control and show how CSF provably recovers the ground-truth features both from states and pixels.

Under Review

Computational Linguistics

On the Emergence and Test-Time Use of Structural Information in Large Language Models

Anonymous ACL submission

Abstract

Learning structural information from observational data is central to producing new knowledge outside the training corpus. This holds for mechanistic understanding in scientific discovery as well as flexible test-time compositional generation. We thus study how language models learn abstract structures and utilize the learnt structural information at test-time. To ensure a controlled setup, we design a natural language dataset based on linguistic structural transformations. We empirically show that the emergence of learning structural information correlates with complex reasoning tasks, and that the ability to perform test-time compositional generation remains limited.

To ensure a controlled synthetic playground, we generate a natural language dataset based on *Transformational Grammar (TG)* (Chomsky, 1957; Radford, 1988). This allows us to analyze *whether* and *how* the model learns the emergence of structure during training, analyze *whether* they can compose learnt structures at test-time, and provide evidence on where in the model this behavior occurs. By doing so, we shine light on how LLMs can generate sentences beyond those directly observed in the corpus. Our contributions are:

- We introduce a natural language dataset based on linguistic structural transformations to formally study structural information in language (Section 3.1).

Under Review

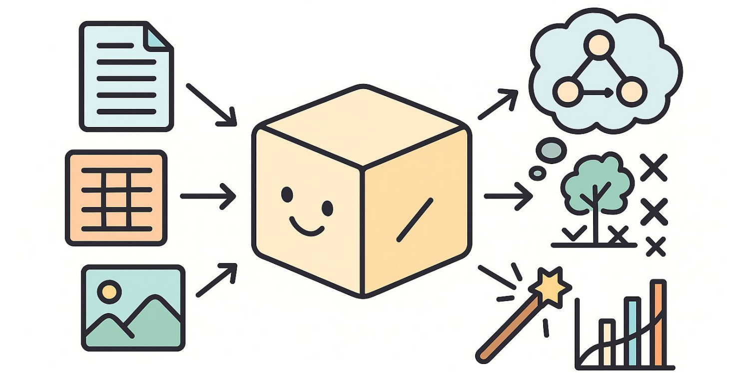
Foundation Models for Causal Inference.

DO-PFN: IN-CONTEXT LEARNING FOR CAUSAL EFFECT ESTIMATION

Jake Robertson[*1,4], **Arik Reuter**[*2], **Siyuan Guo**[2,3], **Noah Hollmann**[5], **Frank Hutter**[†4,1,5], **Bernhard Schölkopf**[†2,1]

[1]ELLIS Institute Tübingen, Tübingen, Germany
[2]Max Planck Institute for Intelligent Systems, Tübingen, Germany
[3]University of Cambridge, Cambridge, United Kingdom
[4]University of Freiburg, Freiburg, Germany
s, Freiburg, Germany

NeurIPS 2025 **spotlight**



UNIVERSITY OF CAMBRIDGE

universität freiburg

PRIOR LABS

MAX PLANCK INSTITUTE FOR INTELLIGENT SYSTEMS

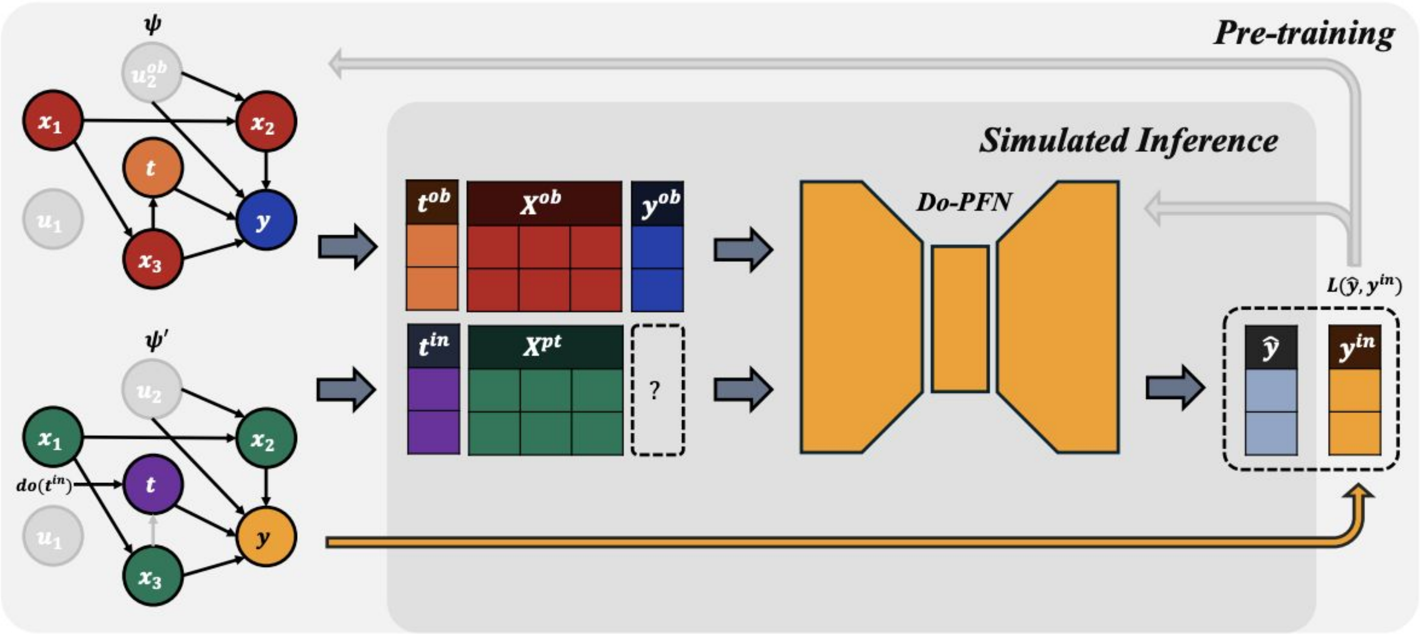


Figure 1: **Do-PFN overview**: Do-PFN performs in-context learning (ICL) for causal effect estimation, predicting conditional interventional distributions (CIDs) based on observational data alone. In pre-training, a large number of structural causal models (SCMs) is sampled. For each SCM, we sample an entire dataset of M^{ob} *observational* data points $\mathcal{D}^{ob} = \{(t_j^{ob}, \mathbf{x}_j^{ob}, y_j^{ob})\}_{j=1}^{M^{ob}}$. We also sample M^{in} *interventional* data points $\mathcal{D}^{in} = \{(t_k^{in}, \mathbf{x}_k^{pt}, y_k^{in})\}_{k=1}^{M^{in}}$. To simulate inference, we input (t^{in}, x^{pt}) along with the entire observational dataset $\mathcal{D}_{ob}$, which can have various sizes and dimensionalities. Subsequently, the transformer makes predictions $\hat{y}$, and we calculate the pre-training loss $L(\hat{y}, y^{in})$ between the predictions $\hat{y}$ and the ground truth interventional outcomes y^{in}. Pre-training repeats this procedure across millions of sampled SCMs to *meta-learn* how to perform causal inference *in context*. In real-world applications, Do-PFN leverages the many simulated interventions it has seen during pre-training to predict CIDs, relying only on observational data and requiring no information about the causal graph.

Synthetic Case Studies.

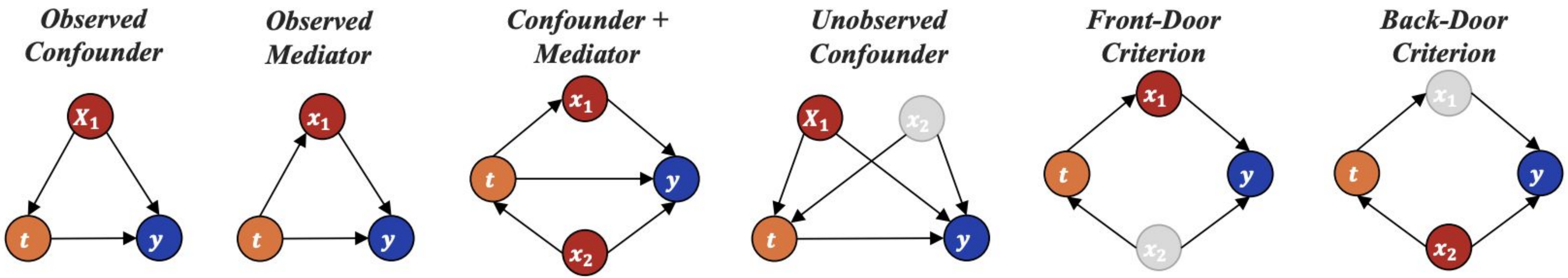
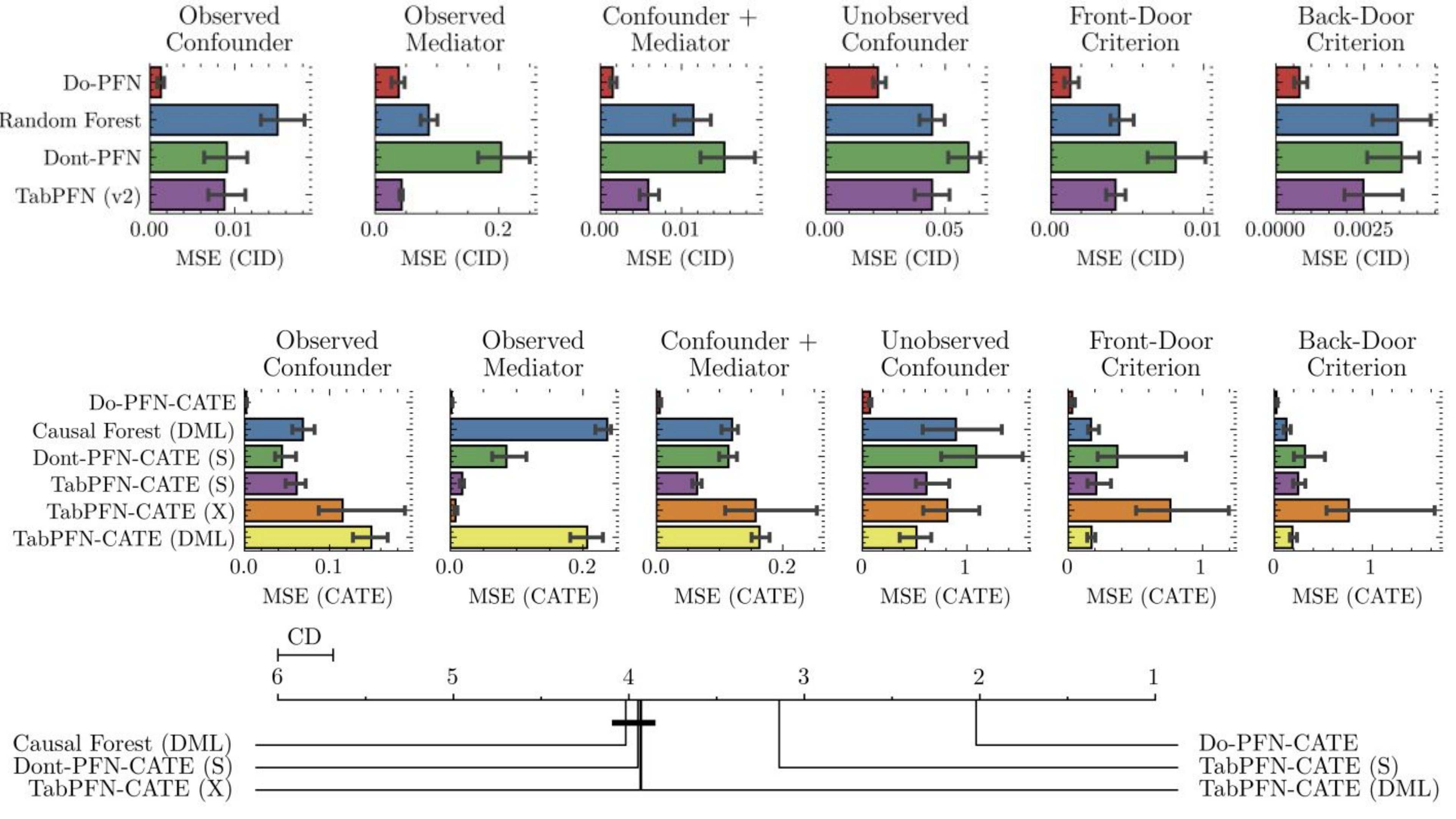


Figure 2: **Case studies**: Visualization of the graph structures of our six causal case studies, requiring Do-PFN to automatically perform adjustment based on the front-door and back-door criteria. Treatment variables t are visualized in orange, covariates $\mathbf{x}$ in red, and outcomes y in blue. Gray variables represent unobservables, not shown to any of the methods yet influencing the generated data.



Observed Confounder
Observed Mediator
Confounder + Mediator
Unobserved Confounder
Front-Door Criterion
Back-Door Criterion
Do-PFN
Random Forest
Dont-PFN
TabPFN (v2)
MSE (CID)
Do-PFN-CATE
Causal Forest (DML)
Dont-PFN-CATE (S)
TabPFN-CATE (S)
TabPFN-CATE (X)
TabPFN-CATE (DML)
MSE (CATE)
CD

Hybrid Synthetic Real-world data

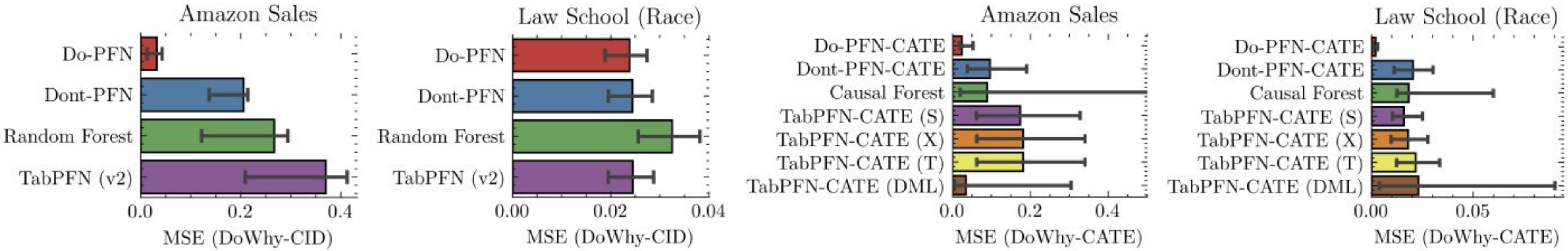
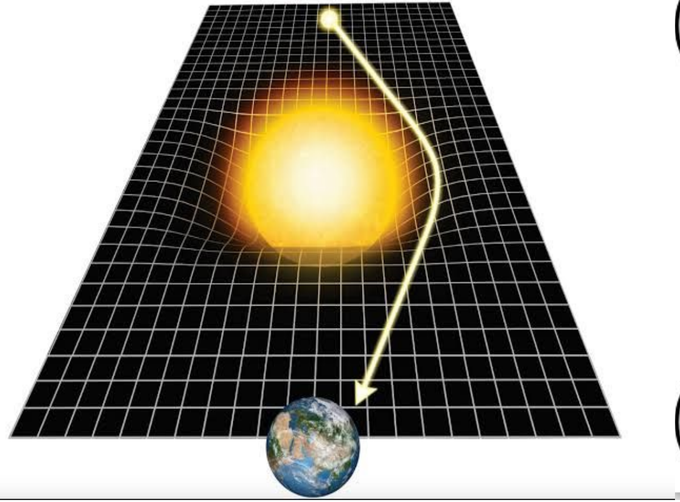
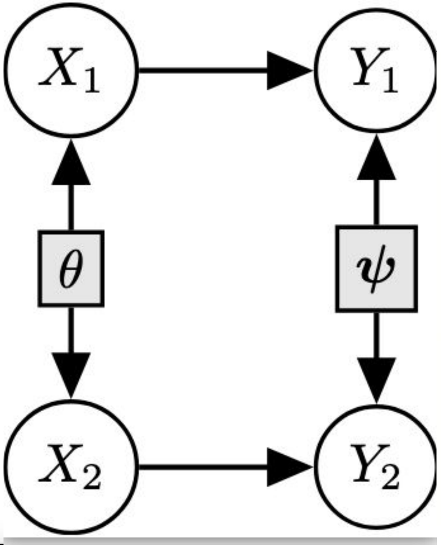


Figure 5: **Hybrid synthetic-real-world data**: Bar-plots with 95% confidence intervals depicting distributions of normalized mean squared error (MSE) of Do-PFN compared to causal and regression baselines in interventional outcome prediction (left) and conditional average treatment effect (CATE) estimation (right). Do-PFN's strong performance in synthetic settings extends to hybrid synthetic-real-world scenarios, especially in CATE estimation.

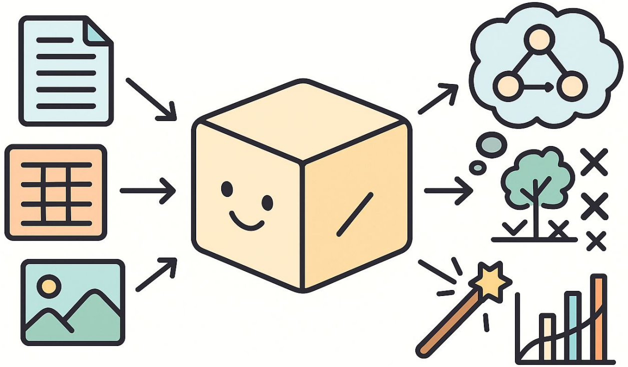
Curious on Intelligence



Learning



Structure



Models

Physics of Learning: A Lagrangian perspective to different learning paradigms. Preprint.
Siyuan Guo and Bernhard Schölkopf

Causal de Finetti: On the Identification of Invariant Causal Structure in Exchangeable Data. NeurIPS 2023.
Siyuan Guo*, Viktor Tóth*, Bernhard Schölkopf, Ferenc Huszár.

Do Finetti: On Causal Effects for Exchangeable Data. NeurIPS 2024 **oral** (acceptance rate 0.46%).
Siyuan Guo, Chi Zhang, Karthika Mohan, Ferenc Huszár, Bernhard Schoelkopf.

Identifiable Exchangeable Mechanisms for Causal Structure and Representation Learning. ICLR 2025 **spotlight** (acceptance rate 5.1%).
Patrik Reizinger*, **Siyuan Guo***, Ferenc Huszár, Bernhard Schölkopf, Wieland Brendel

Counterfactual reasoning: an analysis of in-context emergence. NeurIPS 2025.
Moritz Miller, Bernhard Schölkopf, **Siyuan Guo**

Skill Learning via Policy Diversity Yields Identifiable Representations for Reinforcement Learning. Under Review.
Patrik Reizinger*, Bálint Mucsányi*, **Siyuan Guo***, Benjamin Eysenbach, Bernhard Schölkopf, Wieland Brendel

On the Emergence and Test-Time Use of Structural Information in Large Language Models. Under Review.
Michelle Chao Chen, Moritz Miller, Bernhard Schölkopf, **Siyuan Guo**

Do-PFN: In-Context Learning for Causal Effect Estimation. NeurIPS 2025 **spotlight** (acceptance rate 3.19%).
Jake Robertson*, Arik Reuter*, **Siyuan Guo**, Noah Hollmann, Frank Hutter, Bernhard Schölkopf

RESEARCH

UNIVERSITY OF CAMBRIDGE
DEPARTMENT OF
COMPUTER SCIENCE AND TECHNOLOGY