MISO 2026 · MAISI

Solomonoff Induction & AIXI

Cole Wyeth, AIXI Labs

Highly indebted to Aram Ebtekar, Marcus Hutter, and Danica Sutherland for ideas, and the Iliad Intensive Curriculum for some content.

Classical statistics is insufficient to understand general intelligence

The Bitter Lesson; Sutton, 2019 : Hand-picked restrictive inductive biases cap learning algorithms at the level of the designer's formalized domain knowledge.

Ebtekar et al., 2026 : Even "expert knowledge" must be learned, so it is not a satisfying explanation for successful generalization (the meta-no-free-lunch argument).

> Superintelligence will (or already does) discover new, better theories contrary to expert knowledge.
>
> Consider hierarchical or learned priors in Bayesian statistics and program induction, domain-specific hypothesis classes with low Vapnik-Chervonenkis dimension, and the art of "feature engineering" in ML which all gave way to more flexible (neural) architectures.

Scaling and computational universality

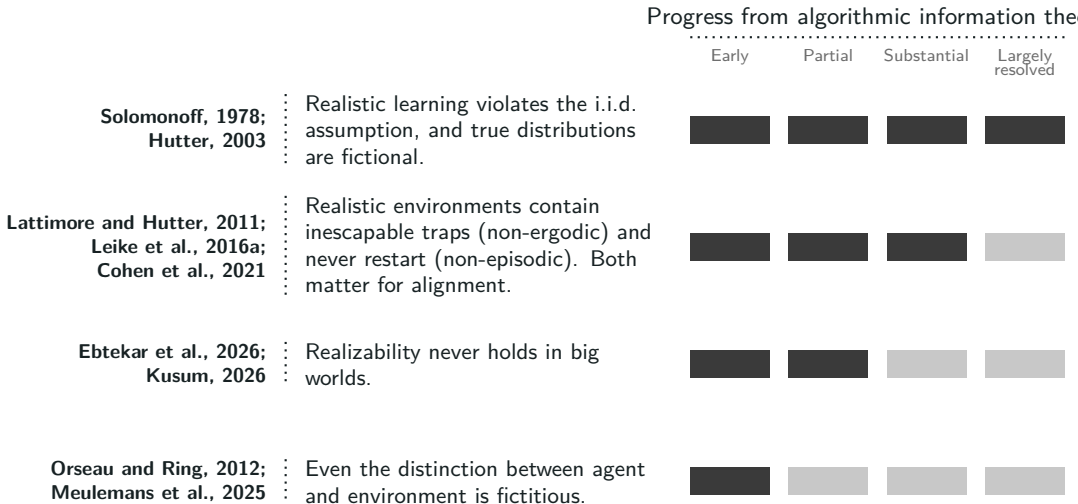
> Taking AGI seriously means recognizing that the scale of intelligence extends far above the human level.

Ebtekar et al., 2026: We need to go beyond a fixed, finitely computable theory of intelligence. A theory should scale with compute, inventing qualitatively new hypotheses as they come within reach.

Big Blob of Compute; Amodei, 2017: Empirically, the most successful ML research serves to minimize the constraints on learned solutions while flexibly leveraging compute.

Even though no computable agent is the best, **computational universality** gives a (rough) direction of improvement under the Church–Turing thesis.

Common tractability assumptions are inappropriate for modeling AGI (alignment)



Intro summary

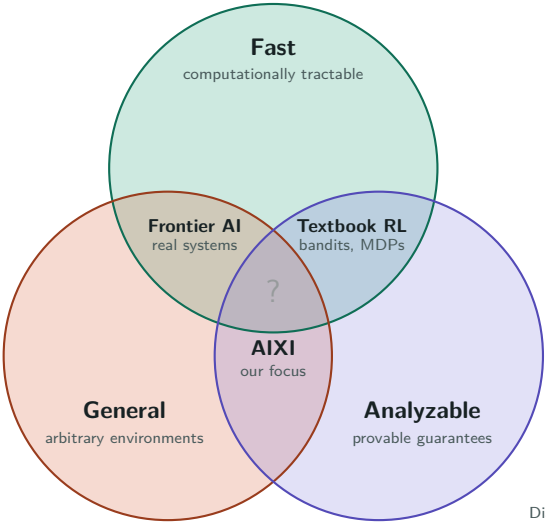


Diagram: aixi.uk

Table of contents

Bayesian prediction

Notation and the Bayesian mixture

Log-loss bound on every sequence

Expected loss bounds

Merging of opinions

Misspecification

Admissibility and Pareto optimality

Solomonoff induction

Definition

Occam's razor

Robustness of the definition

Non-probabilistic loss bound

Accessibility justification

Choice of UTM

Anytime predictors

Merging analogues

Philosophy papers

Bayesian RL

Setup and definitions

Pareto optimality and self-optimization

AIXI

Definition of AIXI

Incomputability

Not worst possible in deterministic environments

Asymptotic optimality

Curiosity killed the cat

Exploration with traps

Safe exploration

Stopping complexity

AIXI Labs · **References**

Bayesian prediction: notation

Symbols used throughout.

- $\mathbb{B} = \{0, 1\}$: the binary alphabet; $\mathbb{B}^*$ all finite binary strings; $\mathbb{B}^n$ length-n strings.
- ϵ: the empty string.
- xy: concatenation. If $x := x_{1:n}$ and $y := y_{1:m}$, then $xy = x_1 \dots x_n y_1 \dots y_m$.
- $x_{i:j} := x_i x_{i+1} \dots x_j$.
- $x_{<t} := x_1 x_2 \cdots x_{t-1}$.
- $\lg := \log_2$.
- ν, ρ: generic environments / predictors. μ: the true (unknown) environment. ξ: a Bayesian mixture of environments.
- $\mathcal{M} = \{\nu_1, \nu_2, \dots\}$: a countable class of candidate environments.
- $w_\nu > 0$ with $\sum_{\nu \in \mathcal{M}} w_\nu = 1$: the prior over $\mathcal{M}$. $w(\nu \mid x_{<t})$: the posterior after observing $x_{<t}$.
- $\Delta\mathcal{X}$: the probability distributions over a finite set $\mathcal{X}$.

[Hutter, 2005, Ch. 3; Hutter et al., 2024, Ch. 3; Quarel and Lang, n.d.]

Environments

- An unknown environment μ generates $x_1 x_2 \ldots$, with $x_t \sim \mu(\cdot \mid x_{<t})$. A predictor ρ sees $x_{<t}$ and outputs a distribution $\rho(\cdot \mid x_{<t})$ over the next symbol.

Definition (Environment). An **environment** ν assigns to each history $x_{<t} \in \mathbb{B}^*$ a predictive distribution $\nu(\cdot \mid x_{<t}) \in \Delta\mathbb{B}$ over the next symbol. The conditional joint probability of a string $x \in \mathbb{B}^*$, conditioned on $y \in \mathbb{B}^*$, is defined as

$$\nu(x_{1:n} \mid y) := \prod_{t=1}^{n} \nu(x_t \mid y\, x_{<t}), \qquad \nu(\epsilon) := 1.$$

The unconditional joint is defined as $\nu(x_{1:n}) := \nu(x_{1:n} \mid \epsilon)$. The true (unknown) environment generating the data is denoted μ.

- **Remark:** $\nu(x_{1:t}) = \nu(x_{<t})\,\nu(x_t \mid x_{<t})$ and $\nu(x) = \nu(x0) + \nu(x1)$ for all $x \in \mathbb{B}^*$.

[Quarel and Lang, n.d.]

Definition (Bayesian mixture ξ). Let $\mathcal{M} = \{\nu_1, \nu_2, \ldots\}$ be a countable class of environments with prior weights $w_\nu > 0$ satisfying $\sum_{\nu \in \mathcal{M}} w_\nu = 1$. The **Bayesian mixture** is the prior-weighted mixture over all environments ν in $\mathcal{M}$:

$$\xi(x_{1:t}) \;:=\; \sum_{\nu \in \mathcal{M}} w_\nu \, \nu(x_{1:t}).$$

Its one-step predictive distribution and posterior weights are

$$\xi(x_t \mid x_{<t}) \;=\; \sum_{\nu \in \mathcal{M}} w(\nu \mid x_{<t}) \, \nu(x_t \mid x_{<t}), \qquad w(\nu \mid x_{<t}) \;:=\; w_\nu \, \frac{\nu(x_{<t})}{\xi(x_{<t})}.$$

- **Multiplicative update:** $w(\nu \mid x_{1:t}) = w(\nu \mid x_{<t}) \cdot \nu(x_t \mid x_{<t}) / \xi(x_t \mid x_{<t})$. The same posterior governs multi-step predictions $\xi(x_{t:m} \mid x_{<t})$.
- **Dominance:** $\xi(x) \geq w_\nu \, \nu(x)$ for every $\nu \in \mathcal{M}$ and $x \in \mathbb{B}^*$. Every bound below rests on this.

[Hutter, 2005, Ch. 3; Quarel and Lang, n.d.]

Definition (Log loss). The **log loss** of a predictor ν on $x = x_{1:n} \in \mathbb{B}^n$ is

$$L(\nu, x) \;:=\; \sum_{t=1}^{n} \lg \frac{1}{\nu(x_t \mid x_{<t})} \;=\; \lg \frac{1}{\nu(x)}.$$

Theorem (log-loss regret bound)

For every n, every sequence $x \in \mathbb{B}^n$, and every $\nu \in \mathcal{M}$,

$$L(\xi, x) \;=\; \lg \frac{1}{\xi(x)} \;\overset{(*)}{\leq}\; \lg \frac{1}{\nu(x)} + \lg \frac{1}{w_\nu} \;=\; L(\nu, x) + \lg \frac{1}{w_\nu}.$$

Proof. $(*)$ follows by taking $-\lg$ of dominance, $\xi(x) \geq w_\nu\, \nu(x)$. The equalities are the definition of L.

Remark. The bound holds on every sequence, with no assumption of a true measure $\mu \in \mathcal{M}$.

[Hutter, 2005, Ch. 3]

Mistake bound on a single sequence

Mistake bound. Predict $\hat{x}_t := \arg\max_x \xi(x \mid x_{<t})$. A mistake means $\xi(x_t \mid x_{<t}) \leq \frac{1}{2}$, so it costs at least one bit of log loss: $\lg(1/\xi(x_t \mid x_{<t})) \geq 1$. By the log-loss bound, on *every* sequence,

$$\#\{t \leq n : \hat{x}_t \neq x_t\} \;\leq\; L(\xi, x_{1:n}) \;\leq\; \min_{\nu \in \mathcal{M}} \Big[L(\nu, x_{1:n}) + \lg \frac{1}{w_\nu} \Big].$$

Deterministic hypotheses. If a deterministic $\nu \in \mathcal{M}$ outputs the sequence, so $\nu(x_{1:n}) = 1$ for all n, then ξ makes at most $\lg(1/w_\nu)$ mistakes *ever*:

$$\#\{t \geq 1 : \hat{x}_t \neq x_t\} \;\leq\; \lg \frac{1}{w_\nu}.$$

[Hutter, 2005, Ch. 3]

Measuring prediction error: KL divergence

Definition (Cumulative and instantaneous KL). For environments / predictors ν, ρ, horizon $n \geq 1$, and history $x_{<t}$:

$$D_n(\nu \,\|\, \rho) := \sum_{x_{1:n} \in \mathbb{B}^n} \nu(x_{1:n}) \ln \frac{\nu(x_{1:n})}{\rho(x_{1:n})},$$

$$d_t(x_{<t}) := \sum_{x_t \in \mathbb{B}} \nu(x_t \mid x_{<t}) \ln \frac{\nu(x_t \mid x_{<t})}{\rho(x_t \mid x_{<t})} \qquad (\nu, \rho \text{ understood}),$$

with $0 \ln \frac{0}{q} := 0$ and $p \ln \frac{p}{0} := \infty$ for $p > 0$.
Write $D_\infty := \lim_{n \to \infty} D_n$ and $D_\infty^\mu := D_\infty(\mu \,\|\, \xi)$.

Chain rule: the joint KL telescopes into expected instantaneous KLs, $D_n(\nu \,\|\, \rho) = \sum_{t=1}^{n} \mathbb{E}_{x_{<t} \sim \nu} \, d_t(x_{<t})$. Each $d_t \geq 0$, so D_∞ exists in $[0, \infty]$.

[Cover and Thomas, 2006; Quarel and Lang, n.d.]

Measuring prediction error: Hellinger distance

Definition (Instantaneous and cumulative squared Hellinger).

$$h_t(x_{<t}) := \sum_{x_t \in \mathbb{B}} \Big(\sqrt{\nu(x_t \mid x_{<t})} - \sqrt{\rho(x_t \mid x_{<t})} \Big)^2 \qquad (\nu, \rho \text{ understood}),$$
$$H_n^2(\nu \,\|\, \rho) := \sum_{t=1}^{n} \mathbb{E}_{x_{<t} \sim \nu}\, h_t(x_{<t}), \qquad H_\infty^2 := \lim_{n\to\infty} H_n^2.$$

Squared Hellinger $\leq$ KL: $h_t(x_{<t}) \leq d_t(x_{<t})$ on every history, so $H_n^2 \leq D_n$. For distributions p, q, by (1) Jensen on $-2\ln(q_x/p_x)^{1/2}$ and (2) $-\ln y \geq 1 - y$:

$$\sum_x p_x \ln \tfrac{p_x}{q_x} \overset{(1)}{\geq} -2 \ln \sum_x \sqrt{p_x q_x} \overset{(2)}{\geq} 2\big(1 - \textstyle\sum_x \sqrt{p_x q_x}\big) \;=\; \sum_x (\sqrt{p_x} - \sqrt{q_x})^2.$$

Why not squared error? $\sum_x (p_x - q_x)^2$ fails data processing: $p = (\frac{1}{2}, \frac{1}{2}, 0)$, $q = (0, 0, 1)$ give $\frac{3}{2}$, but merging the first two outcomes gives 2. Squared Hellinger is an f-divergence: the squared chord between $\sqrt{p}$ and $\sqrt{q}$ on the unit sphere.

[Hutter, 2005, Ch. 3]

Theorem (cumulative expected bounds), for $\mu \in \mathcal{M}$

$$H_\infty^2(\mu \,\|\, \xi) \overset{\text{Jensen}}{\leq} D_\infty(\mu \,\|\, \xi) \;\leq\; \ln \frac{1}{w_\mu}.$$

Corollary of the log-loss bound: set $\nu = \mu$ and average over $x_{1:n} \sim \mu$: $D_n(\mu \,\|\, \xi) = \mathbb{E}_\mu\big[\ln \frac{\mu(x_{1:n})}{\xi(x_{1:n})}\big] \leq \ln \frac{1}{w_\mu}$. The true environment enters only through this average.

The expected total is finite, so μ-almost surely $\sum_t \sum_x \big(\sqrt{\mu(x \mid x_{<t})} - \sqrt{\xi(x \mid x_{<t})}\big)^2 < \infty$, and $\xi(\cdot \mid x_{<t}) \to \mu(\cdot \mid x_{<t})$ with μ-probability 1.

No i.i.d., stationarity, or ergodicity assumption, only realizability: $\mu \in \mathcal{M}$.

[Hutter, 2005, Ch. 3; Quarel and Lang, n.d.]

Theorem (Blackwell–Dubins)

Let P, Q be probability measures on $\mathbb{B}^\infty$ with $P \ll Q$.

Then P-almost surely, $\sup_A \left| P(A \mid x_{<t}) - Q(A \mid x_{<t}) \right| \to 0$ as $t \to \infty$, where A ranges over all measurable sets of infinite continuations.

- Dominance $\xi \geq w_\mu \mu$ gives $\mu \ll \xi$, with density at most $1/w_\mu$. So ξ merges with μ: its forecasts of *every* future event converge to the truth, not just its next-symbol forecasts.
- This will matter when we discuss learning from interaction, because general-purpose agents must be able to optimize a wide variety of goals (e.g. not only log loss).
- Finite KL divergence implies merging, but not conversely.

[Blackwell and Dubins, 1962; Cover and Thomas, 2006, Lemma 11.6.1]

Theorem (misspecified bound)

For every $\hat{\mu} \in \mathcal{M}$, $\quad D_n(\mu \,\|\, \xi) \;\leq\; -\ln w_{\hat{\mu}} \;+\; D_n(\mu \,\|\, \hat{\mu})$.

Proof: the log-loss bound with $\nu = \hat{\mu}$, plus $\ln \mu(x_{1:n})$ on both sides, averaged over $x_{1:n} \sim \mu$.

- $-\ln w_{\hat{\mu}}$ is constant in n: the cost of not knowing which model is best.
- $D_n(\mu \,\|\, \hat{\mu}) = \sum_{t \leq n} \mathbb{E}_\mu d_t(x_{<t})$ generally grows linearly in n: the cost of no model being right. So ξ is only as good as the best model in the class, and $D_n(\mu \,\|\, \xi)$ diverges in general.
- The bound holds for every $\hat{\mu}$, so take the infimum. Which $\hat{\mu}$ attains it depends on the horizon and the prior.

[Hutter, 2005, §3.2.8; Quarel and Lang, n.d.]

Definition (Pareto optimality). Fix a horizon n and a loss $\mathcal{F}(\nu, \rho) \geq 0$ of predictor ρ when the environment is ν. A predictor ρ' *Pareto-dominates* ρ if

$$\mathcal{F}(\nu, \rho') \leq \mathcal{F}(\nu, \rho) \text{ for all } \nu \in \mathcal{M}, \quad \text{with strict inequality for some } \nu \in \mathcal{M}.$$

ρ is *Pareto optimal* if no predictor Pareto-dominates it.

Losses: the KL divergence D_n, and the squared error

$$S_n(\nu \,\|\, \rho) := \sum_{t=1}^{n} \mathbb{E}_{x_{<t} \sim \nu} \sum_{x_t \in \mathbb{B}} \big(\nu(x_t \mid x_{<t}) - \rho(x_t \mid x_{<t})\big)^2.$$

Theorem (Pareto optimality of ξ)

If $w_\nu > 0$ for every $\nu \in \mathcal{M}$ (full support), then the Bayesian mixture ξ is Pareto optimal under D_n and under S_n, for every n.

[Hutter, 2003; Hutter, 2005, Ch. 3; Quarel and Lang, n.d.]

Admissibility and full support

General Bayesian predictors minimize the (subjective) expected value of the loss.

Assuming their priors have full support, they are Pareto optimal under very general assumptions. However, not all losses incentivize the mixture prediction. In statistics, **Pareto optimality is called admissibility**.

Ferguson, 1967 : Complete class theorems by Wald establish that in finite hypothesis spaces, all admissible predictors are Bayesian.

Ryabko, 2019 : The situation is subtler in infinite hypothesis spaces, but a similar result holds asymptotically.

Notice that these optimality properties assume realizability.

Ryabko, 2019 : "A very general recommendation for choosing a model can be made based on these results: it is better to take a model large enough to make sure it includes the process that generates the data, even if it entails positive asymptotic average loss, for otherwise any combination of predictors in the model class may be useless."

[Wald, 1950; Ferguson, 1967; Ryabko, 2019]

Solomonoff induction: Intro

Solomonoff induction takes Ryabko's suggestion very seriously, by adding every possible predictor to the hypothesis class.

This bolsters the case for Bayesian prediction by explaining how to ensure full prior support (for admissibility).

Ebtekar et al., 2026; Schurz, 2019; Sterkenburg, 2018, 2019b

Specifically, Solomonoff induction attempts to include (and outperform) every prediction method that we "could have used".

Under the **Church–Turing thesis**, the accessible predictors are the computable ones.

Solomonoff induction upper-bounds the performance of any machine learning system, and provides a widely applicable **simplified model** of superintelligent sequence prediction.

Sources: Solomonoff, 1964; Solomonoff, 1978; Hutter, 2005

Solomonoff induction: Definition

Definition (Semimeasure). A **semimeasure** is a function $\nu : \mathbb{B}^* \to [0,1]$ with $\nu(\epsilon) \leq 1$ and $\nu(x) \geq \nu(x0) + \nu(x1)$ for all $x \in \mathbb{B}^*$. It is **lower semicomputable** (l.s.c.) if $\nu(x) = \lim_k \phi(x,k)$ for a computable ϕ nondecreasing in k. $\mathcal{M}_U$: all l.s.c. semimeasures.

Definition (Universal mixture). Let $w_\nu > 0$ be an l.s.c. prior on $\mathcal{M}_U$ with $\sum_{\nu \in \mathcal{M}_U} w_\nu \leq 1$. Then $\xi_U(x) := \sum_{\nu \in \mathcal{M}_U} w_\nu \, \nu(x)$, and **Solomonoff induction** predicts with $\xi_U(x_t \mid x_{<t}) := \xi_U(x_{1:t}) / \xi_U(x_{<t})$.

Why semimeasures: unlike the computable measures, $\mathcal{M}_U$ is effectively enumerable, and $\xi_U \in \mathcal{M}_U$. The missing mass $\nu(x) - \nu(x0) - \nu(x1)$ is the probability of no further output.

Definition (Kolmogorov complexity). For a universal prefix-free machine U, $K(x) := \min\{\ell(p) : U(p) = x\}$. For an effective enumeration (ν_j) of $\mathcal{M}_U$, $K(\nu) := \min_{j:\, \nu_j = \nu} K(j)$. Conventionally, $w_\nu := 2^{-K(\nu)}$.

[Solomonoff, 1964; Hutter, 2005, Ch. 2]

Inevitability of Occam's razor

The 2^{-K}-mixture ξ' is l.s.c., so $\xi' = \nu_i$ for some i, inside *any* universal mixture ξ_U:

$$\xi_U \;=\; w_1\,\nu_1 \;+\; w_2\,\nu_2 \;+\; \cdots \;+\; \boxed{w_i\,\xi'} \;+\; \cdots$$

$$\xi' \;:=\; 2^{-K(1)}\,\nu_1 \;+\; \cdots \;+\; \underbrace{\boxed{2^{-K(j)}\,\nu_j}}_{\nu_j = \nu,\;\; K(j) = K(\nu)} \;+\; \cdots$$

So $\xi_U(x) \geq w_i\,\xi'(x) \geq 2^{-K(\nu)-\lg(1/w_i)}\,\nu(x)$: a penalty of $K(\nu) + O(1)$ bits, the same $O(1)$ slack as changing U in K.

Coding theorem (Li and Vitányi, 2019, Thm. 4.3.3) : 2^{-K} is the slowest-decaying l.s.c. prior. For every l.s.c. w with $\sum_j w_j \leq 1$, $w_j \leq c_w\,2^{-K(j)}$.

OPEN PROBLEM

Each ν occurs at infinitely many indices j, and $K(\nu) := \min_{j:\,\nu_j=\nu} K(j)$. Can its *total* weight $\sum_{j:\,\nu_j=\nu} 2^{-K(j)}$ exceed $2^{-K(\nu)}$ by more than a constant factor? The coding-theorem proof would need to recognize the indices of ν, which is generally not computable.

[Li and Vitányi, 2019, Ch. 4; Lang, 2025]

Robustness of the definition

A **probabilistic Turing machine** T reads fair coin flips and writes output bits one at a time, never erasing them. Its output distribution is

$$\lambda_T(x) \;:=\; \Pr\big[\, T \text{ outputs a string beginning with } x \,\big].$$

Equivalently: $\mathcal{M}_U = \{\lambda_T : T \text{ a probabilistic Turing machine}\}$.

Definition (Solomonoff prior M). Fix a universal monotone machine U; write $U(p) = x*$ if U on input p outputs a string beginning with x. Summing over minimal such p,

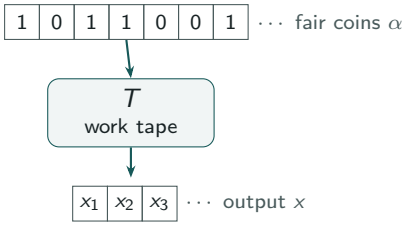
$$M(x) \;:=\; \lambda_U(x) \;=\; \sum\nolimits_{p\,:\,U(p)=x*} 2^{-\ell(p)}.$$

- $M \in \mathcal{M}_U$, and M and ξ_U dominate each other up to multiplicative constants. Changing U also changes M only by a multiplicative constant.
- In fact, Solomonoff priors and universal mixtures (normalized at ϵ) define *identical* classes of semimeasures. The universally dominant semimeasures form a strictly larger class.

[Li and Vitányi, 2019; Hutter, 2005, Ch. 2; Wood et al., 2013]

Probabilistic Turing machines

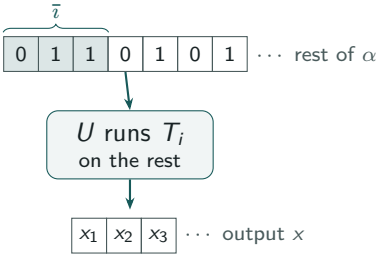
A probabilistic machine T



read-only & one-way in, write-only & one-way out

$$\lambda_T(x) = \Pr_\alpha\left[T(\alpha) \text{ begins with } x\right]$$

The universal machine U



α begins with the code $\bar{\imath}$ of T_i w.p. $2^{-|\bar{\imath}|}$

$$M(x) = \lambda_U(x) = \sum_i 2^{-|\bar{\imath}|}\,\lambda_{T_i}(x)$$

So M is itself a universal mixture: the λ_{T_i} range over all of $\mathcal{M}_U$, with prior weights $2^{-|\bar{\imath}|}$.

[Ebtekar et al., 2026; Li and Vitányi, 2019]

Optimality and the non-probabilistic loss bound

Definition (Information vantage point). Fix a universal monotone machine U with an **oracle** tape $q \in \mathbb{B}^\infty$ (e.g. a snapshot of the Web) and a program tape. With $\alpha \in \mathbb{B}^\infty$ uniformly random and $U(p) = x*$ as before,

$$\nu_p(x) := \Pr_\alpha\big[U^\alpha(p) = x*\big], \qquad M^q(x) := \Pr_\alpha\big[U^q(\alpha) = x*\big],$$
$$KM^q(x) := \lg \tfrac{1}{M^q(x)}, \qquad MI^q(p : x) := KM^q(x) - KM^{p\oplus q}(x).$$

ν_p is the semimeasure *programmed* by p; M^q is **oracle Solomonoff induction**.

Theorem. For all $p, q \in \mathbb{B}^\infty$ and *every* $x \in \mathbb{B}^*$,

$$L(M^q, x) - L(\nu_p, x) \overset{+}{\leq} MI^q(p : x) \overset{+}{\leq} KM^q(p).$$

No true environment, no realizability: the regret is the information p carries about x that q lacks. A random seed in p makes $KM^q(p) = \infty$, but leaves $MI^q(p : x)$ small.

[Ebtekar et al., 2026]

The accessibility justification

We need not beat a clairvoyant predictor that hardcodes the test data, only methods **accessible** to us.

- A short program r, run from our vantage point q with random bits β, outputs a predictor $p := U^q(r\beta)$.
- From a Bash shell with Web access, r can be "download a pretrained model and fine-tune it". Any other computer is an interpreter inside q, so the choice of U stops mattering.

Theorem (No information ex nihilo). For every $x \in \mathbb{B}^*$,

$$\mathbb{E}_p\left[Ml^q(p:x)\right] \;\leq\; \lg \mathbb{E}_p\left[2^{Ml^q(p:x)}\right] \;\overset{+}{\leq}\; |r|.$$

- By Markov, $\Pr_p\left[Ml^q(p:x) \geq |r| + c + k\right] \leq 2^{-k}$: *whatever* x is, an accessible process rarely beats M^q by much more than $|r|$ bits.
- Under the physical Church–Turing thesis, the engineer is such a process too, so M^q upper-bounds every machine learning system we could build.

[Ebtekar et al., 2026]

Choice of UTM

The optimality proofs compile.

Every $\overset{+}{\leq}$ above hides the code length $|\overline{i}|$ of a few fixed, explicit machines T_i: copy the program to the output, ignore an oracle, swap program and oracle, read a self-delimiting prefix, ...

So we can **pick** U to give these machines short codes, making every constant small.

Contrast the invariance theorem.

Changing U changes K by $O(1)$, but that constant is uncontrolled, and a badly chosen U can make a universal prior behave arbitrarily badly on any finite horizon.

No commitment to one computer is needed: any other computer we have access to is an interpreter stored in the oracle q, a short pointer away.

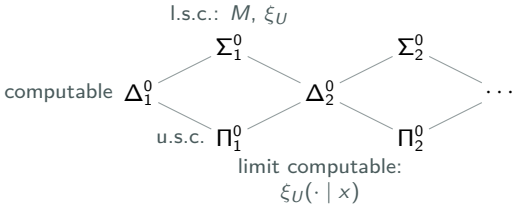
[Ebtekar et al., 2026; Li and Vitányi, 2019, Sec. 3.9; Leike and Hutter, 2015; Almond, 2005]

Anytime predictors

Diagonalization. SI's predictions $\xi_U(\cdot \mid x_{<t})$ are limit computable, so a sequence that always takes a bit of conditional probability $\leq \frac{1}{2}$ can be chosen limit computable too, and on it $\xi_U(x_{1:n}) \leq 2^{-n}$. The same diagonal argument defeats *every* limit computable predictor: none is universal.

Response. Run an anytime algorithm p for t steps: the result $\nu_{p,t} = \nu_{rp}$ is computable, with r encoding t, so $L(M^q, x) - L(\nu_{p,t}, x) \overset{+}{\leq} |r| + MI^q(p : x)$. With t the kth busy beaver number, $|r| \approx k$: k extra bits of information cost $\mathrm{BB}(k)$ steps, e.g. waiting for the slowest programs to halt in order to diagonalize SI.

Other universal predictors are still worth exploring, e.g. it is (nontrivially) possible to mix all computable predictors, avoiding ξ_U's semimeasure loss, while remaining within Δ_2^0 (Wyeth et al., unpublished).

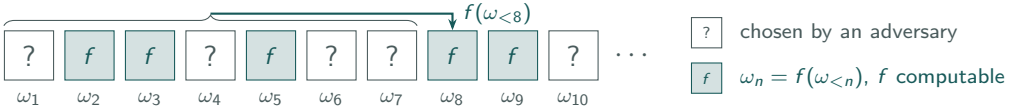


[Sterkenburg, 2019a; Ebtekar et al., 2026, App. A.2]

Prediction of selected bits

Selected bits

Without a true measure, we can still ask the predictor to learn every computable pattern in the data. SI does not always succeed on adversarial sequences:



Lattimore et al., 2011

Some (uncomputable) ω has $\omega_{2n} = \omega_{2n-1}$ for all n, yet $\liminf_n M(\omega_{2n} \mid \omega_{<2n}) < 1$. Normalized M learns deterministic sub-patterns, but whether it learns stochastic ones remains open.

Rates

Even after normalizing, convergence on the selected bits is slow.

Bauwens and Kuznetsov, 2026

For some computable pattern, the total absolute error $\sum_n M_{\mathrm{norm}}(\bar{\omega}_n \mid \omega_{<n})$ on the bits it fixes is unbounded. The sum of squared errors is bounded, but only exponentially in the pattern's program length.

Partial models

The repeated-bit pattern is a *partial model*: it says nothing about the odd bits.

We may want to beat every such partial model.

Fallenstein et al., 2015a,b; Leike et al., 2016b; Wyeth et al., 2025	⋮	Reflective oracles give a universal, Bayesian approach: a mixture over reflective-oracle-computable environments
Kusum, 2026	:	...which include partial models.
see also: Meulemans et al., 2025	:	Further game theoretic applications.

Other approaches

- Defensive forecasting (Vovk et al., 2005).
- Forecasting with incomplete models (Kosoy, 2017).
- Logical induction (Garrabrant et al., 2016).

More bounds

Below, $\overset{+}{\leq}$ and $\overset{\times}{\leq}$ hide additive and multiplicative constants, μ is a computable measure, and h_t is the Hellinger distance between $M(\cdot \mid x_{<t})$ and $\mu(\cdot \mid x_{<t})$.

- **Future bounds.** If the past $x_{1:n}$ says a lot about μ, few errors are still to come: $\sum_{t>n} \mathbb{E}\big[h_t \mid x_{1:n}\big] \overset{+}{\leq} \big[K(\mu \mid x_{1:n}) + K(n)\big] \ln 2$.
- **i.i.d. environments.** For a mixture ξ over a class $\mathcal{M}$ of i.i.d. measures, the error of ξ at each step decays like $1/n$: $\mathbb{E}[h_n] \overset{\times}{\leq} \frac{1}{n} \ln \frac{1}{w_\mu}$, i.e. $\frac{1}{n} K(\mu) \ln 2$ for $w_\mu = 2^{-K(\mu)}$.
- **Weak instantaneous bounds.** With monotone complexity $Km(x) := \min\{\ell(p) : U(p) = x*\}$, for every n, every $x_{1:n}$, and $\bar{x}_n \neq x_n$,

$$2^{-K(n)} \overset{\times}{\leq} M(\bar{x}_n \mid x_{<n}) \overset{\times}{\leq} 2^{2Km(x_{1:n}) - K(n)}.$$

- **Magic instance numbers.** $M(0 \mid 1^n) \overset{\times}{=} 2^{-K(n)} \to 0$, but it spikes up at simple n: M is cautious at "magic" instance numbers.

[Hutter, n.d., slide 164]

Incomputability means Solomonoff induction is useless?

A confused objection.

Any reasonable model of superintelligent prediction must be parameterized by available computational resources, so it should be an anytime algorithm lying in Δ_2^0. Additional compute always helps!

Scale.

If someone claims to have a theory of AGI that does not require massive scale (e.g. stacking GPUs), one wonders why they are not much richer than Anthropic ($\sim$\$1.5 trillion).

Efficient AGI may exist.

It is very likely that sufficient understanding of cognition, or particular facts about our universe, would in principle allow AGI implementations much more efficient than modern ML, e.g. some AGI or even superintelligence that runs on an edge device. However, it is less clear whether a human could actually write down the code for such an AGI with non-vanishing probability.

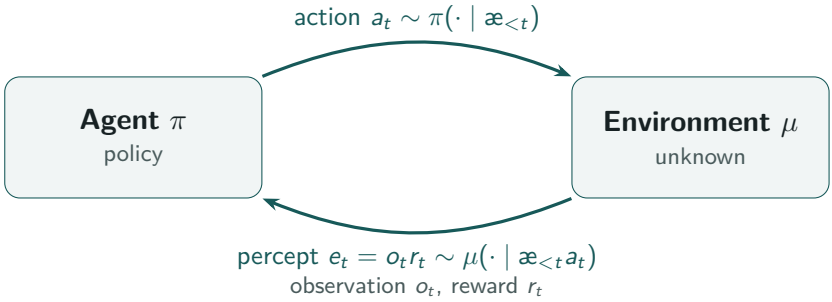
Bounded versions.

Finally, it is not hard to imagine computationally bounded versions of this theory, though the ones invented so far have been somewhat clumsy or inelegant. Our experience suggests that the development of bounded models should be guided by concrete scientific questions.

Philosophical library on Solomonoff induction

- **Putnam (1963):** *"Degree of confirmation" and inductive logic.*
- **Hutter (2007):** *On universal prediction and Bayesian confirmation.*
- **Rathmanner and Hutter (2011):** *A philosophical treatise of universal induction.*
- **Sterkenburg (2016):** *Solomonoff prediction and Occam's razor.*
- **Sterkenburg (2018):** *Universal prediction: a philosophical investigation.*
- **Sterkenburg (2019a):** *Putnam's diagonal argument and the impossibility of a universal learning machine.*
- **Sterkenburg (2019b):** *The metainductive justification of induction: the pool of strategies.*
- **Schurz (2019):** *Hume's problem solved: the optimality of meta-induction.*
- **Neth (2023):** *A dilemma for Solomonoff prediction.*
- **Ebtekar et al. (2026):** *Understanding generalization requires universal induction.*
- **Sterkenburg (2026):** *Solomonoff induction.*

Bayesian reinforcement learning: Setup



history $æ_{<t}$: | a_1 | e_1 | a_2 | e_2 | a_3 | e_3 | $\cdots$ no resets, no assumed ergodicity

The agent's goal is to maximize (discounted) future reward, learning about μ only through this one interaction history.

[Hutter, 2005, Ch. 4]

Finite **actions** $\mathcal{A}$, **observations** $\mathcal{O}$, **rewards** $\mathcal{R} \subset [0,1]$; **percepts** $e_t = o_t r_t \in \mathcal{E} := \mathcal{O} \times \mathcal{R}$; **histories** $æ_{<t} := a_1 e_1 \cdots a_{t-1} e_{t-1}$.

Definition (Policy, environment). A **policy** $\pi : \mathcal{H} \to \Delta\mathcal{A}$ maps each history to a distribution over actions. An **environment** $\nu : \mathcal{H} \times \mathcal{A} \to \Delta\mathcal{E}$ maps each history–action pair to a distribution over percepts, with $\nu(æ_{1:t}) := \prod_{k=1}^{t} \nu(e_k \mid æ_{<k} a_k)$. Together they induce the **interaction measure** $\nu^\pi(æ_{i:j} \mid æ_{<i}) := \prod_{k=i}^{j} \pi(a_k \mid æ_{<k})\, \nu(e_k \mid æ_{<k} a_k)$.

Definition (Value). For summable discounts $\gamma_k \geq 0$ and $\Gamma_t := \sum_{k \geq t} \gamma_k$:

$$V_\nu^\pi(æ_{<t}) := \frac{1}{\Gamma_t} \mathbb{E}_\nu^\pi \Big[\sum\nolimits_{k \geq t} \gamma_k r_k \;\Big|\; æ_{<t} \Big], \qquad \xi(æ_{1:t}) := \sum\nolimits_{\nu \in \mathcal{M}} w_\nu\, \nu(æ_{1:t}).$$

$V_\nu^* := \sup_\pi V_\nu^\pi$, and the **Bayes-optimal policy** is $\pi_\xi^* \in \arg\max_\pi V_\xi^\pi$.

[Quarel and Lang, n.d.; Hutter, 2002; Hutter, 2005, Ch. 4–5]

Let $\mathcal{M}$ be countable, with **realizability** $\mu \in \mathcal{M}$. Fix a **horizon** m ($\gamma_k = 1$ for $k \le m$, else 0) and evaluate at the empty history, so $V_\nu^\pi = \frac{1}{m}\mathbb{E}_\nu^\pi[r_1 + \cdots + r_m]$; let $\pi^*_{\xi,m}$ be Bayes-optimal.

Theorem (Pareto optimality)

No policy π has $V_\nu^\pi \ge V_\nu^{\pi^*_{\xi,m}}$ for all $\nu \in \mathcal{M}$, with strict inequality for some ν.

Proof. Otherwise $V_\xi^\pi = \sum_\nu w_\nu V_\nu^\pi > \sum_\nu w_\nu V_\nu^{\pi^*_{\xi,m}} = V_\xi^*$.

Definition (Self-optimizing). A sequence of policies $\tilde{\pi}_m$ is **self-optimizing** for $\mathcal{M}$ if $V_\nu^* - V_\nu^{\tilde{\pi}_m} \to 0$ as $m \to \infty$, for every $\nu \in \mathcal{M}$.

Theorem (self-optimization, fixed horizon)

If $V_\nu^* - V_\nu^{\tilde{\pi}_m} \le \Delta(m)$ for all $\nu \in \mathcal{M}$, then $V_\mu^* - V_\mu^{\pi^*_{\xi,m}} \le \Delta(m)/w_\mu$. So if any sequence is self-optimizing for $\mathcal{M}$, then so is $\pi^*_{\xi,m}$.

Proof. As $V_\xi = \sum_\nu w_\nu V_\nu$ and $\pi^*_{\xi,m}$ maximizes it,

$w_\mu[V_\mu^* - V_\mu^{\pi^*_{\xi,m}}] \le \sum_\nu w_\nu[V_\nu^* - V_\nu^{\pi^*_{\xi,m}}] \le \sum_\nu w_\nu[V_\nu^* - V_\nu^{\tilde{\pi}_m}] \le \Delta(m)$.

[Hutter, 2002, Thms. 2, 4]

A fixed horizon gives a different policy for each m. We want *one* policy that converges to optimality.

Definition (Self-optimizing, off-policy). Fix summable discounts γ_k and a **historic policy** π that generates the past actions. A policy $\tilde{\pi}$ is **self-optimizing** for $\mathcal{M}$ if for every $\nu \in \mathcal{M}$, $V_\nu^*(æ_{<t}) - V_\nu^{\tilde{\pi}}(æ_{<t}) \to 0$ as $t \to \infty$, ν^π-a.s.

Theorem (self-optimization)

If some policy is self-optimizing for $\mathcal{M}$, then so is π_ξ^*.

- The same chain with posterior weights gives $V_\mu^* - V_\mu^{\pi_\xi^*} \le \sum_\nu \frac{w(\nu|æ_{<t})}{w(\mu|æ_{<t})}\left[V_\nu^* - V_\nu^{\tilde{\pi}}\right]$. Each gap vanishes ν^π-a.s., but showing that their posterior average vanishes μ^π-a.s. takes a long argument (omitted).
- The necessary condition is sufficient. E.g. ergodic MDPs admit self-optimizing policies, so π_ξ^* is self-optimizing for them.

[Hutter, 2002; Hutter et al., 2024, Thm. 7.3.7]

Definition of AIXI

Definition (AIXI). AIXI is the Bayes-optimal policy $\pi^*_{\xi_U}$ for the class $\mathcal{M}_U$ of all lower semicomputable (chronological) semimeasures, with prior $w_\nu := 2^{-K(\nu)}$. Via the **recursive value function** (Bellman equation), it takes the action $a_t \in \arg\max_a V^*_{\xi_U}(æ_{<t}a)$, where $V^*_{\xi_U}(æ_{<t}) := \max_a V^*_{\xi_U}(æ_{<t}a)$ and

$$V^*_{\xi_U}(æ_{<t}a_t) = \frac{1}{\Gamma_t}\sum_{e_t} \xi_U(e_t \mid æ_{<t}a_t)\big(\gamma_t r_t + \Gamma_{t+1} V^*_{\xi_U}(æ_{1:t})\big).$$

Realizability.

It contains every computable environment, so $\mu \in \mathcal{M}_U$ reduces to "the universe is computable," a realizability assumption that we preferred to avoid in the predictive setting, but most of our results use in the interactive setting.

Caveat.

The constants depend on U, and adversarial choices of U can make AIXI behave arbitrarily badly (Leike and Hutter, 2015).

[Hutter, 2000; Hutter, 2005; Leike, 2016, Sec. 4.2; Quarel and Lang, n.d.]

How incomputable is AIXI?

Computability of optimal and ε-optimal policies, with the recursive value function and $\nu \in \mathcal{M}_U$:

Agent	**Optimal**	**ε-optimal**
AIMU (μ a computable measure)	Δ_2^0	Δ_1^0
AINU (any $\nu \in \mathcal{M}_U$)	Δ_3^0, Π_2^0-hard for some ν	Δ_2^0, Σ_1^0-hard for some ν
AIXI ($\nu = \xi_U$)	Δ_3^0, not computable for any U	Δ_2^0, Σ_1^0-hard for some U

OPEN PROBLEM

Is AIXI limit computable (Δ_2^0) for some U and tie-breaking rule? Contrary to common belief, this is not settled: the Π_2^0-hardness, which comes from detecting exact argmax ties, is proven only for special *non-universal* environments ν.

[Leike, 2016, Ch. 6]

AIXI is not worst possible in deterministic environments

Theorem. If $\mu \in \mathcal{M}$ is deterministic and $V_\mu^*(æ_{<t}) > \varepsilon > 0$ for all t along the history generated by μ and π_ξ^*, then $V_\xi^*(æ_{<t}) \geq w_\mu \varepsilon > 0$ for all t.

Proof. By linearity of the value in the environment, using the policy π_μ^*,

$$V_\xi^*(æ_{<t}) \;\geq\; V_\xi^{\pi_\mu^*}(æ_{<t}) \;=\; \sum_{\nu \in \mathcal{M}} w(\nu \mid æ_{<t})\, V_\nu^{\pi_\mu^*}(æ_{<t}) \;\geq\; w(\mu \mid æ_{<t})\, V_\mu^*(æ_{<t}).$$

Since μ is deterministic, $\mu(æ_{<t}) = 1$ on its own history, so $w(\mu \mid æ_{<t}) = w_\mu / \xi(æ_{<t}) \geq w_\mu$. □

- AIXI never comes to believe that a good deterministic world is worthless.
- The guarantee is a constant lower bound $w_\mu \varepsilon$, not near-optimal value.

[Hutter et al., 2024, Thm. 7.4.10; Quarel and Lang, n.d.]

Asymptotic optimality is achievable

Definition. π is **asymptotically optimal (a.o.)** in $\mathcal{M}$ if $V^*_\mu(æ_{<t}) - V^\pi_\mu(æ_{<t}) \to 0$ on μ^π-histories for every $\mu \in \mathcal{M}$: **strongly** if a.s., **in mean** if in expectation, **weakly** if the Cesàro averages converge a.s.

AIXI is not a.o. (Orseau, 2013); no deterministic policy is strongly a.o. in computable environments (Lattimore and Hutter, 2011). Deliberate explorers succeed:

Agent	Rule	Guarantee
Optimism (Sunehag and Hutter, 2012)	Follow π^*_ν for the most optimistic consistent $\nu \in \mathcal{M}$.	Finite deterministic $\mathcal{M}$: optimal eventually.
BayesExp (Lattimore, 2013)	Follow π^*_ξ, with bursts of exploration when information gain $> \varepsilon_t$.	Weakly a.o. in countable $\mathcal{M}$.
Thompson (Leike et al., 2016a)	Sample $\nu \sim w(\cdot \mid æ_{<t})$, follow π^*_ν for an effective horizon, resample.	A.o. in mean in countable $\mathcal{M}$.
Inq (Cohen et al., 2019)	Explore at random, with odds set by expected information gain.	Strongly a.o. in all computable $\mathcal{M}$ (bounded horizon).

[Hutter, n.d.; Leike, 2016, Ch. 5]

Unsafe exploration

Definition. An event E (a measurable set of infinite histories) is **accessible infinitely often** if there are a computable deterministic policy π^E, a decidable set $\tau \subseteq \mathcal{H}$ of histories that occurs infinitely often, a horizon k and $\varepsilon > 0$ such that, from every $æ_{<t} \in \tau$, running π^E for k steps makes E happen with μ-probability $> \varepsilon$. Otherwise E **becomes inaccessible**.

Theorem. Let $\mathcal{M}$ be the computable environments, and suppose μ "has no heaven": $\limsup_t V^*_\mu(æ_{<t}) < 1$ a.s. If π is strongly a.o. in $\mathcal{M}$, then for every event E, μ^π-a.s., E happens or becomes inaccessible.

- With $E =$ "the agent is destroyed": every a.o. agent is destroyed or incapacitated. Weakly a.o. agents: the same, with "regularly accessible".
- **Proof idea.** An a.o. agent tries everything (every computable deterministic policy, from every recurring decidable context, infinitely often), or it could not learn an environment that agrees with μ except for rewarding the untried behavior maximally.

[Cohen et al., 2021]

Exploration with traps

Example (heaven or hell). In the computable environment ν_b ($b \in \{0,1\}$), the first action $a_1 = b$ leads to reward 0 forever and $a_1 = 1 - b$ to reward 1 forever. Every policy π has $V^*_{\nu_b} - V^\pi_{\nu_b} = \pi(a_1 = b) \geq \frac{1}{2}$ for some b, for any horizon or discount.

So no sequence of policies is self-optimizing for $\mathcal{M}_U$, and the fixed-horizon theorem says nothing about AIXI.

AIXI is not a.o. (Orseau, 2013), and it may stop exploring once further exploration is not worth the risk (Leike and Hutter, 2015). Given the theorem, this is reasonable.

> Bayesian agents trade off subjectively assessed risks and benefits of exploratory actions, which is the best one can do in novel situations.

[Hutter, 2002; Cohen et al., 2021]

Safe exploration

Rather than optimality in every environment, aim to do at least as well as a **mentor** (a human, or a known-safe policy) whom the agent may defer to:

Cohen et al., 2021

Mentee follows π^*_ξ, but defers to the mentor with probability set by the expected information gain from doing so. Deferral vanishes; value approaches at least the mentor's. Exploration is safe, but exploitation need not be.

Cohen and Hutter, 2020

Pessimism: maximize the worst-case expected reward over the most plausible world-models, deferring to the mentor when unsure. With probability $1 - \delta$ it causes no *unprecedented* events; its value approaches at least the mentor's while deferral vanishes.

Ebtekar and Cohen, 2026

Golden handcuffs: allow a large negative reward $-L$ in the agent's subjective reward range, and yield to the mentor when predicted value drops below a threshold. Sublinear regret against the best mentor, and no decidable low-complexity predicate is triggered before a mentor triggers it.

Stopping complexity

Cohen's safety condition is deferring to a mentor instead of violating any simple rule (or predicate) for the first time.

A sufficiently cautious agent will follow all simply-describable precedent, which is a brute-force approach to avoiding catastrophe.

The histories where a rule is first violated form a prefix-free set.

The complexity of such sets can be viewed as a universal "novelty/out-of-distribution" measure, which has led to new and independently interesting problems in algorithmic probability.

[Cohen and Hutter, 2020; Ebtekar and Cohen, 2026; Vovk and Pavlovic, 2017; Andreev et al., 2018; Mironov et al., 2026]

Developing the mathematics to model superintelligent agents is **obviously very important** for AI safety.

Algorithmic information theory is often dismissed as arcane or nerd-snipey. This is mostly based on misconceptions that are rebutted in these slides, such as:

- Dependence on the UTM or assuming Occam's razor
- Incomputable theory $\implies$ no practical implications
- Questioning various assumptions that we do not actually rely on

Algorithmic Information Theory is the **deepest theory grappling with the hard problems.**

There are many open problems in this space. If you want to solve some:

- AIXI Labs is hiring! (See aixi.uk)
- Other groups that seem to be attacking the core problems include: CORAL, Resolution's agent foundations team, and Principled Agents, among others!

References (1/4)

P. Almond. What is a low level language. Essay, paul-almond.com, 2005.

D. Amodei. The big blob of compute hypothesis. Internal memo, 2017. First 10 pages published in K. Roose, "The essay that started the AI race," Substack, 6 Oct. 2026, `https://kevinroose.substack.com/p/the-essay-that-started-the-ai-race`.

M. Andreev, G. Posobin, and A. Shen. Plain stopping time and conditional complexities revisited. In *MFCS*, 2018.

B. Bauwens and G. Kuznetsov. Solomonoff induction of selected bits and online classification. arXiv:2610.18632, 2026.

D. Blackwell and L. Dubins. Merging of opinions with increasing information. *Annals of Mathematical Statistics*, 33(3):882–886, 1962.

M. K. Cohen, E. Catt, and M. Hutter. A strongly asymptotically optimal agent in general environments. In *IJCAI*, 2019.

M. K. Cohen and M. Hutter. Pessimism about unknown unknowns inspires conservatism. In *COLT*, 2020.

M. K. Cohen, E. Catt, and M. Hutter. Curiosity killed or incapacitated the cat and the asymptotically optimal agent. *IEEE Journal on Selected Areas in Information Theory*, 2, 2021.

T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 2nd edition, 2006.

A. Ebtekar and M. K. Cohen. Golden handcuffs make safer AI agents. arXiv:2604.13609, 2026.

A. Ebtekar, M. Hutter, and D. J. Sutherland. Understanding generalization requires universal induction. arXiv:2609.34458, 2026.

B. Fallenstein, J. Taylor, and P. F. Christiano. Reflective oracles: a foundation for game theory in artificial intelligence. In *LORI*, 2015.

B. Fallenstein, N. Soares, and J. Taylor. Reflective variants of Solomonoff induction and AIXI. In *AGI*, 2015.

References (2/4)

T. S. Ferguson. *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press, 1967.

S. Garrabrant, T. Benson-Tilsen, A. Critch, N. Soares, and J. Taylor. Logical induction. arXiv:1609.03543, 2016.

M. Hutter. A theory of universal artificial intelligence based on algorithmic complexity. arXiv:cs/0004001, 2000.

M. Hutter. Self-optimizing and Pareto-optimal policies in general environments based on Bayes-mixtures. In *COLT*, 2002.

M. Hutter. Optimality of universal Bayesian sequence prediction for general loss and alphabet. *Journal of Machine Learning Research*, 4, 2003.

M. Hutter. *Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability*. Springer, 2005.

M. Hutter. On universal prediction and Bayesian confirmation. *Theoretical Computer Science*, 384:33–48, 2007.

M. Hutter. Universal artificial intelligence (lecture slides). `http://www.hutter1.net/ai/suaibook.pdf`.

M. Hutter, D. Quarel, and E. Catt. *An Introduction to Universal Artificial Intelligence*. CRC Press, 2024.

V. Kosoy. Forecasting using incomplete models. arXiv:1705.04630, 2017.

D. Kusum. Solomonoff induction with reflective oracles. arXiv:2611.14454, 2026.

L. Lang. A technical introduction to Solomonoff induction without K-complexity. LessWrong, 2025.

T. Lattimore. *Theory of General Reinforcement Learning*. PhD thesis, Australian National University, 2013.

T. Lattimore and M. Hutter. Asymptotically optimal agents. In *ALT*, 2011.

T. Lattimore, M. Hutter, and V. Gavane. Universal prediction of selected bits. In *ALT*, 2011.

J. Leike and M. Hutter. Bad universal priors and notions of optimality. In *COLT*, 2015.

References (3/4)

J. Leike, T. Lattimore, L. Orseau, and M. Hutter. Thompson sampling is asymptotically optimal in general environments. In *UAI*, 2016.

J. Leike, J. Taylor, and B. Fallenstein. A formal solution to the grain of truth problem. In *UAI*, 2016.

J. Leike. *Nonparametric General Reinforcement Learning*. PhD thesis, Australian National University, 2016.

M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer, 4th edition, 2019.

A. Meulemans, R. Nasser, et al. Embedded universal predictive intelligence: a coherent framework for multi-agent learning. arXiv:2511.22226, 2025.

M. Mironov, A. Ebtekar, and C. Wyeth. Failure of the coding theorem for randomized stopping machines. LessWrong, 19 September 2026.

S. Neth. A dilemma for Solomonoff prediction. *Philosophy of Science*, 90(2), 2023.

L. Orseau. Asymptotic non-learnability of universal agents with computable horizon functions. *Theoretical Computer Science*, 473, 2013.

L. Orseau and M. Ring. Space-time embedded intelligence. In *AGI*, 2012.

H. Putnam. "Degree of confirmation" and inductive logic. In P. A. Schilpp, editor, *The Philosophy of Rudolf Carnap*. Open Court, 1963.

D. Quarel and L. Lang. Solomonoff induction. Iliad Intensive, module D.3.1. `https://iliad-intensive.org/agency/solomonoff-induction/`.

S. Rathmanner and M. Hutter. A philosophical treatise of universal induction. *Entropy*, 13(6), 2011.

D. Ryabko. On asymptotic and finite-time optimality of Bayesian predictors. *Journal of Machine Learning Research*, 20:1–24, 2019.

G. Schurz. *Hume's Problem Solved: The Optimality of Meta-Induction*. MIT Press, 2019.

R. J. Solomonoff. A formal theory of inductive inference, Parts I and II. *Information and Control*, 7, 1964.

References (4/4)

R. J. Solomonoff. Complexity-based induction systems: comparisons and convergence theorems. *IEEE Transactions on Information Theory*, 24(4), 1978.

T. F. Sterkenburg. Solomonoff prediction and Occam's razor. *Philosophy of Science*, 83(4), 2016.

T. F. Sterkenburg. *Universal Prediction: A Philosophical Investigation*. PhD thesis, University of Groningen, 2018.

T. F. Sterkenburg. Putnam's diagonal argument and the impossibility of a universal learning machine. *Erkenntnis*, 84(3), 2019.

T. F. Sterkenburg. The metainductive justification of induction: the pool of strategies. *Philosophy of Science*, 86(5), 2019.

T. F. Sterkenburg. Solomonoff induction. arXiv:2603.20274, 2026.

P. Sunehag and M. Hutter. Optimistic agents are asymptotically optimal. In *Australasian Joint Conference on AI*, 2012.

R. S. Sutton. The bitter lesson. Blog post, `http://www.incompleteideas.net/IncIdeas/BitterLesson.html`, 2019.

V. Vovk and D. Pavlovic. Universal probability-free prediction. *Annals of Mathematics and Artificial Intelligence*, 81(1), 2017.

V. Vovk, A. Takemura, and G. Shafer. Defensive forecasting. In *AISTATS*, 2005.

A. Wald. *Statistical Decision Functions*. Wiley, 1950.

I. Wood, P. Sunehag, and M. Hutter. (Non-)equivalence of universal priors. In *Solomonoff 85th Memorial Conference*, LNCS 7070. Springer, 2013.

C. Wyeth, M. Hutter, J. Leike, and J. Taylor. Limit-computable grains of truth for arbitrary computable extensive-form (un)known games. arXiv:2508.16245, 2025.