On The Independence Postulate

A Finitary Church Turing Thesis

samepst@jptheorygroup.org
Samuel Epstein


The Independence Postulate

Finitary Church Turing Thesis	[Lev84, Lev13]
Logic	[Lev13]
Outliers	[Eps21, Eps22c, Eps23c]
Induction	[Lev16, Eps19]
Games	[Eps22a, Eps22c]
Machine Learning	[Eps21, Eps23b]
Physics	[Eps23a]

Statement

IP: *Let* α *be a sequence defined with an n-bit mathematical statement (e.g., in Peano Arithmetic), and a sequence* β *can be located in the physical world with a k-bit instruction set (e.g., ip-address). Then* $\mathbf{I}(\alpha : \beta) < k + n + c$*, for some small absolute constant* c*.*

Statement

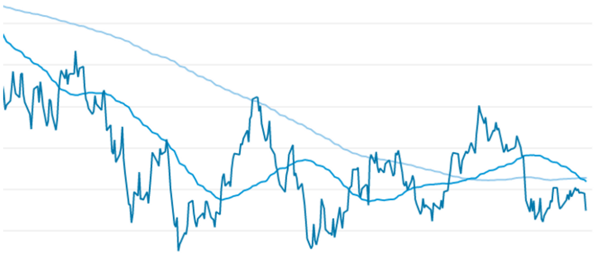
IP: *Let α be a sequence defined with an n-bit mathematical statement (e.g., in Peano Arithmetic), and a sequence β can be located in the physical world with a k-bit instruction set (e.g., ip-address). Then $\mathbf{I}(\alpha : \beta) < k + n + c$, for some small absolute constant c.*

$\mathbf{I}(x : y) = \mathbf{K}(x) + \mathbf{K}(y) - \mathbf{K}(x, y)$.
$\mathbf{I}(x; \mathcal{H}) = \mathbf{K}(x) - \mathbf{K}(x|\mathcal{H})$.
$\mathbf{I}(\alpha : \beta) = \log \sum_{x,y} \mathbf{m}(x|\alpha)\mathbf{m}(y|\beta)2^{\mathbf{I}(x:y)}$.

Independence Postulate

From [Lev13]:

Thus, a (physical) sequence of all mathematical publications has little information about the (mathematical) sequence of all true statements of arithmetic. This is of little concern because the latter has, in turn, little information about the stock market (a physical sequence).



Finitary Church-Turing Thesis

$\mathcal{H}_n = \mathcal{H}[0 \ldots 2^n]$

Singleton language $\{\mathcal{H}_n\}$ is computable.

Finitary Church-Turing Thesis

$\mathcal{H}_n = \mathcal{H}[0 \ldots 2^n]$

Singleton language $\{\mathcal{H}_n\}$ is computable.

By **IP**,

$$\begin{aligned} n =^{\log} \mathbf{K}(\mathcal{H}_n) =^{\log} \mathbf{I}(\mathcal{H}_n : \mathcal{H}_n) <^{\log} \mathbf{NR}(\mathcal{H}_n) + \mathbf{Addr}(\mathcal{H}_n), \\ \mathbf{NR}(\mathcal{H}_n) = O(\log n), \\ n <^{\log} \mathbf{Addr}(\mathcal{H}_n). \end{aligned}$$

Recursive Sequences

Finite **recursive** sequences: algorithmic descriptions $\mathbf{K}(x)$ as short as any their "higher-level" math descriptions $\mathbf{NR}(x)$.

IP $\Rightarrow$ only recursive sequences exist in reality.

Recursive Sequences

Finite **recursive** sequences: algorithmic descriptions $\mathbf{K}(x)$ as short as any their "higher-level" math descriptions $\mathbf{NR}(x)$.

IP $\Rightarrow$ only recursive sequences exist in reality.

$$\mathbf{K}(x) =^{+} \mathbf{I}(x:x) <^{+} \mathbf{NR}(x) + \mathbf{Addr}(x)$$
$$\mathbf{K}(x) - \mathbf{NR}(x) <^{+} \mathbf{Addr}(x).$$

Mathematical Sequences

Mathmatical strings have $\mathbf{NR}(x) \ll \|x\|$.

Such physically obtainable x must be algorithmic:

$$\mathbf{K}(x) \approx \mathbf{NR}(x) \ll \|x\|.$$

$9999^{1,000,000}$ is

1. mathematical
2. algorithmic
3. physically obtainable

Conservation Inequalities

[Lev74, Lev84, Eps22c, Eps22b, Ver21]

Theorem

1. $\mathbf{I}(f(x):y) <^{+} \mathbf{I}(x:y)$.
2. $\Pr_{x\sim p}[\mathbf{I}(p:y) < \mathbf{I}(x:y)+m] \overset{*}{<} 2^{-m}$.

Target Sequences

[Lev13]:

> *"IP is much more comprehensive: CT prohibits only generating the target math sequence itself; IP bars all strings with any significant information about it. So, IP can be applied where CT cannot."*

Example:

CT: cannot compute $\mathcal{H}$.

IP: cannot generate string x with high $\mathbf{I}(x;\mathcal{H})$,

$\alpha = \mathcal{H}$, $\beta = x$, $n = O(1)$.

$\mathbf{I}(x;\mathcal{H}) <^{+} \mathbf{Addr}(x)$.

Logic

There is no complete, consistent extension of PA based on a recursively enumerable set of axioms.

Gödel had doubts.

Logic

There is no complete, consistent extension of PA based on a recursively enumerable set of axioms.

Gödel had doubts.

Theorem ([Lev13])

Every consistent completion α *of PA has* $\mathbf{I}(\alpha : \mathcal{H}) = \infty$.

Thus by **IP**: $\mathbf{Addr}(\alpha) = \infty$.

Induction

Complete environment: huge string x.

Observation: $D \subset \{0,1\}^*$, $x \in D$.

Example: macro parameters pressure, temperature,

Induction: find the hidden part x of the environment.

Induction

Complete environment: huge string x.

Observation: $D \subset \{0,1\}^*$, $x \in D$.

Example: macro parameters pressure, temperature,

Induction: find the hidden part x of the environment.

Apriori distribution p

$$\arg\max_{x \in D} p(x).$$

Induction

Complete environment: huge string x.

Observation: $D \subset \{0,1\}^*$, $x \in D$.

Example: macro parameters pressure, temperature,

Induction: find the hidden part x of the environment.

Apriori distribution p

$$\arg\max_{x \in D} p(x).$$

$\mathbf{m}(x) = 2^{-\mathbf{K}(x)}$, $O(1)\mathbf{m} > p$, $\mathbf{d}(x|\mathbf{m}) = O(1)$.

No refutation to "x is generated by $\mathbf{m}$".

$$\arg\min_{x \in D} \mathbf{K}(x).$$

Induction

There could exist $G \subset D$ of hypotheses representing a concept (such as a more detailed description of particles) with

$$\max_{x \in D} \mathbf{m}(x) \ll \mathbf{m}(G).$$

Induction

There could exist $G \subset D$ of hypotheses representing a concept (such as a more detailed description of particles) with

$$\max_{x \in D} \mathbf{m}(x) \ll \mathbf{m}(G).$$

Only a mathermatical construction

EL Theorem states:

$$\min_{x \in D} \mathbf{K}(x) <^{\log} -\log \mathbf{m}(D) + \mathbf{I}(D; \mathcal{H}).$$

By **IP**:

$$\min_{x \in D} \mathbf{K}(x) + \log \mathbf{m}(D) \lessapprox \mathbf{Addr}(D).$$

Occam's razor is the prefered strategy.

Game:

Player **p** and environment **q** exchange actions ($\mathbb{N}$)

q can declare **p** the winner.

Games can never end.

Game:

Player **p** and environment **q** exchange actions ($\mathbb{N}$)

q can declare **p** the winner.

Games can never end.

Example

q has $\textsc{k-Sat}$ formula, **p** has n variables assignments.

At each round, **q** $\Rightarrow$ unsatisfied clauses $\Rightarrow$ **p**.

p $\Rightarrow$ k changed variables $\Rightarrow$ **q**.

If all clauses satisfied, **p** wins.

Games

Theorem

If randomized player $\mathbf{p}$ *wins with probability* p *with* $\mathbf{q}$*, then exists deterministic winning player with complexity* $<^{\log} \mathbf{K}(\mathbf{p}) - \log p + \mathbf{I}((\mathbf{p}, p, \mathbf{q}); \mathcal{H})$.

Corollary

For $k > 5$*, there exists a player that can beat the* $\mathrm{K\text{-}SAT}$ *environment* $\mathbf{q}$ *in* $1.1n/k$ *turns, with complextiy* $<^{\log} \mathbf{I}(\mathbf{q}; \mathcal{H})$.

By **IP**:

> *If an physically realizable enviroment has a good simple randomized player then it has a simple winning deterministic player.*

Machine Learning

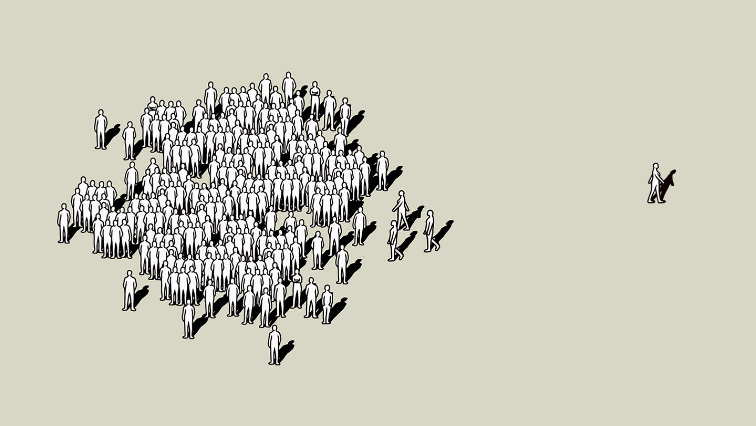
Theorem

Given a sample $\{(x_i, b_i)\}_{i=1}^n$, *there is a function* $f : \{0,1\}^* \to \{0,1\}$ *such that* $f(x_i) = b_i$, *for* $i = 1, \ldots, n$, *and* $\mathbf{K}(f) <^{\log} n + \mathbf{I}(\{(x_i, b_i)\}; \mathcal{H})$.

By **IP**:

> *Given a realizable sample* $\{(x_i, b_i)\}_{i=1}^n$, *there is a function* $f : \{0,1\}^* \to \{0,1\}$ *such that* $f(x_i) = b_i$, *for* $i = 1, \ldots, n$, *and* $\mathbf{K}(f) <^{\log} n$.

Outliers



Outliers

Hawkins (1980): An observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism.

Cambridge Dictionary: a person, thing, or fact that is very different from other people, things, or facts, so that it cannot be used to draw general conclusions.

Oxford English Dictionary: An observation whose value lies outside the set of values considered likely according to some hypothesis (usually one based on other observations); an isolated point.

Deficiency of Randomness

Finite sequences x, probabilities p over $\mathbb{N}$

$$\mathbf{d}(x|p) = -\log p(x) - \mathbf{K}(x|p).$$

Infinite sequences α, probabilities P over $\{0,1\}^\infty$.

$$\mathbf{D}(\alpha|P) = \sup_n \left(-\log P(\alpha_n) - \mathbf{K}(\alpha_n|P)\right).$$

See [G2́1] for more information.

All Sampling Methods Produce Outliers

Definition

A discrete sampling method A is a computable function that maps an integer N with probability 1 to a set containing N unique strings.

Theorem

$$\Pr\left(\max_{a \in A(2^n)} \mathbf{d}(a|p) < n - k - c \log n\right) \leq 2e^{-2^k}.$$

All Sampling Methods Produce Outliers

Definition

A continuous sampling method B is a computable function that maps, with probability 1, an integer N to an infinite encoding of N different infinite sequences.

Theorem

$$\Pr\Big(\max_{\alpha \in B(2^n)} \mathbf{D}(\alpha|P) < n - k - c\log n\Big) \leq 2.5e^{-2^k}.$$

Outliers from Probabilistic Algorithms

μ: computable measure over $\{0,1\}^\infty$.

λ: non-atomic computable measure over $\{0,1\}^\infty$.

λ: output of randomized algorithm.

Theorem

$\lambda\{\alpha : \mathbf{D}(\alpha|\mu) > n\} > 2^{-n-\mathbf{K}(n,\mu,\lambda)-O(1)}$.

Ergodic Dynamics: Cantor Space

Let λ, μ be computable measures over $\{0,1\}^\infty$.

Let T be ergodic and measure preserving over probability space $(\{0,1\}^\infty, \mathcal{B}, \lambda)$.

Theorem

Starting λ-almost everywhere, $> 2^{-n-\mathbf{K}(n,\mu,\lambda)-O(1)}$ states α visited by n iterations of T, as $n \to \infty$, have $\mathbf{D}(\alpha|\mu) > n$.

Dynamics: Computable Metric Spaces

$\mathfrak{X}$: computable metric space.

μ: computable measure.

$\mathbf{t}_\mu$: universal uniform test.

G^t: 1D transformation group acting on $\mathfrak{X}$.

Theorem

Let $\alpha \in \mathcal{X}$, with finite mutual information with $\mathcal{H}$. There is a constant c with $\mathrm{Leb}\{t \in [0,1] : \mathbf{t}_\mu(G^t\alpha) > 2^n\} > 2^{-n-\mathbf{K}(n)-c}$.

Outliers in Complex Systems

Not computable, but it has an address!

Outliers in Complex System

Process is infinite sequence of infinite sequences $\gamma \in \{0,1\}^{\infty^{\mathbb{N}}}$.

- $\gamma[n]$ is the first 2^n sequences of γ.

Theorem

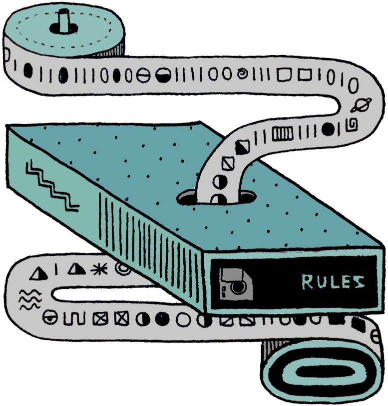
$$t = \sup_n (n - \mathbf{K}(n) - \max_{\alpha \in \gamma(n)} \mathbf{D}(\alpha|P)).$$
$$t <^{\log} \mathbf{I}(\langle\gamma\rangle : \mathcal{H}) + O(\log \mathbf{K}(P)),$$

By **IP**:

$$t <^{+} \mathbf{Addr}(\gamma) + +O(\log \mathbf{K}(P)).$$

It's hard to find observations with small anomalies and impossible to find observations with no anomalies.

Conclusion



S. Epstein.
On the algorithmic probability of sets.
CoRR, abs/1907.04776, 2019.

Samuel Epstein.
All sampling methods produce outliers.
IEEE Transactions on Information Theory, 67(11):7568–7578, 2021.

S. Epstein.
22 examples of solution compression via derandomization.
CoRR, abs/2208.11562, 2022.

S. Epstein.
The kolmogorov birthday paradox.
CoRR, abs/2208.11237, 2022.

S. Epstein.
The outlier theorem revisited.
CoRR, abs/2203.08733, 2022.

S. Epstein.

A Complication for the Many Worlds Theory.
CoRR, abs/2302.07649, 2023.

S. Epstein.
Regression and Algorithmic Information Theory.
CoRR, abs/2304.07825, 2023.

S. Epstein.
Uniform Tests and Algorithmic Thermodynamic Entropy.
CoRR, abs/2303.05619, 2023.

Peter Gács.
Lecture notes on descriptional complexity and randomness.
CoRR, abs/2105.04704, 2021.

L. A. Levin.
Laws of Information Conservation (Non-growth) and Aspects of the Foundations of Probability Theory.
Problemy Peredachi Informatsii, 10(3):206–210, 1974.

L. A. Levin.

Randomness conservation inequalities; information and independence in mathematical theories.
Information and Control, 61(1):15–37, 1984.

L. A. Levin.
Forbidden information.
J. ACM, 60(2), 2013.

L. A. Levin.
Occam bound on lowest complexity of elements.
Annals of Pure and Applied Logic, 167(10):897–900, 2016.
And also: S. Epstein and L.A. Levin, Sets have simple members, arXiv preprint arXiv:1107.1458, 2011.

N. Vereshchagin.
Proofs of conservation inequalities for levin's notion of mutual information of 1974.
Theoretical Computer Science, 856, 2021.