A Short Proof on the Existence of Anomalies

Samuel Epstein*

July 20, 2022

Abstract

The Independence Postulate (IP) is a finitary Church-Turing Thesis, saying mathematical sequences are independent from physical ones. IP implies the existence of anomalies.

Anomalies

$\mathbf{K}(x|y)$ is the conditional prefix Kolmogorov complexity. For probability p over $\mathbb{N}$, randomness deficiency is $\mathbf{d}(a|p,b) = \lfloor -\log p(a) \rfloor - \mathbf{K}(a|b)$ and measures the extent of the refutation against the hypothesis p given the result a [G21]. $\mathbf{d}(a|p) = \mathbf{d}(a|p,\emptyset)$. $\mathbf{I}(a;\mathcal{H}) = \mathbf{K}(a) - \mathbf{K}(a|\mathcal{H})$, where $\mathcal{H}$ is the halting sequence. $<^+ f$ is $< f+O(1)$ and $<^{\log} f$ is $< f+O(\log(f+1))$. Stochasticity is $\Lambda(a|b) = \min\{\mathbf{K}(Q|b) + 3\log\max\{\mathbf{d}(a|Q,b),1\} : Q$ has finite support and a range in $\mathbb{Q}\}$. $\Lambda(a|b) < \Lambda(a) + O(\log \mathbf{K}(b))$. The following definition is from [Lev74] .

Definition 1 (Information) $\mathbf{I}(\alpha:\beta)= \log\sum_{x,y} 2^{\mathbf{K}(x)+\mathbf{K}(y)-\mathbf{K}(x,y)-\mathbf{K}(x|\alpha)-\mathbf{K}(y|\beta)}$.

The Independence Postulate states [Lev13]:

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

There are many proofs in the literature that stochastic numbers have high mutual information with the halting sequence. One such detailed proof is in [Eps21].

Lemma 1 $\Lambda(x) <^{\log} \mathbf{I}(x;\mathcal{H})$.

Lemma 2 *For probability p over $\mathbb{N}$, $D\subset\mathbb{N}$, $|D| = 2^s$, $s < \max_{a\in D}\mathbf{d}(a|p) + \Lambda(D) + \mathbf{K}(s,p) + O(\log\mathbf{K}(s,p))$.*

Proof. We relativize the universal Turing machine to $\langle s,p\rangle$. Let Q be a probability measure that realizes $\Lambda(D)$, with $d = \max\{\mathbf{d}(D|Q),1\}$. Let $F \subseteq \mathbb{N}$ be a random set where each element $a\in\mathbb{N}$ is selected independently with probability $cd2^{-s}$, where $c\in\mathbb{N}$ is chosen later. $\mathbf{E}[p(F)] \leq cd2^{-s}$. Furthermore

$$\mathbf{E}[Q(\{G : |G| = 2^s, G\cap F = \emptyset\})] \leq \sum_G Q(G)(1-cd2^{-s})^{2^s} < e^{-cd}.$$

Thus finite $W \subset \mathbb{N}$ can be chosen such that $p(W) \leq 2cd2^{-s}$ and $Q(\{G : |G| = 2^s, G\cap W = \emptyset\}) \leq e^{1-cd}$. $D\cap W \neq \emptyset$, otherwise, using the Q-test, $t(G) = e^{cd-1}$ if $(|G| = 2^s, G\cap W = \emptyset)$ and $t(G) = 0$ otherwise, we have

$$\begin{aligned}
&\mathbf{K}(D|Q,d,c) <^+ -\log Q(D) - (\log e)cd\\
&\quad(\log e)cd <^+ -\log Q(D) - \mathbf{K}(D|Q) + \mathbf{K}(d,c)\\
&\quad(\log e)cd <^+ d + \mathbf{K}(d,c),
\end{aligned}$$

which is a contradiction for large enough c. Thus there is an $a\in D\cap W$, where

$$\begin{aligned}
&\mathbf{K}(a) <^+ -\log p(a) + \log d - s + \mathbf{K}(d) + \mathbf{K}(Q)\\
&\quad s <^+ \mathbf{d}(a|p) + \Lambda(D).
\end{aligned}$$

Making the relativization of $\langle s,p\rangle$ explicit,

$$\begin{aligned}
&s < -\log p(a) - \mathbf{K}(a|s,p) + \Lambda(D|s,p)\\
&s < \max_{a\in D}\mathbf{d}(a|p) + \Lambda(D) + \mathbf{K}(s,p)\\
&\quad + O(\log\mathbf{K}(s,p)). \ \square
\end{aligned}$$

Let $\tau\in\mathbb{N}^{\mathbb{N}}$ represent a series of observations. Assuming τ has an infinite amount of unique numbers, $\tau(n)$ is the first 2^n unique numbers of τ.

*JP Theory Group. samepst@jptheorygroup.org

Theorem 1 *For probability p over $\mathbb{N}$, $\tau \in \mathbb{N}^{\mathbb{N}}$, let $s_{\tau,p} = \sup_n \left(n - 3\mathbf{K}(n) - \max_{a\in\tau(n)} \mathbf{d}(a|p)\right)$. Then $s_{\tau,p} <^{\log} \mathbf{I}(\langle\tau\rangle : \mathcal{H}) + \mathbf{K}(p)$.*

Proof. By Lemmas 1 and 2, and the fact that $\mathbf{I}(x;\mathcal{H}) <^{+} \mathbf{I}(\alpha : \mathcal{H}) + \mathbf{K}(x|\alpha)$,

$$
\begin{aligned}
n &< \max_{a\in\tau(n)} \mathbf{d}(a|p) + \mathbf{I}(\tau(n);\mathcal{H}) + \mathbf{K}(p) + \mathbf{K}(n) \\
&\quad + O(\log \mathbf{I}(\tau(n);\mathcal{H})\mathbf{K}(p)\mathbf{K}(n)), \\
n &< \max_{a\in\tau(n)} \mathbf{d}(a|p) + 2\mathbf{K}(n) + \mathbf{I}(\langle\tau\rangle : \mathcal{H}) + \mathbf{K}(p) \\
&\quad + O(\log \mathbf{I}(\langle\tau\rangle : \mathcal{H})\mathbf{K}(p)\mathbf{K}(n)), \\
n - 3\mathbf{K}(n) &- \max_{a\in\tau(n)} \mathbf{d}(a|p) <^{\log} \mathbf{I}(\langle\tau\rangle : \mathcal{H}) + \mathbf{K}(p). \square
\end{aligned}
$$

Let k be a physical address of τ. $\mathcal{H}$ can be described by a small mathematical statement. By Theorem 1 and IP,

$$s_{\tau,p} <^{\log} \mathbf{I}(\langle\tau\rangle : \mathcal{H}) + \mathbf{K}(p) <^{\log} k + c + \mathbf{K}(p).$$

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

References

[Eps21] Samuel Epstein. All sampling methods produce outliers. *IEEE Transactions on Information Theory*, 67(11):7568–7578, 2021.

[Gá21] Peter Gács. Lecture notes on descriptional complexity and randomness. *CoRR*, abs/2105.04704, 2021.

[Lev74] 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.

[Lev13] L. A. Levin. Forbidden information. *J. ACM*, 60(2), 2013.