A Small Theorem for Small $\mathbf{m}$

Samuel Epstein*

August 22, 2023

Abstract

If a semi measure is greater than the universal semi-measure $\mathbf{m}$ up to a multiplicative constant, then it is exotic in that it has infinite mutual information with the halting sequence. This result has applications to Neutral Measures.

1 Introduction

In this note, we show that semi measures that majorize the algorithmic probability have infinite mutual information with the halting sequence. This means all neutral measures have infinite mutual information with the halting sequence.

For a probability p over $\{0,1\}^*$, $[p] \subset \{0,1\}^\infty$ is the set of infinite sequences $\beta \in [p]$ such that $U_x(\beta)$ outputs the bit representation of $p(x)$. The algorithm U is a standard universal Turing machine. $\mathbf{K}(x|y)$ is the prefix Kolmogorov complexity. $\mathbf{m}$ is the algorithmic probability. $\mathbf{I}(x:y) = \mathbf{K}(x) + \mathbf{K}(y) - \mathbf{K}(x,y)$ is the mutual information between two strings. For infinite sequences $\alpha, \beta \in \{0,1\}^\infty$, $\mathbf{I}(\alpha:\beta) = \log \sum_{x,y\in\{0,1\}^*} \mathbf{m}(x|\alpha)\mathbf{m}(y|\beta)2^{\mathbf{I}(x:y)}$ [Lev74]. The halting sequence is $\mathcal{H}$. The amount of mutual information between a probability p and $\mathcal{H}$ is $\mathbf{I}(p:\mathcal{H}) = \inf_{\beta\in[p]} \mathbf{I}(\beta:\mathcal{H})$. At the risk of a simple proof that $\mathbf{w}$ computes $0'$, the main theorem is as follows.

Theorem. *If $\mathbf{w}$ is a semimeasure on $\{0,1\}^*$ and $\mathbf{m} < O(1)\mathbf{w}$ then $\mathbf{I}(\mathbf{w}:\mathcal{H}) = \infty$.*

The amount of information that $\mathcal{H}$ has about $x \in \{0,1\}^*$ is $\mathbf{I}(x;\mathcal{H}) = \mathbf{K}(x) - \mathbf{K}(x|\mathcal{H})$. For positive real functions f, by $<^+ f$, $>^+ f$, $=^+ f$, and $<^{\log} f$, $>^{\log} f$, $\sim f$ we denote $\leq f + O(1)$, $\geq f - O(1)$, $= f \pm O(1)$ and $\leq f + O(\log(f+1))$, $\geq f - O(\log(f+1))$, $= f \pm O(\log(f+1))$. Furthermore, $\overset{*}{<} f$, $\overset{*}{>} f$ denotes $< O(1)f$ and $> f/O(1)$. The term and $\overset{*}{=} f$ is used to denote $\overset{*}{>} f$ and $\overset{*}{<} f$. The chain rule states $\mathbf{K}(x) + \mathbf{K}(y|x,\mathbf{K}(x)) =^+ \mathbf{K}(x,y)$.

2 Kolmogorov Complexity is Exotic

We cover material on busy beaver functions. Let $\Omega = \sum\{2^{-\|p\|} : U(p) \text{ halts}\}$ be Chaitin's Omega, $\Omega_n \in \mathbb{Q}_{\geq 0}$ be be the rational formed from the first n bits of Ω, and $\Omega^t = \sum\{2^{-\|p\|} : U(p) \text{ halts in time } t\}$. For $n \in \mathbb{N}$, let $\mathbf{bb}(n) = \min\{t : \Omega_n < \Omega^t\}$. $\mathbf{bb}^{-1}(m) = \arg\min_n\{\mathbf{bb}(n-1) < m \leq \mathbf{bb}(n)\}$. Let $\Omega[n] \in \{0,1\}^*$ be the first n bits of Ω. For $t \in \mathbb{N}$ define the function $\mathbf{m}^t(x) = \sum\{2^{-\|p\|} : U(p) = x \text{ in } t \text{ steps}\}$ and for $n \in \mathbb{N}$, we have $\mathbf{m}_n(x) = \sum\{2^{-\|p\|} : U(p) = x \text{ in } \mathbf{bb}(n) \text{ steps}\}$.

*JP Theory Group. samepst@jptheorygroup.org

Lemma 1 *For* $n = \mathbf{bb}^{-1}(m)$, $\mathbf{K}(\Omega[n]|m,n) = O(1)$.

Proof. For a string x, let $BB(x) = \inf\{t : \Omega^t > 0.x\}$. Enumerate strings of length n, starting with 0^n, and return the first string x such that $BB(x) \geq m$. This string x is equal to $\Omega[n]$, otherwise let y be the largest common prefix of x and $\Omega[n]$. Thus $BB(y) = \mathbf{bb}(\|y\|) \geq BB(x) \geq m$, which means $\mathbf{bb}^{-1}(m) \leq \|y\| < n$, causing a contradiction. □

Lemma 2 $\mathbf{I}(f(a);\mathcal{H}) <^+ \mathbf{I}(a;\mathcal{H}) + \mathbf{K}(f)$.

Proof.

$$\mathbf{I}(a;\mathcal{H}) = \mathbf{K}(a) - \mathbf{K}(a|\mathcal{H}) >^+ \mathbf{K}(a, f(a)) - \mathbf{K}(a, f(a)|\mathcal{H}) - \mathbf{K}(f).$$

The chain rule applied twice results in

$$\begin{aligned}
\mathbf{I}(a;\mathcal{H}) + \mathbf{K}(f) &>^+ \mathbf{K}(f(a)) + \mathbf{K}(a|f(a), \mathbf{K}(f(a))) - (\mathbf{K}(f(a)|\mathcal{H}) + \mathbf{K}(a|f(a), \mathbf{K}(f(a)|\mathcal{H}), \mathcal{H})\\
&=^+ \mathbf{I}(f(a);\mathcal{H}) + \mathbf{K}(a|f(a), \mathbf{K}(f(a))) - \mathbf{K}(a|f(a), \mathbf{K}(f(a)|\mathcal{H}), \mathcal{H})\\
&=^+ \mathbf{I}(f(a);\mathcal{H}) + \mathbf{K}(a|f(a), \mathbf{K}(f(a))) - \mathbf{K}(a|f(a), \mathbf{K}(f(a)), \mathbf{K}(f(a)|\mathcal{H}), \mathcal{H})\\
&>^+ \mathbf{I}(f(a);\mathcal{H}).
\end{aligned}$$

□

Lemma 3 *A relation* $X = \{(x_i, c_i)\}_{i=1}^{2^n} \subset \{0,1\}^* \times \mathbb{N}$, $|\mathbf{K}(x_i) - c_i| \leq s$, *has* $n <^{\log} 2s + 2\mathbf{I}(X;\mathcal{H})$.

Proof. We relativize the universal Turing machine to (n, s), which can be done due to the precision of the theorem. Let $T = \min\{t : \lceil -\log \mathbf{m}_t(x_i) \rceil - c_i < s + 1\}$. Let $N = \mathbf{bb}^{-1}(T)$ and $M = \mathbf{bb}(N)$. So for all x_i, $-\log \mathbf{m}_M(x_i) - \mathbf{K}(x_i) <^+ 2s$. Let Q be an elementary probability measure that realizes $\mathbf{Ks}(X)$ and $d = \max\{\mathbf{d}(X|Q), 1\}$. Without loss of generality, the support of Q is restricted to binary relations $B \subset \{0,1\}^* \times \mathbb{N}$ of size 2^n. Let $B_1 = \bigcup\{y : (y, c) \in B\}$. Let $S = \bigcup\{B_1 : B \in \mathrm{Support}(Q)\}$. We randomly select each string in S to be in a set R independently with probability $d2^{-n}$. Thus $\mathbf{E}[\mathbf{m}_M(R)] \leq d2^{-n}$. For $B \in \mathrm{Support}(Q)$,

$$\mathbf{E}_R \mathbf{E}_{B\sim Q}[[R \cap B_1 = \emptyset]] = \mathbf{E}_{B\sim Q} \Pr(R \cap B_1 = \emptyset) = (1 - d2^{-n})^{2^n} < e^{-d}.$$

Thus there exists a set $R \subseteq S$ such that $\mathbf{m}_M(R) \leq 2{\cdot}2^{-n}$ and $\mathbf{E}_{B\sim Q}[[R \cap B_1 = \emptyset]] < 2e^{-d}$. Let $t(B) = .5[R \cap B_1 = \emptyset]2^d$. t is a Q-test, with $\mathbf{E}_{B\sim Q}[t(B)] \leq 1$. It must be that $t(X) \neq 0$, otherwise,

$$1.44d - 1 < \log t(X) <^+ \mathbf{d}(X|Q) + \mathbf{K}(t|Q) <^+ d + \mathbf{K}(d),$$

which is a contradiction for large enough d, which one can assume without loss of generality. Thus $t(X) \neq 0$ and $R \cap X_1 \neq \emptyset$. Furthermore, if $y \in R$, $\mathbf{K}(y) <^+ -\log \mathbf{m}_M(x) - n + \log d + \mathbf{K}(d, M, R)$. So for $x \in R \cap X_1$, .

$$\begin{aligned}
\mathbf{K}(x) &<^+ -\log \mathbf{m}_M(x) - n + \log d + \mathbf{K}(d, M, R)\\
\mathbf{K}(x) &<^+ \mathbf{K}(x) + 2s - n + \log d + \mathbf{K}(M) + \mathbf{K}(R, d)\\
n &<^+ 2s + \mathbf{K}(M) + \log d + \mathbf{K}(Q, d)\\
n &<^+ 2s + \mathbf{K}(\Omega[N]) + \mathbf{Ks}(X)\\
n &<^+ 2s + \mathbf{K}(\Omega[N]) + \mathbf{I}(X;\mathcal{H}) \qquad (1)
\end{aligned}$$

From Lemma 1, $\mathbf{K}(\Omega[N]|T,N) =^+ \mathbf{K}(\Omega[N]|X,N) = O(1)$. Furthermore it is well known for the bits of Chaitin's Omega, $N <^+ \mathbf{K}(\Omega[N])$ and $\mathbf{K}(\Omega[N]|\mathcal{H}) <^+ \mathbf{K}(N)$. So, using Lemma 2,

$$N <^+ \mathbf{K}(\Omega[N]) <^{\log} \mathbf{I}(\Omega[N];\mathcal{H}) <^{\log} \mathbf{I}(X;\mathcal{H}) + \mathbf{K}(N) <^{\log} \mathbf{K}(X;\mathcal{H}). \tag{2}$$

So combining Equations 1 and 2, one gets

$$n <^{\log} 2s + 2\mathbf{I}(X;\mathcal{H}).$$

□

3 Results

Theorem 1 *If $\mathbf{w}$ is a semimeasure on $\{0,1\}^*$ and $\mathbf{m} \overset{*}{<} \mathbf{w}$ then $\mathbf{I}(\mathbf{w}:\mathcal{H}) = \infty$.*

Proof. Note that $\mathbf{w}$ has full support since $\mathbf{m}$ does. One can also assume that for each $x \in \{0,1\}^*$, $-\log \mathbf{w}(x) \in \mathbb{N}$. Let $N \subseteq \mathbb{N}$ be a set of numbers n such that $\mathbf{w}(\{0,1\}^n) < 1/n$. Obviously $|N| = \infty$. Fix $n \in N$. We have $X_n = \{x : \mathbf{w}(x) < 2^{-n-\log n+O(1)}\}$. Some simple math shows that $|X_n| \overset{*}{>} 2^n$. So for each $x \in X_n$, $\mathbf{K}(x) >^+ -\log \mathbf{w}(x) >^+ n + \log n$. We also have that for each $x \in \{0,1\}^n$, $\mathbf{K}(x) <^+ n + \mathbf{K}(n)$. Let $Y_n = \{(x, n+\log n) : x \in X_n\}$. So for each $(x,c) \in Y_n$, $|\mathbf{K}(x) - c| <^+ \log n$. So applying Lemma 3 to Y_n, we get $n <^{\log} \mathbf{I}(Y_n;\mathcal{H}) <^{\log} \mathbf{I}(\mathbf{w}:\mathcal{H}) + \mathbf{K}(n) <^{\log} \mathbf{I}(\mathbf{w}:\mathcal{H})$. Since $|N| = \infty$, $\mathbf{I}(\mathbf{w}:\mathcal{H}) = \infty$. □

A measure M over $\{0,1\}^\infty$ is neutral if $\mathbf{t}(M,\alpha) < 1$ for all $\alpha \in \{0,1\}^\infty$, where $\mathbf{t}$ is a universal uniform test, defined by any of the references [HR09, G2́1, DJ13]. For a measure μ over $\{0,1\}^\infty$, let $\mathbf{I}(\mu:\mathcal{H})$ be defined analagously to the definition of information of semi measures with $\mathcal{H}$ in the introduction.

Corollary 1 *If M is a neutral measure then $\mathbf{I}(M:\mathcal{H}) = \infty$.*

Proof. Let $t(\alpha) = \sup_n \log \mathbf{m}(\alpha[1..n])/M(\alpha[1..n])$ for all $\alpha \in \{0,1\}^\infty$. Using any of these definitions for $\mathbf{t}$, since $t(\alpha) = O(1)$, $\mathbf{m}(\langle x\rangle)/M(\langle x\rangle) = O(1)$ for all $x \in \{0,1\}^*$. Let $\mathbf{w}(x) = M(\langle x\rangle)$. Thus $\mathbf{w}(x)$ is a semi-measure computable from M and $\mathbf{m} \overset{*}{<} \mathbf{w}$. Thus due to Theorem 1, $\infty = \mathbf{I}(\mathbf{w}:\mathcal{H}) = \mathbf{I}(M:\mathcal{H})$

There exists another (stronger) means to prove this result. A measure M is weakly neutral if $\mathbf{t}(M,\alpha) < \infty$ for all $\alpha \in \{0,1\}^\infty$. In [DJ13] it was proved that any representation of weakly neutral M computes a PA degree and due to [Lev13] this implies $\mathbf{I}(M:\mathcal{H}) = \infty$.

References

[DJ13] A. Day and Miller. J. Randomness for non-computable measures. *Transactions of the American Mathematical Society*, 365(7):3575–3591, 2013.

[G2́1] Peter Gács. Lecture notes on descriptional complexity and randomness. *CoRR*, abs/2105.04704, 2021.

[HR09] M. Hoyrup and C. Rojas. Computability of probability measures and martin-löf randomness over metric spaces. *Information and Computation*, 207(7):830–847, 2009.

[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.