A Chain Rule for Randomness Deficiency

Samuel Epstein*

October 13, 2023

Abstract

This paper is an exposition of the addition equality theorem for algorithmic entropy in [Ǵ01], applied to the specific case of infinite sequences. This application implies that randomness deficiency of infinite sequences obeys the chain rule, analgous to the finite Kolmogorov complexity case. This is a generalization of van Lambalgen's Theorem. It is unclear whether this result is folklore, but in any case this paper presents a dedicated proof of the equality.

1 Introduction

Prefix free Kolmogorov complexity, $\mathbf{K}$, obeys the chain rule, with for $x, y \in \{0,1\}^*$,

$$\mathbf{K}(x,y) =^+ \mathbf{K}(x) + \mathbf{K}(y|x, \mathbf{K}(x)).$$

In this paper, we apply the addition equality theorem for algorithmic entropy in [Ǵ01] to the specific case of infinite sequences. The consequence to this is a result about randomness deficiency $\mathbf{D}$, where for computable probability μ, for infinite sequences, $\mathbf{D}(\alpha|\mu, x) = \sup_n -\log \mu(\alpha[0..n] - \mathbf{K}(\alpha[0..n]|x)$. The randomness deficiency over the space $\{0,1\}^\infty \times \{0,1\}^\infty$, is $\mathbf{D}(\alpha, \beta|\mu, \nu) = \sup_n -\log \mu(\alpha[0..n]) - \log \nu(\beta[0..n]) - \mathbf{K}(\alpha[0..n]\beta[0..n])$. The discrete case for $\mathbf{d}(x|p) = -\log p(x) - \mathbf{K}(x)$ is trivial. The result detailed in this paper is as follows.

Theorem. **([Ǵ01])** *Relativized to probabilities μ and ν over $\{0,1\}^\infty$,*

$$\mathbf{D}(\alpha, \beta|\mu, \nu) =^+ \mathbf{D}(\alpha|\mu) + \mathbf{D}(\beta|\nu, (\alpha, \lceil \mathbf{D}(\alpha|\mu) \rceil)).$$

This is a generalization of van Lambalgen's Theorem, which states (α, β) is ML random iff α is ML random and β is ML random with respect to α. If one were to take the complexities of the probabilities μ and ν into account (that is, they are no longer $O(1)$) then the theorem statement and proofs become more nuanced. This generalization can be seen in [Ǵ01].

2 Results

As shown in [Ǵ01], $2^{\mathbf{D}(\alpha|\mu)} \stackrel{*}{=} \mathbf{t}_\mu(\alpha)$ where $\mathbf{t}_\mu$ is a universal lower computable μ-test. Furthermore, similar arguments can be used to show that $2^{\mathbf{D}(\alpha,\beta|\mu,\nu)} \stackrel{*}{=} \mathbf{t}_{\mu,\nu}(\alpha, \beta)$, where $\mathbf{t}_{\mu,\nu}$ is a universal lower computable test over $\{0,1\}^\infty \times \{0,1\}^\infty$. For measure μ and lower continuous function f over $\{0,1\}^\infty$, we use the notation $\mu^x f(x) = \int_{x \in \{0,1\}^\infty} f(x) d\mu(x)$. Throughout this section, the universal Turing machine is assumed to be relativized to probabilities μ and ν over $\{0,1\}^\infty$. This means that there is an $O(1)$ sized program that can compute $\mu(x\{0,1\}^\infty)$ uniformly in $x \in \{0,1\}^*$, and similarly for ν.

*JP Theory Group. samepst@jptheorygroup.org

Proposition 1 $-\mathbf{D}(x|\mu) <^+ -\log \nu^y 2^{\mathbf{D}(x,y|\mu,\nu)}$.

Proof. Let $f(x,\mu,\nu) = -\log \nu^y 2^{\mathbf{D}(x,y|\mu,\nu)}$. The function f is upper computable and has $\mu^x 2^{-f(x,\mu,\nu)} \leq 1$. The proposition follows from the universal properties of $\mathbf{t}_\mu$, where $2^{-f} \stackrel{*}{<} \mathbf{t}_\mu$. □

Proposition 2 *For a computable function* $f : N^2 \to \mathbb{N}$,

$$-\mathbf{D}(x|\mu, y) <^+ \mathbf{K}(z) - \mathbf{D}(x|\mu, f(y,z)).$$

Proof. The function

$$g_\mu(x,y) = \sum_z 2^{\mathbf{D}(x|\mu,f(y,z))-\mathbf{K}(z)},$$

is lower computable and $\mu^x g_\mu(x,y) \leq \sum_z 2^{-\mathbf{K}(z)} \leq 1$. So $g_\mu(x,y) \stackrel{*}{<} 2^{\mathbf{D}(x|\mu,y)}$. The left hand side is a summation, so the inequality holds for each element of the sum, proving the proposition. □

Proposition 3 *If* $i < j$, *then*

$$i - \mathbf{D}(x|\mu, i) <^+ j - \mathbf{D}(x|\mu, j).$$

Proof. Using Proposition 2, with $f(i,n) = i + n$, we have

$$-\mathbf{D}(x|\mu, i) + \mathbf{D}(x|\mu, j) <^+ \mathbf{K}(j-i) <^+ j - i.$$

□

Definition 1 *Let* $F : \{0,1\}^\infty \to \mathbb{Z}\cup\{-\infty,\infty\}$ *be an upper semicomputable function. An* (μ, F)*-test is a function* $t : \{0,1\}^\infty \times \{0,1\}^\infty \to \mathbb{Z} \cup \{-\infty,\infty\}$ *that is lower semicomputable and* $\mu^x t(x,y) \leq 2^{-F(y)}$. *There exists a maximal* (μ, F) *test,* $\mathbf{t}_{(\mu,F)}$, *such that* $t \stackrel{*}{<} \mathbf{t}_{(\mu,F)}$.

Proposition 4 *Let* $F : \{0,1\}^\infty \to \mathbb{Z}| \cup \{-\infty,\infty\}$ *be an upper semicomputable function,. For all* x *and with* $\mathbf{t}_{(\nu,F)}(y) > -\infty$,

$$\mathbf{t}_{(\nu,F)}(x,y) \stackrel{*}{=} 2^{-F(y)} \mathbf{t}_\nu(x|y, F(y)).$$

Proof. To prove the inequality $\stackrel{*}{>}$, let $g(x,y,m) = \max_{i\geq m} 2^{-i} \mathbf{t}_\nu(x|y,i)$. This function is lower computable, and decreasing in m. Let $g(x,y) = g_\nu(x,y,F(y))$ is lower semicomputable since F is upper semi-computable. The multiplicative form of Proposition 3 implies

$$g(x,y,m) \stackrel{*}{=} 2^{-m} \mathbf{t}_\nu(x|y,m)$$
$$g(x,y) \stackrel{*}{=} 2^{-F(y)} \mathbf{t}_\nu(x|y,F(y)).$$

Since $\mathbf{t}_\nu$ is a test:

$$\nu^x 2^{-m} \mathbf{t}_\nu(x|y,m) \leq 2^{-m}$$
$$\nu^x g(x,y) \stackrel{*}{<} 2^{-F(y)},$$

which implies $g(x,y) \stackrel{*}{<} \mathbf{t}_{(\nu,F)}(x,y)$ by the optimality of $\mathbf{t}_{(\nu,F)}$. We now consider the upper bound. Let $\mathbf{t}'_{(\nu,F)}(x,y,m)$ be the modification of $\mathbf{t}_{(\nu,F)}$, which is a lower computable function such that

$\nu^x \mathbf{t}'_{(\nu,F)}(x,y,m) \leq 2^{-m+1}$ and if $\nu^x \mathbf{t}_{(\nu,F)}(x,y) \leq 2^{-m}$ then $\mathbf{t}'_{(\nu,F)}(x,y,m) = \mathbf{t}_{(\nu,F)}(x,y)$. The function $2^{m-1}\mathbf{t}'_{(\nu,F)}(x,y,m)$ is a test conditioned on y,m so it has $\stackrel{*}{<} \mathbf{t}_\nu(x|y,m)$. Substituting $F(y)$ for m, we have that $\nu^x \mathbf{t}_{(\nu,F)} \leq 2^{-m}$ and so

$$\mathbf{t}_{(\nu,F)}(x,y) = \mathbf{t}'_{(\nu,F)}(x,y,F_\nu(y)) \stackrel{*}{<} 2^{-F(y)+1}\mathbf{t}_\nu(x|y,F(y)).$$

□

Theorem 1 ([GÓ1]) *Relativized to computable probabilities μ and ν over $\{0,1\}^\infty$,*

$$\mathbf{D}(x,y|\mu,\nu) =^+ \mathbf{D}(x|\mu) + \mathbf{D}(y|\nu,(x,\lceil \mathbf{D}(x|\mu)\rceil)).$$

Proof. We first prove the $<^+$ inequality. Let $G(x,y,m) = \min_{i\geq m} i - \mathbf{D}((y|\nu,(x,i))$, which is upper computable and increasing in m. So the function

$$G(x,y) = G(x,y,\lceil -\mathbf{D}(x|\mu)\rceil).$$

which is also upper computable because m is replaced with an upper computable function $\lceil -\mathbf{D}(x|\mu)\rceil$. Proposition 2 implies

$$G(x,y,m) =^+ m - \mathbf{D}(y|\nu,(x,m)),$$
$$G(x,y) =^+ -\mathbf{D}(x|\mu) - \mathbf{D}(y|\nu,(x,\lceil -\mathbf{D}_\mu(x|\nu)\rceil)).$$

So

$$\nu^y 2^{-m+\mathbf{H}(y|\nu,(x,m))} \leq 2^{-m}$$
$$\nu^y 2^{-G(x,y)} \stackrel{*}{<} 2^{\mathbf{D}(x|\mu)}.$$

Integrating over x gives $\mu^x\nu^y 2^{-G(x,y)} \stackrel{*}{<} 1$, implying $-\mathbf{D}(x,y|\mu,\nu) <^+ G(x,y)$.

To prove the $>^+$ inequality, let $f(x,y) = 2^{\mathbf{D}(x,y|\mu,\nu)}$. Proposition 1 implies there exists $c \in \mathbb{N}$ with $\nu^y f(x,y) \leq 2^{\mathbf{D}(x|\mu)+c}$. Let $F(x,\mu) = \lceil -\mathbf{D}(x|\mu)\rceil$. Note that if h is a lower computable function such that $\nu^y h(x,y) \stackrel{*}{<} 2^{\mathbf{D}(x|\mu)}$, then $\mu^x\nu^y h(x,y) \stackrel{*}{<} \mu^x \mathbf{t}_\mu(x) \stackrel{*}{<} 1$, so $h \stackrel{*}{<} f$, so f is a universal F-test. Proposition 4 (substituting y for x and (x,μ) for y) gives

$$-\mathbf{D}(x,y|\mu,\nu) = -\log f(x,y) >^+ F(x) - \mathbf{D}(y|\nu,(x,F(x))).$$

□

References

[GÓ1] P. Gács. Quantum Algorithmic Entropy. *Journal of Physics A Mathematical General*, 34(35), 2001.