

# 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 [G01], applied to the specific case of infinite sequences. This application implies that randomness deficiency of infinite sequences obeys the chain rule, analogous 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 [G01] to the specific case of infinite sequences. The consequence to this is a result about randomness deficiency  $\mathbf{D}$ , where for computable probability  $\mu$ , 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.** ([G01]) *Relativized to probabilities  $\mu$  and  $\nu$  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  $(\alpha, \beta)$  is ML random iff  $\alpha$  is ML random and  $\beta$  is ML random with respect to  $\alpha$ . If one were to take the complexities of the probabilities  $\mu$  and  $\nu$  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 [G01].

## 2 Results

As shown in [G01],  $2^{\mathbf{D}(\alpha|\mu)} \stackrel{*}{=} \mathbf{t}_\mu(\alpha)$  where  $\mathbf{t}_\mu$  is a universal lower computable  $\mu$ -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  $\{0, 1\}^\infty \times \{0, 1\}^\infty$ . This can be seen in Proposition 1. For measure  $\mu$  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  $\mu$  and  $\nu$  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  $\nu$ . The following proposition generalizes Theorem 2.3.4 to the  $\{0, 1\}^\infty \times \{0, 1\}^\infty$  space.

---

\*JP Theory Group. samepst@jpththeorygroup.org

**Proposition 1**  $2^{d(\alpha, \beta | \mu, \nu)} \stackrel{*}{=} \mathbf{t}_{\mu, \nu}(\alpha, \beta)$ .

**Proof.** For the  $<^*$  inequality, note that

$$\begin{aligned} f_{\mu, \nu}(\alpha, \beta) &= \sum_{n=1}^{\infty} \sum_{x, y \in \{0, 1\}^n} 1_{x\{0, 1\}^\infty, y\{0, 1\}^\infty}(\alpha, \beta) \frac{\mathbf{m}(x, y)}{\mu(x)\nu(y)} \\ &= \sum_{n=1}^{\infty} \sum_{x, y \in \{0, 1\}^n} \frac{\mathbf{m}(\alpha[0..n]\beta[0..n])}{\mu(x)\nu(y)} \\ &\geq \sup_n \frac{\mu(\alpha[0..n]\beta[0..n])}{\mu(\alpha[0..n])\nu(\beta[0..n])}. \end{aligned}$$

So

$$\log \mathbf{t}_{\mu, \nu}(\alpha, \beta) >^+ \log f_{\mu, \nu}(\alpha, \beta) >^+ \mathbf{D}(\alpha, \beta | \mu, \nu).$$

The  $>^+$  inequality is as follows. Note that  $1_{x, y}(\alpha, \beta) = 1_{x\{0, 1\}^\infty, y\{0, 1\}^\infty}(\alpha, \beta)$ . Since  $d(\alpha, \beta) = \lceil \log \mathbf{t}_{\mu, \nu}(\alpha, \beta) \rceil$  takes only integer values and is lower computable,

$$2^{d(\alpha, \beta)} = \sup_i 2^{k_i} 1_{x_i, y_i}(\alpha, \beta) \geq \frac{1}{2} \sum_i 2^{k_i} 1_{x_i, y_i}(\alpha, \beta),$$

with the property that for all  $i$ ,  $\|x_i\| = \|y_i\|$ , and if  $1_{x_i, x_j}(\alpha, \beta) = 1_{x_j, x_j}(\alpha, \beta) = 1$  and  $i < j$  then  $k_i < k_j$ . Let  $\gamma(x, y) = \sum_{y_i=y, x_i=x} 2^{k_i}$ .

$$\sum_i 2^{k_i} 1_{x_i, y_i}(\alpha, \beta) = \sum_{n=1}^{\infty} \sum_{x, y \in \{0, 1\}^n} 1_{x, y}(\alpha, \beta) \gamma(x, y).$$

Since  $\mu^x \nu^y 2^{d(\alpha, \beta)} \leq 2$ , we have that  $\sum_{x, y \in \{0, 1\}^*} \mu(x)\nu(y)\gamma(x, y) \leq 2$ , so  $\mu(x)\nu(y)\gamma(x, y) \stackrel{*}{<} \mathbf{m}(x, y)$  and  $\gamma(x, y) = 0$  if  $\|x\| \neq \|y\|$ . So

$$2^{d(\alpha, \beta)} \stackrel{*}{<} \sup_{n \geq 1, x, y \in \{0, 1\}^n} 1_{x, y}(\alpha, \beta) \frac{\mathbf{m}(x, y)}{\mu(x)\nu(y)} \stackrel{*}{=} \sup_n \frac{\mathbf{m}(\alpha[0..n]\beta[0..n])}{\mu(\alpha[0..n])\nu(\beta[0..n])}.$$

□

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

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

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

**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 4** If  $i < j$ , then

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

**Proof.** Using Proposition 3, 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  $F : \{0, 1\}^\infty \rightarrow \mathbb{Z} \cup \{-\infty, \infty\}$  be an upper semicomputable function. An  $(\mu, F)$ -test is a function  $t : \{0, 1\}^\infty \times \{0, 1\}^\infty \rightarrow \mathbb{Z} \cup \{-\infty, \infty\}$  that is lower semicomputable and  $\mu^x t(x, y) \leq 2^{-F(y)}$ . There exists a maximal  $(\mu, F)$  test,  $\mathbf{t}_{(\mu, F)}$ , such that  $t <^* \mathbf{t}_{(\mu, F)}$ .

**Proposition 5** Let  $F : \{0, 1\}^\infty \rightarrow \mathbb{Z} \cup \{-\infty, \infty\}$  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 4 implies

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

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

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

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

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

□

The following Theorem is a specific case of Theorem 4.5.2 in [Gó1], to the Cantor space and with  $O(1)$  complexities for the probabilities.

**Theorem 1** Relativized to computable probabilities  $\mu$  and  $\nu$  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 3 implies

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

So

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

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 2 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  $h <^* f$ , so  $f$  is a universal  $F$ -test. Proposition 5 (swapping  $x$  and  $y$ ) gives

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

□

## References

- [G01] P. Gács. Quantum Algorithmic Entropy. *Journal of Physics A Mathematical General*, 34(35), 2001.