A Chaitin Ω number based on compressible strings

Kohtaro Tadaki

Research and Development Initiative, Chuo University
JST CREST
Tokyo, Japan

Supported by SCOPE from the Ministry of Internal Affairs and Communications of Japan
and by the Ministry of Economy, Trade and Industry of Japan

Introduction

Definition [Chaitin Ω number] Let U be an optimal prefix-free machine.

$$\Omega = \sum_{p \in \mathrm{Dom}\, U} 2^{-|p|}.$$

Chaitin proved Ω to be random by discovering the fact that the first n bits of Ω can solve the halting problem of U for inputs of length at most n.

Chaitin also defined variants of Ω as follows, and showed they are also random:

$$\sum_{s \in \{0,1\}^*} 2^{-H(s)},$$

where, $H(s)$ is the program-size complexity of s, and

$$\sum_{p \in U^{-1}(A)} 2^{-|p|} \qquad \text{and} \qquad \sum_{s \in A} 2^{-H(s)},$$

where A is an arbitrary infinite r.e. subset of $\{0,1\}^*$.

Introduction

In this talk, we introduce a new variant Θ of Chaitin Ω number as follows.

Definition

$$\Theta = \sum_{s \text{ is compressible}} 2^{-|s|}.$$

Preliminaries: Program-size Complexity

Definition [prefix-free machine] A partial recursive function $M\colon \{0,1\}^* \to \{0,1\}^*$ is called a *prefix-free machine* if $\operatorname{Dom} M$ is a prefix-free set. □

Definition For any prefix-free machine M and any $s \in \{0,1\}^*$,

$$H_M(s) := \min\left\{\,|p| \;\middle|\; p \in \{0,1\}^* \;\&\; M(p) = s\,\right\}.$$ □

Definition [optimal prefix-free machine] A prefix-free machine U is called *optimal* if, for each prefix-free machine M, there exists $d \in \mathbb{N}$ such that, for every $s \in \{0,1\}^*$,

$$H_U(s) \leq H_M(s) + d.$$ □

Definition [program-size complexity] We choose a particular optimal prefix-free machine U as a standard one. Then the *program-size complexity* (or *Kolmogorov complexity*) $H(s)$ of $s \in \{0,1\}^*$ is defined by

$$H(s) := H_U(s).$$ □

Preliminaries: Chaitin Ω Number

Definition [Chaitin randomness, Chaitin 1975]

We say $\alpha \in \mathbb{R}$ is *Chaitin random* if $n \le H(\alpha\restriction_n) + O(1)$ for all $n \in \mathbb{N}^+$.

Here, $\alpha\restriction_n$ is the first n bits of the base-two expansion of α. □

Definition [Chaitin Ω number, Chaitin 1975]

$$\Omega := \sum_{p \in \mathrm{Dom}\, U} 2^{-|p|}.$$ □

- If $\Omega\restriction_n$ is given, then one can calculate the list of all halting inputs for U of length at most n (i.e., $\mathrm{Dom}\, U\restriction_n$).
- If $\Omega\restriction_n$ is given, then one can calculate a string $s_n \in \{0,1\}^*$ with $H(s_n) > n$.

Theorem [Chaitin 1975] Ω is Chaitin random. □

New Variant of Chaitin Ω Number

Compressible Strings

Definition [compressible string] A string $s \in \{0,1\}^*$ is called *compressible* if $H(s) < |s|$ (i.e., if s is 1-compressible: $H(s) \leq |s| - 1$). □

Fact For every $n \in \mathbb{N}$, there exists an incompressible string of length n.

Proof)

The number of strings of length less than n is $2^n - 1$ while the number of strings of length n is 2^n. □

New Variant of Chaitin Ω Number

Definition [new variant Θ of Chaitin Ω number]

$$\Theta := \sum_{H(s)<|s|} 2^{-|s|},$$

where the sum is over all compressible strings s. □

First of all, we have to check the convergence of Θ.

$$\Theta < \sum_{H(s)<|s|} 2^{-H(s)} \leq \sum_{s\in\{0,1\}^*} 2^{-H(s)} \leq \sum_{p\in \mathrm{Dom}\, U} 2^{-|p|} = \Omega < 1.$$

Remark Note that

$$\sum_{H(s)\geq|s|} 2^{-|s|} = \infty,$$

where the sum is over all incompressible strings s. This is because

$$\sum_{H(s)<|s|} 2^{-|s|} + \sum_{H(s)\geq|s|} 2^{-|s|} = \sum_{s\in\{0,1\}^*} 2^{-|s|} = \sum_{n=0}^{\infty} 2^n 2^{-n} = \infty.$$ □

Randomness of New Variant I

• If $\Theta\restriction_n$ is given, then one can calculate the list of all compressible strings of length at most n.

• If $\Theta\restriction_n$ is given, then one can calculate a string $s_n \in \{0,1\}^*$ with $H(s_n) > n$.

Theorem Θ is Chaitin random. □

Randomness of New Variant II

Proof of the theorem) Let $s_1, s_2, s_3, \ldots\ldots$ be a particular recursive enumeration of the r.e. set $\{\, s \mid H(s) < |s| \,\}$. Then $\Theta = \sum_{i=1}^{\infty} 2^{-|s_i|}$.

Procedure: Given $\Theta\restriction_n$, one can effectively find k_0 which satisfies

$$0.(\Theta\restriction_n) < \sum_{i=1}^{k_0} 2^{-|s_i|}.$$

This is possible because $0.(\Theta\restriction_n) < \Theta$ and $\lim_{k\to\infty} \sum_{i=1}^{k} 2^{-|s_i|} = \Theta$. It follows that

$$\sum_{i=k_0+1}^{\infty} 2^{-|s_i|} < 2^{-n}.$$

Hence, $n < |s_i|$ for every $i > k_0$. Thus,

$$\{\, s \mid s \text{ is compressible of length } \le n \,\} = \{\, s_1, s_2, \ldots, s_{k_0} \} \cap \{0,1\}^{\le n}.$$

Since an incompressible n bits string exists, by picking any n bits string t which is not in the above set, one can obtain $t \in \{0,1\}^*$ such that

$$H(t) \ge |t| = n.$$

$\square$

Distribution of Compressible Strings

It would be important to evaluate how many compressible n bits strings exist, i.e., to evaluate the number of elements in the set

$$\{\, s \in \{0,1\}^* \mid |s| = n \;\&\; H(s) < n \,\}.$$

Theorem

$$\#\{\, s \in \{0,1\}^* \mid |s| = n \;\&\; H(s) < n \,\} = 2^{n-H(n)+O(1)}.$$

□

Remark Solovay (1975) showed that

$$\#\{\, s \in \{0,1\}^* \mid H(s) < n \,\} = 2^{n-H(n)+O(1)}.$$

The above theorem slightly improves Solovay's result. □

A Genararization of Θ

Definition

$$\Theta_a := \sum_{H(s) \le |s| - a} 2^{-|s|}$$

for any $a \in \mathbb{Z}$. Here, the sum is over all a-compressible strings s. □

In the case of $a = 1$, $\Theta_1 = \Theta$.

Let $a \in \mathbb{Z}$. It is easy to show that, for all sufficiently large $n \in \mathbb{N}$, there exists an n bits string s such that $H(s) > |s| - a$ (i.e., s is a-incompressible).

For example, this follows from the Solovay's result.

Thus, based on this fact, we can show the following theorem in the same manner as the proof of the randomness of Θ.

Theorem Θ_a is Chaitin random for every $a \in \mathbb{Z}$. □

Another Proof of the Randomness of Θ_a based on Universal Martin-Löf Test

Universal Martin-Löf Test

For any subset G of $\{0,1\}^*$, the subset $I(G)$ of $[0,1)$ is defined by

$$I(G) = \bigcup_{s \in G} I(s),$$

where $I(s) = [0.s, 0.s + 2^{-|s|})$.

Definition [Martin-Löf randomness] A subset $\mathcal{C}$ of $\mathbb{N}^+ \times \{0,1\}^*$ is called a *Martin-Löf test* if $\mathcal{C}$ is an r.e. set and

$$\forall n \in \mathbb{N}^+ \quad \mathcal{L}(I(\mathcal{C}_n)) \leq 2^{-n},$$

where $\mathcal{L}$ is Lebesgue measure on $\mathbb{R}$ and $\mathcal{C}_n = \left\{ s \;\middle|\; (n,s) \in \mathcal{C} \right\}$.

For any $\alpha \in \mathbb{R}$, we say that α is *Martin-Löf random* if for every Martin-Löf test $\mathcal{C}$, there exists $n \in \mathbb{N}^+$ such that $\alpha - \lfloor \alpha \rfloor \notin I(\mathcal{C}_n)$. □

Definition [universal Martin-Löf test] A Martin-Löf test $\mathcal{U}$ is called *universal* if

$$\bigcap_{n=1}^{\infty} I(\mathcal{C}_n) \subset \bigcap_{n=1}^{\infty} I(\mathcal{U}_n)$$

for every Martin-Löf test $\mathcal{C}$. □

Another Proof I

We give another proof of the randomness of Θ_n with $n \in \mathbb{N}^+$, based on the property of universal Martin-Löf test.

On the one hand,

Theorem [Kučera & Slaman 2001] Let $\mathcal{U}$ be a universal Martin-Löf test. Then $\mathcal{L}(I(\mathcal{U}_n))$ is Chaitin random for every $n \in \mathbb{N}^+$. □

On the other hand,

Theorem [Schnorr 1973] For every $\alpha \in \mathbb{R}$, α is Martin-Löf random if and only if α is Chaitin random. □

In other words,

Theorem [Calude's book, Nies's book] The set

$$\mathcal{R} = \{\, (n, s) \in \mathbb{N}^+ \times \{0, 1\}^* \mid H(s) \leq |s| - n \,\}$$

is a universal Martin-Löf test. □

Another Proof II

Theorem Θ_n is Chaitin random for every $n \in \mathbb{N}^+$. □

Another proof of the above theorem)

Let $n \in \mathbb{N}^+$. By the previous two theorems, $\mathcal{L}(I(\mathcal{R}_n))$ is Chaitin random. Since the set of all n-compressible strings does form a prefix-free set, note that

$$\Theta_n := \sum_{H(s) \le |s| - n} 2^{-|s|}$$

differs from

$$\mathcal{L}(I(\mathcal{R}_n)) := \mathcal{L}(I(\{\, s \mid H(s) \le |s| - n \,\})).$$

However, we can show that

$$\Theta_n = \mathcal{L}(I(\mathcal{R}_n)) + \gamma$$

for some left-computable real γ. Since Θ_n and $\mathcal{L}(I(\mathcal{R}_n))$ are left-computable, the result follows. □

Two Generalizations of Θ to a Partial Random Real

Partial Randomness

$$\Theta := \sum_{H(s)<|s|} 2^{-|s|}$$ is a random real.

By introducing real parameter T with $0 < T \leq 1$ to Θ, we can introduce partial random reals whose compression rate is T in the following two manner.

Generalization $\Theta(T)$ of Θ to a Partial Random Real

Definition [frist generalization of Θ]

$$\Theta(T) := \sum_{H(s)<|s|} 2^{-\frac{|s|}{T}}$$

for each real $T > 0$. □

In the case of $T = 1$, $\Theta(1) = \Theta$.

Theorem

(i) If $0 < T < 1$ and T is computable, then

$$H(\Theta(T)\restriction_n) = Tn + O(1)$$

and therere

$$\lim_{n\to\infty} \frac{H(\Theta(T)\restriction_n)}{n} = T.$$

(ii) If $1 < T$, then $\Theta(T)$ diverges to ∞. □

Generalization $\overline{\Theta}(T)$ of Θ to a Partial Random Real

Definition [second generalization of Θ]

$$\overline{\Theta}(T) := \sum_{H(s)<T|s|} 2^{-|s|}$$

for each real $T > 0$. □

In the case of $T = 1$, $\overline{\Theta}(1) = \Theta$.

Theorem

(i) If $0 < T < 1$ and T is computable, then

$$H(\overline{\Theta}(T)\restriction_n) = Tn + O(1)$$

and therere

$$\lim_{n\to\infty} \frac{H(\overline{\Theta}(T)\restriction_n)}{n} = T.$$

(ii) If $1 < T$, then $\overline{\Theta}(T)$ diverges to ∞. □

Summary

Definition [new variant Θ of Chaitin Ω number]

$$\Theta = \sum_{H(s)<|s|} 2^{-|s|}.$$

Theorem Θ is Chaitin random.