Properties of fibers of optimal prefix-free machines

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

Abstract

The optimal prefix-free machine U is a universal decoding algorithm used to define the notion of program-size complexity $H(s)$.

In this talk, we investigate the fiber of an optima prefix-free machine U at a string $s \in \{0,1\}^*$, i.e., the set

$$U^{-1}(s) := \{\, p \mid U(p) = s \,\}.$$

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 $\mathrm{Dom}\, M$ 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 $s \in \{0,1\}^*$,

$$H_U(s) \le H_M(s) + d.$$ □

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

$$H(s) := H_U(s).$$ □

Aim of This Talk

The fiber $U^{-1}(s)$ *of an optimal prefix-free machine* U *at a string* $s \in \{0,1\}^*$, is defined by

$$U^{-1}(s) := \{\, p \in \mathrm{Dom}\, U \mid U(p) = s \,\}.$$

The fiber $U^{-1}(s)$ is considered to be important. For example, the definition of $H(s)$ is directly related to the fiber by

$$H(s) = \min \left\{ \, |p| \;\middle|\; p \in U^{-1}(s) \, \right\}.$$

In what follows, we call elements in the fiber $U^{-1}(s)$ *programs*.

In this talk,
we first investigate the number $\#U^{-1}(s)$ of programs in the fiber $U^{-1}(s)$.
We next investigate the distribution of programs in the fiber $U^{-1}(s)$.

Number of Programs in Fibers

Number of Programs in Fibers I

Note that the fiber $U^{-1}(s)$ is an infinite set, in general.

However, our first result shows that, while keeping $H(s)$ unchanged for all s, we can modify U so that each $U^{-1}(s)$ is a finite set and moreover the number of programs in $U^{-1}(s)$ is bounded by some total recursive function $f\colon \{0,1\}^* \to \mathbb{N}$.

Theorem

For every optimal prefix-free machine V, there exists an optimal prefix-free machine W for which the following (i) and (ii) hold:

(i) $H_V(s) = H_W(s)$ for every $s \in \{0,1\}^*$.

(ii) There exists a total recursive function $f\colon \{0,1\}^* \to \mathbb{N}$ such that

$$\#W^{-1}(s) \leq f(s)$$

for every $s \in \{0,1\}^*$. □

Number of Programs in Fibers II

Theorem [posted again]

For every optimal prefix-free machine V, there exists an optimal prefix-free machine W for which the following (i) and (ii) hold:

(i) $H_V(s) = H_W(s)$ for every $s \in \{0,1\}^*$.

(ii) There exists a total recursive function $f\colon \{0,1\}^* \to \mathbb{N}$ such that

$$\#W^{-1}(s) \leq f(s)$$

for every $s \in \{0,1\}^*$.

Proof)

Let $(p_1, s_1), (p_2, s_2), (p_3, s_3), \dots$ be a particular recursive enumeration of Graph(V). Consider the total recursive function $g(s) = \min\{\, i \mid s_i = s \,\}$. Define a prefix-free machine W by

$$W^{-1}(s) = \{\, p \in V^{-1}(s) \mid |p| \leq |p_{g(s)}| \,\}.$$

By counting the number of binary strings of length at most $|p_{g(s)}|$, we see that $\#W^{-1}(s) \leq 2^{|p_{g(s)}|+1} - 1$.

It follows from $H_V(s) \leq |p_{g(s)}|$ that $H_W(s) = H_V(s)$ and therefore W is optimal. □

Number of Programs in Fibers III

The following theorem shows that the total recursive upper bound $f(s)$ in the previous theorem cannot be chosen to be tight at all.

Theorem

Let V be an optimal prefix-free machine, and let $f\colon \{0,1\}^* \to \mathbb{N}$ be a total recursive function. Suppose that $\#V^{-1}(s) \le f(s)$ for all $s \in \{0,1\}^*$. Then

$$\lim_{s\to\infty} \left\{ f(s) - \#V^{-1}(s) \right\} = \infty.$$

Here we identify $\{0,1\}^*$ with $\mathbb{N}$.

Proof) We use the technique by Meyer and Loveland to show that, for every $\alpha \in \mathbb{R}$, if $C(\alpha\restriction_n \mid n) = O(1)$ then α is computable.

Let us assume contrarily that $\lim_{s\to\infty} \left\{ f(s) - \#V^{-1}(s) \right\} \neq \infty$. Then,

$$\alpha := \liminf_{s\to\infty} \left\{ f(s) - \#V^{-1}(s) \right\}$$

is finite and is in $\mathbb{N}$. Hence, there exists $s_0 \in \{0,1\}^*$ such that $\alpha \le f(s) - \#V^{-1}(s)$ for all $s \ge s_0$ and $\alpha = f(s) - \#V^{-1}(s)$ for infinitely many $s \ge s_0$. Based on this, one can compute a string with arbitrary high program-size complexity. □

Number of Programs in Fibers IV

As a result, the following holds in particular.

Corollary Let V be an optimal prefix-free machine. Suppose that $V^{-1}(s)$ is a finite set for all $s \in \{0,1\}^*$. Then the function $\#V^{-1}(s)$ of $s \in \{0,1\}^*$ is not bounded to the above. $\square$

Number of Programs in Fibers V

On the other hand, while keeping $H(s)$ unchanged for all s, we can modify U so that each $U^{-1}(s)$ is an infinite set.

Theorem

For every optimal prefix-free machine V, there exists an optimal prefix-free machine W for which the following conditions (i) and (ii) hold:

(i) $H_V(s) = H_W(s)$ for every $s \in \{0,1\}^*$.

(ii) $W^{-1}(s)$ is an infinite set for every $s \in \{0,1\}^*$. □

Distribution of Programs in Fibers

Distribution of Programs in Fibers I

Next, we investigate the distribution of programs in the fiber $V^{-1}(s)$ of an optimal prefix-free machine V at a string $s \in \{0,1\}^*$

In 1975 Solovay showed the following result for the distribution of all programs in $\operatorname{Dom} V$ for an optimal prefix-free machine V.

Theorem [Solovay 1975]

Let V be an optimal prefix-free machine, and let

$$S_V(n) = \{\, p \mid |p| \le n \;\&\; p \in \operatorname{Dom} V \,\}.$$

Then

$$\#S_V(n) = 2^{n-H(n)+O(1)},$$

i.e., there exists $d \in \mathbb{N}$ such that

(i) $\#S_V(n) \le 2^{n-H(n)+d}$ for all $n \in \mathbb{N}$, and

(ii) $\#S_V(n) \ge 2^{n-H(n)-d}$ for all $n \in \mathbb{N}$ with $n - H(n) - d \ge 0$. □

In what follows, we replace $\operatorname{Dom} V$ by $V^{-1}(s)$ in $S_V(n)$ and investigate the properties of the resultant set $S_V(n,s)$.

Distribution of Programs in Fibers II

We define

$$S_M(n,s) := \{\, p \mid |p| \le n \;\&\; p \in M^{-1}(s) \,\}$$

for any prefix-free machine M, $n \in \mathbb{N}$, and $s \in \{0,1\}^*$.

Theorem

Let M be a prefix-free machine. Then

$$\#S_M(n,s) \le 2^{n-H(n,s)+O(1)}.$$

Here $H(n,s)$ is defined as $H(b(n,s))$ with a particular bijective total recursive function $b\colon \mathbb{N} \times \{0,1\}^* \to \{0,1\}^*$.

Proof) We use the property of universal probability (i.e., maximal c.e. discrete semimeasure).

First define $f\colon \{0,1\}^* \to [0,\infty)$ by $f(b(n,s)) = \#S_M(n,s)2^{-n-1}$. Then f is shown to be a c.e. discrete semimeasure. Since the mapping $t \mapsto 2^{-H(t)}$ is a universal probability, $f(t) \le 2^{-H(t)+O(1)}$ by its definition. $\square$

Distribution of Programs in Fibers III

It is natural to consider prefix-free machines which attain the upper bound

$$2^{n-H(n,s)+O(1)}$$

in the previous theorem. Thus, we introduce the following notion to prefix-free machines.

Definition [maximal redundancy]
A prefix-free machine M is called *maximally redundant* if

$$2^{n-H(n,s)+O(1)} \leq \#S_M(n,s),$$

i.e., if there exists $d \in \mathbb{N}$ such that

$$2^{n-H(n,s)-d} \leq \#S_M(n,s)$$

for all $n \in \mathbb{N}$ and $s \in \{0,1\}^*$ with $n-H(n,s)-d \geq 0$. Note that this condition is equivalent to the condition that there exists $d \in \mathbb{N}$ for which

$$\lfloor 2^{n-H(n,s)-d} \rfloor \leq \#S_M(n,s) \text{ for all } n \text{ and } s. \qquad \square$$

Proposition For every prefix-free machine M, M is maximally redundant if and only if $\#S_M(n,s) = 2^{n-H(n,s)+O(1)}$. $\square$

Distribution of Programs in Fibers IV

Theorem For every prefix-free machine M, if M is maximally redundant then M is optimal. □

Lemma $H(s)+n-H(H(s)+n,s)$ diverges to ∞ as $n\to\infty$ uniformly on $s\in\{0,1\}^*$. □

Proof of the theorem)

Since M is maximally redundant, there exists $d\in\mathbb{N}$ such that

$$2^{n-H(n,s)-d}\le\#S_M(n,s)$$

for all n and s with $n-H(n,s)-d\ge 0$.

On the other hand, by the above lemma, there exists $n_0\in\mathbb{N}$ such that, for every $s\in\{0,1\}^*$,

$$H(s)+n_0-H(H(s)+n_0,s)\ge d.$$

Hence,

$$1\le\#S_M(H(s)+n_0,s).$$

This implies that $H_M(s)\le H(s)+n_0$ and therefore M is optimal. □

Distribution of Programs in Fibers V

Remark Not every optimal prefix-free machine is maximally redundant.

There are many such counter examples. We saw the optimal prefix-free machine W for which $W^{-1}(s)$ is finite for all $s \in \{0,1\}^*$.

An interesting example might be the optimal prefix-free machine V such that

$$\#S_V(n,s) = \begin{cases} n - H_V(s) + 1 & \text{if } n \geq H_V(s), \\ 0 & \text{otherwise} \end{cases}$$

(Chaitin 75, Figueira, Stephan, & Wu 2005). The machine V satisfies

$$2^{-H_V(s)+1} = \Omega_V$$

(Figueira, Stephan, & Wu 2005), where

$$\Omega_V := \sum_{p \in \mathrm{Dom}\, V} 2^{-|s|}.$$

□

Distribution of Programs in Fibers VI

Theorem

There exists a maximally redundant prefix-free machine.

Proof)

Define a partial recursive function $V\colon \{0,1\}^* \to \{0,1\}^*$ by the condition that $V(p) = s$ if and only if there exists a prefix q of p such that $U(q) = b(|p|, s)$. Recall here that U is the optimal prefix-free machine used to define $H(s)$. Then we can show that V is a pref-free machine. Moreover, if $n - H(n,s) \geq 0$ then $2^{n-H(n,s)} \leq \#S_V(n,s)$. $\square$

Distribution of Programs in Fibers VII

Theorem [Calude, Hertling, Khoussainov and Wang 2001, Kučera and Slaman 2001] For every $\alpha \in (0,1)$, the following conditions are equivalent:

(i) α is left-computable and Martin-Löf random.

(ii) $\alpha = \Omega_V$ for some optimal prefix-free machine V. □

We can give a characterization of left-computable random reals by maximally redundant prefix-free machines as follows.

Theorem

For every $\alpha \in (0,1)$, the following conditions are equivalent:

(i) α is left-computable and Martin-Löf random.

(ii) $\alpha = \Omega_V$ for some maximally redundant prefix-free machine V. □