CDMTCS
Research
Report
Series

Properties of Optimal Prefix-Free Machines as Instantaneous Codes

Kohtaro Tadaki
Chuo University, Japan

CDMTCS-389
July 2010

Centre for Discrete Mathematics and
Theoretical Computer Science

Properties of Optimal Prefix-Free Machines as Instantaneous Codes

Kohtaro Tadaki
Research and Development Initiative, Chuo University
1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan
Email: tadaki@kc.chuo-u.ac.jp WWW: http://www2.odn.ne.jp/tadaki/

***Abstract*—The optimal prefix-free machine U is a universal decoding algorithm used to define the notion of program-size complexity $H(s)$ for a finite binary string s. Since the set of all halting inputs for U is chosen to form a prefix-free set, the optimal prefix-free machine U can be regarded as an instantaneous code for noiseless source coding scheme. In this paper, we investigate the properties of optimal prefix-free machines as instantaneous codes. In particular, we investigate the properties of the set $U^{-1}(s)$ of codewords associated with a symbol s. Namely, we investigate the number of codewords in $U^{-1}(s)$ and the distribution of codewords in $U^{-1}(s)$ for each symbol s, using the toolkit of algorithmic information theory.**

I. Introduction

Algorithmic information theory (AIT, for short) is a framework for applying information-theoretic and probabilistic ideas to recursive function theory. One of the primary concepts of AIT is the *program-size complexity* (or *Kolmogorov complexity*) $H(s)$ of a finite binary string s, which is defined as the length of the shortest binary input for a universal decoding algorithm U, called an *optimal prefix-free machine*, to output s. By the definition, $H(s)$ can be thought of as the information content of the individual finite binary string s. In fact, AIT has precisely the formal properties of normal information theory (see Chaitin [1]). On the other hand, $H(s)$ can also be thought to represent the amount of randomness contained in a finite binary string s, which cannot be captured in a computational manner. In particular, the notion of program-size complexity plays a crucial role in characterizing the *randomness* of an infinite binary string, or equivalently, a real.

The optimal prefix-free machine U is chosen so as to satisfy that the set $\operatorname{dom} U$ of all halting inputs for U forms a prefix-free set. Therefore, as considered in Chaitin [1], we can think of the optimal prefix-free machine U as a decoding equipment at the receiving end of a noiseless binary communication channel. We can regard its programs (i.e., finite binary strings in $\operatorname{dom} U$) as codewords and can regard the result of the computation by U, which is a finite binary string, as a decoded "symbol." Since $\operatorname{dom} U$ is a prefix-free set, such codewords form what is called an "instantaneous code," so that successive symbols sent through the channel in the form of concatenation of codewords can be separated.[1]

Thus, from the point of view of information theory, it is important to investigate the properties of optimal prefix-free machine as an instantaneous code. In this paper, in particular, we investigate the properties of the set $U^{-1}(s)$ of codewords associated with a symbol s, where $U^{-1}(s) = \{\, p \mid U(p) = s \,\}$. Unlike for instantaneous codes in normal information theory, the codeword p associated with each symbol s by $s = U(p)$ is not necessarily unique for optimal prefix-free machines U in AIT. We investigate this property from various aspects.

After the preliminary section, in Section III we investigate the number of codewords in $U^{-1}(s)$. We show the following: (i) While keeping $H(s)$ unchanged for all s, we can modify U so that each $U^{-1}(s)$ is a finite set, where the number of codewords in $U^{-1}(s)$ is bounded to the above by some total recursive function $f(s)$, i.e., by some computable function $f(s)$. (ii) This upper bound $f(s)$ cannot be chosen to be tight at all. (iii) As a result, even in the case where all $U^{-1}(s)$ are a finite set, the number of codewords in $U^{-1}(s)$ is not bounded to the above on all finite binary strings s. (iv) While keeping $H(s)$ unchanged for all s, we can modify U so that each $U^{-1}(s)$ is an infinite set. In Section IV, we then investigate the distribution of codewords in $U^{-1}(s)$. We estimate the distribution using the notion of program-size complexity, and then show that the estimation is tight.

II. Preliminaries

A. Basic Notation

We start with some notation about numbers and strings which will be used in this paper. $\#S$ is the cardinality of S for any set S. $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ is the set of natural numbers, and $\mathbb{N}^+$ is the set of positive integers. $\mathbb{Q}$ is the set of rationals, and $\mathbb{R}$ is the set of reals. Normally, $O(1)$ denotes any function $f\colon \mathbb{N}^+ \to \mathbb{R}$ such that there is $C \in \mathbb{R}$ with the property that $|f(n)| \le C$ for all $n \in \mathbb{N}^+$.

$\{0,1\}^* = \{\lambda, 0, 1, 00, 01, 10, 11, 000, \dots\}$ is the set of finite binary strings where λ denotes the *empty string*, and $\{0,1\}^*$ is ordered as indicated. We identify any string in $\{0,1\}^*$ with a natural number in this order, i.e., we consider $\varphi\colon \{0,1\}^* \to \mathbb{N}$ such that $\varphi(s) = 1s - 1$ where the concatenation $1s$ of strings 1 and s is regarded as a dyadic integer, and then we identify s with $\varphi(s)$. For any $s \in \{0,1\}^*$, $|s|$ is the *length* of s. A subset S of $\{0,1\}^*$ is called *prefix-free* if no string in S is a prefix of another string in S. For any function f, the domain of definition of f is denoted by $\operatorname{dom} f$. We write "r.e." instead of "recursively enumerable."

[1] Note that AIT does not assume the existence of an encoding algorithm E such that $E(s) = p$ if and only if $U(p) = s$.

B. Algorithmic Information Theory

In the following we concisely review some definitions and results of AIT [1], [3], [6], [4]. A *prefix-free machine* is a partial recursive function $C\colon \{0,1\}^* \to \{0,1\}^*$ such that $\operatorname{dom} C$ is a prefix-free set. For each prefix-free machine C and each $s \in \{0,1\}^*$, $H_C(s)$ is defined by

$$H_C(s) = \min \left\{ |p| \mid p \in \{0,1\}^* \ \&\ C(p) = s \right\} \quad \text{(may be } \infty\text{)}.$$

A prefix-free machine U is said to be *optimal* if for each prefix-free machine C there exists $d \in \mathbb{N}$ with the following property; if $p \in \operatorname{dom} C$, then there is q for which $U(q) = C(p)$ and $|q| \le |p| + d$. Note that a prefix-free machine U is optimal if and only if for each prefix-free machine C there exists $d \in \mathbb{N}$ such that, for every $s \in \{0,1\}^*$, $H_U(s) \le H_C(s) + d$. It is easy to see that there exists an optimal prefix-free machine. We choose a particular optimal prefix-free machine U as the standard one for use, and define $H(s)$ as $H_U(s)$, which is referred to as the *program-size complexity* of s or the *Kolmogorov complexity* of s. It follows that for every prefix-free machine C there exists $d \in \mathbb{N}$ such that, for every $s \in \{0,1\}^*$,

$$H(s) \le H_C(s) + d. \tag{1}$$

Based on this we can show that, for every partial recursive function $\Psi\colon \{0,1\}^* \to \{0,1\}^*$, there exists $d \in \mathbb{N}$ such that, for every $s \in \operatorname{dom} \Psi$,

$$H(\Psi(s)) \le H(s) + d. \tag{2}$$

Based on (1) we can also show that there exists $c \in \mathbb{N}$ such that, for every $n \in \mathbb{N}^+$,

$$H(n) \le 2\log_2 n + c. \tag{3}$$

For any $s \in \{0,1\}^*$, we define s^* as $\min\{ p \in \{0,1\}^* \mid U(p) = s\}$, i.e., the first element in the ordered set $\{0,1\}^*$ of all strings p such that $U(p) = s$. Then, $|s^*| = H(s)$ for every $s \in \{0,1\}^*$. For any $s, t \in \{0,1\}^*$, we define $H(s,t)$ as $H(b(s,t))$, where $b\colon \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$ is a particular bijective total recursive function.

AIT has precisely the formal properties of normal information theory, as demonstrated by Chaitin [1]. The program-size complexity $H(s)$ corresponds to the notion of entropy in information theory, while $H(s,t)$ corresponds to the notion of joint entropy in information theory.

The program-size complexity $H(s)$ is originally defined using the notion of program-size, as in the above. However, it is possible to define $H(s)$ without referring to such a notion. Namely, as in the following, we first introduce a *universal probability* m, and then define $H(s)$ as $-\log_2 m(s)$. A universal probability is defined as follows.

Definition 1 (universal probability, Zvonkin and Levin [8])**.** *A function $r\colon \{0,1\}^* \to [0,1]$ is called a lower-computable semi-measure if $\sum_{s\in\{0,1\}^*} r(s) \le 1$ and the set $\{(a,s) \in \mathbb{Q} \times \{0,1\}^* \mid a < r(s)\}$ is r.e. We say that a lower-computable semi-measure m is a universal probability if for every lower-computable semi-measure r, there exists $c \in \mathbb{N}^+$ such that, for all $s \in \{0,1\}^*$, $r(s) \le cm(s)$.* ■

The following theorem can be then shown (see e.g. Chaitin [1, Theorem 3.4] for its proof).

Theorem 2. *For every optimal prefix-free machine V, the function $2^{-H_V(s)}$ of s is a universal probability.* ■

For each universal probability m, by Theorem 2 we see that $H(s) = -\log_2 m(s) + O(1)$ for all $s \in \{0,1\}^*$. Thus it is possible to define $H(s)$ as $-\log_2 m(s)$ with a particular universal probability m instead of as $H_U(s)$. Note that the difference up to an additive constant is nonessential to AIT.

Normally, for each prefix-free machine C and each $s \in \{0,1\}^*$, the set $C^{-1}(s)$ is defined by

$$C^{-1}(s) = \{ p \in \operatorname{dom} C \mid C(p) = s \}.$$

Note that $V^{-1}(s) \neq \emptyset$ for every optimal prefix-free machine V and every $s \in \{0,1\}^*$.

III. The Number of Codewords

In this section, we investigate the properties of the number $\#V^{-1}(s)$ of codewords in $V^{-1}(s)$ for an optimal prefix-free machine V. In Theorem 4 below we show that, while keeping $H_V(s)$ unchanged for all s, we can modify V so that each $V^{-1}(s)$ is a finite set, where $\#V^{-1}(s)$ is bounded to the above by some total recursive function $f(s)$. Before that, we prove a more general theorem for prefix-free machines in general, as follows.

Theorem 3. *For every prefix-free machine C, there exists a prefix-free machine D for which the following conditions (i), (ii), and (iii) hold:*

(i) *$H_D(s) = H_C(s)$ for every $s \in \{0,1\}^*$.*
(ii) *$D^{-1}(s)$ is a finite set for every $s \in \{0,1\}^*$.*
(iii) *Moreover, there exists a partial recursive function $f\colon \{0,1\}^* \to \mathbb{N}^+$ such that $\#D^{-1}(s) \le f(s)$ for every $s \in \operatorname{dom} f$ and $\operatorname{dom} f = \{ s \in \{0,1\}^* \mid D^{-1}(s) \neq \emptyset \}$.*

Proof: Let C be an arbitrary prefix-free machine. We define the *graph* $\operatorname{Graph}(C)$ of C by

$$\operatorname{Graph}(C) = \{ (p,s) \in \{0,1\}^* \times \{0,1\}^* \mid C(p) = s \}.$$

Note that $\operatorname{Graph}(C)$ is an r.e. set, since $C\colon \{0,1\}^* \to \{0,1\}^*$ is a partial recursive function. In the case where $\operatorname{Graph}(C)$ is a finite set, the set $\{ s \in \{0,1\}^* \mid C^{-1}(s) \neq \emptyset \}$ is finite and the set $C^{-1}(s)$ is finite for every $s \in \{0,1\}^*$. Thus, in this case, by setting $D = C$ and $f(s) = \#C^{-1}(s)$, the conditions (i), (ii), and (iii) hold, and therefore the result follows. Hence, in what follows we assume that $\operatorname{Graph}(C)$ is an infinite set.

Let $(p_1, s_1), (p_2, s_2), (p_3, s_3), \dots$ be a particular recursive enumeration of the infinite r.e. set $\operatorname{Graph}(C)$. It is then easy to show that there exists a partial recursive function $g\colon \{0,1\}^* \to \mathbb{N}^+$ which satisfies the following two conditions:

(a) $\operatorname{dom} g = \{ s \mid \exists i \in \mathbb{N}^+\ s_i = s \}$.
(b) $g(s) = \min\{ i \in \mathbb{N}^+ \mid s_i = s \}$ for every $s \in \operatorname{dom} g$.

We then define a partial recursive function $D\colon \{0,1\}^* \to \{0,1\}^*$ by the condition that

$$D^{-1}(s) = \left\{ p_i \mid i \in \mathbb{N}^+ \ \&\ s_i = s \ \&\ |p_i| \le |p_{g(s)}| \right\}$$

if $s \in \operatorname{dom} g$ and $D^{-1}(s) = \emptyset$ otherwise. It is easy to see that such a partial recursive function D exists. By counting the number of binary strings of length at most $|p_{g(s)}|$, we see that, for each $s \in \operatorname{dom} g$, $\#D^{-1}(s) \leq 2^{|p_{g(s)}|+1} - 1$ and therefore $D^{-1}(s)$ is a finite set. Thus, the condition (ii) holds for D. Moreover, by defining a partial recursive function $f\colon \{0,1\}^* \to \mathbb{N}^+$ by the conditions that $\operatorname{dom} f = \operatorname{dom} g$ and $f(s) = 2^{|p_{g(s)}|+1} - 1$ for every $s \in \operatorname{dom} f$, the condition (iii) holds for D.

Next, we show that D is a prefix-free machine. It follows from the definition of D that

$$D^{-1}(s) \subset C^{-1}(s) \tag{4}$$

for every $s \in \{0,1\}^*$. Therefore we see that

$$\operatorname{dom} D = \bigcup_{s\in\{0,1\}^*} D^{-1}(s) \subset \bigcup_{s\in\{0,1\}^*} C^{-1}(s) = \operatorname{dom} C.$$

Thus, since $\operatorname{dom} C$ is prefix-free, its subset $\operatorname{dom} D$ is also prefix-free. Hence D is a prefix-free machine.

Finally, we show that the condition (i) holds for D. Let us assume that $C(p) = s$ and $|p| = H_C(s)$. Then $(p,s) \in \operatorname{Graph}(C)$ and therefore $s \in \operatorname{dom} g$. Since $C(p_{g(s)}) = s$, we see that $|p| \leq |p_{g(s)}|$ and therefore $p \in D^{-1}(s)$. Hence, $D(p) = s$ and therefore $H_D(s) \leq |p|$. Thus we have

$$H_D(s) \leq H_C(s) \tag{5}$$

for every $s \in \{0,1\}^*$. On the other hand, (4) implies that

$$H_D(s) \geq H_C(s) \tag{6}$$

for every $s \in \{0,1\}^*$. It follows from (5) and (6) that the condition (i) holds for D. ∎

Theorem 4. *For every optimal prefix-free machine V, there exists an optimal prefix-free machine W for which the following conditions (i), (ii), and (iii) hold:*

(i) *$H_W(s) = H_V(s)$ for every $s \in \{0,1\}^*$.*
(ii) *$W^{-1}(s)$ is a finite set for every $s \in \{0,1\}^*$.*
(iii) *Moreover, 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 V be an arbitrary optimal prefix-free machine. Then it follows from Theorem 3 that there exists a prefix-free machine W for which the following conditions (a), (b), and (c) hold:

(a) $H_W(s) = H_V(s)$ for every $s \in \{0,1\}^*$.
(b) $W^{-1}(s)$ is a finite set for every $s \in \{0,1\}^*$.
(c) Moreover, there exists a partial recursive function $f\colon \{0,1\}^* \to \mathbb{N}^+$ such that $\#W^{-1}(s) \leq f(s)$ for every $s \in \operatorname{dom} f$ and $\operatorname{dom} f = \{\, s \in \{0,1\}^* \mid W^{-1}(s) \neq \emptyset \,\}$.

Therefore, the conditions (i) and (ii) hold obviously. Since V is optimal, W is also optimal by the above condition (a). On the other hand, since W is optimal, $W^{-1}(s) \neq \emptyset$ for every $s \in \{0,1\}^*$. Thus, the condition (iii) holds. ∎

Through Theorems 6 and 7 below, we show that the upper bound $f(s)$ in Theorem 4 cannot be chosen to be tight at all. We first show a weaker result, Theorem 6. Then, based on this, we show a stronger result, Theorem 7. The underlying idea of the proofs of Theorems 6 and 7 is due to A. R. Meyer and D. W. Loveland [5, pp. 525–526] (see also Chaitin [1, Theorem 5.1 (f)]). In order to prove Theorem 6, we need Lemma 5 below. It is a well-known fact and follows from the inequality $\#\{\, s \in \{0,1\}^* \mid H(s) < n \,\} \leq 2^n - 1$.

Lemma 5. *Let R be an infinite subset of $\{0,1\}^*$. Then the function $H(s)$ of $s \in R$ is not bounded to the above.* ∎

A function $f\colon \{0,1\}^* \to \mathbb{N}$ is called *right-computable* if the set $\{\, (s,n) \in \{0,1\}^* \times \mathbb{N} \mid f(s) \leq n \,\}$ is r.e. Obviously, every total recursive function $f\colon \{0,1\}^* \to \mathbb{N}$ is right-computable.

Theorem 6. *Let V be an optimal prefix-free machine, and let $f\colon \{0,1\}^* \to \mathbb{N}$. Suppose that $\#V^{-1}(s) \leq f(s)$ for all $s \in \{0,1\}^*$ and f is right-computable. Then $\#V^{-1}(s) < f(s)$ for all but finitely many $s \in \{0,1\}^*$.*

Proof: We define a function h by the following two conditions:

(a) $\operatorname{dom} h = \{\, s \in \{0,1\}^* \mid \#V^{-1}(s) = f(s) \,\}$.
(b) $h(s) = \min\{\, |p| \mid p \in V^{-1}(s), \}$ for every $s \in \operatorname{dom} h$.

Note first that $V^{-1}(s) \neq \emptyset$ for every $s \in \{0,1\}^*$ since V is optimal. Therefore $\min\{\, |p| \mid p \in V^{-1}(s), \}$ is well-defined as a natural number for every $s \in \{0,1\}^*$. Since $\#V^{-1}(s) \leq f(s)$ for all $s \in \{0,1\}^*$ and f is right-computable, it is easy to see that the above two conditions (a) and (b) define a partial recursive function $h\colon \{0,1\}^* \to \mathbb{N}$. On the other hand, it follows from the condition (b) that

$$h(s) = H(s) \tag{7}$$

for every $s \in \operatorname{dom} h$.

Now, let us assume contrarily that $\#V^{-1}(s) = f(s)$ for infinitely many $s \in \{0,1\}^*$. Then, obviously, $\operatorname{dom} h$ is an infinite set. It follows from Lemma 5, the function h is not bounded to the above. Thus, given $n \in \mathbb{N}^+$, by enumerating the graph of the partial recursive function h, one can find $s \in \operatorname{dom} h$ such that $n \leq h(s)$.

Hence, combined with (7), we see that there exists a partial recursive function $\Psi\colon \mathbb{N}^+ \to \{0,1\}^*$ such that $n \leq H(\Psi(n))$. Using (2), we then see that $n \leq H(n) + O(1)$ for all $n \in \mathbb{N}^+$. It follows from (3) that $n \leq 2\log_2 n + O(1)$ for all $n \in \mathbb{N}^+$. Dividing by n and letting $n \to \infty$ we have $1 \leq 0$, a contradiction. This completes the proof. ∎

Theorem 7. *Let V be an optimal prefix-free machine, and let $f\colon \{0,1\}^* \to \mathbb{N}$. Suppose that $\#V^{-1}(s) \leq f(s)$ for all $s \in \{0,1\}^*$ and f is right-computable. Then*

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

Recall here that we identify $\{0,1\}^*$ with $\mathbb{N}$.

Proof: We denote by Q the set of all $k \in \mathbb{Z}$ such that $k \leq f(s) - \#V^{-1}(s)$ for all but finitely many $s \in \{0,1\}^*$. Note that $0 \in Q$ and therefore $Q \neq \emptyset$. This is because $\#V^{-1}(s) \leq f(s)$ for all $s \in \{0,1\}^*$.

Now, let us assume contrarily that $f(s) - \#V^{-1}(s)$ does not diverge to ∞ as $s \to \infty$. Then there exists $M \in \mathbb{N}$ such that, for infinitely many $s \in \{0,1\}^*$, $f(s) - \#V^{-1}(s) \le M$. It is then easy to see that $k \le M$ for all $k \in Q$. Thus, since Q is a nonempty subset of $\mathbb{Z}$ bounded to the above, Q has the maximum element k_0. Since $k_0 \in Q$,

$$k_0 \le f(s) - \#V^{-1}(s) \tag{8}$$

for all but finitely many $s \in \{0,1\}^*$. If $k_0 < f(s) - \#V^{-1}(s)$ for all but finitely many $s \in \{0,1\}^*$, then $k_0 + 1 \in Q$ and this contradicts the fact that k_0 is the maximum element of Q. Thus, $k_0 \ge f(s) - \#V^{-1}(s)$ for infinitely many $s \in \{0,1\}^*$. Hence, it follows from (8) that there exists a finite subset E of $\{0,1\}^*$ such that $k_0 \le f(s) - \#V^{-1}(s)$ for all $s \in \{0,1\}^* \setminus E$ and $k_0 = f(s) - \#V^{-1}(s)$ for infinitely many $s \in \{0,1\}^* \setminus E$.

We define a function $g\colon \{0,1\}^* \to \mathbb{N}$ by $g(s) = \#V^{-1}(s)$ if $s \in E$ and $g(s) = f(s) - k_0$ otherwise. Then, obviously, $\#V^{-1}(s) \le g(s)$ for all $s \in \{0,1\}^*$ and g is right-computable. Moreover, $\#V^{-1}(s) = g(s)$ for infinitely many $s \in \{0,1\}^*$. However, this contradicts Theorem 6, and the proof is completed. ■

Corollary 8. *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.*

Proof: Assume contrarily that the function $\#V^{-1}(s)$ of $s \in \{0,1\}^*$ is bounded to the above. Then there exists $M \in \mathbb{N}$ such that, for every $s \in \{0,1\}^*$, $\#V^{-1}(s) \le M$. We define a function $f\colon \{0,1\}^* \to \mathbb{N}$ by $f(s) = M$. Then, obviously, $\#V^{-1}(s) \le f(s)$ for all $s \in \{0,1\}^*$ and f is right-computable. It follows from Theorem 7 that $\lim_{s\to\infty} \{f(s) - \#V^{-1}(s)\} = \infty$. However, this contradicts the fact that $f(s) - \#V^{-1}(s) \le M$ for all $s \in \{0,1\}^*$. This completes the proof. ■

Theorem 9. *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) *$H_W(s) = H_V(s)$ for every $s \in \{0,1\}^*$.*
(ii) *$W^{-1}(s)$ is an infinite set for every $s \in \{0,1\}^*$.*

Proof: Let V be an arbitrary optimal prefix-free machine. We first show that $V^{-1}(s_0)$ has at least two elements for some $s_0 \in \{0,1\}^*$. In the case where $V^{-1}(s_0)$ is an infinite set for some $s_0 \in \{0,1\}^*$, obviously $V^{-1}(s_0)$ has at least two elements. Thus, we assume that $V^{-1}(s_0)$ is a finite set for all $s_0 \in \{0,1\}^*$, in what follows.

First, it follows from Corollary 8 that $\#V^{-1}(s_0) \ge 2$ for some $s_0 \in \{0,1\}^*$. Thus, some $V^{-1}(s_0)$ has two elements q and r with $|q| \ge |r|$. Let $b\colon \{0,1\}^* \times \mathbb{N} \to \mathbb{N}$ be a particular bijective total recursive function. We then define a partial recursive function $W\colon \{0,1\}^* \to \{0,1\}^*$ by the condition that $W^{-1}(s) = \left(V^{-1}(s) \setminus \{q\}\right) \cup \{\, q0^{b(s,i)}1 \mid i \in \mathbb{N} \,\}$ if $s = s_0$ and $W^{-1}(s) = V^{-1}(s) \cup T(s)$ otherwise, where $T(s) = \{\, q0^{b(s,i)}1 \mid i \in \mathbb{N} \;\&\; H_V(s) \le |q| + b(s,i) + 1 \,\}$. Since the set $\{\, (s,n) \in \{0,1\}^* \times \mathbb{N} \mid H_V(s) \le n \,\}$ is r.e., it is easy to see that such a partial recursive function W exists.

Since b is a bijection, the set $\{\, q0^{b(s_0,i)}1 \mid i \in \mathbb{N} \,\}$ is infinite and the set $T(s)$ is infinite for every $s \ne s_0$. Therefore the condition (ii) holds for W. On the other hand, it follows that

$$\begin{aligned} \operatorname{dom} W &= \bigcup_{s\in\{0,1\}^*} W^{-1}(s) \\ &\subset \Big(\Big(\bigcup_{s\in\{0,1\}^*} V^{-1}(s)\Big) \setminus \{q\}\Big) \cup \{\, q0^k1 \mid k \in \mathbb{N} \,\} \\ &= (\operatorname{dom} V \setminus \{q\}) \cup \{\, q0^k1 \mid k \in \mathbb{N} \,\}. \end{aligned} \tag{9}$$

Thus, since $\operatorname{dom} V$ is prefix-free and $q \in \operatorname{dom} V$, the most right-hand side of (9) is prefix-free. Hence its subset $\operatorname{dom} W$ is also prefix-free, and therefore W is a prefix-free machine.

Finally, we show that the condition (i) holds for W. In the case of $s = s_0$, since $|q| < \left|q0^k1\right|$ for all $k \in \mathbb{N}$ and there is $r \in \operatorname{dom} V$ with $|r| \le |q|$, we have $H_W(s) = H_V(s)$. In the case of $s \ne s_0$, since the set $T(s)$ does not contain any string of length less than $H_V(s)$, we have $H_W(s) = H_V(s)$ again. Thus, the condition (i) holds for W. ■

IV. The Distribution of Codewords

In this section, we investigate the distribution of codewords in $V^{-1}(s)$ for each optimal prefix-free machine V and each $s \in \{0,1\}^*$. Solovay [7] showed the following result for the distribution of all codewords $\operatorname{dom} V$ for an optimal prefix-free machine V.

Theorem 10. *Let V be an optimal prefix-free machine. Then*

$$\#\{\, p \in \{0,1\}^* \mid |p| \le n \;\&\; p \in \operatorname{dom} V \,\} = 2^{n - H(n) + O(1)}.$$

Namely, there exists $d \in \mathbb{N}$ such that

(i) $\#\{\, p \in \{0,1\}^* \mid |p| \le n \;\&\; p \in \operatorname{dom} V \,\} \le 2^{n-H(n)+d}$ *for all $n \in \mathbb{N}$, and*
(ii) $2^{n-H(n)-d} \le \#\{\, p \in \{0,1\}^* \mid |p| \le n \;\&\; p \in \operatorname{dom} V \,\}$ *for all $n \in \mathbb{N}$ with $n - H(n) \ge d$.* ■

Note that $\lim_{n\to\infty} \{n - H(n)\} = \infty$ by (3). We refine Theorem 10 to a certain extent. For that purpose, we define

$$S_C(n,s) = \{\, p \in \{0,1\}^* \mid |p| \le n \;\&\; C(p) = s \,\}$$

for each prefix-free machine C, each $n \in \mathbb{N}$, and each $s \in \{0,1\}^*$. We can then show the following theorem.

Theorem 11. *Let C be a prefix-free machine. Then $\#S_C(n,s) \le 2^{n-H(n,s)+O(1)}$.*

Proof: We show that there exists $d \in \mathbb{N}$ such that $\#S_C(n,s) \le 2^{n-H(n,s)+d}$ for all $n \in \mathbb{N}$ and all $s \in \{0,1\}^*$. For that purpose, we define a function $f\colon \mathbb{N} \to [0,\infty)$ by $f(b(n,s)) = \#S_C(n,s)2^{-n-1}$. Recall here that $b\colon \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$ is a particular bijective total recursive function. It is easy to see that the set $\{(a,k) \in \mathbb{Q} \times \mathbb{N} \mid a < f(k)\}$ is r.e. On the other hand,

$$\sum_{k=0}^{\infty} f(k) = \sum_{s\in\{0,1\}^*} \sum_{n=0}^{\infty} \#S_C(n,s)2^{-n-1}$$

$$= \sum_{s\in\{0,1\}^*} \sum_{n=0}^{\infty} \sum_{l=0}^{n} \#\overline{S_C}(l,s)2^{-n-1}$$
$$= \sum_{s\in\{0,1\}^*} \sum_{l=0}^{\infty} \sum_{n=l}^{\infty} \#\overline{S_C}(l,s)2^{-n-1}$$
$$= \sum_{s\in\{0,1\}^*} \sum_{l=0}^{\infty} \#\overline{S_C}(l,s)2^{-l} = \sum_{p\in \operatorname{dom} C} 2^{-|p|} \le 1,$$

where $\overline{S_C}(n,s) = \{ p \in \{0,1\}^* \mid |p| = n \ \& \ C(p) = s \}$. Thus, f is a lower-computable semi-measure. It follows from Theorem 2 that there exists $d' \in \mathbb{N}$ such that $f(k) \le 2^{d'}2^{-H(k)}$ for all $k \in \mathbb{N}$. Therefore we have $\#S_C(n,s)2^{-n-1} \le 2^{d'-H(b(n,s))}$ for all $n \in \mathbb{N}$ and all $s \in \{0,1\}^*$, which implies that $\#S_C(n,s) \le 2^{n-H(n,s)+d'+1}$ for all $n \in \mathbb{N}$ and all $s \in \{0,1\}^*$, as desired. ■

Theorem 13 below shows that the upper bound $2^{n-H(n,s)+O(1)}$ in Theorem 11 is tight among all optimal prefix-free machines. In order to prove Theorems 13 and 14 below, we need the following lemma.

Lemma 12. *$H(s) + n - H(H(s) + n, s)$ diverges to ∞ as $n \to \infty$ uniformly on $s \in \{0,1\}^*$. Namely, for every $M \in \mathbb{N}$, there exists $n_0 \in \mathbb{N}$ such that, for every $n \ge n_0$ and every $s \in \{0,1\}^*$, $H(s) + n - H(H(s) + n, s) \ge M$.*

Proof: Let us consider a prefix-free machine C such that, for every $p, q \in \operatorname{dom} U$, $C(pq) = b(|p| + U(q), U(p))$, where $U(q)$ is regarded as a natural number based on our identification of $\{0,1\}^*$ with $\mathbb{N}$. It is easy to see that such a prefix-free machine exists. For each $s \in \{0,1\}^*$ and each $n \in \mathbb{N}$, we see that $C(s^*n^*) = b(H(s) + n, s)$ and therefore $H_C(b(H(s)+n, s)) \le |s^*n^*| = H(s)+H(n)$. It follows from (1) that there exists $d \in \mathbb{N}$ such that, for every $s \in \{0,1\}^*$ and every $n \in \mathbb{N}$, $H(b(H(s) + n, s)) \le H(s) + H(n) + d$. Using (3) we then see that there exists $d' \in \mathbb{N}$ such that, for every $s \in \{0,1\}^*$ and every $n \in \mathbb{N}^+$, $H(s)+n-H(H(s)+n, s) \ge n - 2\log_2 n - d'$. Hence, the result follows. ■

Theorem 13. *There exists an optimal prefix-free machine V which satisfies that $\#S_V(n,s) = 2^{n-H(n,s)+O(1)}$. Namely, there exist an optimal prefix-free machine V and $d \in \mathbb{N}$ such that*

(i) *$\#S_V(n,s) \le 2^{n-H(n,s)+d}$ for all $n \in \mathbb{N}$ and all $s \in \{0,1\}^*$, and*
(ii) *$2^{n-H(n,s)-d} \le \#S_V(n,s)$ for all $n \in \mathbb{N}$ and all $s \in \{0,1\}^*$ with $n - H(n,s) \ge d$.*

Proof: By Theorem 11, it is enough to show that the condition (ii) holds for some optimal prefix-free machine V and some $d \in \mathbb{N}$ (in fact, d can be chosen to be 0 in the following construction of V).

Let us consider a partial recursive function $V \colon \{0,1\}^* \to \{0,1\}^*$ such that, for every $p, s \in \{0,1\}^*$, $V(p) = s$ if and only if there exist $q, t \in \{0,1\}^*$ for which $p = qt$ and $U(q) = b(|p|, s)$. Since U is a prefix-free machine and $b \colon \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$ is a bijective total recursive function, it is easy to see that such a partial recursive function $V \colon \{0,1\}^* \to \{0,1\}^*$ exists. Since $\operatorname{dom} U$ is prefix-free and b is an injective function, we can also check that $\operatorname{dom} V$ is prefix-free. Thus V is a prefix-free machine.

We show that $2^{n-H(n,s)} \le \#S_V(n,s)$ for all $n \in \mathbb{N}$ and all $s \in \{0,1\}^*$ with $n - H(n,s) \ge 0$. For each $n \in \mathbb{N}$ and $t \in \{0,1\}^*$, if $|t| = n - H(n,s)$, then $|b(n,s)^*t| = n$ and $V(b(n,s)^*t) = s$. Recall here that $|b(n,s)^*| = H(n,s)$. Thus, for each $n \in \mathbb{N}$, if $n - H(n,s) \ge 0$ then $2^{n-H(n,s)} \le \#S_V(n,s)$, as desired.

Finally, we show that V is optimal. By Lemma 12, we see that there exists $n_0 \in \mathbb{N}$ such that, for every $s \in \{0,1\}^*$, $H(s)+n_0-H(H(s)+n_0, s) \ge 0$. Hence, for each $s \in \{0,1\}^*$, $|b(H(s) + n_0, s)^*t| = H(s) + n_0$ and therefore $V(b(H(s) + n_0, s)^*t) = s$, where $t = 0^{H(s)+n_0-H(H(s)+n_0,s)}$. Thus, we see that $H_V(s) \le H(s)+n_0$ for all $s \in \{0,1\}^*$, which implies that V is optimal. This completes the proof. ■

As a complement to Theorem 13, the following theorem shows that only an optimal prefix-free machine can attain the upper bound $2^{n-H(n,s)+O(1)}$ in Theorem 11.

Theorem 14. *Let C be a prefix-free machine. Suppose that $2^{n-H(n,s)+O(1)} \le \#S_C(n,s)$, namely, suppose that there exists $d \in \mathbb{N}$ such that $2^{n-H(n,s)-d} \le \#S_C(n,s)$ for all $n \in \mathbb{N}$ and all $s \in \{0,1\}^*$ with $n - H(n,s) \ge d$. Then C is optimal.*

Proof: It follows from Lemma 12 that there exists $n_0 \in \mathbb{N}$ such that, for every $s \in \{0,1\}^*$, $H(s) + n_0 - H(H(s) + n_0, s) \ge d$. By the assumption, we see that, for each $s \in \{0,1\}^*$, $1 \le 2^{H(s)+n_0-H(H(s)+n_0,s)-d} \le \#S_C(H(s)+n_0, s)$. Thus, for each $s \in \{0,1\}^*$, $S_C(H(s) + n_0, s) \ne \emptyset$ and therefore there exists $p \in \{0,1\}^*$ such that $|p| \le H(s) + n_0$ and $C(p) = s$. Hence, we see that $H_C(s) \le H(s) + n_0$ for all $s \in \{0,1\}^*$, which implies that C is optimal. ■

Acknowledgments

This work was supported by KAKENHI, Grant-in-Aid for Scientific Research (C) (20540134), by SCOPE from the Ministry of Internal Affairs and Communications of Japan, and by CREST from Japan Science and Technology Agency.

References

[1] G. J. Chaitin, "A theory of program size formally identical to information theory," *J. Assoc. Comput. Mach.*, vol. 22, pp. 329–340, 1975.
[2] G. J. Chaitin, "Algorithmic entropy of sets," *Computers & Mathematics with Applications*, vol. 2, pp. 233–245, 1976.
[3] G. J. Chaitin, *Algorithmic Information Theory*. Cambridge University Press, Cambridge, 1987.
[4] R. G. Downey and D. R. Hirschfeldt, *Algorithmic Randomness and Complexity*. Springer-Verlag, To appear.
[5] D. W. Loveland, "A variant of the Kolmogorov concept of complexity," *Inform. and Contr.*, vol. 15, pp. 510–526, 1969.
[6] A. Nies, *Computability and Randomness*. Oxford University Press Inc., New York, 2009.
[7] R. M. Solovay, "Draft of a paper (or series of papers) on Chaitin's work ... done for the most part during the period of Sept.–Dec. 1974," unpublished manuscript, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, May 1975, 215 pp.
[8] A. K. Zvonkin and L. A. Levin, "The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms," *Russian Math. Surveys*, vol. 25, no. 6, pp. 83–124, 1970.