Almost Sure Convergence Theorems of Rate of Coin Tosses for Random Number Generation by Interval Algorithm

Tomohiko UYEMATSU
Dept. of Computer Science
Tokyo Institute of Technology
Ookayama, Meguro-ku, Tokyo 152-8552, Japan
e-mail: uematsu@ss.titech.ac.jp

Fumio KANAYA
Dept. of Information Science
Shonan Institute of Technology
Fujisawa-shi, Kanagawa 251-0046, Japan
e-mail: fkanaya@info.shonan-it.ac.jp

Abstract — **This paper deals with the interval algorithm proposed by Han and Hoshi for random number generation, and evaluates the efficiency of the algorithm for each sample path instead of evaluating overall expectation. We show a theorem in the almost-sure sense to give bounds on the sup generating rate as well as on the inf generating rate for each sample of input and output processes.**

I. Introduction

This paper deals with the most general random number generation problem by interval algorithm [1] where the process of repeated coin tosses and that of repeated random number generations are general processes subject to neither stationarity nor ergodicity but consistency restrictions. We are concerned with the case in which the target process should be generated exactly subject to the prescribed probability measure, and concentrate on the almost sure asymptotic property of the generating rate of each sample, i.e. the number of coin tosses per output sample of the general process. To this end, we introduce the minimum length function to indicates the length of the shortest prefix of sample $x \in \mathcal{A}^\infty$ from the general source with which the interval algorithm generates the n-length prefix of some sample $y \in \mathcal{A}^\infty$ subject to the target probability measure. Then we define *sup generating rate* and *inf generating rate* of each input sample. As a result, we prove a theorem in the almost-sure sense to give bounds on the sup generating rate as well as on the inf generating rate for each sample of input and output processes.

II. Basic Definitions

(a) General sources

Let $\mathcal{A}$ be a finite set and $(\mathcal{A}^\infty, \mathcal{F})$ a measurable space, where $\mathcal{A}^\infty$ is the set of all strings of infinite length that is formed from the symbols in $\mathcal{A}$, and $\mathcal{F}$ is a σ-field of subsets of $\mathcal{A}^\infty$. Let μ be a probability measure defined on $(\mathcal{A}^\infty, \mathcal{F})$. Then we call $(\mathcal{A}^\infty, \mathcal{F}, \mu)$ a probability space. We call μ a general process [2]. Throughout this article, we assume for μ neither stationarity nor ergodicity but consistency restrictions.

An extension of the interval algorithm for general sources was indicated in [1, Remark 12]. So, we omit the description of the algorithm.

(b) Inf generating rate and sup generating rate

The minimum length function $L_I^n : \mathcal{A}^\infty \to \mathbf{N}$ is defined as the length of the shortest prefix of sample $x \in \mathcal{A}^\infty$ from the general source ν with which the interval algorithm generates the n-length prefix of some sample $y \in \mathcal{A}^\infty$ subject to the target probability μ. Here it should be understood that $L_I^n(x)$ is defined as $+\infty$ if the above set is empty. We call $L_I^n(x)$ the minimum length of x. Further, we define the *sup generating rate* for any source sample x as

$$\bar{l}_I(x) = \limsup_{n\to\infty} \frac{1}{n} L_I^n(x) \quad \forall x \in \mathcal{A}^\infty.$$

Similarly, the *inf generating rate* is defined as

$$\underline{l}_I(x) = \liminf_{n\to\infty} \frac{1}{n} L_I^n(x) \quad \forall x \in \mathcal{A}^\infty.$$

III. Main results

We require the following hypotheses to prove the theorem as well as the consistency restrictions for μ and ν :

H1: There exists a positive number α such that

$$\liminf_{n\to\infty} \frac{1}{n} \log \frac{1}{\nu(x^n)} \geq \alpha \quad \nu\text{-a.s.}$$

H2: There exists a positive number β such that

$$\liminf_{n\to\infty} \frac{1}{n} \log \frac{1}{\mu(x^n)} \geq \beta \quad \mu\text{-a.s.}$$

Suppose that for the input sample $x \in \mathcal{A}^\infty$, the output sample $y \in \mathcal{A}^\infty$ is generated by the interval algorithm. Then, the following theorem holds.

Theorem :

$$\frac{\underline{h}_\mu(y)}{\overline{h}_\nu(x)} \leq \underline{l}_I(x) \leq \frac{\overline{h}_\mu(y)}{\overline{h}_\nu(x)} \quad \text{a.s.}$$

$$\frac{\underline{h}_\mu(y)}{\underline{h}_\nu(x)} \leq \bar{l}_I(x) \leq \frac{\overline{h}_\mu(y)}{\underline{h}_\nu(x)} \quad \text{a.s.}$$

where $\underline{h}_\nu(x)$ and $\overline{h}_\nu(x)$ (resp. $\underline{h}_\mu(y)$ and $\overline{h}_\mu(y)$) are *inf ν-complexity rate* and *sup ν-complexity rate* (resp. inf and sup μ-complexity rates) defined in [2]. Especially, if both processes ν and μ are stationary ergodic, then

$$\underline{l}_I(x) = \bar{l}_I(x) = \frac{h_\mu}{h_\nu} \quad \text{a.s.}$$

where h_ν (resp. h_μ) denotes the entropy rate of the process ν (resp. μ).

It should be noted this theorem is an extension of the results in [3] where we only deal with i.i.d. processes.

References

[1] T. S. Han and M. Hoshi, "Interval algorithm for random number generation," *IEEE Trans. Inform. Theory*, vol. 43, pp. 599-611, 1997.

[2] J. Muramatsu and F. Kanaya, "Almost-sure variable-length source coding theorems for general sources," *IEEE Trans. Inform. Theory*, vol. 45, pp. 337-342, 1999.

[3] T. Uyematsu and F. Kanaya, "Channel simulation by interval algorithm: A performance analysis of interval algorithm," *IEEE Trans. Inform. Theory*, vol. 45, no.6, pp.2121-2129, 1999.

0-7803-5857-0/00/$10.00 ©2000 IEEE.