The 26th Symposium on Information Theory
and Its Applications (SITA2003)
Higashiura, Hyogo, Japan, Dec. 15–18, 2003

# 語頭条件を満たさないWord-Valued Sourceモデルに関する一考察
# A Note on Models of a Non-Prefix-Free Word-Valued Source

石田 崇* Takashi ISHIDA　松嶋 敏泰* Toshiyasu MATSUSHIMA　平澤 茂一* Shigeichi HIRASAWA

**Abstract**— *Word-valued source* has been proposed as a new class of source models. A word-valued source is defined as a pair of an ergodic source with a countable alphabet and a function form each symbol into a word over a finite alphabet. T.Ishida et al. have derived the lower bound on the entropy rate of the source limited to a non-prefix-free i.i.d. word-valued source by approximate evaluation to typical sequences.

In this paper, we give a strict proof of the lower bound on the entropy rate of the source theoretically and the validity of the lower bound is shown.

**Keywords**— source coding, word-valued source, entoropy rate.

## 1　はじめに

情報源符号化における情報源モデルとして，“言語アルファベット情報源 (Word-Valued Source)”が提案されている [1],[2].

西新らは，i.i.d. 言語アルファベット情報源を，可算アルファベット $\mathcal{Y}$ 上の i.i.d.（定常無記憶）情報源と，$\mathcal{Y}$ から有限アルファベット $\mathcal{X}$ の有限系列への写像 $\phi$ によって定義し，この情報源の漸近等分割性 (AEP) とエントロピー・レートを示した [1]．後藤らはこれを定常エルゴード言語アルファベット情報源に一般化し，同様の結果を示した [2]．これらの結果は，写像 $\phi$ が prefix-free であるという条件もとで導かれている．

一方，写像 $\phi$ が prefix-free でない場合は，エントロピー・レートの存在性については明らかではなく，情報源のエントロピー密度レート [4] に対して上界 [1],[3] と，下界 [3] が示されている．

石田らは [3] において，標準系列について漸近的な近似手法を用いて情報源のエントロピー密度レートの下界を導出しているが，本稿では理論的に厳密な下界の証明を与え，下界の正当性を示すことを目的とする．

## 2　言語アルファベット情報源 [1],[2]

西新らは単語単位で定常無記憶 (i.i.d.) を仮定した “i.i.d. 言語アルファベット情報源 ( i.i.d. Word-Valued Source )” を提案し，情報源のエントロピーレートを導出している [1]．後に後藤らそれを単語単位で定常エルゴードである “定常エルゴード言語アルファベット情報源” に一般化し，同様の結果を導いている [2].

**定義 1. (言語アルファベット情報源 [1],[2])** 確率変数 $Y$ の無限系列 $\boldsymbol{Y} = (Y_1, Y_2, Y_3, \cdots)$ を可算アルファベット $\mathcal{Y}$ 上に値をとる定常エルゴード情報源とする[1].

$\mathcal{X}$ を有限アルファベットとし，$\mathcal{X}^*$ は $\mathcal{X}$ 上の有限系列すべての集合を表すものとする．写像 $\phi$ を $\phi : \mathcal{Y} \to \mathcal{X}^*$ で定める．さらに確率変数 $W = \phi(Y)$ を定義し，$W$ は $\mathcal{W}(\subseteq \cup_{i=0}^{\infty}\mathcal{X}^i)$ 上に値をとるものとする．このとき $W$ の実現値 $w$ を単語，$\mathcal{W}$ を単語集合と呼ぶ．$W$ の無限系列を $\boldsymbol{W} = (W_1, W_2, W_3, \cdots)$ とする．

さらに，確率変数 $\phi(Y_1)(= W_1), \phi(Y_2)(= W_2), \phi(Y_3)(= W_3), \cdots$ の連接 $\phi(Y_1)\phi(Y_2)\phi(Y_3)\cdots$ によって得られる $X$ の確率変数列を $\phi(\boldsymbol{Y})$ と表す．このとき，定常エルゴード言語アルファベット情報源は以下のように定義される．

$$\boldsymbol{X} \stackrel{\text{def}}{=} \phi(\boldsymbol{Y}) = (X_1, X_2, X_3, \cdots). \tag{1}$$

□

$n = 1, 2, 3, \cdots$ に対して長さ $n$ の $X$ の系列を $X^n = (X_1, X_2, \cdots, X_n)$，その実現値を $x^n = (x_1, x_2, \cdots, x_n)$ と書く．同様に $m = 1, 2, 3, \cdots$ に対して，長さ $m$ の確率変数列とその実現値をそれぞれ $Y^m = (Y_1, Y_2, \cdots, Y_m)$, $y^m = (y_1, y_2, \cdots, y_m)$, $W^m = (W_1, W_2, \cdots, W_m)$, $w^m = (w_1, w_2, \cdots, w_m)$ とする．さらに $\phi(Y_1), \phi(Y_2), \cdots, \phi(Y_m)$ の連接 $\phi(Y_1)\phi(Y_2)\cdots\phi(Y_m)$ からなる $X$ の有限系列を $\phi(Y^m)$ と書く．

言語アルファベット情報源からは，系列は単語単位で出力されるが，それぞれの単語が連接された長さ $n$ の $\mathcal{X}$ 上の系列 $x^n$ として観測される．すなわち $x^n = \phi(y^m)$ であり[2]，このとき任意の $m$ に対して $n = |\phi(y_1)| + |\phi(y_2)| + \cdots + |\phi(y_m)| = |w_1| + |w_2| + \cdots + |w_m|$ である．ここで $|w|$ は単語 $w$ の長さを表す．本稿ではこれ以降，$w^m$ を単語系列，$x^n$ をシンボル系列とよぶことにする．$\phi$ が prefix-free であれば観測されたシンボル系列 $x^n$ から単語系列 $w^m$ を一意に決定することが可能であるが，prefix-free でなければ一般に $w^m$ を一意に決定できない．

次に，系列 $Y^m$，$W^m$ の確率分布を

$$P_{Y^m}(y^m) \stackrel{\text{def}}{=} \Pr\{Y^m = y^m\} \tag{2}$$

$$P_{W^m}(w^m) \stackrel{\text{def}}{=} \Pr\{W^m = w^m\} \tag{3}$$

で表す．このとき，

$$P_{W^m}(w^m) = \sum_{\{y^m : w^m = (\phi(y_1), \phi(y_2), \cdots \phi(y_m))\}} P_{Y^m}(y^m) \tag{4}$$

が成り立つ．$m = 1$ のとき $P_{Y^1}(y^1) = P_Y(y)$，$P_{W^1}(w^1) = P_W(w)$ と書く．$w^m = (\phi(y_1), \phi(y_2), \cdots \phi(y_m))$ となる $w^m, y^m$ について，一般に $P_{W^m}(w^m) \geq P_{Y^m}(y^m)$ となる．また，系列 $X^n$ の確率分布を次のように表す．

$$P_{X^n}(x^n) \stackrel{\text{def}}{=} \Pr\{X^n = x^m\} \tag{5}$$

$\boldsymbol{X}$ に対し，$-\frac{1}{n}\log P_{X^n}(X^n)$ を情報源のエントロピー密度レートという [4]．エントロピー密度レートの期待値について

$$H(\boldsymbol{X}) = \lim_{n\to\infty}\left[-\frac{1}{n}\sum_{x^n\in\mathcal{X}^n} P_{X^n}(x^n)\log P_{X^n}(x^n)\right] \tag{6}$$

が存在するとき $H(\boldsymbol{X})$ を言語アルファベット情報源 $\boldsymbol{X}$ のエントロピーレートとよぶ [1],[2].

一方，単語系列 $\boldsymbol{W}$ に対してエントロピーレートは以下で与えられる．

$$H(\boldsymbol{W}) = \lim_{m\to\infty}\left[-\frac{1}{m}\sum_{w^m\in\mathcal{W}^m} P_{W^m}(w^m)\log P_{W^m}(w^m)\right] \tag{7}$$

また，平均単語長レート $E[|W|]$ を以下で定義する [2].

$$E[|W|] \stackrel{\text{def}}{=} \lim_{m\to\infty}\frac{1}{m}E_{P_{W^m}}\left[\sum_{i=1}^{m}|W_i|\right] \tag{8}$$

---

* 早稲田大学 理工学部 経営システム工学科，〒169-8555 東京都新宿区大久保 3-4-1. School of Science and Engineering, Waseda University, 3-4-1 Okubo Shinjyuku-ku, Tokyo 169-8555, Japan. E-mail: ishida@hirasa.mgmt.waseda.ac.jp

[1] 西新ら [1] は $\boldsymbol{Y}$ を i.i.d. 情報源に限定してしており，これを i.i.d. 言語アルファベット情報源とよぶ．

[2] $\phi(y) = w$ より，$\phi(y^m)$ は単語 $w$ の連接によって得られるシンボル系列を意味する．

$E_P[\cdot]$ は確率分布 $P$ による期待値を表す．

**補題 1．(i.i.d. 言語アルファベット情報源のエントロピー密度レートの上界式と AEP[1])** $\boldsymbol{Y}$ を i.i.d. 情報源とする．$\boldsymbol{X}=\phi(\boldsymbol{Y})$，$H(\boldsymbol{Y})<\infty$，$E[|W|]<\infty$ のとき次式が成り立つ．

$$\lim_{n\to\infty}\left[-\frac{1}{n}\log P_{X^n}(X^n)\right]\leq\frac{H(\boldsymbol{Y})}{E[|W|]},\ a.s. \tag{9}$$

特に，$\phi$ が prefix-free ならば，$H(\boldsymbol{X})$ が存在して

$$H(\boldsymbol{X})=\frac{H(\boldsymbol{W})}{E[|W|]},\ a.s. \tag{10}$$

で与えられる．ただし，

$$H(\boldsymbol{W}) = \sum_{w\in\mathcal{W}}P_W(w)\log P_W(w), \tag{11}$$

$$E[|W|] = \sum_{w\in\mathcal{W}}|w|\cdot P_W(w). \tag{12}$$

□

補題 1 について，$\boldsymbol{Y}$ を定常エルゴード情報源とした場合 (定常エルゴード言語アルファベット情報源) にも同様の結果が得られている [3][3]．

**定理 1．(定常エルゴード言語アルファベット情報源のエントロピー密度レートの上界 [3])** $\boldsymbol{Y}$ を定常エルゴード情報源とする．$\boldsymbol{X}=\phi(\boldsymbol{Y})$，$H(\boldsymbol{Y})<\infty$，$E[|W|]<\infty$ のとき，次式が成り立つ．

$$\limsup_{n\to\infty}\left[-\frac{1}{n}\log P_{X^n}(X^n)\right]\leq\frac{H(\boldsymbol{Y})}{E[|W|]},\ a.s. \tag{13}$$

□

定理 1 で，$\phi$ が prefix-free である場合，言語アルファベット情報源のエントロピーレート $H(\boldsymbol{X})$ が存在することが後藤ら [2] によって証明されている．

# 3 言語アルファベット情報源のエントロピー密度レートの下界式

$\phi$ が prefix-free ではない場合，言語アルファベット情報源のエントロピーレート $H(\boldsymbol{X})$ が存在することは示されていない．エントロピー密度レートの上界式を，$\boldsymbol{Y}$ が i.i.d. の場合について西新ら [1] が，定常エルゴードの場合を石田ら [3] が示している．また，[3] では単語集合を有限集合に制限した i.i.d. 言語アルファベットに対して，典型系列における近似的な評価を用いてエントロピー密度レートの下界式を導出した．

## 3.1 単語集合が語頭条件を満たさない i.i.d. 言語アルファベット情報源モデル

[3] では語頭条件を満たさない有限単語集合 $\mathcal{W}$ 上の i.i.d. 言語アルファベット情報源モデルを対象している．

**定義 2．(non-prefix-free i.i.d. 言語アルファベット情報源 [3])** 有限アルファベット $\mathcal{Y}$ 上に値をとる確率変数列 $\boldsymbol{Y}$ を i.i.d. 情報源，$\mathcal{X}=\{0,1\}$ とする．$\mathcal{Y}$ から $\mathcal{W}$ への写像 $\phi$ を 1 対 1 写像とし，$\phi:\mathcal{Y}\to\mathcal{W}=\cup_{i=1}^{k}\mathcal{X}^i$ とし[4]，$\boldsymbol{X}=\phi(\boldsymbol{Y})$ を non-prefix-free i.i.d. 言語アルファベット情報源という． □

このとき，$P_Y(y)=P_W(w)$，$H(\boldsymbol{W})=H(\boldsymbol{Y})$ が成り立つ．$\mathcal{W}$ に対して確率分布 $P_W(w)$ の与え方によって様々な構造を持つ情報源モデルとなるが，一般に $\mathcal{W}$ は語頭条件を満たなさい．

[3] 石田ら [3] による定理の証明に不備があったため，本稿付録 A で改めて証明を与える．

[4] $\phi$ は 1 対 1 写像だが prefix-free とは限らない．

## 3.2 non-prefix-free i.i.d. 言語アルファベット情報源エントロピー密度レートの下界

**定理 2．(non-prefix-free i.i.d. 言語アルファベット情報源のエントロピー密度レートの下界 [3])** non-prefix-free i.i.d. 言語アルファベット情報源 $\boldsymbol{X}$ について，

$$\liminf_{n\to\infty}\left[-\frac{1}{n}\log P_{X^n}(x^n)\right]\geq\frac{H(\boldsymbol{W})}{E[|W|]}-\frac{H(S)}{E[|W|]},\ a.s. \tag{14}$$

が成り立つ．ここで，$H(S)$ は単語の長さの確率分布 $P_S(s)$ のエントロピーを表し，以下で与えられる．

$$H(S)=-\sum_{s=1}^{K}P_S(s)\log P_S(s), \tag{15}$$

$$P_S(s)=\sum_{\{w:|w|=s\}}P_W(w). \tag{16}$$

□

本稿では [3] で示された定理 2 に対して，理論的に厳密な証明を与えることによって，下界式の正当性を示す．

## 3.3 定理 2 の証明

まず，証明に用いるいくつかの定義と補題を示す．

$W^m$ に対して $N_m$, $M_n$ を

$$N_m \stackrel{\text{def}}{=} N(W^m)=\sum_{i=1}^{m}|W_i|, \tag{17}$$

$$M_n \stackrel{\text{def}}{=} \min_{m\geq 1}\{m|N_m\geq n\}, \tag{18}$$

と定義する．ここで，$M_n$ は $W^m$ において $N_m\geq n$ となる最小の $m$ を表しており，$N_m$, $M_n$ は，ともに任意の $m\geq 1$ について $W^m$ に依存する確率変数である．また，定義より $M_n$ は $\frac{n}{|w|_{\max}}\leq M_n\leq n$ を満たす．ただし，$|w|_{\max}=\max_{w\in\mathcal{W}}|w|$ である．

ここで，与えられた $n$ について $\underline{m_{(n,\varepsilon)}}$, $\overline{m_{(n,\varepsilon)}}$ を任意の $\varepsilon>0$ を用いて

$$\overline{m_{(n,\varepsilon)}} \stackrel{\text{def}}{=} n\left(\frac{1}{E[|W|]}+\varepsilon\right), \tag{19}$$

$$\underline{m_{(n,\varepsilon)}} \stackrel{\text{def}}{=} n\left(\frac{1}{E[|W|]}-\varepsilon\right), \tag{20}$$

と定義すると，$M_n$ について，次の補題が成り立つ．

**補題 2．** $m\to\infty$ のとき次式が成り立つ．

$$\underline{m_{(n,\varepsilon)}}\leq M_n\leq\overline{m_{(n,\varepsilon)}},\ a.s. \tag{21}$$

( 証明 ) 文献 [1],[2] において

$$\lim_{n\to\infty}M_n = \infty, \tag{22}$$

$$\lim_{n\to\infty}\frac{M_n}{n} = \frac{1}{E[|W|]},\ a.s. \tag{23}$$

が示されており，これは式 (21) を意味する． □

確率分布 $P_W(w)$ $(w\in\mathcal{W})$ に従う i.i.d. 情報源からの出力系列 $W^m$ を考える．任意の $\varepsilon'>0$ を用いて

$$\overline{P_{(m,\varepsilon')}} \stackrel{\text{def}}{=} 2^{-m(H(W)+\varepsilon')}, \tag{24}$$

$$\underline{P_{(m,\varepsilon')}} \stackrel{\text{def}}{=} 2^{-m(H(W)-\varepsilon')}, \tag{25}$$

と定義すると，式 (22) より次の補題が成り立つ．

**補題 3．($\boldsymbol{W}$ の漸近等分割性 [7])** $n\to\infty$ のとき次式が成り立つ．

$$\underline{P_{(M_n,\varepsilon')}}\leq P_{W^{M_n}}(W^{M_n})\leq\overline{P_{(M_n,\varepsilon')}},\ a.s. \tag{26}$$

□

系列 $w^m$ の中に出現する $w \in \mathcal{W}$ の個数を $N(w|w^m)$ によって表すとき，次の補題が成り立つ．

**補題 4. (重複対数の法則 [6])**　$m \to \infty$ のとき次式が成り立つ．

$$P_W(w) - \delta_m \le \frac{N(w|w^m)}{m} \le P_W(w) + \delta_m,\ a.s. \quad (27)$$

ただし，$\delta_m = O\left(\sqrt{\frac{\log\log m}{m}}\right)$ である．□

単語系列 $w^m$ 中に含まれる長さ $s$ $(s = 1, 2, \cdots, K)$ の単語の数 $L_m(s)$ とその確率分布 $P_S(s)$ を

$$L_m(s) \stackrel{\mathrm{def}}{=} \sum_{\{w:|w|=s\}} N(w|w^m), \quad (28)$$

$$P_S(s) \stackrel{\mathrm{def}}{=} \sum_{\{w:|w|=s\}} P_W(w), \quad (29)$$

と定義する．式 (22) と，補題 4 より任意の $s$ $(s = 1, 2, \cdots, K)$ に対して以下が成り立つ．

$$\begin{aligned} &M_n\left(P_S(s) - \delta_{M_n}\right) \\ &\quad \le L_{M_n}(s) \le M_n\left(P_S(s) + \delta_{M_n}\right),\ a.s. \end{aligned} \quad (30)$$

ここで，$m$ を固定して $m = \sum_{s=1}^K L_m(s)$ とし，

$$V_m \stackrel{\mathrm{def}}{=} \frac{m!}{L_m(1)!L_m(2)!\cdots L_m(K)!}, \quad (31)$$

を定義する．$V_m$ は，長さ $s$ の単語をそれぞれ $L_m(s)$ 個ずつ含んでいる単語系列 $w^m$ において，単語 $w$ の長さのみに着目したときの単語の並べ方の総数を表している．

$\overline{V_{(m,\delta_m)}}$, $\underline{V_{(m,\delta_m)}}$ を以下のように定義する．

$$\overline{V_{(m,\delta_m)}} \stackrel{\mathrm{def}}{=} \frac{m!}{\prod_{s=1}^K m\left(P_S(s) - \delta_m\right)!}, \quad (32)$$

$$\underline{V_{(m,\delta_m)}} \stackrel{\mathrm{def}}{=} \frac{m!}{\prod_{s=1}^K m\left(P_S(s) + \delta_m\right)!}. \quad (33)$$

このとき，式 (22) より $n \to \infty$ のとき次式を満たす．

$$\underline{V_{M_n}} \le V_{M_n} \le \overline{V_{M_n}},\ a.s. \quad (34)$$

$\pi\{x^n\}$ によって，系列 $x^n$ の prefix 全体からなる集合を表すものとし，$\mathcal{G}_\phi$ を以下のように定義する．

$$\mathcal{G}_\phi(x^n) \stackrel{\mathrm{def}}{=} \left\{w^{M_n} :\ x^n \in \pi\{\phi(y^{M_n})\}\right\}. \quad (35)$$

このとき，系列 $x^n$ の出現確率 $P_{X^n}(x^n)$ は

$$P_{X^n}(x^n) = \sum_{w^{M_n} \in \mathcal{G}_\phi(x^n)} P_{W^{M_n}}(w^{M_n}), \quad (36)$$

で与えられる．

以上の準備のもとで，定理 2 を証明する．

[定理 2 の証明]　式 (36) より，

$$\begin{aligned} P_{X^n}(X^n) &= \sum_{w^{M_n} \in \mathcal{G}_\phi(X^n)} P_{W^{M_n}}(w^{M_n}), \\ &\le |\mathcal{G}_\phi(X^n)| \cdot \max_{w^{M_n} \in \mathcal{G}_\phi(X^n)} \left[P_{W^{M_n}}(w^{M_n})\right]. \end{aligned} \quad (37)$$

ここで，$|\mathcal{G}_\phi(X^n)|$ は $\mathcal{G}_\phi(X^n)$ に含まれる系列 $w^{M_n}$ の個数を表す．

任意の $n$ に対して $X^n$ が単語を $M_n$ 個含むような，$X^n$ の単語の切れ目の入れ方の総数は $V_{M_n}$ で与えられる．一方，$X^n$ に対して単語の切れ目の入れ方を一つ固定した場合を考える．このような単語の切れ方をもつ各々の単語系列 $W^{M_n}$ において，$W_1, W_2, \cdots, W_{M_n}$ の連接によって得られるシンボル系列 $X^{N_{M_n}}$ はすべて異なる系列となる．したがって，$\mathcal{G}_\phi(X^n)$ に含まれる系列の総数は，起こりうる単語の切れ方の総数によって上から抑ることができる．すなわち，

$$|\mathcal{G}_\phi(X^n)| \le \sum_{\{M_n: \frac{n}{|w|_{\max}} \le M_n \le n\}} V_{M_n}, \quad (38)$$

が成り立つ．ここで，任意の $\varepsilon > 0$ について

$$\mathcal{M}_\varepsilon^n = \left\{m:\ \left|m - \frac{n}{E[|W|]}\right| \le \varepsilon\right\}, \quad (39)$$

を定義すると，

$$|\mathcal{G}_\phi(X^n)| \le \sum_{\{M_n:\ M_n \in \mathcal{M}_\varepsilon^n\}} V_{M_n} + \sum_{\{M_n:\ M_n \notin \mathcal{M}_\varepsilon^n\}} V_{M_n}, \quad (40)$$

と表すことができる．ここで，

$$P_{W^{M_n}}^{\max}(w^{M_n}) = \max_{w^{M_n} \in \mathcal{G}_\phi(X^n)} \{P_{W^{M_n}}(w^{M_n})\}, \quad (41)$$

と表すことにすると，式 (37),(40) より次式が成り立つ．

$$\begin{aligned} &\liminf_{n\to\infty}\left[-\frac{1}{n}\log P_{X^n}(X^n)\right] \\ &\ge \liminf_{n\to\infty}\left[-\frac{1}{n}\log\left(\sum_{\{M_n:\ M_n \in \mathcal{M}_\varepsilon^n\}} V_{M_n} + \sum_{\{M_n:\ M_n \notin \mathcal{M}_\varepsilon^n\}} V_{M_n}\right)\right] \\ &\quad + \liminf_{n\to\infty}\left[-\frac{1}{n}\log P_{W^{M_n}}^{\max}(w^{M_n})\right]. \end{aligned} \quad (42)$$

ここで，式 (22), (23)，および補題 3 より

$$\begin{aligned} &\liminf_{n\to\infty}\left[-\frac{1}{n}\log P_{W^{M_n}}^{\max}(w^{M_n})\right] \\ &\quad \ge \liminf_{n\to\infty}\left[-\frac{1}{n}\log \overline{P_{(M_n,\varepsilon')}}\right] \\ &\quad = \liminf_{n\to\infty}\left[-\frac{1}{n}\log(2^{-M_n(H(W)+\varepsilon')})\right] \\ &\quad = \liminf_{n\to\infty}\left[\frac{M_n}{n}(H(\boldsymbol{W})+\varepsilon'))\right] \\ &\quad = \frac{H(W)}{E[|W|]} + \frac{\varepsilon'}{E[|W|]} \\ &\quad = \frac{H(W)}{E[|W|]} + \varepsilon'',\ a.s. \end{aligned} \quad (43)$$

を得る．ただし，$\varepsilon'' = \frac{\varepsilon'}{E[|W|]}$ である．一方，

$$V_{M_n}^{\max} = \max_{\{M_n:\ M_n \in \mathcal{M}_\varepsilon^n\}} V_{M_n} \quad (44)$$

と書くことにすると，式 (22), 補題 2 より

$$\liminf_{n\to\infty}\left[-\frac{1}{n}\log\left(\sum_{\{M_n:\ M_n \in \mathcal{M}_\varepsilon^n\}} V_{M_n} + \sum_{\{M_n:\ M_n \notin \mathcal{M}_\varepsilon^n\}} V_{M_n}\right)\right]$$

$$\geq \liminf_{n\to\infty}\left[-\frac{1}{n}\log\left(\sum_{\{M_n:\ M_n\in\mathcal{M}_\varepsilon^n\}} V_{M_n}\right)\right]$$
$$\geq \liminf_{n\to\infty}\left[-\frac{1}{n}\log\left(\left(\overline{m_{(n,\varepsilon)}}-\underline{m_{(n,\varepsilon)}}+1\right)\cdot V_{M_n}^{\max}\right)\right]$$
$$\geq \liminf_{n\to\infty}\left[-\frac{1}{n}\log\left((2n\varepsilon+1)\cdot V_{M_n}^{\max}\right)\right]$$
$$\geq \liminf_{n\to\infty}\left[-\frac{1}{n}\log\overline{V_{M_n}}\right]$$
$$= \liminf_{n\to\infty}\left[-\frac{1}{n}\log\frac{M_n!}{\prod_{s=1}^K M_n\left(P_S(s)-\delta_{M_n}\right)!}\right],\ a.s. \quad (45)$$

と書けるが，stirling の公式 [6] から

$$-\frac{1}{n}\log\frac{M_n!}{\prod_{s=1}^K M_n\left(P_S(s)-\delta_{M_n}\right)!}$$
$$=\gamma_{M_n}+\frac{M_n}{n}\left(\sum_{s=1}^K (P_S(s)-\delta_{M_n})\log(P_S(s)-\delta_{M_n})\right), \quad (46)$$

が成り立つ．ただし，

$$\gamma_{M_n}=\frac{(K-1)\log\sqrt{2\pi M_n}}{n}-\frac{(\theta_{M_n}-\sum_{s=1}^K\theta_{M_n(P_S(s)-\delta_{M_n})})\cdot\log e}{n}+\sum_{s=1}^K\frac{\log\sqrt{P_S(s)-\delta_{M_n}}}{n}-\frac{M_n}{n}\sum_{s=1}^K\delta_{M_n}\log\frac{M_n}{e}, \quad (47)$$

であり，$n\to\infty$ のとき $\gamma_{M_n}\to 0$ を満たす．ここで，$\theta_m$ は $m\to\infty$ のとき $\theta_m\to 0$ となる項である．いま，

$$H(S)\stackrel{\mathrm{def}}{=}-\sum_{s=1}^K P_S(s)\log P_S(s), \quad (48)$$

を定義すると，式 (22), (23) より

$$\liminf_{n\to\infty}\left[-\frac{1}{n}\log\frac{M_n!}{\prod_{s=1}^K M_n\left(P_S(s)-\delta_{M_n}\right)!}\right]$$
$$=\liminf_{n\to\infty}\left[\frac{M_n}{n}\left(\sum_{s=1}^K (P_S(s)-\delta_{M_n})\log(P_S(s)-\delta_{M_n})\right)\right]$$
$$=-\frac{H(S)}{E[|W|]},\ a.s. \quad (49)$$

が得られる．以上，式 (42),(43),(49) より

$$\liminf_{n\to\infty}\left[-\frac{1}{n}\log P_{X^n}(X^n)\right]\geq\frac{H(\boldsymbol{W})}{E[|W|]}-\frac{H(P_s)}{E[|W|])}+\varepsilon'',\ a.s. \quad (50)$$

が得られ，定理 2 が示された．

## 4 まとめ

本稿では，[3] において石田らが導出した，non-prefix-free i.i.d. 言語アルファベット情報源のエントロピー密度レート下界式に対して厳密な証明を与えることにより，その正当性を理論的に示すことができた．

今後の課題としては，単語集合が有限の non-prefix-free i.i.d. 言語アルファベット情報源に限定していた下界式を，定常エルゴード言語アルファベット情報源や，単語集合が可算無限に拡張して検証することが挙げられる．また，non-prefix-free 言語アルファベット情報源のエントロピーレートが存在するのか，漸近等分割性が成り立つのか議論する必要がある．


### 謝辞

著者の 1 人石田は，本研究を行うにあたり日頃よりご助言をいただいている武蔵工業大学 後藤正幸 先生，湘南工科大学 小林学 先生，及び平澤研究室の皆様に深く感謝いたします．


## 参考文献


[1] M.Nishiara and H.Morita, “On the AEP of word-valued sources,” *IEEE Trans. Inform. Theory*, vol.IT-46, no.3, pp.1116-1120, May 2000.

[2] M.Goto, T.Matsushima, and S.Hirasawa, “A source model with probability distribution over word set and recurrence time theorem,” IEICE Trans. Fundamentals, vol.E86-A, no.2, pp.2517-2525, Oct. 2003.

[3] 石田 崇，後藤 正幸，松嶋敏泰，平澤 茂一，“単語単位で系列を出力する情報源の性質について，” 第 25 回情報理論とその応用シンポジウム予稿集 (SITA2002), pp.695-698, 2002.

[4] 韓 太舜，情報理論における情報スペクトル的方法，培風館,1998.

[5] 韓 太舜，小林 欣吾，情報の符号化の数理，岩波講座応用数学，岩波書店,1994.

[6] William Feller, *An Introduction to Probability Theory and Its Applications*, Volume 1, 3rd Ed., John Wiley Sons, Inc., 1968.

[7] C.E.Shannon, “A mathematical theory of communication,” Bell Sys. Tech. Journal, no.27, pp.379-423, pp.623-656, 1948.

[8] B.McMillan, “The basic theorems of information theory,” Ann. Math. Stat., no.24, pp.196-219, 1953.

[9] L.Breiman, “The individual ergodic theorems of information theory,” Ann. Math. Stat., no.28, pp.809-811, 1957.

[10] P.Billingsley, *Ergodic theory and information*, John Wiley & Sons, New York, 1965.


付　録

## A 定理 1 の証明

$N_m$，$M_n$ を式 (17), (18) のように定義する．$\boldsymbol{Y}$ は定常エルゴード情報源なので，Shannon-McMillan-Breiman の定理 [7],[8],[9] より

$$\lim_{m\to\infty}-\frac{1}{m}\log P_{Y^m}(Y^m)=H(\boldsymbol{Y}),\ a.s. \quad (51)$$

が成り立つ．また，エルゴード定理 [10] より

$$\lim_{m\to\infty}N_m/m=E[|W|],\ a.s. \quad (52)$$

が成り立つ．式 (22) より，$\lim_{n\to\infty}M_n=\infty$ が成り立つので，

$$\limsup_{n\to\infty}\left[-\frac{1}{M_n}\log P_{X^{N_{M_n}}}(X^{N_{M_n}})\right]$$
$$\leq\limsup_{n\to\infty}\left[-\frac{1}{M_n}\log P_{W^{M_n}}(W^{M_n})\right]$$
$$\leq\lim_{n\to\infty}\left[-\frac{1}{M_n}\log P_{Y^{M_n}}(Y^{M_n})\right]=H(\boldsymbol{Y}),\ a.s. \quad (53)$$

を満たす．ここで，$n\leq N_{M_n}$ より，

$$-\frac{1}{n}\log P_{X^n}(X^n)\quad\leq\quad-\frac{1}{n}\log P_{X_{N_{M_n}}}(X^{N_{M_n}})=-\frac{M_n}{n}\frac{1}{M_n}\log P_{X^{N_{M_n}}}(X^{N_{M_n}}), \quad (54)$$

と書けるが，式 (23) より，

$$\limsup_{n\to\infty}\left[-\frac{1}{n}\log P_{X^n}(X^n)\right]$$
$$\leq\limsup_{n\to\infty}\left[-\frac{M_n}{n}\frac{1}{M_n}\log P_{X^{N_{M_n}}}(X^{N_{M_n}})\right]=\frac{H(\boldsymbol{Y})}{E[|W|]}, \quad (55)$$

となり，定理 1 が示された．