KYBERNETIKA — VOLUME 23 (1987), NUMBER 5

ON A CODING THEOREM CONNECTED WITH 'USEFUL' ENTROPY OF ORDER α AND TYPE β

PRITI JAIN, R. K. TUTEJA

In this paper two generalizations of 'useful' mean length of the codewords of the outputs by a memoryless source are given. Lower and upper bounds are also derived for these generalized 'useful' mean lengths in terms of order α and type β 'useful' entropy.

1. INTRODUCTION

Let S be a finite memoryless stationary source which takes values in a finite set $A = (a_1, a_2, \ldots, a_n)$ according to a probability distribution $P = (p_1, p_2, \ldots, p_n)$; where $p_i > 0$ for all i and $\sum_{i=1}^{n} p_i = 1$. Belis and Guiaşu [1] remarked that a source is not completely specified by the probability distribution P over the source alphabet A. They enriched the usual description of the information source (i.e., a finite alphabet and a finite probability distribution) by introducing an additional parameter measuring the utility associated with an event. Let $U = (u_1, u_2, \ldots, u_n)$ be a set of positive real numbers, where u_i in general is independent of p_i, the probability of occurrence. They derived the following function

$$H(P,U) = -\sum_{i=1}^{n} u_i p_i \log p_i \tag{1.1}$$

for two distributions P and U and called it 'useful' entropy. The function (1.1) can be taken as a satisfactory measure for the average quantity of valuable or useful information provided by a source letter.

Guiaşu and Picard [2] considered the problem of encoding the letter output by the source S by means of a single letter prefix code whose codewords $b_1, b_2, \ldots, b_n$ have the lengths $l_1, l_2, \ldots, l_n$ satisfy the Kraft's inequality

$$\sum_{i=1}^{n} D^{-l_i} \leq 1\,, \tag{1.2}$$

where D is the size of the code alphabet. They defined the following quantity

$$(1.3) \qquad L_u = \frac{\sum_{i=1}^{n} l_i u_i p_i}{\sum_{i=1}^{n} u_i p_i}$$

and called it 'useful' mean length of the code.

Longo [3] interpreted (1.3) as the average cost of transmitting letters a_i with probability p_i and utility u_i and gave some practical interpretation of this length. Lower and upper bounds for the cost function (1.3) in terms of (1.1) are also derived by him.

2. TWO MEASURES OF THE COST

In the derivation of the cost function (1.3) clearly the assumption that the cost varies linearly as with the code length is taken. But this is not always the case. However there may be occasions when the cost does not vary linearly with code lengths. For example, this may be the case when the cost of encoding or decoding equipment were an important factor. In this case cost is more nearly an exponential function of l_i. Such types of functions occur frequently in market equilibrium and growth models in economics. Obviously linear dependence is the limiting case of this measure (exponential measure). Thus some times, it is more appropriate to minimize the quantities

$$(2.1) \qquad c_1 = \sum_{i=1}^{n} u_i p_i D^{t l_i},$$

and

$$(2.2) \qquad c_2 = \sum_{i=1}^{n} u_i p_i^{\alpha} D^{(\alpha-1) l_i},$$

where t and α are some parameters related to the cost.

In order to make the result of this paper more directly comparable with the usual noiseless coding theorem, instead of minimizing (2.1) and (2.2), we shall minimize

$$(2.3) \qquad L_u^{\beta}(t) = \frac{1}{(2^{1-\beta}-1)\log D}\left[\left\{\frac{\sum_{i=1}^{n} u_i p_i D^{t l_i}}{\sum_{i=1}^{n} u_i p_i}\right\}^{(1-\beta)/t} - 1\right] \qquad \begin{array}{l}\beta > 0,\\ \beta \neq 1,\\ 0 < t < \infty\end{array}$$

and

$$(2.4) \qquad L_u^{\beta}(\alpha) = \frac{1}{(2^{1-\beta}-1)\log D}\left[\left\{\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i^{\alpha} D^{(\alpha-1) l_i}}\right\}^{(\beta-1)/(\alpha-1)} - 1\right] \qquad \begin{array}{l}\beta > 0,\ \beta \neq 1,\\ \alpha > 0,\ \alpha \neq 1,\\ \alpha \neq \beta\end{array}$$

being monotone functions of c_1 and c_2 respectively. (2.3) would be termed as order t and type β cost length, and (2.4) as order α and type β cost length.

Clearly (2.3) and (2.4) are the generalizations of the 'useful' average codeword length (1.3). Clearly as $\beta \to 1$ and $t \to 0$, (2.3) reduces to (1.3). In the limiting case as $\alpha \to 1$ and $\beta \to 1$, (2.4) also reduces to (1.3).

We derive the bounds for the measure (2.3) and (2.4) in terms of generalized 'useful' entropy of order α and type β studied by Picard [4], which is given by

$$H_\alpha^\beta(P, U) = \frac{1}{(2^{1-\beta} - 1)} \left[\left(\frac{\sum_{i=1}^n u_i p_i^\alpha}{\sum_{i=1}^n u_i p_i} \right)^{(\beta-1)/(\alpha-1)} - 1 \right] \tag{2.5}$$

$$\alpha \neq 1, \quad \beta \neq 1, \quad \alpha \neq \beta, \quad \alpha, \beta > 0$$

under the condition

$$\sum_{i=1}^n u_i D^{-l_i} \leqq \sum_{i=1}^n u_i p_i . \tag{2.6}$$

Clearly the inequality (2.6) is the generalization of Kraft's inequality (1.2). When $u_i = 1$ for each i, (2.6) reduces to (1.2). A code satisfying the Kraft's inequality (2.6) would be termed as 'useful' code. Also we note that when u_i's are ignored, (2.5) reduces to an entropy of order α and type β.

3. CODING THEOREM FOR 'USEFUL' CODES

In this section we derive the lower and upper bounds for the generalized 'useful' mean lengths (2.3) and (2.4) in terms of (2.5). In the following theorem, we find lower bound for $L_u^\beta(t)$.

Theorem 1. If $\{u_i\}_{i=1}^n$, $\{p_i\}_{i=1}^n$ and $\{l_i\}_{i=1}^n$ satisfy (2.6), then

$$L_u^\beta(t) \geqq \frac{H_\alpha^\beta(P, U)}{\log D}, \tag{3.1}$$

where $\alpha = 1/(1 + t)$.

Proof. By Holder's inequality

$$\left(\sum_{i=1}^n x_i^p\right)^{1/p} \left(\sum_{i=1}^n y_i^q\right)^{1/q} \leqq \sum_{i=1}^n x_i y_i , \tag{3.2}$$

where $1/p + 1/q = 1$, $p < 1$ and $x_i, y_i > 0$.

In (3.2), put

$$p = -t, \quad x_i = \left[\frac{u_i p_i}{\sum_{i=1}^n u_i p_i} \right]^{-(1/t)} D^{-l_i}, \quad q = 1 - \alpha, \quad y_i = \left[\frac{u_i p_i^\alpha}{\sum_{i=1}^n u_i p_i} \right]^{1/(1-\alpha)},$$

we get

$$(3.3)\qquad \left[\frac{\sum_{i=1}^{n} u_i p_i D^{l_i t}}{\sum_{i=1}^{n} u_i p_i}\right]^{-(1/t)} \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}\right]^{1/(1-\alpha)} \leqq \frac{\sum_{i=1}^{n} u_i D^{-l_i}}{\sum_{i=1}^{n} u_i p_i}.$$

Using (2.6) in (3.3), we get

$$(3.4)\qquad \left[\frac{\sum_{i=1}^{n} u_i p_i D^{l_i t}}{\sum_{i=1}^{n} u_i p_i}\right]^{-(1/t)} \leqq \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}\right]^{1/(\alpha-1)}.$$

Now consider two cases:
(i) Let $0 < \beta < 1$. Raising both sides of (3.4) to the power $(\beta - 1)$, we get

$$(3.5)\qquad \left[\frac{\sum_{i=1}^{n} u_i p_i D^{l_i t}}{\sum_{i=1}^{n} u_i p_i}\right]^{(1-\beta)/t} \geqq \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}\right]^{(\beta-1)/(\alpha-1)}.$$

Since $2^{1-\beta} - 1 > 0$ for $\beta < 1$, we get from (3.5) the inequality (3.1).
(ii) Let $\beta > 1$. The proof follows on same lines. □

It is clear that the equality in (3.1) is true if and only if

$$D^{-l_i} = \frac{p_i^{\alpha}}{\dfrac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}}$$

or

$$(3.6)\qquad l_i = -\alpha \log_D p_i + \log_D \Big[\sum_{i=1}^{n} u_i p_i^{\alpha} \Big/ \sum_{i=1}^{n} u_i p_i\Big].$$

Thus it is always possible to have a codeword satisfying the requirement

$$(3.7)\qquad p_i^{-\alpha}\left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}\right] \leqq D^{l_i} < D p_i^{-\alpha}\left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}\right].$$

In the following theorem, we give an upper bound for $L_u^{\beta}(t)$ in terms of $H_{\alpha}^{\beta}(P, U)$.

Theorem. By properly choosing the lengths $l_1, l_2, \ldots, l_n$ in the code of Theorem 1, $L_u^{\beta}(t)$ can be made to satisfy the following inequality

$$(3.8)\qquad L_u^{\beta}(t) < \frac{H_{\alpha}^{\beta}(P, U)}{\log D} D^{1-\beta} + \frac{D^{1-\beta} - 1}{(2^{1-\beta} - 1)\log D}.$$

Proof. Choose the code words length l_i as to satisfy (3.7). Now from the left hand inequality of (3.7), we have

$$D^{l_i} < Dp_i^{-\alpha}\left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}\right]. \tag{3.9}$$

Raising both sides of (3.9) to the power t, we have

$$D^{l_i t} < D^t p_i^{-\alpha t}\left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}\right]^t. \tag{3.10}$$

Now multiplying both sides of (3.10) by $u_i p_i / \sum_{i=1}^{n} u_i p_i$ and summing over i and using the value of t on RHS, we get

$$\frac{\sum_{i=1}^{n} u_i p_i D^{l_i t}}{\sum_{i=1}^{n} u_i p_i} < D^t\left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}\right]^{1/\alpha}. \tag{3.11}$$

Again raising both sides of (3.11) to the power $1/t$, we have

$$\left[\frac{\sum_{i=1}^{n} u_i p_i D^{l_i t}}{\sum_{i=1}^{n} u_i p_i}\right]^t < D\left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}\right]^{1/(1-\alpha)}. \tag{3.12}$$

Now consider two cases:

(i) Let $0 < \beta < 1$. Raising both sides of (3.12) to the power $1 - \beta > 0$, we get

$$\left[\frac{\sum_{i=1}^{n} u_i p_i D^{t l_i}}{\sum_{i=1}^{n} u_i p_i}\right]^{(1-\beta)/t} < D^{(1-\beta)}\left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i}\right]^{(\beta-1)/(\alpha-1)}. \tag{3.13}$$

Since $2^{1-\beta} - 1 > 0$ for $\beta < 1$, a simple manipulation proves (3.8).

(ii) Let $\beta > 1$. The proof follows on the same lines. □

In the next theorem we obtain lower bound for $L_u^{\beta}(\alpha)$ in terms of $H_{\alpha}^{\beta}(P, U)$.

Theorem 3. Let $\{u_i\}_{i=1}^n$, $\{p_i\}_{i=1}^n$ and $\{l_i\}_{i=1}^n$ satisfy the inequality (1.6), then

$$L_u^{\beta}(\alpha) \geqq \frac{H_{\alpha}^{\beta}(P, U)}{\log D}. \tag{3.14}$$

Proof. We have two possibilities.

(1) Let $0 < \alpha < 1$. By using Holder's inequality and (1.6), we have

$$\sum_{i=1}^{n} u_i p_i^\alpha D^{(\alpha-1)l_i} \leqq \sum_{i=1}^{n} u_i p_i \,. \tag{3.15}$$

Obviously (3.15) implies

$$\left[\frac{\sum_{i=1}^{n} u_i p_i^\alpha}{\sum_{i=1}^{n} u_i p_i^\alpha D^{(\alpha-1)l_i}}\right]^{1/(\alpha-1)} \leqq \left[\frac{\sum_{i=1}^{n} u_i p_i^\alpha}{\sum_{i=1}^{n} u_i p_i}\right]^{1/(\alpha-1)} . \tag{3.16}$$

Now consider two cases:

(a) Let $0 < \beta < 1$. Raising both sides of (3.16) to the power $\beta - 1$, we get

$$\left[\frac{\sum_{i=1}^{n} u_i p_i^\alpha}{\sum_{i=1}^{n} u_i p_i^\alpha D^{(\alpha-1)l_i}}\right]^{(\beta-1)/(\alpha-1)} \geqq \left[\frac{\sum_{i=1}^{n} u_i p_i^\alpha}{\sum_{i=1}^{n} u_i p_i}\right]^{(\beta-1)/(\alpha-1)} . \tag{3.17}$$

Since $2^{1-\beta} - 1 > 0$ for $\beta < 1$, we get from (3.17), the equation (3.14).

(b) Let $\beta > 1$. The proof follows similarly.

(ii) If $1 < \alpha < \infty$, the proof follows on the same lines as for $0 < \alpha < 1$. □

It is clear that equality in (3.14) is true if and only if

$$D^{-l_i} = p_i \,,$$

which implies that

$$l_i = \log_D (1/p_i) \,. \tag{3.18}$$

Thus, it is always possible to have a code word satisfying the requirement

$$\log_D \frac{1}{p_i} \leqq l_i < \log_D \frac{1}{p_i} + 1 \,,$$

which is equivalent to

$$\frac{1}{p_i} \leqq D^{l_i} < \frac{D}{p_i} . \tag{3.19}$$

Next we obtain a result giving the upper bound to the 'useful' mean code length $L_u^\beta(\alpha)$.

Theorem 4. By properly choosing the lengths $l_1, l_2, \ldots, l_n$ in the code of Theorem 3, $L_u^\beta(\alpha)$ can be made to satisfy the following

$$L_u^\beta(\alpha) < \frac{D^{[-\alpha(\beta-1)]/(\alpha-1)} H_\alpha^\beta(P, U)}{\log D} + \left[\frac{D^{[-\alpha(\beta-1)]/(\alpha-1)} - 1}{2^{1-\beta} - 1}\right] \frac{1}{\log D} \tag{3.20}$$

$$\text{for} \quad \alpha > 1 \,, \quad \beta > 0 \,.$$

Proof. From (3.19), it is clear that $p_i D^{l_i} < D$. Consequently

$$p_i^\alpha D^{(\alpha-1)l_i} < D^\alpha D^{-l_i} \,. \tag{3.21}$$

Multiplying both sides by u_i and then summing over i and using (2.6), we get

$$\sum_{i=1}^{n} u_i p_i^{\alpha} D^{(\alpha-1)l_i} < D^{\alpha} \sum_{i=1}^{n} u_i p_i \,. \tag{3.22}$$

Obviously (3.22) can be written as

$$\left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{u} u_i p_i^{\alpha} D^{(\alpha-1)l_i}}\right]^{1/(\alpha-1)} > \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{D^{\alpha} \sum_{i=1}^{n} u_i p_i}\right]^{1/(\alpha-1)} . \tag{3.23}$$

We have two possibilities

(i) Let $0 < \beta < 1$. Raising both sides of (3.23) to the power, $\beta - 1$, we get

$$\left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{\sum_{i=1}^{n} u_i p_i^{\alpha} D^{(\alpha-1)l_i}}\right]^{(\beta-1)/(\alpha-1)} < \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha}}{D^{\alpha} \sum_{i=1}^{n} u_i p_i}\right]^{(\beta-1)/(\alpha-1)} . \tag{3.24}$$

(ii) Let $\beta > 1$. The proof follows similarly. □

Special Case. When $\beta \to 1$ and $\alpha \to 1$ both the inequalities (3.1) and (3.14) give

$$L_u \geqq \frac{H(P, U)}{\bar{u} \log D}, \tag{3.25}$$

and the inequalities (3.8) and (3.20) give

$$L_u < \frac{H(P, U)}{\bar{u} \log D} + 1 \,. \tag{3.26}$$

(3.25) and (3.26) can be written as

$$\frac{H(P, U)}{\bar{u} \log D} \leqq L_u < \frac{H(P, U)}{\bar{u} \log D} + 1 \,,$$

where L_u is the 'useful' mean length as defined by (1.3). Longo [3] gave the lower and upper bounds on L_u as follows

$$\frac{H(P, U) + \overline{u \log u} - \bar{u} \log \bar{u}}{\bar{u} \log D} \leqq L_u < \frac{H(P, U) + \overline{u \log u} - \bar{u} \log \bar{u}}{\bar{u} \log D} + 1 \,,$$

where the bar means the mean value with respect to the probability distribution $P = (p_1, p_2, \ldots, p_n)$. Since $x \log x$ is a convex U function, the inequality $\overline{u \log u} \geqq \bar{u} \log \bar{u}$ holds and therefore $H(P, U)$ does not seem to be as basic in (3.26) as in (3.25).

(Received March 11, 1986.)

REFERENCES

[1] M. Belis and S. Guiaşu: A quantitative-qualitative measure of information in cybernetic systems. IEEE Trans. Inform. Theory *14* (1968), 593–594.

[2] S. Guiaşu and C. F. Picard: Borne inférieure de la longueur utile de certains codes. C.R. Acad. Sci. Paris *273* (1971), 248–251.

[3] G. Longo: A noiseless coding theorem for source having utilities. SIAM J. Appl. Math. *30* (1971), 4, 739–748.

[4] C. F. Picard: Weighted probabilistic information measures. J. Combin. Inform. System Sci. *4* (1979), 4, 343–356.

Prof. Dr. R. K. Tuteja, Department of Mathematics, Maharshi Dayanand University, Rohtak — 124 001. India.

Dr. Priti Jain, Department of Mathematics, Maharshi Dayanand University, Rohtak — 124 001. India.