

情報理論レポート課題 5(符号構成法)

提示：2009/12/09(水) 提出：2009/12/16(水)

次の情報源 S に関して、問いに答えよ。

$$S = \left[\begin{array}{cccccccc} s_1 & , & s_2 & , & s_3 & , & s_4 & , & s_5 & , & s_6 & , & s_7 & , & s_8 \\ \frac{12}{40} & , & \frac{8}{40} & , & \frac{5}{40} & , & \frac{5}{40} & , & \frac{3}{40} & , & \frac{3}{40} & , & \frac{2}{40} & , & \frac{2}{40} \end{array} \right]$$

1. シャノン・ファノ符号化 (算術符号化)

(1) 情報源 S に対して、シャノン・ファノ符号 ϕ_S を求めよ。

各情報源 $s_i, 1 \leq i \leq 8$ に対して、符号長 l_i を求める。 l_i は各記号が生成されたときの自己情報量の切り上げで求められる。すなわち、次式で求める。

$$l_i = \lceil -\log p_i \rceil$$

よって、次のように求められる。

i	p_i	$-\log p_i$	l_i
1	0.3	1.7	2
2	0.2	2.3	3
3	0.125	3.0	3
4	0.125	3.0	3
5	0.075	3.7	4
6	0.075	3.7	4
7	0.05	4.3	5
8	0.05	4.3	5

よって、シャノン・ファノ符号 ϕ_S の符号長 (のベクトル) が $L_S = (2, 3, 3, 3, 4, 4, 5, 5)$ と求められる。

次に、各情報源 s_i に対して、累積確率 $p_i^+ = \sum_{j=0}^{i-1} p_j$ を求める。ただし、 $p_0 = p_1 = 0$ とす

る。また、この累積確率の2進数 $(p_i^+)_2$ を上記の桁数だけ求める。

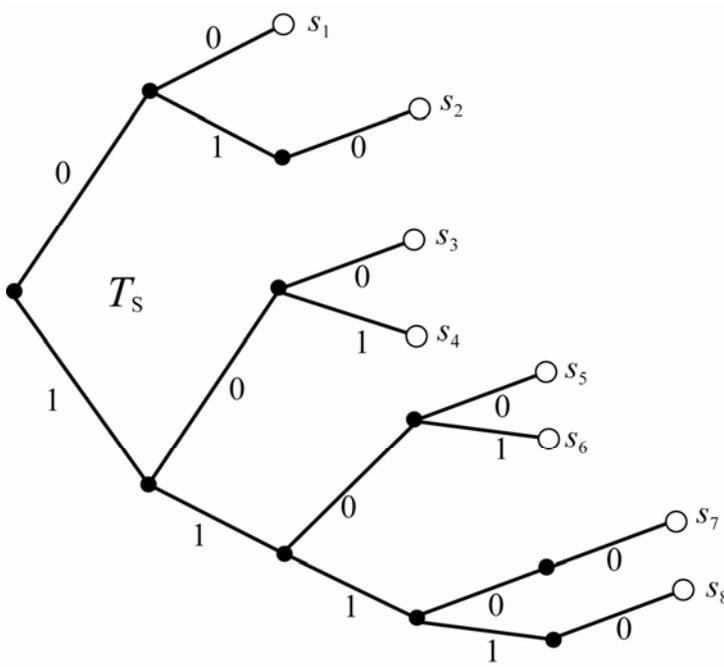
i	p_i	$(p_i^+)_{10}$	$(p_i^+)_{2}$
1	0.3	$(0.0)_{10}$	$(0.00)_2$
2	0.2	$(0.3)_{10}$	$(0.010)_2$
3	0.125	$(0.5)_{10}$	$(0.100)_2$
4	0.125	$(0.625)_{10}$	$(0.101)_2$
5	0.075	$(0.75)_{10}$	$(0.1100)_2$
6	0.075	$(0.825)_{10}$	$(0.1101)_2$
7	0.05	$(0.9)_{10}$	$(0.11100)_2$
8	0.05	$(0.95)_{10}$	$(0.11110)_2$

以上より、次のようにシャノンファノ符号 ϕ_s が構成できる。

$$\phi_s = \{s_1 \mapsto 00, s_2 \mapsto 010, s_3 \mapsto 100, s_4 \mapsto 101, \\ s_5 \mapsto 1100, s_6 \mapsto 1101, s_7 \mapsto 11100, s_8 \mapsto 11110\}$$

(2) 符号 ϕ_s に対して符号の木 T_s を示せ。

次図のようになる。



(3) 符号 ϕ_s の平均符号長 \bar{L}_s と効率 e_s を求めよ。

平均符号長 \bar{L}_s は次のように求められる。

$$\begin{aligned} \overline{L}_S &= \sum_{i=1}^8 p_i \cdot l_i \\ &= \frac{12}{40} \cdot 2 + \frac{8}{40} \cdot 3 + \frac{5}{40} \cdot 3 + \frac{5}{40} \cdot 3 + \frac{3}{40} \cdot 4 + \frac{3}{40} \cdot 4 + \frac{2}{40} \cdot 5 + \frac{2}{40} \cdot 5 \\ &= 3.05 \end{aligned}$$

情報源 S のエントロピー $H(S)$ は次のように求められる。

$$\begin{aligned} H(S) &= -\sum_{i=1}^8 p_i \log p_i \\ &= -\frac{12}{40} \log \frac{12}{40} - \frac{8}{40} \log \frac{8}{40} - \frac{5}{40} \log \frac{5}{40} - \frac{5}{40} \log \frac{5}{40} \\ &\quad - \frac{3}{40} \log \frac{3}{40} - \frac{3}{40} \log \frac{3}{40} - \frac{2}{40} \log \frac{2}{40} - \frac{2}{40} \log \frac{2}{40} \\ &\simeq 2.73 \quad [\text{bit} / \text{記号}] \end{aligned}$$

よって、効率 e_s は次のように求められる。

$$e_s = \frac{H(S)}{\overline{L}_S} \simeq \frac{2.73}{3.05} \simeq 0.89$$

2. ハフマン符号化 (コンパクト符号化)

(1) 情報源 S に対して、ハフマン符号 ϕ_H を求めよ。

次のように縮退情報源の系列が求められる。

$$S_0 = S = \left[\begin{array}{cccccccc} s_1 & , & s_2 & , & s_3 & , & s_4 & , & s_5 & , & s_6 & , & s_7 & , & s_8 \\ \frac{12}{40} & , & \frac{8}{40} & , & \frac{5}{40} & , & \frac{5}{40} & , & \frac{3}{40} & , & \frac{3}{40} & , & \frac{2}{40} & , & \frac{2}{40} \end{array} \right]$$

$$S_1 = \left[\begin{array}{ccccccc} s_1 & , & s_2 & , & s_3 & , & s_4 & , & A & , & s_5 & , & s_6 \\ \frac{12}{40} & , & \frac{8}{40} & , & \frac{5}{40} & , & \frac{5}{40} & , & \frac{4}{40} & , & \frac{3}{40} & , & \frac{3}{40} \end{array} \right], A = \{s_7, s_8\}$$

$$S_2 = \left[\begin{array}{cccccc} s_1 & , & s_2 & , & B & , & s_3 & , & s_4 & , & A \\ \frac{12}{40} & , & \frac{8}{40} & , & \frac{6}{40} & , & \frac{5}{40} & , & \frac{5}{40} & , & \frac{4}{40} \end{array} \right], B = \{s_5, s_6\}$$

$$S_3 = \left[\begin{array}{ccccc} s_1 & , & C & , & s_2 & , & B & , & s_3 \\ \frac{12}{40} & , & \frac{9}{40} & , & \frac{8}{40} & , & \frac{6}{40} & , & \frac{5}{40} \end{array} \right], C = \{s_4, A\}$$

$$S_4 = \left[\begin{array}{cccc} s_1 & , & D & , & C & , & s_2 \\ \frac{12}{40} & , & \frac{11}{40} & , & \frac{9}{40} & , & \frac{8}{40} \end{array} \right], D = \{B, s_3\}$$

$$S_5 = \left\{ \begin{array}{ccc} E & , & s_1 & , & D \\ \frac{17}{40} & , & \frac{12}{40} & , & \frac{11}{40} \end{array} \right\}, E = \{C, s_2\}$$

$$S_6 = \left\{ \begin{array}{cc} F & , & E \\ \frac{23}{40} & , & \frac{17}{40} \end{array} \right\}, F = \{s_1, D\}$$

$$S_7 = \left\{ \begin{array}{c} G \\ 1 \end{array} \right\}, G = \{F, E\}$$

ここから、縮退された情報源記号に、繰り返し符号記号を割り振ることにより、ハフマン

符号を得ることができる。

$$S_7 = \left\{ \begin{array}{c} G \\ 1 \end{array} \right\}, G = \{F, E\}, \phi_H^7 = \{G \mapsto \quad\}, \phi_H^G = \{F \mapsto 1, E \mapsto 0\}$$

$$S_6 = \left\{ \begin{array}{cc} F & , & E \\ \frac{23}{40} & , & \frac{17}{40} \end{array} \right\}, F = \{s_1, D\}, \phi_H^6 = \{F \mapsto 1, E \mapsto 0\}, \phi_H^F = \{s_1 \mapsto 1, D \mapsto 0\}$$

$$S_5 = \left\{ \begin{array}{ccc} E & , & s_1 & , & D \\ \frac{17}{40} & , & \frac{12}{40} & , & \frac{11}{40} \end{array} \right\}, E = \{C, s_2\},$$

$$\phi_H^5 = \{E \mapsto 0, s_1 \mapsto 11, D \mapsto 10\}, \phi_H^E = \{C \mapsto 1, s_2 \mapsto 0\}$$

$$S_4 = \left\{ \begin{array}{cccc} s_1 & , & D & , & C & , & s_2 \\ \frac{12}{40} & , & \frac{11}{40} & , & \frac{9}{40} & , & \frac{8}{40} \end{array} \right\}, D = \{B, s_3\},$$

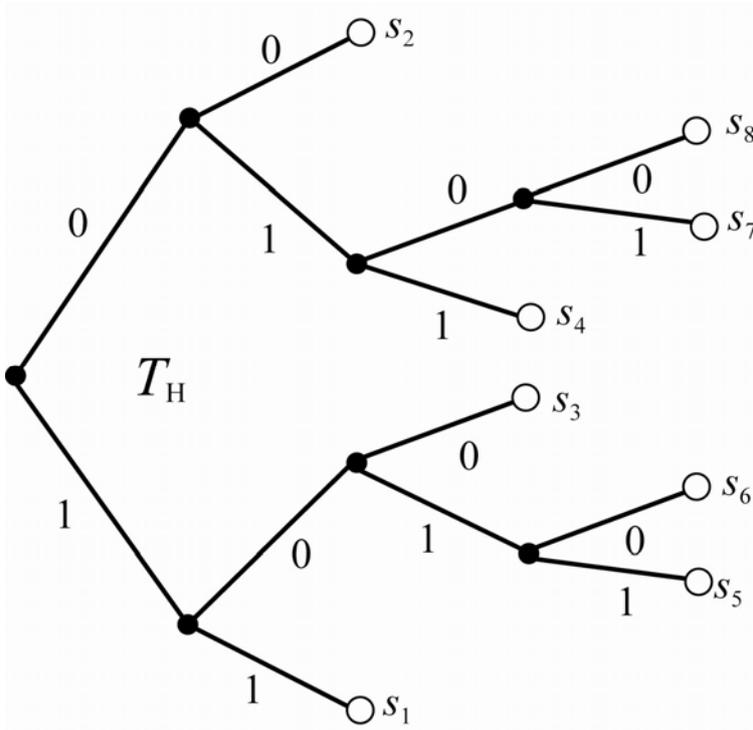
$$\phi_H^4 = \{s_1 \mapsto 11, D \mapsto 10, C \mapsto 01, s_2 \mapsto 00\}, \phi_H^D = \{B \mapsto 1, s_3 \mapsto 0\}$$

$$S_3 = \left\{ \begin{array}{ccccc} s_1 & , & C & , & s_2 & , & B & , & s_3 \\ \frac{12}{40} & , & \frac{9}{40} & , & \frac{8}{40} & , & \frac{6}{40} & , & \frac{5}{40} \end{array} \right\}, C = \{s_4, A\},$$

$$\phi_H^3 = \{s_1 \mapsto 11, C \mapsto 01, s_2 \mapsto 00, B \mapsto 101, s_3 \mapsto 100\}, \phi_H^C = \{s_4 \mapsto 1, A \mapsto 0\}$$

$$S_2 = \left\{ \begin{array}{cccccc} s_1 & , & s_2 & , & B & , & s_3 & , & s_4 & , & A \\ \frac{12}{40} & , & \frac{8}{40} & , & \frac{6}{40} & , & \frac{5}{40} & , & \frac{5}{40} & , & \frac{4}{40} \end{array} \right\}, B = \{s_5, s_6\},$$

$$\phi_H^2 = \{s_1 \mapsto 11, s_2 \mapsto 00, B \mapsto 101, s_3 \mapsto 100, s_4 \mapsto 011, A \mapsto 010\}, \phi_H^B = \{s_5 \mapsto 1, s_6 \mapsto 0\}$$



(3)符号 ϕ_H の平均符号長 \overline{L}_H と効率 e_H を求めよ。

ハフマン符号の符号長 (のベクトル) は、 $L_H = (2, 2, 3, 3, 4, 4, 4, 4)$ となる。よって、平均

符号長 \overline{L}_H は次のように求められる。

$$\begin{aligned} \overline{L}_H &= \sum_{i=1}^8 p_i \cdot l_i \\ &= \frac{12}{40} \cdot 2 + \frac{8}{40} \cdot 2 + \frac{5}{40} \cdot 3 + \frac{5}{40} \cdot 3 + \frac{3}{40} \cdot 4 + \frac{3}{40} \cdot 4 + \frac{2}{40} \cdot 4 + \frac{2}{40} \cdot 4 \\ &= 2.75 \end{aligned}$$

よって、効率 e_H は次のように求められる。

$$e_H = \frac{H(S)}{\overline{L}_H} \simeq \frac{2.73}{2.75} \simeq 0.99$$