情報理論レポート課題 5(符号構成法)解答例

次の情報源Sに関して、問いに答えよ。

$$S=\begin{Bmatrix} a & , & b & , & c & , & d & , & e & , & f \\ \dfrac{10}{32} & , & \dfrac{8}{32} & , & \dfrac{6}{32} & , & \dfrac{5}{32} & , & \dfrac{2}{32} & , & \dfrac{1}{32} \end{Bmatrix}$$

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$\alpha \in S$に対して、符号長l_αを求める。l_α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。

$$l_\alpha = \left\lceil -\log P\left(\alpha\right) \right\rceil$$

よって、次のように求められる。

α	$P(\alpha)$	$-\log P(\alpha)$	l_α
a	0.3125	1.68	2
b	0.25	2.00	2
c	0.1875	2.42	3
d	0.15625	2.68	3
e	0.0625	4.00	4
f	0.03125	5.00	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$L_S = (2,2,3,3,4,5)$と求められる。

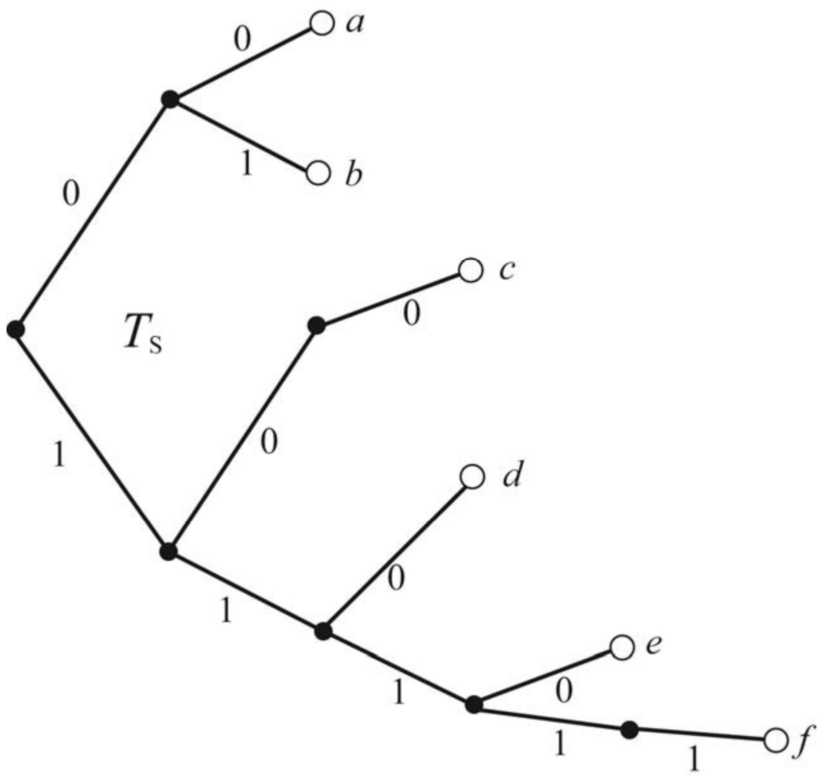
次に、各情報源s_iに対して、累積確率$p_\alpha^+ = \sum_{\beta は \alpha 以前} P\left(\beta\right)$を求める。ただし、$P\left(a\right) = 0.0$とする。また、この累積確率の 2 進数$\left(p_\alpha^+\right)_2$を上記の桁数だけ求める。

α	$P(\alpha)$	$(p_\alpha)_{10}$	$(p_\alpha)_2$
a	0.3125	0.00	0.00
b	0.25	0.3125	0.01
c	0.1875	0.5625	0.100
d	0.15625	0.75	0.110
e	0.0625	0.90625	0.1110
f	0.03125	0.96875	0.11111

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。

$$\phi_S = \left\{ a \mapsto 00, b \mapsto 01, c \mapsto 100, d \mapsto 110, e \mapsto 1110, f \mapsto 11111 \right\}$$

(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$\overline{L_S}$と効率e_Sを求めよ。

平均符号長$\overline{L_S}$は次のように求められる。

$$\overline{L_S} = \sum_{\alpha \in S} P\left(\alpha\right) \bullet l_\alpha$$
$$= \frac{10}{32} \bullet 2 + \frac{8}{32} \bullet 2 + \frac{6}{32} \bullet 3 + \frac{5}{32} \bullet 3 + \frac{2}{32} \bullet 4 + \frac{1}{32} \bullet 5$$
$$\simeq 2.56$$

情報源Sのエントロピー$H(S)$は次のように求められる。

$$
\begin{aligned}
H(S) &= -\sum_{\alpha \in S} P\left(\alpha\right) \log P\left(\alpha\right) \\
&= -\frac{10}{32}\log\frac{10}{32} - \frac{8}{32}\log\frac{8}{32} - \frac{6}{32}\log\frac{6}{32} \\
&\qquad -\frac{5}{32}\log\frac{5}{32} - \frac{2}{32}\log\frac{2}{32} - \frac{1}{32}\log\frac{1}{32} \\
&\simeq 2.30 \quad [bit\ /\ \text{記号}]
\end{aligned}
$$

よって、効率e_Sは次のように求められる。

$$
e_S = \frac{H(S)}{\bar{L}_S} \simeq \frac{2.30}{2.56} \simeq 0.90
$$

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。

$$
S_0 = S = \begin{Bmatrix} a & , & b & , & c & , & d & , & e & , & f \\ \frac{10}{32} & , & \frac{8}{32} & , & \frac{6}{32} & , & \frac{5}{32} & , & \frac{2}{32} & , & \frac{1}{32} \end{Bmatrix}
$$

$$
S_1 = \begin{Bmatrix} a & , & b & , & c & , & d & , & A \\ \frac{10}{32} & , & \frac{8}{32} & , & \frac{6}{32} & , & \frac{5}{32} & , & \frac{3}{32} \end{Bmatrix}, A = \{e, f\}
$$

$$
S_2 = \begin{Bmatrix} a & , & b & , & B & , & c \\ \frac{10}{32} & , & \frac{8}{32} & , & \frac{8}{32} & , & \frac{6}{32} \end{Bmatrix}, B = \{d, A\}
$$

$$
S_3 = \begin{Bmatrix} C & , & a & , & b \\ \frac{14}{32} & , & \frac{10}{32} & , & \frac{8}{32} \end{Bmatrix}, C = \{B, c\}
$$

$$
S_4 = \begin{Bmatrix} D & , & C \\ \frac{18}{32} & , & \frac{14}{32} \end{Bmatrix}, D = \{a, b\}
$$

$$
S_5 = \begin{Bmatrix} E \\ 1 \end{Bmatrix}, E = \left\{D, C\right\}
$$

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。

$$S_5 = \left\{\begin{matrix} E \\ 1 \end{matrix}\right\}, E = \left\{D, C\right\},$$
$$\phi_H^5 = \left\{E \mapsto \quad \right\}, \phi_H^E = \left\{D \mapsto 0, C \mapsto 1\right\}$$

$$S_4 = \left\{\begin{matrix} D & , & C \\ \frac{18}{32} & , & \frac{14}{32} \end{matrix}\right\}, D = \{a, b\},$$
$$\phi_H^4 = \left\{D \mapsto 0, C \mapsto 1\right\}, \phi_H^D = \left\{a \mapsto 0, b \mapsto 1\right\}$$

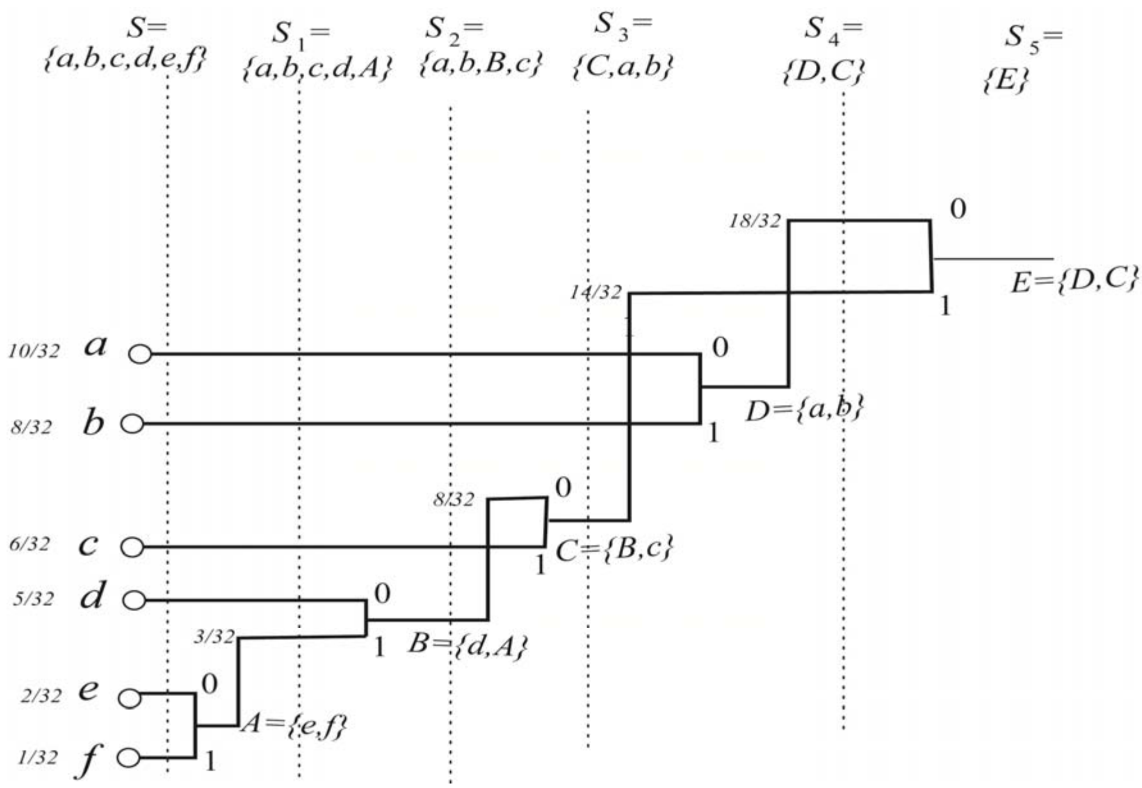
$$S_3 = \left\{\begin{matrix} C & , & a & , & b \\ \frac{14}{32} & , & \frac{10}{32} & , & \frac{8}{32} \end{matrix}\right\}, C = \{B, c\},$$
$$\phi_H^3 = \left\{C \mapsto 1, a \mapsto 00, b \mapsto 01\right\}, \phi_H^C = \left\{B \mapsto 0, c \mapsto 1\right\}$$

$$S_2 = \left\{\begin{matrix} a & , & b & , & B & , & c \\ \frac{10}{32} & , & \frac{8}{32} & , & \frac{8}{32} & , & \frac{6}{32} \end{matrix}\right\}, B = \{d, A\}$$
$$\phi_H^2 = \left\{a \mapsto 00, b \mapsto 01, B \mapsto 10, c \mapsto 11\right\}, \phi_H^B = \left\{d \mapsto 0, A \mapsto 1\right\}$$

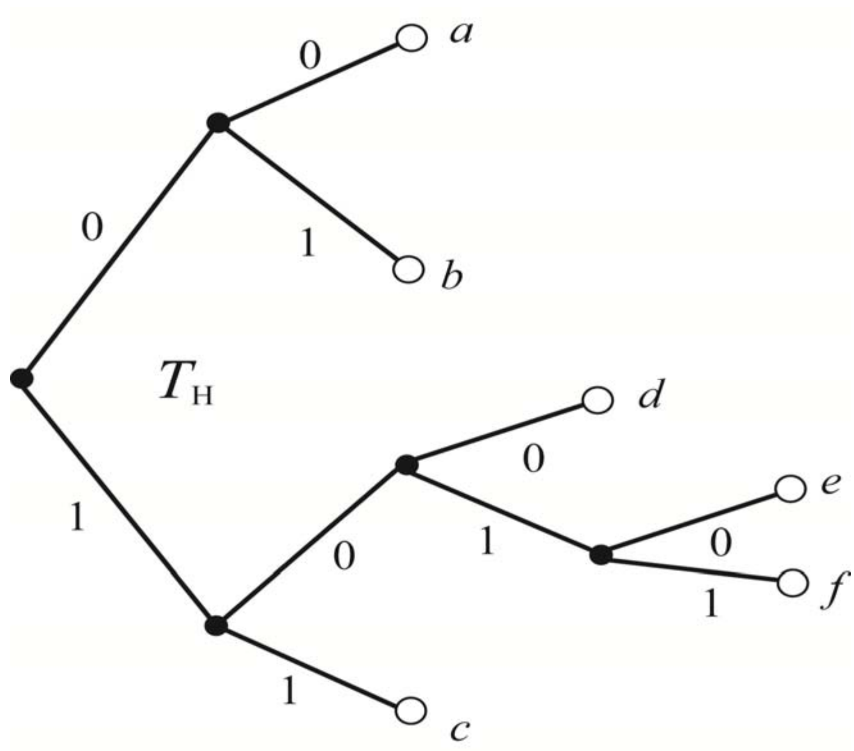
$$S_1 = \left\{\begin{matrix} a & , & b & , & c & , & d & , & A \\ \frac{10}{32} & , & \frac{8}{32} & , & \frac{6}{32} & , & \frac{5}{32} & , & \frac{3}{32} \end{matrix}\right\}, A = \{e, f\}$$
$$\phi_H^1 = \left\{a \mapsto 00, b \mapsto 01, c \mapsto 11, d \mapsto 100, A \mapsto 101\right\}, \phi_H^A = \left\{e \mapsto 0, f \mapsto 1\right\}$$

$$S_0 = S = \left\{\begin{matrix} a & , & b & , & c & , & d & , & e & , & f \\ \frac{10}{32} & , & \frac{8}{32} & , & \frac{6}{32} & , & \frac{5}{32} & , & \frac{2}{32} & , & \frac{1}{32} \end{matrix}\right\}$$
$$\phi_H = \left\{a \mapsto 00, b \mapsto 01, c \mapsto 11, d \mapsto 100, e \mapsto 1000, f \mapsto 1001\right\}$$

なお、符号の木を葉から順に求める図的な解法でもよい。この場合、次のような木が描ける。



(2)符号 ϕ_H に対して符号の木 T_H を示せ。下図のようになる。



(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、$L_H = \left(2,2,2,3,4,4\right)$ 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。

$$\overline{L_H} = \sum_{\alpha \in S} P(\alpha) l_\alpha$$
$$= \frac{10}{32} \bullet 2 + \frac{8}{32} \bullet 2 + \frac{6}{32} \bullet 2 + \frac{5}{32} \bullet 3 + \frac{2}{32} \bullet 4 + \frac{1}{32} \bullet 4$$
$$\simeq 2.34$$

また、設問 1 より、情報源 S エントロピーは $H(S) \simeq 2.30$ である。よって、効率 e_H は次のように求められる。

$$e_H = \frac{H(S)}{\overline{L}_H} \simeq \frac{2.30}{2.34} \simeq 0.98$$