

H20 院試 10. 情報理論 回答

(1) $\alpha = a_3a_2a_1, \beta = a_4a_4$ としたとき, $\alpha \neq \beta$ であるにもかかわらず $f_1(\alpha) = f_1(\beta)$ だから

(2)

(2-1) f_2 は符号語の系列が $011b_1$ のときに先頭 (001) を瞬時に a_3 へ復号することができない (b_1 の先頭が 0 であるか 1 であるかを確認しなければ復号できない). f_3 は符号語の系列中に対応する符号語が現れた瞬間に瞬時に復号することができる. つまり, f_2 は瞬時に復号可能ではなく, f_3 は瞬時に復号可能なので f_3 の方が優れている.

(2-2)

$$\begin{aligned}f_4(a_1) &= 0 \\f_4(a_2) &= 10 \\f_4(a_3) &= 110 \\f_4(a_4) &= 1110 \\f_4(a_5) &= 1111\end{aligned}$$

先頭から見て, 0 が現れるまでの 1 の個数を数える. 以下を繰り返せば一意に復号できる.

- 0 が現れる前に 1 が 4 個連続で現れたらそれを a_5 に復号
- 1 が $i (0 \leq i \leq 3)$ 個現れた後 0 が現れたらそれを a_{i+1} に復号

また, f_2, f_3 による符号化の平均符号語長はともに

$$2p_1 + 2p_2 + 2p_3 + 4p_4 + 4p_5$$

であり, f_4 による平均符号語長は

$$p_1 + 2p_2 + 2p_3 + 4p_4 + 4p_5$$

である. よって p_i の値にかかわらず f_4 による平均符号語長は f_2, f_3 による符号化の平均符号語長よりも短い.

(3)

(3-1)

(i) $s = 2$ のとき明らかに $N_{f_5}(s-1) = 2$ である.

(ii) $s > 2$ のとき $N_{f_5}(s-1) \neq 0$ より長さ $s-1$ のものが少なくとも 1 つあるので, 符号化の木は図 1(これは一つの例) のように枝分かれば各深さで丁度一回起こる. 従って $N_{f_5}(s-1) = 1$ である.

(3-2) $p_1 = \frac{1}{3}, p_2 = \frac{1}{3}, p_3 = \frac{1}{6}, p_4 = \frac{1}{6}$

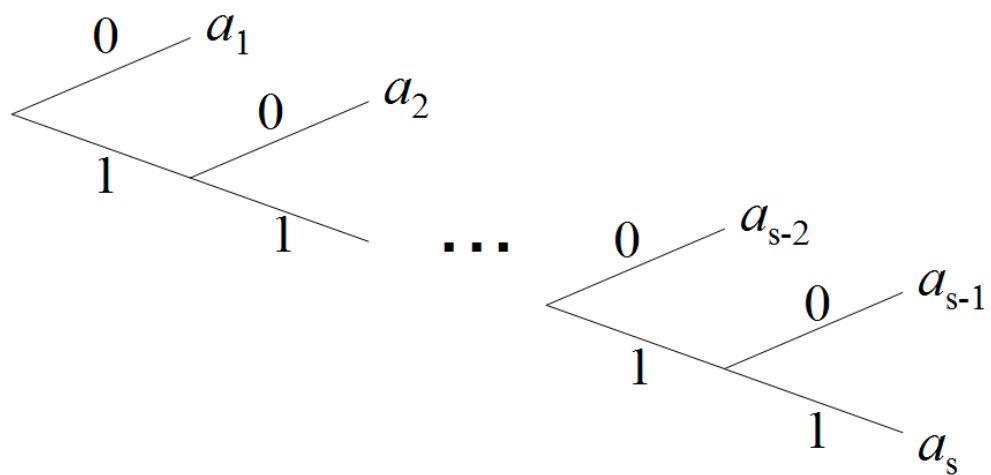


図 1: 条件を満たすハフマン符号化の木

$$\begin{aligned}
 f_6(a_1) &= 0 \\
 f_6(a_2) &= 10 \\
 f_6(a_3) &= 110 \\
 f_6(a_4) &= 1111
 \end{aligned}$$

$$\begin{aligned}
 f_7(a_1) &= 00 \\
 f_7(a_2) &= 01 \\
 f_7(a_3) &= 10 \\
 f_7(a_4) &= 11
 \end{aligned}$$

f_6 は $(a_1, (a_2, (a_3, a_4)))$, f_7 は $((a_1, a_2), (a_3, a_4))$ と縮約している .