

情報理論

問題 1. 下記の無記憶情報源 S を考える.

$$S = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ p_1 & p_2 & p_3 & p_4 \end{pmatrix}$$

p_i は記号 α_i の生起確率であり, $p_1 \geq p_2 \geq p_3 \geq p_4 > 0$ とする. S に対して, 二元ハフマン符号化したときの α_i に対する符号語の長さ (符号語長) を ℓ_i とおく. このとき, 符号語長の組 $(\ell_1, \ell_2, \ell_3, \ell_4)$ が $(1, 2, 3, 3)$ と $(2, 2, 2, 2)$ の二種類のハフマン符号が得られたとする.

- (1) $(\ell_1, \ell_2, \ell_3, \ell_4) = (1, 2, 3, 3)$ のときの平均符号長を p_i を用いて表せ.
- (2) p_2 がとり得る範囲 (上限と下限) を示せ.
- (3) このようなハフマン符号が得られる生起確率の組 (p_1, p_2, p_3, p_4) の例を一つあげよ.

問題 2. 情報源アルファベットが $\{a, b, c\}$ の単純マルコフ情報源 X を考える. 記号 x を出力した後, 記号 y を出力する確率を $p(y|x)$ と書くとき,

$$\begin{aligned} p(a|a) &= \frac{1}{2}, & p(b|a) &= \frac{1}{2}, & p(c|a) &= 0 \\ p(a|b) &= \frac{1}{4}, & p(b|b) &= 0, & p(c|b) &= \frac{3}{4} \\ p(a|c) &= \frac{1}{4}, & p(b|c) &= 0, & p(c|c) &= \frac{3}{4} \end{aligned}$$

とする. 解答の際には, 情報量の単位は bit とし, 整数にできる対数はできる限り整数にすること.

- (1) この情報源 X のエントロピーを求めよ.
- (2) この情報源 X の随伴情報源のエントロピーを求めよ.
- (3) (2) で求めたエントロピーが (1) で求めたエントロピーより小さくない理由を述べよ.