

第13回 情報源符号化

(教科書 3.3 情報源符号化定理)

符号の効率

情報源 S のエントロピー $H(S)$

2元符号 $C = \{c_1, c_2, \dots, c_M\}$,

符号長 l_1, l_2, \dots, l_M

平均符号長 L

$$L = \sum_{i=1}^M P(c_i) l_i \text{ [bit/記号]}$$

情報源符号化 — 符号の効率化 \Rightarrow 平均符号長をできるだけ短く

2元符号の効率 e

$$e = \frac{H(S)}{L}$$

$0 \leq e \leq 1$ ($L = H(S)$ のとき最大 $e = 1$)

(r 元符号に対しては $H(S)/\log_2 r$)

情報源符号化定理 (シャノンの第1基本定理) (source coding theorem)

情報源のエントロピーを $H(S)$ とすると, r 元符号を用いて符号化する場合, 平均符号長 L は

$$\frac{H(S)}{\log_2 r} \quad (2\text{元符号 } r = 2 \text{ の場合は } H(S))$$

にいくらでも近づけることができる。

参考文献

- [1] 久保田一, 大石邦夫, 福本昌弘, "3.3 情報源符号化定理", C言語による情報理論入門, pp.42-47, コロナ社, 2007 (ISBN978-4-339-02521-0).