慶應義塾大学大学院理工学研究科　入試
2008 年　電気電子工学　D5　情報工学

1.

文字の現れる確率は、

$$p(A) = \frac{1}{2}$$

$$p(B) = \frac{1}{4}$$

$$p(C) = \frac{1}{4}$$

である。

情報源のエントロピー $H(x)$ は、

$$H(x) = -p(A)\log_2 p(A) - p(B)\log_2 p(B) - p(C)\log_2 p(C) = \frac{3}{2}$$

である。

情報量、エントロピー

情報量は確率の逆数の対数で表現される。

$$(\text{情報量}) = -\log_2 p(i)$$

エントロピー H とは、平均の情報量、つまり情報量の期待値のことであり、

$$H = \sum_i p(i)(\text{情報量}) = -\sum_i p(i)\log_2 p(i)$$

である。

2.

ハフマンの符号化法によって、文字に最適な符号を決定する。

$$p(A) \geq p(B) \geq p(C)$$

であるから、文字 B に 0 の符号を、文字 C に 1 の符号を与える。

$$p(A) \geq p(B \cap C)$$

であるから、文字 A に 0 の符号を、文字 B と C に 1 の符号を与える。
まとめると、

- A : 0
- B : 10
- c : 11

である。

3.

「01 符号」の 0 と 1 の符号の長さが等しいとして考える。このとき、符号容量は $C = 1$ である。
上で求めた符号の平均符号長 τ_1 は、

$$\tau_1 = 1p(A) + 2p(B) + 2p(C) = \frac{3}{2}$$

である。したがって、通信速度 R_1 は、

$$R_1 = \frac{H(x)}{\tau_1} = 1$$

である。
符号を、

- A : 00
- B : 01
- c : 10

とした場合の平均符号長 τ_2 は、

$$\tau_2 = 2p(A) + 2p(B) + 2p(C) = 2$$

である。したがって、通信速度 R_2 は、

$$R_2 = \frac{H(x)}{\tau_2} = \frac{3}{4}$$

である。
符号化の効率は R/C で得られる。前者の効率は、

$$(\text{効率})_1 = \frac{R_1}{C} = 1$$

後者の効率は、

$$(\text{効率})_2 = \frac{R_2}{C} = \frac{3}{4}$$

である。

平均符号長、通信速度、符号容量

平均符号長、または平均通信時間は、符号の長さの期待値である。

$$\tau = \sum_i \tau_i p(i)$$

通信速度は、単位時間当たりの平均情報量である。

$$R = \frac{H}{\tau}$$

符号容量、または通信路容量は、通信速度の最大値である。

$$C = max\{R\}$$

2 元符号の符号容量

通信速度の対数をとって最大化すると符号容量は簡単に計算できる。確率の総和は 1 であるという条件があるので、ラグランジュの未定係数法を用いる。

2 元符号の 0 と 1 の通信時間が等しく 1sec とすると、

$$W^{-1} + W^{-1} = 1$$

を解くことで符号容量が求められる。この方程式を解くと、$W = 2$ であるから、$C = \log_2 W = 1$ として符号容量が得られる。

符号化の効率

符号化の効率は、符号容量に対する通信速度で定義される。

$$(\text{効率}) = \frac{R}{C}$$

4.

失われる情報量は、あいまい度を求めることで得られる。

問題文より、

$$p(y_A|x_A) = p(y_C|x_C) = 1$$

$$p(y_B|x_A) = p(y_C|x_A) = p(y_A|x_B) = p(y_A|x_C) = p(y_B|x_C) = 0$$

$$p(y_B|x_B) = \frac{3}{4}$$

$$p(y_C|x_B) = \frac{1}{4}$$

である。ここで、x は送信側であることを意味し、y は受信側であることを意味する。

したがって、

$$p(x_A|y_A) = p(x_B|y_B) = 1$$

$$p(x_B|y_A) = p(x_C|y_A) = p(x_A|t_B) = p(x_C|y_B) = p(x_A|y_C) = 0$$

$$p(x_B|y_C) = \frac{1}{5}$$

$$p(x_C|y_C) = \frac{4}{5}$$

である。また、

$$p(y_A) = \frac{1}{2}$$

$$p(y_B) = \frac{3}{16}$$

$$p(y_C) = \frac{5}{16}$$

である。

あいまい度 $H(x|y)$ は、

$$\begin{aligned}
H(x|y) =& p(y_A)H(x|y_A) + p(y_B)H(x|y_B) + p(y_C)H(x|y_C) \\
=& p(y_C)(-p(x_B|y_C)\log_2 p(x_B|y_C) - p(x_C|y_C)\log_2 p(x_C|y_C) \\
=& \frac{5}{16}\log_2 5 - \frac{1}{2} \\
=& 0.725
\end{aligned}$$

である。

これを情報源のエントロピーと比較して、

$$\frac{H(x|y)}{H(x)} = 48.4\%$$

である。

5.

最初から誤りを検出したり訂正できる符号を使う方法と、既にある符号に誤りを検出したり訂正したりするための符号を添える方法がある。

前者としては、ハミング距離を一定にした符号を使うことで目的が達成される。たとえば、ハミング距離が互いに 2 の符号を用いて通信するとする。受信した符号のハミング距離が 1 だった場合、誤りが生じていることが検出できる。ハミング距離の大きな符号を使うことで、より多くの誤りが検出できたり、誤りを訂正できたりする。

後者としては、パリティビットを添えることで目的が達成される。符号中の 1 の数が奇数であれば、パリティビットとして 1 を、符号中の 1 の数が偶数であれば、パリティビットとして 0 を添える。これによって、符号中の 1 の数は必ず偶数になる。受信した符号の 1 の数が奇数になっていた場合、誤りが生じていることが検出できる。パリティビットを複数用いることによって、より多くの誤りが検出できたり、誤りを訂正できたりする。

慶應義塾大学大学院理工学研究科　入試
2007年　電気電子工学　D5　情報工学

(1)

通信速度を求める問題である。情報源のエントロピーは、

$$H = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1\mathrm{bit}$$

平均の通信時間は、

$$\tau = \frac{1}{2}1\mathrm{sec} + \frac{1}{2}2\mathrm{sec} = \frac{3}{2}\mathrm{sec}$$

であるから、通信速度は、

$$R = \frac{H}{\tau} = \frac{2}{3}\mathrm{bit/sec}$$

である。

(2)

求める最大値を C として、

$$2^{-C} + 2^{-2C} = 1$$

を解けばよい。

(3)

$$W = 2^C$$

と置くと、方程式は、

$$W^{-1} + W^{-2} = 1$$

となる。

両辺に W^2 をかけると、

$$W^2 - W - 1 = 0$$

$$\therefore W = \frac{1+\sqrt{5}}{2} = 1.62$$

である。2 次方程式を解く際に複号が生じるが、最大値を求める問題なので、複号が正の場合のみを考える。

したがって、

$$C = \log_2 1.62 \approx -\log_2 0.618 \approx 0.694$$

(4)

○の割合を p、×の割合を q とすると、

$$p + q = 1$$

である。

通信速度は、

$$R = \frac{H}{\tau} = \frac{-p\log_2 p - q\log_2 q}{p + 2q}$$

であり、これを上の条件のもとで最大化する p と q を求める。

通信速度の対数をとり、ラグランジュの未定係数法を用いて、通信速度を最大化することを考える。つまり、

$$\log_2 R = \log_2 H - \log_2 \tau + \lambda(p + q - 1)$$

を最大化する p と q を求める。

上式を p と q で偏微分すると、

$$\frac{\partial}{\partial p}\log_2 R = \frac{1}{H_0}(-1 - \log_2 p) - \frac{1}{\tau_0} + \lambda = 0$$

$$\frac{\partial}{\partial q}\log_2 R = \frac{1}{H_0}(-1 - \log_2 q) - \frac{2}{\tau_0} + \lambda = 0$$

である。この 2 式から、λ を消去すると、

$$\frac{1}{H_0}\log\frac{q}{p} + \frac{1}{\tau_0} = 0$$

$$\frac{q}{p} = 2^{-C}$$

$$\therefore \frac{q}{p} = 2^{-0.281} \approx 0.823$$

である。これと、$p + q = 1$ を連立させて解くと、

$$p = 0.618$$

$$q = 0.382$$

である。

慶應義塾大学大学院理工学研究科　入試
2007 年　電気電子工学　D5　情報工学

(1)

アルファベットの送信側を x、受信側を y と表現することにする。

$$H(x) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{4}\log_2\frac{1}{4} = \frac{7}{4}$$

(2)

$$p(y_A) = p(x_A)p(y_A|x_A) + p(x_B)p(y_A|x_B) + p(x_C)p(y_A|x_C) + p(x_D)p(y_A|x_D) = \frac{1}{2}\cdot 1 + 0 + 0 + 0 = \frac{1}{2}$$

同様にして、

$$p(y_B) = \frac{7}{32}$$

$$p(y_C) = \frac{1}{8}$$

$$p(y_D) = \frac{5}{32}$$

(3)

あいまい度 $H(x|y)$ を求めて (1) で求めたエントロピーと比較する。

$$p(x_A|y_A) = p(x_C|y_C) = 1$$

$$p(x_B|y_A) = p(x_C|y_A) = p(x_D|y_A) = p(x_A|y_B) = p(x_C|y_B) = 0$$

$$p(x_A|y_C) = p(x_B|y_C) = p(x_D|y_C) = p(x_A|y_D) = p(x_C|y_D) = 0$$

$$p(x_B|y_B) = \frac{6}{7}$$

$$p(x_D|y_B) = \frac{1}{7}$$

$$p(x_B|y_D) = \frac{2}{5}$$

$$p(x_D|y_D) = \frac{3}{5}$$

であるから、

$$\begin{aligned}
H(x|y) =& p(y_A)H(x|y_A) + p(y_B)H(x|y_B) + p(y_C)H(x|y_C) + p(y_D)H(x|y_D) \\
=& \frac{7}{32}\left(-\frac{6}{7}\log_2\frac{6}{7} - \frac{1}{7}\log_2\frac{1}{7}\right) + \frac{5}{32}\left(-\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5}\right) \\
=& \frac{7}{32}\left(\log_2 7 - \frac{6}{7} - \frac{6}{7}\log_2 3\right) + \frac{5}{32}\left(\log_2 5 - \frac{3}{5}\log_2 3 - \frac{2}{5}\right) \\
=& 0.2823
\end{aligned}$$

したがって、

$$\frac{H(x|y)}{H(x)} = \frac{0.2823}{7/4} \fallingdotseq 16\%$$

慶應義塾大学大学院理工学研究科　入試
2005 年　電気電子工学　D5　情報工学

(1)

$$H = -\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3} = \log_2 3 - \frac{2}{3} = 0.918\mathrm{bit}$$

(2)

点灯と消灯の割合が 1/2、1/2 のとき、

$$H_m = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1\mathrm{bit}$$

(3)

電球 1 つで知らせることができる情報量は、(2) で求めた $H_m = 1\mathrm{bit}$ である。一方、10 段階の運転状況の情報量は、

$$-\log_2\frac{1}{10} = 1 + \log_2 5 = 3.322\mathrm{bit}$$

である。したがって、3.322 以上の整数である、4 個以上の電球が必要。

(4)

平均通信時間は、

$$\tau = \frac{1}{3}\cdot 1\mathrm{sec} + \frac{2}{3}\cdot 2\mathrm{sec} = \frac{5}{3}\mathrm{sec}$$

である。したがって、1 秒あたりの平均送信情報量は、

$$R = \frac{H}{\tau} = \frac{0.918}{5/3} = 0.551\mathrm{bit/sec}$$

(5)

点灯の割合を p、消灯の割合を q として、

$$p + q = 1$$

である。平均情報量と、平均通信時間は、

$$H = -p \log_2 p - q \log_2 q$$

$$\tau = p + 2q$$

である。これを用いて、通信速度は、

$$R = \frac{H}{\tau}$$

である。

通信速度の対数を最大化することを考える。

$$\log_2 R = \log_2 H - \log_2 \tau + \lambda(p + q - 1)$$

を、p と q について偏微分し、0 となるようにおく。

$$\frac{\partial}{\partial p} \log_2 R = \frac{1}{H_0}(-1 - \log_2 p) - \frac{1}{\tau_0} + \lambda = 0$$

$$\frac{\partial}{\partial q} \log_2 R = \frac{1}{H_0}(-1 - \log_2 q) - \frac{2}{\tau_0} + \lambda = 0$$

ここで、H_0 と τ_0 は通信速度を最大にした場合のエントロピーと平均通信時間であり、通信路容量 $C = H_0/\tau_0$ である。

この 2 式にそれぞれ p と q をかけてから足し合わせると、

$$\frac{1}{H_0}(-1 + H_0) - \frac{\tau_0}{\tau_0} + \lambda = 0$$

$$\therefore \lambda = \frac{1}{H_0}$$

これを元の 2 式に代入すると、

$$\frac{1}{H_0}(-\log_2 p) - \frac{1}{\tau_0} = 0$$

$$\therefore p = 2^{-C}$$

$$\frac{1}{H_0}(-\log_2 q) - \frac{2}{\tau_0} = 0$$

$$\therefore q = 2^{-2C}$$

これを $p + q = 1$ に代入すると、

$$2^{-C} + 2^{-2C} = 1$$

である。$W = 2^C$ とおくと、

$$W^{-1} + W^{-2} = 1$$

$$\therefore W^2 - W - 1 = 0$$

$$\therefore W = \frac{1 + \sqrt{5}}{2} = 1.618$$

したがって、平均情報量の最大値は、

$$C = \log_2 W = -\log_2 0.618 = -0.694$$

である。

平均情報量が最大となる場合の p と q は、

$$p = 2^{-C} = 0.618$$

$$q = 1 - p = 0.382$$

である。

慶應義塾大学大学院理工学研究科　入試
2004年　電気電子工学　D5　情報工学

(1)

$$H = -\frac{1}{4}\log_2\frac{1}{4}\times 3 - \frac{1}{8}\log_2\frac{1}{8}\times 2 = \frac{9}{4}\mathrm{bit}$$

(2)

$$\tau = \frac{1}{4}\cdot 2\mathrm{sec}\times 3 + \frac{1}{8}\cdot 1\mathrm{sec}\times 2 = \frac{13}{8}\mathrm{sec}$$

$$\therefore R = \frac{H}{\tau} = \frac{18}{13}$$

(3)

$$p_A = p_B = p_C = 2^{-2C}$$

$$p_D = p_E = 2^{-C}$$

より、

$$3\cdot 2^{-2C} + 2\cdot 2^{-C} = 1$$

$$\therefore 2^C = 3$$

$$\therefore C = \log_2 3 = 1.59$$

(3-1)

$$\therefore p_A = p_B = p_C = 2^{-3.18} = 0.11$$

$$p_D = p_E = 2^{-1.59} = 0.33$$

(3-2)

$$C = 1.59$$