平均符号長の限界と情報源符号化定理

情報源の1次エントロピー

- 情報源S:
  - 情報源記号$A=\{a_1, \cdots, a_M\}$
  - 各情報源記号の発生確率　$P(a_i)=p_i$
- 情報源符号化
  - 情報源記号a_iの符号化$K(a_i)=c_i$
  - 1情報源記号あたりの平均符号長 $L=l_1p_1+\cdots+l_Mp_M$
    ここで, $l_i=|c_i|$
- 情報源Sの1次エントロピー

$$H_1(S) = -\sum_{x \in A} P(x)\log_2 P(x) = -\sum_{i=1}^{M} p_i \log_2 p_i = \sum_{i=1}^{M} p_i \log_2 \frac{1}{p_i}$$

情報源の1次エントロピー

【補助定理3】

$p_1, \cdots, p_M$ を $p_1 + \cdots + p_M = 1$ となる非負の数，$q_1, \cdots, q_M$ を $q_1 + \cdots + q_M \leqq 1$ となる非負の数とする．$p_i \neq 0$ のとき，$q_i \neq 0$ とすれば，

$$-\sum_{i=1}^{M} p_i \log_2 q_i \geq -\sum_{i=1}^{M} p_i \log_2 p_i = H_1(S)$$

が成立する．等号はすべてのi について$q_i = p_i$ のときに限り成立する．

【補助定理3】の適用例

		$-\sum_{i=1}^{M} p_i \log_2 q_i$
p_1, p_2, p_3, p_4, p_5	0.6, 0.2, 0.1, 0.07, 0.03	
q_1, q_2, q_3, q_4, q_5	0.2, 0.2, 0.2, 0.2, 0.2	2.32193
	0.9 0.025 0.025 0.025 0.025	2.21997
	0.6, 0.2, 0.1, 0.07, 0.03	1.65908

情報源の1次エントロピー

【補助定理3】

$p_1, \cdots, p_M$ を $p_1 + \cdots + p_M = 1$ となる非負の数，$q_1, \cdots, q_M$ を $q_1 + \cdots + q_M \leqq 1$ となる非負の数とする．$p_i \neq 0$ のとき，$q_i \neq 0$ とすれば，

$$-\sum_{i=1}^{M} p_i \log_2 q_i \geq -\sum_{i=1}^{M} p_i \log_2 p_i = H_1(S)$$

が成立する．等号はすべてのi について$q_i=p_i$ のときに限り成立する．

【補助定理3】の証明の骨子

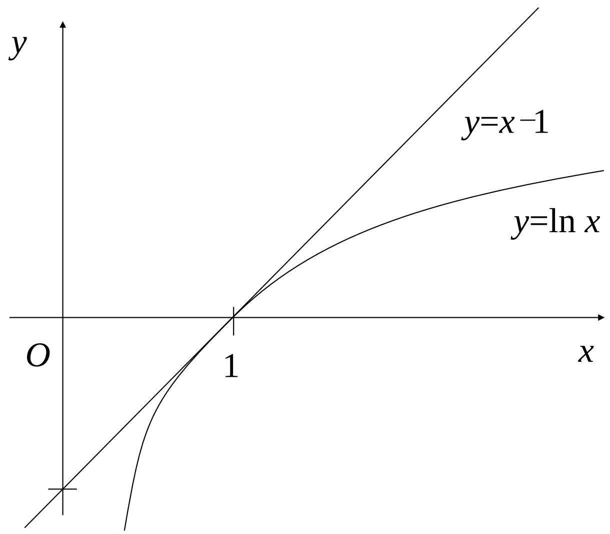
$$D = -\sum_{i=1}^{M} p_i \log_2 q_i + \sum_{i=1}^{M} p_i \log_2 p_i = -\sum_{i=1}^{M} p_i \log_2 \frac{q_i}{p_i} = -\sum_{i=1}^{M} \frac{p_i}{\ln 2} \ln \frac{q_i}{p_i}$$

と置く． $\ln x \leq x - 1$ であることを利用すると，

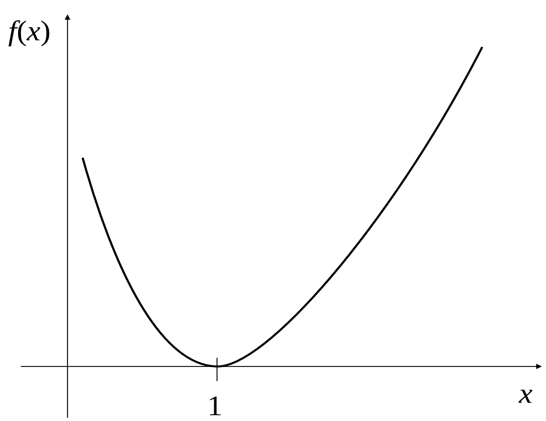
$$D = -\sum_{i=1}^{M} \frac{p_i}{\ln 2} \ln \frac{q_i}{p_i} \geq -\sum_{i=1}^{M} \frac{p_i}{\ln 2} \left(\frac{q_i}{p_i} - 1 \right) = \frac{1}{\ln 2} \sum_{i=1}^{M} (p_i - q_i) = \frac{1}{\ln 2} \left(\sum_{i=1}^{M} p_i - \sum_{i=1}^{M} q_i \right) \geq 0$$

■

情報源の1次エントロピー



$$\ln x \le x - 1$$



$$f(x) = x - 1 - \ln x$$

$$f(1) = 0,\ f'(x) = 1 - \frac{1}{x}$$

情報源の1次エントロピー

【定理】

情報源Sの各情報源記号を一意復号可能な2元符号に符号化すると，平均符号長Lは，$H_1(S) \leqq L$を満足する.

また，平均符号長Lが$L < H_1(S) + 1$となる瞬時符号を作ることができる.

情報源の1次エントロピー

【定理】の証明の骨子　（前半：$L \geq H_1(S)$ の証明）

- $q_i = 2^{-l_i}$ と置く（q_iを導入する）.
- すると，一意復号可能だから，$q_i > 0$ かつ, $l_1, l_2, \ldots, l_M$ はマクミランの不等式を満足するので,$2^{-l_1} + 2^{-l_2} + \cdots + 2^{-l_M} \leq 1$ つまり，$q_1 + \cdots + q_M \leq 1$ が満足される.
- 従って補助定理から，$-\sum_{i=1}^{M} p_i \log_2 q_i \geq -\sum_{i=1}^{M} p_i \log_2 p_i = H_1(S)$ となる.
- ここで左辺は，$-\sum_{i=1}^{M} p_i \log_2 q_i = -\sum_{i=1}^{M} p_i \log_2 2^{-l_i} = \sum_{i=1}^{M} p_i l_i = L$ である．等号が成立するのは，全てのiについて，$p_i = 2^{-l_i}$ のときである．■

情報源の1次エントロピー

【定理】の証明の骨子 （後半： $L \le H_1(S)+1$ を満足する瞬時符号が構成可能）

- $-\log_2 p_i \le l_i < -\log_2 p_i + 1$ なる整数l_iが一意に決まる.
- 得られたl_iに対して $2^{-l_i} \le 2^{\log_2 p_i} = p_i$ が成りたつ.
- $\sum 2^{-l_i} \le \sum p_i = 1$ となり，クラフトの不等式が満たされるので，l_iを符号長セットとする瞬時符号をつくることができる.
- この符号の平均符号長は，$L = \sum_{i=1}^{M} p_i l_i$ であるので，
  $-p_i \log_2 p_i \le p_i l_i < -p_i \log_2 p_i + p_i$ つまり，$H_1(S) \le L < H_1(S)+1$
  が満足される. ■

定理の意味すること

情報源S: ⟨A: 0.6, B : 0.2, C: 0.1, D: 0.07 , E: 0.03⟩

$$
\begin{aligned}
H_1(S) &= -\left(0.6\log_2 0.6 + 0.2\log_2 0.2 + 0.1\log_2 0.1 + 0.07\log_2 0.07 + 0.03\log_2 0.03\right) \\
&\approx 1.65908
\end{aligned}
$$

平均符号長が1.65908～2.65908となる瞬時符号を構成可能.

しかしこの定理は，そのような瞬時符号の構成法を教えてくれるわけではない.

⇒　ハフマン符号化などを用いてコンパクト符号を計算する.

⟨A→0, B→10, C→110, D→1110, E→1111⟩はコンパクト符号で平均符号長1.7

情報源記号ごとに符号化するかぎり，平均符号長を理論的下限に近づけることはできない

拡大情報源

- q元情報源Sのn次の拡大情報源S^n: S の連続する n 個の情報源記号列を情報源記号とするq^n元情報源.

- 無記憶情報源Sの連続するn個の出力は互いに独立であるから，その結合確率分布は，

$$P(x_0, x_1, \cdots, x_{n-1}) = P(x_0)P(x_1)\cdots P(x_{n-1})$$

となる.

- 無記憶情報源Sのn次拡大情報源をS^nとすれば，

$$H_1(S^n) = nH_1(S)$$

が成り立つ.

拡大情報源

$H_1(S^n) = nH_1(S)$ になるわけ $\Rightarrow$

$$
\begin{aligned}
&H_1(S^3) \\
&= -\sum_{x_0}\sum_{x_1}\sum_{x_2} P(x_0, x_1, x_2)\log_2 P(x_0, x_1, x_2) \\
&= -\sum_{x_0}\sum_{x_1}\sum_{x_2} P(x_0)P(x_1)P(x_2)\log_2\left(P(x_0)P(x_1)P(x_2)\right) \\
&= -\sum_{x_0}\sum_{x_1}\sum_{x_2} P(x_0)P(x_1)P(x_2)\log_2 P(x_0) \\
&\quad -\sum_{x_0}\sum_{x_1}\sum_{x_2} P(x_0)P(x_1)P(x_2)\log_2 P(x_1) \\
&\quad -\sum_{x_0}\sum_{x_1}\sum_{x_2} P(x_0)P(x_1)P(x_2)\log_2 P(x_2) \\
&= -\sum_{x_0} P(x_0)\log_2 P(x_0) \\
&\quad -\sum_{x_1} P(x_1)\log_2 P(x_1) \\
&\quad -\sum_{x_2} P(x_2)\log_2 P(x_2) \\
&= 3H_1(S)
\end{aligned}
$$

計算の様子 (1/2)

$$\sum_{x_1\{A,B,C\}}\sum_{x_2\{D,E,F\}}\sum_{x_3\in\{G,H,I\}}P(x_0)P(x_1)P(x_2)\log_2\left(P(x_0)P(x_1)P(x_2)\right)$$

$$
=\sum\begin{array}{l}
P(A)P(D)P(G)\log_2(P(A)P(D)P(G))\\
P(A)P(D)P(H)\log_2(P(A)P(D)P(H))\\
P(A)P(D)P(I)\log_2(P(A)P(D)P(I))\\
P(A)P(E)P(G)\log_2(P(A)P(E)P(G))\\
P(A)P(E)P(H)\log_2(P(A)P(E)P(H))\\
P(A)P(E)P(I)\log_2(P(A)P(E)P(I))\\
P(A)P(F)P(G)\log_2(P(A)P(F)P(G))\\
P(A)P(F)P(H)\log_2(P(A)P(F)P(H))\\
P(A)P(F)P(I)\log_2(P(A)P(F)P(I))\\
P(B)P(D)P(G)\log_2(P(B)P(D)P(G))\\
P(B)P(D)P(H)\log_2(P(B)P(D)P(H))\\
P(B)P(D)P(I)\log_2(P(B)P(D)P(I))\\
P(B)P(E)P(G)\log_2(P(B)P(E)P(G))\\
P(B)P(E)P(H)\log_2(P(B)P(E)P(H))\\
P(B)P(E)P(I)\log_2(P(B)P(E)P(I))\\
P(B)P(F)P(G)\log_2(P(B)P(F)P(G))\\
P(B)P(F)P(H)\log_2(P(B)P(F)P(H))\\
P(B)P(F)P(I)\log_2(P(B)P(F)P(I))\\
P(C)P(D)P(G)\log_2(P(C)P(D)P(G))\\
P(C)P(D)P(H)\log_2(P(C)P(D)P(H))\\
P(C)P(D)P(I)\log_2(P(C)P(D)P(I))\\
P(C)P(E)P(G)\log_2(P(C)P(E)P(G))\\
P(C)P(E)P(H)\log_2(P(C)P(E)P(H))\\
P(C)P(E)P(I)\log_2(P(C)P(E)P(I))\\
P(C)P(F)P(G)\log_2(P(C)P(F)P(G))\\
P(C)P(F)P(H)\log_2(P(C)P(F)P(H))\\
P(C)P(F)P(I)\log_2(P(C)P(F)P(I))
\end{array}
=\sum\begin{array}{l}
P(A)P(D)P(G)(\log_2P(A)+\log_2P(D)+\log_2P(G))\\
P(A)P(D)P(H)(\log_2P(A)+\log_2P(D)+\log_2P(H))\\
P(A)P(D)P(I)(\log_2P(A)+\log_2P(D)+\log_2P(I))\\
P(A)P(E)P(G)(\log_2P(A)+\log_2P(E)+\log_2P(G))\\
P(A)P(E)P(H)(\log_2P(A)+\log_2P(E)+\log_2P(H))\\
P(A)P(E)P(I)(\log_2P(A)+\log_2P(E)+\log_2P(I))\\
P(A)P(F)P(G)(\log_2P(A)+\log_2P(F)+\log_2P(G))\\
P(A)P(F)P(H)(\log_2P(A)+\log_2P(F)+\log_2P(H))\\
P(A)P(F)P(I)(\log_2P(A)+\log_2P(F)+\log_2P(I))\\
P(B)P(D)P(G)(\log_2P(B)+\log_2P(D)+\log_2P(G))\\
P(B)P(D)P(H)(\log_2P(B)+\log_2P(D)+\log_2P(H))\\
P(B)P(D)P(I)(\log_2P(B)+\log_2P(D)+\log_2P(I))\\
P(B)P(E)P(G)(\log_2P(B)+\log_2P(E)+\log_2P(G))\\
P(B)P(E)P(H)(\log_2P(B)+\log_2P(E)+\log_2P(H))\\
P(B)P(E)P(I)(\log_2P(B)+\log_2P(E)+\log_2P(I))\\
P(B)P(F)P(G)(\log_2P(B)+\log_2P(F)+\log_2P(G))\\
P(B)P(F)P(H)(\log_2P(B)+\log_2P(F)+\log_2P(H))\\
P(B)P(F)P(I)(\log_2P(B)+\log_2P(F)+\log_2P(I))\\
P(C)P(D)P(G)(\log_2P(C)+\log_2P(D)+\log_2P(G))\\
P(C)P(D)P(H)(\log_2P(C)+\log_2P(D)+\log_2P(H))\\
P(C)P(D)P(I)(\log_2P(C)+\log_2P(D)+\log_2P(I))\\
P(C)P(E)P(G)(\log_2P(C)+\log_2P(E)+\log_2P(G))\\
P(C)P(E)P(H)(\log_2P(C)+\log_2P(E)+\log_2P(H))\\
P(C)P(E)P(I)(\log_2P(C)+\log_2P(E)+\log_2P(I))\\
P(C)P(F)P(G)(\log_2P(C)+\log_2P(F)+\log_2P(G))\\
P(C)P(F)P(H)(\log_2P(C)+\log_2P(F)+\log_2P(H))\\
P(C)P(F)P(I)(\log_2P(C)+\log_2P(F)+\log_2P(I))
\end{array}
$$

$$
=\sum\begin{array}{l}
P(A)P(D)P(G)\log_2P(A)\\
P(A)P(D)P(H)\log_2P(A)\\
P(A)P(D)P(I)\log_2P(A)\\
P(A)P(E)P(G)\log_2P(A)\\
P(A)P(E)P(H)\log_2P(A)\\
P(A)P(E)P(I)\log_2P(A)\\
P(A)P(F)P(G)\log_2P(A)\\
P(A)P(F)P(H)\log_2P(A)\\
P(A)P(F)P(I)\log_2P(A)\\
P(B)P(D)P(G)\log_2P(B)\\
P(B)P(D)P(H)\log_2P(B)\\
P(B)P(D)P(I)\log_2P(B)\\
P(B)P(E)P(G)\log_2P(B)\\
P(B)P(E)P(H)\log_2P(B)\\
P(B)P(E)P(I)\log_2P(B)\\
P(B)P(F)P(G)\log_2P(B)\\
P(B)P(F)P(H)\log_2P(B)\\
P(B)P(F)P(I)\log_2P(B)\\
P(C)P(D)P(G)\log_2P(C)\\
P(C)P(D)P(H)\log_2P(C)\\
P(C)P(D)P(I)\log_2P(C)\\
P(C)P(E)P(G)\log_2P(C)\\
P(C)P(E)P(H)\log_2P(C)\\
P(C)P(E)P(I)\log_2P(C)\\
P(C)P(F)P(G)\log_2P(C)\\
P(C)P(F)P(H)\log_2P(C)\\
P(C)P(F)P(I)\log_2P(C)
\end{array}
+\sum\begin{array}{l}
P(A)P(D)P(G)\log_2P(D)\\
P(A)P(D)P(H)\log_2P(D)\\
P(A)P(D)P(I)\log_2P(D)\\
P(A)P(E)P(G)\log_2P(E)\\
P(A)P(E)P(H)\log_2P(E)\\
P(A)P(E)P(I)\log_2P(E)\\
P(A)P(F)P(G)\log_2P(F)\\
P(A)P(F)P(H)\log_2P(F)\\
P(A)P(F)P(I)\log_2P(F)\\
P(B)P(D)P(G)\log_2P(D)\\
P(B)P(D)P(H)\log_2P(D)\\
P(B)P(D)P(I)\log_2P(D)\\
P(B)P(E)P(G)\log_2P(E)\\
P(B)P(E)P(H)\log_2P(E)\\
P(B)P(E)P(I)\log_2P(E)\\
P(B)P(F)P(G)\log_2P(F)\\
P(B)P(F)P(H)\log_2P(F)\\
P(B)P(F)P(I)\log_2P(F)\\
P(C)P(D)P(G)\log_2P(D)\\
P(C)P(D)P(H)\log_2P(D)\\
P(C)P(D)P(I)\log_2P(D)\\
P(C)P(E)P(G)\log_2P(E)\\
P(B)P(E)P(H)\log_2P(E)\\
P(B)P(E)P(I)\log_2P(E)\\
P(B)P(F)P(G)\log_2P(F)\\
P(B)P(F)P(H)\log_2P(F)\\
P(B)P(F)P(I)\log_2P(F)
\end{array}
+\sum\begin{array}{l}
P(A)P(D)P(G)\log_2P(G)\\
P(A)P(D)P(H)\log_2P(H)\\
P(A)P(D)P(I)\log_2P(I)\\
P(A)P(E)P(G)\log_2P(G)\\
P(A)P(E)P(H)\log_2P(H)\\
P(A)P(E)P(I)\log_2P(I)\\
P(A)P(F)P(G)\log_2P(G)\\
P(A)P(F)P(H)\log_2P(H)\\
P(A)P(F)P(I)\log_2P(I)\\
P(B)P(D)P(G)\log_2P(G)\\
P(B)P(D)P(H)\log_2P(H)\\
P(B)P(D)P(I)\log_2P(I)\\
P(B)P(E)P(G)\log_2P(G)\\
P(B)P(E)P(H)\log_2P(H)\\
P(B)P(E)P(I)\log_2P(I)\\
P(B)P(F)P(G)\log_2P(G)\\
P(B)P(F)P(H)\log_2P(H)\\
P(B)P(F)P(I)\log_2P(I)\\
P(C)P(D)P(G)\log_2P(G)\\
P(C)P(D)P(H)\log_2P(H)\\
P(C)P(D)P(I)\log_2P(I)\\
P(C)P(E)P(G)\log_2P(G)\\
P(C)P(E)P(H)\log_2P(H)\\
P(C)P(E)P(I)\log_2P(I)\\
P(C)P(F)P(G)\log_2P(G)\\
P(C)P(F)P(H)\log_2P(H)\\
P(C)P(F)P(I)\log_2P(I)
\end{array}
$$

計算の様子 (1/2)

$$
\sum \begin{matrix}
P(A)P(D)P(G)\log_2 P(A)\\
P(A)P(D)P(H)\log_2 P(A)\\
P(A)P(D)P(I)\log_2 P(A)\\
P(A)P(E)P(G)\log_2 P(A)\\
P(A)P(E)P(H)\log_2 P(A)\\
P(A)P(E)P(I)\log_2 P(A)\\
P(A)P(F)P(G)\log_2 P(A)\\
P(A)P(F)P(H)\log_2 P(A)\\
P(A)P(F)P(I)\log_2 P(A)\\
P(B)P(D)P(G)\log_2 P(B)\\
P(B)P(D)P(H)\log_2 P(B)\\
P(B)P(D)P(I)\log_2 P(B)\\
P(B)P(E)P(G)\log_2 P(B)\\
P(B)P(E)P(H)\log_2 P(B)\\
P(B)P(E)P(I)\log_2 P(B)\\
P(B)P(F)P(G)\log_2 P(B)\\
P(B)P(F)P(H)\log_2 P(B)\\
P(B)P(F)P(I)\log_2 P(B)\\
P(C)P(D)P(G)\log_2 P(C)\\
P(C)P(D)P(H)\log_2 P(C)\\
P(C)P(D)P(I)\log_2 P(C)\\
P(C)P(E)P(G)\log_2 P(C)\\
P(C)P(E)P(H)\log_2 P(C)\\
P(C)P(E)P(I)\log_2 P(C)\\
P(C)P(F)P(G)\log_2 P(C)\\
P(C)P(F)P(H)\log_2 P(C)\\
P(C)P(F)P(I)\log_2 P(C)
\end{matrix}
+\sum \begin{matrix}
P(A)P(D)P(G)\log_2 P(D)\\
P(A)P(D)P(H)\log_2 P(D)\\
P(A)P(D)P(I)\log_2 P(D)\\
P(A)P(E)P(G)\log_2 P(E)\\
P(A)P(E)P(H)\log_2 P(E)\\
P(A)P(E)P(I)\log_2 P(E)\\
P(A)P(F)P(G)\log_2 P(F)\\
P(A)P(F)P(H)\log_2 P(F)\\
P(A)P(F)P(I)\log_2 P(F)\\
P(B)P(D)P(G)\log_2 P(D)\\
P(B)P(D)P(H)\log_2 P(D)\\
P(B)P(D)P(I)\log_2 P(D)\\
P(B)P(E)P(G)\log_2 P(E)\\
P(B)P(E)P(H)\log_2 P(E)\\
P(B)P(E)P(I)\log_2 P(E)\\
P(B)P(F)P(G)\log_2 P(F)\\
P(B)P(F)P(H)\log_2 P(F)\\
P(B)P(F)P(I)\log_2 P(F)\\
P(C)P(D)P(G)\log_2 P(D)\\
P(C)P(D)P(H)\log_2 P(D)\\
P(C)P(D)P(I)\log_2 P(D)\\
P(C)P(E)P(G)\log_2 P(E)\\
P(B)P(E)P(H)\log_2 P(E)\\
P(B)P(E)P(I)\log_2 P(E)\\
P(B)P(F)P(G)\log_2 P(F)\\
P(B)P(F)P(H)\log_2 P(F)\\
P(B)P(F)P(I)\log_2 P(F)
\end{matrix}
+\sum \begin{matrix}
P(A)P(D)P(G)\log_2 P(G)\\
P(A)P(D)P(H)\log_2 P(H)\\
P(A)P(D)P(I)\log_2 P(I)\\
P(A)P(E)P(G)\log_2 P(G)\\
P(A)P(E)P(H)\log_2 P(H)\\
P(A)P(E)P(I)\log_2 P(I)\\
P(A)P(F)P(G)\log_2 P(G)\\
P(A)P(F)P(H)\log_2 P(H)\\
P(A)P(F)P(I)\log_2 P(I)\\
P(B)P(D)P(G)\log_2 P(G)\\
P(B)P(D)P(H)\log_2 P(H)\\
P(B)P(D)P(I)\log_2 P(I)\\
P(B)P(E)P(G)\log_2 P(G)\\
P(B)P(E)P(H)\log_2 P(H)\\
P(B)P(E)P(I)\log_2 P(I)\\
P(B)P(F)P(G)\log_2 P(G)\\
P(B)P(F)P(H)\log_2 P(H)\\
P(B)P(F)P(I)\log_2 P(I)\\
P(C)P(D)P(G)\log_2 P(G)\\
P(C)P(D)P(H)\log_2 P(H)\\
P(C)P(D)P(I)\log_2 P(I)\\
P(C)P(E)P(G)\log_2 P(G)\\
P(C)P(E)P(H)\log_2 P(H)\\
P(C)P(E)P(I)\log_2 P(I)\\
P(C)P(F)P(G)\log_2 P(G)\\
P(C)P(F)P(H)\log_2 P(H)\\
P(C)P(F)P(I)\log_2 P(I)
\end{matrix}
=
$$

$$
\sum\begin{pmatrix} P(A)\log_2 P(A)\\ P(B)\log_2 P(B)\\ P(C)\log_2 P(C)\end{pmatrix}
\times
\sum\begin{pmatrix} P(D)P(G)\\ P(D)P(H)\\ P(D)P(I)\\ P(E)P(G)\\ P(E)P(H)\\ P(E)P(I)\\ P(F)P(G)\\ P(F)P(H)\\ P(F)P(I)\end{pmatrix}
+
\sum\begin{pmatrix} P(D)\log_2 P(D)\\ P(E)\log_2 P(E)\\ P(F)\log_2 P(F)\end{pmatrix}
\times
\sum\begin{pmatrix} P(A)P(G)\\ P(A)P(H)\\ P(A)P(I)\\ P(B)P(G)\\ P(B)P(H)\\ P(B)P(I)\\ P(C)P(G)\\ P(C)P(H)\\ P(C)P(I)\end{pmatrix}
+
\sum\begin{pmatrix} P(G)\log_2 P(G)\\ P(H)\log_2 P(H)\\ P(I)\log_2 P(I)\end{pmatrix}
\times
\sum\begin{pmatrix} P(A)P(D)\\ P(A)P(E)\\ P(A)P(F)\\ P(B)P(D)\\ P(B)P(E)\\ P(B)P(F)\\ P(C)P(D)\\ P(C)P(E)\\ P(C)P(F)\end{pmatrix}
$$

$$
= \sum\begin{pmatrix} P(A)\log_2 P(A)\\ P(B)\log_2 P(B)\\ P(C)\log_2 P(C)\end{pmatrix}
+ \sum\begin{pmatrix} P(D)\log_2 P(D)\\ P(E)\log_2 P(E)\\ P(F)\log_2 P(F)\end{pmatrix}
+ \sum\begin{pmatrix} P(G)\log_2 P(G)\\ P(H)\log_2 P(H)\\ P(I)\log_2 P(I)\end{pmatrix}
$$

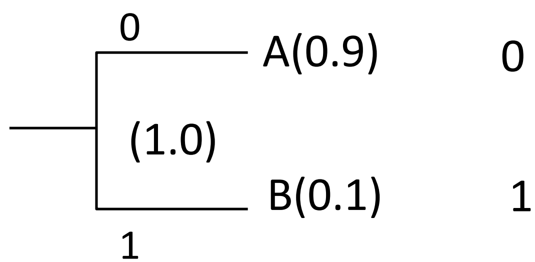
$$
\boxed{\sum\begin{pmatrix} P(D)P(G)\\ P(D)P(H)\\ P(D)P(I)\\ P(E)P(G)\\ P(E)P(H)\\ P(E)P(I)\\ P(F)P(G)\\ P(F)P(H)\\ P(F)P(I)\end{pmatrix} = \big(P(D)+P(E)+P(F)\big)\times\big(P(G)+P(H)+P(I)\big)=1}
$$

ブロック符号化

- ブロック符号化：情報源から発生する記号をまとめて符号化する方法.
- ブロック符号化をすることによって，平均符号長を情報源エントロピーに近づけることができる.

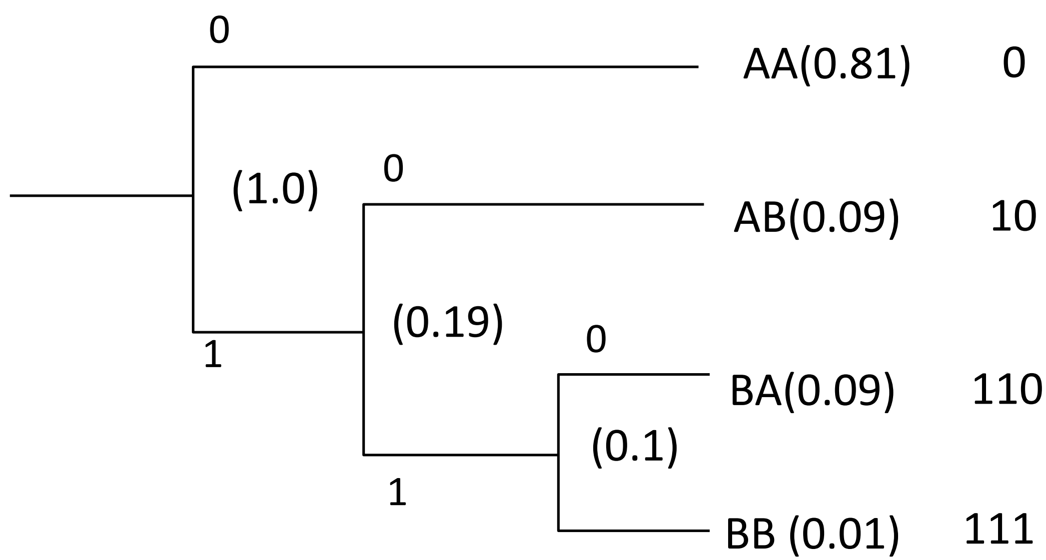
ブロック符号化

〈A: 0.9, B: 0.1〉に対するハフマン符号化



⇒ 情報源記号あたりの平均符号長1

〈AA: 0.81, AB: 0.09, BA: 0.09, BB: 0.01〉に対するハフマン符号化



⇒ 情報源記号あたりの平均符号長0.645

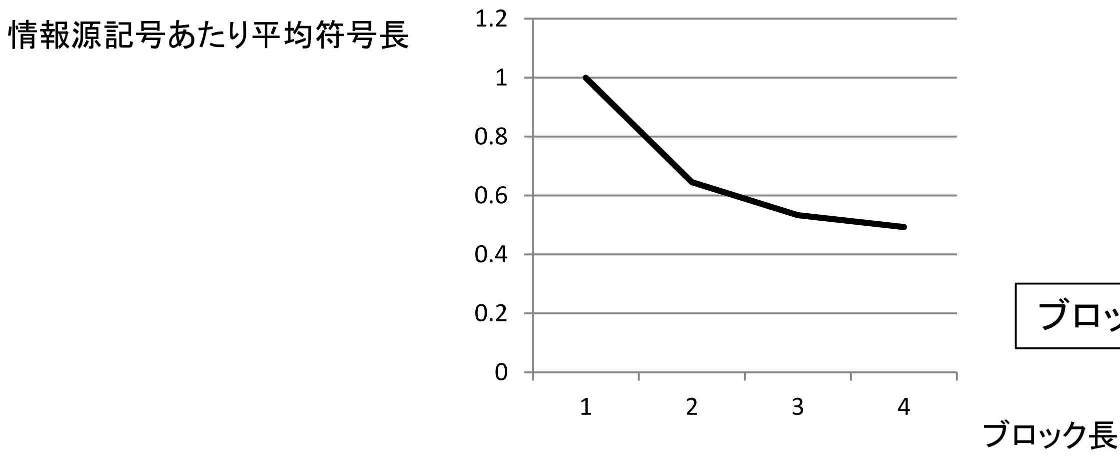
ブロック符号化

〈A: 0.9, B: 0.1〉の3次拡大情報源に対するハフマン符号化

AAA: 0.729　0
AAB: 0.081　100
ABA: 0.081　101
BAA: 0.081　110
ABB: 0.009　11100
BAB: 0.009　11101
BBA: 0.009　11110
BBB: 0.001　11111

⇒ 情報源記号あたりの平均符号長約0.53267

ブロック符号化の威力



ブロック長を大きくしたらどうなるか？

情報源符号化定理

情報源 Sについて，

$$H_1(S) \le L < H_1(S) + 1$$

が成立する．

情報源 Sのn次の拡張 S^nについても，

$$H_1(S^n) \le L_n < H_1(S^n) + 1$$

ここで，L_n は S^nの1記号あたりの平均符号長．Sが無記憶であれば

$$H_1(S^n) = nH_1(S)$$

また，$L_n = nL$ なので，

$$H_1(S) \le L = \frac{L_n}{n} < H_1(S) + \frac{1}{n}$$

$\rightarrow$ブロック符号化によって，平均符号長を $H_1(S)$に限りなく近づけられる

情報源符号化定理

【情報源符号化定理】
情報源 Sは，任意の正数 εに対して，１情報源記号あたりの平均符号長 Lが，

$$H(S) \le L < H(S) + \varepsilon$$

となるような2元瞬時符号に符号化できる.

しかし，どのような一意復号可能な2元符号を用いても，平均符号長がこの式の左辺より小さくなる符号化はできない.

情報源符号化定理

情報源Sのエントロピー$H(S)$:

$$H(S) \equiv \lim_{n\to\infty} H_n(S) = \lim_{n\to\infty} \frac{H_1(S^n)}{n}$$

ここで， $H_1(S^n) = -\sum \cdots \sum P(x_0, \cdots, x_{n-1}) \log_2 P(x_0, \cdots, x_{n-1})$

はSのn次の拡大情報源 S^nの1次エントロピー.

$H_n(S) \equiv \dfrac{H_1(S^n)}{n}$ はSの1情報源記号あたりのn次エントロピー.

一般には， $H(S) \le H_n(S)$ であることが知られている.

まとめ

- 情報源の1次エントロピー
- 補助定理3
- 情報源記号ごとに瞬時符号を構成した場合の平均符号長の下限
- 拡大情報源とその1次エントロピー
- ブロック符号化：情報源から発生する記号をまとめて符号化する方法.
- ブロック符号化をすることによって，平均符号長を情報源エントロピーに近づけることができる.
- 情報源符号化定理