符号への要請

- **一意符号:** $\mathcal{C}^* : S^* \longrightarrow T^*$ **: 単射**
- **瞬時符号:** $\mathcal{C}(\boldsymbol{x}) = \mathcal{C}(s)\boldsymbol{w} \Longrightarrow \boldsymbol{x} = s\boldsymbol{y}$
  (最初に届いた符号語で最初の文字が復元できる)
  (以上は生起確率 P には依らない)
- **効率が良い** $\cdots$ **平均符号長** $L(\mathcal{C})$ **が小さい**
  (これは生起確率 P に依る)

Kraft の不等式

$S = \{s_1, \ldots, s_k\}, \qquad \#T = r$

自然数列 $(\ell_1, \ldots, \ell_k)$ **に対し、**

各符号語長 $|\mathcal{C}(s_i)| = \ell_i$ **なる**

r **元 瞬時符号 が存在**

$$\iff \qquad \sum_{i=1}^{k} \frac{1}{r^{\ell_i}} \leq 1$$

McMillan の不等式

$$S = \{s_1, \ldots, s_k\}, \qquad \#T = r$$

自然数列 $(\ell_1, \ldots, \ell_k)$ **に対し、**

各符号語長 $|\mathcal{C}(s_i)| = \ell_i$ **なる**

r **元 一意符号 が存在**

$$\iff \qquad \sum_{i=1}^{k} \frac{1}{r^{\ell_i}} \leq 1$$

次に、生起確率を考慮に入れて、

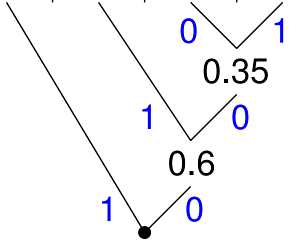
平均符号長の小さい符号の構成を考えよう

$\longrightarrow$ **Huffman 符号**

平均符号長の小さい符号の構成 (Huffman 符号)

例: $\#S = 4 = 2^2$, **平均符号長:** $1.95 (< 2)$

	1	01	000	001
S	a	b	c	d
P	0.4	0.25	0.2	0.15



Huffman 符号

- **平均符号長 $L(\mathcal{C})$ の最小値を実現**
  $\cdots$ **「最適符号 (optimal code)」**

- **各文字の生起確率 $P(s)$ の**
  “ばらつき” が大きいほど効果的

$\#S = 2$ **だったら、**

どうやっても

(生起確率に関わらず) $L(\mathcal{C}) = 1$ **か？**

(何かうまい手はないか)

$\longrightarrow$ **"拡大情報源" を考える**

$\#S = 2$ **だったら、**

どうやっても

(生起確率に関わらず) $L(\mathcal{C}) = 1$ **か？**

(何かうまい手はないか)

$\longrightarrow$ **“拡大情報源” を考える**

問題: 次の生起確率を持つ情報源 $\mathcal{S} = (S, P), S = \{a, b\}$ について、

S	a	b
P	0.8	0.2

(1) 2 **次の拡大情報源** $\mathcal{S}^2 = (S^2, P^{\otimes 2})$ **に対する Huffman 符号** $\mathcal{C}_2$ **を構成し、“**1 **文字当たりの平均符号長”** $L(\mathcal{C}_2)/2$ **を求めよ。**

(2) (時間に余裕があれば) 3 **次の拡大情報源** $\mathcal{S}^3 = (S^3, P^{\otimes 3})$ **に対しても同様の計算をせよ。**

n **次の拡大情報源** $\mathcal{S}^n = (S^n, P^{\otimes n})$

の符号 $\mathcal{C}_n$ **を利用すると、**

情報源 alphabet 1 **文字当たりの平均符号長**

$$\frac{L(\mathcal{C}_n)}{n}$$

はどこまで下げられるか？

$\longrightarrow$ **情報源** $\mathcal{S}$ **が本来持っている "情報の量" より小さくはならないだろう。**

"エントロピー (entropy)"

n **次の拡大情報源** $\mathcal{S}^n = (S^n, P^{\otimes n})$
の符号 $\mathcal{C}_n$ **を利用すると、**

情報源 alphabet 1 **文字当たりの平均符号長**

$$\frac{L(\mathcal{C}_n)}{n}$$

はどこまで下げられるか？

$\longrightarrow$ **情報源** $\mathcal{S}$ **が本来持っている "情報の量"**
より小さくはならないだろう。

"エントロピー (entropy)"

n 次の拡大情報源 $\mathcal{S}^n = (S^n, P^{\otimes n})$
の符号 $\mathcal{C}_n$ を利用すると、

情報源 alphabet 1 文字当たりの平均符号長

$$\frac{L(\mathcal{C}_n)}{n}$$

はどこまで下げられるか？

$\longrightarrow$ 情報源 $\mathcal{S}$ が本来持っている "情報の量"
より小さくはならないだろう。

"エントロピー (entropy)"

“情報の量”

「或る事象 P が起こる」

という “情報の価値” は

どう評価したら良いか？

「事象 P が起こる」という "情報の価値" $I(P)$

要請:

(1) 生起確率 p のみに依る

$$\longrightarrow I(p) := I(P)$$

(2) 独立な事象 P_1, P_2 に対して、

$$I(P_1 \wedge P_2) = I(P_1) + I(P_2)$$

$$\longrightarrow I(p_1 p_2) = I(p_1) + I(p_2)$$

(3) $I : [0,1] \longrightarrow \boldsymbol{R}_{\geq 0}$ ： 連続関数

$$\longrightarrow \boxed{I(p) = C \log \frac{1}{p} = -C \log p \quad (C \geq 0)}$$

「事象 P が起こる」という "情報の価値" $I(P)$

要請:

(1) **生起確率 p のみに依る**

$$\longrightarrow I(p) := I(P)$$

(2) **独立な事象 P_1, P_2 に対して、**

$$I(P_1 \wedge P_2) = I(P_1) + I(P_2)$$
$$\longrightarrow I(p_1 p_2) = I(p_1) + I(p_2)$$

(3) $I : [0,1] \longrightarrow \boldsymbol{R}_{\geq 0}$ **: 連続関数**

$$\longrightarrow \boxed{I(p) = C \log \frac{1}{p} = -C \log p \quad (C \geq 0)}$$

$$I(p) = C \log \frac{1}{p} = -C \log p \quad (C \geq 0)$$

定数 C の取り方
$\longleftrightarrow$ $\log$ **の底の取り方**
$\longleftrightarrow$ **“情報量” の単位の取り方**

通常 2 を底に取って ($I(\frac{1}{2}) := 1$)、

“情報量” の単位とする: bit (binary digit)

$$I(p) = C \log \frac{1}{p} = -C \log p \quad (C \geq 0)$$

定数 C の取り方
$\longleftrightarrow$ $\log$ **の底の取り方**
$\longleftrightarrow$ **“情報量” の単位の取り方**

通常 2 を底に取って ($I(\frac{1}{2}) := 1$)、

“情報量” の単位とする: bit (binary digit)

情報源 $\mathcal{S} = (S, P)$ の
1 文字から得られる平均情報量

$$\boxed{H(\mathcal{S}) := \sum_{s \in S} P(s) I(P(s))}$$

：情報源 $\mathcal{S}$ の <u>エントロピー (entropy)</u>

$S = \{s_1, \ldots, s_k\}, P(s_i) = p_i$ **の時は、**

$$H(\mathcal{S}) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i} = -\sum_{i=1}^{k} p_i \log p_i$$

特に $\#S = 2$ **の時、**

- **起きる確率** p
- **起きない確率** $1 - p =: \overline{p}$

$$H(p) := H(\mathcal{S}) = p \log \frac{1}{p} + \overline{p} \log \frac{1}{\overline{p}}$$

: <u>2 **値エントロピー関数**</u>

問: $y = I(p), y = H(p)$ **のグラフの概形を描け。(最大値をとる** p **とその時の値は ?)**