

情報源を効率良く符号化する話は一段落。

雑音の入る通信路を介して

これを誤りなく (効率良く) 伝達するには?

→ 符号理論 (誤り訂正符号)

通常、我々が通信するとき、

一般には雑音が入って正しくは伝わらない。

- それでも正しく伝えるにはどうするか？
→ 繰り返し言う・別の言い方をする・...
- 「何かおかしい」と気付けるのは何故か？
→ あり得ないから !!

「あり得ない」とは？

符号 $C : S \rightarrow T^+$ で、
受信した語 $y \in T^+$ が $y \notin \text{Im}C^*$
(誤り検出)

正しくは何だったのか？

y に “一番近い” $x \in \text{Im}C^*$ だろう !!
(誤り訂正)

受信語 $y \notin \text{Im}C^* \subset T^*$ で誤り検出

→ T^* 全部は使わない

→ 冗長度を持たせて誤り検出・訂正

とは言え

- より効率良く (冗長度少なめ)
- より高い誤り対処性能を持つ
(誤りが沢山あっても大丈夫)

ものが望ましい。

以下では、

- 生起確率の違いを考慮しない
- 等長符号のみを考える
(全ての符号語が同じ長さ)

効率良い情報源符号で符号化された文字列を、
一定の個数毎に切って、再符号化 (通信路符号)

符号 $C : S \longrightarrow V := T^n$ (n : 符号語長)
の像 $\text{Im}C =: U \subset V$ のみが大事

→ 寧ろ、
像 U をも単に C と書き、符号と呼ぶ。

誤り訂正符号の性能を表すには:

- 符号長 n (当座固定)
- 符号語数 $M := \#C$ (大きいほど良い)
- 訂正出来る誤りの数 t (大きいほど良い)

しかし一般に、

「 $M \rightarrow$ 大」と「 $t \rightarrow$ 大」とは

相反する要求!!

符号 $\mathcal{C} \subset V = T^n$ で、

- 受信語 $y \notin \mathcal{C}$ によって誤り検出
- y に “一番近い” $x \in \mathcal{C}$ が正しい、
として誤り訂正

→ “一番近い” とは？

→ V に “距離” を導入
(通常 **Hamming 距離** を用いる)

距離の公理

X : 集合

$d : X \times X \longrightarrow \mathbf{R}_{\geq 0}$: X 上の距離 (metric)

であるとは、

- $d(x, y) = 0 \iff x = y$
- $d(x, y) = d(y, x)$
- $d(x, y) + d(y, z) \geq d(x, z)$
(三角不等式 (triangle inequality))

Hamming 距離

$V = T^n$ 上に次で定まる距離

$$d : V \times V \longrightarrow \mathbf{R}$$

を V 上の **Hamming 距離** と呼ぶ:

$\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_n) \in V$ に対し、

$$d(\mathbf{x}, \mathbf{y}) := \#\{i \mid x_i \neq y_i\}$$

$$\mathcal{C} \subset V = T^n$$

n 箇所のうち何箇所違っても訂正できるか？

= 距離が幾ら以内なら訂正できるか？

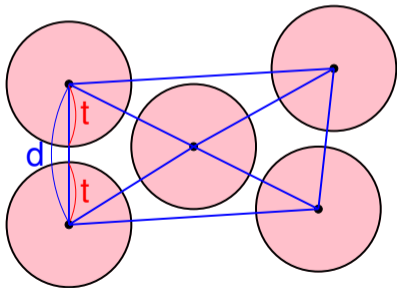
t 箇所違っても一意に訂正できる

$\iff \forall \mathbf{y} \in V$ に対し

$d(\mathbf{x}, \mathbf{y}) \leq t$ なる $\mathbf{x} \in \mathcal{C}$ は高々 1 つ
($\#\{\mathbf{x} \in \mathcal{C} \mid d(\mathbf{x}, \mathbf{y}) \leq t\} \leq 1$)

$d := \min\{d(\mathbf{x}, \mathbf{x}') \mid \mathbf{x}, \mathbf{x}' \in \mathcal{C}, \mathbf{x} \neq \mathbf{x}'\}$
: \mathcal{C} の最小距離 (minimum distance)

$d \geq 2t+1$ なら一意に訂正可能 $\longrightarrow t = \left\lfloor \frac{d-1}{2} \right\rfloor$



訂正性能は最小距離で計れる!!

誤り訂正符号の性能を表すには:

- 符号長 n (当座固定)
- 符号語数 $M := \#C$ (大きいほど良い)
- 最小距離 d (大きいほど良い)

→ (n, M, d) -符号

「 $M \longrightarrow \text{大}$ 」と「 $d \longrightarrow \text{大}$ 」とは
相反する要求!!

問題: n, d を固定した時の M の上限は ?

$\longrightarrow n, d$ と $q := \#T$ とで評価する