**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?**

$\longrightarrow$ **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？**
  **$\longrightarrow$ 繰返し言う・別の言い方をする・$\cdots$**

- **「何かおかしい」と気付けるのは何故か？**
  **$\longrightarrow$ あり得ないから !!**

## 「あり得ない」とは？

**符号** $\mathcal{C} : S \longrightarrow T^+$ **で、**

**受信した語** $y \in T^+$ **が** $y \notin \mathrm{Im}\mathcal{C}^*$

**(誤り検出)**

**正しくは何だったのか？**

$y$ **に “一番近い”** $x \in \mathrm{Im}\mathcal{C}^*$ **だろう !!**

**(誤り訂正)**

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

$\longrightarrow$ **$T^*$ 全部は使わない**

$\longrightarrow$ **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** $\mathcal{C} : S \longrightarrow V := T^n$ **(**$n$**：符号語長)**

**の像** $\mathrm{Im}\mathcal{C} =: U \subset V$ **のみが大事**

$\longrightarrow$ **寧ろ、**

**像** $U$ **をも単に** $\mathcal{C}$ **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 $M := \#\mathcal{C}$ (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、**

**「$M \longrightarrow$ 大」と「$t \longrightarrow$ 大」とは**

**相反する要求!!**

**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 $M := \#\mathcal{C}$ (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、**

**「$M \longrightarrow$ 大」と「$t \longrightarrow$ 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**

**$\longrightarrow$ “一番近い” とは？**

**$\longrightarrow$ $V$ に “距離” を導入**
**(通常 Hamming 距離 を用いる)**

## 距離の公理

$X$：集合

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

**であるとは、**

- $d(x, y) = 0 \Longleftrightarrow x = y$
- $d(x, y) = d(y, x)$
- $d(x, y) + d(y, z) \geq d(x, z)$

**(三角不等式 (triangle inequality))**

## Hamming 距離

$V = T^n$ 上に次で定まる距離

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

を $V$ 上の **Hamming 距離** と呼ぶ:

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

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

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

$n$ **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**

$\Longleftrightarrow \forall \boldsymbol{y} \in V$ **に対し**

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

$(\#\{\boldsymbol{x} \in \mathcal{C} | d(\boldsymbol{x}, \boldsymbol{y}) \leq t\} \leq 1)$

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

$n$ **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**

$\Longleftrightarrow \forall \boldsymbol{y} \in V$ **に対し**

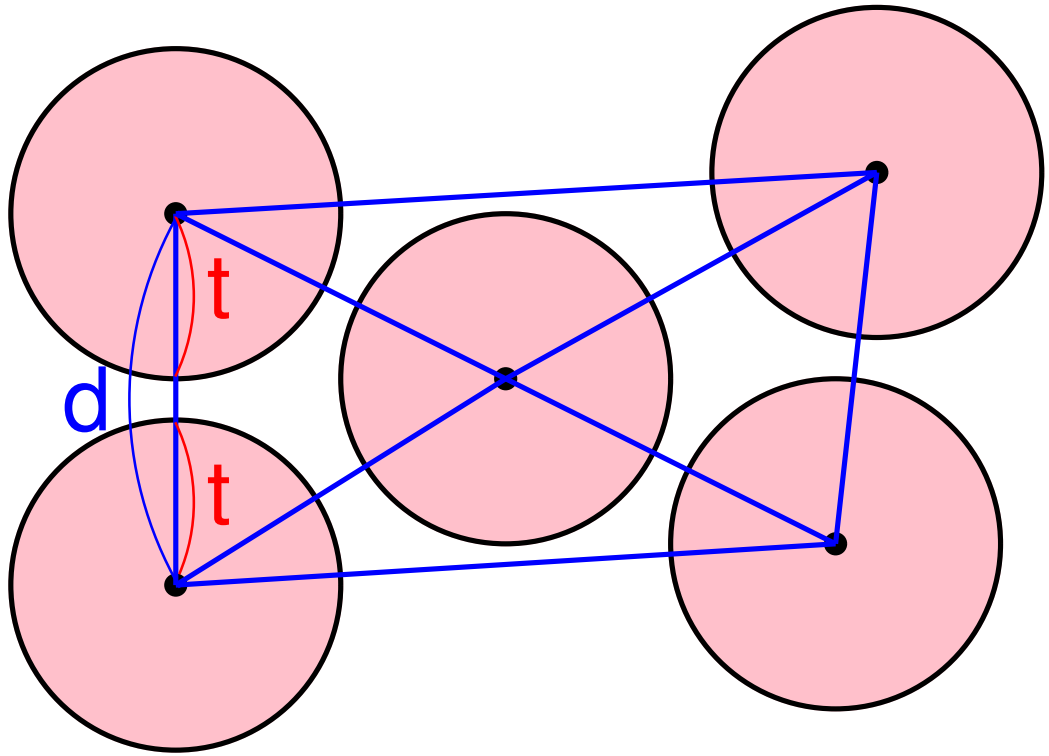
$d(\boldsymbol{x}, \boldsymbol{y}) \leq t$ **なる** $\boldsymbol{x} \in \mathcal{C}$ **は高々** $1$ **つ**

$(\#\{\boldsymbol{x} \in \mathcal{C} | d(\boldsymbol{x}, \boldsymbol{y}) \leq t\} \leq 1)$

$d := \min\{d(\boldsymbol{x}, \boldsymbol{x}') | \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{C}, \boldsymbol{x} \neq \boldsymbol{x}'\}$

**: $\mathcal{C}$ の最小距離 (minimum distance)**

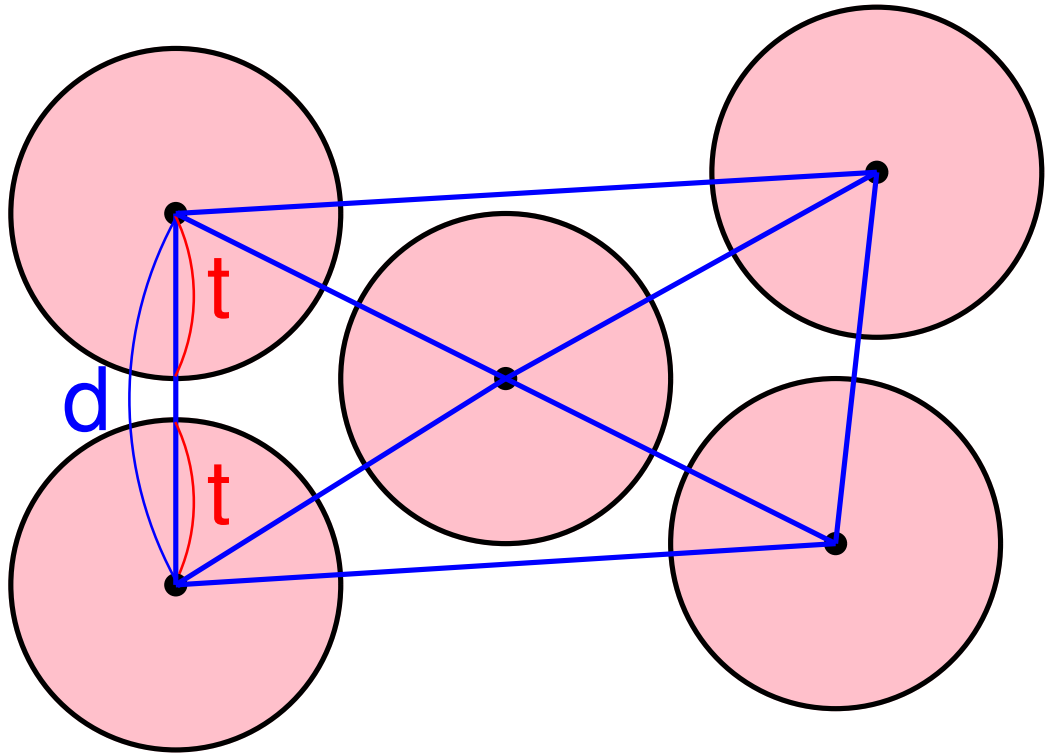
$d \geq 2t+1$ **なら一意に訂正可能** $\longrightarrow t = \left\lfloor \dfrac{d-1}{2} \right\rfloor$



$d := \min\{d(\boldsymbol{x}, \boldsymbol{x}') | \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{C}, \boldsymbol{x} \neq \boldsymbol{x}'\}$

**: $\mathcal{C}$ の最小距離 (minimum distance)**

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



**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** $M := \#\mathcal{C}$ **(大きいほど良い)**
- **最小距離** $d$ **(大きいほど良い)**

$$\longrightarrow \underline{(n, M, d)\text{-符号}}$$

**「$M \longrightarrow$ 大」と「$d \longrightarrow$ 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**