誤り訂正符号 (Error-correcting codes)

情報通信にはノイズ (雑音) が妨げとなる

Noises are obstructive in communication.

雑音の入る通信路を介して情報通信を行なう際、
通信途中での誤りに如何に対処するか?

When trying to communicate
with noisy channel,
how can one overcome the errors ?

誤りに対処しつつ、如何に効率的に伝達するか?

Dealing with the errors,
how efficiently can one communicate ?

$\longrightarrow$ **Coding theory(符号理論)**

- **Error-correcting codes(誤り訂正符号)**

誤り訂正符号 (Error-correcting codes)

情報通信中に誤り (らしきこと) に出遭ったら ?
How should we do when we meet
a (suspicious) error in communication ?

- **聞き直す (より安全なプロトコルの採用)**
  Ask again (more secure protocol)
- **見当を付ける (誤りの自動訂正)**
  Make a guess (error-correcting)

$\longrightarrow$ **冗長性を利用して安全性を確保**
Use of redundancy to keep certainty

誤り訂正符号 (Error-correcting codes)

「誤りの自動訂正」が出来るように
予め適切に冗長性を持たせて通信する
Give some controlled redundancy beforehand
to do automatic error-correction.

- **誤り訂正性能は高く**
  (High error-correction ability)
- **とは言えなるべく効率的に**
  (As efficient as possible)

$\longrightarrow$ 有限体上の線型代数・代数幾何などの利用
(Use of linear algebra and algebraic geometry
over finite fields)

誤り訂正符号 (Error-correcting codes)

$\mathcal{C}^* : S^* \longrightarrow \mathsf{T}^*$: a code

$y \in \mathsf{T}^*$: a received word

- **誤り検出 (error-detection):**
  Detected an error if $y \notin \mathrm{Im}\mathcal{C}^*$.

- **誤り訂正 (error-correction):**
  Guess the "nearest" $x \in \mathrm{Im}\mathcal{C}^*$ to y as the correct word.

誤り訂正符号 (Error-correcting codes)

Error-detection(誤り検出): $y \notin \mathrm{Im}\mathcal{C}^* \subset \mathrm{T}^*$

$\longrightarrow$ $\mathcal{C}^* : S^* \longrightarrow \mathrm{T}^*$ **should not be surjective.**

$\longrightarrow$ **適切に冗長度を持たせて誤り検出・訂正**
Give some controlled redundancy for error-correction.

Requirement(要請):

- **More efficient (less redundant)**
- **Higher error-correcting ability (can correct many errors)**

These are conflicting !!

誤り訂正符号 (Error-correcting codes)

Error-detection(誤り検出): $\mathrm{y} \notin \mathrm{Im}\mathcal{C}^* \subset \mathrm{T}^*$

$\longrightarrow$ $\mathcal{C}^* : S^* \longrightarrow \mathrm{T}^*$ **should not be surjective.**

$\longrightarrow$ **適切に冗長度を持たせて誤り検出・訂正**
Give some controlled redundancy for error-correction.

Requirement(要請):

- **More efficient (less redundant)**
- **Higher error-correcting ability (can correct many errors)**

These are conflicting !!

誤り訂正符号 (Error-correcting codes)

Error-detection(誤り検出): $\mathrm{y} \notin \mathrm{Im}\mathcal{C}^* \subset \mathrm{T}^*$

$\longrightarrow$ $\mathcal{C}^* : S^* \longrightarrow \mathrm{T}^*$ **should not be surjective.**

$\longrightarrow$ **適切に冗長度を持たせて誤り検出・訂正**
Give some controlled redundancy for error-correction.

Requirement(要請):

- **More efficient (less redundant)**
- **Higher error-correcting ability (can correct many errors)**

These are conflicting !!

誤り訂正符号 (Error-correcting codes)

In the following, we

- **do not consider the occurence probability, (生起確率の違いを考慮しない)**
- **consider only "block codes" (等長符号) (all code-words have the same length).**

誤り訂正符号 (Error-correcting codes)

状況設定:
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する、と想定

Situation:
We already have a sequence
encoded using suitable source-coding.
Cut it into blocks with a fixed length,
and then re-encode them for error-correction.

$\longrightarrow$ **channel coding (通信路符号)**

誤り訂正符号 (Error-correcting codes)

$\mathcal{C} : S \longrightarrow \mathrm{V} := \mathrm{T}^{\mathrm{n}}$: **a block code (等長符号)**
(n **: code-word length)**

符号 $\mathcal{C}$ **の誤り訂正性能は、**
その像 $\mathrm{Im}\mathcal{C} =: \mathrm{U} \subset \mathrm{V}$ **のみに依る。**
The ability of error-correction of $\mathcal{C}$
depends only on $\mathrm{Im}\mathcal{C} =: \mathrm{U} \subset \mathrm{V}$**.**

$\longrightarrow$ **We denote the image** U **by** $\mathcal{C}$ **simply,**
and call it a "code": $\mathcal{C} \subset \mathrm{V}$**.**

誤り訂正性能 (the ability of error-correction)

- the code-word length (符号長) n
  (temporarily fix)
- the number of code-words (符号語数)
  $M := \#\mathcal{C}$ (larger, better)
- the number of correctable errors t
  (larger, better)

But in general,

"larger M" and "larger t"

are conflicting requests !!

誤り訂正性能 (the ability of error-correction)

- the code-word length (符号長) n
  (temporarily fix)
- the number of code-words (符号語数)
  $\mathrm{M} := \#\mathcal{C}$ (larger, better)
- the number of correctable errors t
  (larger, better)

But in general,

"larger M" and "larger t"

are conflicting requests !!

誤りの訂正 (Error-correction)

符号 $\mathcal{C} \subset V = T^n$ で、

- **受信語 $y \notin \mathcal{C}$ によって誤り検出**
  Detect an error if $y \notin \mathrm{Im}\mathcal{C}^*$.
- **y に “一番近い” $x \in \mathcal{C}$ が正しい、**
  として誤り訂正
  Guess the “nearest” $x \in \mathrm{Im}\mathcal{C}^*$ to y
  as the correct word.

  $\longrightarrow$ **What is “nearest” ?**

$\longrightarrow$ **Introduce a “metric” into V**
(Hamming metric, usually)

誤りの訂正 (Error-correction)

符号 $\mathcal{C} \subset V = T^n$ で、

- **受信語 $y \notin \mathcal{C}$ によって誤り検出**
  Detect an error if $y \notin \mathrm{Im}\mathcal{C}^*$.
- **y に “一番近い” $x \in \mathcal{C}$ が正しい、**
  として誤り訂正
  Guess the “nearest” $x \in \mathrm{Im}\mathcal{C}^*$ to y
  as the correct word.
  $\longrightarrow$ What is “nearest” ?

$\longrightarrow$ Introduce a “metric” into V
(Hamming metric, usually)

誤りの訂正 (Error-correction)

符号 $\mathcal{C} \subset V = \mathrm{T}^n$ で、

- **受信語 $y \notin \mathcal{C}$ によって誤り検出**
  Detect an error if $y \notin \mathrm{Im}\mathcal{C}^*$.
- **y に “一番近い” $x \in \mathcal{C}$ が正しい、**
  として誤り訂正
  Guess the “nearest” $x \in \mathrm{Im}\mathcal{C}^*$ to y
  as the correct word.
  $\longrightarrow$ What is “nearest” ?

$\longrightarrow$ Introduce a “metric” into V
(Hamming metric, usually)

距離の公理 (Axiom of metric)

X : a set

$d : X \times X \longrightarrow \mathbf{R}_{\geq 0}$: a **metric (distance)** on X
(X 上の**距離**)

if d satisfies the following:

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

Hamming 距離 (Hamming distance)

Hamming distance on $V = T^n$

$$d : V \times V \longrightarrow \mathbf{R}_{\geq 0}$$

is defined as follows:

for $\boldsymbol{x} = (x_1, \ldots, x_n), \boldsymbol{y} = (y_1, \ldots, y_n) \in V$**,**

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

誤り訂正性能 (the ability of error-correction)

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

n 箇所のうち何箇所違っても訂正できるか？
(How many errors can one correct ?)

$=$ 距離が幾ら以内なら訂正できるか？
(Within what distance can one correct ?)

uniquely correctable for t errors

$\Longleftrightarrow$ for any $y \in V$,
at most one $x \in \mathcal{C}$ satisfies $d(x, y) \leq t$
$(\#\{x \in \mathcal{C} | d(x, y) \leq t\} \leq 1)$

誤り訂正性能 (the ability of error-correction)

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

n 箇所のうち何箇所違っても訂正できるか？
(How many errors can one correct ?)

$=$ 距離が幾ら以内なら訂正できるか？
(Within what distance can one correct ?)

uniquely correctable for t errors

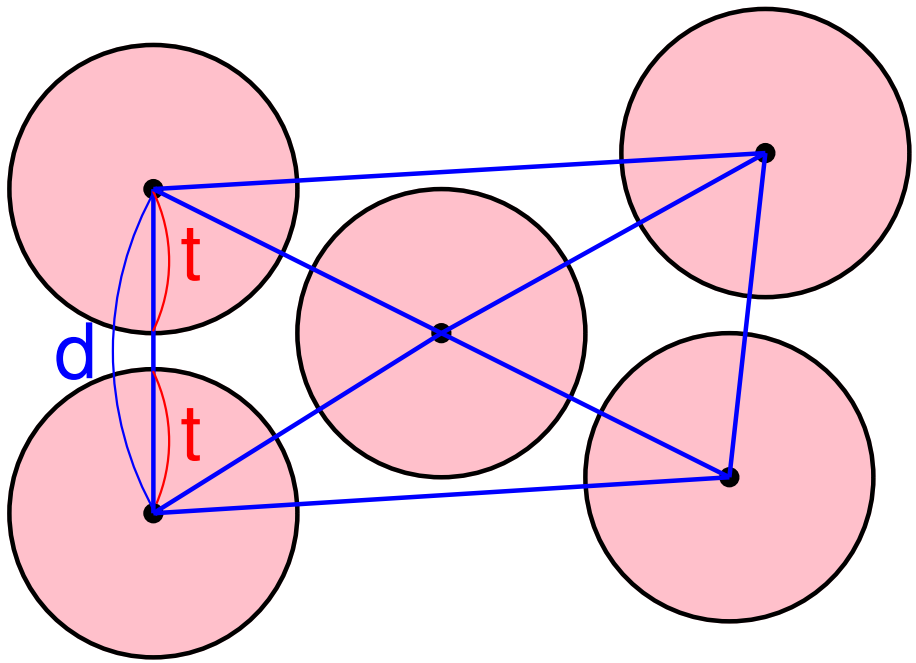
$\Longleftrightarrow$ for any $y \in V$,
at most one $x \in \mathcal{C}$ satisfies $d(x, y) \leq t$
$(\#\{x \in \mathcal{C} \,|\, d(x, y) \leq t\} \leq 1)$

誤り訂正性能 (the ability of error-correction)

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

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

t-error-correctable if $d \geq 2t + 1 \longrightarrow t = \left\lfloor \frac{d-1}{2} \right\rfloor$

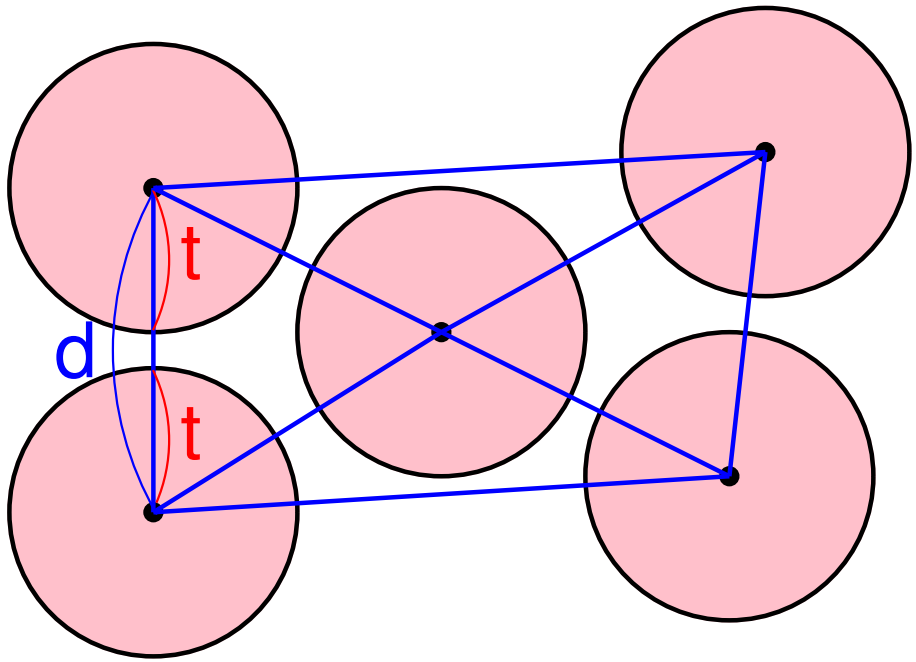


誤り訂正性能 (the ability of error-correction)

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

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

t-error-correctable if $\mathrm{d} \geq 2\mathrm{t}+1 \longrightarrow \mathrm{t} = \left\lfloor \dfrac{\mathrm{d}-1}{2} \right\rfloor$



誤り訂正性能 (the ability of error-correction)

- the code-word length (符号長) n
  (temporarily fix)
- the number of code-words (符号語数)
  $\mathrm{M} := \#\mathcal{C}$ (larger, better)
- the number of correctable errors t
  (larger, better)

$\longrightarrow$ Error-correctability t can be seen
from the minimum distance d !!

誤り訂正性能 (the ability of error-correction)

- the code-word length (符号長) n
  (temporarily fix)
- the number of code-words (符号語数)
  $M := \#\mathcal{C}$ (larger, better)
- the **minimum distance** (**最小距離**) d
  (larger, better)

$\longrightarrow$ (n, M, d)-**code**

誤り訂正性能 (the ability of error-correction)

- the code-word length (符号長) n
  (temporarily fix)
- the number of code-words (符号語数)
  $\mathrm{M} := \#\mathcal{C}$ (larger, better)
- the **minimum distance** (**最小距離**) d
  (larger, better)

"larger M" and "larger d" are conflicting !!

What is the supremum of M for fixed n, d ?