2010 年度春期

応用数学 I (数学科)

情報数学特論 (理工学専攻情報学領域)

(Information Mathematics)

(担当: 角皆)

本講義の概要 (Course Description)

- **情報通信の数理 (Math. in Communication)**
  - ⋆ 情報理論 (Information Theory)
  - ⋆ 符号理論 (Coding Theory)
  - ⋆ 暗号理論 (Cryptography)
- **土台の数学 (Mathematical background)**
  - ⋆ 有限体とその上の線型代数・Galois 理論 (Finite fields, Linear algebra, Galois theory)
  - ⋆ 計算量の理論 (Theory of Complexity)

情報通信を行なう際の要請
(What is required in communication ?)

- **効率的に (efficiently)**
  $\longrightarrow$ 情報理論 (Information Theory)

- **確実に (certainly)**
  $\longrightarrow$ 符号理論 (Coding Theory)

- **安全に (safely)**
  $\longrightarrow$ 暗号理論 (Cryptography)

情報理論 (情報源符号化・情報量の理論)

(Information Theory, Source coding)

- 伝えるべき情報をより効率良く伝えるには
  (How to communicate more efficiently)
- 「効率の良さ」を計る
  (quantify (formulate) "efficiency")
  - ⋆ 伝えるべき「情報の量」を計る
    (quantify "information")
  - ⋆ 伝える為の「手間」を計る
    (quantify "cost to transmit")

$\longrightarrow$ Shannon
「情報の量は伝えるのに必要な手間と一致」
("quantity of information = minimum cost")

符号理論 (誤り訂正符号)
(Coding theory, Error-correcting codes)

- 通信路での雑音による誤りを
  検出・訂正するための符号方式
  (Coding system to detect/correct errors
  caused by noise in transmission)
  - ⋆ 「冗長性」を持たせる (Add "redundancy")
  - ⋆ しかしなるべく効率良く (Keep efficiency)

$\longrightarrow$ 効率の良い符号の構成のために
様々な代数的性質を利用
(線型符号・代数幾何符号など)
(Use various algebraic properties)

暗号理論 (共通鍵・公開鍵暗号)
(Cryptography (Secret Key / Public Key))

- **安全な情報生活の為に**
  (For secure communication)
  - ⋆ **秘密通信 (secret communication)**
  - ⋆ **ディジタル認証・署名 (digital signature)**
  - ⋆ **秘密分散 (secret sharing)**
  - ⋆ **鍵共有 (key exchange)**
- **安全な暗号の実現 (RSA 暗号・楕円曲線暗号等)**
  (Construction of secure cryptosystem)
- **「安全さ」を計る (計算量の理論)**
  (quantify "security" (Theory of Complexity))

基礎となる数理の予備知識

代表的には、例えば次のような事柄

<table>
<tr><th></th><th>基礎編</th><th>初級編</th></tr>
<tr><td>情報理論</td><td colspan="2">微分積分・線型代数・確率論</td></tr>
<tr><td>符号理論</td><td>有限体上の
線型代数</td><td rowspan="3">整数論・群論・
代数幾何の初歩</td></tr>
<tr><td>暗号理論</td><td>初等整数論
(素数の話)</td></tr>
</table>

他に、計算の理論 (計算可能性・計算量) など

情報通信を行なう際の要請
(What is required in communication ?)

- **効率的に (efficiently)**
  $\longrightarrow$ 情報理論 (Information Theory)
- **確実に (certainly)**
  $\longrightarrow$ 符号理論 (Coding Theory)
- **安全に (safely)**
  $\longrightarrow$ 暗号理論 (Cryptography)

情報理論 (情報源符号化・情報量の理論)

(Information Theory, Source coding)

- 伝えるべき情報をより効率良く伝えるには
  (How to communicate more efficiently)
- 「効率の良さ」を計る
  (quantify (formulate) "efficiency")
  - ⋆ 伝えるべき「情報の量」を計る
    (quantify "information")
  - ⋆ 伝える為の「手間」を計る
    (quantify "cost to transmit")

$\longrightarrow$ Shannon
「情報の量は伝えるのに必要な手間と一致」
("quantity of information = minimum cost")

情報の符号化 (Coding)

Analog data (continuous data)

⇓ sampling

Digital data (discrete, finite)
↑ “情報 (information)” to be treated here
(情報源, source)

⇓ ⟵ “符号化 (coding)” to be treated here

Digital data for transmission (伝送用データ)
· · · 特定 (一般には少数) の種類の文字の列
(a sequence of specific alphabets)

文字 (characters) $\longrightarrow$ ASCII code

	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	(control characters)															
1	(control characters)															
2		!	"	#	$	%	&	’	(	)	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
6	‘	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{	|	}	~	

7 bits for each character

(8 bits in actual implementation)

ASCII code

A	01000001
B	01000010
C	01000011
D	01000100
E	01000101
F	01000110
G	01000111
H	01001000
I	01001001

J	01001010
K	01001011
L	01001100
M	01001101
N	01001110
O	01001111
P	01010000
Q	01010001
R	01010010

S	01010011
T	01010100
U	01010101
V	01010110
W	01010111
X	01011000
Y	01011001
Z	01011010

モールス符号 (Morse code)

A	・ ー
B	ー ・・・
C	ー ・ ー ・
D	ー ・・
E	・
F	・・ ー ・
G	ー ー ・
H	・・・・
I	・・

J	・ ー ー ー
K	ー ・ ー
L	・ ー ・・
M	ー ー
N	ー ・
O	ー ー ー
P	・ ー ー ・
Q	ー ー ・ ー
R	・ ー ・

S	・・・
T	ー
U	・・ ー
V	・・・ ー
W	・ ー ー
X	ー ・・ ー
Y	ー ・ ー ー
Z	ー ー ・・

モールス符号 (Morse code)

1 文字のための符号長が区々
(various length for each character)

$\longleftarrow$ 頻度の高い文字は短く、低い文字は長く
(Shorter if frequent, longer if rare)

$\longrightarrow$ 頻度まで考慮して符号長の期待値を短く
(Shorten the expectation length)

$\cdots$ 頻度 (出現確率) も考慮した符号効率の定式化
(Formulate efficiency considering frequency)

モールス符号 (Morse code)

1 文字のための符号長が区々
(various length for each character)

$\longrightarrow$ 1 文字毎の区切りは判るのか ?
(Can one detect an end of each character ?)

$\longrightarrow$ 実はモールス符号では、
文字間・単語間の送信間隔が定められている
(The interval between characters are fixed.)

A	・ー
B	ー ・・・
C	ー ・ ー ・
D	ー ・・
E	・
F	・・ ー ・
G	ー ー ・

H	・・・・
I	・・
J	・ ー ー ー
K	ー ・ ー
L	・ ー ・・
M	ー ー
N	ー ・

O	ー ー ー
P	・ ー ー ・
Q	ー ー ・ ー
R	・ ー ・
S	・・・
T	ー

U	・・ ー
V	・・・ ー
W	・ ー ー
X	ー ・・ ー
Y	ー ・ ー ー
Z	ー ー ・・

Ex. the cat and the dog

ー○ ・・・・ ○ ・ ○○○ー ・ ー ・ ○ ・ ー○ー○○○ ・ ー
○ー ・○ー ・・○○○ー○ ・・・・○ ・○○○ー ・・○ー ー ー
○ー ー ・

(○ is a space (interval) with length as one ・)

⟶ ーと・と○とを用いて符号化している

(Coding with ー, ・, and ○)

情報源符号化の定式化
(Formulation of source coding)

source alphabet S **: a finite set**

$S^+ := \bigsqcup_{n \geq 1} S^n$ **:** S **の元の** 1 **個以上の列全体**

ε **: 空語 (the empty word),** $S^0 := \{\varepsilon\}$

$S^* := \bigsqcup_{n \geq 0} S^n$ **:** S **の元の** 0 **個以上の列全体**

$= S^+ \sqcup \{\varepsilon\}$

$w \in S^n$ **に対し、** $|w| := n$

(the length of a sequence)