

情報通信システム概論

第5回
～情報量基礎・デシベル～

2020-10-17

情報システム工学科
福田 浩



ゴールイメージ

- ・ 情報量・情報エントロピーの考え方を説明できる
- ・ 符号化について、具体例を用いて説明できる
- ・ デシベル計算ができる

- ・ 情報量定義の動機
- ・ 情報量の定量化
- ・ 情報量と情報エントロピー
- ・ データ量を減らす手順 標本化，量子化，符号化
- ・ 一意復号可能性と瞬時復号，ハフマン符号
- ・ デシベル表記
- ・ 第3回演習問題の解説

価値ある情報を高速に正確に送る
“情報理論の父”，
クロード・エルウッド・シャノン
(Claude Elwood Shannon, 1916~2001)



C. E. Shannon

Reprinted with corrections from *The Bell System Technical Journal*,
Vol. 27, pp. 379–423, 623–656, July, October, 1948.

通信の数学的理論

A Mathematical Theory of Communication

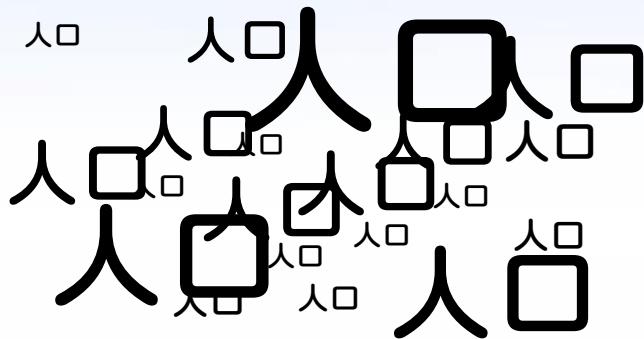
By C. E. SHANNON

INTRODUCTION

THE recent development of various methods of modulation such as PCM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is contained in the important papers of Nyquist¹ and Hartley² on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.



情報量は多いか，少ないか...



A word cloud forming the pieces of a chessboard. The words are arranged in a grid pattern, with some words being larger than others. The words include King, Queen, Bishop, Knight, Rook, and Pawn, all in various colors (blue, green, black).



情報量は多いか、少ないか...

人口

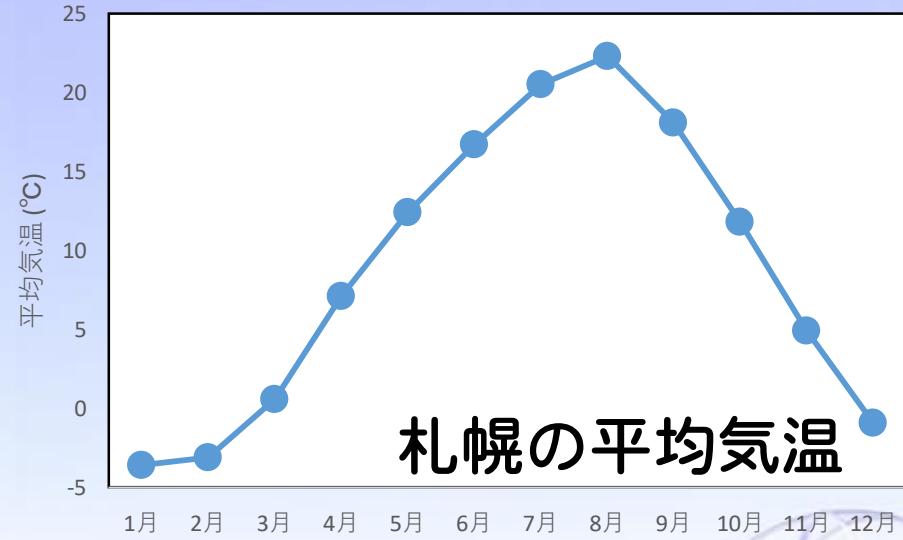
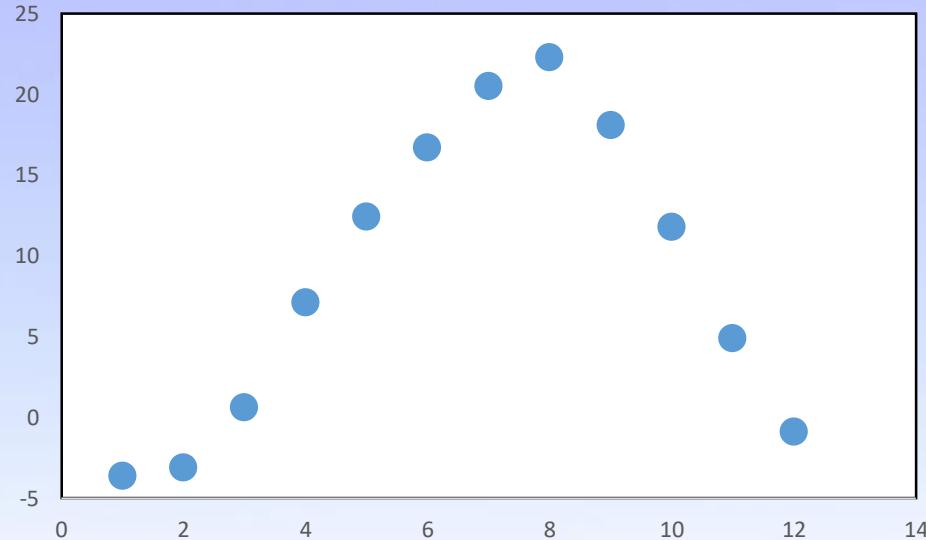


情報量は多いか，少ないか...



札幌市	1,970,057
仙台市	1,090,263
さいたま市	1,307,931
千葉市	980,203
横浜市	3,748,781
川崎市	1,530,457
相模原市	722,828
新潟市	796,500
静岡市	691,185
浜松市	791,770
名古屋市	2,327,557
京都市	1,466,264
大阪市	2,740,202
堺市	827,971
神戸市	1,522,944
岡山市	720,865
広島市	1,199,359
北九州市	940,141
福岡市	1,592,657
熊本市	739,393

情報量は多いか、少ないか...



- ・ 情報通信システム概論の中間試験はレポートに替える
 - 「レポートに替える」はフェイクニュースらしい
 - ・ 「〇〇らしい」では、どう判断すればよいかわからず、**貴重でない情報**
 - 「レポートに替える」ことは第1回の授業で明言されている
 - ・ 証拠があるなら確実な情報と言えるので、**貴重な情報**

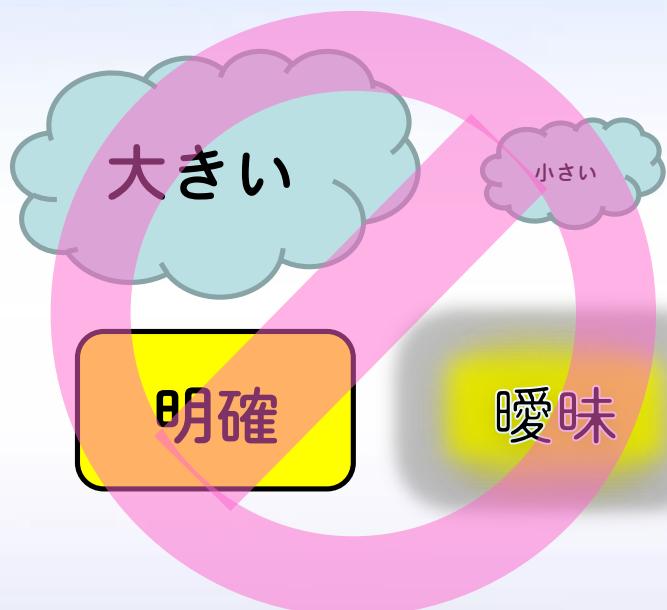
- ・ 情報通信システム概論は最終試験を行う
 - 情報通信システム概論の最終試験は、今後行われる
 - ・ 最終試験というくらいなので、今行われることが無いことは当然であり、**曖昧な情報**
 - 情報通信システム概論の最終試験は最終日に行われる
 - ・ 日時がわかるので、**明確な情報**

情報の定量化

- ・ シャノンが拘ったこと
 - 情報を高速に送る
 - 情報を正確に送る

そのためにも、

- 情報を定量化する



情報量 直感的な解釈

例1) 情報通信システム概論の最終試験は、今後行われる
「当たり前のこと」は情報量が無い(少ない)

例2-1) 情報通信システム概論の最終試験は、1月に行われる

例2-2) 情報通信システム概論の最終試験は、1月20日に行われる

例2-3) 情報通信システム概論の最終試験は、1月20日午前10:45に行われる
「限定的なこと」は情報量が多い

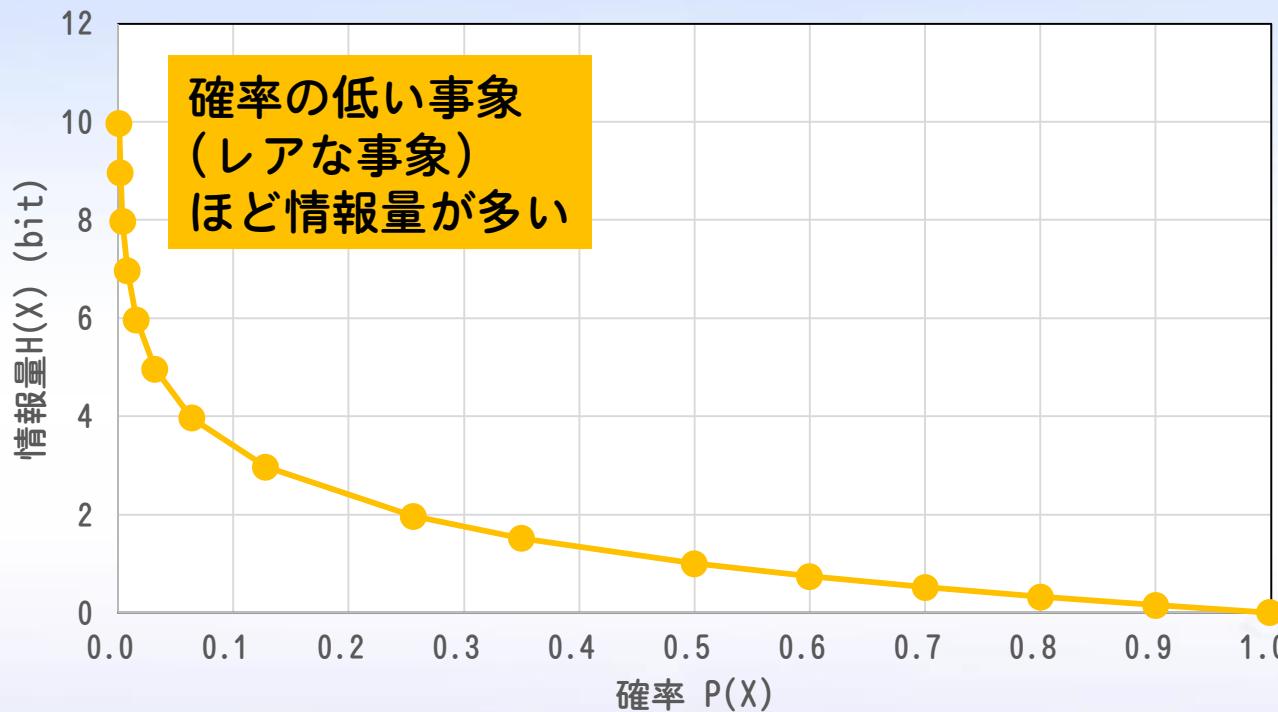
情報量はそれが関連する事象が起こる確率(の逆数)に関係しそう

情報量 定量的な定義

(自己)情報量

ある事象 X が起こる確率が $P(X)$ のときの情報量 $H(X)$

$$H(X) = \log_2 \frac{1}{P(X)} = -\log_2 P(X)$$





情報エントロピー 直感的な解釈

当たり前のこととは情報量は少ないが、起こる確率は高い
限定的なことは情報量は多いが、起こる確率は低い

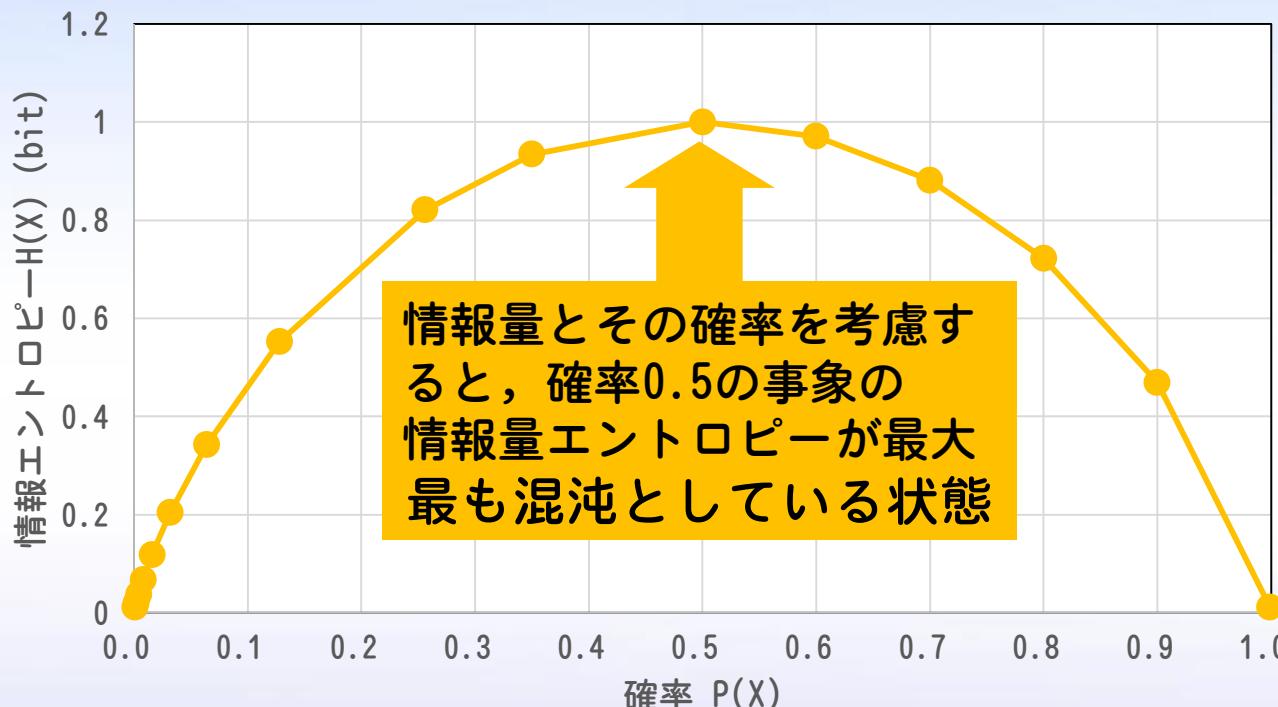
情報量に起こる確率という重みを掛けて、総和をとることで
その時の状態を表現できそう

情報エントロピー 定量的な定義

情報源 X の情報エントロピー $H(X)$

$$H(X) = - \sum_{i=1}^M p_i \log_2 p_i$$

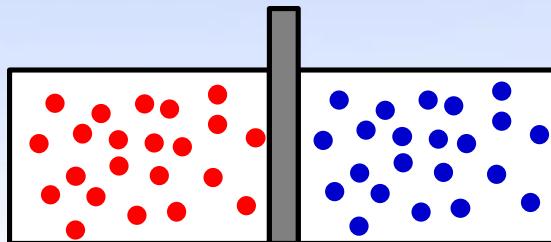
$$p_1 + p_2 + p_3 + \cdots + p_M = 1$$



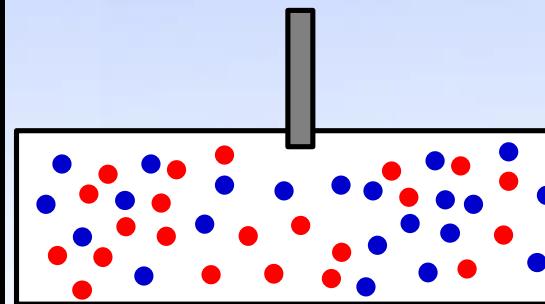
エントロピー増大則

断熱系において不可逆変化が生じた場合、その系のエントロピーは増大する

エントロピー小



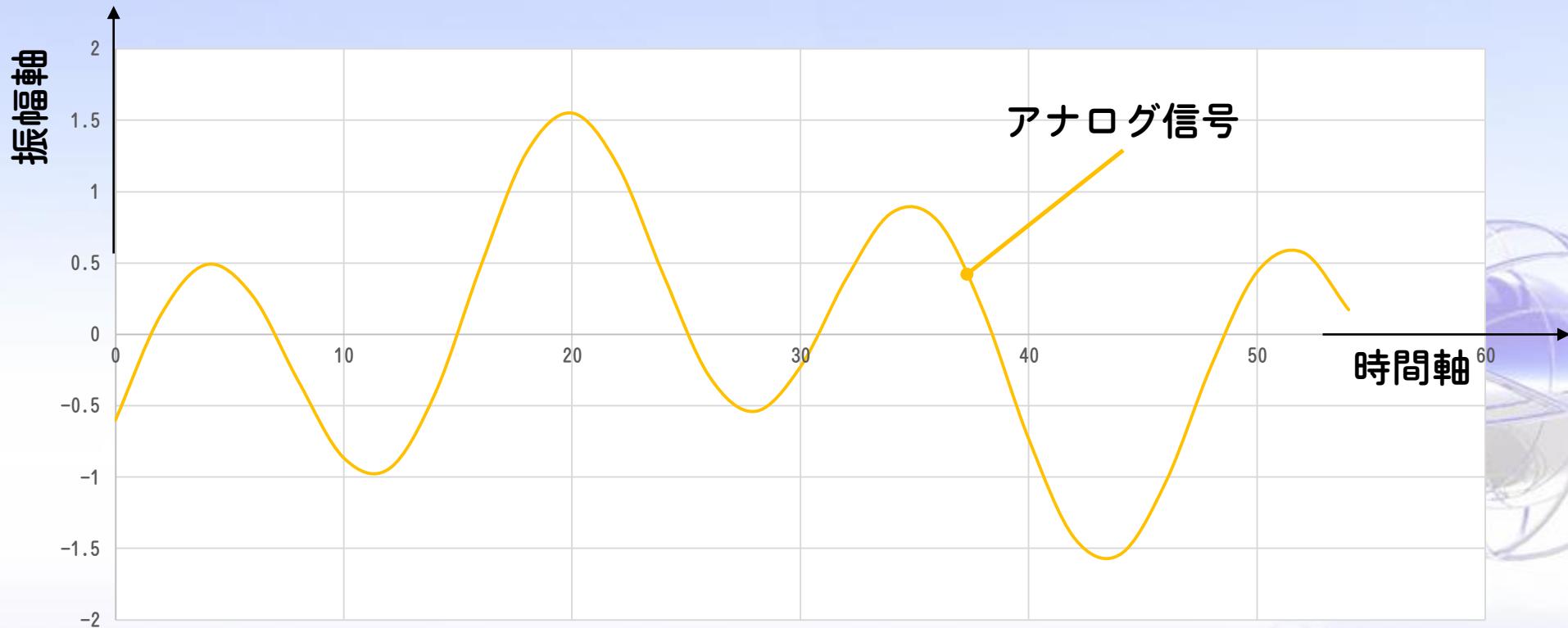
エントロピー大



Designed by macrovector / Freepik

データ量を減らす

高速に正確に情報を送る為に、情報量を損なわずにデータ量を削減
アナログからデジタルへの変換 標本化 -> 量子化 -> 符号化



標本化 (sampling)

アナログ信号を一定時間間隔で測定し、離散信号として収集すること

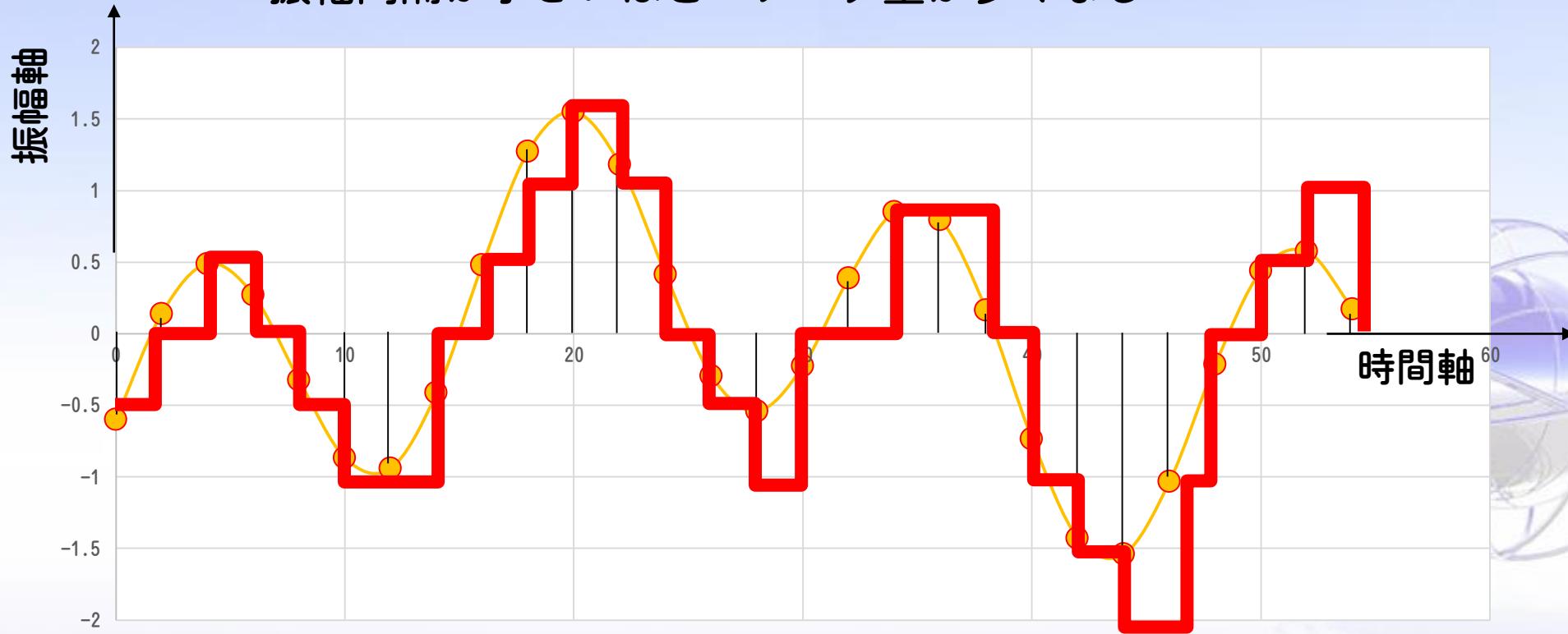
- 時間間隔が短いほど アナログ信号を忠実に再現できる
- × 時間間隔が短いほど データ量が多くなる



量子化 (quantization)

振幅の大きさを離散的な値で近似的に表すこと

- 振幅間隔が小さいほど アナログ信号を忠実に再現できる
- × 振幅間隔が大きいほど データ量が多くなる



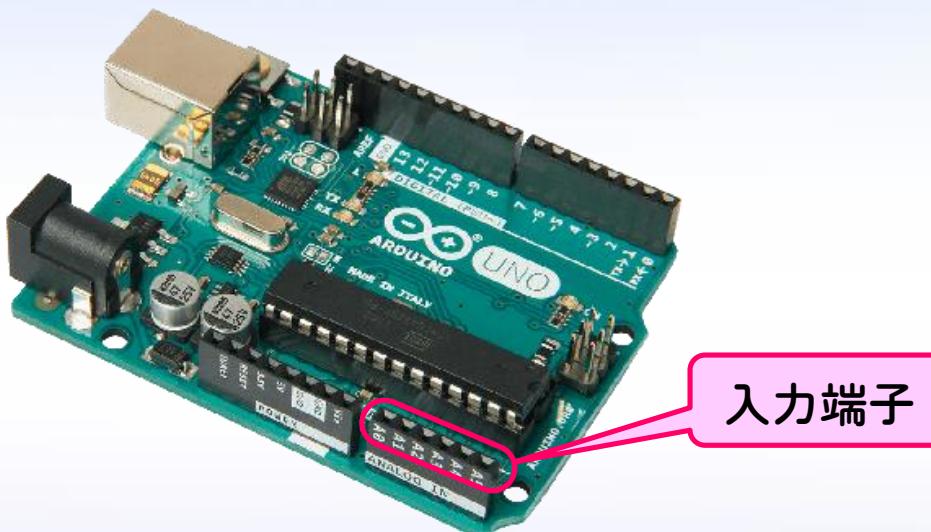
符号化 (encoding)

デジタル信号に特定の方法で、後に元の(あるいは近似)信号に戻せるような変換を加えること

例：マイコンArduinoUnoの入力端子での符号化

0Vを00000000とし、5Vを11111111とする

0~5Vの間は256分割して(量子化)，8bitの2進数で表す



※3年次のプロジェクト科目で使用します

符号化の例

モールス符号

モールス符号

- 標準的な英文におけるアルファベットの出現頻度に応じて符号化
- 出現頻度が高い文字は短い符号
- 例えば “E” , “T” は1符号, 出現頻度が少ない “Q” や “J” は4符号

CHITOSE

1001100

000 10

10 10

00

INSTITUTE

010101010

000 0 0

0 1

OF

10

10

11

0

SCIENCE

0100110

000 00

010 1

0 0

AND

011

100

0

TECHNOLOGY

1010110111

00011110

10 10101

00 0 1

A	0	1			
B	1	0	0	0	
C	1	0	1	0	
D	1	0	0		
E	0				
F	0	0	1	0	
G	1	1	0		
H	0	0	0	0	
I	0	0			
J	0	1	1	1	
K	1	0	1		
L	0	1	0	0	
M	1	1			
N	1	0			
O	1	1	1		
P	0	1	1	0	
Q	1	1	0	1	
R	0	1	0		
S	0	0	0		
T	1				
U	0	0	1		
V	0	0	0	1	
W	0	1	1		
X	1	0	0	1	
Y	1	0	1	1	
Z	1	1	0	0	

一意復号可能

- 必ず一意に復号できること
- 符号語の長さが等しい符号は、一意復号可能
- 符号語の長さが異なっても、一意復号可能な符号を生成可能

一意復号不可	一意復号可能
A	1
B	01
C	011
D	111

BA(011)とC(011)が同じ
符号になってしまい、
一意に複合できない

瞬時符号

- ・ 符号語の最後のbitを読んだ時点で(遅れることなく)復号できる符号
- ・ 瞬時符号は必ず一意復号可能

非瞬時符号	瞬時符号
A	0
B	01
C	011
D	111

モールス符号はそのままでは非瞬時符号

文字間に3つぶんの空白を挿入することで対応

“遅れることなく”復号できるわけではないので非瞬時符号

効率的であることから広く使われている瞬時符号
1952年にデビッド・ハフマンによって開発

作成手順(アルゴリズム)

1. 複数個のメッセージを確率の大きい順に並べる
2. 確率の小さいメッセージ2個をまとめ，これを1つの複合メッセージとして2つの確率の和を求める
3. 2.でまとめたものを1個のメッセージとみなし，再度確率の大きいもの順に並べなおす
4. 確率1の複合メッセージになるまで2. 3.を繰り返す
5. Tree構造を形成し，Treeを遡りながら符号化する

ハフマン符号化の例

CHITOSE

1011011

1110010

1010001

001 1

INSTITUTE

101111011

101010100

110010001

1 1 1 1

0

OF

00

01

00

1

1

SCIENCE

1111011

1110010

0111111

101 0

AND

000

101

111

0 0

0 1

TECHNOLOGY

1110000001

0011001010

0110101011

00 1 11

0 11

ハフマン符号化の例

CHITOSE	INSTITUTE	OF
4443343	434343533	35
SCIENCE	AND	TECHNOLOGY
4443343	535	3344335355

符号長

ハフマン符号による符号長：141 bit

アルファベットを 5 bit であらわした場合：190 bit

75%以下に圧縮可能

比較的高効率であり
JPEG等の画像圧縮
符号化に使用

通信では比率の大きな減少を扱うことが多く、底が10の対数で表現することが多い

$$\frac{P_o}{P_{in}} = \frac{|E_o|^2}{|E_{in}|^2} = 10^{-\rho}$$

無次元
(単位無し)

$$\rho = -\log_{10}\left(\frac{P_o}{P_{in}}\right) \quad (\text{B}): \text{ベル}$$

$$y = 10\rho = -10\log_{10}\left(\frac{P_o}{P_{in}}\right) \quad (\text{dB}): \text{デシベル}$$

デシベル(dB)	電力比	電圧比
30	1000.00	31.62
20	100.00	10.00
10	10.00	3.16
9	7.94	2.82
8	6.31	2.51
7	5.01	2.24
6	3.98	2.00
5	3.16	1.78
4	2.51	1.58
3	2.00	1.41
2	1.58	1.26
1	1.26	1.12
0	1.00	1.00
-3	0.50	0.71
-10	0.10	0.32
-20	0.01	0.10

よく使うデシベル

頻繁に使うdBは記憶し， それらの組合せで概算できるようにしておくこと

dB	電力比	電圧比
30	1000.00	31.62
20	100.00	10.00
10	10.00	3.16
7	5.01	2.24
6	3.98	2.00
5	3.16	1.78
3	2.00	1.41

dB	電力比	電圧比
-3	0.50	0.71
-5	0.32	0.56
-6	0.25	0.50
-7	0.20	0.45
-10	0.10	0.32
-20	0.01	0.10
-30	0.001	0.03

【概算例 1】

電力比-23dBは， -20dBと-3dBの組合せなので， $0.01 \times 0.5 = 0.005$ ($=1/200$)

【概算例 2】

電圧比+26dBは， +20dBと+6dBの組合せなので， $10 \times 2 = 20$

デシベルに関する注意点

- ・ 読み方
 - 正式には「デシベル」，ディービーと読む場合も多い
 - 「デシ」と略す場合もあるが，この読み方は推奨しない
- ・ 絶対量としてのデシベル
 - 基準となる物理量を予め決めてあれば，物理量をデシベル表現可能
 - 通信分野での例
 - ・ dBm: 1 mWに対するパワー比，0 dBmは 1 mW，20 dBmは 100 mW

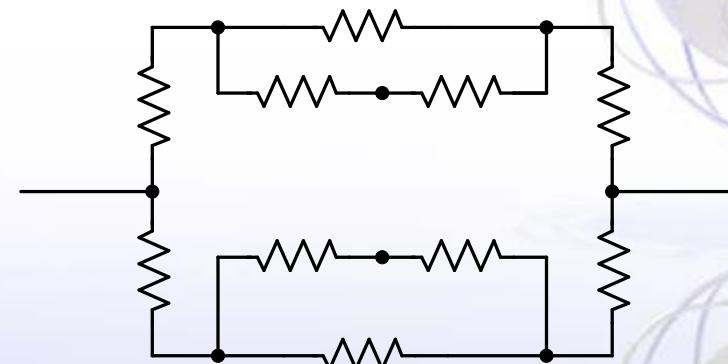
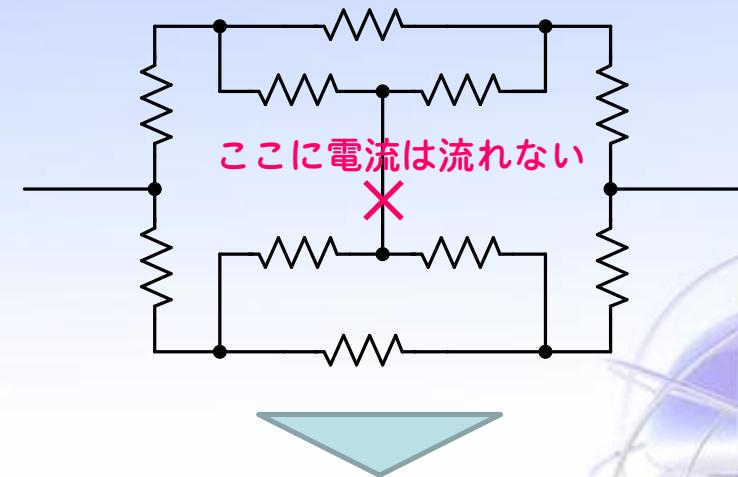
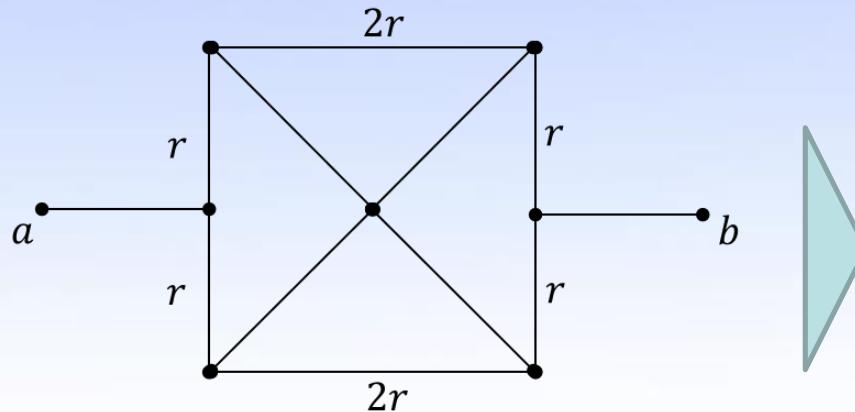


$$\begin{aligned}-50 \text{ (dBm)} &= 0.00001 \text{ (mW)} \\ &= 10 \text{ (nW)}\end{aligned}$$

- ・ 動機：価値ある情報を高速に正確に送る
 - そのためにも、情報を定量化する
- ・ (自己)情報量 $H(X) = -\log_2 P(X)$
- ・ 情報エントロピー $H(X) = -\sum_{i=1}^M p_i \log_2 p_i$
- ・ アナログ/デジタル変換 標本化 \rightarrow 量子化 \rightarrow 符号化
 1. アナログ信号を一定時間間隔で測定し、離散信号として収集する
 2. 振幅の大きさを離散的な値で近似的に表す
 3. デジタル信号に特定の方法で、後に元の(あるいは近似)信号に戻せるような変換を加える(例:ハフマン符号化)
- ・ デシベル：比率を底が10の対数で表現
 - 基準となる物理量を予め決めてあれば、物理量をデシベル表現可能

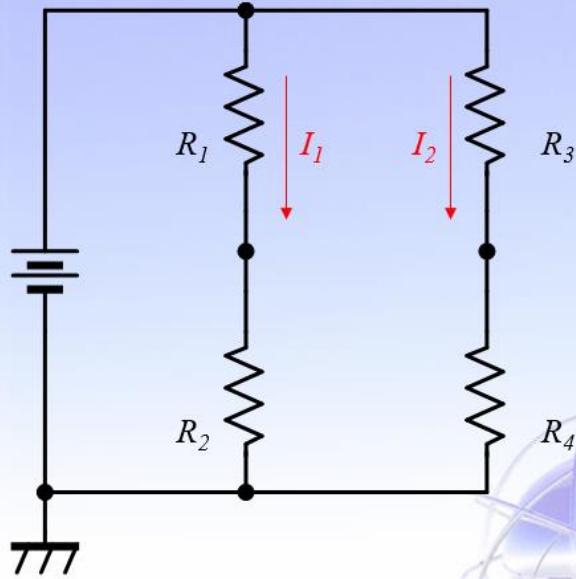
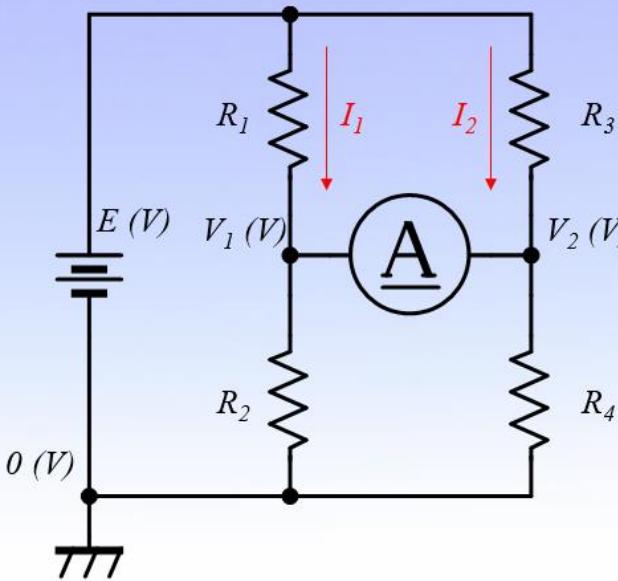
第3回演習問題解説

- 図に示すように、1辺が $2r$ (Ω)の電熱線で作った正方形の対角線をそれぞれ同じ種類の電熱線で結んだ回路のa-b間の合成抵抗を求めよ。計算結果だけでなく、その過程も示すこと。【令和元年度 第2回 電気通信主任者試験問題一部改】



第3回演習問題解説

復習：ブリッジ回路



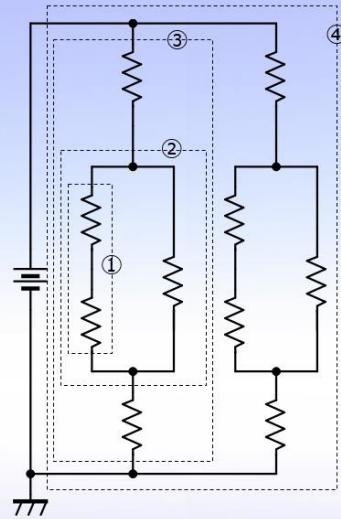
$$V_1 = E - I_1 R_1, V_2 = E - I_2 R_3$$

$I_1 R_1 = I_2 R_3$ のとき， $V_1 = V_2$ となり
電流計には電流は流れない

第3回講義資料より

第3回演習問題解説

復習：合成抵抗



©2020 h-fukuda copyrights all rights reserved.

$$\textcircled{1} \quad \sqrt{2}r + \sqrt{2}r = 2\sqrt{2}r$$

$$\textcircled{2} \quad \frac{2\sqrt{2} \times 2 r^2}{(2\sqrt{2}+2)r} = \frac{2\sqrt{2}}{\sqrt{2}+1} r = \frac{2\sqrt{2}(\sqrt{2}-1)}{(\sqrt{2}+1)(\sqrt{2}-1)} r = (4 - 2\sqrt{2})r$$

$$\textcircled{3} \quad (4 - 2\sqrt{2})r + r + r = (6 - 2\sqrt{2})r$$

$$\textcircled{4} \quad \frac{(6-2\sqrt{2})r \times (6-2\sqrt{2})r}{(6-2\sqrt{2})r + (6-2\sqrt{2})r} = (3-\sqrt{2})r \quad \boxed{\text{解答}}$$

第3回講義資料より

第3回演習問題解説

