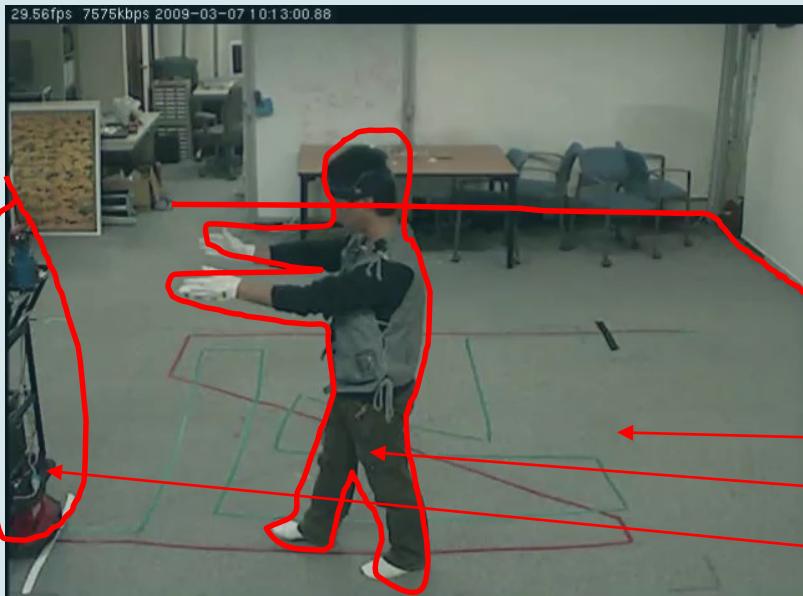


# 情報符号理論の枠組み

# 情報符号理論の枠組み



## 物理理論

質点, 力, 位置, 速度, 加速度, 光, 電気, 電子, ...

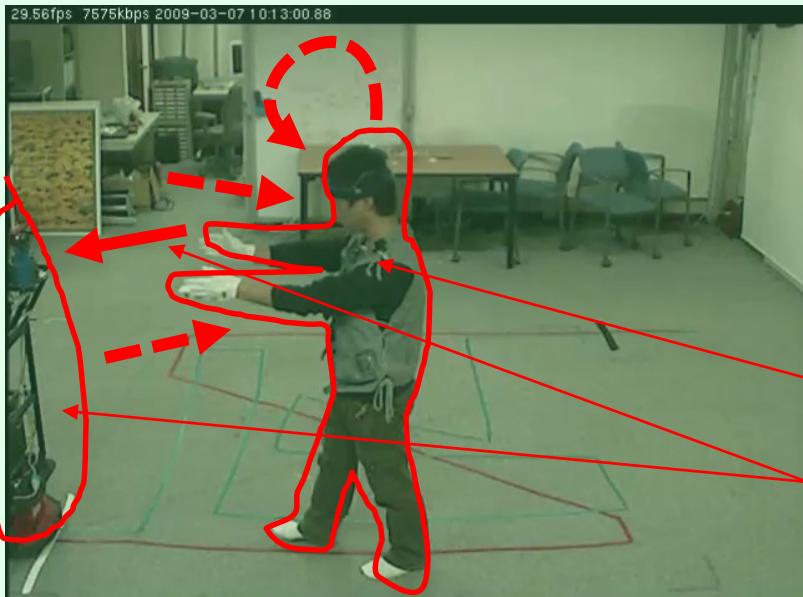
## 物理的説明

- 平面
- 物体1
- 物体2

物理学者



# 情報符号理論の枠組み



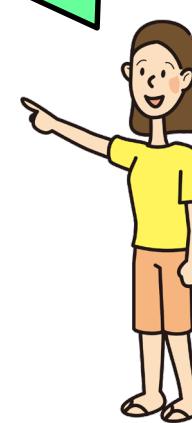
情報の理論

情報源, 情報, ...

情報科学的説明

- 情報源
- 受信者
- 情報

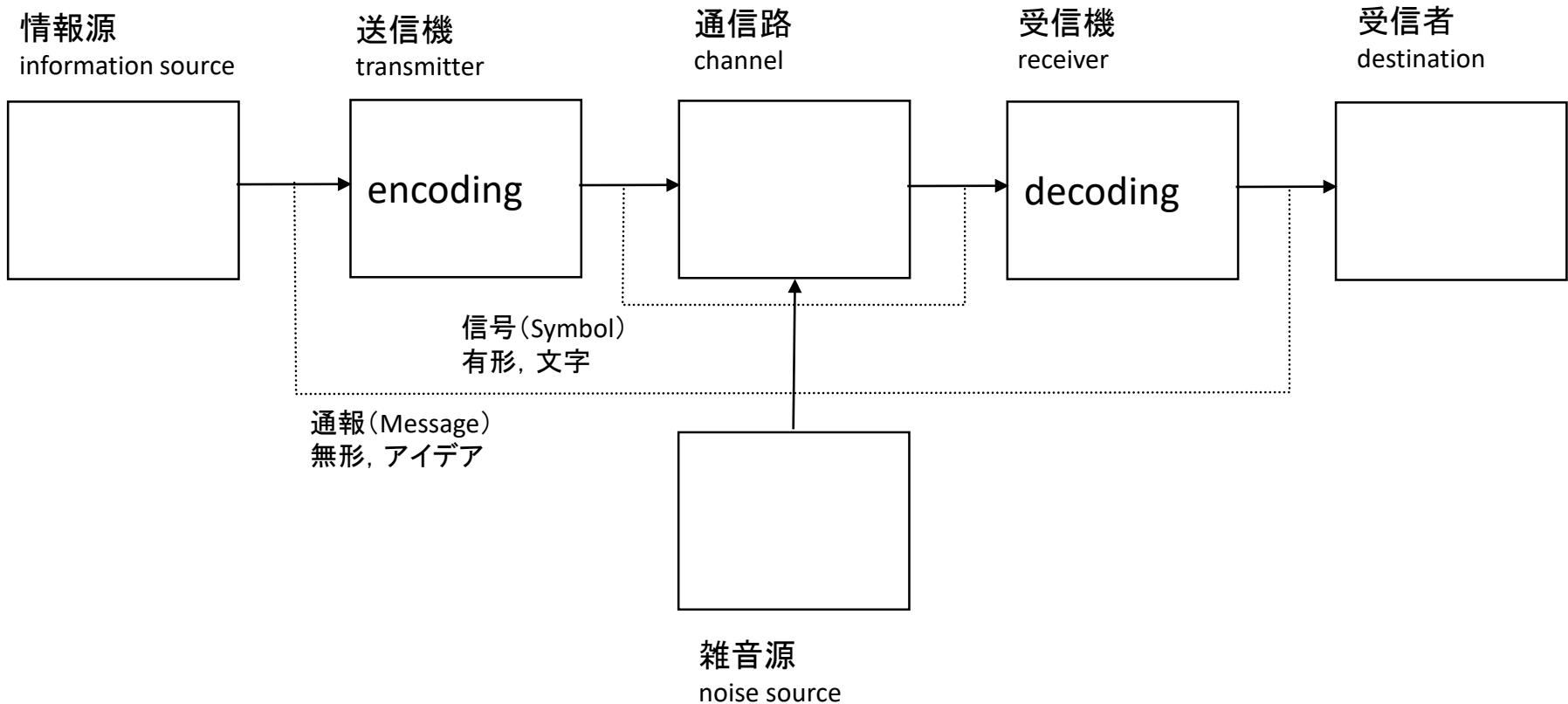
情報科学者



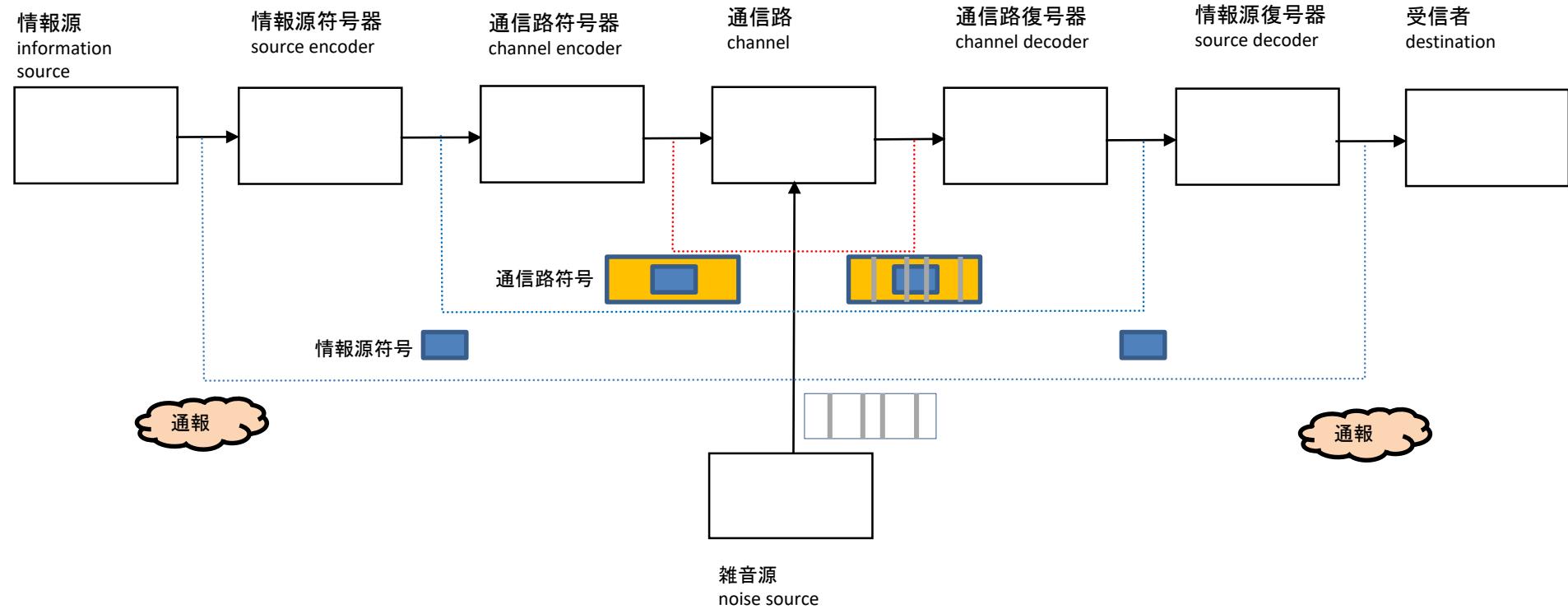
# 情報とは何か？

- 情報=不確かさを減らすもの
- 生命: 環境のなかの様々な情報に反応して複雑な行動をする。
- 機械: 環境内の情報を検出して組み込まれた機序にしたがって行動する。

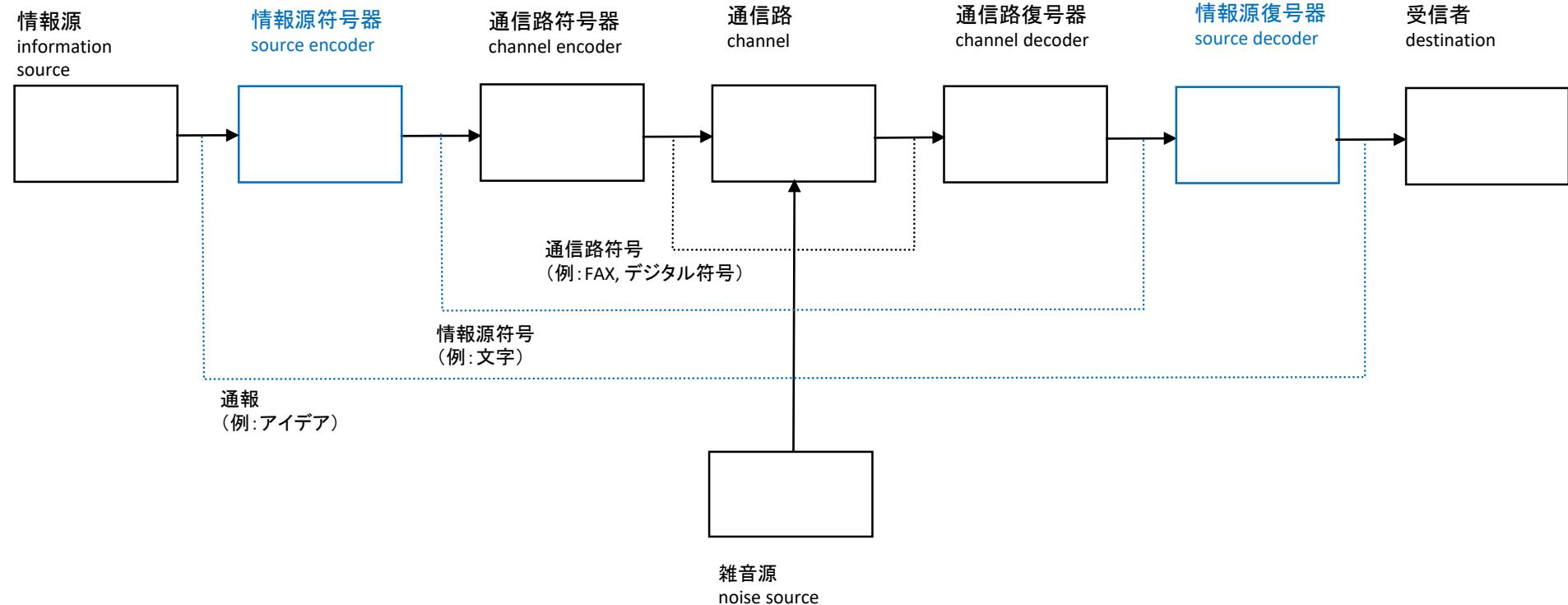
# Shannonモデル



# Fanoによる修正Shannonモデル



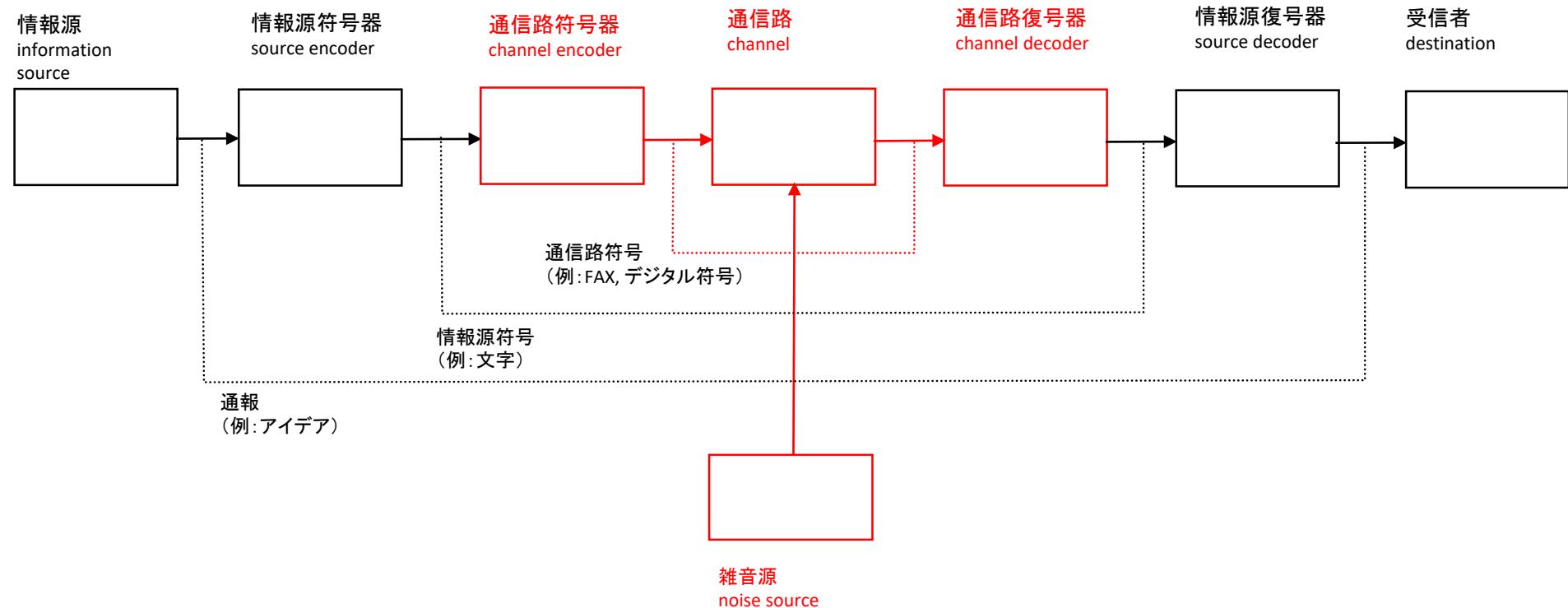
# 本科目の構成



## ①(離散系)情報源符号化とその限界

瞬時復号可能性を保ったまま、平均符号長をどれだけ短くできるか？

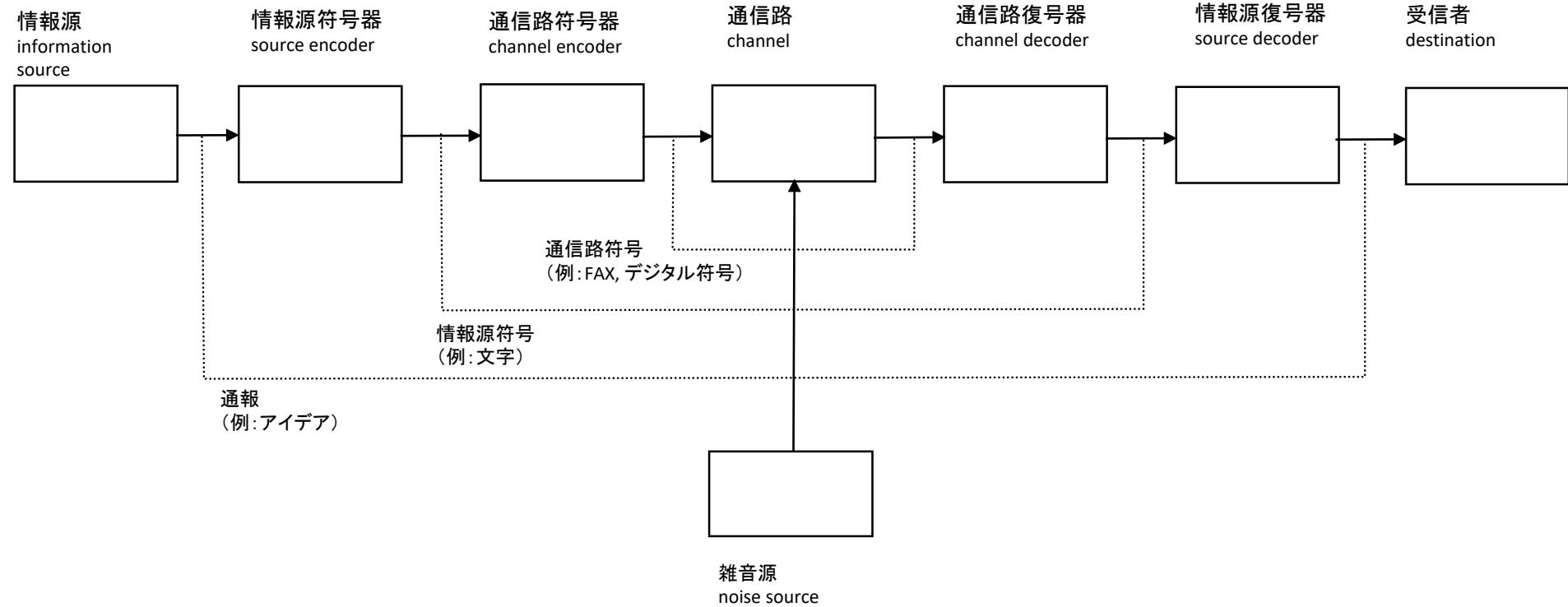
# 本科目の構成



## ②(離散系)通信路符号化とその限界

通信路の品質から定まる通信速度の範囲内なら、復号誤りを任意に小さくできる

# 本科目の構成



- ③離散系の誤り訂正／検出符号構成法
- ④アナログ系

# 情報とは何か？

不確かさを減らす度合いが大きければ、より価値のある情報



# 情報量

2つの事象 $\{e_1, e_2\}$ のいずれかが起きるとする.

$e_1$ の生起確率を $p_1$ ,  $e_2$ の生起確率を $p_2$ とする.

- $e_1$ が起きたことを知らせるメッセージの情報量:

$$\log_2 \frac{1}{p_1} = -\log_2 p_1 \text{ ビット}$$

- $e_2$ が起きたことを知らせるメッセージの情報量:

$$\log_2 \frac{1}{p_2} = -\log_2 p_2 \text{ ビット}$$

- メッセージの平均的な情報量:

$$p_1 \log_2 \frac{1}{p_1} + p_2 \log_2 \frac{1}{p_2} = -(p_1 \log_2 p_1 + p_2 \log_2 p_2) \text{ ビット}$$

# 情報量

出現確率 $p$ の事象 $e$ が起きたことを知らせるメッセージの情報量

$$\log_2 \frac{1}{p} = -\log_2 p \quad \text{ビット}$$

(例) 左へ行った確率が $1/4$ であるとき、実際にそうであるということを伝えるメッセージの情報量：

$$\log_2 \frac{1}{\frac{1}{4}} = -\log_2 4 = 2 \quad \text{ビット}$$

# 情報量

$n$ 個の事象 $\{e_1, \dots, e_n\}$ が与えられたとき,  $e_i$ の生起確率を $p_i$ とすると,  $e_i$ の起きたことを知らせるメッセージの平均的な情報量:

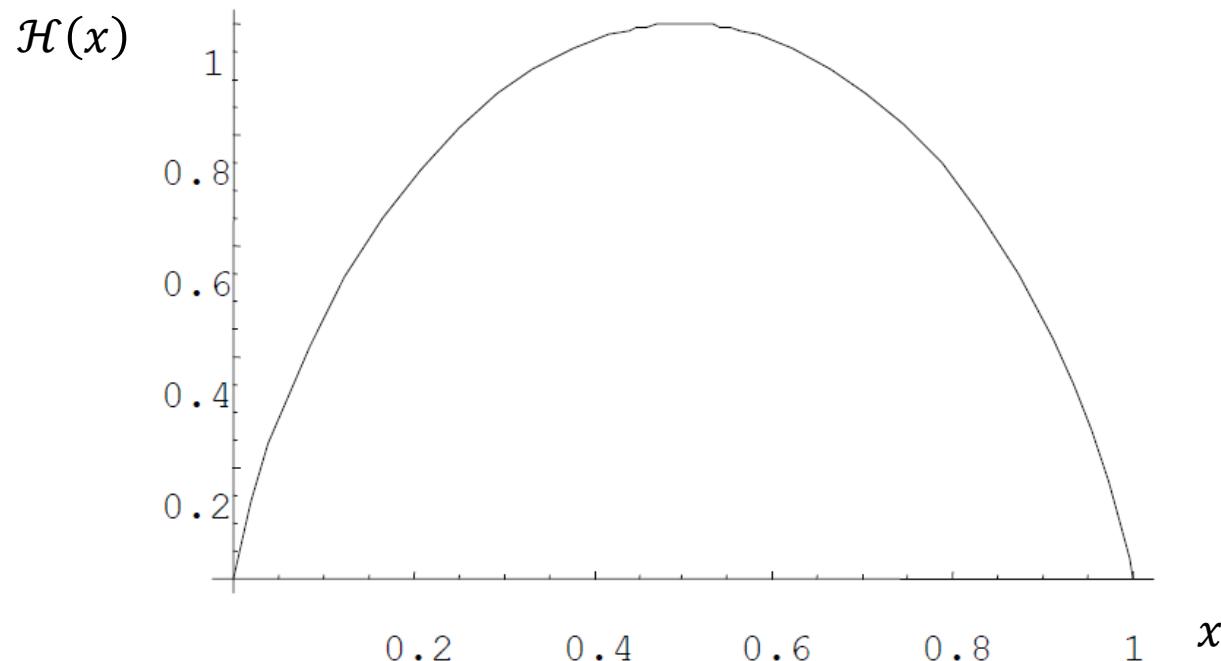
$$p_1 \log_2 \frac{1}{p_1} + \cdots + p_n \log_2 \frac{1}{p_n} = - \sum_{i=1}^n p_i \log_2 p_i \quad \text{ビット}$$

- 例1:{左: 1/2, 右: 1/2} … 左でも, 右でも1ビットの情報が得られる. 平均情報量は1ビット.
- 例2:{左: 1/4, 右: 3/4} } … 1/4の確率で $\log_2 4 = 2$ ビットの情報が得られるが, 残り3/4の確率で得られる情報量は $\log_2 \frac{4}{3} \approx 0.415$ ビットだけである. 平均すれば,

$$\frac{1}{4} \log_2 \frac{1}{\frac{1}{4}} + \frac{3}{4} \log_2 \frac{1}{\frac{3}{4}} \approx \frac{1}{4} \times 2 + \frac{3}{4} \times 0.415 \approx 0.811 \quad \text{ビット}$$

# 情報量

事象が2個のとき、そのうちの1個の生起確率を $x$ とすれば、他方の事象の生起確率は $1 - x$ なので、事象の生起を知らせるメッセージのもたらす平均的な情報量を $\mathcal{H}(x)$ とすれば、 $\mathcal{H}(x) = -(x \log_2 x + (1 - x) \log_2(1 - x))$ となる。 $\mathcal{H}(x)$ はエントロピー関数と呼ばれる。



# 相互情報量の概念

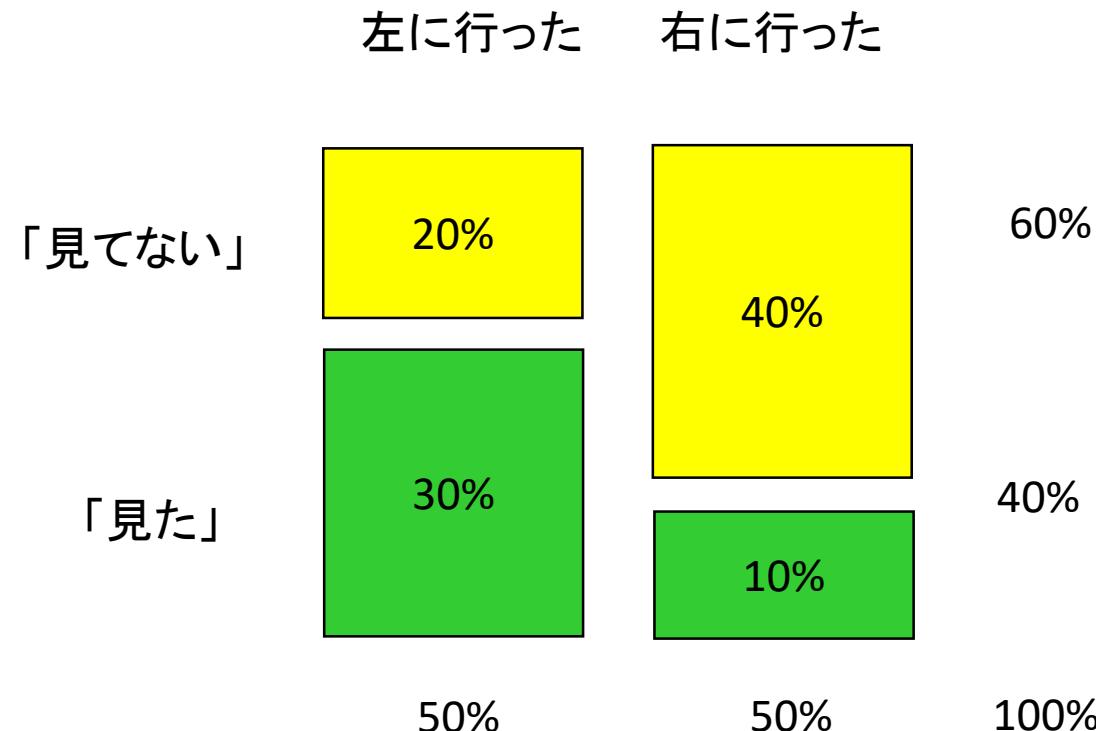


	左	右
見てない	20%	40%
見た	30%	10%

# 相互情報量の概念

## ■ 不完全な答えからも情報は得られる

「見てない／見た」と「左／右に行った」の関わりについての知識



# 相互情報量

- 「左に行った」, 「右に行った」というメッセージのもたらす価値は文脈に依存する.
- 何も知らないときは, 「左に行った」は1ビットの情報量, 「右に行った」も1ビットの情報量をもつから, メッセージは平均1ビットの情報量をもつことになる.
- 目撃有無情報を得たとしよう. 「見てない」という回答だったとすれば,
  - 「左に行った」というメッセージは,  $\log_2 \frac{1}{\frac{1}{3}} = \log_2 3 \approx 1.585$ ビットの情報をもたらす.
  - 「右に行った」というメッセージは,  $\log_2 \frac{1}{\frac{2}{3}} = \log_2 \frac{3}{2} \approx 0.585$ ビットの情報をもたらす.
  - 平均すると,  $[-(p \log_2 p + (1 - p) \log_2(1 - p))]_{p=\frac{1}{3}} \approx 0.918$ ビットの情報がもたらされる.
  - 同様に, 「見た」という回答だったとすれば, メッセージの平均情報量は
$$[-(p \log_2 p + (1 - p) \log_2(1 - p))]_{p=\frac{1}{4}} \approx 0.811 \text{ ビット}$$
- 以上より, 目撃有無情報から得られる平均情報量は
$$0.918 \times \frac{3}{5} + 0.811 \times \frac{2}{5} \approx 0.8755 \text{ ビット}$$

# 相互情報量

確率変数 $X$ のエントロピー(平均的な不確かさ)を $H(X)$ とすると、

$$H(\text{左右}) = \mathcal{H}\left(\frac{1}{2}\right) = 1$$

$$H(\text{左右}|\text{見てない}) = \mathcal{H}\left(\frac{1}{3}\right) \approx 0.918296$$

$$H(\text{左右}|\text{見てる}) = \mathcal{H}\left(\frac{3}{4}\right) \approx 0.811278$$

$$H(\text{左右}|\text{目撃有無}) = \frac{3}{5}\mathcal{H}\left(\frac{1}{3}\right) + \frac{2}{5}\mathcal{H}\left(\frac{3}{4}\right) \approx 0.8755$$

$$H(\text{左右}) - H(\text{左右}|\text{目撃有無}) \approx 1 - 0.8755 = 0.1245$$

一方、

$$H(\text{目撃有無}) = \mathcal{H}\left(\frac{3}{5}\right) \approx 0.970951$$

$$H(\text{目撃有無}|\text{左}) = \mathcal{H}\left(\frac{2}{5}\right) \approx 0.970951$$

$$H(\text{目撃有無}|\text{右}) = \mathcal{H}\left(\frac{4}{5}\right) \approx 0.721928$$

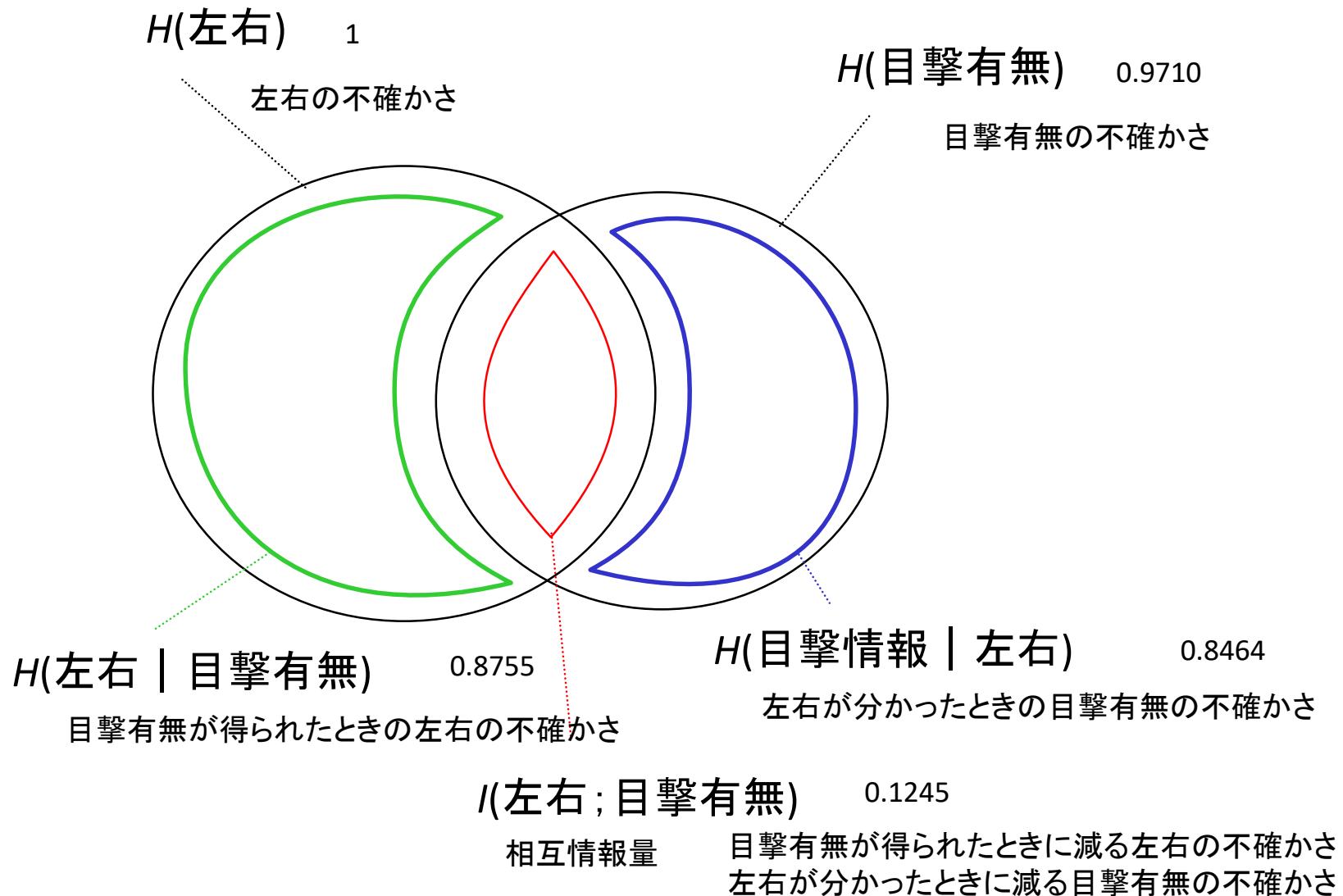
$$H(\text{目撃有無}|\text{左右}) = \frac{1}{2}\mathcal{H}\left(\frac{2}{5}\right) + \frac{1}{2}\mathcal{H}\left(\frac{4}{5}\right) \approx 0.846439$$

$$H(\text{目撃有無}) - H(\text{目撃有無}|\text{左右}) = \mathcal{H}\left(\frac{3}{5}\right) - \frac{1}{2}\mathcal{H}\left(\frac{2}{5}\right) - \frac{1}{2}\mathcal{H}\left(\frac{4}{5}\right) \approx 0.1245$$

このように、

$$H(\text{左右}) - H(\text{左右}|\text{目撃有無}) = H(\text{目撃有無}) - H(\text{目撃有無}|\text{左右})$$

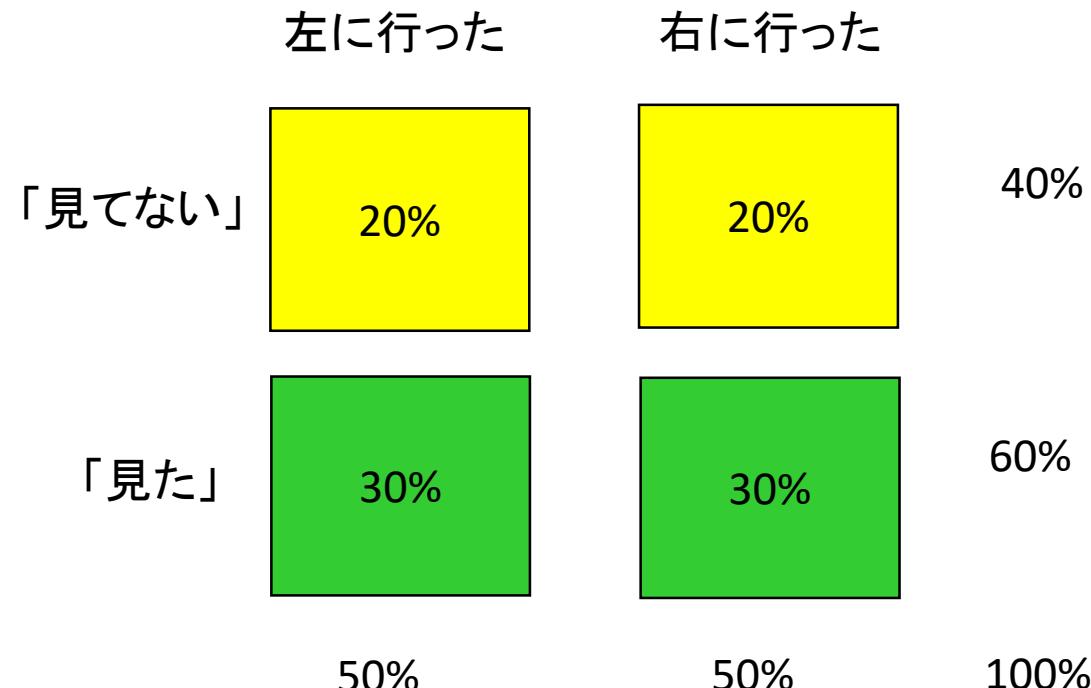
# 相互情報量の概念



# 相互情報量の概念

## ■ 不完全な答えからも情報は得られるか？

「見てない／見た」と「左／右に行った」の関わりについての知識  
... もう一つのケース



# 課題

- 情報符号理論では、情報をどのように単純化しているか？基礎になっている仮定は何か？

# 講義プラン

## 1. 情報符号理論の枠組み

情報符号理論の基礎となる情報量の概念を中心に、情報符号理論の枠組みを紹介する。

## 2. 情報源符号化とその限界

情報源符号化法、情報源符号化定理、情報源のエントロピーについて述べる。

## 3. 通信路符号化とその限界

情報量の概念を導入し、相互情報量、通信路容量、通信路符号化、通信路符号化定理について述べる。

## 4. 通信路符号化法

誤り検出・訂正が可能な符号の構成法について、パリティ検査符号、ハミング符号、巡回符号、BCH符号などを中心に述べる

## 5. アナログ情報源の扱い

フーリエ級数、フーリエ変換など、アナログ情報の扱いについて基本的事項を述べる。