この講義について

- 情報理論 : 金曜1限@工学部3号館・341講義室
  - 秋1期, 全8回
  - 10月5, 12, 19, 26日, 11月2, 9, 16, 27(火)日

    (11/27(火)は金曜授業の実施日)

- 符号理論 : 金曜3限@ IB電子情報館・IB011講義室
  - 秋2期, 全8回
  - 11月30日, 12月7, 14, 21日, 1月11, 25日, 2月1, 8日

    (1/18はセンター試験準備のため休講)

2科目で1セット
… 両方の受講を前提に講義内容を構成

講義に関する情報

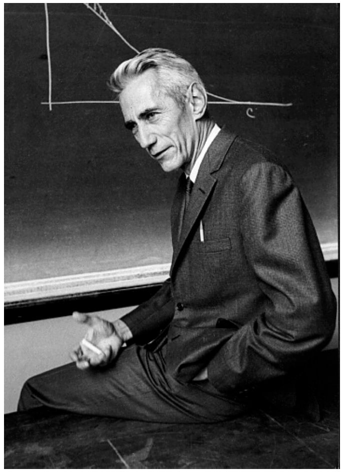
- 担当
  - 梶 勇一（かじ ゆういち）
  - kaji@icts.nagoya-u.ac.jp

- 出欠は取らない
- スライドを使った講義
- 講義で使用するスライド…NUCT で事前に公開（予定）
  - PC等の持ち込みOK（画面で資料を参照することを推奨）
  - （教科書は…買わなくてもなんとかなる）

情報理論

- 1948年の C. E. Shannon の1本の論文からスタート
- 「情報」を科学的に取り扱った最初の研究

- 情報を正確に，効率よく伝えるための理論と技術
- 今日のデジタル技術に多大な影響
  - 有線・無線の通信・放送技術
  - CD/DVD/HDD等のデータ記録技術
  - データ圧縮
  - 暗号，言語学，バイオ情報学，ゲーム理論，...



Claude E. Shannon
1916-2001

講義の構成

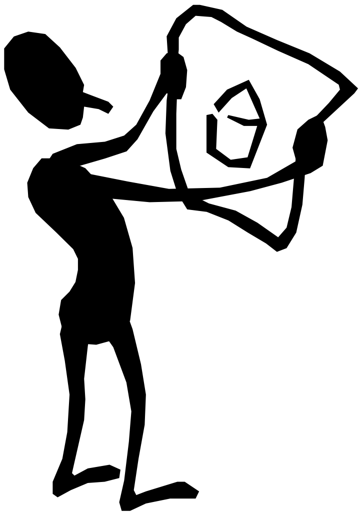
最初の能書き ＋ 3つの章:

- 能書き: 講義内容全体の予告編

- chapter 1: 情報を測る
- chapter 2: 情報をコンパクトに表現する
- chapter 3: エラーから情報を守る

秋1期
情報理論

秋2期
符号理論



NTT ＝日本電信電話株式会社

シャノン当時の時代背景を知る

1940年代の通信技術…

- 広い用途で電信が一般的に使われていた
- モールス符号：「トン（・）」と「ツー （－）」の記号の組み合わせ

● ● ● ▬ ● ▬ ● ▬ ● ▬　　▬ ● ▬ ▬ ▬ ● ▬ ▬

- トン = 1 単位時間, ツー = 3 単位時間
- 記号と記号の間は，1単位時間の空白
- 英文字間は3単位，英単語間は8単位時間の空白

1010100011100010111000101110100011 1

00000001110100011101110111000 1011101110

ある意味で，「デジタル通信」が既に用いられていた

情報処理の自動化・機械化

通信の一部を自動化する「装置」が発達

Teletype model 14-KTR, 1940
http://www.baudot.net/teletype/M14.htm

Enigma machine
http://enigma.wikispaces.com/

機械…人間より高速で，ミスを犯さない（と思われていた）

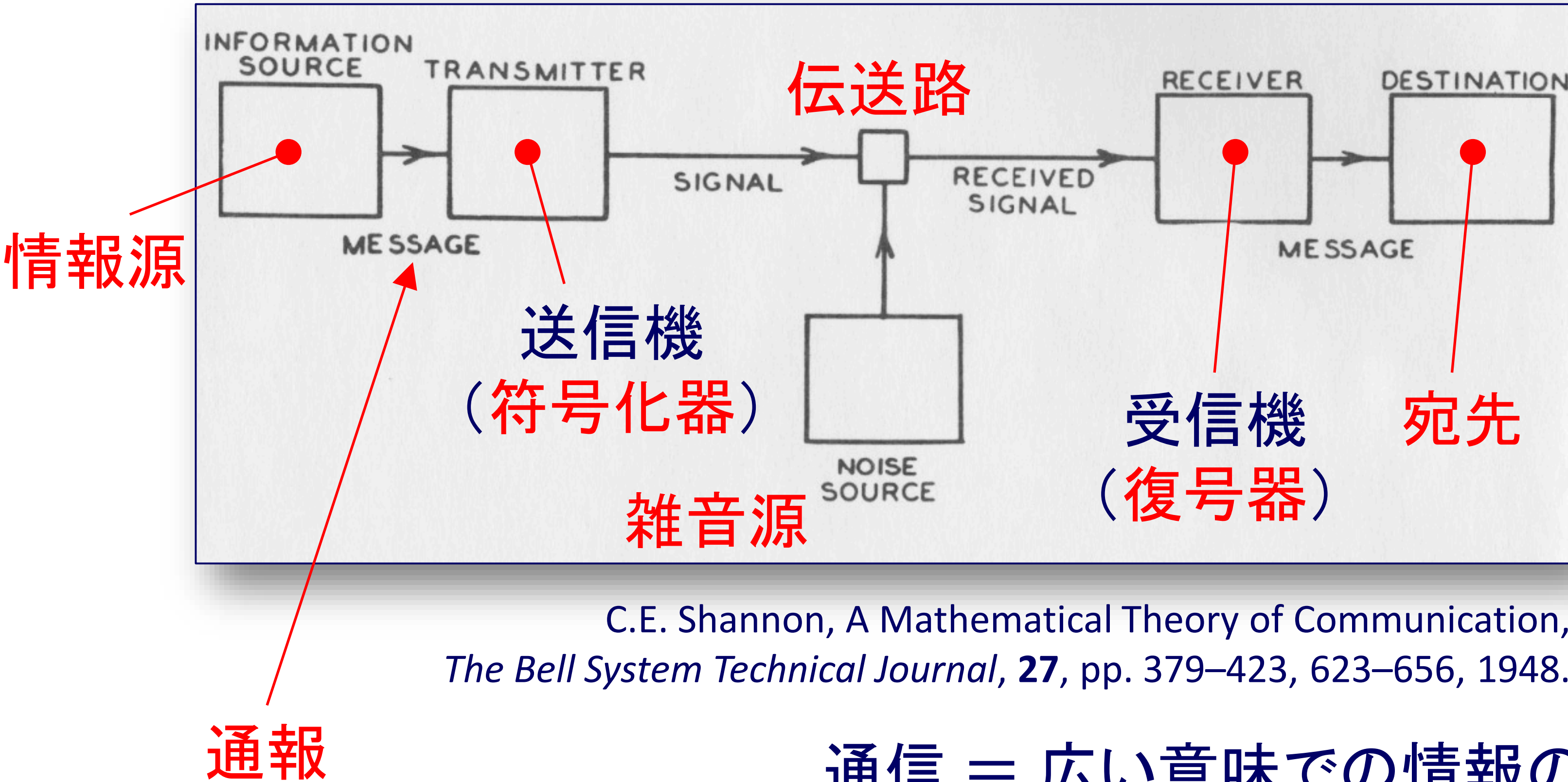
当時の興味の方向性：限られた資源（時間，通信路）の中で…

- 【効率の問題】どれだけ多くの情報を伝えることができるか
- 【信頼性の問題】どれだけ正確に情報を伝えることができるか

通信のモデル：THE figure 1

シャノンのアプローチ...

個別の通信システムではなく，数学的にモデル化して考える



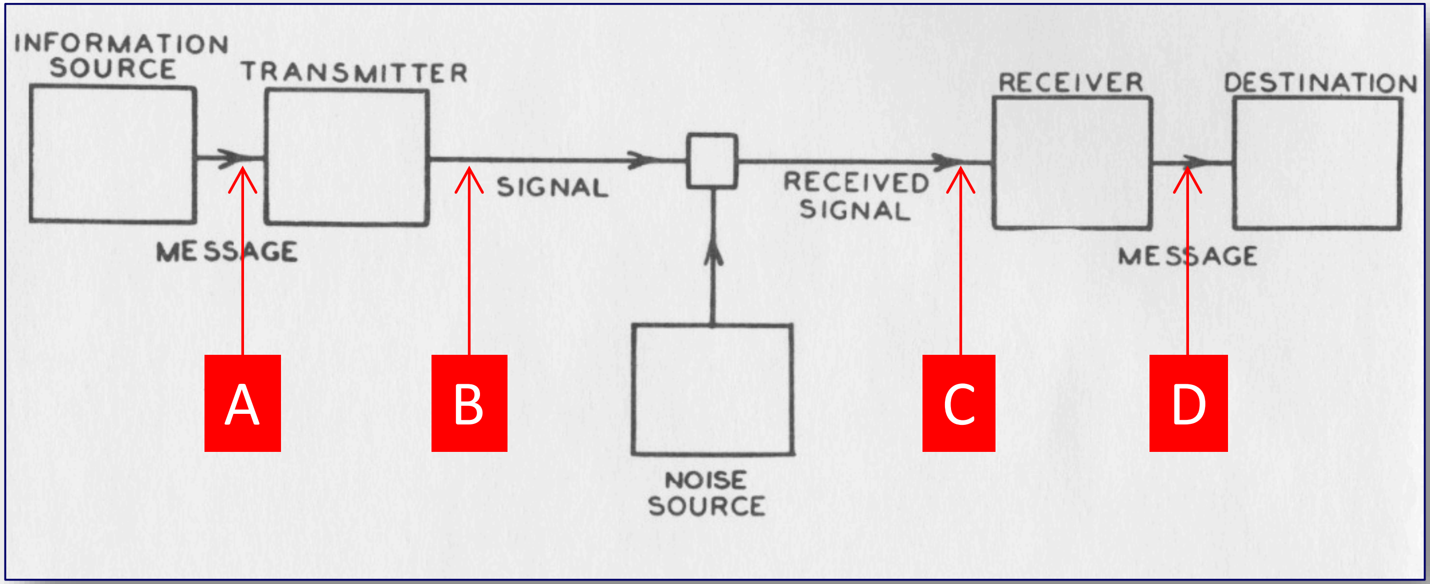
C.E. Shannon, A Mathematical Theory of Communication,
The Bell System Technical Journal, **27**, pp. 379–423, 623–656, 1948.

通信 ＝ 広い意味での情報の伝達

効率的であるとは

通信を効率化する ＝ B のサイズを小さくする

- ただし A = D（または A ≈ D）の必要あり
- 通信路に雑音あり(B ≠ C), 雑音なし (B = C)の２つのケース

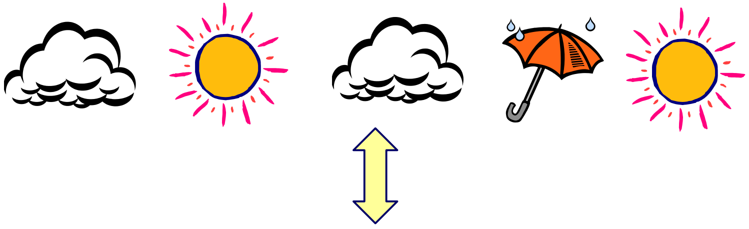


問題その1：効率性

例：天気を毎日記録したい（情報源 ＝天気）

- 通報 = {晴, 曇, 雨}
- 記録には “0” と “1”だけが使用可能（空白等の使用はNG）

天気	符号語
晴	00
曇	01
雨	10

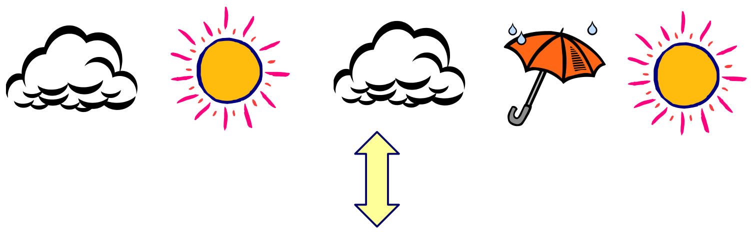


0100011000

- 1日当たり2ビットの記号を記録することになる
- 記録すべきビット数を減らすことができれば，効率改善が可能に

良い符号はあるか？

天気	符号 A	符号 B
晴	00	00
曇	01	01
雨	10	1



符号 A...0100011000
符号 B...010001100

符号 B のほうが，よりコンパクトに情報を表現できる

- 符号語の長さが違っているが，正しく復号できるか?
  - 先頭から処理すれば問題ナシ

- 符号 B よりも良い符号はあるか?
  - Yes でもあり，No でもある(→ 次ページ)

「平均」で考える

天気の発生確率は，一般には均等でない...

天気	確率	符号 A	符号 B	符号 C
晴	0.5	00	00	1
曇	0.3	01	01	01
雨	0.2	10	1	00

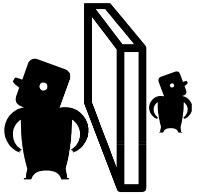
天気の記録に必要なビット数の期待値は

- 符号 A : 2.0 bit
- 符号 B : $2\times0.5 + 2\times0.3 + 1\times0.2 =$ 1.8 bit
- 符号 C : $1\times0.5 + 2\times0.3 + 2\times0.2 =$ 1.5 bit

...「工夫次第で，効率的な符号を作ることができる」

最良の符号

たとえば，一日あたり，平均 0.0000000001 bit で表現できる？
…無理っぽい



- 「どこかに限界がある」ことは，直感的にわかる
- シャノン：「どこに限界があるのかを数学的に解明したい」

→ この確率分布では，一日あたり 1.485 ビットが絶対に必要

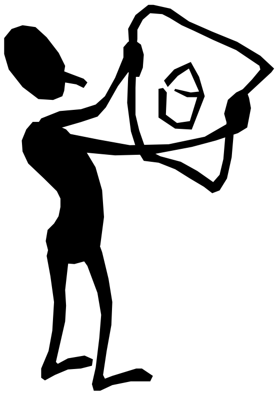
天気	確率
晴	0.5
曇	0.3
雨	0.2

天気の情報そのものの「量」

「情報を格納する容器（符号語）のサイズは，
格納される情報の量よりも小さくできない」

本講義の前半部分について

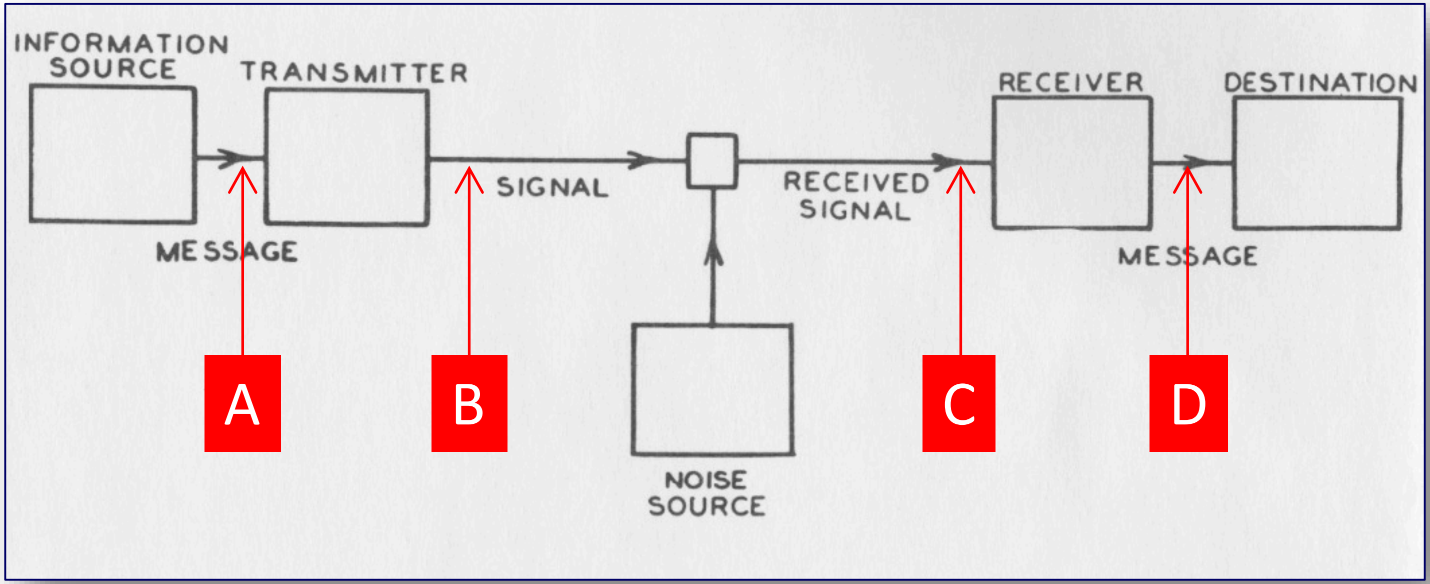
- 能書き：講義内容全体の予告編
- chapter 1: 情報を測る
  - 情報を定量的に測るための技術について学ぶ
- chapter 2: 情報をコンパクトに表現する
  - 情報をコンパクトに表現するための技術と限界について学ぶ
- chapter 3:エラーから情報を守る



信頼性の高さとは

通信の信頼性を上げる ＝「A = D（または A ≈ D）」を保証する

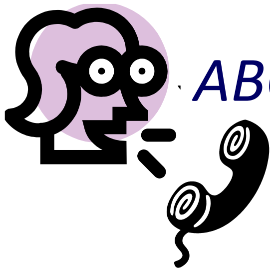
- 雑音の影響により，B ≠ Cとなるおそれがある
- B のサイズをあまり大きくせず，A = Dとなる確率を上げたい



問題その2：信頼性

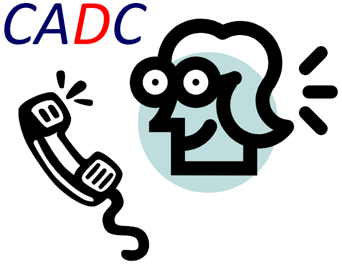
伝送路は，必ずしも信頼できるものではない

- 送信情報 ≠ 受信情報



ABCABC

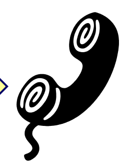
ABCADC



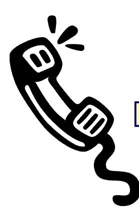
- 伝送路上での誤りを根絶することは難しい

- 日常会話では…「符丁」の利用により問題を回避

ABC ⇨ Alpha, Bravo, Charlie ⇨



あさひの「あ」
いろはの「い」



⇨ Alpha, Bravo, Charlie ⇨ ABC

符丁とは

Alpha

$\underbrace{\text{A}}_{\text{送りたい通報}}\underbrace{\text{lpha}}_{\text{誤り対策のため，やむを得ず付加する冗長な記号}}$

冗長 = 必要のない余分な

- 符丁では，冗長な記号を故意に付加する
- 冗長記号が「文脈」を作り出し，誤りの発見・訂正を助ける

→ これと同じ仕組みを，0-1 データ上で実現したい

冗長性について

Q. どうやって 0-1 データに冗長性を付加するか？

A. パリティビットを使えばよい

パリティビットとは...

データの中の1 の個数を偶数にするための「追加ビット」

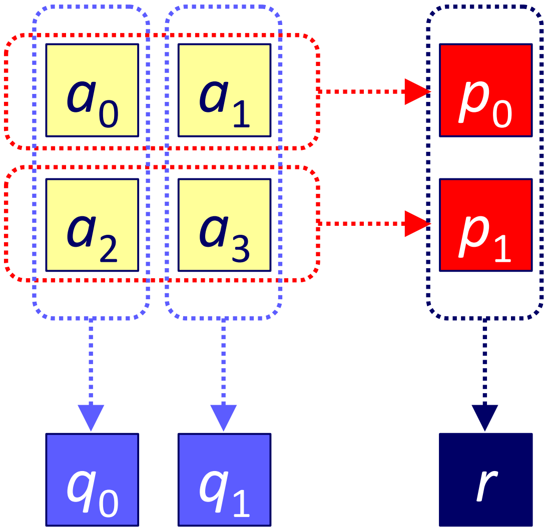
- 00101 → 001010 (2個の1 → 2個の 1)
- 11010 → 110101 (3個の1 → 4個の 1)

パリティビットを一個使うと，奇数個のビット誤りを検出可能

誤りを訂正するには？

パリティビットを複数使うと，誤りを訂正できる（場合もある）

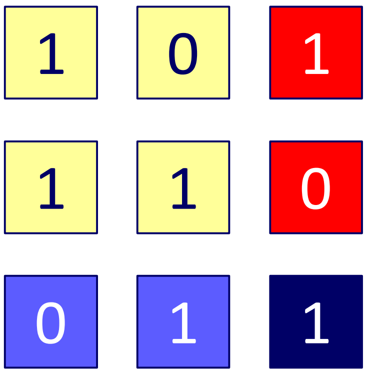
例: 4ビットデータ (a_0, a_1, a_2, a_3)に対し，パリティビットを5個付加



符号語 =
(a_0, a_1, a_2, a_3, p_0, p_1, q_0, q_1, r)

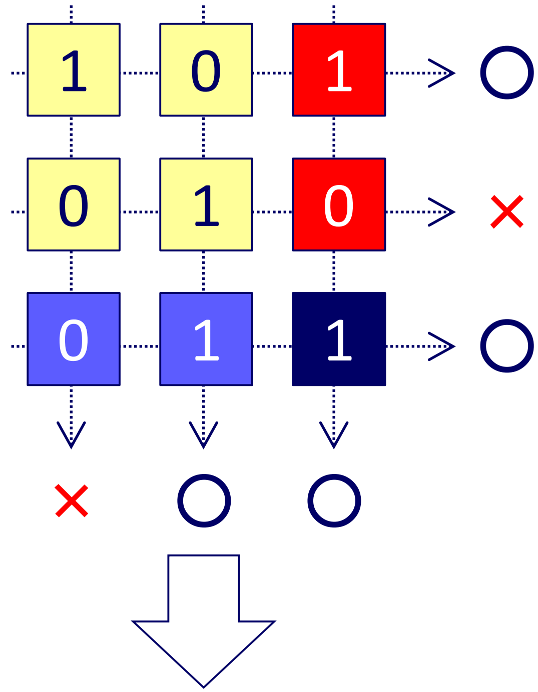
誤り訂正の例

- 1011 を送信する...



符号語 = 1 0 1 1 1 0 0 1 1

- 100110011 が受信された...



3ビット目が怪しい...

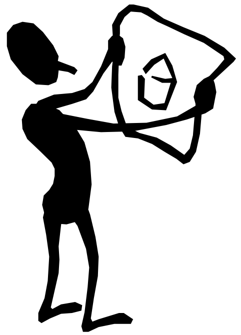
「送信されたのは101110011 だろう」

1ビット誤りを訂正可能
(だが，あまりにも安直)

本講義の後半部分について

- 能書き : 講義内容全体の予告編
- chapter 1: 情報を測る
- chapter 2: 情報をコンパクトに表現する

- chapter 3:エラーから情報を守る
  - 誤りを発見し，訂正するための技術について学ぶ



授業日程

- 情報理論: 金曜1限@工学部3号館・341講義室
  - 秋1期, 全8回
  - 10月5, 12, 19, 26日, 11月2, 9, 16, 27(火)日

(11/27(火)は金曜授業の実施日)

- 符号理論: 金曜3限@ IB電子情報館・IB011講義室
  - 秋2期, 全8回
  - 11月30日, 12月7, 14, 21日, 1月11, 25日, 2月1, 8日

(1/18はセンター試験準備のため休講)

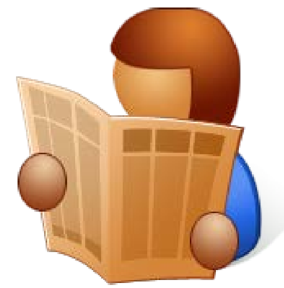
- 途中で演習を実施し, レポート提出を課す
- 各講義の最終回に試験を実施

chapter 1:
情報を測る

測るべき「情報」

情報とは，何かを伝えるもの．ただし…

- まったく興味のないことを教わっても，「情報」とは思わない
- 既知のことを教わっても，「情報」とは思わない



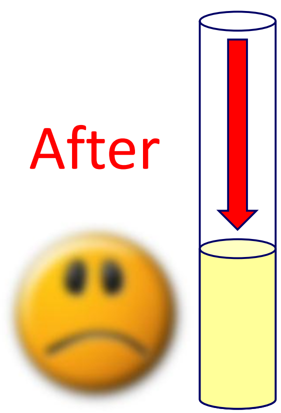
情報とは…

不確実性を持つ興味対象について，その不確実さを減らすもの

Before

不確実さが
大きい

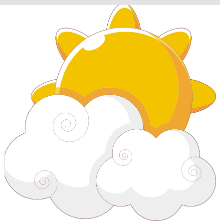
After

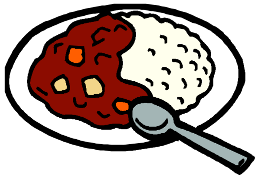


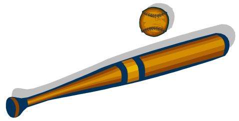
不確実さが
小さい

興味対象の「表現」

- 興味対象は様々
  - 明日の天気，野球の試合結果，テストに出る問題，友人の予定，夕食のおかず …

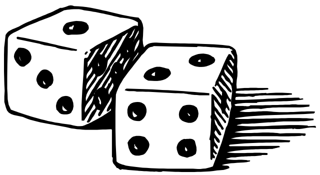






現実の細部はバッサリと切り落とし，確率・統計の世界で考える

- 興味対象は，確率変数の値
  - どれくらいの確率で，どの値を取るかはわかっている
  - 実際に発生する（発生した）値は，いまのところ不明
  - 「サイコロの目」が典型例



確率変数とは

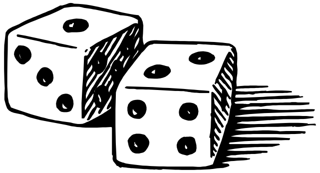
確率変数 X : 中身を覗けない「箱」のようなもの

- 箱の中には，$v_1, \dots v_M$ のどれか一個が入っている
  - $D(X) = \{v_1, \dots, v_M\}$と書く
- 何が入っているかは，箱を開けてみないとわからない
  - 箱を開けたときに出てくる値：確率変数の実現値

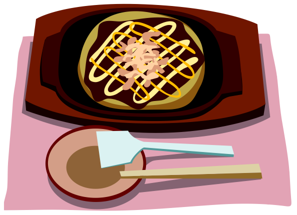
- $X = v_i$である確率が p_i のとき
  - $P_X(v_i) = p_i$ と書く
  - $p_1 + \cdots + p_M = \sum_{v \in D(X)} P_X(v) = 1$

確率変数の例

- 「サイコロの目を，確率変数Xで表す」
  - Xの値は$1, 2, \ldots, 6$ のどれか，全部同じ確率
  - $D(X) = \{1,2,3,4,5,6\}$
  - $P_X(1) = P_X(2) = P_X(3) = P_X(4) = P_X(5) = P_X(6) = 1/6$

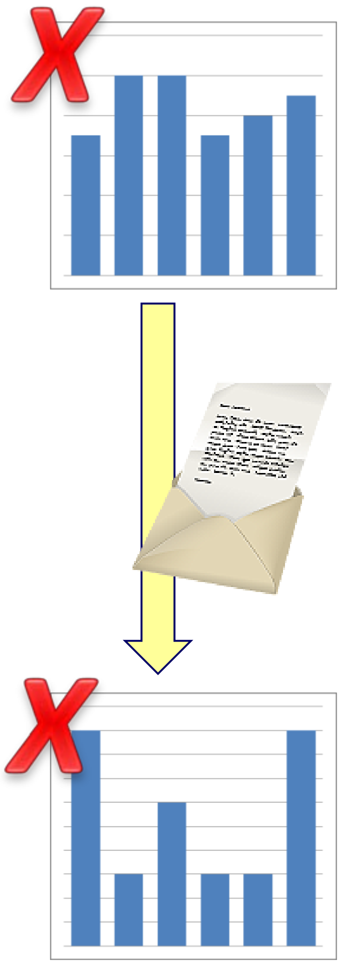


- 「今夜のメニューを確率変数Xで表す」
  - $D(X) = \{$カレー$,$とんかつ$,$ラーメン$, \ldots\}$
  - $P_X($カレー$) = 1/6, P_X($とんかつ$) = 1/4, \ldots$



情報の伝達と確率変数

- 確率変数 Xの実現値を知りたい
  - Xの実現値の集合や，確率分布は既知
  - 実際に X が取った値は不明

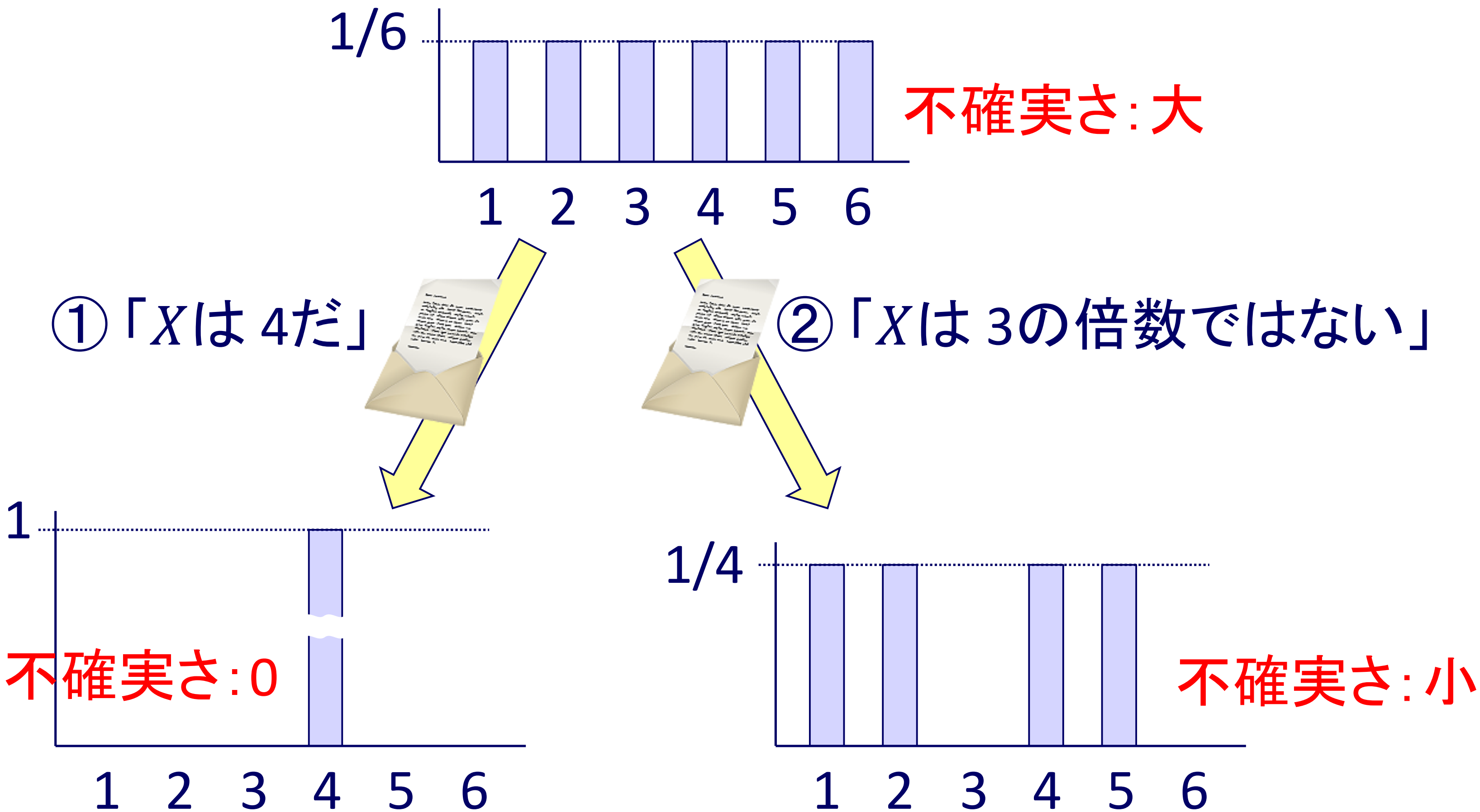


- Xの値について，なんらかの情報を得る

- Xの確率分布が変化する
  - 正確で完全な情報が得られれば…
    - ⇒ X の値が一意に定まる
  - 不正確，不完全な情報が得られれば…
    - ⇒ 多少の不確実さが残る

情報伝達の例

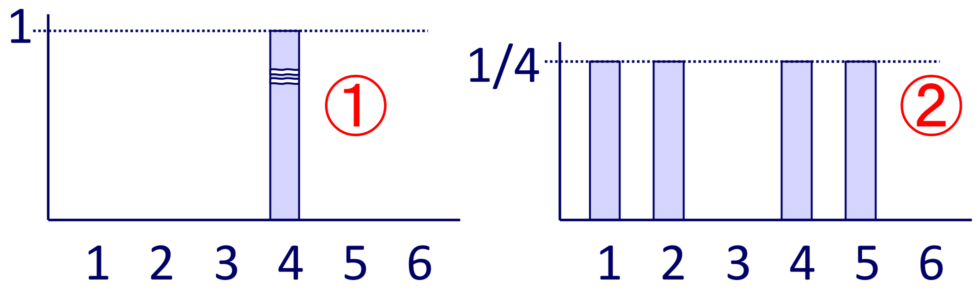
Xはサイコロの目を表す確率変数，$P_X(1) = \cdots = P_X(6) = 1/6$



情報の「量」と不確実さ

①「Xは 4だ」

②「Xは 3の倍数ではない」



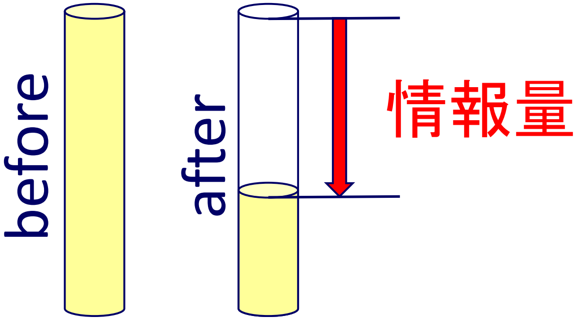
- 直感的には ...

  ①のほうが②よりも大きな「情報量」を持つ，ように思われる

  - ① ... 不確実さを大きく削減
  - ② ... 不確実さを少しだけ削減

「情報量 ＝ 不確実さの削減量」

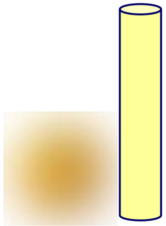
として定義するのが自然



この後のシナリオ

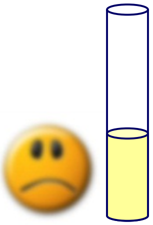
最終目標：**「情報」の量を測る定量的指標**を導入する

- step 1: 確率変数の「エントロピー」を定義
  - エントロピー大 ⇔ 不確実さ大



今回

次回以降



- step 2: エントロピーの変化により，情報量を定義
  - 情報量= (BEFORE エントロピー) – (AFTER エントロピー)

エントロピーの定義

確率変数X... 以下の値と確率分布を持つ

値	v_1	v_2	$\cdots$	v_M
確率	p_1	p_2	$\cdots$	p_M

（値は，あまり重要でない）
（確率値が重要）

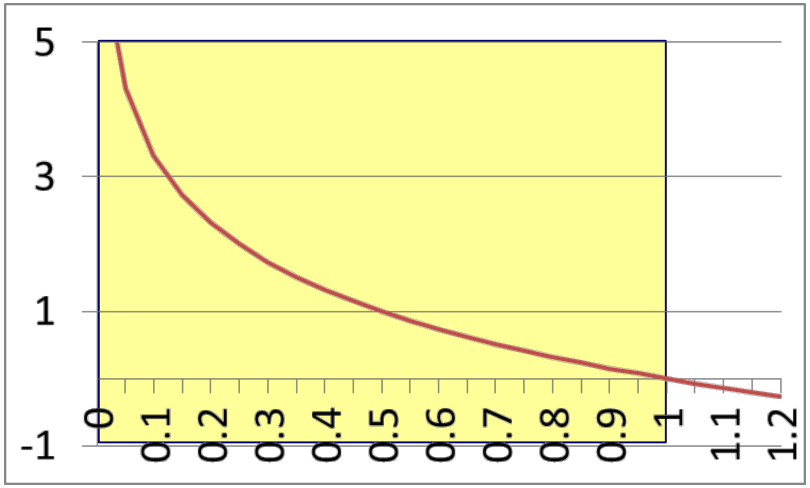
- X の（一次）エントロピー

$$H_1(X) = \sum_{i=1}^{M} -p_i \log_2 p_i = \sum_{v \in D(X)} -P_X(v) \log_2 P_X(v) \ (\mathrm{bit})$$

（ただし，$0\log_2 0 = 0$とする）

- $-\log_2 p_i$ の平均（期待値）と考えることもできる
- $-\log_2 p_i$ を，値v_iの自己情報量と呼ぶ場合も

自己情報量の直感的意味付け



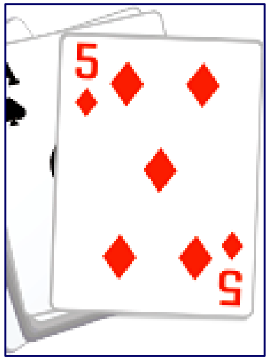
自己情報量 $-\log_2 p$

… 確率pの出来事が起こったと知ったときの「驚き」の量

- pに対して単調減少
  … 滅多にないことが起こる（pが小さい）と，驚きが大きい
- $p > 0$で連続
  … 同程度の確率であれば，驚きも同程度
- $p = q_1 q_2$ならば，$-\log p = -\log q_1 - \log q_2$
  … 驚きの「加法性」に対応している（次ページ）

驚きの加法性

トランプのカードを一枚引く



- N_1=「ダイヤの5だった」... 1/52の確率
- N_2=「ダイヤだった」... 1/4の確率
- N_3=「5だった」... 1/13の確率

N_1を知ったときの驚き ＝ N_2を知ったときの驚き ＋N_3を知ったときの驚き

$$-\log_2 \frac{1}{52} = -\log_2 \frac{1}{4} - \log_2 \frac{1}{13}$$

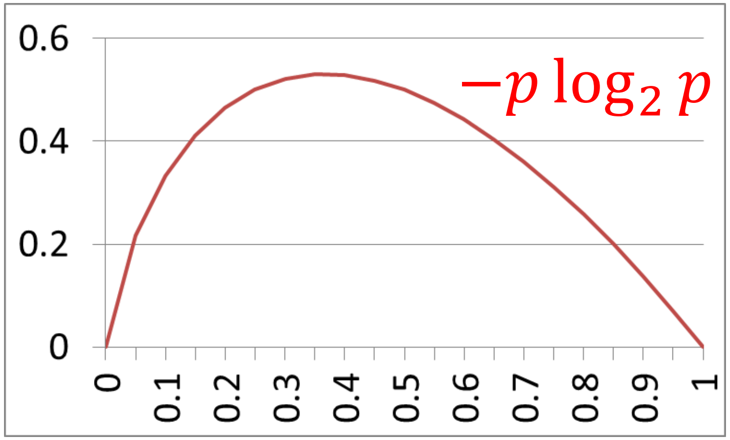
自己情報量は，我々の直感的な理解と良く対応している

エントロピーの定義(再)

- X の(一次)エントロピー

$$H_1(X) = \sum_{i=1}^{M} -p_i \log_2 p_i \quad \text{(bit)}$$

  - 確率で重み付けした，自己情報量の平均値
  - 確率変数の値が与える「驚き」の平均値 ＝ 不確実さ



エントロピー計算の例(1)

- コインを投げて出た面を確率変数Xで表す
  - Xの取りうる値は「表」か「裏」の2種類
  - $P_X(表) = P_X(裏) = 1/2$

$$\begin{aligned} H_1(X) &= -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} \\ &= -\log_2\frac{1}{2} = \log_2 2 = 1 \text{ bit} \end{aligned}$$

  - 1bit の情報は，２進数１桁で表現できる ⇒ Chapter 2

エントロピー計算の例(２)

- ２枚の異なるコインを投げる
  - $X \in \{(表,\ 表),\ (表,\ 裏),\ (裏,\ 表),\ (裏,\ 裏)\}$
  - $P_X((表, 表)) = \cdots = P_X((裏, 裏)) = 1/4$

$$
\begin{aligned}
H_1(X) &= -\frac{1}{4}\log_2\frac{1}{4} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{4}\log_2\frac{1}{4} \\
&= -\log_2\frac{1}{4} = \log_2 2^2 = 2\ \mathrm{bit}
\end{aligned}
$$

  - コイン１枚のときの2倍のエントロピー …不確実さが「2倍」

エントロピー計算の例(3)

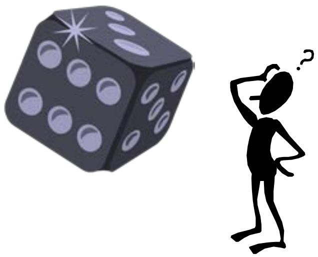
- サイコロ投げ
  - Xの取りうる値は 1, 2, 3, 4, 5, 6
  - $P_X(1) = P_X(2) = \cdots = P_X(6) = 1/6$

$$H_1(X) = -\frac{1}{6}\log_2\frac{1}{6} - \frac{1}{6}\log_2\frac{1}{6} \cdots - \frac{1}{6}\log_2\frac{1}{6}$$
$$= -\log_2\frac{1}{6} = \log_2 6 = 2.585 \text{ bit}$$

  - コイン投げのときと同じ尺度で比較ができる

エントロピー計算の例（4）

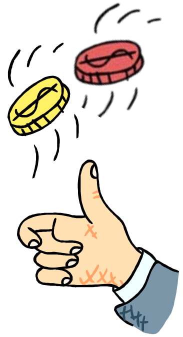
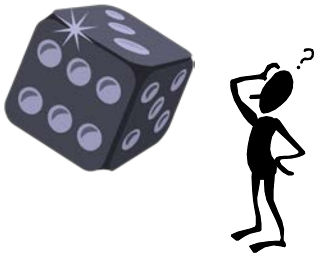
- 公正でないサイコロ
  - Xの取りうる値は 1, 2, 3, 4, 5, 6
  - $P_X(1) = 0.9, P_X(2) = \cdots = P_X(6) = 0.02$



$$
\begin{aligned}
H_1(X) &= -0.9 \log_2 0.9 - 0.02 \log_2 0.02 \cdots - 0.02 \log_2 0.02 \\
&= 0.701 \text{ bit}
\end{aligned}
$$

- コインを1枚投げるときより，不確実さが小さい

唯一尺度としてのエントロピー

			
$H_1(X) = 1$	$H_1(X) = 2$	$H_1(X) = 2.585$	$H_1(X) = 0.701$

- 様々な現象に対し，エントロピーを計算できる
- 違ったタイプの現象の「比較」ができる

- エントロピーが何を意味するのか ... これから議論

今回のまとめ

- 講義概要
  - 情報理論が生まれた背景
  - 講義計画
- エントロピーの定義
  - 「驚きの量」の定式化
  - いくつかの例