北海道大学
Hokkaido University

情報理論

第4回　情報源のモデル (1)

工学部 情報エレクトロニクス学科

情報科学研究院 情報理工学部門
大規模知識処理研究室

堀山 貴史

教科書の構成

- 中村、喜田、湊、廣瀬 著
  基礎から学ぶ 情報理論（第2版）、ムイスリ出版

今回
＋
次回

1. 情報理論とは
2. 情報量とエントロピー

概要と準備

3. **情報源のモデル**
4. 情報源符号化の限界
5. 情報源符号化法

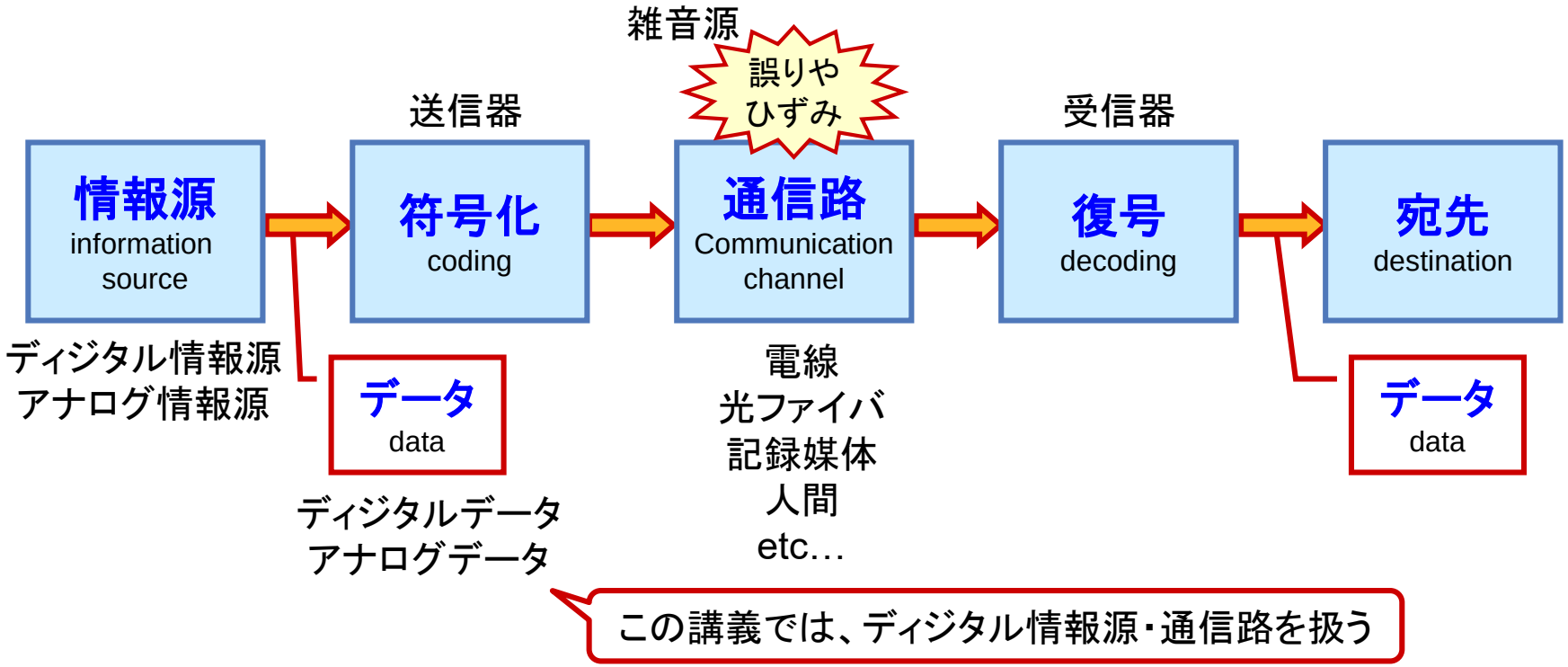
情報源とその符号化

6. 通信路のモデル
7. 通信路符号化の限界
8. 通信路符号化法

通信路とその符号化

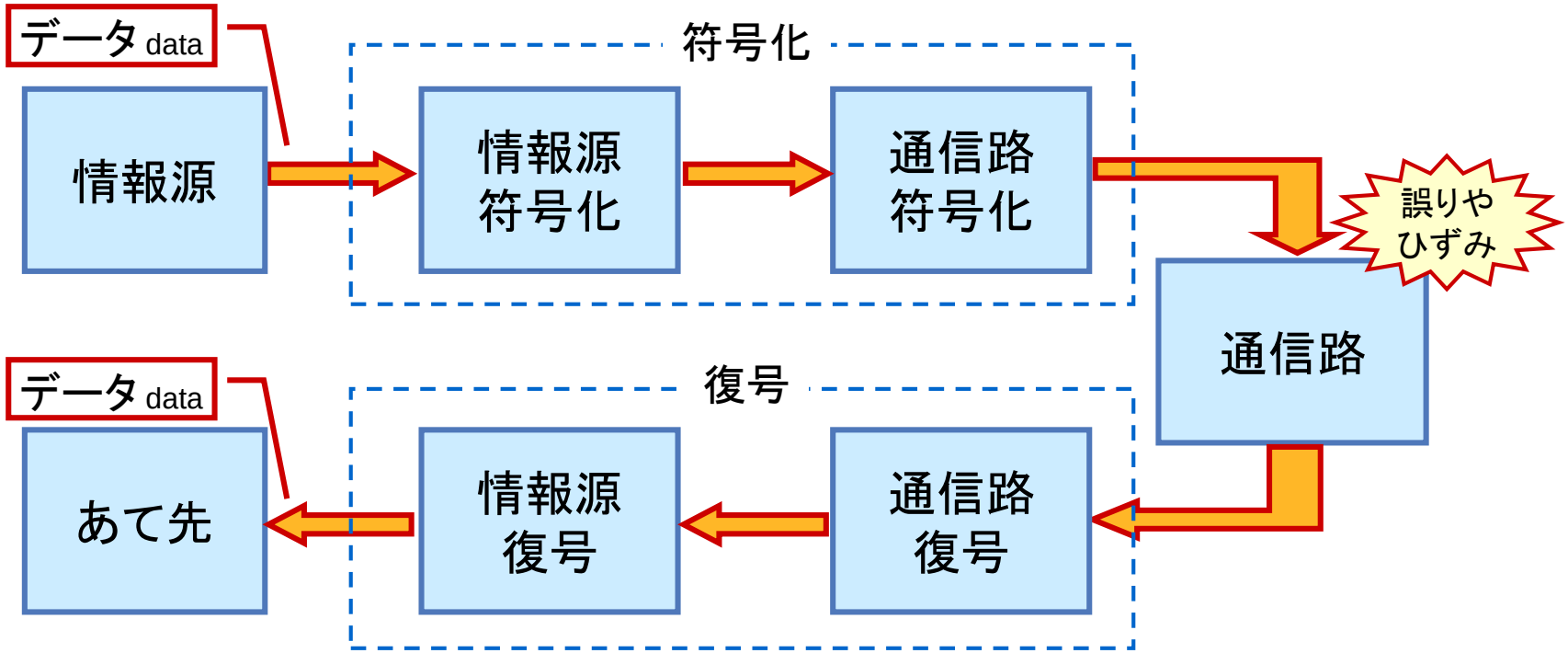
第1回の復習: 通信システムのモデル

- ある通信モデルでの情報伝達の効率や信頼性について考える
  - 情報源から発生するデータを、通信路を通して宛先へ伝達する
    - ただし、通信路では雑音やひずみが存在する



第1回の復習: 通信システムのモデル

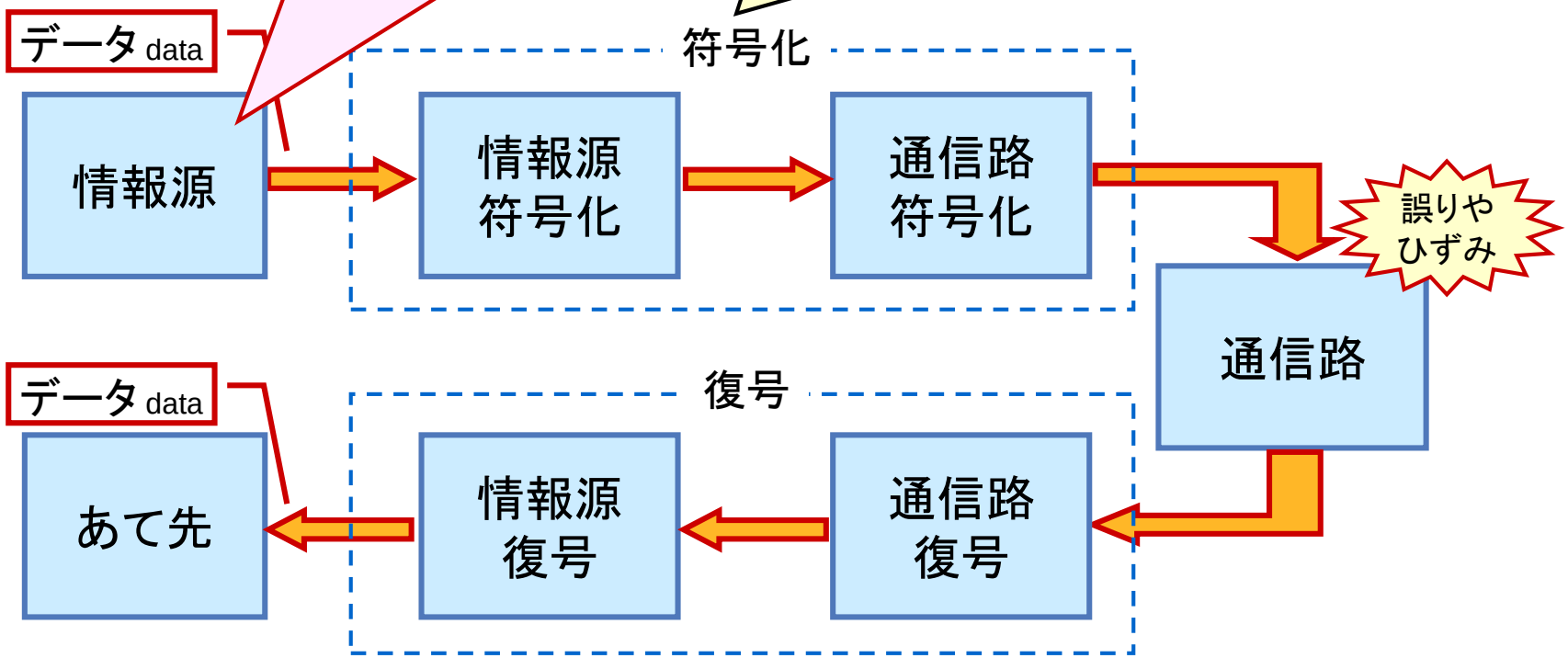
この図は第2回の冒頭に掲載したもの。符号化が「情報源符号化」と「通信路符号化」に分けられることは、第1回で説明済み



これからやること

その前に、ここを
数学的に定義しないと
議論が始まらない
（統計的なモデルで記述する）

この部分の性質を
数学的に考えたい



情報源の統計的表現： 離散的 M 元情報源

- M 個の元からなる記号の有限集合 $A = \{a_1, a_2, \ldots, a_M\}$ を考える
  - A を **情報源アルファベット**、
    各元 a_i $(\in A)$ を **情報源記号**（記号）とよぶ
- 時点 0 より毎時点で（時点は整数値）、
  A 上の情報源記号を、ある確率にしたがって1個ずつ出力する

- 情報源から出力されるデータ = 情報源記号が並んだ記号列
  （**情報源系列**とよぶ）

離散的46元情報源
A={あ, い, う, …, ん}

0	1	2	3	4	5	6	7	8	9	
き	よ	う	の	て	ん	き	く	も	り	…

情報源系列

情報源の統計的表現: 情報源系列の確率分布

- 時点 0 から時点 $n-1$ まで(長さ n)の情報源系列
  - 時点 i での情報源の出力を、確率変数 X_i で表す
    つまり、X_i は $a_1, a_2, \ldots, a_M$ のいずれか
  - 情報源系列は $X_0\, X_1 \cdots X_{n-1}$

$X_0, X_1, \ldots, X_{n-1}$ の**結合確率分布**

$$P_{X_0\, X_1 \cdots X_{n-1}}(x_0, x_1, \ldots, x_{n-1})$$
$$= [X_0 = x_0, X_1 = x_1, \ldots, X_{n-1} = x_{n-1} \text{となる確率}]$$

$p(x_0, x_1, \ldots, x_{n-1})$ とも書く

結合確率分布の例

添え字を省略している
$P_{X_0 X_1 X_2}(x_0, x_1, x_2)$ の意

- 情報源アルファベット A = {0, 1}
- 結合確率分布 $P(x_0, x_1, x_2)$ の例を右の表に示す

結合確率分布から、時点0,1,2での
統計的性質は完全に定まる

x_0	x_1	x_2	$P(x_0, x_1, x_2)$
0	0	0	0.008
0	0	1	0.032
0	1	0	0.032
0	1	1	0.128
1	0	0	0.032
1	0	1	0.128
1	1	0	0.128
1	1	1	0.512

結合確率分布の例

添え字を省略している
$P_{X_0 X_1 X_2}(x_0, x_1, x_2)$ の意

- 情報源アルファベット A = {0, 1}
- 結合確率分布 $P(x_0, x_1, x_2)$ の例を右の表に示す。

結合確率分布から、時点0,1,2での
統計的性質は完全に定まる

x_0	x_1	x_2	$P(x_0, x_1, x_2)$
0	**0**	**0**	**0.008**
0	**0**	**1**	**0.032**
0	**1**	**0**	**0.032**
0	**1**	**1**	**0.128**
1	0	0	0.032
1	0	1	0.128
1	1	0	0.128
1	1	1	0.512

例）$X_0 = 0$ となる確率 $P_{X_0}(0)$ を求める

$$P_{X_0}(0) = \sum_{x_1=0}^{1} \sum_{x_2=0}^{1} P(0, x_1, x_2)$$
$$= 0.008+0.032+0.032+0.128$$
$$= 0.2$$

条件付き確率分布の求め方

- **条件付き確率分布** $P_{X_1|X_0}(x_1 \mid x_0) = P_{X_0X_1}(x_0, x_1) / P_{X_0}(x_0)$
  - $X_0 = x_0$ という条件の下での $X_1 = x_1$ となる確率（割合）

  例）$X_0 = 0$ という条件の下での $X_1 = 0$ となる条件付き確率 $P_{X_1|X_0}(0|0)$

$$P_{X_1|X_0}(0|0) = \frac{P_{X_0X_1}(0, 0)}{P_{X_0}(0)} = \frac{0.008+0.032}{0.2} = 0.2$$

x_0	x_1	x_2	$P(x_0, x_1, x_2)$
0	**0**	**0**	**0.008**
0	**0**	**1**	**0.032**
0	1	0	0.032
0	1	1	0.128
1	0	0	0.032
1	0	1	0.128
1	1	0	0.128
1	1	1	0.512

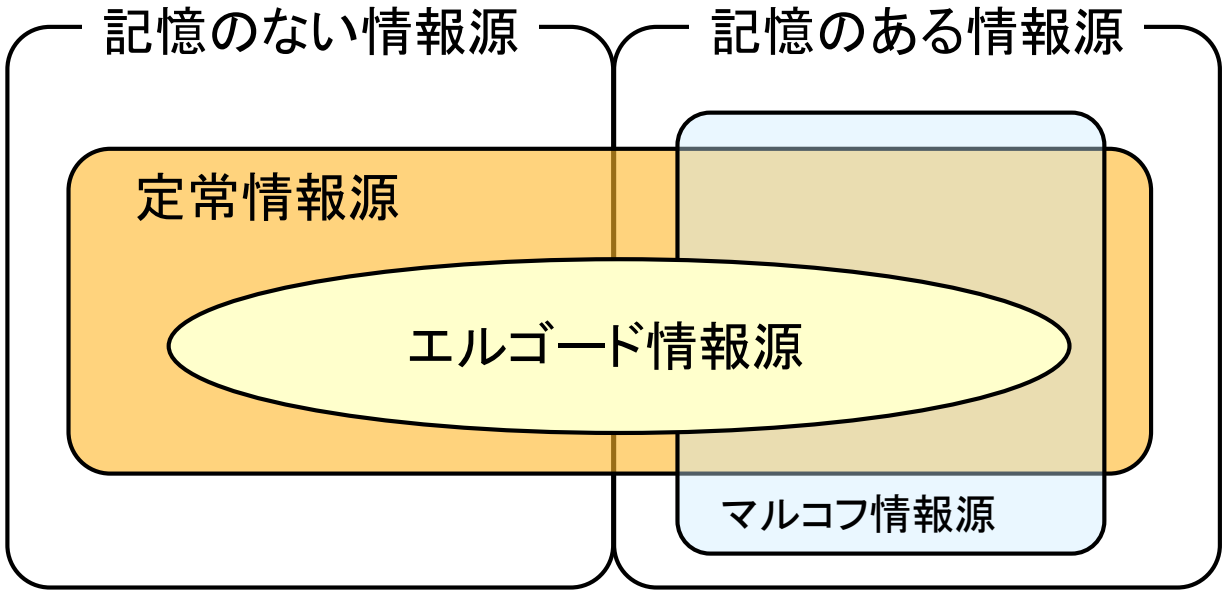
0.2

さらに、X_0とX_1で条件を付けたX_2の条件付き確率分布

$$P_{X_2|X_0,X_1}(x_2 \mid x_0, x_1) = \frac{P(x_0, x_1, x_2)}{P(x_0, x_1)}$$

情報源の各種モデル

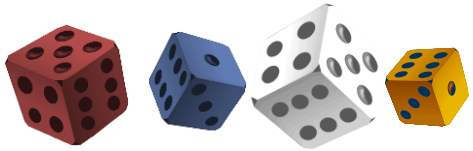
- どんな大きい n についても、$X_0, X_1, \ldots, X_{n-1}$ の結合確率分布を与えることができれば、この情報源の統計的性質は完全に記述できる。しかし、一般には困難。
- 議論を進めるためには、情報源に何らかの**扱いやすい性質**を仮定する必要がある



記憶のない定常情報源（定義3.1）

- **記憶のない情報源**（memoryless information source）
  - 各時点における情報源記号の発生が、**他の時点と独立**である情報源

0 1 2 3

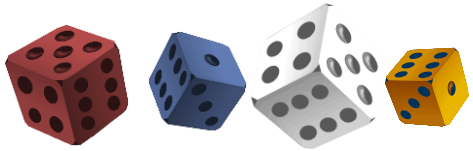


…

記憶のない定常情報源（定義3.1）

- **記憶のない情報源**（memoryless information source）
  - 各時点における情報源記号の発生が、**他の時点と独立**である情報源

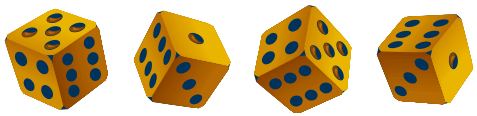
0 1 2 3



…

- **記憶のない定常情報源**（i.i.d.(independently and identically distributed)）
  - 記憶のない情報源において，各時点における情報源記号の発生が**同一の確率分布** $P_X(x)$ に従う情報源

0 1 2 3



…

時点 i によらないので、X の添え字が無い

記憶のない定常情報源の結合確率分布

記憶のない定常情報源における長さnの系列の結合確率分布

$$P_{X_0 \cdots X_{n-1}}(x_0, \ldots, x_{n-1}) = \boxed{}$$

記憶のない定常情報源の結合確率分布

記憶のない定常情報源における長さnの系列の結合確率分布

$$P_{X_0 \cdots X_{n-1}}(x_0, \ldots, x_{n-1}) = \prod_{i=0}^{n-1} P_X(x_i)$$

Σは、その右の和
Πは、その右の積

記憶のない定常情報源の結合確率分布

記憶のない定常情報源における長さnの系列の結合確率分布

$$P_{X_0 \cdots X_{n-1}}(x_0, \dots, x_{n-1}) = \prod_{i=0}^{n-1} P_X(x_i)$$

例題3.3）偏りの無いサイコロの目の出方を調べる。目が 6 ならば a、それ以外ならば b を出力する記憶のない定常情報源 S が系列 $aaba$ を出力する確率はいくらか？

⇒ $\mathrm{A} = \{a, b\}$ の2元情報源で、かつ a, b の発生確率が $\frac{1}{6}$ と $\frac{5}{6}$ となる記憶のない定常情報源になる。よって、

$$P(a, a, b, a) = \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} \times \frac{1}{6} = \frac{5}{1296}$$

（一般の）定常情報源の定義

つまり、記憶があってもよい

- 定常情報源（stationary information source）
  - **時間をずらしても、統計的性質の変わらない情報源**
  - すなわち、任意の正整数 n と i および情報源アルファベットの任意の元 $x_0, x_1, \ldots, x_{n-1}$ に対し、

$$P_{X_0 X_1 \ldots X_{n-1}}(x_0, x_1, \ldots, x_{n-1}) = P_{X_i X_{i+1} \ldots X_{i+n-1}}(x_0, x_1, \ldots, x_{n-1})$$

    が成立するとき、この情報源を**定常情報源**という

  - 定常情報源の出力は、各時点において同一の確率分布に従う。この確率分布を**定常分布**と呼ぶ。

エルゴード情報源

- エルゴード情報源（ergodic information source）
  - 十分長い任意の出力系列に、その情報源の統計的性質が完全に現れている定常情報源
  - エルゴード的であれば、十分に長い任意の系列は、必ず一定の統計的性質を満たす

集合平均と時間平均の統計的性質が一致

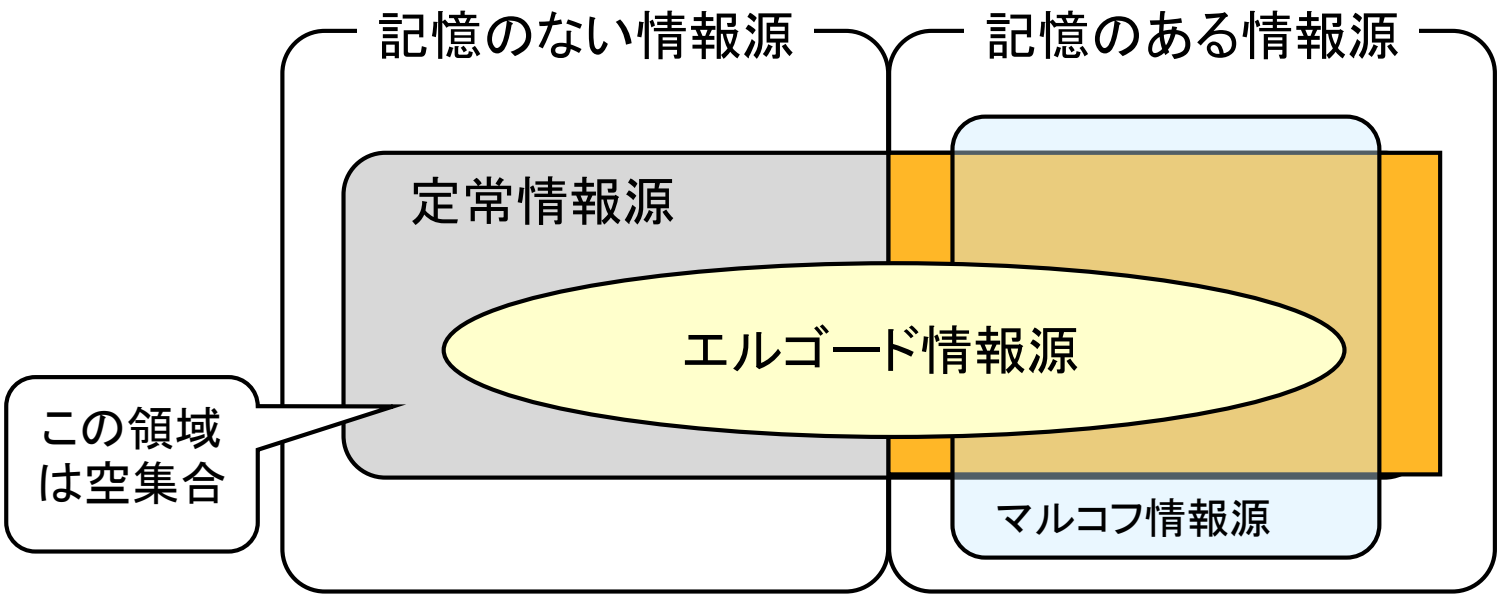
一つの系列を長く観察すれば統計的性質が分かる

同一の情報源

情報源	0	1	0	0	0	0	0	0	1	・	・	・
情報源	0	0	1	0	0	0	1	0	0	・	・	・
情報源	1	0	0	0	0	0	0	0	0	・	・	・
⋮			⋮									
情報源	0	0	0	0	1	1	0	0	0	・	・	・
⋮			⋮									

情報源の各種モデル

- 記憶のない定常情報源は必ずエルゴード情報源になる
- 記憶のある情報源でエルゴード的なものはある（例：正規マルコフ情報源）
- 定常情報源であってもエルゴード的でないものもある
  例）確率1/2で0だけからなる系列か、1だけからなる系列を発生する情報源
  - 最初の１回だけサイコロを振って、２回目以降は１回目と同じ目を出す情報源
- **情報理論では、ほとんどの場合、エルゴード情報源を扱う**



ちょっと休憩

マルコフ情報源（定義3.3）

- マルコフ情報源
  - 記憶のある情報源の基本的なモデル
- **m 重マルコフ情報源**（m-th order Markov information source）
  - n を m以上の任意の整数とする
    任意の時点 i について、その直前の n 個の出力 $X_{i-1}, X_{i-2}, \ldots, X_{i-n}$ で条件を付けた X_i の条件付確率分布が、直前の m個の出力 $X_{i-1}, X_{i-2}, \ldots, X_{i-m}$ だけで条件を付けた X_i の条件付確率と一致するとき、すなわち

$$P_{X_i \mid X_{i-1} \ldots X_{i-n}}(x_i \mid x_{i-1}, \ldots, x_{i-n}) = P_{X_i \mid X_{i-1} \ldots X_{i-m}}(x_i \mid x_{i-1}, \ldots, x_{i-m}) \qquad (n \geqq m)$$

となる

系列中の**過去 m 文字**の並びにしたがって、次の文字の生起確率が決まるモデル

Andrey Andreyevich Markov
(1856 – 1922)

1重マルコフ情報源（単純マルコフ情報源）の例

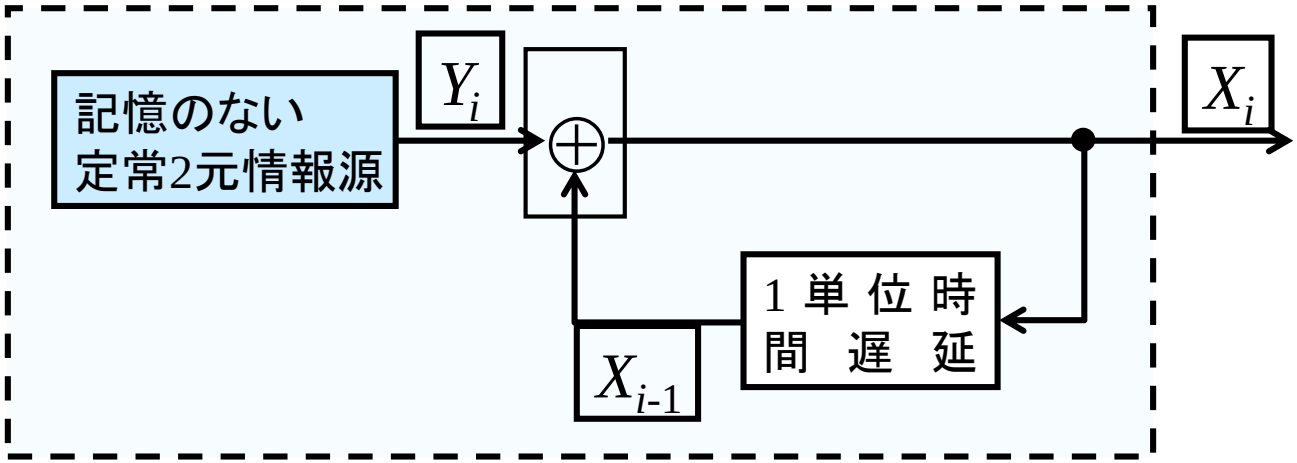
過去１文字

例）1, 0 をそれぞれ確率 p, $1-p$ で発生する記憶のない 2元情報源の出力 Y_i と、1時点前のこの情報源の出力 X_{i-1} とから

$$X_i = X_{i-1} \oplus Y_i$$

⊕ は、排他的論理和
$0 \oplus 0 = 0$、$0 \oplus 1 = 1$、$1 \oplus 1 = 0$

により、現時点の出力 X_i が定まる2元情報源



$$P_{X_i \mid X_{i-1}X_{i-2}}(0 \mid 0\mathbf{0}) = P_{X_i \mid X_{i-1}X_{i-2}}(0 \mid 0\mathbf{1})$$
$$= P_{X_i \mid X_{i-1}}(0 \mid 0) = 1-p$$

2つ前の出力には左右されない

1重マルコフ情報源（単純マルコフ情報源）の例

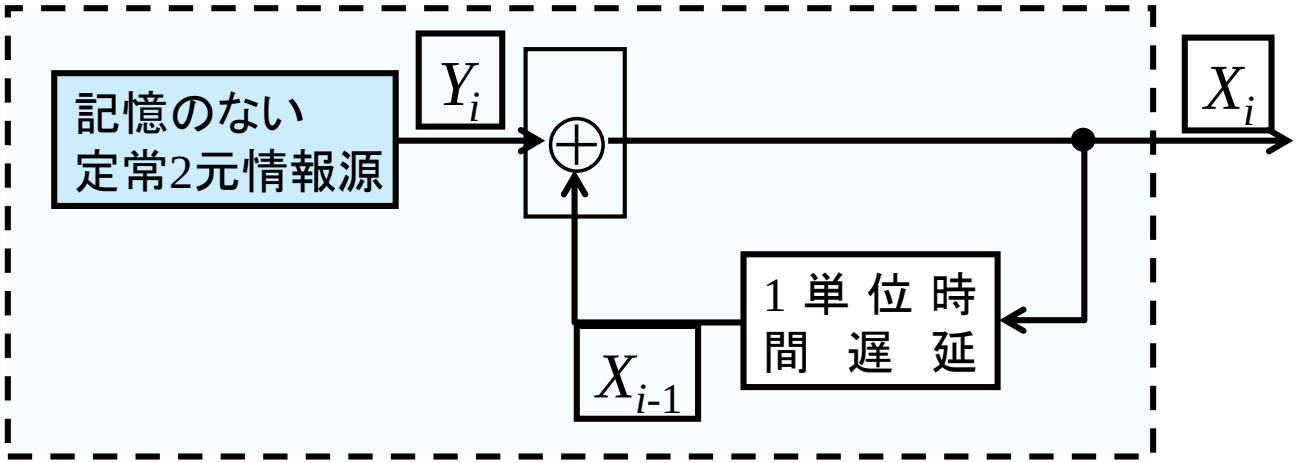
過去１文字

例）1, 0 をそれぞれ確率 p, $1-p$ で出力する記憶のない 2元情報源の出力 Y_i と、1時点前のこの情報源の出力 X_{i-1} とから

$$X_i = X_{i-1} \oplus Y_i$$

⊕ は、排他的論理和
$0 \oplus 0 = 0$、$0 \oplus 1 = 1$、$1 \oplus 1 = 0$

により、現時点の出力 X_i が定まる2元情報源



他の $P_{X_i|X_{i-1}}(x_i \mid x_{i-1})$ は？

$$P_{X_i|X_{i-1}}(1 \mid 0) = p \qquad P_{X_i|X_{i-1}}(1 \mid 1) = 1-p$$

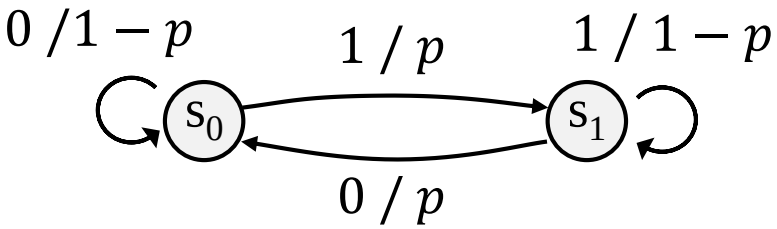
$$P_{X_i|X_{i-1}}(0 \mid 0) = 1-p \qquad P_{X_i|X_{i-1}}(0 \mid 1) = p$$

m 重マルコフ情報源の状態遷移図

前ページのモデルの例

離散的q元情報源を考える
これがm重マルコフ情報源だとする

- q^m個の**状態**を持ち，各状態において記号の出力確率が定まっているとみなせる
- 出力を一つ発生するたびにq^m個の状態の中を**遷移**していく



1重マルコフ情報源

$$P_{X_i \mid X_{i-1}}(1 \mid 0) = p \qquad P_{X_i \mid X_{i-1}}(1 \mid 1) = 1 - p$$

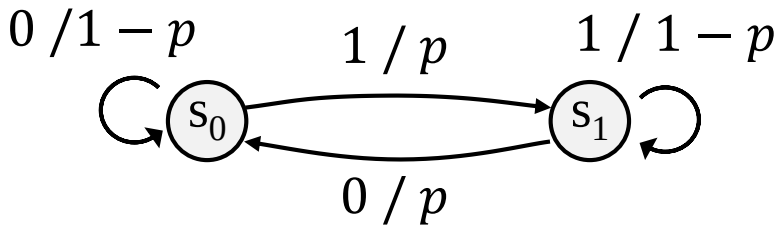
$$P_{X_i \mid X_{i-1}}(0 \mid 0) = 1 - p \qquad P_{X_i \mid X_{i-1}}(0 \mid 1) = p$$

m重マルコフ情報源の状態遷移図

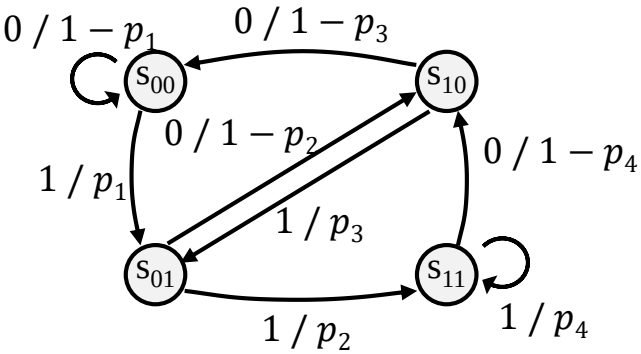
前ページのモデルの例

離散的q元情報源を考える
これがm重マルコフ情報源だとする

- q^m個の**状態**を持ち，各状態において記号の出力確率が定まっているとみなせる
- 出力を一つ発生するたびにq^m個の状態の中を**遷移**していく

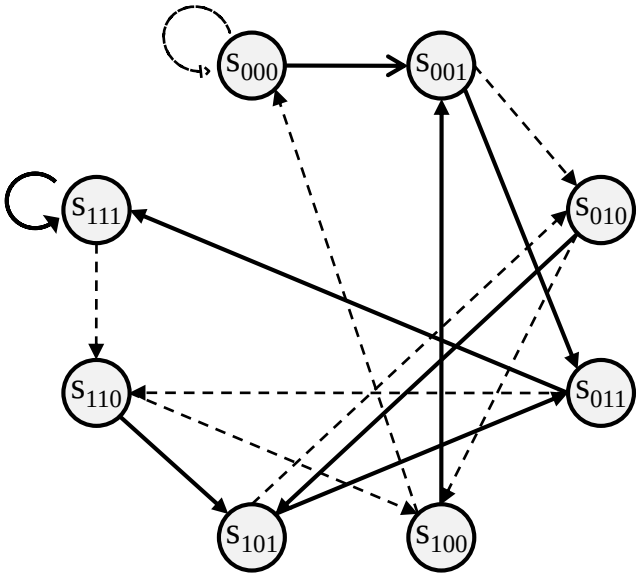


1重マルコフ情報源



2重マルコフ情報源

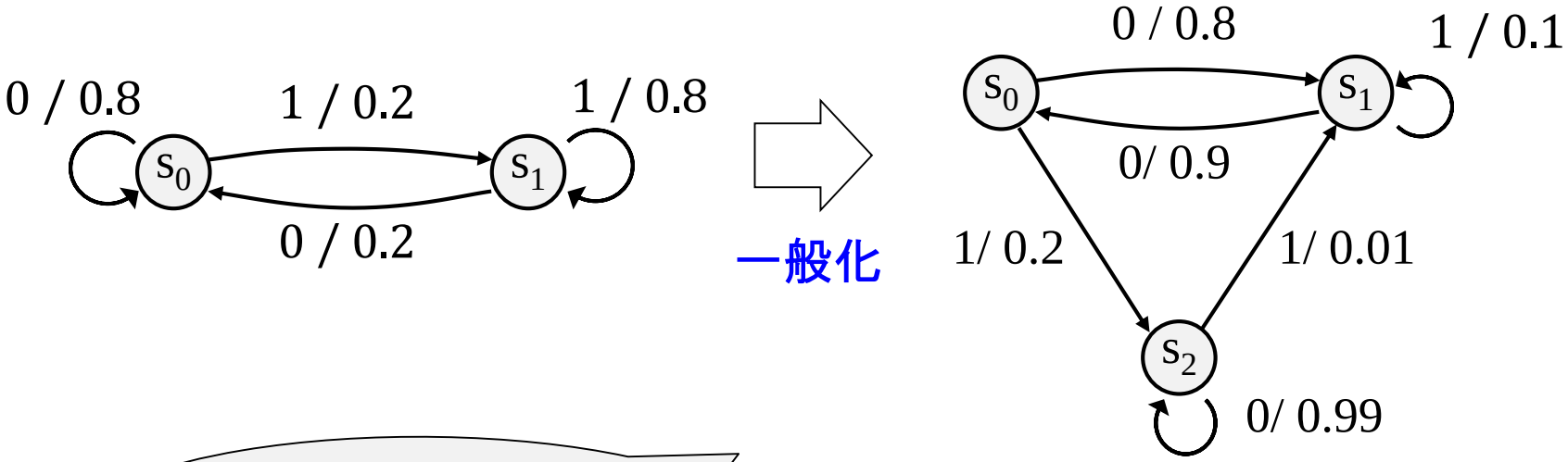
→ 1の遷移
--→ 0の遷移
※確率は省略



3重マルコフ情報源

一般化されたマルコフ情報源

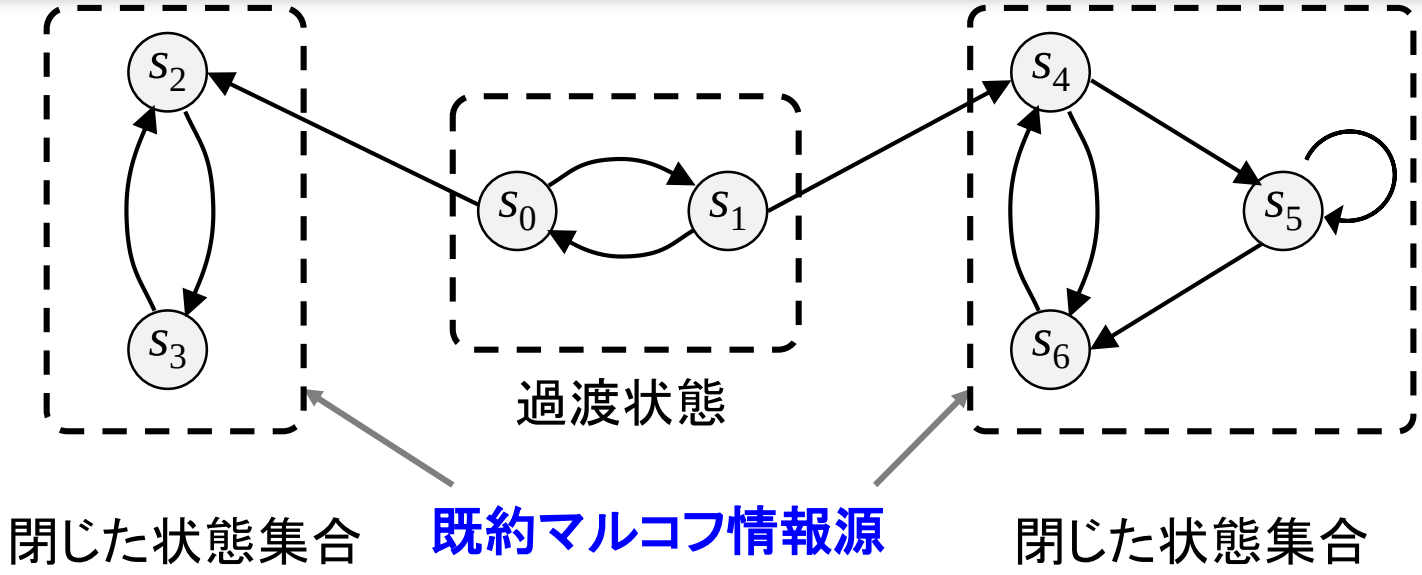
- 単に、マルコフ情報源ともよぶ
- 「直前のm個の出力に対応する」という条件をなくし、より抽象的に（自由な状態図で）情報源を定義できる



マルコフ連鎖（Markov chain）モデルとも言う

一般化されたマルコフ情報源の状態遷移図（m重マルコフとは限らない）

状態の分類

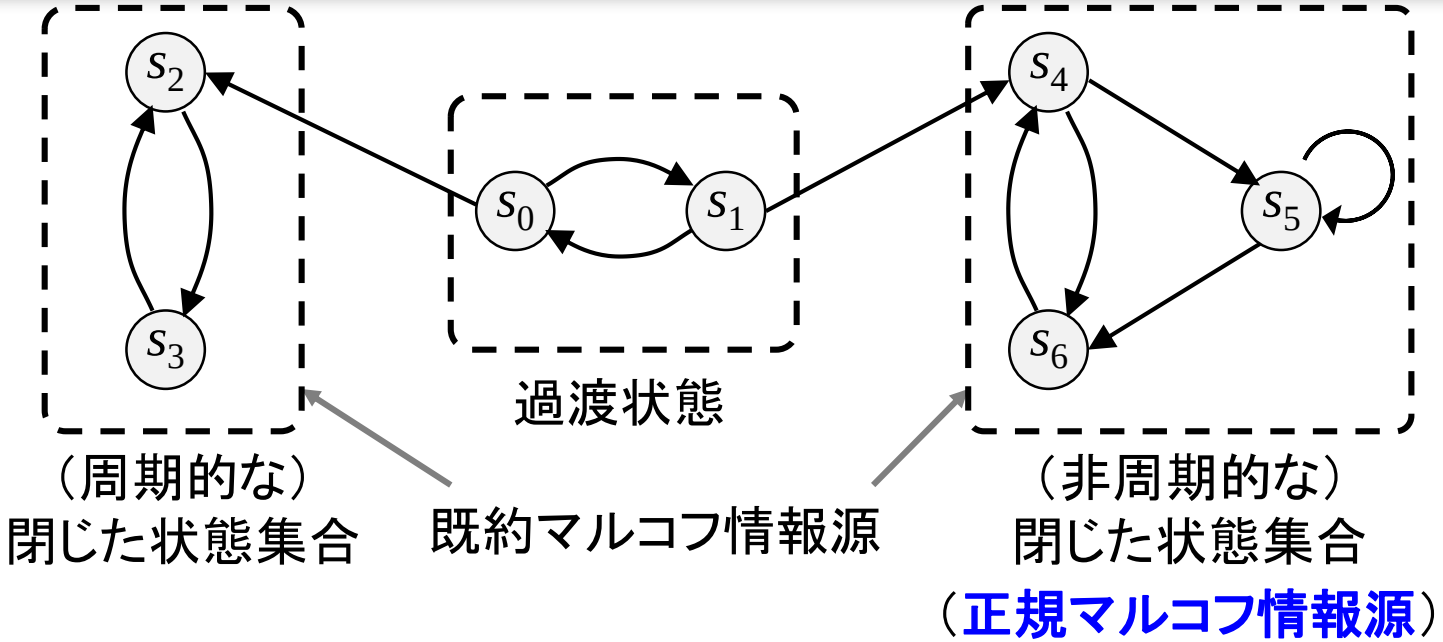


過渡状態：十分時間が経過すれば抜け出てしまい、戻ることのない状態

閉じた状態集合：任意の状態から任意の状態へ遷移可能な
（矢印をたどっていくことのできる）状態の集合

既約マルコフ情報源：
任意の状態から任意の状態へ遷移可能なマルコフ情報源

状態の分類



非周期的状態集合：閉じた状態集合で、ある時間が経過した後の任意の時点において、どの状態にある確率も0でないようなもの

周期的状態集合：閉じた状態集合がいくつかの部分集合に分割され、そのおのおのが周期的な時点においてのみ現れ得るもの

正規マルコフ情報源：
非周期的状態集合のモデルで定義されるマルコフ情報源
十分な時間が経過した後に、定常情報源として扱えて有益（次回）

今日のまとめ

- 情報源のモデル
  - 情報源の統計的性質は**結合確率分布**によって決定される
  - 一般的には、結合確率分布を知ることは難しい
- 情報源の各種モデルについて
  - 記憶のない情報源
  - 定常情報源
  - エルゴード情報源
  - **マルコフ情報源**

    （m 重マルコフ、一般化マルコフ、正規マルコフ）