

HMM って僕にも分かりますか？

篠崎隆宏

東工大

日本音響学会ビギナーズセミナー2013秋

はじめに(確率の復習)

- 確率
 - 事象の起こり易さを0から1の数値で表したものの
- 確率変数
 - 試行によってランダムに値が決まる変数
- ベイズの定理

$$P(Y | X) = \frac{P(X | Y)P(Y)}{P(X)}$$



Example

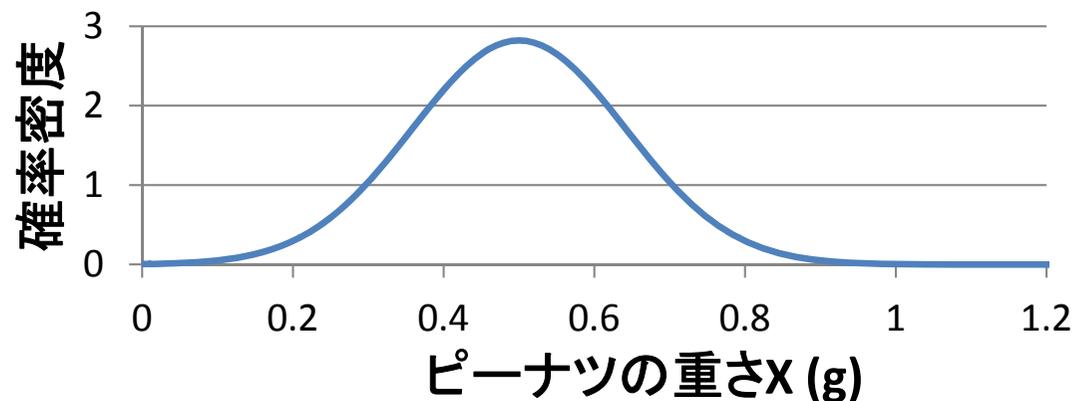
- サイコロの目

- 確率変数 X : サイコロを振った時に出る目
- 確率分布

	$X=1$	$X=2$	$X=3$	$X=4$	$X=5$	$X=6$
$\Pr(X=x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

- ピーナッツの重さ

- 確率変数 X : ピーナッツ1粒の重さ
- 確率密度分布



HMMとは？

Hidden Markov Model
(隠れマルコフモデル)
の略です



では、Markov Modelとは？

- 時間と共に確率分布が変化する確率変数の系列(X_1, X_2, X_3, \dots)のモデルで、各時刻の確率分布がその直前の変数値によって決まるもの

$$\begin{aligned} & P(X_t | X_{t-1}, X_{t-2}, X_{t-3}, \dots) \\ = & P(X_t | X_{t-1}) \end{aligned}$$

X_{t-1} の値が与えられれば、
他の変数は関係ない



Andrey
Markov
1856-1922

図で描くと

$$P(X_1, X_2, X_3, \dots, X_T)$$



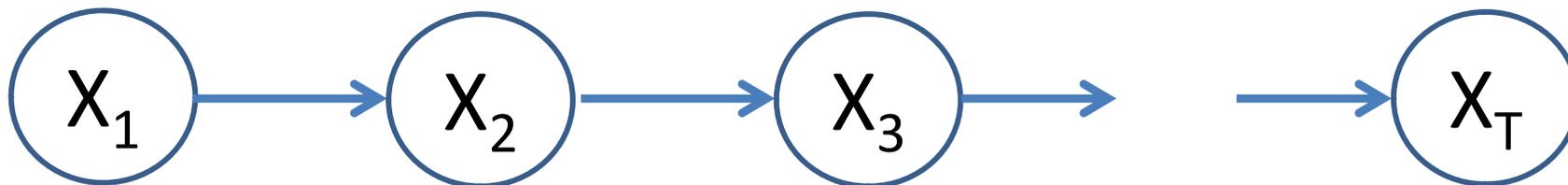
ベイズの定理(入れ子で使用)

$$= P(X_1)P(X_2 | X_1)P(X_3 | X_1, X_2) \cdots P(X_T | X_1, X_2, \dots, X_{T-1})$$



Markovモデルの仮定

$$= P(X_1)P(X_2 | X_1)P(X_3 | X_2) \cdots P(X_T | X_{T-1})$$

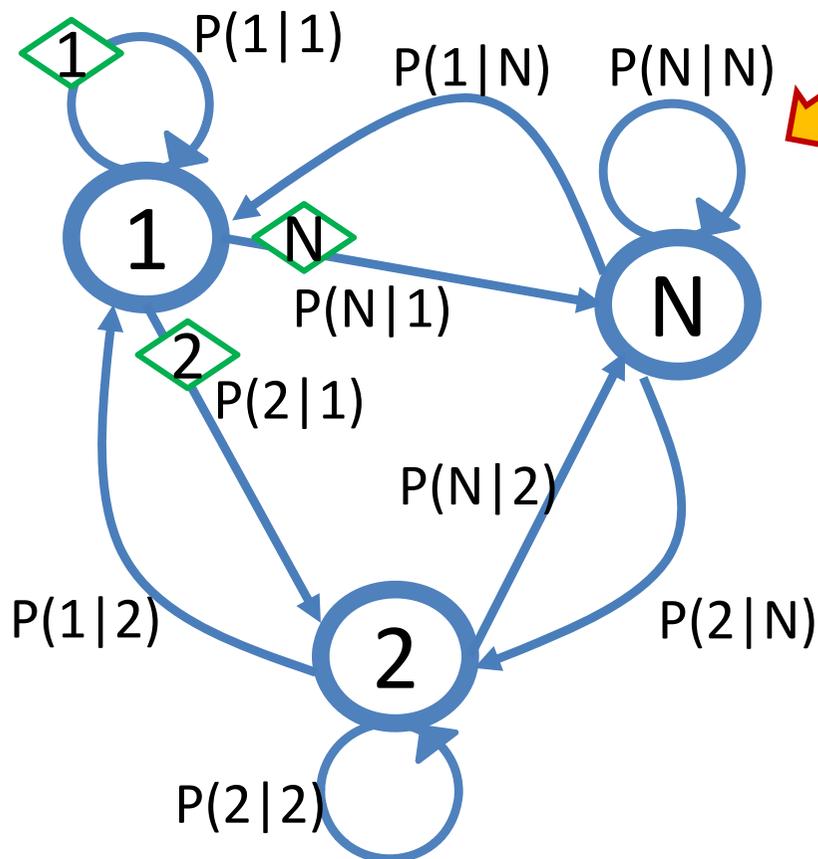


変数をノード、依存関係を枝で表すと、
線形なグラフになる

変数値が有限離散の場合

オートマトンや行列
として表現可能

$X_t = \{1, 2, \dots, N\}$ とすると...

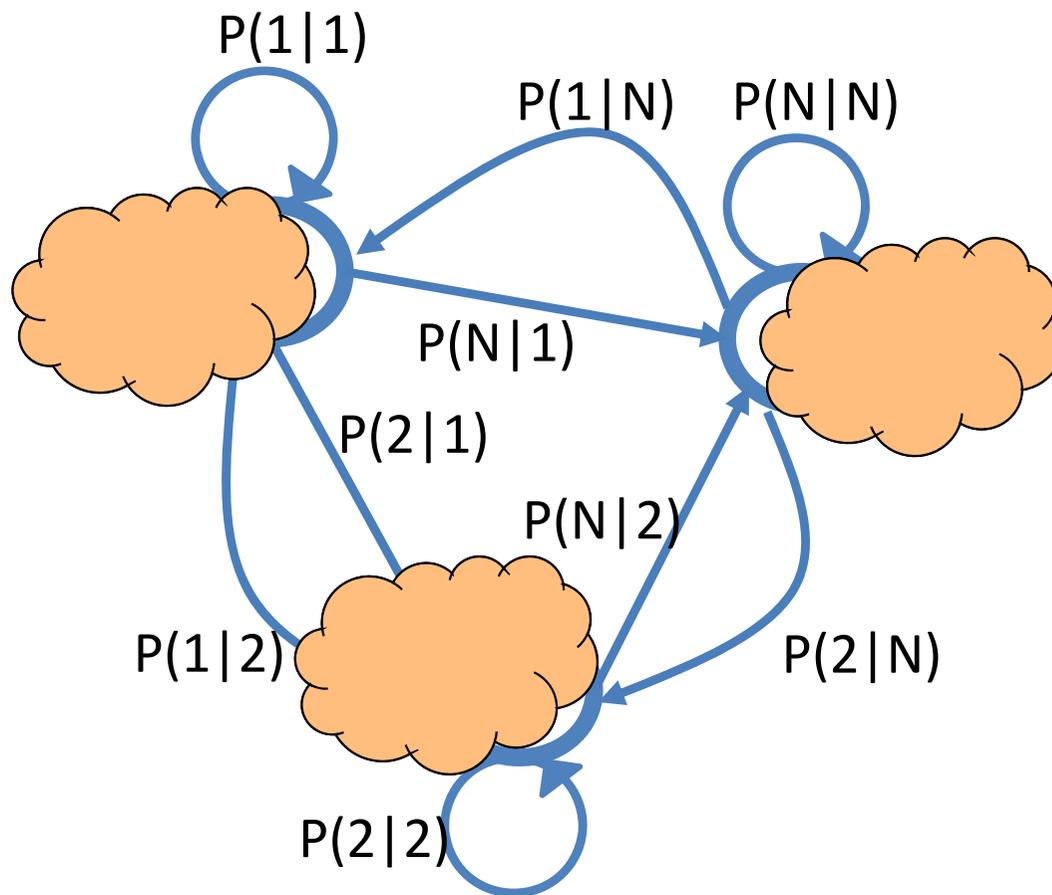


$P(X_t X_{t-1})$	$X_t=1$	$X_t=2$...	$X_t=N$
$X_{t-1}=1$	$P(1 1)$	$P(2 1)$		$P(N 1)$
$X_{t-1}=2$	$P(1 2)$	$P(2 2)$		$P(N 2)$
\vdots				
$X_{t-1}=N$	$P(1 N)$	$P(2 N)$		$P(N N)$

→ X_t が単語の場合はN-gramと呼ばれる

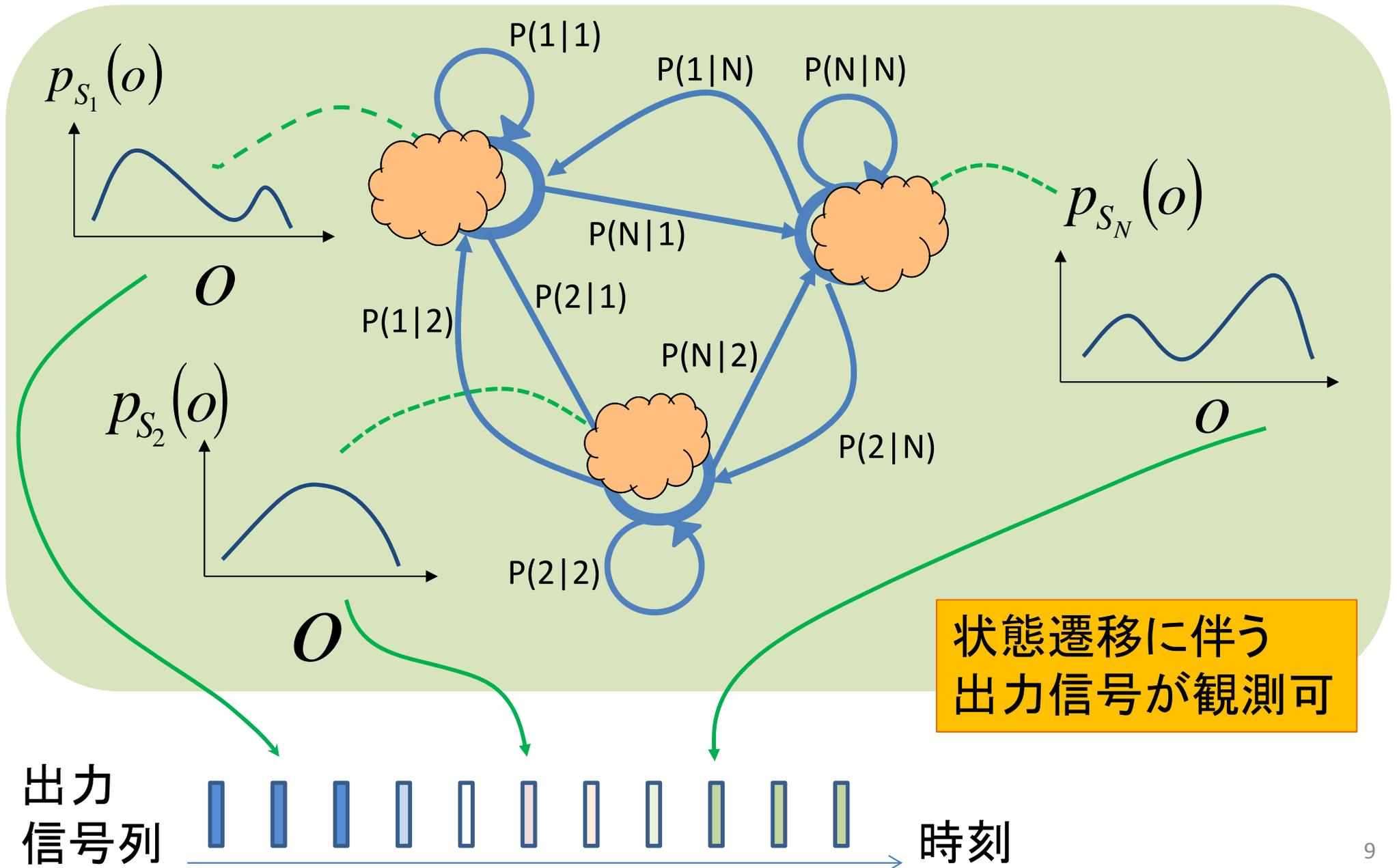
状態が観測できない(Hiddenな)場合

どの状態
にいるのか
見れない



これだけでは余りどうしようもない。。。

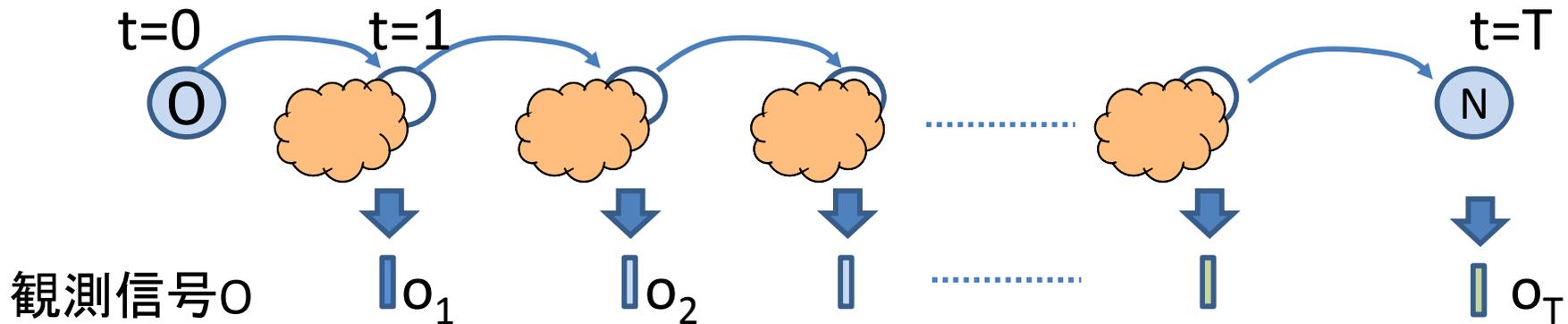
隠れマルコフモデル (HMM)



状態遷移の推定

前提

- 初期状態は s_0 、最終状態は s_N で、既知とする
- T 回の状態遷移を行い、遷移のたびに遷移先状態の出力確率分布に従い信号を出力した



最大事後確率状態系列 $K(= \langle K(0), K(1), \dots, K(T) \rangle)$

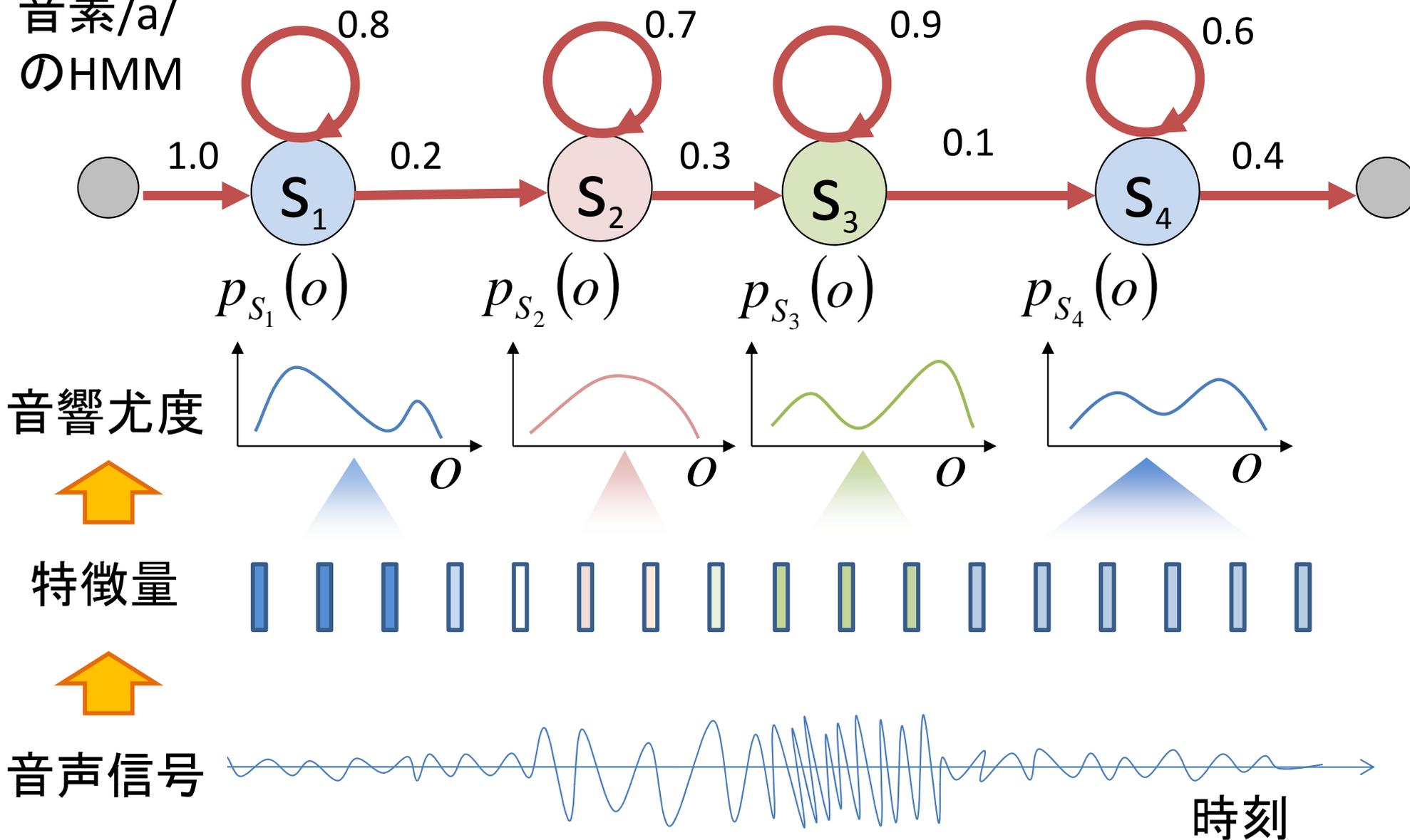
$$= \arg \max_K P(K | \mathbf{O}, HMM)$$

$$= \arg \max_K \prod_{t=1}^T \{P(K(t-1) \rightarrow K(t) | \text{遷移分布}[K(t-1)])P(o_t | \text{出力分布}[K(t)])\}$$

$$= \langle S_0, S_2, S_1, S_1, S_5, S_2, S_3, \dots, S_N \rangle$$

音声認識での利用

音素/a/
のHMM



まとめ

- マルコフモデル
 - 確率変数の系列のモデルで、各時刻の状態遷移確率分布がその直前の時刻の変数の値のみによって決まると仮定したもの
- 隠れマルコフモデル
 - 状態を表す確率変数を外部から直接観測することは出来ない(隠されている)とした、マルコフモデル
 - 状態により決まる確率分布に従った出力信号の観測は可
 - 音声認識では隠された状態変数を音素や単語、出力信号を音声に対応させる