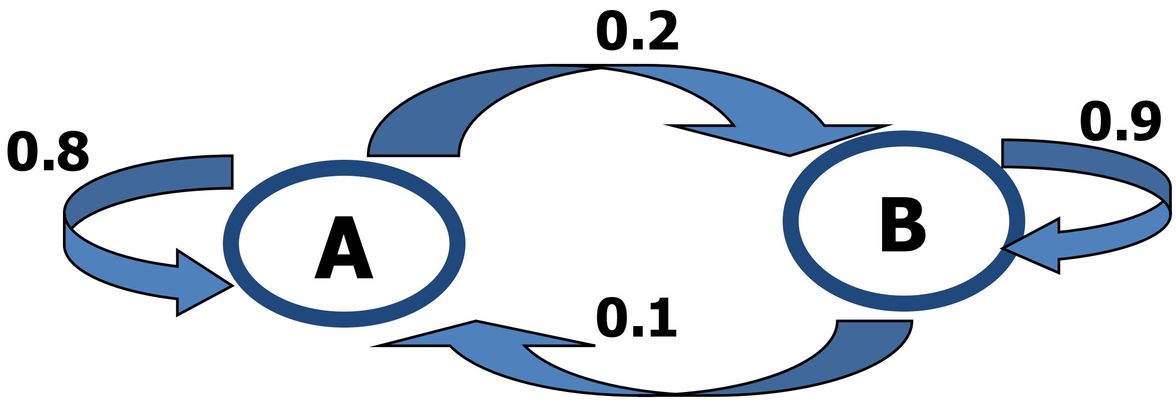
Hidden Markov model and Viterbi's decoding algorithm

Minghua Deng and Dongbo Bu

2012

HMM—韦小宝的骰子

- 两种骰子，开始以2/5的概率出千。
  - 正常A：以1/6的概率出现每个点
  - 不正常B：5,6出现概率为3/10,其它为1/10
- 出千的随机规律



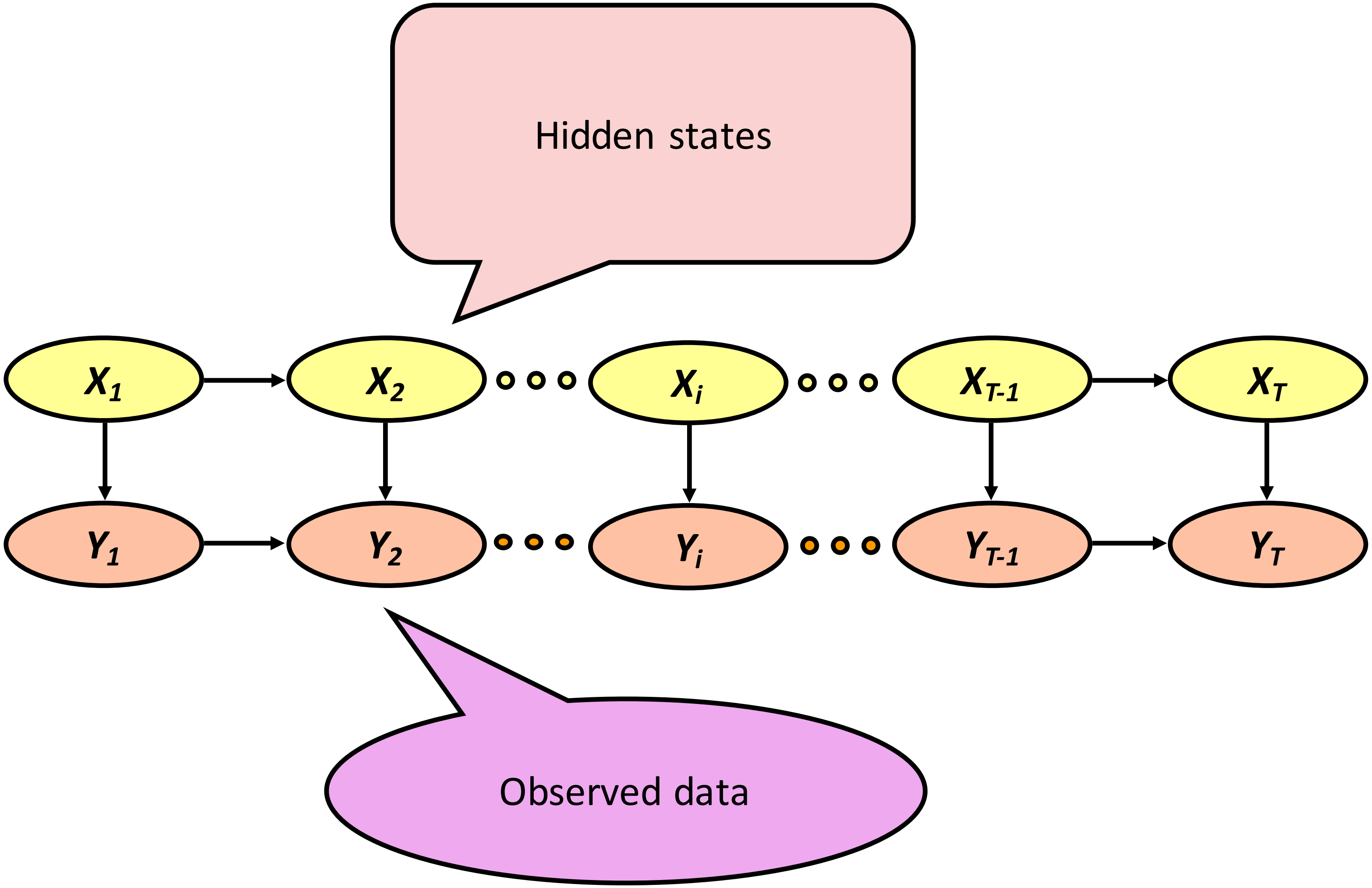
HMM例1—韦小宝的骰子

- 观测到其一次投掷结果

$$O = (1, 3, 4, 5, 5, 6, 6, 3, 2, 6)$$

- 问题：请判断韦小宝什么时候出千了？

Hidden Markov Models - HMM



隐马氏模型的数学模型

- 隐过程为$X=\{X_1,\cdots,X_T\}$
- 观察过程为$Y=\{Y_1,\cdots,Y_T\}$
- 模型参数$\boldsymbol{\lambda} = \{\boldsymbol{\pi}, \boldsymbol{A}, \boldsymbol{B}\}$
  - 初始分布$\boldsymbol{\pi}=(\pi_i)$, $\pi_i=P\{X_1=i\}$
  - 转移矩阵$\mathbf{A}=(a_{ij})$, $a_{ij} = P(X_{n+1}=j \mid X_n= i)$
  - 给定某个时间的隐状态的条件下，观测的分布矩阵$\mathbf{B}=(b_{il})$, $b_{il} = P(Y_n=l \mid X_n= i)$ 。

解码问题(I)

- 问题： 给定观测序列 $Y=(y_1,y_2,\cdots,y_T)$，如何给出隐状态序列 $X^0=(x_1^0,x_2^0,\cdots,x_T^0)$.

- 路径最优指：对任意的$X^0=(x_1,x_2,\cdots,x_T)$ 有

$$
\begin{aligned}
&Pr(x_1',x_2',\cdots,x_T'|y_1,\cdots,y_T)\\
&\geq Pr(x_1,x_2,\cdots,x_T|y_1,\cdots,y_T)
\end{aligned}
$$

解码问题(II)

- 由Bayesian公式有

$$Pr(x'_1, x'_2, \cdots, x'_T | y_1, \cdots, y_T) = \frac{Pr(x_1, x_2, \cdots, x_T, y_1, \cdots, y_T)}{Pr(y_1, \cdots, y_T)}$$

- 又由于序列 **Y** 给定, 问题等价于找最优的X^{0}使联合概率 $Pr(x_1, \cdots, x_T; y_1, \cdots, y_T)$ 最大。

一个相关问题

$$\delta_t(i) = \max_{x_1, \cdots, x_{t-1}} Pr(x_1, \cdots, x_{t-1}, x_t = i, y_1, \cdots, y_t | \lambda)$$

- 求解过程：一系列的决策
- 在每个决策步，确定一个xi是A还是B
- 确定决策项之后，剩下的是求解一个更小的子问题

Viterbi算法(I)

- 算法的思想动态规划的递推算法。
- 递推变量为

$$\delta_t(i) = \max_{x_1,\cdots,x_{t-1}} Pr(x_1,\cdots,x_{t-1},x_t=i,y_1,\cdots,y_t|\lambda)$$

- 我们有递推公式

$$\begin{aligned}\delta_{t+1}(i) &= \max_{x_1,\cdots,x_t} Pr(x_1,\cdots,x_t,x_{t+1}=i,y_1,\cdots,y_{t+1}|\lambda) \\ &= \left(\max_j \delta_t(j)a_{ji}\right) b_i(y_{t+1})\end{aligned}$$

- 以$\psi_t(i)$记录t时刻时使$\delta_t(j)a_{ji}$最大的状态j。

Viterbi算法(II)

- 初始化

$$\delta_1(i) = \pi_i b_i(y_1),$$
$$\psi_1(i) = 0, \quad i = 1, 2, \cdots, N.$$

- 迭代

$$\delta_t(j) = \left(\max_{1\le i\le N} \delta_{t-1}(i) a_{ij}\right) b_j(y_t)$$
$$\psi_t(j) = \operatorname*{Argmax}_{1\le i\le N} \left(\delta_{t-1}(i) a_{ij}\right)$$
$$t = 2, \cdots, T;\ j = 1, \cdots, N.$$

Viterbi算法(III)

- 终止

$$p^\star = \max_{1 \le i \le N} \delta_T(i)$$

$$x_T^\star = \operatorname*{Argmax}_{1 \le i \le N} \delta_T(i)$$

- Backtracking

$$x_t^\star = \psi_{t+1}(x_{t+1}^\star)$$

$$t = T-1, T-2, \cdots, 1.$$

Viterbi算法实例(I)

- 转移概率以及初概率

	A	B
A	0.8	0.2
B	0.1	0.9
初概率	0.6	0.4

- 条件概率(Emission Probability)

	V_1	V_2	V_3	V_4	V_5	V_6
A	1/6	1/6	1/6	1/6	1/6	1/6
B	0.1	0.1	0.1	0.1	0.3	0.3

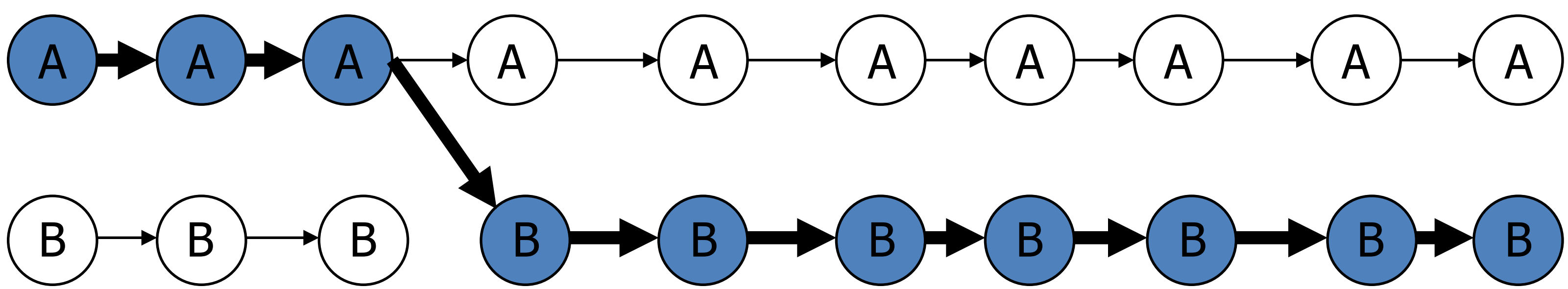
Viterbi算法实例(II)

	y_t	$\delta_t(A)$	$\psi_t(A)$	$\delta_t(B)$	$\psi_t(B)$
t=1	1	1.000×10^{-1}	-	4.000×10^{-2}	-
t=2	3	1.333×10^{-2}	A	3.600×10^{-3}	B
t=3	4	1.778×10^{-3}	A	3.240×10^{-4}	B
t=4	5	3.370×10^{-4}	A	1.067×10^{-4}	A
t=5	5	3.161×10^{-4}	A	2.880×10^{-5}	B
t=6	6	4.214×10^{-6}	A	7.776×10^{-6}	B
t=7	6	5.619×10^{-7}	A	2.100×10^{-6}	B
t=8	3	7.492×10^{-8}	A	1.890×10^{-7}	B
t=9	2	9.989×10^{-9}	A	1.701×10^{-8}	B
t=10	6	1.322×10^{-9}	A	4.592×10^{-9}	B

Viterbi算法实例(III)

观测序列为：

1 3 4 5 5 6 6 3 2 6



解码出来的状态序列为：

A A A B B B B B B B

参考文献

- 钱敏平，龚光鲁。《应用随机过程》，北京大学出版社，1998。
- David W. Mount. Bioinformatics, Sequence and Genome Analysis. Cold Spring Harbor Laboratory Press, 2002.
- Amy N. Langville and Carl D. Meyer. Deeper Inside PageRank, Internet Mathematics Vol. 1, No. 3: 335-380, 2004.
- L. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", Proceedings of the IEEE, Vol. 77, No. 2, Feb. 1989
- On-line tutorial: http://www.comp.leeds.ac.uk/roger/HiddenMarkovModels/html_dev/main.html