

Hidden Markov Model

Nhóm 06

Nội dung

1. Ôn tập xác suất

- Xác suất có điều kiện
- Giả định Markov

2. Bài toán đơn giản

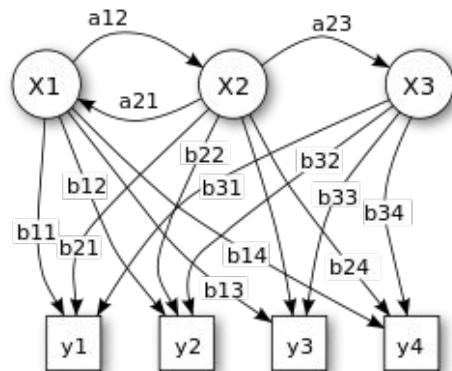
- Bài toán đoán thời tiết
- Mô hình Markov
- Naive solution

3. Tổng quát hóa bài toán

- Viterbi algorithm
- Beam Search

4. Ứng dụng, ưu nhược điểm của HMM

5. Bài toán POS Tagging



Ôn tập xác suất

	Male	Female
B1	50	50
B2	65	55
B3	80	70

+ A: biến cố chọn 1 bạn nữ
+ B1, B2, B3: lần lượt là biến cố chọn 1 bạn trong lớp B1, B2, B3

- 1/Tính xác suất chọn một bạn nữ?
- 2/Tính xác suất một bạn nữ trong lớp B3?

Công thức xác suất đầy đủ

	Male	Female
B1	50	50
B2	65	55
B3	80	70

$$P(B1) = 100/370$$

$$P(B2) = 120/370$$

$$P(B3) = 150/370$$

$$P(A|B1) = 50/100$$

$$P(A|B2) = 55/120$$

$$P(A|B3) = 70/150$$

$$P(A) = \sum_{i=1}^n P(A|B_i) * P(B_i)$$

$$P(A) = \frac{50}{100} * \frac{100}{370} + \frac{55}{120} * \frac{120}{370} + \frac{70}{150} * \frac{150}{370} = \frac{35}{74}$$

Xác suất có điều kiện (Bayes Theory)

	Male	Female
B1	50	50
B2	65	55
B3	80	70

$$P(B_k|A) = \frac{P(B_k) * P(A|B_k)}{P(A)}$$

$$P(B_3|A) = \frac{P(B_3) * P(A|B_3)}{P(A)} = \frac{\frac{150}{370} * \frac{70}{150}}{\frac{35}{74}} = \frac{2}{5}$$

Giả định markov

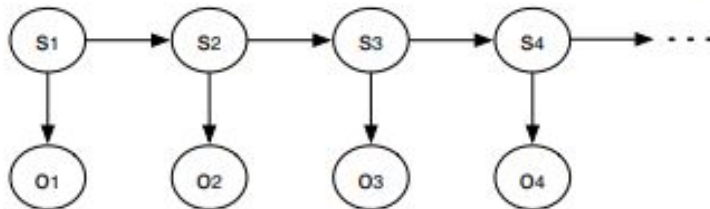
Cho chuỗi thời gian rời rạc $t = 0, 1, 2, \dots$

+ $S_t = s$: Trạng thái ẩn tại thời điểm t

+ $O_t = o$: Quan sát tại thời điểm t

Hệ thống chuyển đổi tại trạng thái

+ $S_{t+1} = P(S_{t+1}|S_t)$, quá trình lặp lại tương tự.



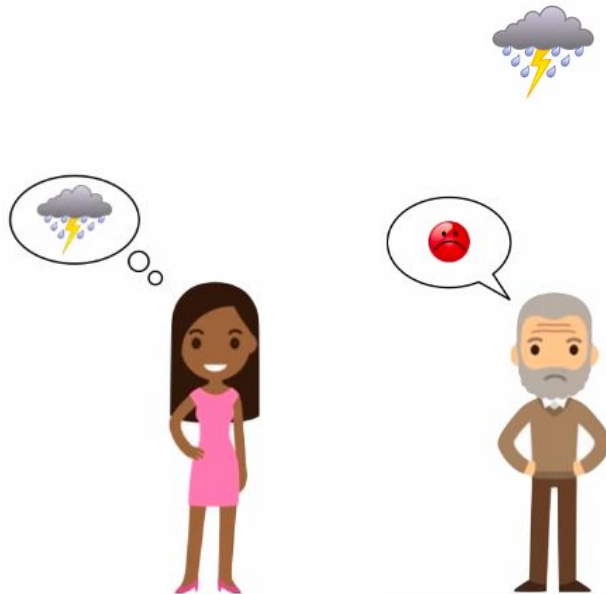
- **Markov assumption:** $S_{t+1} \perp\!\!\!\perp S_{t-1} | S_t$ (future is independent of the past given the present)

Bài toán đơn giản

Bob và Alice là sống ở xa nhau và họ nói chuyện với nhau mỗi ngày qua điện thoại. Bob nói rằng, tâm trạng của anh ta sẽ thay đổi theo thời tiết:

- Nếu trời **nắng**, anh ta sẽ **vui**
- Nếu trời **mưa**, anh ta sẽ **không vui**

Dựa vào những gì Bob mô tả, Alice muốn dự đoán thời tiết nơi Bob ở diễn ra như thế nào. Giả sử thời tiết chỉ có 2 loại là **nắng** hoặc **mưa** và Bob sẽ nói anh ta **vui** hay **không vui**.



Bài toán đơn giản



S



R







H







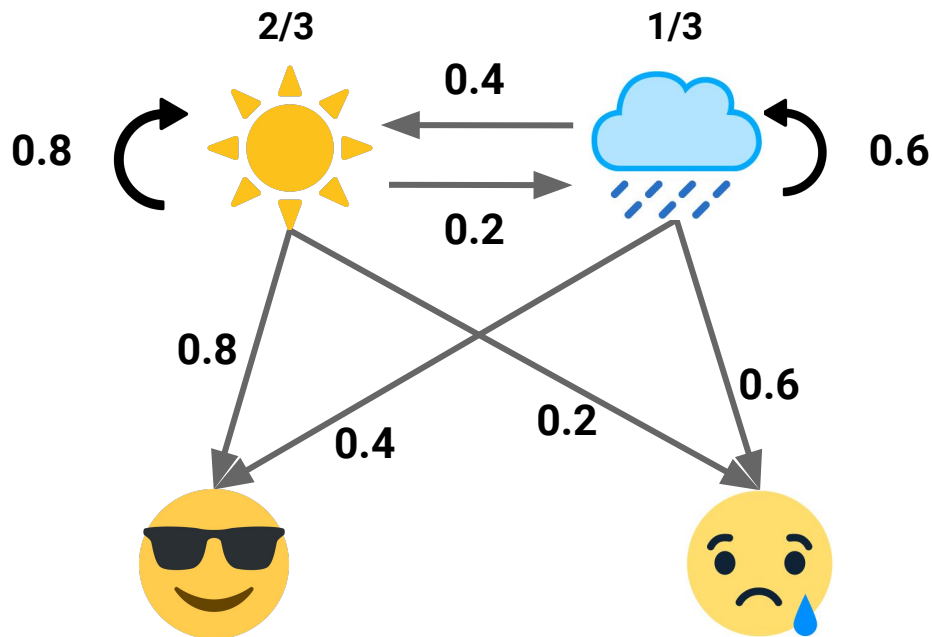
G

Input	Output
H	S
H G H G	?
H H G G G H	?

Hidden Markov Model

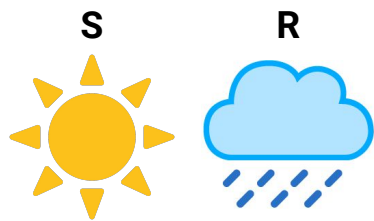
		
	0.8	0.2
	0.4	0.6

		
	0.8	0.2
	0.4	0.6



Hidden Markov Model

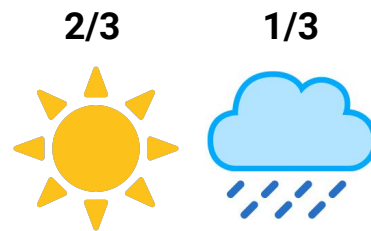
Hidden Markov Model





States = { S, R }







Observations = { H, G }

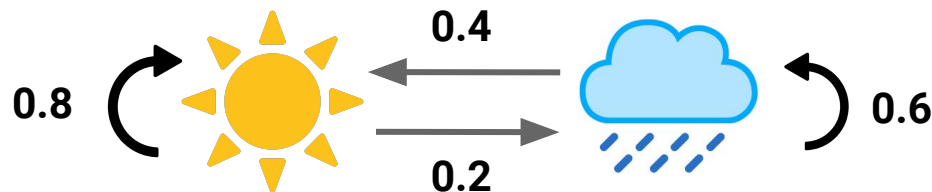


Start probabilities

	2/3
	1/3





Hidden Markov Model

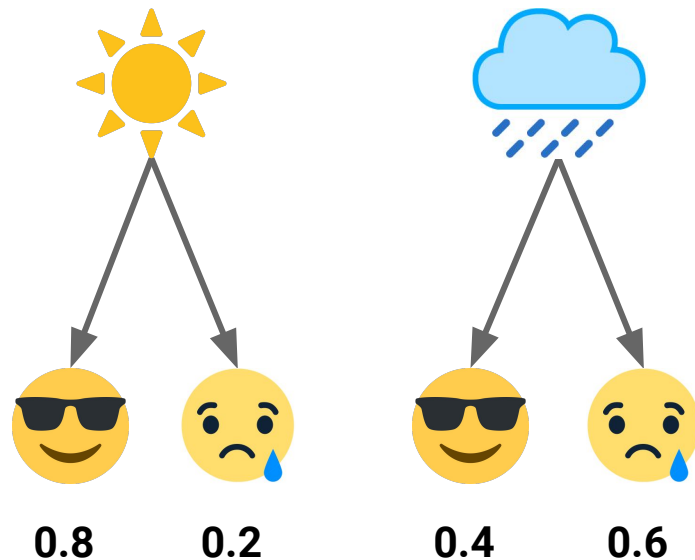
		
	0.8	0.2
	0.4	0.6



Transition Probabilities

Hidden Markov Model





		
	0.8	0.2
	0.4	0.6

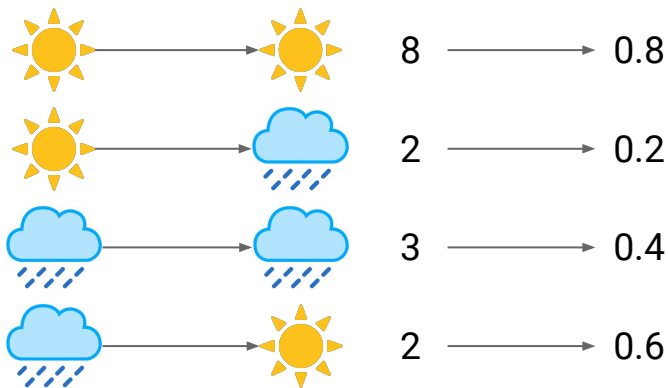


Emission Probabilities

Time Transition Probabilities & Emission Probabilities







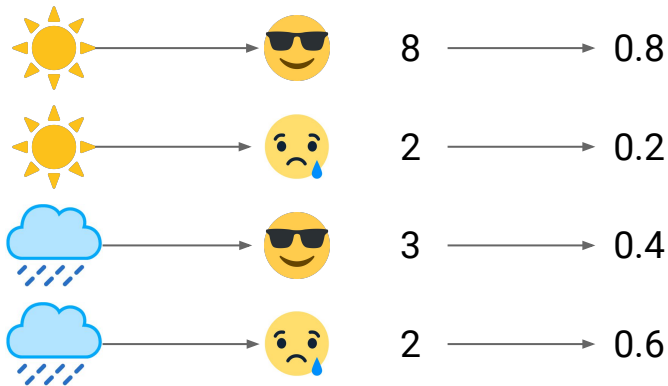
		
	0.8	0.2
	0.4	0.6



Time Transition Probabilities & Emission Probabilities



		
	0.8	0.2
	0.4	0.6





Naive Solution

Input:

Naive Solution



Input: _

	$2/3$
	$1/3$

Start Probabilities

Naive Solution





Input: __

	2/3
	1/3

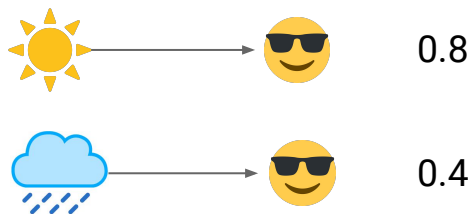
Output: S

Naive Solution

Input: H





		
	0.8	
	0.4	

Emission Probabilities





Naive Solution

Input: H

		
	0.8	
	0.4	

Emission Probabilities





	$\frac{2}{3}$
	$\frac{1}{3}$

Start Probabilities





Naive Solution

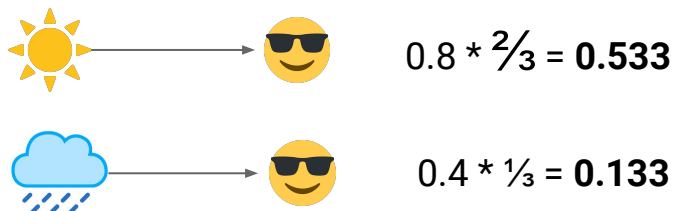
Input: H

		
	0.8	
	0.4	

Emission Probabilities

	$\frac{2}{3}$
	$\frac{1}{3}$





Start Probabilities







Output: S

Naive Solution

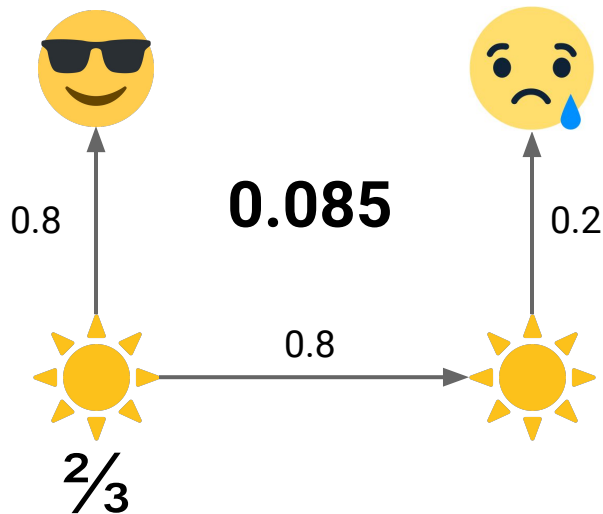
Input: H G

		
	0.8	0.2
	0.4	0.6

Emission Probabilities





		
	0.8	0.2
	0.4	0.6

Transition Probabilities







Naive Solution

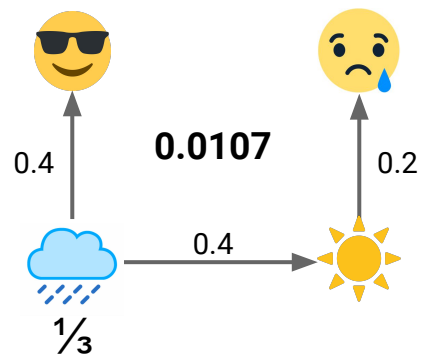
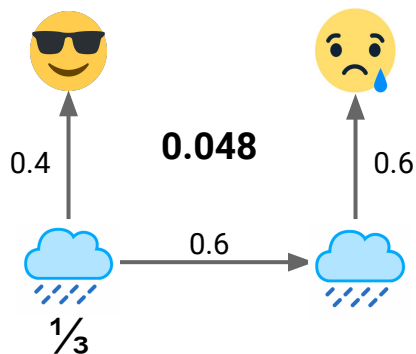
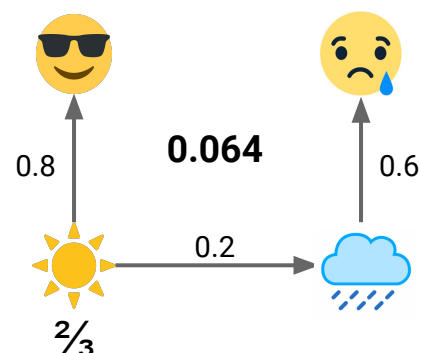
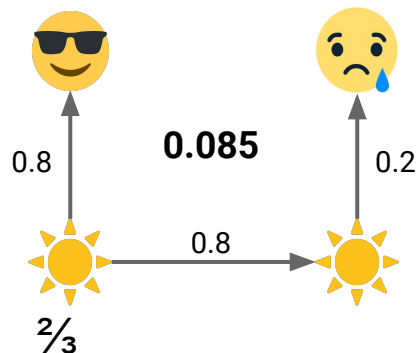
Input: H G

		
	0.8	0.2
	0.4	0.6

Emission Probabilities





		
	0.8	0.2
	0.4	0.6

Transition Probabilities







Naive Solution

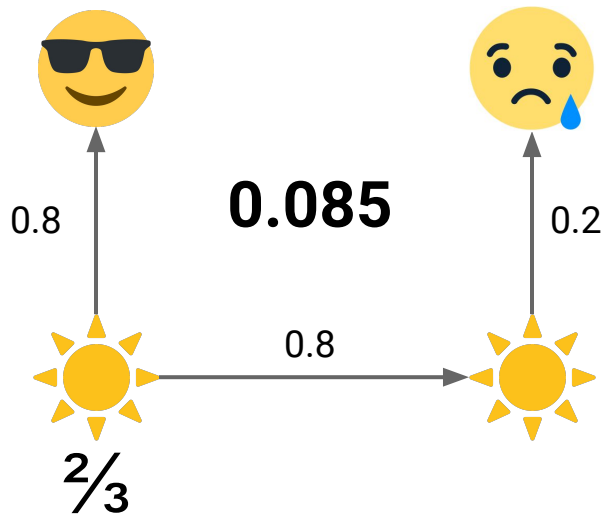
Input: H G

		
	0.8	0.2
	0.4	0.6

Emission Probabilities

		
	0.8	0.2
	0.4	0.6

Transition Probabilities



Output: S S

Naive Solution Complexity

- $M = |S|$
- $N = |\text{input}|$

$$\Rightarrow O(N * (N^M))$$



Phát biểu tổng quát của Hidden Markov Model

1. Markov chain.

Là mô hình giúp chúng ta mô tả, tính toán xác suất của một dãy các biến ngẫu nhiên (dãy các trạng thái). Mô hình này tính toán dựa trên giả định Markov, tức xác suất của bất cứ trạng thái nào chỉ dựa vào xác suất của một trạng thái liền trước nó.

Với q_i là trạng thái có được tại thời điểm i :

Markov Assumption: $P(q_i = a | q_1 \dots q_{i-1}) = P(q_i = a | q_{i-1})$

Phát biểu tổng quát của Hidden Markov Model

1. Markov chain.

Một mô hình markov chain được xác định bởi các tham số sau:

$$Q = [q_1, q_2, q_3 \dots q_n]$$

Tập các trạng thái có thể nhận được tại mỗi bước.







$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

Transition Probability Matrix: a_{ij} cho biết xác suất chuyển từ trạng thái thứ i sang trạng thái thứ j . ($P(q_j|q_i)$)

$$\pi = [\pi_1, \pi_2, \pi_3, \dots \pi_n]$$

Initial Probability Distribution: π_i cho biết xác suất chuỗi Markov sẽ bắt đầu bằng trạng thái thứ i .

Phát biểu tổng quát của Hidden Markov Model

 		
		
	0.8	0.2
	0.4	0.6
1	0	

Phát biểu tổng quát của Hidden Markov Model

2. Hidden Markov Model.

Trong một số trường hợp, ta không chỉ quan tâm tới xác suất xảy ra trạng thái nào đó tại một thời điểm i , mà còn quan tâm tới các xác suất của các trạng thái hệ quả (gọi là observation).

Một mô hình Hidden Markov model được xác định bởi các tham số sau:

$o = [o_1, o_2, o_3, \dots o_t]$ Tập các observation

$B_{n \times t}$

Emission probability matrix: b_{ij} cho xác suất observation thứ j xuất hiện khi trạng thái i xuất hiện ($P(o_j | q_i)$).

Phát biểu tổng quát của Hidden Markov Model

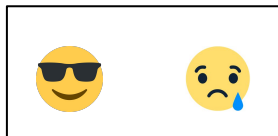
2. Hidden Markov Model.





Giả định markov đối với HMM.

Output Independence: $P(o_i | q_1 \dots q_i, \dots, q_T, o_1, \dots, o_i, \dots, o_T) = P(o_i | q_i)$

Phát biểu tổng quát của Hidden Markov Model

2. Hidden Markov Model.



		
	0.8	0.2
	0.4	0.6

Phát biểu tổng quát của Hidden Markov Model

3. Hai bài toán đặt ra trong Hidden Markov Model.

- a) Decoding : Cho biết tất cả các thông số của HMM. Và một dãy các observation, tìm dãy trạng thái tương ứng có xác suất cao nhất.
- b) Learning: Cho một tập dữ liệu gồm bộ hai dãy Observation và Trạng thái tương ứng. Tìm Model, tức tất cả các tham số.

Phát biểu tổng quát của Hidden Markov Model

4. Decoding - Viterbi Algorithm

Decoding: Given as input an HMM $\lambda = (A, B)$ and a sequence of observations $O = o_1, o_2, \dots, o_T$, find the most probable sequence of states $Q = q_1 q_2 q_3 \dots q_T$.

Dựa theo tư tưởng Quy Hoạch Động:

Gọi $v_t(j)$ là xác suất tốt nhất của những dãy các trạng thái có độ dài t và trạng thái cuối cùng là j , ta có:

$$v_t(j) = \max_{i=1}^N (v_{t-1}(i) * A_{ij} * b_{jt})$$

Phát biểu tổng quát của Hidden Markov Model

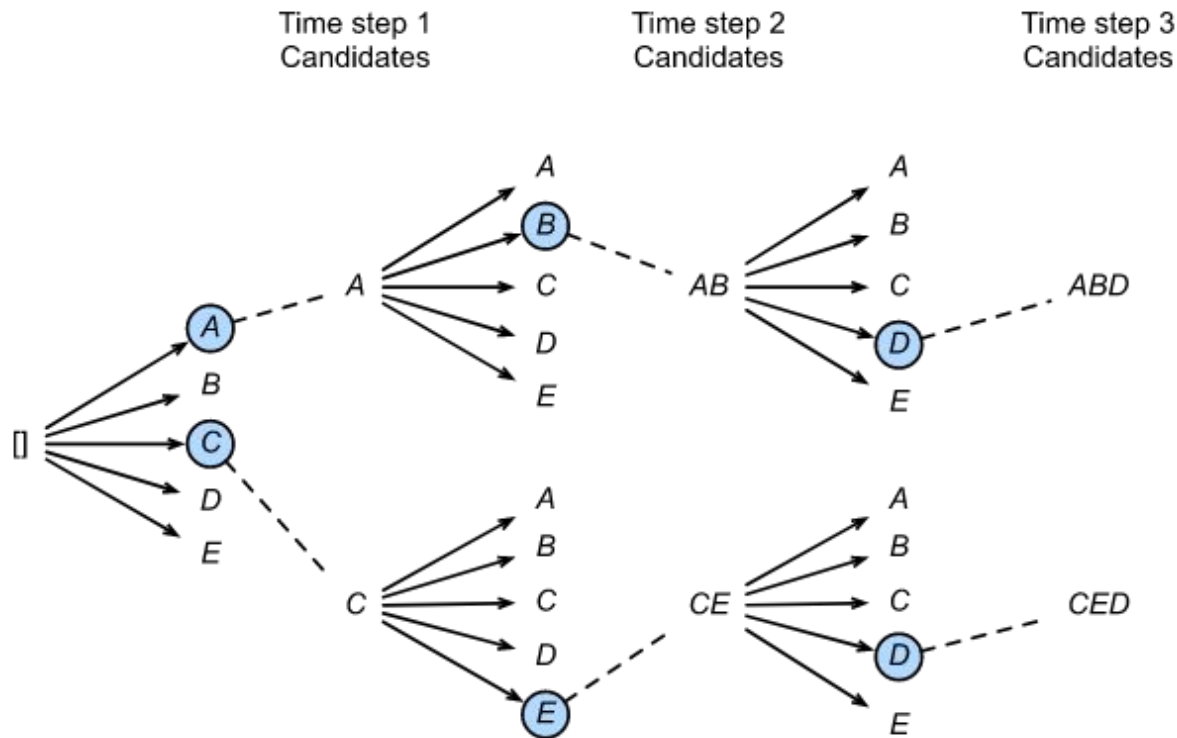
4. Decoding - Beam Search

Xuất phát từ thuật toán tham lam sau, tại mỗi thời điểm t , ta chọn trạng thái cho xác suất của dãy cao nhất.

- Gọi v_t là xác suất của dãy và w_t là trạng thái được chọn tại thời điểm t , ta có:

$$v_t = \max_{j=1}^N (v_{t-1} * A_{w_{t-1}j} * b_{jt})$$

Phát biểu tổng quát của Hidden Markov Model



Phát biểu tổng quát của Hidden Markov Model

5. Learning

Khi đã có đầy đủ bộ dữ liệu, mô hình HMM được xây dựng một cách đơn giản

Giả sử bộ dữ liệu của chúng ta có **M điểm**, mỗi điểm có dạng:

$$x^i = (s_1^i, o_1^i, s_2^i, o_2^i, \dots, s_t^i, o_t^i)$$

Phát biểu tổng quát của Hidden Markov Model

5. Learning

- **Q**(tập các trạng thái) và **o**(tập các observation) được suy ra dễ dàng bằng cách đếm các phần tử phân biệt trong bộ dữ liệu.
- π (phân phối xác suất khởi đầu), được tính như sau:

$$\pi(a) = \frac{\sum_{i=1}^m (s_1^i = a)}{m}$$

Phát biểu tổng quát của Hidden Markov Model

5. Learning

- Ma trận A (**Transition Probability Matrix**)

$$A_{pq} = \frac{\text{Số lần xuất hiện phép chuyển từ } p \rightarrow q}{\text{Số lần xuất hiện trạng thái } p} = \frac{\sum_{i=1}^m \sum_{j=1}^{t^i-1} (s_j^i = p \text{ and } s_{j+1}^i = q)}{\sum_{i=1}^m \sum_{j=1}^{t^i-1} (s_j^i = p)}$$

Phát biểu tổng quát của Hidden Markov Model

5. Learning

- Ma trận B (**Emission Probability Matrix**)

$$B_{qw} = \frac{\text{Số lần xuất hiện phép chuyển từ } q \rightarrow w}{\text{Số lần xuất hiện trạng thái } q} = \frac{\sum_{i=1}^m \sum_{j=1}^{t^i} (s_j^i = q \text{ and } o_j^i = w)}{\sum_{i=1}^m \sum_{j=1}^{t^i} (s_j^i = q)}$$

Ưu khuyết điểm

- Advantages
 - Effective
 - Can handle variations in record structure
 - Optional fields
 - Varying field ordering
- Disadvantages
 - Requires training using annotated data
 - Not completely automatic
 - May require manual markup
 - Size of training data may be an issue

Ứng dụng

- Nhận dạng giọng nói
- Xử lý văn bản (có tính đến chuỗi từ)
- Phân tích DNA
- Theo dõi nhịp tim
- Dự báo thị trường tài chính
- Xử lý cảm biến robot di động