Lecture 35
Evaluation of HMMs
ECEN 5283 Computer Vision

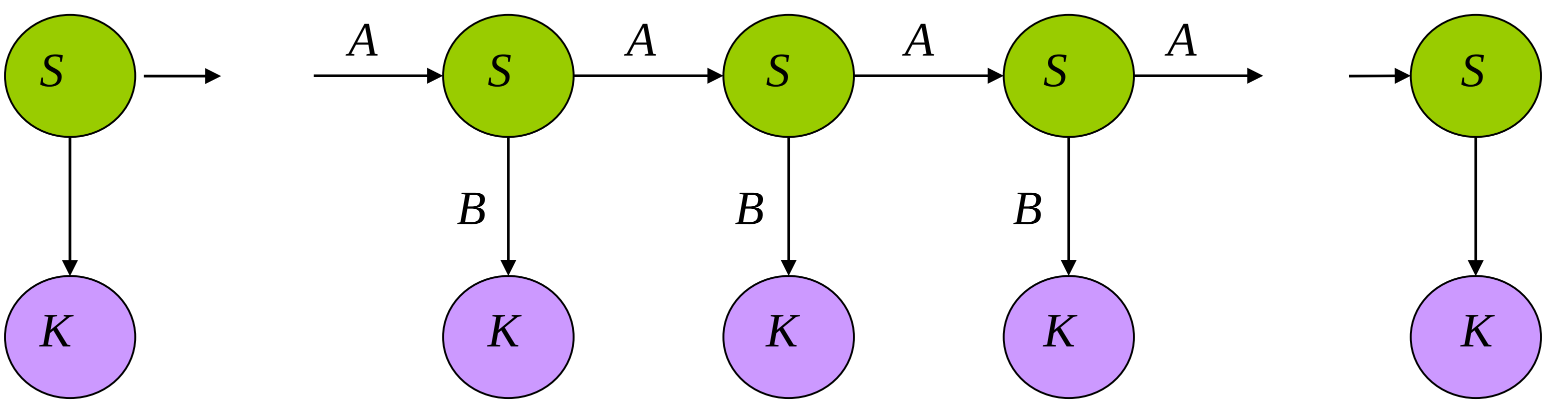
Dr. Guoliang Fan
School of Electrical and Computer Engineering
Oklahoma State University

Goals

- To review the basic issues of HMMs
- To find the probability of an observation sequence against a specific HMM

HMM Parameterization

- $\mu=\{S, K, P, A, B\}$
- $S : \{s_1 \ldots s_N\}$ are the values for the hidden states
- $K : \{k_1 \ldots k_M\}$ are the values for the observations
- $P = \{\pi_i\}$ are the initial state probabilities
- $A = \{a_{ij}\}$ are the state transition probabilities
- $B = \{b_{ik}\}$ are the observation state probabilities.
  - **B** can also be a set of PDFs related to different states. Then observation will be continuous variables.



Inferences in HMMs

- **Evaluation:** Compute the probability of a given observation sequence $\mathbf{O} = \{o_1, o_2, \ldots, o_T\}$ and a HMM μ

$$P(\mathbf{O} \mid \mu) = ?$$

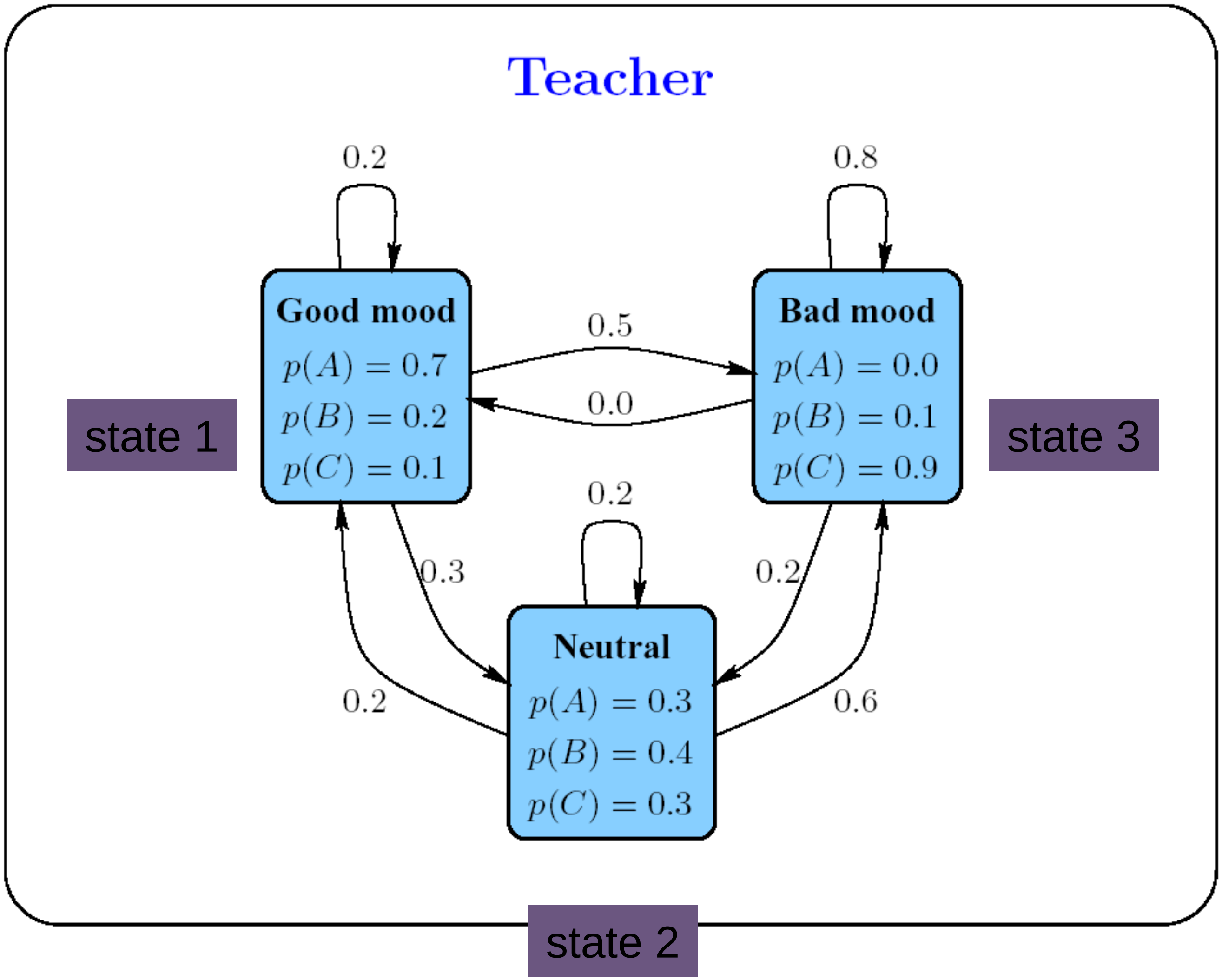
- **Decoding:** Given an observation sequence **O** and a HMM μ, compute the most likely hidden state sequence

$$X_{\{1,\ldots T\}} = \max_{x_1 \ldots x_T} P(X | \mathbf{O}, \mu)$$

- **Learning:** Given an observation sequence **O** and set of possible models, which model most closely fits the data?

$$\mu = \arg\max_{\mu} P(\mathbf{O} \mid \hat{\mu})$$

Teacher HMM: Parameterization



$$\pi_g = \pi_n = \pi_b = 1/3$$

$$\mathbf{A} = \begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0.2 & 0.2 & 0.6 \\ 0 & 0.2 & 0.8 \end{bmatrix}$$

$$b_{1,1} = b_1(A) = 0.7$$
$$b_{1,2} = b_1(B) = 0.2$$
$$b_{1,3} = b_1(C) = 0.1$$

$$b_{2,1} = b_2(A) = 0.3$$
$$b_{2,2} = b_2(B) = 0.4$$
$$b_{2,3} = b_2(C) = 0.3$$

$$b_{3,1} = b_3(A) = 0.0$$
$$b_{3,2} = b_3(B) = 0.1$$
$$b_{3,3} = b_3(C) = 0.9$$

Teacher HMM: Questions

Questions Teacher

One week, your teacher gave the following homework assignments:

Monday:	A
Tuesday:	C
Wednesday:	B
Thursday:	A
Friday:	C

Questions:

- What did his mood curve look like most likely that week?
- What is the probability that he would assign this order of homework assignments?
- What is the probability that he was in a good mood on Thursday?

Review of Probability Theorem

- Total Probability Theorem
  - Let $B_1, B_2, \ldots, B_n$ be a set of mutually *exclusive* and *exhaustive* events

$$\Pr(A) = \sum_{i=1}^{n} \Pr(A, B_i) = \sum_{i=1}^{n} \Pr(A \mid B_i) \Pr(B_i)$$

- Probability Chain Rule $\Pr(A, B) = \Pr(A \mid B) \Pr(B)$

$$\Pr(x_1, x_2, \ldots, x_M) = \Pr(x_M \mid x_1, \ldots, x_{M-1}) \ldots \Pr(x_3 \mid x_1, x_2) \Pr(x_2 \mid x_1) \Pr(x_1)$$

$$P(A, B \mid C) = P(A \mid B, C) P(B \mid C)$$

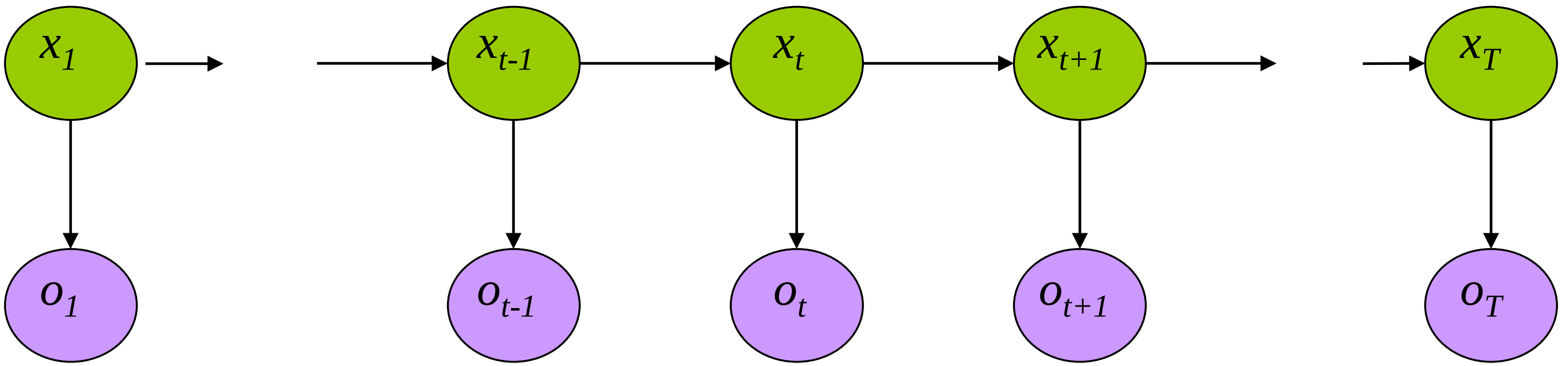
- Markov chain rule

$$\Pr(x_1, x_2, \ldots, x_M) = \Pr(x_M \mid x_{M-1}) \ldots \Pr(x_3 \mid x_2) \Pr(x_2 \mid x_1) \Pr(x_1)$$

$$= a_{x_{M-1}, x_M} \ldots a_{x_2, x_3} a_{x_1, x_2} \pi_{x_1}$$

Evaluation of HMMs

$$\Pr(A) = \sum_{i=1}^{n} \Pr(A \mid B_i)\Pr(B_i)$$



$$P(O \mid \mu) = \sum_{X} P(O, X \mid \mu) = \sum_{X} P(O \mid X, \mu) P(X \mid \mu)$$

$O = \{o_1, o_2, o_3, \ldots o_T\}$

All possible state sequences

$$P(O \mid X, \mu) = b_{x_1 o_1} b_{x_2 o_2} \ldots b_{x_T o_T}$$

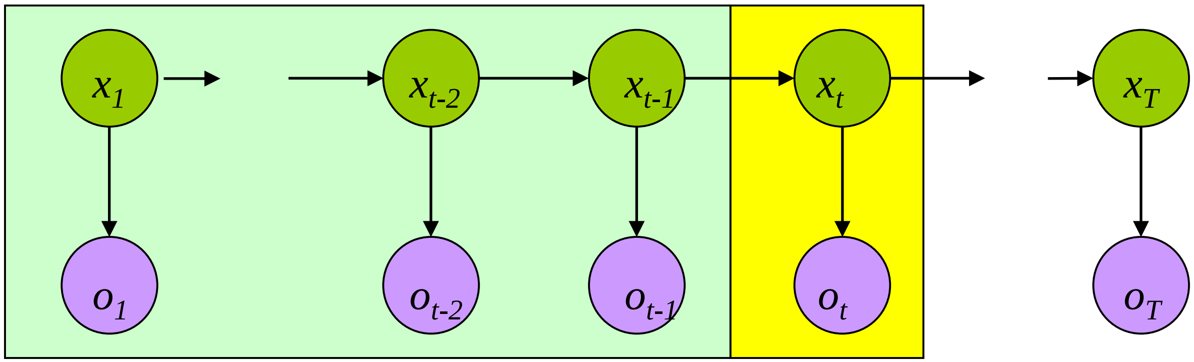
$$P(X \mid \mu) = \pi_{x_1} a_{x_1 x_2} a_{x_2 x_3} \ldots a_{x_{T-1} x_T}$$

$$P(O \mid \mu) = \sum_{X} \underline{\pi_{x_1} b_{x_1 o_1}} \prod_{t=1}^{T-1} \underline{a_{x_t x_{t+1}} b_{x_{t+1} o_{t+1}}}$$

N^T

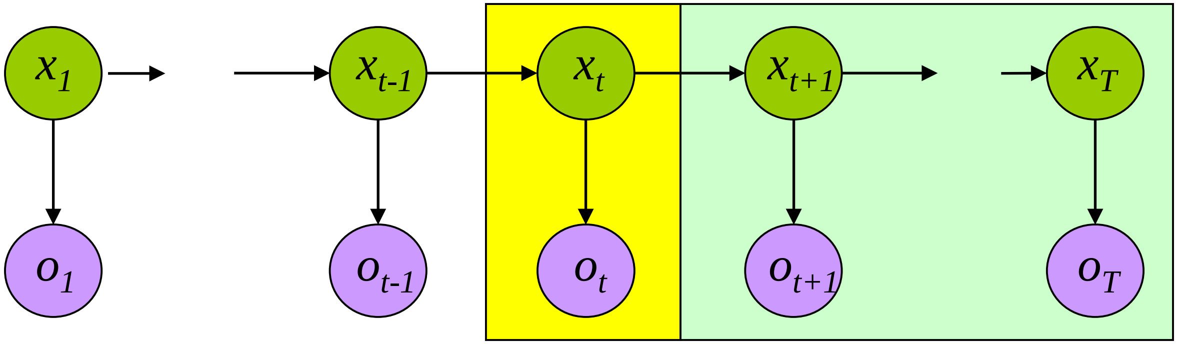
How many possible state sequences?

Forward and Backward Processes



Probability of the first part of observations and the current state at *t*

$$\alpha_t(i) = P(o_1 \ldots o_t, x_t = i)$$



Probability of the rest of the observations given the state at *t*

$$\beta_t(i) = P(o_{t+1} \ldots o_T \mid x_t = i)$$

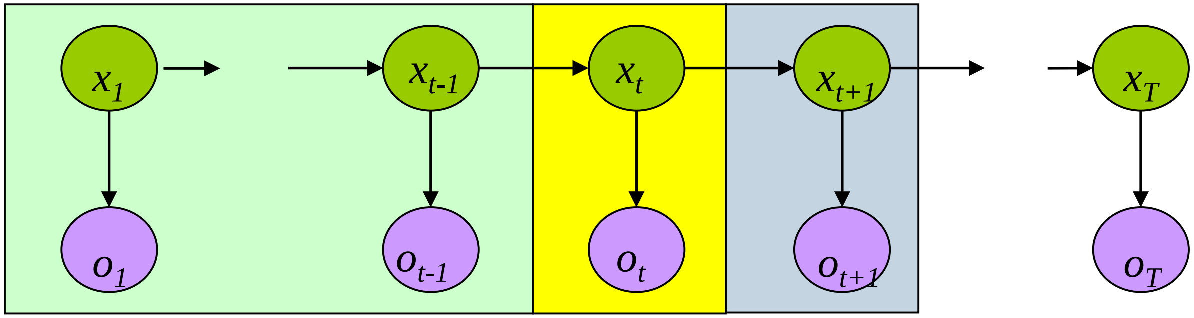
The key is to compute the two terms recursively.

$$P(A,B)P(C \mid B) = P(A,B)P(C \mid A,B) = P(A,B,C)$$

(Markovian property : Given B, A and C are independent.)

$$\alpha_t(i)\beta_t(i) = P(o_1 \ldots o_t, o_{t+1} \ldots o_T, x_t = i) \rightarrow P(O \mid \mu) = \sum_{i=1}^{N} \alpha_t(i)\beta_t(i)$$

Evaluation of HMMs: Forward Process



$$\alpha_t(i) = P(o_1 \ldots o_t, x_t = i)$$

$$\alpha_t(i) \rightarrow \alpha_{t+1}(j)$$

$$(i, j = 1, \ldots, N)$$

$$\alpha_1(i) = P(o_1, x_1 = i) = P(o_1 \mid x_1 = i) P(x_1 = i) = \pi_i b_{i,o_1}$$

$$\alpha_{t+1}(j) = P(o_1 \ldots o_{t+1}, x_{t+1} = j) = \sum_{i=1 \ldots N} P(o_1 \ldots o_{t+1}, x_t = i, x_{t+1} = j)$$

$$\Pr(A) = \sum_{i=1}^{n} \Pr(A \mid B_i) \Pr(B_i)$$

$$= \sum_{i=1,\ldots,N} P(o_1 \ldots o_t, , o_{t+1}, x_{t+1} = j | x_t = i) \, P(x_t = i)$$

$$= \sum_{i=1,\ldots,N} P(o_1 \ldots o_t, | x_t = i) \, P(o_{t+1}, x_{t+1} = j | x_t = i) P(x_t = i)$$

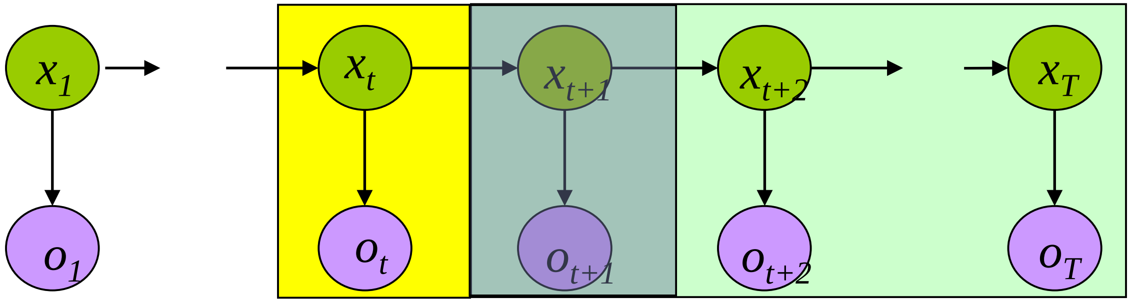
$$= \sum_{i=1,\ldots,N} P(o_1 \ldots o_t, x_t = i) \, P(o_{t+1}, x_{t+1} = j | x_t = i)$$

$$P(A, B | C) = P(A | B, C) P(B | C)$$

$$= \sum_{i=1,\ldots,N} P(o_1 \ldots o_t, x_t = i) \, P(o_{t+1} | \boxed{x_t = i,} x_{t+1} = j) P(x_{t+1} = j | x_t = i)$$

$$= \sum_{i=1,\ldots,N} \alpha_t(i) a_{i,j} b_{j o_{t+1}}$$

Evaluation of HMMs: Backward Process



$$\beta_t(i) = P(o_{t+1} \dots o_T \mid x_t = i)$$

$$\beta_t(i) \leftarrow \beta_{t+1}(j)$$

$$(i, j = 1, \dots, N)$$

$$\Pr(A) = \sum_{i=1}^{n} \Pr(A \mid B_i) \Pr(B_i)$$

$$\beta_T(i) = 1$$

$$\beta_t(i) = P(o_{t+1} \dots o_T \mid x_t = i) = \sum_{j=1,\dots,N} P(x_{t+1} = j, o_{t+1} \dots o_T \mid x_t = i)$$

$$P(A, B|C) = P(A|B, C)P(B|C)$$

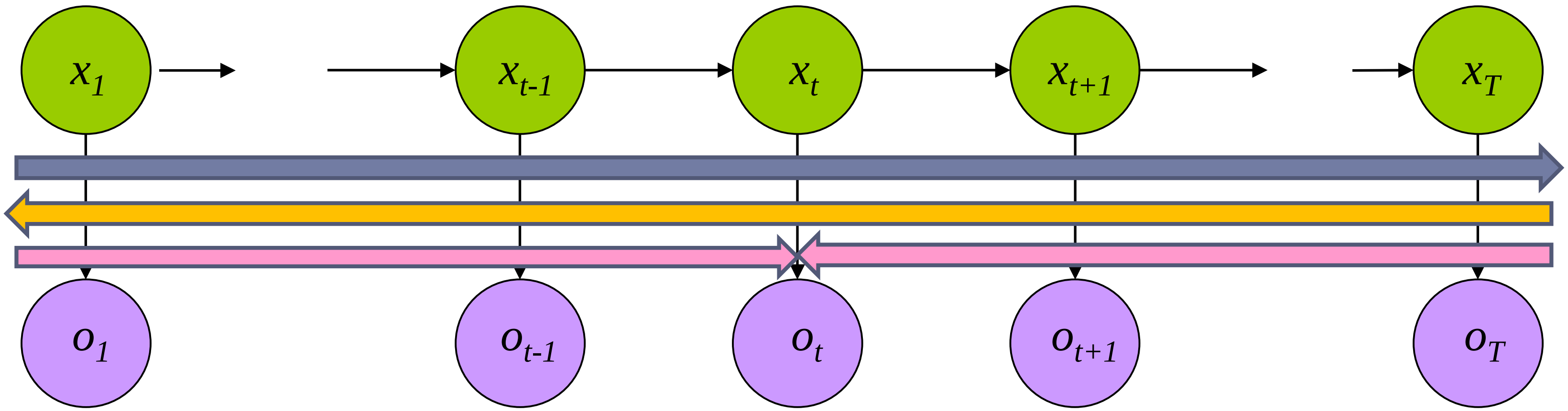
$$= \sum_{j=1}^{N} P(o_{t+1} o_{t+2} \dots o_T \mid x_t = i, x_{t+1} = j)\, P(x_{t+1} = j \mid x_t = i)$$

$$= \sum_{j=1}^{N} P(o_{t+1} o_{t+2} \dots o_T \mid x_{t+1} = j)\, P(x_{t+1} = j \mid x_t = i)$$

$$= \sum_{j=1}^{N} P(o_{t+2} \dots o_T \mid x_{t+1} = j) P(o_{t+1} \mid x_{t+1} = j)\, P(x_{t+1} = j \mid x_t = i)$$

$$= \sum_{j=1 \dots N} a_{ij} b_{j o_{t+1}} \beta_{t+1}(j)$$

Evaluation of HMMs: Forward and Backward Process



$\alpha_1(i) = \pi_i b_{i,o_1}$

$$P(O \mid \mu) = \sum_{i=1}^{N} \alpha_T(i)$$ **Forward Procedure**

$\alpha_t(i) \to \alpha_{t+1}(j)$
$(i, j = 1, \ldots, N)$

$\beta_t(i) \leftarrow \beta_{t+1}(j)$
$(i, j = 1, \ldots, N)$

$$P(O \mid \mu) = \sum_{i=1}^{N} \pi_i b_{i,o_1} \beta_1(i)$$ **Backward Procedure**

$\beta_T(i) = 1$

$\alpha_t(i) \to \alpha_{t+1}(j)$
$(i, j = 1, \ldots, N)$

$$P(O \mid \mu) = \sum_{i=1}^{N} \alpha_t(i) \beta_t(i)$$ **Combination (at *t*)**

$\beta_t(i) \leftarrow \beta_{t+1}(j)$
$(i, j = 1, \ldots, N)$

Why do we need both forward and backward processes for HMM evaluation?

Teacher HMM: Forward Algorithm

Forward Algorithm / Teacher

- Initialization : $\alpha_1(i) = \pi_i b_i(A)$

	A	C	B	A	C
good	0.23				
neutral	0.1				
bad	0.0				

$$\pi_g = \pi_n = \pi_b = 1/3$$

$$A = \begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0.2 & 0.2 & 0.6 \\ 0 & 0.2 & 0.8 \end{bmatrix}$$

$$b_1(A) = 0.7$$
$$b_1(B) = 0.2$$
$$b_1(C) = 0.1$$

$$b_2(A) = 0.3$$
$$b_2(B) = 0.4$$
$$b_2(C) = 0.3$$

$$b_3(A) = 0.0$$
$$b_3(B) = 0.1$$
$$b_3(C) = 0.9$$