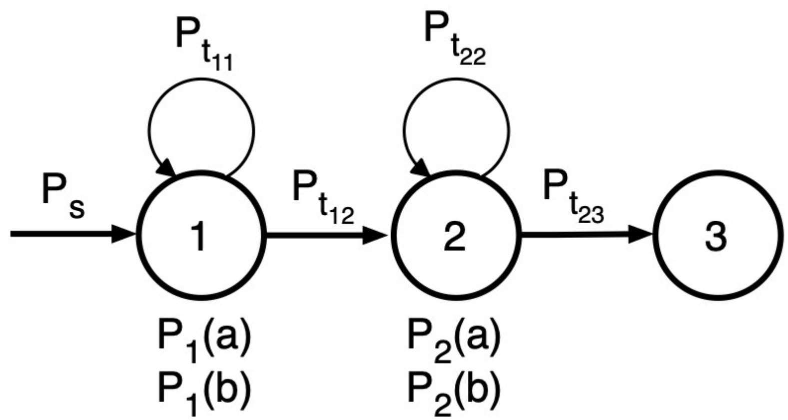
Hidden Markov Model 2

learning HMM parameters

- when we know the state path for each training sequence, learning the model parameters is simple
  - estimate the parameters by counting
  - normalize to get probability

- when we do not know the path for each training sequence, how can we determine the counts
  - estimate the counts by considering every path weight by its probability

Review: Hidden Markov model



$M = (S, A, P_e, P_{tr}, P_s)$

1. States $S = \{S_i | 1 \leq i \leq N\}$.
2. Symbols $A = \{A_j | 1 \leq j \leq M\}$.
3. Emission probability distribution $P_e = \{E_{ij} | P(O_t = A_j | Q_t = S_i), 1 \leq i \leq N, 1 \leq j \leq M\}$, which is the probabilities of emitting A_j in a state S_i at any time t.
4. Transition probability distribution, $P_{tr} = \{T_{ik} | P(Q_t = S_k | Q_{t-1} = S_i), 1 \leq i \leq N, 1 \leq k \leq N\}$, which is the probabilities of transiting from hidden state S_i to S_k.
5. Initial state probability distribution, $P_s = \{\pi_i | P(Q_1 = S_i), 1 \leq i \leq N\}$, which is the probabilities that the initial state is the state S_i.

Review: Hidden Markov model

Forward algorithm

1) Initialization

$$\alpha_1(i) = \pi_i b_i(O_1),\ 1 \leq i \leq N$$

2) Induction

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1})$$

$$1 \leq t \leq T-1$$
$$1 \leq j \leq N$$

Perform for all states for given t, then advance t.

3) Termination

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

Hidden Markov model

Backward algorithm

1) **Initialization** (arbitrarily define $\beta_T(i)$ to be 1 for all i)

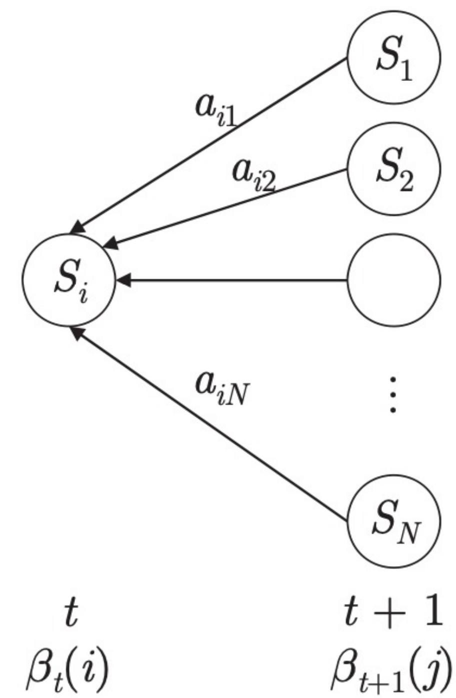
$$\beta_T(i) = 1,\ 1 \leq i \leq N$$

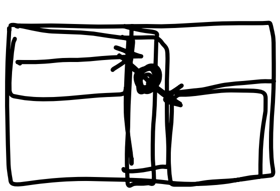
2) **Induction**

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)$$

$$t = T-1, T-2, \ldots, 1$$

$$1 \leq i \leq N$$





learning HMM parameters

Supervised

X = THTHHHTHTTH

S = F F F F B B B B B B B

Unsupervised

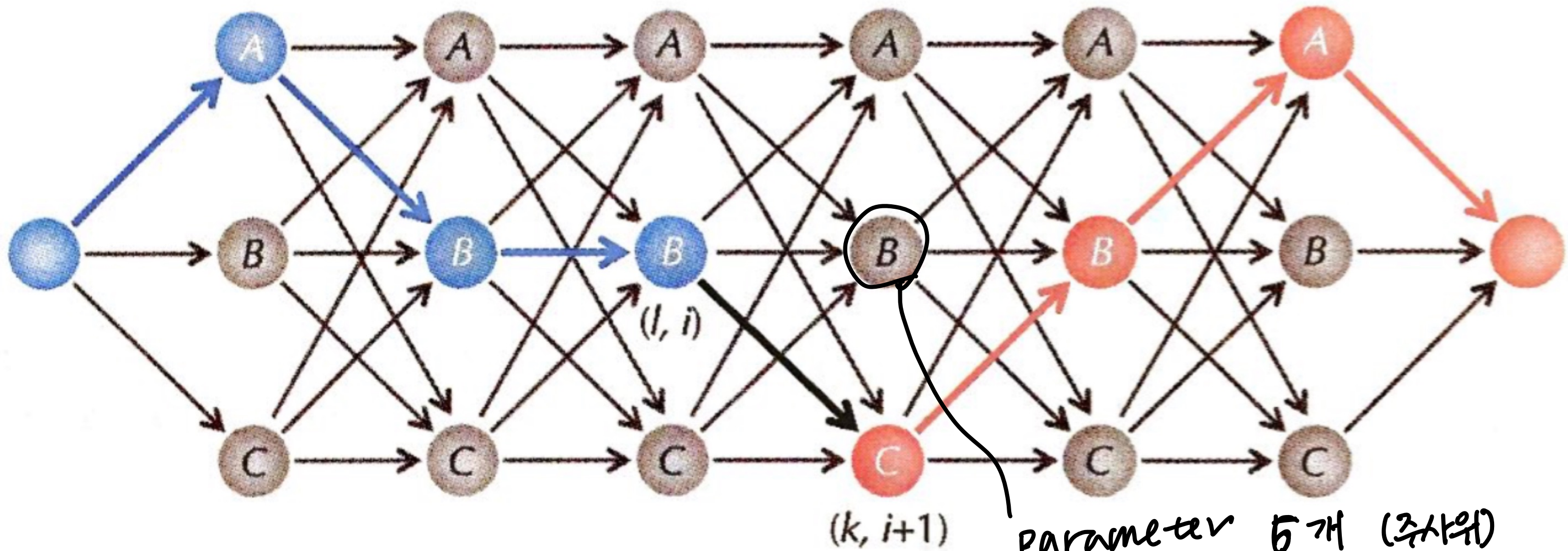
M step

$b_B(H)$ = ? 어떻게 계산?

B state에서 H symbol 나올 확률

X = THTHHHTHTTH

learning HMM parameters by soft decision



$E_B(3)$

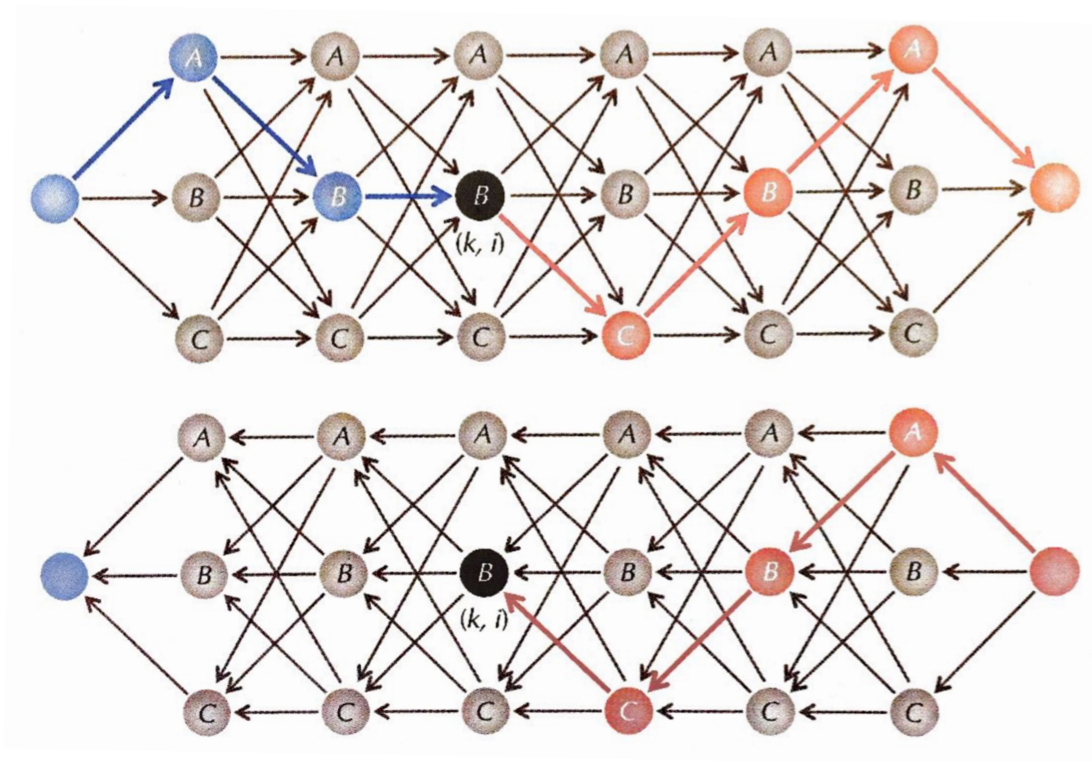
most probable state path가
B를 지난다

$E_B(1)$ $E_B(3)$
$E_B(2)$

어떤 state 지나는지 알려져 있는지
(label)
찾아서
'most probable path 찾기'

M
모델 파라미터 구하기
E

learning HMM parameters by soft decision



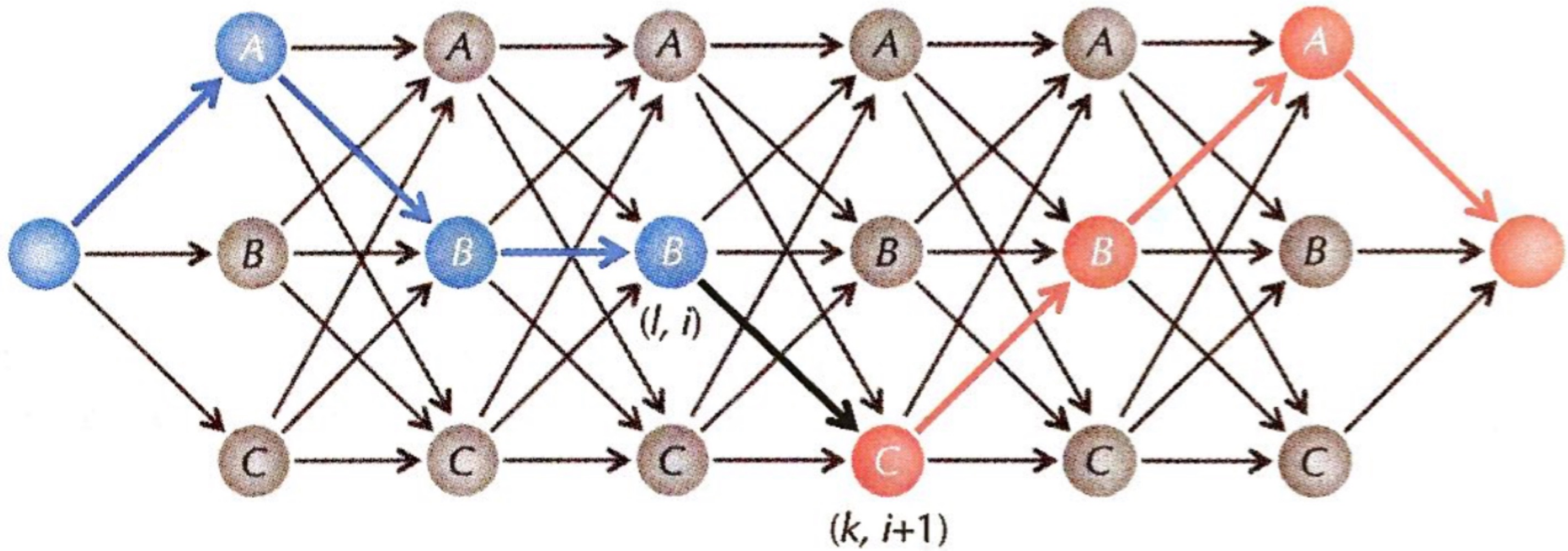
$E_k^i(b)$

$$Pr(\pi_i = k, x) = \sum_{all\ paths\ \pi\ with\ \pi_i = k} Pr(x, \pi)$$

$$= \sum_{all\ paths\ \pi\ _{before}} Pr(\pi_{before}) \sum_{all\ paths\ \pi\ _{after}} Pr(\pi_{after})$$

forward $_{k,i}$ *backward* $_{k,i}$

learning HMM parameters by soft decision



$E_k^i(b)$

		T	H	T	H	H	H	T	H	T	T	H
K=F :	***F***	0.636	0.593	0.600	0.533	0.515	0.544	0.627	0.633	0.692	0.686	0.609
K=B :	***B***	0.364	0.407	0.400	0.467	0.485	0.456	0.373	0.367	0.308	0.314	0.391

$E_B(H) = ?$ = 앞부분 모든 path와 뒷부분 모든 path를 더한다

= B state에서 H가 emit 될 확률

learning HMM parameters without hidden state

$$E_k^i(b) = \begin{cases} 1 \; if \;\; \pi_i = k \; and \; x_i = b \\ 0 \; otherwise \end{cases}$$

$$E_k(b) = \sum_{i=1}^{n} E_k^i(b)$$

for hard decision

$$E_k^i(b) = \begin{cases} Pr(\pi_i = k|x) \; if \;\; x_i = b \\ 0 \; otherwise \end{cases}$$

$$E_k(b) = \sum_{i=1}^{n} E_k^i(b)$$

for soft decision

$$emission_k(b) = \frac{E_k(b)}{\sum_{all\ symbols\ c\ in\ the\ alphabet} E_k(c)}$$

learning HMM parameters without hidden state

E-step $\longrightarrow$ M-step

각 포지션마다 레이블 결정

→ MLE

Estimate T and E for each position

(and find optimal hidden path)

$$transition_{l.k} = \frac{T_{l,k}}{\sum_{all\ states\ j} T_{l,j}}$$

$$emission_k(b) = \frac{E_k(b)}{\sum_{all\ symbols\ c\ in\ the\ alphabet} E_k(c)}$$

learning HMM parameters without hidden state

- Forward-Backward algorithm
- Expectation Maximization algorithm
  - the hidden state is the path that best explains each training sequence
  - E-step: calculate the expected transition or emission is used
  - M-step: estimate the parameters to maximize the likelihood of these expected values

Baum-Welch 알고리즘
→ hidden state가 있는 시퀀스로
HMM의 parameter를 estimate

hidden state를 찾고 MLE로 모델 파라미터를 찾아내는 과정

HMM: supervised

– label 없으면 EM 또는 self-supervised
새로운 task 만듦

10문제

- OX
- 설명하는 문제
- 계산하는 문제 - 공식 암기 X 방법 이해

- HMM
- Markov Chain
- EM (k-means)
- Hierarchical clustering
- global / local alignment, DP
- Motif finding 문제 접근 방식