1. (i) Transmission parameter $a_{u,v}$ and Emission parameter $b_u(v)$ are the parameters associated with the HMM.

(ii) Transmission parameter $a_{u,v}$

u \ v	x	y	z	STOP
START	0.4	0	0.6	0
x	0	0.5	0.333	0.167
y	0.167	0	0.167	0.667
z	0.5	0.5	0	0

Emission parameter $b_u(o)$

u \ o	a	b	c
x	0.167	0.5	0.333
y	0.333	0	0.667
z	0.167	0.333	0.5

$$a_{START,x} = \frac{count(START;x)}{count(START)} = \frac{2}{5} = 0.4$$

$$a_{START,y} = \frac{count(START;y)}{count(START)} = \frac{0}{5} = 0$$

$$a_{START,z} = \frac{count(START;z)}{count(START)} = \frac{3}{5} = 0.6$$

$$a_{START,STOP} = \frac{count(START;STOP)}{count(START)} = \frac{0}{5} = 0$$

$$a_{x,x} = \frac{count(x;x)}{count(x)} = \frac{0}{6} = 0$$

$$a_{x,y} = \frac{count(x;y)}{count(x)} = \frac{3}{6} = 0.5$$

$$a_{x,z} = \frac{count(x;z)}{count(x)} = \frac{2}{6} = 0.333$$

$$a_{x,STOP} = \frac{count(x;STOP)}{count(x)} = \frac{1}{6} = 0.167$$

$$a_{y,x} = \frac{count(y;x)}{count(y)} = \frac{1}{6} = 0.167$$

$$a_{y,y} = \frac{count(y;y)}{count(y)} = \frac{0}{6} = 0$$

$$a_{y,z} = \frac{count(y;z)}{count(y)} = \frac{1}{6} = 0.167$$

$$a_{y,stop} = \frac{count(y;STOP)}{count(y)} = \frac{4}{6} = 0.667$$

$$a_{z,x} = \frac{count(z;x)}{count(z)} = \frac{3}{6} = 0.5$$

$$a_{z,y} = \frac{count(z;y)}{count(z)} = \frac{3}{6} = 0.5$$

$$a_{z,z} = \frac{count(z;z)}{count(z)} = \frac{0}{6} = 0, \quad a_{z,stop} = \frac{count(z;STOP)}{count(z)} = \frac{0}{6} = 0$$

$$b_x(a) = \frac{count(x \to a)}{count(x)} = \frac{1}{6} = 0.167$$

$$b_x(b) = \frac{count(x \to b)}{count(x)} = \frac{3}{6} = 0.5$$

$$b_x(c) = \frac{count(x \to c)}{count(x)} = \frac{2}{6} = 0.333$$

$$b_y(a) = \frac{count(y \to a)}{count(y)} = \frac{2}{6} = 0.333$$

$$b_y(b) = \frac{count(y \to b)}{count(y)} = \frac{0}{6} = 0$$

$$b_y(c) = \frac{count(y \to c)}{count(y)} = \frac{4}{6} = 0.667$$

$$b_z(a) = \frac{count(z \to a)}{count(z)} = \frac{1}{6} = 0.167$$

$$b_z(b) = \frac{count(z \to b)}{count(z)} = \frac{2}{6} = 0.333$$

$$b_z(c) = \frac{count(z \to c)}{count(z)} = \frac{3}{6} = 0.5$$

2.

	"b"	"b"	
	x	x	
START	y	y	STOP
	z	z	

$\pi(0, START) = 1$

$\pi(1, x) = \max\{\pi(0, START) \times 0.4 \times 0.5\} = 0.2$

$\pi(1, y) = \max\{\pi(0, START) \times 0 \times 0\} = 0$

$\pi(1, z) = \max\{\pi(0, START) \times 0.6 \times 0.333\} = 0.1998$

$\pi(2, x) = \max\{\pi(1, x) \times b_x(\text{"b"}) \times a_{x,x}, \pi(1, y) \times b_x(\text{"b"}) \times a_{y,x},$
$\pi(1, z) \times b_x(\text{"b"}) \times a_{z,x}\}$

$= \max\{0.2 \times 0.5 \times 0, 0, 0.1998 \times 0.5 \times 0.5\} = 0.1$

$\pi(2, y) = \max\{\pi(1, x) \times b_y(\text{"b"}) \times a_{x,y}, \pi(1, y) \times b_y(\text{"b"}) \times a_{y,y},$
$\pi(1, z) \times b_y(\text{"b"}) \times a_{z,y}\} = 0$

$\pi(2, z) = \max\{\pi(1, x) \times b_z(\text{"b"}) \times a_{x,z}, \pi(1, y) \times b_z(\text{"b"}) \times a_{y,z},$
$\pi(1, z) \times b_z(\text{"b"}) \times a_{z,z}\}$

$= \max\{0.2 \times 0.333 \times 0.333, 0, 0.1998 \times 0.333 \times 0\} = 0.0222$

$\therefore y_2^* = \arg\max\{\pi(2, x) \times a_{x,stop}, \pi(2, y) \times a_{y,stop}, \pi(2, z) \times a_{z,stop}\}$

$= \arg\max\{0.1 \times 0.167, 0, 0.0222 \times 0\} = x$

$y_1^* = \arg\max\{\pi(1, x) \times a_{x,x}, \pi(1, y) \times a_{y,x}, \pi(1, z) \times a_{z,x}\}$

$= \arg\max\{0.2 \times 0, 0, 0.1998 \times 0.5\} = z$

$\Rightarrow$ The most probable state sequence $= (z/\text{"b"}, x/\text{"b"})$

3. $y^* = \arg\max_{y_1 \dots y_n} p(x_1, \dots, x_n, y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n \mid y_i \neq v)$

Base case:

$$\pi(0,v) = \begin{cases} 1 & \text{if } v = \text{START} \\ 0 & \text{otherwise} \end{cases}$$

Then:

For $j = 0, \dots, i-2$ for each $u \in T$,

$$\pi(j+1, u) = \max_v \{\pi(j, v) \times b_u(x_{j+1}) \times a_{v,u}\}$$

For $j = i-1$, for each $u \in T$ and $u \neq v$,

$$\pi(j+1, u) = \max_v \{\pi(j, v) \times b_u(x_{j+1}) \times a_{v,u}\}$$
$$= \max_v \{\pi(i-1, v) \times b_u(x_i) \times a_{v,u}\}$$

For $j = i, \dots, n-1$, for each $u \in T$,

$$\pi(j+1, u) = \max_v \{\pi(j, v) \times b_u(x_{j+1}) \times a_{v,u}\}$$

Final step:

$$\pi(n+1, \text{STOP}) = \max_v \{\pi(n, v) \times a_{v,\text{STOP}}\}$$

Followed by backward tracking:

$y_n^* = \arg\max_u \{\pi(n, u) \cdot a_{u,\text{stop}}\}$

For $j = n-1, \dots, 0$

$$y_j^* = \arg\max_u \{\pi(j, u) \cdot a_{u, y_{j+1}^*}\}$$

4. ① Estimate the most probable state sequence

- Initialization: $\pi(0,u) = \begin{cases} 1 & \text{if } u = \text{START} \\ 0 & \text{otherwise} \end{cases}$

- then $\pi(j+1, u) = \max_u \{\pi(j,v) \times b_u((x_{j+1}, y_{j+1})) \times a_{v,u}\}$

  for $j = 0, \ldots, n-1$ for each $u \in T$.

- Final step: $\pi(n+1, \text{STOP}) = \max_v \{\pi(n,v) \times a_{v,\text{stop}}\}$
- Now every node stores a score of the highest scoring path from START the that node (j,u).
- Back track the score:

  $y^*_n = \arg\max_u \{\pi(n,u) \cdot a_{u,\text{stop}}\}$

  $y^*_j = \arg\max_u \{\pi(j,u) \cdot a_{u,y^*_{j+1}}\}$ for $j = n-1, \ldots, 0$

② $P(z_j = u, x, y; \theta) = P(x_1, x_2, \ldots x_n, y_1 \ldots y_n, z_j = u; \theta)$

$= P(x_1, \ldots x_{j-1}, y_1, \ldots, y_{j-1}, z_j = u; \theta) \cdot P(x_j, \ldots, x_n, y_j, \ldots y_n \mid x_1, \ldots, x_{j-1}, y_1, \ldots, y_{j-1}, z_j = u; \theta)$

$= P(x_1, \ldots, x_{j-1}, y_1, \ldots y_{j-1}, z_j = u; \theta) \cdot P(x_j \ldots x_n, y_j \ldots y_n \mid z_j = u; \theta)$

$= P((x_1, y_1), \ldots, (x_{j-1}, y_{j-1}), z_j = u; \theta) \cdot P((x_j, y_j) \ldots (x_n, y_n) \mid z_j = u; \theta)$

let the forward probability $\alpha_u(j) = P((x_1, y_1), \ldots, (x_{j-1}, y_{j-1}), z_j = u; \theta)$

let the backward probability $\beta_u(j) = P((x_j, y_j) \ldots (x_n, y_n) \mid z_j = u; \theta)$

∴ Forward Algorithm:

- Base case: $\alpha_u(1) = a_{\text{START},u}$
- Recursive case: $\alpha_u(j+1) = \sum_v \alpha_v(j)\, a_{v,u}\, b_v((x_j, y_j))$

∴ Backward Algorithm:

- Base case: $\beta_u(n) = a_{u,\text{stop}}\, b_u((x_n, y_n))$
- Recursive case: $\beta_u(j) = \sum_v a_{u,v}\, b_u((x_j, y_j))\, \beta_v(j+1)$

③ Time complexity = $O(nT^2)$ for both forward and backward algorithm.
For each node in one layer, it has to visit all nodes in next layer.
$\Rightarrow |T^2|$

There are n layers in total.
$\Rightarrow O(nT^2)$

Wang Zijia
1002885