

10-601B HMMs / CRFs [11/21/16]

[Forward - Backward]

Define : $\alpha_k(t) \triangleq p(x_1, \dots, x_k, y_k = t) = \sum_{y_1, \dots, y_{k-1}} p(x_1, \dots, x_k, y_1, \dots, y_{k-1}, y_k = t)$

$\beta_k(t) \triangleq p(x_1, \dots, x_k | y_k = t)$

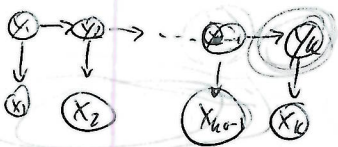
Base Case : $\alpha_0(\text{START}) = 1$ $y_0 = \text{START}$
 $\beta_{K+1}(\text{END}) = 1$ $y_{K+1} = \text{END}$

Recursion : $\alpha_k(t) = \sum_s p(x_k | y_k = t) \alpha_{k-1}(s) p(y_k = t | y_{k-1} = s)$

$\beta_k(t) = \sum_s \beta_{k+1}(s) p(x_{k+1} | y_{k+1} = s) p(y_{k+1} = s | y_k = t)$

Derivation of Forward Alg. $\alpha_{k-1}(t) \triangleq p(x_1, \dots, x_{k-1}, y_{k-1} = t)$

$$\begin{aligned}
 \alpha_k(t) &\triangleq p(x_1, \dots, x_k, y_k = t) \\
 &= p(x_1, \dots, x_k | y_k = t) p(y_k = t) \\
 &= p(x_k | y_k = t) p(x_1, \dots, x_{k-1} | y_k = t) p(y_k = t) \\
 &= p(x_k | y_k = t) p(x_1, \dots, x_{k-1}, y_k = t) \\
 &= p(x_k | y_k = t) \sum_{y_{k-1}} p(x_1, \dots, x_{k-1}, y_{k-1}, y_k = t) \\
 &= p(x_k | y_k = t) \sum_{y_{k-1}} p(x_1, \dots, x_{k-1}, y_{k-1} | y_k = t) p(y_{k-1} | y_k = t) \\
 &= p(x_k | y_k = t) \sum_{y_{k-1}} p(x_1, \dots, x_{k-1} | y_{k-1}) p(y_k | y_{k-1}) p(y_{k-1}) \\
 &= p(x_k | y_k = t) \sum_{y_{k-1}} p(x_1, \dots, x_{k-1} | y_{k-1}) p(y_k | y_{k-1}) \\
 &= p(x_k | y_k = t) \sum_{y_{k-1}} \alpha_{k-1}(s) p(y_k = t | y_{k-1} = s)
 \end{aligned}$$



11/16

Algorithm

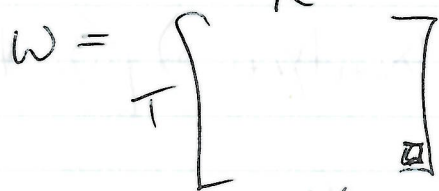
$$w_k(t) \triangleq \max_{y_1, \dots, y_{k-1}} p(x_1, \dots, x_k, y_1, \dots, y_{k-1} | y_k = t)$$

$$w_0(\text{START}) = 1$$

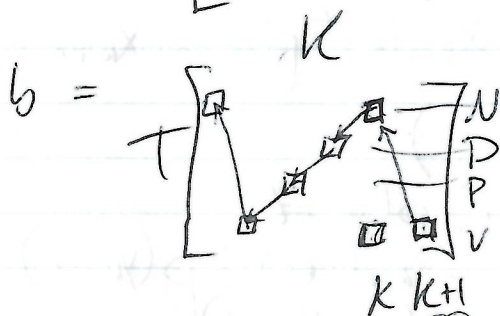
Recurse: $w_k(t) = \max_s p(x_k | y_k = t) w_{k-1}(s) p(y_k = t | y_{k-1} = s)$

$$b_k(t) = \operatorname{argmax}_s p(x_k | y_k = t) w_{k-1}(s) p(y_k = t | y_{k-1} = s)$$

↑
K backpointer



$w_{k+1}(\text{END})$



START

N

time

V

flies

P

like

P

an

N

arrow

END

V

$$\hat{y} = \operatorname{argmax}_{y_1, \dots, y_{k-1}} p(x_1, \dots, x_k, y_1, \dots, y_{k-1}, y_k = v)$$

Log. Reg. class-specific parameters
features

$$p_{\vec{\theta}}(y|\vec{x}) = \frac{\exp(\vec{\theta}_y \cdot \vec{x})}{Z(\vec{x})}$$

$$\vec{\theta} = \begin{bmatrix} \vec{\theta}_1 \\ \vec{\theta}_2 \\ \vdots \\ \vec{\theta}_T \end{bmatrix} = \frac{\exp(\vec{\theta} \cdot \vec{f}(\vec{x}, y))}{Z(\vec{x})}$$

$$\vec{f} = \begin{bmatrix} \vec{\theta} \\ x_1 \\ \vdots \\ 1 \end{bmatrix}_{y=2}$$

CRF:

$$p(\vec{y}|\vec{x}) = \frac{1}{Z(\vec{x})} \cdot \exp(\vec{\theta} \cdot \sum_{k=1}^K \vec{f}_k)$$

$$= \frac{\exp(\vec{\theta} \cdot (\sum_{k=1}^K f_{em}(x_k|y_k) + f_{tr}(y_k|y_{k-1})))}{Z(\vec{x})}$$

Use F-B

$$\begin{aligned} p(y_k | \vec{x}) &= \frac{\alpha_k(y_k) \beta_k(y_k)}{p(\vec{x})} \\ &= \frac{\alpha_k(y_k) \beta_k(y_k)}{\alpha_{K_{tr}}(END)} \end{aligned}$$

