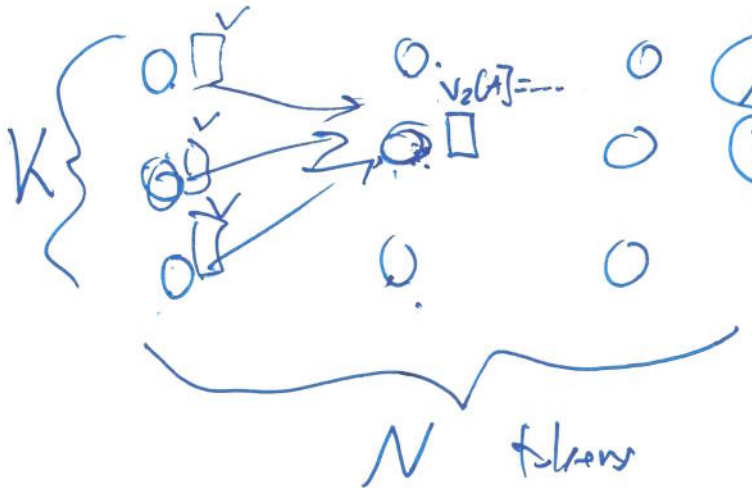
Midterm Review Session

10/24/17 UMass CS 585

Questions from practice midterm questions



N → for $t = 1 \ldots N$

K → for $k = 1 \ldots K$

$V_2[k] := \underset{j=1..K}{\arg\max}\ V_{t-1}[j] + A(j,k) + B(k)$

K

$\Rightarrow O(NK^2)$ time [2.1]

(b) $O(NK)$ vs ~~K~~

$B_1[Adj_1]$ $B_2[Adj_1]$ $\ldots$ $B_N[A]$

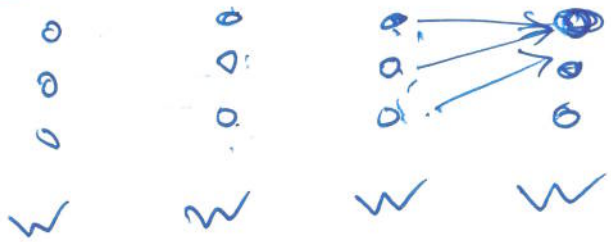
$B_1[Non]$ $B_2[Non]$ $\ldots$ $B_N[Non]$

(c) Enumeration: $O(K^N)$

K K K

Decoder Runtime

Viterbi: Why optimal?



$$\max_{\vec{y}} P(y_1, y_2, y_3, y_4 | \vec{w})$$

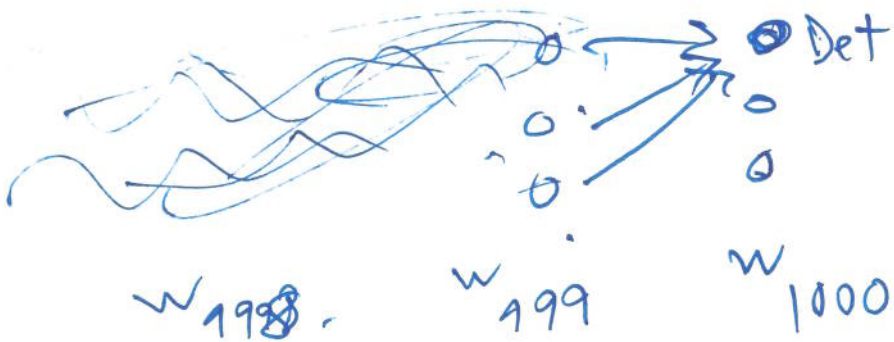
$$\max_{y_1,y_2,y_3} \max_{y_4} P(y_1, y_2, y_3, y_4 | \vec{w})$$

$$G(\vec{y}) = A(y_1, y_2) + A(y_2, y_3) + A(y_3, y_4) + \cancel{B(y_1) + B_{y_2} + B_{y_3} + B_{y_4}}$$

Given

$$(y_1^*, y_2^*) = \underset{(y_1, y_2)}{\operatorname{argmax}} A(y_1, y_2)$$

$$y_3, y_4 ? \quad \underset{y_3, y_4}{\operatorname{argmax}} A(y_1^*, y_2^*) + A(y_2^*, y_3) + A(y_3, y_4)$$



Why is Viterbi optimal? Attempt #2

$V_{1000}[Noun] = \ldots$

$V_{1000}[Adj] = \ldots$

Best $\vec{y}$, where $y_{1000} = Det$?

$$\max_{y_1 \ldots y_{999}} G(y_1, \ldots, y_{999}, y_{1000} = Det) = V_{1000}[Det]$$

$$\ldots A(y_{998}, y_{999}) + \ldots B_{999}(y_{999}) + A(y_{999}, y_{1000} = Det) + B_{1000}(Det)$$

$$= \max_{y_1 \ldots y_{998}} \max_{y_{999}} G(\vec{y})$$

$$= \max_{y_1 \ldots y_{998}} \left[A(y_1, y_2) + \ldots A(y_{997}, y_{998}) + \left(\max_{y_{999}} + A(y_{999}, y_{1000} = Det) + B_{\ldots} \right) \right]$$

Greedy vs. Viterbi

$y_1^* = Det$ $y_2^* = Adj$ $y_3^* = ?$

the happy dog . . . . .

$y_3 = Noun$? $\Rightarrow A(Adj, Noun) + B_3(Noun)$

$y_3 = Det$? $\Rightarrow A(Adj, Det) + B_3(Det)$

$\underset{k \in 1..K}{\text{argmax}}$

N? V? Pronoun

Attack him

Viterbi: optimal!

Pro

$f_{vit}(\vec{x}) \Rightarrow \vec{y}^*$

$y^* = \underset{\vec{y}}{\text{argmax}}\ G(\vec{y})$

N V V

Attack was

Can of Viterbi?

pSparse? Vit time = NK^2 | Vit Space = NK

Greedy time = NK | Greedy Space = $N+K$

N-Gram LM

$w_0 = \text{START}$,

$P(w_1 = A, w_2 = B, w_3 = B, w_4 = \text{STOP/END})$

$= P(A \mid \text{START})\, P(B \mid A)\, P(B \mid B)\, P(\text{END} \mid B)$ ← Model Assump.

$= \frac{\#(\text{START}, A)}{\#(\text{START})} \frac{\#(A, B)}{\#(A)} \frac{\#(B, B)}{\#(B)} \cdots$

$= \left(\frac{1}{1}\right)\left(\frac{1}{2}\right)\left(\frac{0}{2}\right)\left(\frac{\cdots}{\cdots}\right)$

$= 0 \quad \therefore$

(3) Pseudo counts

$= \left(\frac{1+\alpha}{1+\alpha V}\right)\left(\frac{1+\alpha}{2+\alpha V}\right)\left(\frac{0+\alpha}{2+\alpha V}\right)\left(\frac{0+\alpha}{2+\alpha V}\right)$ $V = 4$

$V = \{A, B, C, \text{END}\}$