

Bi-gram Viterbi

0

1

2

3

4

5

6

I

can

can

a

can

$$\delta_{\text{PRO}}(1) = 0.05$$

$$\delta_{\text{MD}}(2) = 0.08$$

$$\delta_{\text{MD}}(3) = 0.002$$

$$\delta_{\text{MD}}(5) = 0.008$$

<s>

PRO

MD

MD

MD

$$\delta_{\text{NN}}(2) = 0.045$$

$$\delta_{\text{NN}}(3) = 0.001$$

$$\delta_{\text{DT}}(4) = 0.035$$

$$\delta_{\text{NN}}(5) = 0.096$$

$$\delta_{\text{<s>}}(6) = 0.000876$$

$$\delta_{\text{PN}}(1) = 0.001$$

$$\delta_{\text{VB}}(2) = 0.03$$

$$\delta_{\text{VB}}(3) = 0.06$$

$$\delta_{\text{VB}}(5) = 0.0012$$

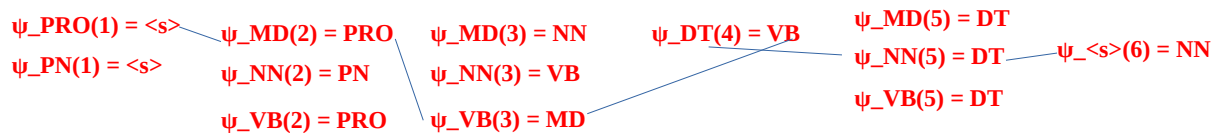
$$\delta_{\text{<s>}}(0) = 1$$

PN

VB

VB

VB



<S> PRO MD VB DT NN

Wie berechnet man Viterbi-WK?

$$\delta_{\text{<s>}}(0) =$$

$$1. \text{ Initialisierung: } \delta_t(0) = \begin{cases} 1 & \text{falls } t = \text{<s>} \\ 0 & \text{sonst} \end{cases}$$

$$\delta_{\text{PRO}}(1) =$$

$$\delta_{\text{PN}}(1) =$$

$$\delta_{\text{MD}}(2) =$$

$$\delta_{\text{NN}}(2) =$$

$$\delta_{\text{VB}}(2) =$$

$$\delta_{\text{MD}}(3) =$$

$$\delta_{\text{MD}}(3) =$$

$$\delta_{\text{MD}}(3) =$$

$$\delta_{\text{DT}}(4) =$$

$$\delta_{\text{MD}}(5) =$$

$$\delta_{\text{NN}}(5) =$$

$$\delta_{\text{VB}}(5) =$$

$$2. \text{ Berechnung: (für } 0 < k \leq n + 1)$$

$$\delta_t(k) = \max_{t'} \delta_{t'}(k-1) p(t|t') p(w_k|t)$$

$$\psi_t(k) = \arg \max_{t'} \delta_{t'}(k-1) p(t|t') p(w_k|t)$$

Tri-gram Viterbi (K=2)

	0	1	2	3	4	5	6
		I	can	can	a	can	
			$\delta_{\text{PRO},\text{MD}(2)}$ $\delta_{\text{PN},\text{MD}(2)}$	$\delta_{\text{MD},\text{MD}(3)}$ $\delta_{\text{NN},\text{MD}(3)}$ $\delta_{\text{VB},\text{MD}(3)}$		$\delta_{\text{DT},\text{MD}(5)}$	
<s>	$\delta_{\text{<s>},\text{<s>(0)}$ <s>	$\delta_{\text{<s>},\text{PRO}(1)}$ PRO	MD	MD	DT	MD	<s>
		$\delta_{\text{<s>},\text{PN}(1)}$ PN	$\delta_{\text{PRO},\text{NN}(2)}$ $\delta_{\text{PN},\text{NN}(2)}$ NN	NN		NN	
			$\delta_{\text{PRO},\text{VB}(2)}$ $\delta_{\text{PN},\text{VB}(2)}$ VB	VB		VB	
			$\psi_{\text{PRO},\text{MD}(2)} = \text{<s>}$ $\psi_{\text{PN},\text{MD}(2)} = \text{<s>}$	$\psi_{\text{MD},\text{MD}(3)} = \text{PRO}$ $\psi_{\text{NN},\text{MD}(3)} = \text{PN}$ $\psi_{\text{VB},\text{MD}(3)} = \text{PRO}$	$\psi_{\text{DT},\text{MD}(5)} = \text{VB}$	$\psi_{\text{MD},\text{<s>(6)} = \text{DT}$	
			$\psi_{\text{PRO},\text{NN}(2)} = \text{<s>}$ $\psi_{\text{PN},\text{NN}(2)} = \text{<s>}$				
			$\psi_{\text{PRO},\text{VB}(2)} = \text{<s>}$ $\psi_{\text{PN},\text{VB}(2)} = \text{<s>}$				

$$\text{best_tag}(5) = \arg\max_{t'} (\text{vit_prob_t}', t(6))$$

Viterbi-Algorithmus (Trigramm-Tagger)

- Beim Trigramm-Tagger entspricht jeder Zustand des Hidden-Markow-Modelles nicht einem einzelnen Tag, sondern einem Tagpaar.
- Übergänge gibt es nur zwischen Zuständen (t, t') und (t'', t''') mit $t' = t''$

1. Initialisierung:
$$\delta_{t',t}(0) = \begin{cases} 1 & \text{falls } t = \langle s \rangle \text{ und } t' = \langle s \rangle \\ 0 & \text{sonst} \end{cases}$$

2. Berechnung: (für $0 < k \leq n + 1$)

$$\begin{aligned} \delta_{t',t}(k) &= \max_{t''} \delta_{t'',t'}(k-1) p(t|t'', t') p(w_k|t) \\ \psi_{t',t}(k) &= \arg \max_{t''} \delta_{t'',t'}(k-1) p(t|t'', t') p(w_k|t) \end{aligned}$$

Wie berechnet man den Tag an der letzten Position?

Sie müssen dazu zunächst über alle Einträge in der letzten Spalte maximieren. Dann können Sie mit der psi-Variablen die beste Tagfolge finden.

Die Alternative dazu ist, k Endetags hinzuzufügen. Dann müssen Sie nicht maximieren. Das ist eigentlich die einfachere Lösung. (Vielleicht überarbeite ich die Folien dazu.)