Homework 2 | Amin Hosseiny Marani

NLP / CS 425

---

(1) * previously I solved this question with three sentences as examples. I later realized we just need to make A matrix, so you can skip the rectangle at the left. or you can see how I come up with these numbers.

Here 3 sentences with DT, V, N, Adj p.s tags.

(1) mary will park her red car.

<S> <N> <V> <V> <DT> <adj> <N> <E>

(2) She wrote a long will.

<S> <DT> <V> <DT> <adj> <N> <E>

(3) NLP excercises are difficult.

<S> <N> <N> <V> <adj> <E>

* <S> means start and <E> means end of a sentence.

Now, we can build a transition matrix with 5 rows and 5 columns:

	N	V	DT	ADJ	<E>
<S>	2/3	0	1/3	0	0
N	1/5	2/5	0	0	2/5
V	1/5	1/5	2/5	1/5	0
DT	0	1/3	0	2/3	0
ADJ	2/3	0	0	0	1/3

(2)

$O = [o_1, o_2, o_3]$

$$Pr([o_1, o_2, o_3]) \Longrightarrow \sum_{Q \in \{q_1, q_2, q_3\}} pr(O, Q \mid A, B, M) \quad (1)$$

using forward algorithm (1) $= \sum_j P(O, q_T = j) = \sum_j \alpha_T(j)$ (2)

so we have to compute $\alpha_3(j)$ for every j

we know $\alpha_1(j) = pr(o_1, q_1 = j) = Pr(q_1 = j) \times Pr(o_1 \mid q_1 = j)$

$$= M(j) \times b_{j o_1} \quad (3)$$

emmison probabilities from B

using $\alpha_1(j)$ (from 3) we can compute $\alpha_t(j) = \sum_k \alpha_{t-1}(k)\, a_{kj}\, b_j(o_t)$ (4)

row k, col j of A

row j col o_t from B

using 3 and 4

so, $\alpha_2(j) = \sum_k \alpha_1(k)\, a_{kj}\, b_j(o_2) = \sum_k [M(k) \times b_k(o_1)] \times a_{kj}\, b_j(o_2)$ (5)

and $\alpha_3(j) = \sum_k \alpha_2(k)\, a_{kj}\, b_j(o_3) = \sum_k \left[\sum_i [M(i) \times b_i(o_1)] \times a_{ik}\, b_k(o_2) \times a_{kj}\, b_j(o_3) \right]$ (6)

Finally $Pr([o_1, o_2, o_3]) = \sum_j \alpha_{T=3}(j) \xrightarrow{\text{from } 6} \sum_j \left[\sum_k \left[\sum_i M(i) \times b_i(o_1) \times a_{ik} \times b_k(o_2) \right] \times a_{kj} \times b_j(o_3) \right]$

(3)

$$Pr(O) = Pr([o_1, o_2, o_3]) \Longrightarrow \sum_j \overbrace{Pr(q_1 = j)}^{M_j}\, b_j(o_1)\, \beta_1(j) \quad (1)$$

$$\beta_t(j) = \sum_k \beta_{t+1}(k)\, \alpha_{jk}\, b_j(o_t) \quad (2)$$

while $\beta_T(j) = \beta_3(j) = Pr(O_{T+1} \mid q_T = j) = Pr(O_4 \mid q_3 = j) = 1$ (3)

using (2) and (3):

$$\beta_2(j) = \sum_k \beta_3(k)\, \alpha_{jk}\, b_j(o_3) = \sum_k \alpha_{jk}\, b_j(o_3) \quad (4)$$

and using (4) and (2) we have:

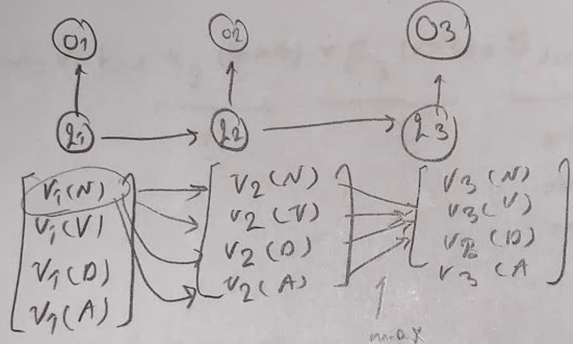
$$\beta_1(j) = \sum_k \beta_2(k)\, \alpha_{jk}\, b_j(o_1) = \sum_k \Big[\sum_i \alpha_{ki}\, b_k(o_3)\Big] \times \alpha_{jk}\, b_j(o_2) \quad (5)$$

finally, using (5) and (1):

$$Pr(O) = Pr([o_1, o_2, o_3]) \Longrightarrow \sum_j M_j\, b_j(o_1) \times \Big[\sum_k \Big[\sum_i \alpha_{ki}\, b_k(o_3)\Big] \times \alpha_{jk}\, b_j(o_2)\Big]$$

(4) $O = [o_1, o_2, o_3]$ POS = Noun/Det/Verb/Adj

we want to compute $V_1(\text{noun}) \to \cdots \to V_3(\text{Verb})$ using Viterbi algorithm:



$V_1(\text{noun}) = \Pi(\text{noun}) \times Pr(O_1 \mid q_1 = \text{noun})$

we have to compute all possible $q_2 = i$ for V_2:

$$V_2(i) = V_1(\text{noun}) \times Pr(q_2 = i \mid q_1 = \text{noun}) \times Pr(O_2 \mid q_2 = i)$$

$$= V_1(\text{noun}) \times a_{\text{noun},i} \times b_i(O_2)$$

Since we know $V_1(i = \text{noun})$ we don't need to compute $\max_{q_1} \Pi(q_1 = j) \times Pr(O_1 \mid q_1 = j)$

we have to compute $V_3(i = \text{verb})$

$$V_3(\text{verb}) = \max_k V_2(k) \times Pr(q_3 = \text{verb} \mid q_2 = k) \times Pr(O_3 \mid q_3 = \text{verb})$$

$$= \max_k V_2(k) \times a_{k,\text{verb}} \times b_{\text{verb}}(O_3)$$

Amin Hosseiny Marani

(A) cont'd

and now to find Q^* we have to try $v_2(j)$ for different j (backtracking)

finally we have to return:

$$Pr(Q^* | O) = \max_{k} v_2(k)\, a_{k,verb}\, b_{verb}(o_3)$$

Amin Hosseiny Marani

5

$\xi_{t=2}(\text{Noun}, \text{verb}) = ?$

$$\xi_t(i,j) = Pr(q_t = i, q_{t+1} = j \mid O, \lambda) = \alpha_t(i) \times \beta_{t+1}(j) \times B_{j(O_{t+1})} \times A_{ij}$$

↓

$t=2$, $i = \text{noun}$, $j = \text{verb}$

↓

I

$$\xi_2(\text{noun}, \text{Verb}) = \underbrace{\alpha_2(\text{Noun})}_{1} \times \underbrace{\beta_3(\text{verb})}_{2} \times \underbrace{B_{\text{verb}}(O_3) \times A_{\text{noun,verb}}}$$

we consider we have these values as we are not asked to compute them.

1: $\alpha_2(\text{Noun}) = \sum_k \alpha_1(k) A_{k,\text{Noun}} B_{\text{Noun}(O_2)}$ 3

$\alpha_1(k) = Pr(q_1 = k) \times Pr(O_1 \mid q_1 = k) = \Pi(k) \times B_{k,(O_1)}$ 4

from 3 and 4: $\alpha_2(\text{Noun}) = \underbrace{\sum_k \Pi(k) \times B_{k,(O_1)} \times A_{k,\text{Noun}} \times B_{\text{Noun},O(2)}}_{5}$

2: $\beta_t(\text{verb}) = pr(O_{t+1} \mid q_t = \text{verb}) \xrightarrow{t=3} \beta_3 = Pr(O_4 \mid q_3 = \text{verb})$

$= pr(\emptyset \mid q_3 = \text{verb}) = 1$ 6

Combing I, 5, and 6: $\xi_2(\text{noun}, \text{verb}) = \left(\sum_k \Pi(k) \times B_{k,(O_1)} \times A_{k,\text{Noun}} \times B_{\text{Noun},O(2)}\right)$

$\times 1 \times B_{\text{verb}}(O_3) \times A_{\text{noun,verb}}$