NLP - ex3

מגישים: אלעד בן דוד
רפאל שקד גרינפלד

נתון כי

$$P(y_1,\dots,y_n \mid x_1,\dots,x_N) = \frac{\prod_{j=1}^{n} \exp(w \cdot f(y_{j-1}, x_1,\dots,x_n, j, y_j))}{Z(x_1,\dots,x_N)}$$

מחוק בייס

$$P(y_i \mid y_{i-1}, x_1,\dots,x_N) = \frac{P(y_i, y_{i-1} \mid x_1,\dots,x_N)}{P(y_{i-1} \mid x_1,\dots,x_N)}$$

כפי שראינו בהרצאה, נגדיר:

$$M_i(y,y') = \exp(f(y, x_1,\dots,x_n, j, y')^T w)$$

$$\alpha_i(y) = \sum_{\substack{y_1,\dots,y_{i-1} \\ y_{i-1}=y}} \prod_{j=1}^{i-1} M_j(y_j, y_{j-1})$$

$$\beta_i(y) = \sum_{\substack{y_1,\dots,y_n \\ y_i=y}} \prod_{i=1}^{n} M_j(y_j, y_{j-1})$$

$$P(y_i, y_{i-1} \mid x_1,\dots,x_N) = \frac{M_i(y,y')\,\alpha_j(y_{j-1})\,\beta_j(y_j)}{Z(x)}$$

נציג את המכנה בחישוב $P(y_{i-1} \mid x_1,\dots,x_N)$
תכנון דינאמי של α ו β

$$P(y_{i-1} \mid x_1,\dots,x_N) = \frac{\alpha_j(y_{j-1})\,\beta_j(y_{j-1})}{Z(x)}$$

נציג:

$$P(y_i \mid y_{i-1}, x_1, \dots, x_N) = \frac{P(y_i, y_{i-1} \mid x_1, \dots, x_N)}{P(y_{i-1} \mid x_1, \dots, x_N)} =$$

$$= \frac{\dfrac{M_i(y, y')\,\alpha_j(y_{j-1})\,\beta_j(y_j)}{Z(x)}}{\dfrac{\alpha_j(y_{j-1})\,\beta_j(y_{j-1})}{Z(x)}} = \frac{M_i(y, y')\,\beta_j(y_j)}{\beta_j(y_{j-1})}$$

כעת אלגוריתם לחישוב β: נניח כי M הוא מספר התגיות ו-N הוא אורך המשפט

```
Def Beta((x_1,...,x_N), w, f, y_j, y_{j-1}, i'):
    init β ← M ∈ M_{M×N}(ℝ)
    For i in range(N-1,...,0):
        For j in range(M-1,...,0):
            β(i,j) = Σ_{k=1}^{M} β(k, k+1) × exp(f(y_{k+1}, x_1,...,x_n, i', y_k)^T w)
    return β
```

וכעת חישוב האלגוריתם

```
Def SectionA((x_1,...,x_N), w, f, y_j, y_{j-1}, i'):
    init output ← M_{M×M}(ℝ)
    β ← Beta((x_1,...,x_N), w, f, y_j, y_{j-1}, i')
    For i in range(1,...,M):
        For j in range(1,...,M):
            output(i,j) = exp(f(y_i, x_1,..., x_n, i', y_j)^T w) × β_i(y_j)/β_i(y_j)
    return output
```

ם) כפי שראינו:

נסמן $\leftarrow \{X\}_1^N$

$$P(y_i | x_1 \ldots x_N) = \frac{\alpha_{i-1}(y_i)\beta_{i+1}(y_i)}{Z(x)}$$

Def Alpha($\{X\}_1^N$, w, f, y_j, y_{j-1}, $\hat{i}$)

0. init $\leftarrow M \in M_{m \times N}(\mathbb{R})$

1. for i in range (1...N)

 1.1 for j in range (1...M)

 1.2 $\alpha(i,j) = \sum_{k=1}^{m} \alpha(k-1,k) \exp(f(y_{k-1}, \{X\}_1^N, \hat{i}, y_k)^T \cdot w)$

2. return α

Def SectionB($\{X\}_1^N$, w, f, y_j, y_{j-1}, $\hat{i}$):

0. init output $\leftarrow M_{m \times N}(\mathbb{R})$

 α = Alpha($\{X\}_1^N$, w, f, y_{j-1}, y_j, $\hat{i}$)

 β = Beta($\{X\}_1^N$, w, f, y_{j+1}, y_j, $\hat{i}$)

 for j in range (1,...,m)

 $$Output[j] = \frac{\alpha(\hat{i}, j)\beta(\hat{i}, j)}{Z}$$