PEE - First Question Set

David Barbas Rebollo

Question 4

- Definition: Given $x = x_1 \dots x_T \in \Sigma^*$ and $A \in N$

$$e(A,i,j) \overset{def}{=} P_\theta(A \overset{*}{\Rightarrow} x_{i+1} \dots x_j)$$

- Initialization: $\forall A \in N; \quad \forall i: 0 \dots T-1;$

$$e(A,i,i+1) = p(A \to b) \cdot \delta(b, x_{i+1})$$

- Recursion: $\forall A \in N; \ \forall l: 2 \dots T; \ \forall i: 0 \dots T-l;$

$$e(A,i,i+l) = \left\{ \sum_{B,C \in N} p(A \to BC) \cdot e(B,i,i+l-1) \cdot e(C,i+l-1,i+l) \right\}$$

- Final Result: Probability of a sentence: $P_\theta(x) = e(S,0,T)$
- Temporal Cost: $O(N^2)$ Being N the length of the sentence

Question 1

- Outside algorithm's initial expression:

$$\forall A \in N; \quad f(A,0,T) = \delta(A,S)$$

Kronecker delta is a function which returns 1 if both variables are equal and 0 otherwise.

$$\delta(A,S) = \begin{cases} 1 & A = S \\ 0 & A \neq S \end{cases}$$

This is done so the substate A is equal to the initial state of the tree. Therefore, its probability is maximum and set to 1.

- Outside algorithm's final expression:

$$P_\theta(x) = \sum_{A \in N} f(A,i,i+1) \cdot p(A \to b) \cdot \delta(b, x_{i+1}) \qquad 0 \leq i \leq T-1$$

The final result is the sum of all right and left productions multiplied by their own A production probability. Recursively, the algorithms computes all outside production's probabilities to determine the probability of sentence x.

Question 2

- Definition: Optimal score and sequence for $x_1 \dots x_t$ ending at $y_t = s \in Y$

$$\varphi_t(s) \overset{\text{def}}{=} \max_{y_1^t ;\, y_t = s} \prod_{i=1}^{t} \Psi_i(y_{i-1}, y_i, x_i)$$

$$b_t(s) \overset{\text{def}}{=} \arg\max_{y_1^t ;\, y_t = s} \prod_{i=1}^{t} \Psi_i(y_{i-1}, y_i, x_i)$$

- Initialization: $\forall s \in Y$; $\varphi_1(s) = \Psi(y_0 = \text{null}, y_1 = s, x_1)$

$$b_1(s) = 0$$

- Recursion: $\forall t = 2 \dots T$; $\forall s \in Y$;

$$\varphi_t(s) = \max_{s' \in Y} \left\{ \varphi_{t-1}(s') \cdot \Psi_t(y_{t-1} = s', y_t = s, x_t) \right\}$$

$$b_t(s) = \arg\max_{s' \in Y} \left\{ \varphi_{t-1}(s') \cdot \Psi_t(y_{t-1} = s', y_t = s, x_t) \right\}$$

Final Result: Optimal score for x:

$$\max_{s} \varphi_T(s)$$

Optimal sequence, $\hat{y}$, for x:

$$y_T = \arg\max_{s} \varphi_T(s)$$

$$\forall t = T \dots 2; \quad y_{t-1} = b_t(y_t)$$

$$\hat{y} = y_1 \dots y_T$$