# Parsing

**Homework 11 (PCFG Viterbi and EM parameter estimation), due 06 July 2020**

Laura Kallmeyer

SS 2020, Heinrich-Heine-Universität Düsseldorf

**Question 1 (PCFG, viterbi)** *Consider the following CFG $G$ with non-terminals $N = \{S, T, A, B\}$, terminals $T = \{a, b\}$, start symbol $S$ and productions*

| | | | |
|---|---|---|---|
| *0.4 (-0.4)* | $S \to AS$ | *1 (0)* | $T \to AS$ |
| *0.05 (-1.3)* | $S \to TB$ | *1 (0)* | $A \to a$ |
| *0.05 (-1.3)* | $S \to a$ | *1 (0)* | $B \to b$ |
| *0.5 (-0.3)* | $S \to b$ | | |

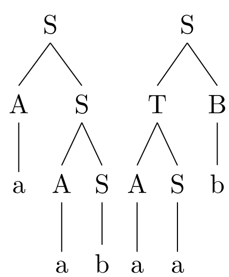
*Preceding each production, its probability is given together with the $log_{10}$ value of the probability in parentheses.*

*We consider the input $w = aab$.*

1. *Give the two parse trees for $w$ together with the respective $log_{10}$ value of their probabilities.*
2. *Give the viterbi chart one obtains when using this PCFG in a CYK-parsing of the input aab. Calculate with the log values instead of the probabilities. This means that instead of multiplying probabilities, you just add their log values.*

   *Note that the higher a probability, the higher the log value and the better the item. (We are using $log_{10}(p)$, not $|log_{10}(p)|$.)*

Solution:

1. 


   with respective weights $-1.1$ and $-2.6$.

2. Chart:

| $l$ | | | | |
|---|---|---|---|---|
| 3 | -1.1:S. -0.7:T | | | |
| 2 | -1.7:S, -1.3:T | -0.7:S, -0.3:T | | |
| 1 | 0:A, -1.3:S | 0:A, -1.3:S | 0:B, -0.3:S | |
| | 1 | 2 | 3 | $i$ |

**Question 2 (PCFG parameter estimation with EM)**

*Consider the PCFG $G = \langle\{S, A, X\}, \{a\}, P, S, p\rangle$ (see course slides) with $P$ and $p$ as follows:*

*0.3:* $S \to AS$ *0.6:* $S \to AX$ *0.1:* $S \to a$ *1:* $X \to SA$ *1:* $A \to a$

*Assume that these probabilities are our starting probabilities for a parameter estimation using EM.*

*Assume that we have a training corpus consisting of 5 sentences, namely 3 sentences aa and 2 sentences aaa.*

*The inside and outside values for the two sentences aa and aaa are as follows:*

*Inside values $\alpha$:*

*aa:*

| *j* | | |
|---|---|---|
| *2* | *($3 \cdot 10^{-2}$,S), (0.1,X)* | *(1,A), (0.1,S)* |
| *1* | *(1,A), (0.1,S)* | |
| | *1* | *2* *i* |

*aaa:*

| *j* | | | |
|---|---|---|---|
| *3* | *($6.9 \cdot 10^{-2}$,S), (0.03,X)* | *($3 \cdot 10^{-2}$,S), (0.1,X)* | *(1,A), (0.1,S)* |
| *2* | *($3 \cdot 10^{-2}$,S), (0.1,X)* | *(1,A), (0.1,S)* | |
| *1* | *(1,A), (0.1,S)* | | |
| | *1* | *2* | *3* *i* |

*Outside values $\beta$ (only values $\neq 0$ are given):*

*aa*

| *j* | | | |
|---|---|---|---|
| *2* | *(1,S)* | *(0.3,S), (0.6,X)* | |
| *1* | *(0.03,A)* | | |
| | *1* | *2* | *i* |

*aaa*

| *j* | | | | |
|---|---|---|---|---|
| *3* | *(1,S)* | *(0.3,S), (0.6,X)* | *($9 \cdot 10^{-2}$,S), (0.18,X), ($3 \cdot 10^{-2}$,A)* | |
| *2* | *(0.03,A)* | *(0.6,S), ($9 \cdot 10^{-3}$, A)* | | |
| *1* | *($6.9 \cdot 10^{-2}$,A)* | | | |
| | *1* | *2* | *3* | *i* |

1. *E-step: Compute the new counts $C_{aa}(A \to \alpha)$ and $C_{aaa}(A \to \alpha)$ and, based on these, the new frequency $f(A \to \alpha)$ for all $A \to \alpha \in P$.*
2. *M-step: Compute the new probabilities $\hat{p}(A \to \alpha)$ for all $A \to \alpha \in P$, based on the previous frequencies.*

Solution:

1. $C_{aa}(S \to AS) = \frac{\beta_{S,1,2}\alpha_{A,1,1}\alpha_{S,2,2}p(S\to AS)}{\alpha_{S,1,2}} = \frac{1\cdot 1\cdot 0.1\cdot 0.3}{0.03} = 1$

   $C_{aa}(S \to AX) = \frac{\beta_{S,1,2}\alpha_{A,1,1}\alpha_{X,2,2}p(S\to AS)}{\alpha_{S,1,2}} = 0$

   $C_{aa}(X \to SA) = 0$

   $C_{aaa}(S \to AS) = \frac{\beta_{S,1,3}\alpha_{A,1,1}\alpha_{S,2,3}p(S\to AS)}{\alpha_{S,1,3}} + \frac{\beta_{S,1,2}\alpha_{A,1,1}\alpha_{S,2,2}p(S\to AS)}{\alpha_{S,1,3}} + \frac{\beta_{S,2,3}\alpha_{A,2,2}\alpha_{S,3,3}p(S\to AS)}{\alpha_{S,1,3}} = \frac{1\cdot 1\cdot 0.03\cdot 0.3+0+0.3\cdot 1\cdot 0.1\cdot 0.3}{0.069} = 0.26$

   $C_{aaa}(S \to AX) = \frac{\beta_{S,1,3}\alpha_{A,1,1}\alpha_{X,2,3}p(S\to AX)}{\alpha_{S,1,3}} + \frac{\beta_{S,1,2}\alpha_{A,1,1}\alpha_{X,2,2}p(S\to AX)}{\alpha_{S,1,3}} + \frac{\beta_{S,2,3}\alpha_{A,2,2}\alpha_{X,3,3}p(S\to AX)}{\alpha_{S,1,3}} = \frac{1\cdot 1\cdot 0.1\cdot 0.6+0+0}{0.069} = 0.87$

   $C_{aaa}(X \to SA) = \frac{\beta_{X,2,3}\alpha_{S,2,2}\alpha_{A,3,3}p(X\to SA)}{\alpha_{S,1,3}} = \frac{0.6\cdot 0.1\cdot 1}{0.069} = 0.87$

   $C_{aa}(S \to a) = \frac{(\beta_{S,1,1}+\beta_{S,2,2})p(S\to a)}{\alpha_{S,1,2}} = \frac{0.3\cdot 0.1}{0.03} = 1$

   $C_{aa}(A \to a) = \frac{(\beta_{A,1,1}+\beta_{A,2,2})p(A\to a)}{\alpha_{S,1,2}} = \frac{0.03}{0.03} = 1$

   $C_{aaa}(S \to a) = \frac{(\beta_{S,1,1}+\beta_{S,2,2}+\beta_{S,3,3})p(S\to a)}{\alpha_{S,1,3}} = \frac{0.69\cdot 0.1}{0.069} = 1$

   $C_{aaa}(A \to a) = \frac{(\beta_{A,1,1}+\beta_{A,2,2}+\beta_{A,3,3})p(A\to a)}{\alpha_{S,1,3}} = \frac{0.069+0.00899+0.03}{0.069} = 1.57$

   $f(S \to AS) = 3\cdot 1 + 2\cdot 0.26 = 3.52$

   $f(S \to AX) = 3\cdot 0 + 2\cdot 0.87 = 1.74$

   $f(X \to SA) = 3\cdot 0 + 2\cdot 0.87 = 1.74$

$f(S \to a) = 3 \cdot 1 + 2 \cdot 1 = 5$

$f(A \to a) = 3 \cdot 1 + 2 \cdot 1.57 = 6.14$

2. $\hat{p}(S \to AS) = \frac{3.52}{3.52+1.74+5} = 0.34$

   $\hat{p}(S \to AX) = \frac{1.74}{3.52+1.74+5} = 0.17$

   $\hat{p}(S \to a) = \frac{5}{3.52+1.74+5} = 0.49$

   $\hat{p}(X \to SA) = \hat{p}(A \to a) = 1$