Machine Learning
Exercises: PCFG 2

Laura Kallmeyer

Summer 2016, Heinrich-Heine-Universität Düsseldorf

Exercise 1 *Consider again the PCFG $G = \langle\{S, A, X\}, \{a\}, P, S, p\rangle$ (see course slides) with P and p as follows:*

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

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.

1. *Give inside and outside values for the two sentences aa and aaa (many of these values are already on the slides 12 and 15).*
2. *E-step: Compute the new counts $C_{aa}(A \rightarrow \alpha)$ and $C_{aaa}(A \rightarrow \alpha)$ and, based on these, the new frequency $f(A \rightarrow \alpha)$ for all $A \rightarrow \alpha \in P$.*
3. *M-step: Compute the new probabilities $\hat{p}(A \rightarrow \alpha)$ for all $A \rightarrow \alpha \in P$, based on the previous frequencies.*

Solution:

1. Inside values α:

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 β (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), ($8.99 \cdot 10^{-3}$,A)		
1	($6.9 \cdot 10^{-2}$,A)			
	1	2	3	*i*

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

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

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

$C_{aaa}(S \to AS) = \frac{\beta_{S,1,3}\alpha_{A,1,1}\alpha_{S,2,3}p(S\to AS)}{\alpha_{S,1,3}} + \frac{\beta_{S,1,2}\alpha_{A,1,1}\alpha_{S,2,2}p(S\to AS)}{\alpha_{S,1,3}} + \frac{\beta_{S,2,3}\alpha_{A,2,2}\alpha_{S,3,3}p(S\to AS)}{\alpha_{S,1,3}} = \frac{1\cdot1\cdot0.03\cdot0.3+0+0.3\cdot1\cdot0.1\cdot0.3}{0.069} = 0.26$

$C_{aaa}(S \to AX) = \frac{\beta_{S,1,3}\alpha_{A,1,1}\alpha_{X,2,3}p(S\to AX)}{\alpha_{S,1,3}} + \frac{\beta_{S,1,2}\alpha_{A,1,1}\alpha_{X,2,2}p(S\to AX)}{\alpha_{S,1,3}} + \frac{\beta_{S,2,3}\alpha_{A,2,2}\alpha_{X,3,3}p(S\to AX)}{\alpha_{S,1,3}} = \frac{1\cdot1\cdot0.1\cdot0.6+0+0}{0.069} = 0.87$

$C_{aaa}(X \to SA) = \frac{\beta_{X,2,3}\alpha_{S,2,2}\alpha_{A,3,3}p(X\to SA)}{\alpha_{S,1,3}} = \frac{0.6\cdot0.1\cdot1}{0.069} = 0.87$

$C_{aa}(S \to a) = \frac{(\beta_{S,1,1}+\beta_{S,2,2})p(S\to a)}{\alpha_{S,1,3}} = \frac{0.1\cdot0.1}{0.069} = 0.14$

$C_{aa}(A \to a) = \frac{(\beta_{A,1,1}+\beta_{A,2,2})p(A\to a)}{\alpha_{S,1,3}} = \frac{0.03}{0.069} = 0.43$

$C_{aaa}(S \to a) = \frac{(\beta_{S,1,1}+\beta_{S,2,2}+\beta_{S,3,3})p(S\to a)}{\alpha_{S,1,3}} = \frac{0.69\cdot0.1}{0.069} = 1$

$C_{aaa}(A \to a) = \frac{(\beta_{A,1,1}+\beta_{A,2,2}+\beta_{A,3,3})p(A\to a)}{\alpha_{S,1,3}} = \frac{0.069+0.00899+0.003}{0.069} = 1.17$

$f(S \to AS) = 3\cdot1 + 2\cdot0.26 = 3.52$

$f(S \to AX) = 3\cdot0 + 2\cdot0.87 = 1.74$

$f(X \to SA) = 3\cdot0 + 2\cdot0.87 = 1.74$

$f(S \to a) = 3\cdot0.14 + 2\cdot0.1 = 2.42$

$f(A \to a) = 3\cdot0.43 + 2\cdot1.17 = 3.63$

3. $\hat{p}(S \to AS) = \frac{3.52}{3.52+1.74+2.42} = 0.46$

   $\hat{p}(S \to AX) = \frac{1.74}{3.52+1.74+2.42} = 0.23$

   $\hat{p}(S \to a) = \frac{2.42}{3.52+1.74+2.42} = 0.32$

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