PCFGs - Inside-outside probabilities

Pawan Goyal

CSE, IIT Kharagpur

Week 5: Lecture 4

How to find the most likely parse?: CKY for PCFG

How to find the most likely parse?: CKY for PCFG

a 1	pilot 2	likes 3	flying 4	planes 5

$S \longrightarrow NP\ VP$		[1.0]
$VP \longrightarrow VBG\ NNS$		[0.1]
$VP \longrightarrow VBZ\ VP$		[0.1]
$VP \longrightarrow VBZ\ NP$		[0.3]
$NP \longrightarrow DT\ NN$		[0.3]
$NP \longrightarrow JJ\ NNS$		[0.4]
$DT \longrightarrow a$		[0.3]
$NN \longrightarrow pilot$		[0.1]
$VBZ \longrightarrow likes$		[0.4]
$VBG \longrightarrow flying$		[0.5]
$JJ \longrightarrow flying$		[0.1]
$NNS \longrightarrow planes$		[.34]

CKY for PCFG

a 1	pilot 2	likes 3	flying 4	planes 5
DT [0.3]	NP [.009]	-	-	S [1.4688x10^{-5}] S [6.12x10^{-6}]
	NN [0.1]	-	-	-
		VBZ [0.4]	-	VP [.001632] VP [.00068]
			JJ [0.1] VBG [0.5]	NP [.0136] VP [.017]
				NNS [.34]

$S \longrightarrow NP\ VP$		[1.0]
$VP \longrightarrow VBG\ NNS$		[0.1]
$VP \longrightarrow VBZ\ VP$		[0.1]
$VP \longrightarrow VBZ\ NP$		[0.3]
$NP \longrightarrow DT\ NN$		[0.3]
$NP \longrightarrow JJ\ NNS$		[0.4]
$DT \longrightarrow a$		[0.3]
$NN \longrightarrow pilot$		[0.1]
$VBZ \longrightarrow likes$		[0.4]
$VBG \longrightarrow flying$		[0.5]
$JJ \longrightarrow flying$		[0.1]
$NNS \longrightarrow planes$		[.34]

0.009 x 0.00068 x 1.0 = 6.12x10^{-6}

Probability of a String

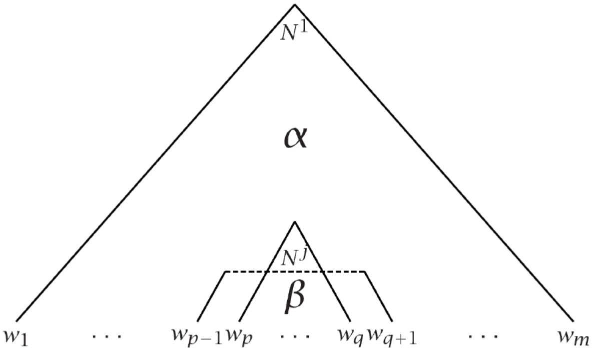
$$P(w_{1m}|G)$$

Probability of a String

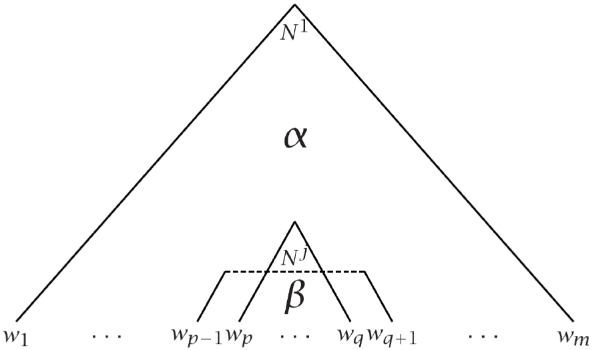
$$P(w_{1m}|G)$$

- In general, simply summing the probabilities of all possible parse trees is not an efficient way to calculate the string probability
- We use *inside algorithm*, a dynamic programming algorithm based on inside probabilities.

Inside and Outside Probabilities

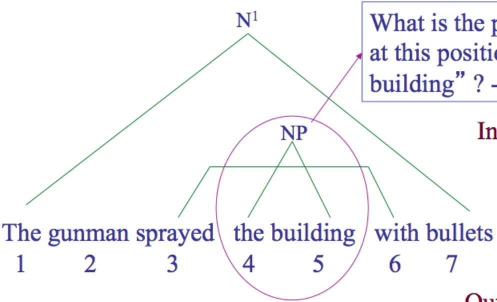


Inside and Outside Probabilities



Outside: $\alpha_j(p,q) = P(w_{1(p-1)}, N^j{}_{pq}, w_{(q+1)m}|G)$
Inside: $\beta_j(p,q) = P(w_{pq}|N^j{}_{pq}, G)$

Inside-outside probabilities



What is the probability that an NP at this position will derive "the building" ? - $\beta_{NP}(4,5)$

Inside Probabilities

Outside Probabilities

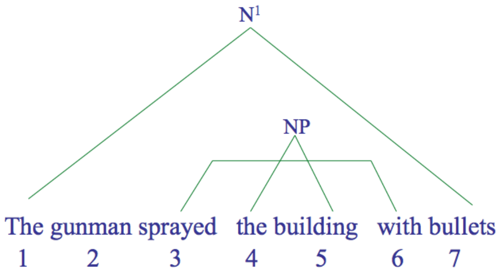
What is the probability of starting from N^1 and deriving "The gunman sprayed", a NP and "with bullets" ? - $\alpha_{NP}(4,5)$

Inside-outside probabilities

$\alpha_{NP}(4,5)$ for "the building"

$= P(\text{The gunman sprayed}, NP_{4,5}, \text{with bullets} \mid G)$

$\beta_{NP}(4,5)$ for "the building" $= P(\text{the building} \mid NP_{4,5}, G)$



Inside Probabilities: Base Step

$$\beta_j(p,q) = P(w_{pq}|N^j{}_{pq}, G)$$

Inside Probabilities: Base Step

$$\beta_j(p,q) = P(w_{pq}|N^j{}_{pq}, G)$$

Base case

$$\beta_j(k,k) = P(w_{kk}|N^j{}_{kk}, G)$$
$$= P(N^j \rightarrow w_k|G)$$

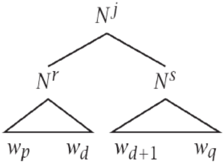
Base case for pre-terminals only

E.g., suppose $N^j = NN$ is being considered and $NN \rightarrow building$ is one of the rules with probability 0.5
$\beta_{NN}(5,5) = P(building|NN_{5,5}, G) = P(NN_{5,5} \rightarrow building|G)$

Inside Probabilities: Induction Step

Assuming Chomsky Normal Form, the first rule must be of the form $N^j \rightarrow N^r N^s$

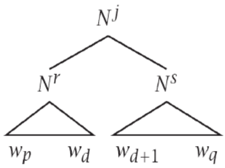
$$\beta_j(p,q) = \sum_{r,s} \sum_{d=p}^{q-1} P(N^j \rightarrow N^r N^s) \beta_r(p,d) \beta_s(d+1,q)$$



Inside Probabilities: Induction Step

Assuming Chomsky Normal Form, the first rule must be of the form $N^j \rightarrow N^r N^s$

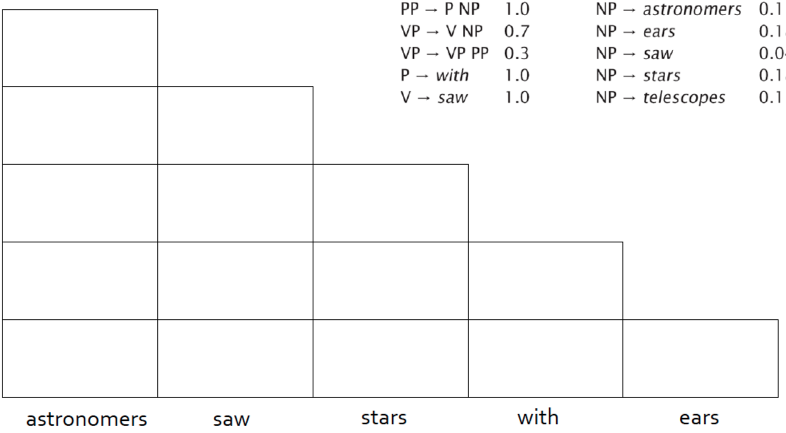
$$\beta_j(p,q) = \sum_{r,s} \sum_{d=p}^{q-1} P(N^j \rightarrow N^r N^s) \beta_r(p,d) \beta_s(d+1,q)$$



- Consider different splits of the words - indicated by d
  E.g., *the huge building*
- Consider different non-terminals to be used in the rule:
  E.g., $NP \rightarrow DT\ NN$, $NP \rightarrow DT\ NNS$

Calculation of inside probabilities

S → NP VP	1.0	NP → NP PP	0.4
PP → P NP	1.0	NP → *astronomers*	0.1
VP → V NP	0.7	NP → *ears*	0.18
VP → VP PP	0.3	NP → *saw*	0.04
P → *with*	1.0	NP → *stars*	0.18
V → *saw*	1.0	NP → *telescopes*	0.1



Calculation of inside probabilities

	1	2	3	4	5
1	$\beta_{NP} = 0.1$		$\beta_S = 0.0126$		$\beta_S = 0.0015876$
2		$\beta_{NP} = 0.04$ $\beta_V = 1.0$	$\beta_{VP} = 0.126$		$\beta_{VP} = 0.015876$
3			$\beta_{NP} = 0.18$		$\beta_{NP} = 0.01296$
4				$\beta_P = 1.0$	$\beta_{PP} = 0.18$
5					$\beta_{NP} = 0.18$
	astronomers	*saw*	*stars*	*with*	*ears*

Outside Probabilities

Compute top-down (after inside probabilities)

Outside Probabilities

Compute top-down (after inside probabilities)

Base Case

Outside Probabilities

Compute top-down (after inside probabilities)

Base Case

$$\alpha_1(1,m) = 1$$

$$\alpha_j(1,m) = 0, j \neq 1$$

Outside Probabilities

Compute top-down (after inside probabilities)

Base Case

$$\alpha_1(1,m) = 1$$

$$\alpha_j(1,m) = 0, j \neq 1$$

Induction

Outside Probabilities

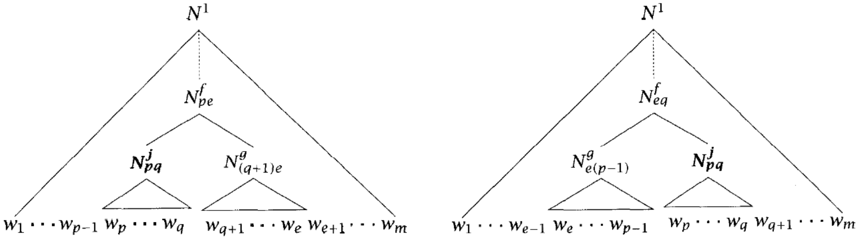
Compute top-down (after inside probabilities)

Base Case

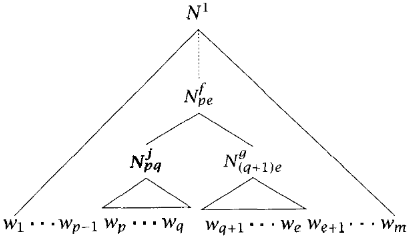
$$\alpha_1(1,m) = 1$$

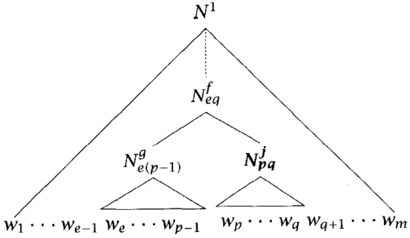
$$\alpha_j(1,m) = 0, j \neq 1$$

Induction



Outside Probabilities: Induction





$$
\begin{aligned}
\alpha_j(p,q) &= \sum_{f,g} \sum_{e=q+1}^{m} \alpha_f(p,e) P(N^f \to N^j N^g) \beta_g(q+1,e) \\
&\quad + \sum_{f,g} \sum_{e=1}^{p-1} \alpha_f(e,q) P(N^f \to N^g N^j) \beta_g(e,p-1)
\end{aligned}
$$

Product of inside-outside probabilities

$$\alpha_j(p,q)\beta_j(p,q) = P(w_{1(p-1)}, {N^j}_{pq}, w_{(q+1)m}|G)P(w_{pq}|{N^j}_{pq}, G) = P(w_{1m}, {N^j}_{pq}|G)$$

Product of inside-outside probabilities

$$\alpha_j(p,q)\beta_j(p,q) = P(w_{1(p-1)}, N^j{}_{pq}, w_{(q+1)m}|G)P(w_{pq}|N^j{}_{pq}, G) = P(w_{1m}, N^j{}_{pq}|G)$$

The probability of the sentence and that there is some consistent spanning from word p to q is given by:

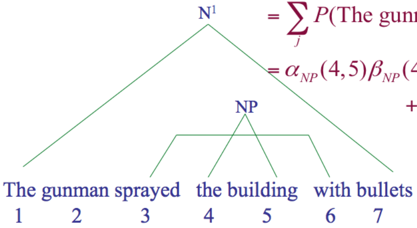
$$P(w_{1m}, N_{pq}|G) = \sum \alpha_j(p,q)\beta_j(p,q) = P(N_1 \rightarrow w_{1m}, N_{pq} \rightarrow w_{pq}|G)$$

Product of inside-outside probabilities

$$\alpha_j(p,q)\beta_j(p,q) = P(w_{1(p-1)}, {N^j}_{pq}, w_{(q+1)m}|G)P(w_{pq}|{N^j}_{pq}, G) = P(w_{1m}, {N^j}_{pq}|G)$$

The probability of the sentence and that there is some consistent spanning from word p to q is given by:

$$P(w_{1m}, N_{pq}|G) = \sum \alpha_j(p,q)\beta_j(p,q) = P(N_1 \rightarrow w_{1m}, N_{pq} \rightarrow w_{pq}|G)$$



$$P(\text{The gunman....bullets}, N_{4,5} \mid G)$$
$$= \sum_j P(\text{The gunman...bullets} \mid N^j_{4,5}, G)$$
$$= \alpha_{NP}(4,5)\beta_{NP}(4,5)$$
$$+ \alpha_{VP}(4,5)\beta_{VP}(4,5) + \ldots$$