Yiwen Zhang

YZ3310

HW3

1(a) For the rules that already follow Chomsky normal form, we keep them unchanged.

For the rules that not follow Chomsky normal form, such as $X \to X_1\ X_2 \dots X_n$ with probability P, we apply transformation function f as following:

$f(X \to X_1\ X_2\ X_3 \dots X_n)$ represents for:

Replace $X \to X_1\ X_2\ X_3 \dots X_n$ with probability P with following rules:

① $X \to X_1\text{-}X_2\text{-}X_3\text{-}\dots\text{-}X_{n-1}\ X_n$ with probability P

② $X_1\text{-}X_2\text{-}X_3\text{-}\dots\text{-}X_{n-1} \to X_1\text{-}X_2\text{-}\dots\text{-}X_{n-2}\ X_{n-1}$ with probability 1

③ $X_1\text{-}X_2\text{-}\dots\text{-}X_{n-2} \to X_1\text{-}X_2\text{-}X_3\text{-}\dots\text{-}X_{n-3}\ X_{n-2}$ with probability 1

④ $X_1\text{-}X_2\text{-}X_3\text{-}\dots\text{-}X_{n-3} \to X_1\text{-}X_2\text{-}\dots\text{-}X_{n-4}\ X_{n-3}$ with probability 1.

$\vdots$

note: all rules omitted here are with probability 1.

(n-1) $X_1\text{-}X_2 \to X_1\ X_2$ with probability 1

The key idea is to make the last symbol itself independent every time and the rest to be a mega symbol, then split the mega symbol step by step.

1(b) We only need to replace the 2nd, the 4th and 5th rule and keep other rules unchanged.

Replace the 2nd rule with $S \to NP\text{-}NP\ VP$ with probability 0.3 and $NP\text{-}NP \to NP\ NP$ with probability 1.

Replace the 4th rule with $VP \to Vt\text{-}NP\ PP$ with probability 0.2 and $Vt\text{-}NP \to Vt\ NP$ with probability 1

Replace the 5th rule with $NP \to DT\text{-}NN\ NN$ with probability 0.3 and $DT\text{-}NN \to DT\ NN$ with probability 1.

∴ The resulting grammar G is following:

$S \to NP\ VP$	0.7	$Vt \to$ saw	1
$S \to NP\text{-}NP\ VP$	0.3	$NN \to$ man	0.7
$NP\text{-}NP \to NP\ NP$	1	$NN \to$ woman	0.2
$VP \to Vt\ NP$	0.8	$NN \to$ telescope	0.1
$VP \to Vt\text{-}NP\ PP$	0.2	$DT \to$ the	1
$Vt\text{-}NP \to Vt\ NP$	1	$IN \to$ with	0.5
$NP \to DT\text{-}NN\ NN$	0.3	$IN \to$ in	0.5
$DT\text{-}NN \to DT\ NN$	1		
$NP \to NP\ PP$	0.7		
$PP \to IN\ NP$	1		

2(a) the PCFG as following:

Σ = {John, said, that, Sally, snored, loudly, declared, Bill, ran, quickly, Fred, pronounced, Jeff, swam, elegantly}

N = {S, NP, VP, V1, SBAR, COMP, ADVP, V2}

S = {S}

~~Re~~ Rules and probability:

S → NP VP 1

SBAR → COMP S 1

VP → V1 SBAR $\frac{1}{3}$

VP → VP ADVP $\frac{1}{3}$

VP → V2 $\frac{1}{3}$

NP → John $\frac{1}{6}$

V1 → said $\frac{1}{3}$

COMP → that 1

NP → Sally $\frac{1}{3}$

V2 → snored $\frac{1}{3}$

ADVP → loudly $\frac{1}{3}$

V1 → declared $\frac{1}{3}$

NP → Bill $\frac{1}{6}$

V2 → ran $\frac{1}{3}$

ADVP → quickly $\frac{1}{3}$

NP → Fred $\frac{1}{6}$

V1 → pronounced $\frac{1}{3}$

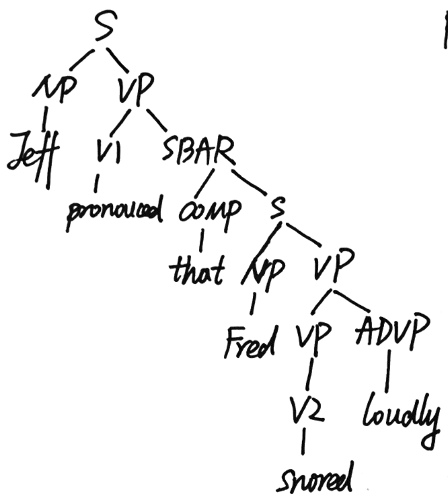
NP → Jeff $\frac{1}{6}$

V2 → swam $\frac{1}{3}$

ADVP → elegantly $\frac{1}{3}$

2(b):

①



probability for tree ①:

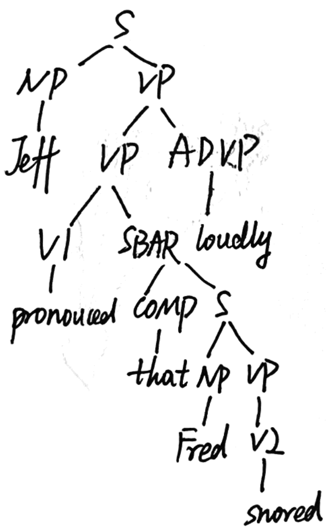
$P_1 = P(S \to NP\ VP) \times P(NP \to Jeff) \times P(VP \to V1\ SBAR) \times P(V1 \to pronouced) \times$
$P(SBAR \to COMP\ S) \times P(COMP \to that) \times P(S \to NP\ VP) \times P(NP \to Fred)$
$\times P(VP \to VP\ ADVP) \times P(VP \to V2) \times P(V2 \to snored) \times P(ADVP \to loudly)$

$= 1 \times 1 \times \frac{1}{3} \times 1 \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{6} \times \frac{1}{3} \times 1 \times \frac{1}{6} \times \frac{1}{3} \times \frac{1}{3}$

$= \frac{1}{3^6} \times \frac{1}{6^2}$

$= \frac{1}{26244}$

②



probability for tree ②:

$P_2 = P(S \to NP\ VP) \times P(VP \to VP\ ADVP) \times P(ADVP \to loudly) \times P(VP \to V1\ SBAR)$
$\times P(SBAR \to COMP\ S) \times P(S \to NP\ VP) \times P(NP \to Jeff) \times P(V1 \to pronouced)$
$\times P(COMP \to that) \times P(NP \to Fred) \times P(VP \to V2) \times P(V2 \to snored)$

$= 1 \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times 1 \times 1 \times \frac{1}{6} \times \frac{1}{3} \times 1 \times \frac{1}{6} \times \frac{1}{3} \times \frac{1}{3}$

$= \frac{1}{3^6} \times \frac{1}{6^2}$

$= \frac{1}{26244}$

2(c) Delete this rule: $VP \to VP\ ADVP$

and Add these 2 rules: $VP \to VP1\ ADVP$ 0.33
$VP1 \to V2$ 1

Q3:

the recursive case as below:

Set $x = \log_2 n$

For $1 \leq i \leq x$:

  For $0 \leq j \leq 2^{x-i} - 1$:

$$\pi(j \cdot 2^i + 1, (j+1) \cdot 2^i, X) = \max_{X \to YZ \in R} \pi(j \cdot 2^i + 1, j \cdot 2^i + 2^{i-1}, Y) \times \pi(j \cdot 2^i + 2^{i-1} + 1, (j+1) \cdot 2^i, Z) \times q(X \to YZ)$$

the running time is $O(|N|^3 n \log_2 n)$, which is more efficient than original CKY — $O(n^3 |N|^3)$.