1

	DT	JJ	NN	VB	CHECK			the	old	man	ships
<S>	0.5	0.1	0.3	0.1	1		DT	1	0	0	0
DT	0	0.5	0.5	0	1		JJ	0	0.8	0.2	0
JJ	0	0.4	0.6	0	1		NN	0	0.2	0.4	0.4
NN	0.1	0	0.2	0.7	1		VB	0	0	0.5	0.5
VB	0.4	0.2	0.2	0.2	1						

<u>STEP - 1</u>

	the	old	man	the	ships
DT	(0.5) (1)				
JJ	(0.1) (0)				
NN	(0.1) (0)				
VB	(0.1) (0)				

	the	old	man	the	ships
DT	0.5	0			
JJ	0	0.5*0.5*0.8			
NN	0	0.5*0.5*0.2			
VB	0	0			

	the	old	man	the	ships
DT	0.5	0	0		
JJ	0	0.2	Max(0.4*0.2*0.2, 0) = Max(0.016, 0)		
NN	0	0.05	Max(0.6*0.4*0.2, 0.2*0.4*0.05) = Max(0.048, 0.004)		
VB	0	0	Max(0, 0.7*0.5*0.05) = Max(0, 0.0175)		

	the	old	man	the	ships
DT	0.5	0	0	Max(0, 0.1*0.048, 0.4*0.0175) = Max(0, 0.0048, 0.07)	
JJ	0	0.2	0.016	0	
NN	0	0.05	0.048	0	
VB	0	0	0.0175	0	

	the	old	man	the	ships
DT	0.5	0	0	0.07	0
JJ	0	0.2	0.016	0	0
NN	0	0.05	0.048	0	0.5*0.4*0.07 = 0.014
VB	0	0	0.0175	0	0

DT	NN	VB	DT	NN

CONSTITUENT : Groups of words behaving as a single unit.

Evidences for constituency :
1. Appear in similar syntactic environments
2. Preposed and postposed constructions.

Identifying constituents in sentences :

A. The cat walked across the porch with a confident air
   It is a constituent as we can change it as follows:
   With a confident air, the cat walked across the porch.

B. They arrived at the concert more quickly than they expected
   It is not a constituent.
   [REPLACEMENT - NO, MOVEMENT - NO]

C. I am very fond of my nephew
   It is not a constituent.
   [REPLACEMENT - NO, MOVEMENT - NO]

Graph based parser consists of two components :

1. Parsing algorithm for inference or searching [Dynamic Programming]
2. Parameter estimation [Machine Learning approaches]

<u>FEATURES :</u>

Each edge score is a weighted sum of features extracted.
It is the relative order of edges not the numbers that matter here.

In considering features we try to capture information about the relationship between heads and their dependents.

Two words u, w highly dependent in one sense of word u and independent in another sense.
Incorporating sense level information in learning weights for arcs can be improve accuracy of the model.

Commonly used features : Wordforms, lemmas, parts of speech of the headword and its dependent.
Along with them we also add features :

Unigram head sense
Unigram dependent sense
Bigram head dependent sense

Linear interpolation between
1. Smoothed bigram model using Knesser Ney Smoothing with β=1
2. Unigram model

c(u,v)	the	man	saw	dog	with	telescope	in	park	SUM	c(u,v)>0
the	0	3	0	2	0	2	0	2	9	4
man	0	0	1	0	1	0	1	0	3	3
saw	3	0	0	0	0	0	0	0	3	1
dog	0	0	1	0	1	0	0	0	2	2
with	2	0	0	0	0	0	0	0	2	1
telescope	0	0	1	0	0	0	0	0	1	1
in	1	0	0	0	0	0	0	0	1	1
park	0	0	0	0	0	0	0	0	0	0
									21	13

<u>DISCOUNTED COUNTS</u>

										$\lambda(w_{i-1})$
c(u,v)	the	man	saw	dog	with	telescope	in	park		
the	0	2	0	1	0	1	0	1		0.44444
man	0	0	0	0	0	0	0	0		1
saw	2	0	0	0	0	0	0	0		0.33333
dog	0	0	0	0	0	0	0	0		1
with	1	0	0	0	0	0	0	0		0.5
telescope	0	0	0	0	0	0	0	0		0.5
in	0	0	0	0	0	0	0	0		1
park	0	0	0	0	0	0	0	0		0

	c(u)	p
the	9	0.38
man	3	0.13
saw	3	0.13
dog	2	0.08
with	2	0.08
telescope	2	0.08
in	1	0.04
park	2	0.08
	24	1

P(dog|the)

= 0.5*[Ps(dog|the)] + 0.5*P(the)
= 0.5*[c_discounted(the, dog)/c(the) + λ(the) * P_cont(dog)] + 0.5*(9/24)

c_discounted(the, dog) = 2-1 = 1
c(the) = 9

λ(the) = 4/9
P_cont(dog) = 1/13

= 0.5* [1/9 + (4/9)*(1/13)] + 0.5 * (9/24)

6

	DT	NN	VB	</S>			the	can	see	</S>
<S>	0.5	0.3	0.2	0		DT	1	0	0	0
DT	0	0.9	0.1	0		NN	0	0.9	0.1	0
NN	0.1	0.2	0.3	0.4		VB	0	0.5	0.5	0
VB	0.4	0.2	0.2	0.2		</S>	0	0	0	1

	can	the	can	see	</s>
DT	0				
NN	0.3*0.9				
VB	0.2*0.5				
</S>	0				

	can	the	can	see	</s>
DT	0	Max(0.1*1*0.27, 0.4*1*0.1) = Max(0.027, 0.04)			
NN	0.27	0			
VB	0.1	0			
</S>	0	0			

	can	the	can	see	</s>
DT	0	0.04	0		
NN	0.27	0	0.9*0.9*0.04		
VB	0.1	0	0.1*0.5*0.04		
</S>	0	0	0		

	can	the	can	see	</s>
DT	0	0.04	0	0	
NN	0.27	0	0.0324	Max(0.2*0.1*0.0324, 0.2*0.1*0.002) = Max(0.000648, 0.00004)	
VB	0.1	0	0.002	Max(0.3*0.5*0.0324, 0.2*0.5*0.002) = Max(0.00486, 0.0002)	
</S>	0	0	0	0	

	can	the	can	see	</s>
DT	0	0.04	0	0	0
NN	0.27	0	0.0324	0.000648	0
VB	0.1	0	0.002	0.00486	0
</S>	0	0	0	0	Max(0.4*0.000648, 0.2*0.00486)= Max(0.0002592, 0.000972)

	can	the	can	see	</s>
DT	0	0.04	0	0	0
NN	0.27	0	0.0324	0.000648	0
VB	0.1	0	0.002	0.00486	0
</S>	0	0	0	0	0.000972

VB	DT	NN	VB	</S>

7

Given rules are not in CNF.

Converting rules to CNF gives us :

NP ---> DT NN 0.4
NP ---> DT NBAR 0.6

NBAR ---> JJ NN 0.12
NBAR ---> NBAR NN 0.096
NBAR ---> NN NBAR
NBAR ---> NN NN

NBAR --> JJ NBAR
NBAR --> NBAR NBAR

	the	blue	fountain	pen	ink
the	DT 1.0	-			
blue		JJ 1.0			
fountain			NN 0.3		
pen				NN 0.4	
ink					NN 0.3

	the	blue	fountain	pen	ink
the	DT 1.0	-			
blue		JJ 1.0	NBAR 0.12*1.0*0.3		
fountain			NN 0.3	NBAR 0.3*0.4*0.096	
pen				NN 0.4	NBAR 0.3*0.4*0.096
ink					NN 0.3

	the	blue	fountain	pen	ink
the	DT 1.0	-	NP 0.6*0.12*0.3	NP, NP	
blue		JJ 1.0	NBAR 0.12*1.0*0.3	NBAR, NBAR	
fountain			NN 0.3	NBAR 0.3*0.4*0.096	NBAR, NBAR
pen				NN 0.4	NBAR 0.3*0.4*0.096
ink					NN 0.3

	the	blue	fountain	pen	ink
the	DT 1.0	-	NP 0.6*0.12*0.3	NP, NP	
blue		JJ 1.0	NBAR 0.12*1.0*0.3	NBAR, NBAR	NBAR, NBAR, NBAR, NBAR
fountain			NN 0.3	NBAR 0.3*0.4*0.096	NBAR, NBAR
pen				NN 0.4	NBAR 0.3*0.4*0.096
ink					NN 0.3

	the	blue	fountain	pen	ink
the	DT 1.0	-	NP 0.6*0.12*0.3	NP, NP	NP, NP, NP, NP
blue		JJ 1.0	NBAR 0.12*1.0*0.3	NBAR, NBAR	NBAR, NBAR, NBAR, NBAR
fountain			NN 0.3	NBAR 0.3*0.4*0.096	NBAR, NBAR
pen				NN 0.4	NBAR 0.3*0.4*0.096
ink					NN 0.3

There are 4 possible parses for the above sentence.

8

Phrase based Machine Translation with the help of Word Sense Disambiguation.

1. In translation, different senses of a word in source language may have different translations in a target language depending on the particular context.
2. So, WSD system can help Machine Translation system to determine the correct translation for ambiguous word.

WSD system for lexical selection in MT by aligning target language words as sense tags.

9

EM to update parameters and find t(cheval|horse)

INITIAL PARAMETERS: [Initialized uniformly]

t	grand	cheval
big	1/2	1/2
horse	1/2	1/2

We have two sentences :

For k = 1 : (m = 2, l=2)
For k = 2 : (m = 1, l=1)

q(j|i,l,m)
j - ENGLISH WORD
i - FRENCH WORD

q(1|1,1,1)	1		
q(1|1,2,2)	1/2	q(2|1,2,2)	1/2
q(1|2,2,2)	1/2	q(2|2,2,2)	1/2

NEXT STEP:

δ(k,i,j)

δ(2,1,1)	1	
δ(1,1,1)	q(1|1,2,2) t(grand|big) / [(q(1|1,2,2) t(grand|big) + q(2|1,2,2) t(grand|horse))] = (1/2)(1/2)/ [(1/2)(1/2) +(1/2)(1/2)]	1/2
δ(1,1,2)	q(2|1,2,2) t(grand|horse) / [(q(1|1,2,2) t(grand|big) + q(2|1,2,2) t(grand|horse))] =(1/2)(1/2)/ [(1/2)(1/2) +(1/2)(1/2)]	1/2
δ(1,2,1)	q(1|2,2,2) t(cheval|big) / [q(1|2,2,2) t(cheval|big) + q(2|2,2,2) t(cheval|horse))] = (1/2)(1/2)/ [(1/2)(1/2) +(1/2)(1/2)]	1/2
δ(1,2,2)	q(2|2,2,2) t(cheval|horse) / [q(1|2,2,2) t(cheval|big) + q(2|2,2,2) t(cheval|horse))] =(1/2)(1/2)/ [(1/2)(1/2) +(1/2)(1/2)]	1/2

c(big, grand)	1/2 + 1		c(big)	2		t(cheval|horse)	c(horse,cheval)/c(horse)
c(big, cheval)	1/2		c(horse)	1			1/2
c(horse, grand)	1/2						
c(horse, cheval)	1/2						

10

	the	can	can	see	the	can
the	DT 1.0					
can		NN 0.9, VB 0.5				
can			NN 0.9, VB 0.5			
see				NN 0.1, VB 0.5		
the					DT 1.0	
can						NN 0.9, VB 0.5

	the	can	can	see	the	can
the	DT 1.0	NP 0.9				
can		NN 0.9, VB 0.5	-			
can			NN 0.9, VB 0.5	-		
see				NN 0.1, VB 0.5	-	
the					DT 1.0	NP 0.9
can						NN 0.9, VB 0.5

	the	can	can	see	the	can
the	DT 1.0	NP 0.9	-			
can		NN 0.9, VB 0.5	-	-		
can			NN 0.9, VB 0.5	-	-	
see				NN 0.1, VB 0.5	-	VP 0.7*0.5*0.9
the					DT 1.0	NP 0.9
can						NN 0.9, VB 0.5

	the	can	can	see	the	can
the	DT 1.0	NP 0.9	-	-		
can		NN 0.9, VB 0.5	-	-	-	
can			NN 0.9, VB 0.5	-	-	VP 0.3*0.315*0.5
see				NN 0.1, VB 0.5	-	VP 0.315
the					DT 1.0	NP 0.9
can						NN 0.9, VB 0.5

	the	can	can	see	the	can
the	DT 1.0	NP 0.9	-	-	-	
can		NN 0.9, VB 0.5	-	-	-	VP 0.3*0.04725*0.5
can			NN 0.9, VB 0.5	-	-	VP 0.04725
see				NN 0.1, VB 0.5	-	VP 0.315
the					DT 1.0	NP 0.9
can						NN 0.9, VB 0.5

	the	can	can	see	the	can
the	DT 1.0	NP 0.9	-	-	-	S 1*0.04725*0.9
can		NN 0.9, VB 0.5	-	-	-	VP 0.0070875
can			NN 0.9, VB 0.5	-	-	VP 0.04725
see				NN 0.1, VB 0.5	-	VP 0.315
the					DT 1.0	NP 0.9
can						NN 0.9, VB 0.5

	the	can	can	see	the	can
the	DT 1.0	NP 0.9	-	-	-	S 1*0.04725*0.9
can		NN 0.9, VB 0.5	-	-	-	VP 0.0070875
can			NN 0.9, VB 0.5	-	-	VP 0.04725
see				NN 0.1, VB 0.5	-	VP 0.315
the					DT 1.0	NP 0.9
can						NN 0.9, VB 0.5

11

Rules to be added for parsing

"I think I like black cats"
"I heard cats like dogs"

X --> think
X --> heard

Considering S as the Simple Declarative Clause of Penn Tree bank tagset :

VB --> think
VB --> heard

VP --> VB Y
Y --> S

If S is considered as a start symbol and can only occur at the start :

VB --> think
VB --> heard

12

Naïve Bayes Classifier using skip gram features:

P(A) = P(B) = P(C) = 1/3

$\mathrm{P(A|the\ cat\ in\ the\ hat)} \propto \mathrm{P(the\ cat\ in\ the\ hat\ |\ A)}$

SKIP GRAM FEATURES : General notation k-skip-n-grams
For simpler calculations we assume features to be bigrams. So, we include 1-skip-2-grams :

This includes all the bigrams + few sub sequences :
So all bigram features with counts for each class can be represented as a matrix as follows:

A	the	cat	sat	on	mat	in	hat	dog	log	fish	dish
the	0	2	2	0	1	0	1	0	0	0	0
cat	0	0	2	1	0	1	0	0	0	0	0
sat	2	0	0	1	0	1	0	0	0	0	0
on	1	0	0	0	1	0	0	0	0	0	0
mat	0	0	0	0	0	0	0	0	0	0	0
in	1	0	0	0	0	0	1	0	0	0	0
hat	0	0	0	0	0	0	0	0	0	0	0
dog	0	0	0	0	0	0	0	0	0	0	0
log	0	0	0	0	0	0	0	0	0	0	0
fish	0	0	0	0	0	0	0	0	0	0	0
dish	0	0	0	0	0	0	0	0	0	0	0

<u>SMOOTHED FEATURE PROBABILITIES GIVEN CLASS A</u>

A	the	cat	sat	on	mat	in	hat	dog	log	fish	dish
the	0.06	0.18	0.18	0.06	0.12	0.06	0.12	0.06	0.06	0.06	0.06
cat	0.07	0.07	0.2	0.13	0.07	0.13	0.07	0.07	0.07	0.07	0.07
sat	0.2	0.07	0.07	0.13	0.07	0.13	0.07	0.07	0.07	0.07	0.07
on	0.15	0.08	0.08	0.08	0.15	0.08	0.08	0.08	0.08	0.08	0.08
mat	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
in	0.15	0.08	0.08	0.08	0.08	0.08	0.15	0.08	0.08	0.08	0.08
hat	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
dog	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
log	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
fish	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
dish	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09

B	the	cat	sat	on	mat	in	hat	dog	log	fish	dish
the	0	1	2	0	0	0	0	2	1	0	0
cat	0	0	0	0	0	0	0	0	0	0	0
sat	2	0	0	2	0	0	0	0	0	0	0
on	2	1	0	0	0	0	0	0	1	0	0
mat	0	0	0	0	0	0	0	0	0	0	0
in	0	0	0	0	0	0	0	0	0	0	0
hat	0	0	0	0	0	0	0	0	0	0	0
dog	0	0	2	2	0	0	0	0	0	0	0
log	0	0	0	0	0	0	0	0	0	0	0
fish	0	0	0	0	0	0	0	0	0	0	0
dish	0	0	0	0	0	0	0	0	0	0	0

SMOOTHED PROBABILITIES FOR FEAURES GIVEN CLASS B

B	the	cat	sat	on	mat	in	hat	dog	log	fish	dish
the	0.06	0.12	0.18	0.06	0.06	0.06	0.06	0.18	0.12	0.06	0.06
cat	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
sat	0.2	0.07	0.07	0.2	0.07	0.07	0.07	0.07	0.07	0.07	0.07
on	0.2	0.13	0.07	0.07	0.07	0.07	0.07	0.07	0.13	0.07	0.07
mat	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
in	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
hat	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
dog	0.07	0.07	0.2	0.2	0.07	0.07	0.07	0.07	0.07	0.07	0.07
log	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
fish	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
dish	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09

C	the	cat	sat	on	mat	in	hat	dog	log	fish	dish
the	0	0	2	0	0	1	1	0	0	2	1
cat	0	0	0	0	0	0	0	0	0	0	0
sat	1	0	0	0	0	1	0	0	0	0	0
on	0	0	0	0	0	0	0	0	0	0	0
mat	0	0	0	0	0	0	0	0	0	0	0
in	2	0	0	0	0	0	1	0	0	0	1
hat	0	0	1	0	0	0	0	0	0	0	0
dog	0	0	0	0	0	0	0	0	0	0	0
log	0	0	0	0	0	0	0	0	0	0	0
fish	1	0	1	0	0	2	0	0	0	0	0
dish	0	0	0	0	0	0	0	0	0	0	0

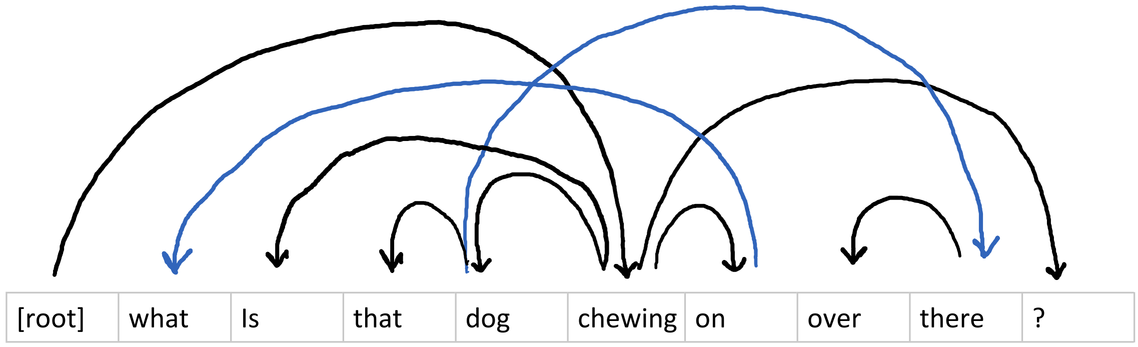
SMOOTHED PROBABILITIES OF FEATURES GIVEN CLASS C

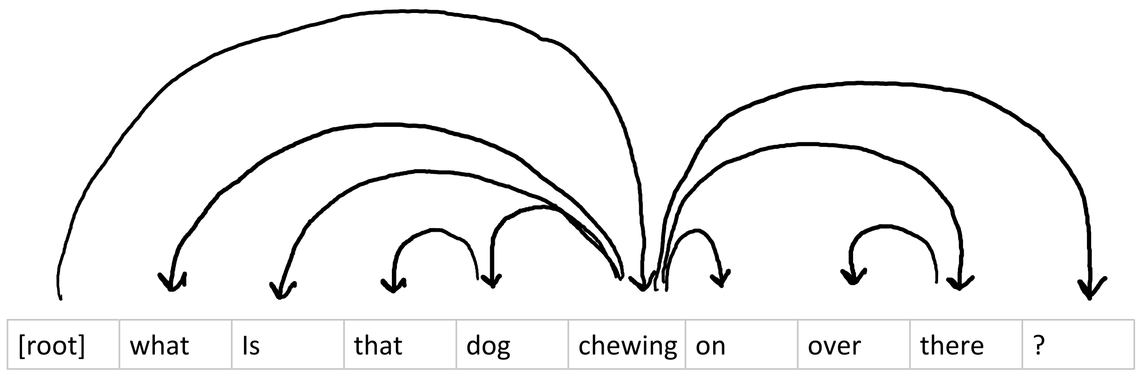
C	the	cat	sat	on	mat	in	hat	dog	log	fish	dish
the	0.06	0.06	0.17	0.06	0.06	0.11	0.11	0.06	0.06	0.17	0.11
cat	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
sat	0.15	0.08	0.08	0.08	0.08	0.15	0.08	0.08	0.08	0.08	0.08
on	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
mat	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
in	0.2	0.07	0.07	0.07	0.07	0.07	0.13	0.07	0.07	0.07	0.13
hat	0.08	0.08	0.17	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
dog	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
log	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
fish	0.13	0.07	0.13	0.07	0.07	0.2	0.07	0.07	0.07	0.07	0.07
dish	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09

Features in the sentence "the cat in the hat"

	A	B	C		log(A)	log(B)	log(C)
the,cat	0.17647	0.12	0.06		-0.7533	-0.929418926	-1.255272505
cat,in	0.13333	0.09	0.09		-0.8751	-1.041392685	-1.041392685
in,the	0.15385	0.09	0.2		-0.8129	-1.041392685	-0.698970004
the,hat	0.11765	0.06	0.11		-0.9294	-1.230448921	-0.954242509
the,in	0.05882	0.06	0.11		-1.2304	-1.230448921	-0.954242509
cat,the	0.06667	0.09	0.09		-1.1761	-1.041392685	-1.041392685
in,hat	0.15385	0.09	0.13		-0.8129	-1.041392685	-0.875061263
					-6.5902	-7.555887509	-6.820574162

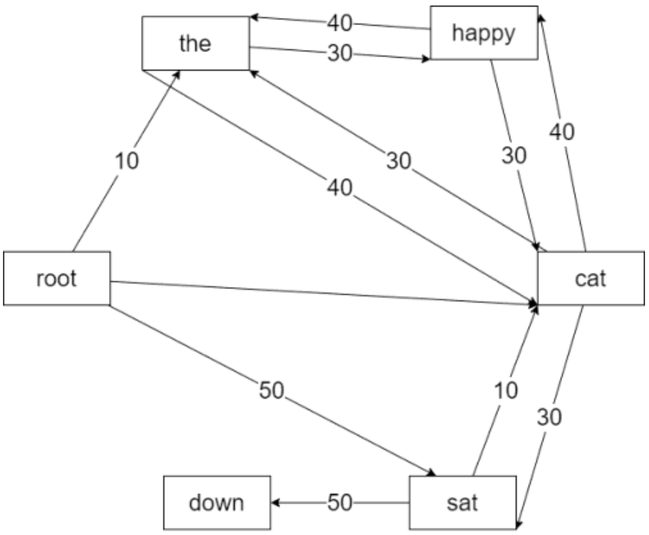
13

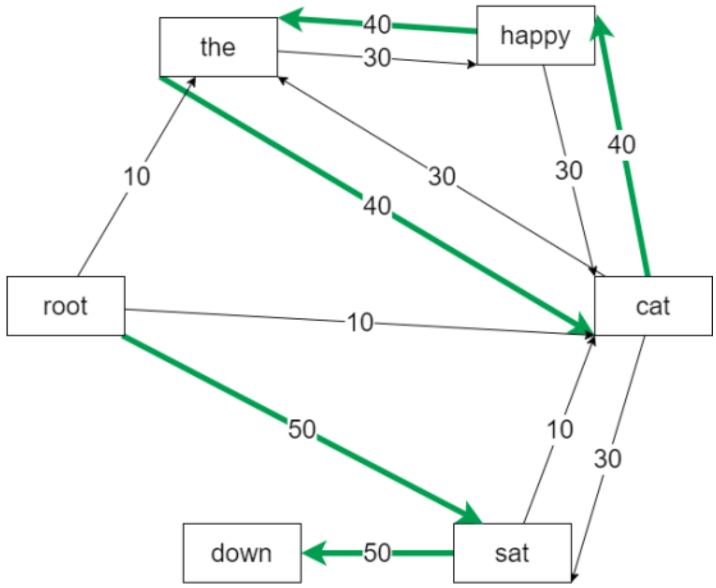

[root] what Is that dog chewing on over there ?

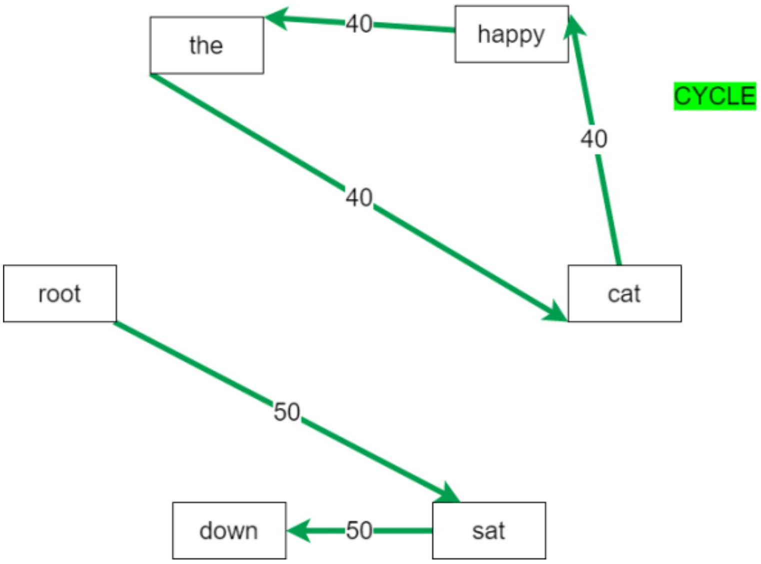

[root] what Is that dog chewing on over there ?

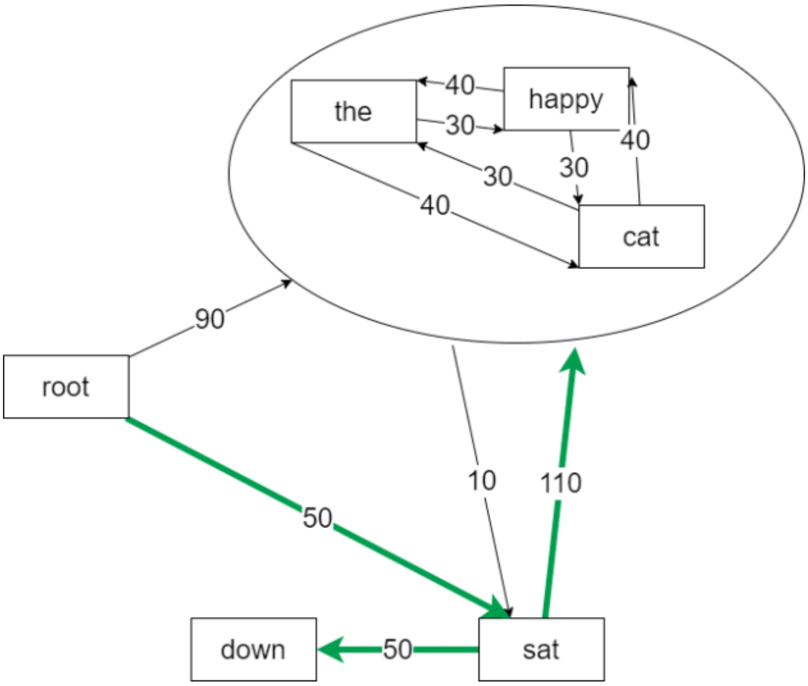
14

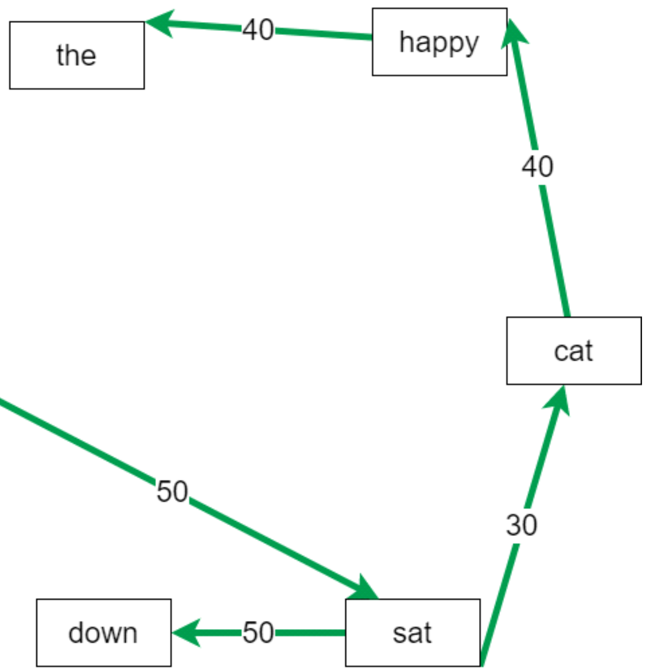
Using Chu-Liu Edmonds algorithm we find the tree with Max score











TOTAL SCORE : 50 + 50 + 30 + 40 + 40 = 210

15

Linear interpolation to handle rare/out-of-vocabulary words:

This strategy is not sufficient.
Linear interpolation extends for cases where higher n-gram probabilities are zeros.
It backs off and helps to generalize to more contexts that the model hasn't learned about.

This strategy falls short for words we haven't seen before.

In open vocabulary system these potential unknown words are modelled using <UNK>.

16

<S>	DT	NN	VB	PP	DT	NN
<S>	DT	NN	VB	PP	DT	NN
<S>	DT	JJ	NN	VB		

<u>TRANSITION PROBABILITIES</u>

	DT	NN	VB	PP	JJ
<S>	1	0	0	0	0
DT	0	0.8	0	0	0.2
NN	0	0	0.6	0	0
VB	0	0	0	0.67	0
PP	1	0	0	0	0
JJ	0	1	0	0	0

<u>UNIGRAM TRANSITION PROBABILITIES</u>

	c	P
<S>	3	0.157895
DT	5	0.263158
NN	5	0.263158
VB	3	0.157895
PP	2	0.105263
JJ	1	0.052632
	19	

<u>SMOOTHED TRANSITION PROBABILITIES</u>

	DT	NN	VB	PP	JJ
<S>	0.578947	0.078947	0.078947	0.078947	0.078947
DT	0.131579	0.531579	0.131579	0.131579	0.231579
NN	0.131579	0.131579	0.431579	0.131579	0.131579
VB	0.078947	0.078947	0.078947	0.413947	0.078947
PP	0.552632	0.052632	0.052632	0.052632	0.052632
JJ	0.026316	0.526316	0.026316	0.026316	0.026316

EMISSION PROBABILITIES

	the	cat	is	in	box	sat	on	mat	black
DT	1	0	0	0	0	0	0	0	0
NN	0	0.6	0	0	0.2	0	0	0.2	0
VB	0	0	0.5	0	0	0.5	0	0	0
PP	0	0	0	0.5	0	0	0.5	0	0
JJ	0	0	0	0	0	0	0	0	1

HMM TO DO POS TAGGING

	the	black	box	cat
DT	0.578			
NN	0			
VB	0			
PP	0			
JJ	0			

	the	black	box	cat
DT	0.578	0		
NN	0	0		
VB	0	0		
PP	0	0		
JJ	0	0.07*1*0.578=0.0405		

	the	black	box	cat
DT	0.578	0	0	
NN	0	0	0.526*0.2*0.0405=0.0043	
VB	0	0	0	
PP	0	0	0	
JJ	0	0.0405	0	

	the	black	box	cat
DT	0.578	0	0	0
NN	0	0	0.0043	0.0003
VB	0	0	0	0
PP	0	0	0	0
JJ	0	0.0405	0	0

DT	JJ	NN	NN

17

Identify constituents among the following:

Underlined phrases in A, B are not constituents

D. Put it over <u>on the table</u> [It is a constituent]

   Can be changed to Put it over there.

18

EM to update parameters and find t(souris|white)

INITIAL PARAMETERS:

t	souris	blanche
white	1/2	1/2
mouse	3/4	1/4

We have two sentences :

For k = 1 : (m = 2, l=2)
For k = 2 : (m = 1, l=1)

Initial q's are initialized uniformly

q(j|i,l,m)
j - ENGLISH WORD
i - FRENCH WORD

q(1|1,1,1)	1		
q(1|1,2,2)	1/2	q(2|1,2,2)	1/2
q(1|2,2,2)	1/2	q(2|2,2,2)	1/2

NEXT STEP:

δ(k,i,j)

δ(2,1,1)	1	
δ(1,1,1)	q(1|1,2,2) t(souris|white) / [(q(1|1,2,2) t(souris|white) + q(2|1,2,2) t(souris|mouse))] = (1/2)(1/2)/ [(1/2)(1/2) + (1/2)(3/4)]	2/5
δ(1,1,2)	q(2|1,2,2) t(souris|mouse) / [(q(1|1,2,2) t(souris|white) + q(2|1,2,2) t(souris|mouse))] =(1/2)(3/4)/ [(1/2)(1/2) + (1/2)(3/4)]	3/5
δ(1,2,1)	q(1|2,2,2) t(blanche|white) / [q(1|2,2,2) t(blanche|white) + q(2|2,2,2) t(blanche|mouse))] = (1/2)(1/2) / [(1/2)(1/2) + (1/2)(1/4)]	2/3
δ(1,2,2)	q(2|2,2,2) t(blanche|mouse) / [q(1|2,2,2) t(blanche|white) + q(2|2,2,2) t(blanche|mouse))] = (1/2)(1/4) / [(1/2)(1/2) + (1/2)(1/4)]	1/3

c(white, souris)	2/5		c(white)	2/5 + 2/3		t(souris|white)	c(white,souris)/c(white)
c(white, blanche)	2/3		c(mouse)	3/5 + 1/3 + 1			(2/5)/ (2/5) + (2/3)
c(mouse, souris)	3/5 + 1						(6/16)
c(mouse, blanche)	1/3						

PROBLEM :
Perform sentiment analysis on unlabelled dataset of restaurant reviews given movie review corpus with gold standard labels.

APPROACH:

1. Consider seed information in the form of keywords associated with gold standard labels from movie review corpus.
2. Bootstrapping initializes a learner with the above seed information.
3. Apply the learner in order to generate labels for the unlabelled restaurant review corpus.
4. Build a new learner from these bootstrapped labels.
5. Repeat 3-4