Q1 Constituent Parsing

20 Points

The problems in this section are based on the material covered in Week 6.

Q1.1 Context Free Grammars

5 Points

Suppose we have the following CFG:

$$\begin{array}{lcl} \text{S} & \rightarrow & \text{NP VP} \\ \text{VP} & \rightarrow & \text{VB NP} \\ \text{NP} & \rightarrow & \text{NP and NP} \\ \text{NP} & \rightarrow & \text{JJ NP} \\ \text{NP} & \rightarrow & \text{NN} \\ \text{JJ} & \rightarrow & \text{black} \\ \text{NN} & \rightarrow & \text{I} \\ \text{NN} & \rightarrow & \text{cats} \\ \text{NN} & \rightarrow & \text{dogs} \\ \text{VB} & \rightarrow & \text{like} \end{array}$$

How many possible parse trees are there for the sentence "I like black cats and dogs"? Show the derivations for each possible parse, ie. the rules used and the order they are used in.

There are 2 possible parse trees for this sentence.

Derivations for the first parse:

Level 0-1: S $\rightarrow$ NP VP
Level 1-2: NP $\rightarrow$ NN   VP $\rightarrow$ VB NP
Level 2-3: NN $\rightarrow I$ VB $\rightarrow like$ NP $\rightarrow$ NP and NP
Level 3-4: NP $\rightarrow$ JJ NP   NP $\rightarrow$ NN
Level 4-5: JJ $\rightarrow black$ NP $\rightarrow$ NN   NN $\rightarrow dogs$
Level 5-6: NN $\rightarrow cats$

The first parse would indicate that I like black cats and I also like dogs.

Derivations for the second parse:

Level 0-1: S $\rightarrow$ NP VP
Level 1-2: NP $\rightarrow$ NN   VP $\rightarrow$ VB NP
Level 2-3: NN $\rightarrow I$ VB $\rightarrow like$ NP $\rightarrow$ JJ NP
Level 3-4: JJ $\rightarrow black$ NP $\rightarrow$ NP and NP
Level 4-5: NP $\rightarrow$ NN   NP $\rightarrow$ NN
Level 5-6: NN $\rightarrow cats$ NN $\rightarrow dogs$

The second parse would indicate that I like cats and dogs that are black.

Q1.2 Context Free Grammars

5 Points

Suppose we have the following CFG:

$$\begin{array}{lcl}
\text{S} & \rightarrow & \text{NP VP} \\
\text{NP} & \rightarrow & \text{DT NN} \\
\text{NP} & \rightarrow & \text{DT NNS} \\
\text{VP} & \rightarrow & \text{VB NP} \\
\text{DT} & \rightarrow & \text{the} \\
\text{NN} & \rightarrow & \text{cat} \\
\text{NN} & \rightarrow & \text{dog} \\
\text{NNS} & \rightarrow & \text{cats} \\
\text{NNS} & \rightarrow & \text{dogs} \\
\text{VB} & \rightarrow & \text{see} \\
\text{VB} & \rightarrow & \text{sees}
\end{array}$$

This grammar can generate ungrammatical sentences, such as "the dog see the cat" and "the dogs sees the cat." What rule(s) would you need to add or change so that the grammar generates only grammatical sentences? List all the rules in your new grammar.

NN can only be followed by sees, and not by see. Similarly, NNS can only be followed by see, and not by sees. Hence, we first need to change VB. Then we need to change the rules VP to accommodate for two VB parses. We will also need to rename the NP rule so that we can change the rule for S such that it can yield grammatically correct sentences.

S $\rightarrow$ NP VPS
S $\rightarrow$ NPS VP
NP $\rightarrow$ DT NN
NPS $\rightarrow$ DT NNS
VP $\rightarrow$ VB NP
VP $\rightarrow$ VB NPS
VPS $\rightarrow$ VBS NP
VPS $\rightarrow$ VBS NPS
DT $\rightarrow$ the
NN $\rightarrow$ cat
NN $\rightarrow$ dog
NNS $\rightarrow$ cats
NNS $\rightarrow$ dogs
VB $\rightarrow$ see
VBS $\rightarrow$ sees

Q1.3 Probabilistic Context Free Grammars

5 Points

Suppose we have the following PCFG:

$$\begin{array}{lcll}
\text{S} & \rightarrow & \text{NP VP} & 1.0 \\
\text{NP} & \rightarrow & \text{DT JJ NN} & 1.0 \\
\text{VP} & \rightarrow & \text{VB} & 1.0 \\
\text{DT} & \rightarrow & \text{the} & 1.0 \\
\text{JJ} & \rightarrow & \text{happy} & 0.5 \\
\text{JJ} & \rightarrow & \epsilon & 0.5 \\
\text{NN} & \rightarrow & \text{cat} & 1.0 \\
\text{VB} & \rightarrow & \text{sat} & 1.0
\end{array}$$

Convert this PCFG into Chomsky normal form. If you need new non-terminal symbols, use X, Y, and Z. Make sure you do not change the probabilities of sentences generated by this grammar.

For CNY, we can have rules of the form: A $\rightarrow$ B C or A $\rightarrow$ a ; where B & C are non-terminals and not the start symbol, and a is a terminal but not empty-string ϵ

In the above grammar, we have 3 rules violating CNF conditions. The given PCFG can be converted to Chomsky Normal Form as following:

$$
\begin{array}{llll}
\text{S} & \rightarrow & \text{NP VB} & 1.0 \\
\text{NP} & \rightarrow & \text{DT X} & 0.5 \\
\text{NP} & \rightarrow & \text{DT NN} & 0.5 \\
\text{X} & \rightarrow & \text{JJ NN} & 1.0 \\
\text{DT} & \rightarrow & the & 1.0 \\
\text{JJ} & \rightarrow & happy & 1.0 \\
\text{NN} & \rightarrow & cat & 1.0 \\
\text{VB} & \rightarrow & sat & 1.0
\end{array}
$$

As we can see, the probabilities of the sentences generated by the CNF are the same.

Q1.4 Probabilistic Context Free Grammars

5 Points

Suppose we have the following PCFG:

$$
\begin{array}{llll}
\text{S} & \rightarrow & \text{NP VP} & 1.0 \\
\text{NP} & \rightarrow & \text{DT NN} & 1.0 \\
\text{VP} & \rightarrow & \text{MD VP} & 0.5 \\
\text{VP} & \rightarrow & \text{VB NP} & 0.5 \\
\text{DT} & \rightarrow & \text{the} & 1.0 \\
\text{NN} & \rightarrow & \text{cat} & 0.5 \\
\text{NN} & \rightarrow & \text{can} & 0.5 \\
\text{MD} & \rightarrow & \text{can} & 1.0 \\
\text{VB} & \rightarrow & \text{see} & 1.0
\end{array}
$$

Use the CKY algorithm to find the highest probability parse of the sentence "The cat can see the can". Show your work, ie. the dynamic programming table.

(Note that CKY for probabilistic PCFGs uses a three-dimensional dynamic programming table. Instead of trying to format a 3D table, simply use a 2D table, as in CKY for non-probabilistic CFGs, and show both the rule(s) being used and the associated probability in each table entry, as we do in lecture.)

	the	cat	can	see	the	can
the	DT(1.0)	NP(0.5)	–	–	–	**S**(0.0625)
cat		NN(0.5)	–	–	–	–
can			NN(0.5) **MD**(1.0)	–	–	VP(0.125)
see				VB(1.0)	–	VB(0.25)
the					DT(1.0)	NP(0.5)
can						**NN**(0.5) MD(1.0)

Using the CKY algorithm, we get the following parse:

S $\rightarrow$ NP VP

  NP $\rightarrow$ DT NN

    DT $\rightarrow$ the

    NN $\rightarrow$ cat

  VP $\rightarrow$ MD VP

    MD $\rightarrow$ can

    VP $\rightarrow$ VB NP

      VB $\rightarrow$ see

      NP $\rightarrow$ DT NN

        DT $\rightarrow$ the

        NN $\rightarrow$ can

Q2 Dependency Parsing

15 Points

The problems in this section are based on the material covered in Week 7.

Consider the sentence "I think I like cats."

Q2.1

5 Points

Parse this sentence by hand and list the dependency relations. You can use arrows, as we do in lecture, or simply list pairs of (head, dependent).

head		dependent
root	$\rightarrow$	think
think	$\rightarrow$	I
think	$\rightarrow$	like
like	$\rightarrow$	I
think	$\rightarrow$	cats

This gives the pairs (root, think), (think, I), (think, like), (like, I), (think, cats)

Q2.2

5 Points

Parse this sentence using shift-reduce parsing, the arc-standard transition system, and no arc labels. You can assume you have a perfet oracle that tells you the correct action to take at each time step. Show the stack, buffer, action, and relation set at each time step (you can use a table, as we do in lecture, or simply show a series of lists).

STACK	BUFFER	ACTION	RELATION
root	I think I like cats	shift	
root, I	think I like cats	shift	
root, I, think	I like cats	LeftArc	I $\leftarrow$ think
root, think	I like cats	shift	
root, think, I	like cats	shift	
root, think, I, like	cats	LeftArc	I $\leftarrow$ like
root, think, like	cats	shift	
root, think, like, cats		RightArc	like $\rightarrow$ cats
root, think, like		RightArc	think $\rightarrow$ like
root, think		RightArc	root $\rightarrow$ think
root		DONE	

Q2.3

5 Points

Parse this sentence again, this time using the arc-eager transition system. Show your work as before.

STACK	BUFFER	ACTION	RELATION
root	I think I like cats	shift	
root, I	think I like cats	LeftArc Pop	I $\leftarrow$ think
root	think I like cats	RightArc Push	root $\rightarrow$ think
root, think	I like cats	shift	
root, think, I	like cats	LeftArc Pop	I $\leftarrow$ like
root, think	like, cats	RightArc Push	think $\rightarrow$ like
root, think, like	cats	RightArc Push	like $\rightarrow$ cats
root, think, like, cats		reduce	
root, think, like		reduce	
root, think		reduce	
root		DONE	

Q3 Machine Translation

10 Points

The problems in this section are based on the material covered in Week 8.

Suppose we have a training corpus of three sentences:

1.	the	dog	2.	little	dog	3.	black	dog
	le	chien		petit	chien		chien	noir

Now suppose we have a word-level statistical machine translation model with the parameters $q(j|i,l,m)$ and $t(f_i|e_j)$ initialized uniformly using this corpus.

Q3.1

5 Points

Perform one iteration of expectation maximization to update the parameters using the training corpus. State the values of the q's and t's at the end of this iteration, plus the intermediate values you calculated for $p(a_i = j|f_i, e_j)$. You can simply state the values; no need to show the arithmetic.

Due to uniform initialization, all $q(j|i)$ values are $1/2$ and all $t(f_i|e_j)$ values are $1/4$

For all 3 sentences:

$p(a_1 = 1) = \frac{1}{2} \quad p(a_1 = 2) = \frac{1}{2} \quad p(a_2 = 1) = \frac{1}{2} \quad p(a_2 = 2) = \frac{1}{2}$

Expected counts:

$c(\text{the}) = 1 \quad c(\text{dog}) = 3 \quad c(\text{little}) = 1 \quad c(\text{black}) = 1$

Expected pair counts:

$c(\text{the, le}) = \frac{1}{2} \quad c(\text{the, chien}) = \frac{1}{2} \quad c(\text{the, petit}) = 0 \quad c(\text{the, noir}) = 0$
$c(\text{dog, le}) = \frac{1}{2} \quad c(\text{dog, chien}) = \frac{3}{2} \quad c(\text{dog, petit}) = \frac{1}{2} \quad c(\text{dog, noir}) = \frac{1}{2}$
$c(\text{little, le}) = 0 \quad c(\text{little, chien}) = \frac{1}{2} \quad c(\text{little, petit}) = \frac{1}{2} \quad c(\text{little, noir}) = 0$
$c(\text{black, le}) = 0 \quad c(\text{black, chien}) = \frac{1}{2} \quad c(\text{black, petit}) = 0 \quad c(\text{black, noir}) = \frac{1}{2}$

$$t_{MLE}(f|e) = \frac{c(e,f)}{c(e)}$$

$t(\text{le}|\text{the}) = \frac{1}{2} \quad t(\text{le}|\text{dog}) = \frac{1}{6} \quad t(\text{le}|\text{little}) = 0 \quad t(\text{le}|\text{black}) = 0$
$t(\text{chien}|\text{the}) = \frac{1}{2} \quad t(\text{chien}|\text{dog}) = \frac{1}{2} \quad t(\text{chien}|\text{little}) = \frac{1}{2} \quad t(\text{chien}|\text{black}) = \frac{1}{2}$
$t(\text{petit}|\text{the}) = 0 \quad t(\text{petit}|\text{dog}) = \frac{1}{6} \quad t(\text{petit}|\text{little}) = \frac{1}{2} \quad t(\text{petit}|\text{black}) = 0$
$t(\text{noir}|\text{the}) = 0 \quad t(\text{noir}|\text{dog}) = \frac{1}{6} \quad t(\text{noir}|\text{little}) = 0 \quad t(\text{noir}|\text{black}) = \frac{1}{2}$

$$q_{MLE}(j|i,l,m) = \frac{p(a_i = j|f_i, e_j)}{c(i,l,m)}$$

$q(1|1) = \frac{1}{2} \quad q(1|2) = \frac{1}{2} \quad q(2|1) = \frac{1}{2} \quad q(2|2) = \frac{1}{2}$

Q3.2

5 Points

Perform a second iteration of expectation maximization to update the parameters again. As before, state the values of the q's and t's at the end of this iteration, plus the intermediate values you calculated for $p(a_i = j|f_i, e_j)$.

For sentence 1:

$p(a_1 = 1) = \frac{1}{2} \quad p(a_1 = 2) = \frac{1}{2} \quad p(a_2 = 1) = \frac{1}{4} \quad p(a_2 = 2) = \frac{3}{4}$

For sentence 2:

$p(a_1 = 1) = \frac{1}{2}$ $p(a_1 = 2) = \frac{1}{2}$ $p(a_2 = 1) = \frac{1}{4}$ $p(a_2 = 2) = \frac{3}{4}$

For sentence 3:
$p(a_1 = 1) = \frac{1}{2}$ $p(a_1 = 2) = \frac{1}{2}$ $p(a_2 = 1) = \frac{3}{4}$ $p(a_2 = 2) = \frac{1}{4}$

$$\text{Updated values for } t_{MLE}(f|e):$$

$$\begin{array}{llll}
t(\text{le}|\text{the}) = \frac{1}{2} & t(\text{le}|\text{dog}) = \frac{1}{6} & t(\text{le}|\text{little}) = 0 & t(\text{le}|\text{black}) = 0 \\
t(\text{chien}|\text{the}) = \frac{1}{2} & t(\text{chien}|\text{dog}) = \frac{1}{2} & t(\text{chien}|\text{little}) = \frac{1}{2} & t(\text{chien}|\text{black}) = \frac{1}{2} \\
t(\text{petit}|\text{the}) = 0 & t(\text{petit}|\text{dog}) = \frac{1}{6} & t(\text{petit}|\text{little}) = \frac{1}{2} & t(\text{petit}|\text{black}) = 0 \\
t(\text{noir}|\text{the}) = 0 & t(\text{noir}|\text{dog}) = \frac{1}{6} & t(\text{noir}|\text{little}) = 0 & t(\text{noir}|\text{black}) = \frac{1}{2}
\end{array}$$

$$\text{Updated values for } q_{MLE}(j|i, l, m):$$

$$q(1|1) = \frac{1}{2} \qquad q(1|2) = \frac{5}{12} \qquad q(2|1) = \frac{1}{2} \qquad q(2|2) = \frac{7}{12}$$

Q4 Late Penalty

0 Points

This problem intentionally left blank.

Homework 3 Written

UNGRADED

STUDENT

Adithya Iyer

TOTAL POINTS

- / 45 pts

QUESTION 1

Constituent Parsing — 20 pts

- 1.1 Context Free Grammars — 5 pts
- 1.2 Context Free Grammars — 5 pts
- 1.3 Probabilistic Context Free Grammars — 5 pts
- 1.4 Probabilistic Context Free Grammars — 5 pts

QUESTION 2

Dependency Parsing — 15 pts

- 2.1 (no title) — 5 pts
- 2.2 (no title) — 5 pts
- 2.3 (no title) — 5 pts

QUESTION 3

Machine Translation — 10 pts

- 3.1 (no title) — 5 pts
- 3.2 (no title) — 5 pts

QUESTION 4

Late Penalty — 0 pts