NLP 1 - Assignment 3

Selene Baez Santamaria

December 1, 2016

Exercise 1. Context free grammar

(a) Convert the grammar to Chomsky Normal Form.

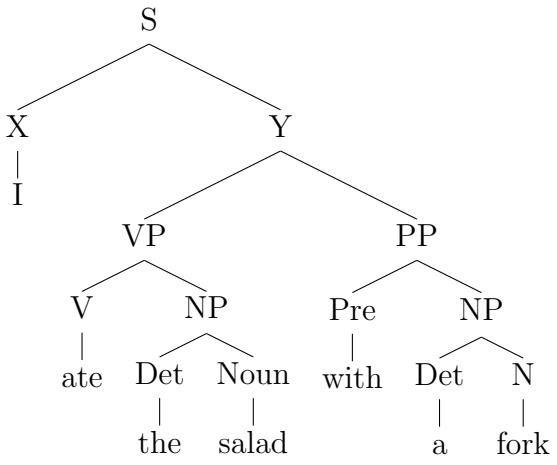
S → NP VP	V → *ate*
S → X Y	Det → *the* | *a*
NP → Det N	N → *fork* | *salad*
VP → V NP	Pre → *with*
VP → V	Y → VP PP
PP → Pre NP	X → *I*

(b) Use the CKY algorithm to parse the sentence, representing the CKY chart in matrix form.
I ate the salad with a fork

I	**ate**	**the**	**salad**	**with**	**a**	**fork**
X → *I*	∅	∅	∅	∅	∅	S → X Y
	V → *ate*	∅	VP → V NP	∅	∅	Y → VP PP
		Det → *the*	NP → Det N	∅	∅	∅
			N → *salad*	∅	∅	∅
				Pre → *with*	∅	PP → Pre PP
					Det → *a*	NP → Det N
						N → *fork*

(c) Parsed trees corresponding to all possible complete analysis of **I ate the salad with a fork**

We get only one complete analysis of S, being:



Exercise 2. Tree corpus

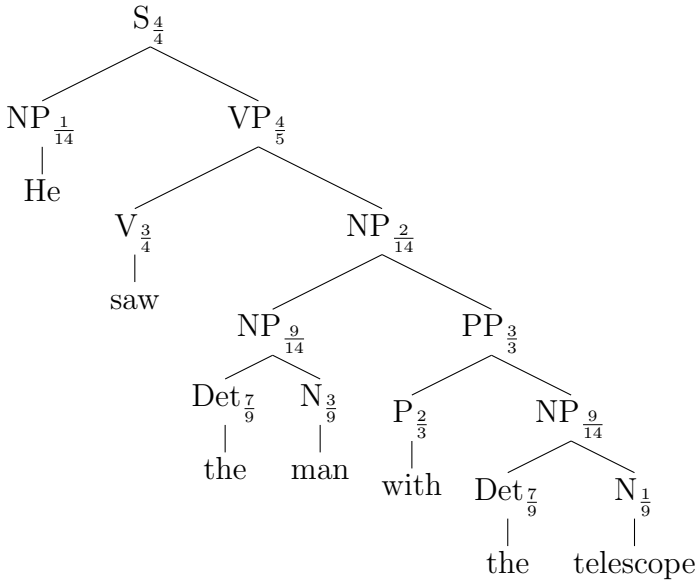
(a) Derive a PCFG. Write down the rules and calculate their probabilities

Written as Chomsky Normal Form:

$$
\begin{aligned}
\text{S} &\rightarrow \text{NP VP} : \tfrac{4}{4} \\
\text{VP} &\rightarrow \text{VP PP} : \tfrac{1}{5} \\
\text{VP} &\rightarrow \text{V NP} : \tfrac{4}{5} \\
\text{NP} &\rightarrow \text{Det N} : \tfrac{9}{14} \\
\text{NP} &\rightarrow \text{NP PP} : \tfrac{2}{14} \\
\text{PP} &\rightarrow \text{P NP} : \tfrac{3}{3} \\
\text{NP} &\rightarrow \textit{She} \mid \textit{Here} \mid \textit{He} : \tfrac{1}{14} \\
\text{V} &\rightarrow \textit{saw} : \tfrac{3}{4} \\
\text{V} &\rightarrow \textit{is} : \tfrac{1}{4}
\end{aligned}
\qquad
\begin{aligned}
\text{Det} &\rightarrow \textit{the} : \tfrac{7}{9} \\
\text{Det} &\rightarrow \textit{a} : \tfrac{2}{9} \\
\text{N} &\rightarrow \textit{man} : \tfrac{3}{9} \\
\text{N} &\rightarrow \textit{girl} : \tfrac{2}{9} \\
\text{N} &\rightarrow \textit{distance} \mid \textit{telescope} \mid \textit{guitar} \mid \textit{flower} : \tfrac{1}{9} \\
\text{P} &\rightarrow \textit{from} : \tfrac{1}{3} \\
\text{P} &\rightarrow \textit{with} : \tfrac{2}{3}
\end{aligned}
$$

(b) Possible trees for **He saw the man with the telescope**

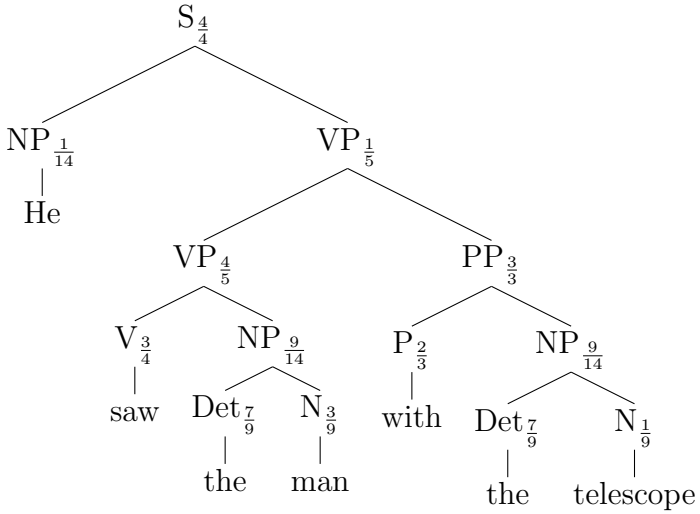
First option:



Probability:

$$P = \frac{4}{4} \cdot \frac{1}{14} \cdot \overbrace{\frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{14}}^{distinct} \cdot \frac{9}{14} \cdot \frac{7}{9} \cdot \frac{3}{9} \cdot \frac{3}{3} \cdot \frac{2}{3} \cdot \frac{9}{14} \cdot \frac{7}{9} \cdot \frac{1}{9} = \frac{1}{5 \cdot 14 \cdot 9 \cdot 14 \cdot 3} = \frac{1}{26460}$$

Second option:

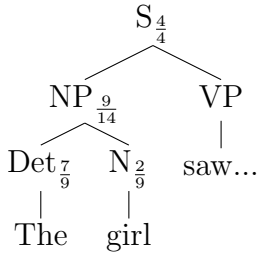


$$P = \underbrace{\frac{4}{4} \cdot \frac{1}{14}}_{sameasbefore} \cdot \overbrace{\frac{1}{5} \cdot \frac{4}{5} \cdot \frac{3}{4}}^{distinct} \cdot \underbrace{\frac{9}{14} \cdot \frac{7}{9} \cdot \frac{3}{9} \cdot \frac{3}{3} \cdot \frac{2}{3} \cdot \frac{9}{14} \cdot \frac{7}{9} \cdot \frac{1}{9}}_{sameasbefore} = \frac{1}{5 \cdot 5 \cdot 14 \cdot 2 \cdot 3 \cdot 9} = \frac{1}{18900}$$

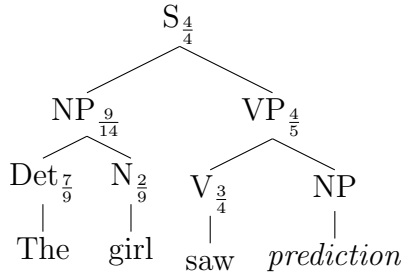
Since $26,460 > 18,900$, the second tree is more likely.

(c) Most likely completion suggestion for **The girl saw**

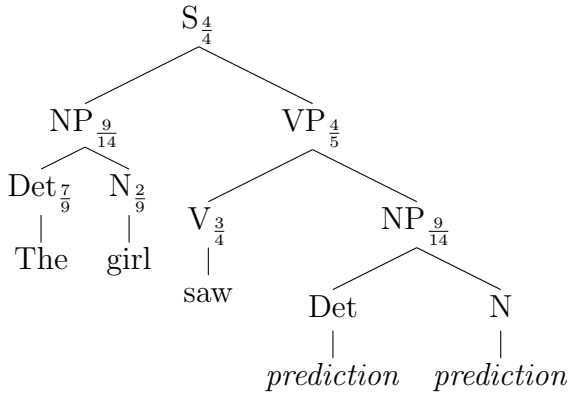
In order to make a prediction, we will first fit the existing words into a plausible tree. First, we expand $S \rightarrow NPVP$, because it is the only possibility. Moreover, the verb must be part of VP and the noun must be part of NP. Lastly, we expand NP with rule $NP \rightarrow DetN$



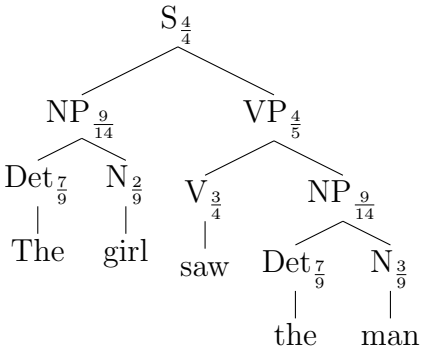
From here on we apply rules for expanding the branches by choosing the ones with highest probability. As such, we expand $VP \rightarrow VNP$ because it scores $\frac{4}{5}$



Then we select $NP \rightarrow DetN$ with $\frac{9}{14}$



Finally, we take $Det \rightarrow the$ and $N \rightarrow man$



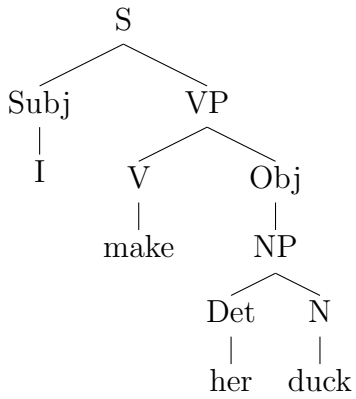
Hence, the most likely suggestion would be **The girl saw the man**. The technique for choosing the most likely rule at each step works because all the rules have probability < 1. Therefore, each rule expansion reduces the likelihood of the sentence, thus favoring short sentences. Furthermore, at each expansion, a larger rule probability produces a larger sentence probability, since both are proportional.

Exercise 3. Probabilistic Context Free Grammar

(a) Find the most probable parse for the sentence **I make her duck**

I	**make**	**her**	**duck**
Subj $\rightarrow$ *I* (0.3)	$\emptyset$	S $\rightarrow$ Subj VP (0.018)	S $\rightarrow$ Subj VP (0.00288) S $\rightarrow$ Subj VP (0.018) S $\rightarrow$ Subj VP (0.00216)
	V $\rightarrow$ *make* (0.6)	VP $\rightarrow$ V Obj (0.06)	VP $\rightarrow$ V Small (0.0096) VP $\rightarrow$ V Obj (0.06) VP $\rightarrow$ V Obj Obj (0.0072)
		Obj $\rightarrow$ *her* (0.2) Det $\rightarrow$ *her* (1.0)	Small $\rightarrow$ Obj V (0.08) NP $\rightarrow$ Det N (0.25) Subj $\rightarrow$ NP (0.175) Obj $\rightarrow$ NP (0.2)
			N $\rightarrow$ *duck* (0.5) V $\rightarrow$ *duck* (0.4) NP $\rightarrow$ N (0.25) Subj $\rightarrow$ NP (0.175) Obj $\rightarrow$ NP (0.2)

The most probable parse for this sentence corresponds to the green parse with this tree:

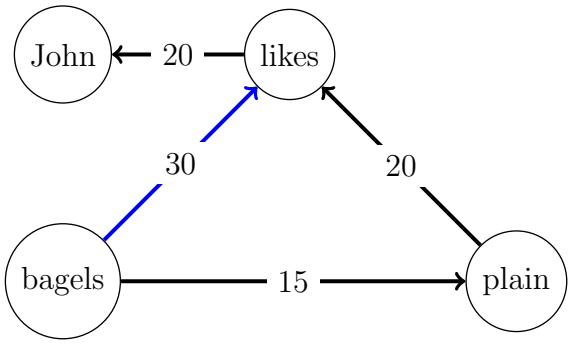


The semantic meaning is equivalent to "I make a duck. The duck is her's".

Exercise 4. Dependency parsing

(a) Explain step by step how the CLE algorithm is applied

(a) Greedily select the incoming edge with the highest score, for each node.

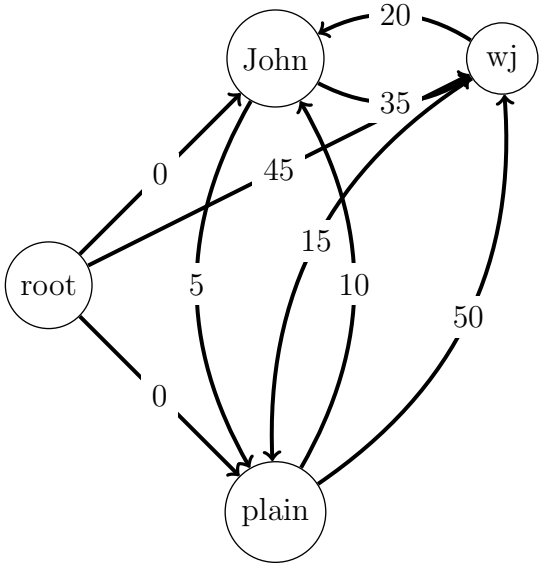


(b) We note there is a cycle and choose to contract the nodes connected by the edge in blue. We call this group w_j and recalculate its incoming and outgoing edges:

Incoming	likes $\rightarrow$ bagels	bagels $\rightarrow$ likes
root $\rightarrow$	$15 + 30 = 45$	$0 + 10 = 10$
John $\rightarrow$	$5 + 30 = 35$	$15 + 10 = 25$
plain $\rightarrow$	$20 + 30 = 50$	$5 + 10 = 15$

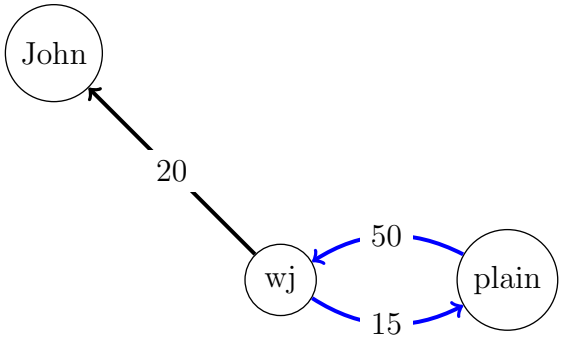
Outcoming	likes	bagels
John $\leftarrow$	20	5
plain $\leftarrow$	5	15

(c) The maximum incoming and outcoming edges per external node are marked in red. The new graph looks as follows:



	root	John	wj	plain
John	0	-	20	10
wj	45	35	-	50
plain	0	5	15	-

(d) We apply CLE recursively and go back to a) with the new graph as basis

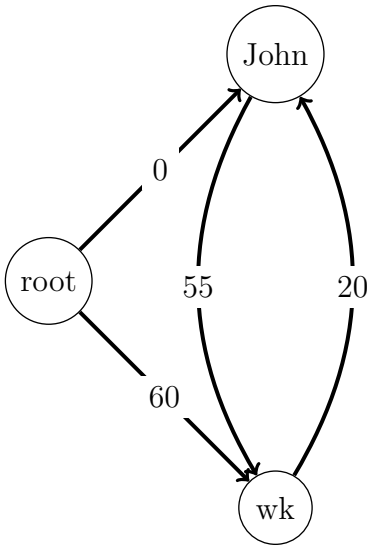


(e) Again, we have a cycle. We contract the nodes selected in blue, call it w_k and recalculate the incoming and outgoing nodes.

Incoming	wj $\rightarrow$ plain	plain $\rightarrow$ wj
root $\rightarrow$	$45 + 15 = 60$	$0 + 50 = 50$
John $\rightarrow$	$35 + 15 = 50$	$5 + 50 = 55$

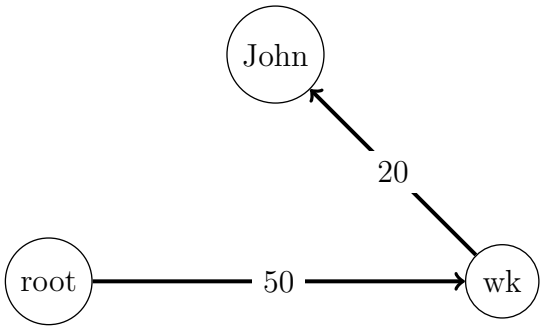
Outcoming	wj	plain
John $\leftarrow$	20	0

(f) The resulting graph is:



	root	John	wk
John	0	-	20
wk	60	55	-

(g) Going back to a) once more:

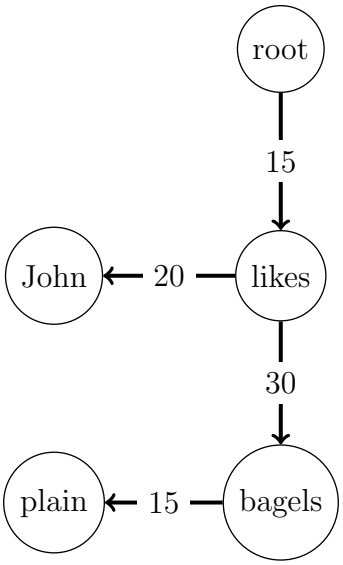


(h) Since there are no more cycles, we are done.

(b) Show the resulting MST

We interpret the previous graph and reconstruct it. Regarding wk outgoing edges, we backtrack that wk → John comes from wj → likes → John. Thus we include **likes** → **John**.

Regarding incoming edges, we backtrack that root → wk comes from root → wj → plain. At the same time, wj → plain comes from bagels → plain. Thus we include **root** → **likes** and **bagels** → **plain**



Exercise 5. Viterbi. Tag the sentence: *The healthy man the lifeboats*

(a) Hand simulate the Viterbi algorithm using the given transition and emission probabilities

(b) Give the joint probability