

Q.1. The simplest augmentation of the context-free grammar is the Probabilistic Context PCFG Free Grammar (PCFG), also known as the Stochastic Context-Free Grammar SCFG (SCFG), first proposed by Booth (1969).

Grammar		Lexicon	
$S \rightarrow NP VP$	[.80]	$Det \rightarrow that [.10] \mid a [.30] \mid the [.60]$	
$S \rightarrow Aux NP VP$	[.15]	$Noun \rightarrow book [.10] \mid flight [.30]$	
$S \rightarrow VP$	[.05]	$\mid meal [.15] \mid money [.05]$	
$NP \rightarrow Pronoun$	[.35]	$\mid flights [.40] \mid dinner [.10]$	
$NP \rightarrow Proper-Noun$	[.30]	$Verb \rightarrow book [.30] \mid include [.30]$	
$NP \rightarrow Det Nominal$	[.20]	$\mid prefer; [.40]$	
$NP \rightarrow Nominal$	[.15]	$Pronoun \rightarrow I [.40] \mid she [.05]$	
$Nominal \rightarrow Noun$	[.75]	$\mid me [.15] \mid you [.40]$	
$Nominal \rightarrow Nominal Noun$	[.20]	$Proper-Noun \rightarrow Houston [.60]$	
$Nominal \rightarrow Nominal PP$	[.05]	$\mid NWA [.40]$	
$VP \rightarrow Verb$	[.35]	$Aux \rightarrow does [.60] \mid can [.40]$	
$VP \rightarrow Verb NP$	[.20]	$Preposition \rightarrow from [.30] \mid to [.30]$	
$VP \rightarrow Verb NP PP$	[.10]	$\mid on [.20] \mid near [.15]$	
$VP \rightarrow Verb PP$	[.15]	$\mid through [.05]$	
$VP \rightarrow Verb NP NP$	[.05]		
$VP \rightarrow VP PP$	[.15]		
$PP \rightarrow Preposition NP$	[1.0]		

Draw the possible parse trees for deriving a sentence “Book the dinner flight”. Explain how PCFG is used for disambiguation.

Q.2.

$S \rightarrow NP VP$	$V \rightarrow swam \mid ran \mid flew$
$VP \rightarrow V NP$	$VP \rightarrow swam \mid ran \mid flew$
$VP \rightarrow VP PP$	$D \rightarrow the \mid a \mid an$
$NP \rightarrow D N$	$N \rightarrow pilot \mid plane$
$NP \rightarrow NP PP$	$NP \rightarrow Edinburgh \mid Glasgow$
$PP \rightarrow P NP$	$P \rightarrow to$

CYK algorithm uses bottom-up parsing technique. Apply the CYK algorithm on the sentence “The pilot flew the plane to Glasgow” using the grammar G and state whether the sentence is recognized or not.

Q.3 Find minimum edit distance between “STALL” and “TABLE”.

Q.4. Classify give test statement using Naïve Bayes algorithm.

	docID	words in document	in c = China?
Training set	1	Chinese Beijing Chinese	yes
	2	Chinese Chinese Shanghai	yes
	3	Chinese Macao	yes
	4	Tokyo Japan Chinese	no
Test set	5	Chinese Chinese Chinese Tokyo Japan	?

Q.5 Find out pos tagging using Viterbi’s algorithm.

Book a car.  
Park the car.  
The book is in the car.  
The car is in a park.