Total Marks | Moderator
Question no. | Re-Assessor

Space for Marks	Question No.	START WRITING HERE

Q1(a) What is Pragmatics

Ans. Pragmatics is the study of how context contributes to meaning. The field of study evaluates how human language is utilized in social interaction, as well as the relationship between the interpreter & the interpreted.

Q1(b) Using the UPenn Treebank tag set, tag the part of speech to the given statement.

The grand jury commented on a number of other topics

Soln The/DT grand/JJ jury/NN commented/VBD on/IN a/DT number/NN of/IN other/JJ topics/NNS

Q1(c) What is semantic analysis

Soln Semantic Analysis is a subfield of Natural language Processing (NLP) that attempts to understand the meaning of natural language.

Total Marks of Question no.	Examiner	
	Moderator	
	Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Q1(d) Convert the following CFG into CNF form

$S \to A | cB | c$

$A \to Bc | b$

$B \to cAA | \epsilon$

Sol^n Remove Null Production $B \to \epsilon$

$S \to A | cB | c$

$A \to Bc | b | c$

$B \to cAA$

Remove Unit Production

$S \to Bc | cB | c | b$

$A \to Bc | b$

$B \to cAA$

Add new variable $X \to AA$, $Y \to c$

$$\boxed{\begin{array}{l} S \to BY | YB | c | b \\ A \to BY | b \\ B \to YX \\ X \to AA \\ Y \to c \end{array}}$$

Total Marks of Question no.	Examiner	
	Moderator	
	Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Q1 (e) List any 4 applications of NLP

Ans 1) Sentiment Analysis
2) Text Summarization
3) Paraphrasing
4) Question - Answer System

Q1(f) What is reference resolution?

Ans. It is defined as the task of determining what entities are referred to by which linguistic expression. Interpretation of the sentences from any discourse is another important task & to achieve this we need to know who or what entity is being talked about, here interpretation reference is the key element

Q1(g) Differentiate between

Top Down Parsing	Bottom-Up Parsing
① Top Down approach starts evaluating the parse tree from the top & move downward for parsing other nodes	① Bottom Up approach starts everything evaluating the parse tree from the lowest level of the tree & move upwards for parsing the node
② Attempts to find left most derivation for a given string	Bottom up parsing attempts to reduce using reverse RMD.

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

③ It performs removal of left recursion & left factoring	③ It does not performs removal of left recursion & left factoring.
④ Eg- Predictive Parser, Recursive Descent Parser	④ Eg- LR Parser, Operator-Precedence Parser

Q1(h) What is syntax analysis?

Sol^n Syntax Analysis / parsing is the process of analyzing natural language with the rules of a formal grammar. Grammatical rules are applied to categories & groups of words, not individual words.

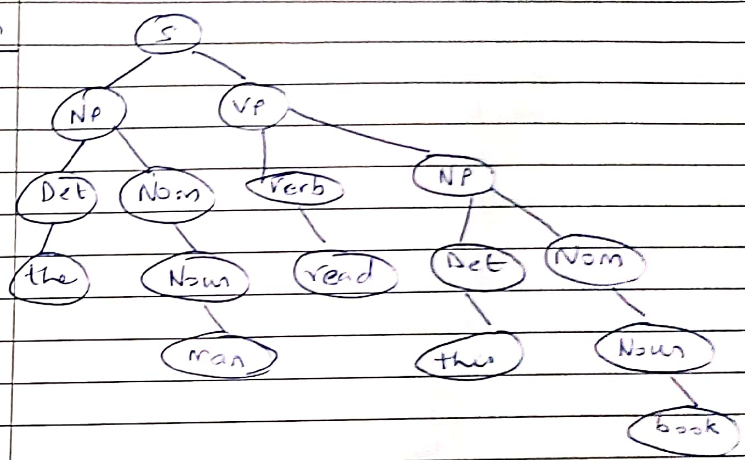
Q2(a) Construct a parse tree for the following sentence using CFG rules — The man read this book

S → NP VP	Det → the, this
S → VP	Noun → book, man
NP → Det Nom	Verb → book, read
Nom → Noun	
VP → verb NP	
Verb → read	

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Sol^n



Q2(b) Explain with suitable example following relationship between word meaning:-

Sol^n ① Homonymy

It is defined as a relation that holds between words that have the same form with unrelated meaning

Eg Bat (wooden stick) vs Bat (flying mammal)

② ~~Homophones are the words with same orthographic form but different meaning~~

② Homophones are the words with the same pronunciation but different spelling

Eg write vs right, piece vs peace

③ Homograph are the lexeme with same orthographic form but different meaning. Eg-Bass

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

(4) Polysemy – Multiple related meaning within a single lexeme
Eg – The bank was constructed in 1875 out of local red brick
I withdrew the money from the bank.

(5) Synonym – Two lexemes are synonym if they can be successfully substituted for each other in all situation.
Eg – sofa / couch
big / large

(6) Antonym – Senses that are opposite with respect to one feature of their meaning.
Eg – dark / light
short / long

Q(c) What is machine translation in application. Explain

Ans Machine translation is a procedure where a computer software translates text from one language to another without human contribution. It using softwares to translate the text from one source language to another target language

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Different type of machine translation in NLP

(1) Statistical Machine translation
It depends on the investigation of huge volumes of bilingual content. It expects to decide the correspondence b/w word from source language & a word from objective language. Eg – Syntax based translation, Hierarchical phrase based translation, word based translation.

(2) Rule-based Machine Translation
It translates the basics of grammatical rules. It directs a grammatical examination of the source language & objective language to create the translated sentence.

(3) Hybrid Machine Translation
It uses a translation memory, making it unquestionably more successful regarding quality.

(4) Neural Machine Translation
It relies upon neural network model to build statistical models with the end goal of translation.

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Q3(a) Based on the given state transition & emission probability matrix assign POS to the statement "Time flies like an arrow"

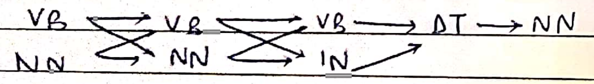
Emission Probability Matrix

	Time	flies	like	an	arrow
VB	0.1	0.2	0.2	0	0
NN	0.1	0.1	0	0	0.1
IN	0	0	0.25	0	0
DT	0	0	0	0.5	0

State Transition Matrix

	VB	NN	IN	DT	</s>
<S>	0.2	0.8	0	0	0
VB	0	0.3	0.2	0.5	0
NN	0.4	0.5	0.1	0	0
IN	0	0.75	0	0.25	0
DT	0	1	0	0	0

Step 1 Time flies like an arrow



Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Step 2 Time as Noun

$P(N \mid Time, <s>) = P(Time \mid N) \times P(N \mid <s>)$
$= 0.1 \times 0.8$
$= 0.08$

Time as Verb

$P(V \mid Time, <s>) = P(Time \mid V) \times P(V \mid <s>)$
$= 0.1 \times 0.2$
$= 0.02$

$\because 0.08 > 0.02$
$\therefore$ Time will be Noun.

Step 3 flies as Noun

$P(N \mid flies, N) = P(flies \mid N) \times P(N \mid N)$
$= 0.1 \times 0.5$
$= 0.05 \times 0.08$
$= 0.04$

flies as verb

$P(V \mid flies, N) = P(flies \mid V) \times P(V \mid N)$
$= 0.2 \times 0.4$
$= 0.08 \times 0.08$
$= 0.0064$

$0.0064 > 0.004$
$\therefore$ flies will be verb

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Step 4 Like as verb

$P(V|like, V) = P(like|V) \times P(V|V)$

$= 0.2 \times 0.01$

$= 0.002 \times 0.0064$

$= 0.0000128$

like as preposition

$P(IN|like, V) = P(like|N) \times P(IN|N)$

$= 0.25 \times 0.2$

$= 0.50 \times 0.0064$

$= 0.00032$

$0.00032 > 0.0000128$

∴ like is Preposition

Time	flies	like	an	arrow
NN	VB	IN	DT	NN

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Q2(b) What are five types of referring expressions? Explain with the help of example

Ans Referring Expression—

1) Indefinite Noun Phrases

It represents the entities that are new to the hearer into the discourse content

Eg- Ram had gone around one day to bring him some food - some is an indefinite references.

2) Definite Noun Phrases

It represents the entities that are not new or identifiable to hearer into the discourse content.

eg- I used to read the Times of India

The Times of India is a definite reference

3) Pronouns

It is a form of definite references. For eg-

Ram laughed as loud as he could.

The word he represents pronoun referring expression.

4) Demonstratives

These demonstrate & behave differently than simple definite pronoun.

For eg- this and that are demonstrative pronouns.

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

(5) Names

It is the simplest type of referring expression. It can be the name of a person, organization & location also. for eg. Ram is the name - referring expression.

Q4(a) Using CKY algorithm, find the possible parse tree for the following statement "A pilot likes flying planes

S → NP VP

VP → VBG NNS

VP → VBZ VP

NP → DT NN

NP → JJ NNS

DT → a

NN → pilot

VBZ → likes

VBG → flying

JJ → flying

NNS → planes

Soln

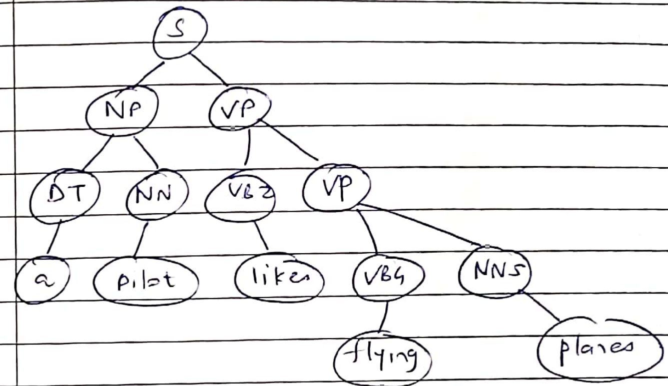
A	pilot	likes	flying	planes
DT x_{01}	NP x_{02}	— x_{03}	— x_{04}	S x_{05}
	NN x_{12}	— x_{22}	— x_{24}	— x_{25}
		VBZ x_{33}	— x_{34}	VP x_{35}
			VBG, JJ x_{44}	VP, NP x_{45}
				NNS x_{55}

$x_{02} \rightarrow x_{01}, x_{12}$

$x_{13} \rightarrow x_{12}\ x_{23}$

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE



Q4(b) Analyse a caseStudy for Sentiment Analysis as an application of NLP

Ans Sentiment analysis also called as opinion mining or emotion detection. It is used to determine the opinion, emotions, & attitude of a writer.

- It uses natural language processing, computational linguistic to systematically extract emotions, sentiments, opinion.
- With the increase in the social media platform, there is an increase in the number of people expressing their sentiments & opinion, labelling these sentiment can be very useful for people who are looking forward to using the opinion in order to improve their product, services etc.

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Sentiment Analysis on Twitter Data has many application — In Business (to understand customers view/feeling) In Politics — used to keep track of political views to detect consistency / inconsistency between statement. In Public Action — is used for monitoring & analyzing social phenomenon for predicting potentially dangerous situation

Steps to perform sentiment analysis —

1. Gather relevant data using twitter streaming API & gather tweets containing keywords, brands, hashtags etc.
2. Clean the data using preprocessing technique like removing emojis, special character, extra blank space.
3. Sentiment analysis using machine learning model
4. Analyze the twitter data using sentiment analysis model
5. Visualise the result using data visualization tools.

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Q4(c) Explain Anaphora Resolution with the help of Hobb's algorithm

Ans Anaphora Resolution is a problem of resolving what a pronoun or a noun phrases refers to.

Eg. 1) John helped Mary
2) He was kind

1 and 2 are utterances and together they form a discourse

— As human, readers and listeners can quickly & unconsciously work out that the pronoun "he" in utterance 2 refers to 'John" in 1. The underlying process of how this is done is yet unclear. especially when we encounter more complex sentences.

— Anaphore resolution is the process of interpreting the link between the anaphor (ie the repeated reference) and its antecedent (ie previous mention of the entity)

— The task of locating all expressions that are coreferential with any of the entities identified in the text is known as coreference resolution, and it occurs when two or more expression in the text relate to the same person or object.

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Gender agreement – In poem Jack & Jill
went up the hill, to fetch pail of water.
Jack fell down & broke his crown &
Jill came tumbling after.
here his refers to Jack not Jill as
Jill is a girl.

- Pronouns can only go a few sentences back, & entities closer to the referring phrases are more important than those further away, which finally leaves us ~~too~~ with the only possible solution is Jack. This property is known as recency property.

- Hobb's algorithm is used for pronoun resolution. The algorithm is mainly based on ~~the~~ syntactic ~~phrase~~ parse tree of the sentences.

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Q5(a) What is word sense disambiguation? Explain different ways to identify correct sense of an ambiguous word.

Ans In natural language, the meaning of a word may vary as per its usage in sentences & the context of the text. Word sense disambiguation involves interpreting the meaning of a word based upon the context of its occurrence in a text.

- For eg – the word 'Bark' may mean "the sound made by a dog" or the outermost layer of a tree. Similarly the word 'rock' may mean 'a stone' or 'a genre of music' so the accurate meaning depends on the context of the sentence.

- Thus the ability of machine to overcome the ambiguity involved in identifying the meaning of a word based on its usage & context is called word sense disambiguation.

- The task of disambiguation is to determine which of the senses of an ambiguous word is invoked in a particular use of word.

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks | Question No | START WRITING HERE

Different methods to find correct sense of words are—

① Knowledge Based Approaches

1→ Overlap Based Approach

- Requires a machine readable dictionary (MRD)
- Find the overlap b/w the features of different senses of an ambiguous word (sense bag) & features of the words in its context (context bag)
- The features could be sense definitions, eg sentences, hypernyms etc.
- The features could also be given weights
- The sense which has the maximum overlap is selected as the contextually appropriate sense

2→ Lesk's Algorithm

It has two groups - Sense bag - contains the words in the definition of a candidate sense of the ambiguous words

Context bag - It contains the words in the definition of each sense of each context word

- Eg on burning coal we get ash.
- Here Ash could have 3 senses —

① Trees of the olive family
② The solid residue left when combustible material is burned
③ To convert into ash.

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks | Question No | START WRITING HERE

Coal

Sense 1 - A piece of glowing carbon or burnt wood

2) charcoal

3) A black solid combustible substance formed by the partial decomposition of vegetable matter used as fuel for burning

So here second definition of ash is having correct sense

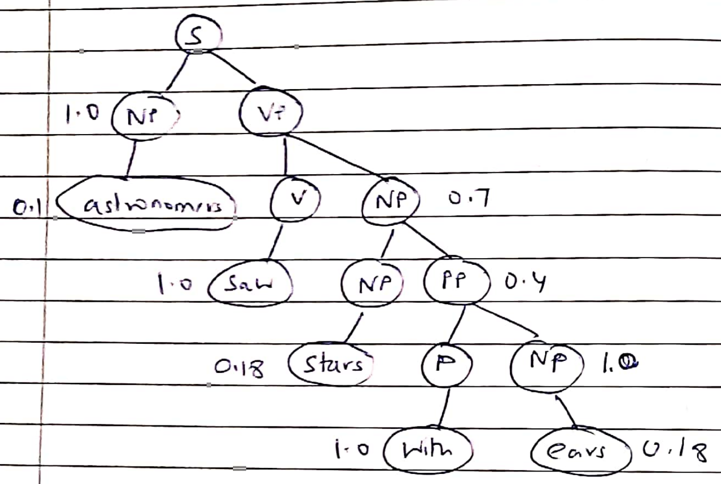
Q5(b) Find the probability of the sentence "Astronomers saw stars with ears" using the given PCFG

S → NP VP 1.0
VP → V NP 0.7
VP → VP PP 0.3
PP → P NP 1.0
P → with 1.0
V → saw 1.0
NP → NP PP 0.4
NP → astronomers 0.1
NP → ears 0.18
NP → saw 0.04
NP → stars 0.18
NP → telescope 0.1

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Sol^n **Parse Tree -1**



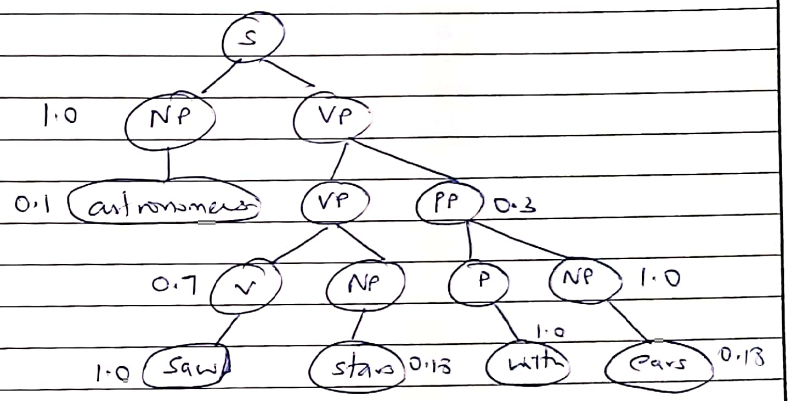
$P1 = 1 \times 0.1 \times 0.7 \times 1 \times 0.4 \times 0.18 \times 1 \times 1 \times 0.18$

$\boxed{P1 = 0.0009072}$

Total Marks of Question no.		Examiner	
		Moderator	
		Re-Assessor	

Space for Marks	Question No.	START WRITING HERE

Parse Tree 2



$P_2 = 1 \times 0.1 \times 0.3 \times 0.7 \times 1.0 \times 1 \times 0.18$
$\times 1.0 \times 0.18$

$\boxed{P_2 = 0.00068}$

So Probability of given statement is

$P(s) = P1 + P_2$

$P(s) = 0.0009072 + 0.00068$

$\boxed{P(s) = 0.00158}$