1.Discuss Lexicon, Mephitic, Orthographic ?

Lexicon: In NLP, a lexicon refers to a collection of words or phrases along with their associated information, such as part-of-speech tags, semantic categories, or pronunciation details. Lexeme is small representation in lexicon.

Mephitic: The term "mephitic" is not commonly used in the context of NLP. However, outside of NLP, "mephitic" typically refers to something that is foul-smelling or toxic. language that is considered unpleasant, offensive, or highly negative in sentiment.

Orthographic: Orthography refers to the conventional system of writing a language using a standardized set of symbols or characters. In NLP, orthographic features are concerned with the visual representation of words or text, such as capitalization, punctuation, word boundaries, and spacing.

2.List out Ambiguity in issues and challenges.

Ambiguity in Issues in Natural Language Processing:

- Lexical Ambiguity
- Syntactic Ambiguity
- Semantic Ambiguity
- Anaphoric Ambiguity
- Pragmatic Ambiguity

Challenges in Natural Language Processing:

- Language Variability
- Context Understanding
- Ambiguity Resolution
- Co-reference Resolution
- Multilingual Processing

3.What is parsing? Describe Parsing in NLP

Parsing in NLP refers to the process of analyzing and interpreting the structure of a sentence or text according to a specific grammar or set of **rules** is called parsing

Parsing in natural language processing is the process of analyzing a sentence or text to determine its grammatical structure and identify the relationships between its components, such as nouns, verbs, and modifiers. It involves breaking down the input into a structured representation that can be further analyzed and used for various language understanding tasks**.**

4.What is Tree Banks? Write an example program.

Treebanks, in the context of Natural Language Processing (NLP), refer to annotated collections of sentences where each sentence is parsed and represented as a syntactic tree structure. These tree structures capture the grammatical relationships between words in a sentence, allowing researchers and NLP models to analyze and understand the syntax of the language.

Program:

```
Import nltk

From nltk.corpus import treebank

Nltk.download('treebank')

Parsed_sentences = treebank.parsed_sents()

For I in range(5):

    Print("Sentence", i+1)

    Print(parsed_sentences[i])

    Print()
```

5.What is Case Grammar? Write an example.

In natural language processing, Case Grammar can be applied to extract meaningful information from sentences and facilitate various language processing tasks such as parsing, information extraction, and semantic role labeling.

Example:

Consider the following sentence: "The cat chased the mouse."

Using Case Grammar, we can assign the following case roles to the sentence constituents:

Agent: "The cat" (the entity performing the action)

Verb: "chased" (the action being performed)

Patient: "the mouse" (the entity affected by the action)

6.Describe Frame based Knowledge Representation.

Frame-based knowledge representation is a method of organizing knowledge by using structured frames or templates that contain slots for storing information about specific objects or concepts, allowing for efficient retrieval and inference.

It provides a hierarchical structure that captures the relationships and attributes of objects, facilitating reasoning and inference in various domains.

7.Explain Stemming with example Program?

Stemming is a technique in natural language processing that reduces words to their root or base form, allowing for better analysis of textual data. For example, stemming would convert words like "running," "runs," and "ran" to the common root "run."

Program:

```
Import nltk

From nltk.stem import PorterStemmer

Def stemming_example():

    Stemmer = PorterStemmer()

    Words = ["running", "runs", "ran", "easily", "fairly"]

    Stemmed_words = [stemmer.stem(word) for word in words]

    For word, stemmed_word in zip(words, stemmed_words):

        Print(f"Original: {word}\tStemmed: {stemmed_word}")

Stemming_example()
```

8.Explain Lemmatization with example Program?

Lemmatization is a technique in natural language processing that reduces words to their base or root form. It helps in normalizing different variations of a word to a common base.

Program:

```
Import nltk

From nltk.stem import WordNetLemmatizer

Lemmatizer = WordNetLemmatizer()

Sentence = "The cats are running and jumping over the fences."

Tokens = nltk.word_tokenize(sentence)

Lemmatized_tokens = [lemmatizer.lemmatize(token) for token in tokens]

Print(lemmatized_tokens)
```

9.What is Chunking explain its types?

Chunking in natural language processing is the process of grouping together related words in a sentence. It helps in identifying and extracting meaningful phrases or chunks from the text.

There are different types of chunking techniques used in NLP.

Noun Phrase (NP) Chunking: This type of chunking focuses on identifying and grouping noun phrases in a sentence. Noun phrases typically consist of a noun and words that modify or describe it.

Verb Phrase (VP) Chunking: VP chunking involves identifying and grouping verb phrases in a sentence. Verb phrases usually consist of a verb and its associated words, such as adverbs, direct objects, or prepositional phrases. "

Prepositional Phrase (PP) Chunking: This type of chunking focuses on identifying and grouping prepositional phrases in a sentence.

Named Entity (NE) Chunking: NE chunking involves identifying and grouping named entities in a sentence, such as names of people, organizations, locations, or dates.

10.Write about Maximum-Likelihood Estimation and Smoothing

Maximum-Likelihood Estimation (MLE) is a statistical method used in natural language processing to estimate the parameters of a probabilistic model by maximizing the likelihood of the observed data.