AMEY THAKUR B-50 Amey.

COMPUTER ENGINEERING DEPARTMENT

ASSIGNMENT NO. 3

Subject: Natural Language Processing

COURSE: B.E Year: 2021-2022 Semester: VIII

DEPT: Computer Engineering

SUBJECT CODE: DLO8012 SUBMISSION DATE: 10/04/2022

Roll No.: 50 Name: Amey Thakur

Class: BE-Comps B Date of Submission: 10/04/2022

NLP Assignment - 3

1. **_____ is the step in which an input sentence is converted into a hierarchical structure that corresponds to the units of meaning in the sentence.**

   A) Graph processing

   B) Syntactic Processing

   C) Semantic Processing

   D) Morphological Processing

   E) Lexical Processing

Answer: c) Semantic Processing

Reason: Graphical processing is for graphical data to identify most relevant elements in text, morphological is for determining morpheme & lexical is for determining meanings of words.

AMEY THAKUR B-50 Amey

2. **Pragmatic ambiguity refers to the situation where the context of a phrase gives it multiple _____.**

A) Assumptions

B) Group of Data

C) Interpretations

D) Classifications

E) Structure

Answer: c) Interpretations

Reason: In layman's term, pragmatic ambiguity occurs when a statement is not explicit.
For example, the phrase "I like you too" might be interpreted as
I like you (Just as you like me)
I like you (Just like someone else does)

3. **Polysemy has the same spelling but _____ meaning.**

A) Different and contextual

B) Simple and Understandable

C) Different and related

D) Same and unrelated

E) Same and related

Answer: c) Different and Related

Reason: A single word with several meanings is polysemy.
In a dictionary, the several meanings are given under a single entry.
For example, the term dish has several meanings. however they are all included in a dictionary under one entry as a noun.

AMEY THAKUR B-50 Amey

4. A context-free grammar (CFG) is a list of _____ that defines the set of all well-formed sentences in a language.

A) Actions

B) Rules

C) Variables

D) Syntax

E) Formula

Answer: B) Rules

Reason: A CFG is a collection of rules that defines a language's set of all well formed sentences. Each rule contains a left side that specifies a syntactic category and a right side that describes alternative component parts, read from left to right.

5. _____ may be defined as the task of determining what entities are referred to by which linguistic expression.

A) Coherence resolution

B) Reference Resolution

C) Discourse resolution

D) Syntactic resolution

E) Semantic resolution

Answer: B) Reference Resolution

Reason: A reference is a language term that denotes an entity or an individual.
For example, Ram, the manager of ABC bank, spotted his friend Shyam in a shop. He went to meet him, the linguistic expressions like Ram, His, He are referenced. Similarly, reference solution may be characterized as the effort of establishing which entities are referred to by which verbal phrase.

AMEY THAKUR B-50 Amey

6. **What is Lexical Ambiguity?**

A) Ambiguity of a single word when it can be used as a verb, noun or adjective.

B) Ambiguity of a sentence

C) Words have many meanings

D) Sentences and words are not aligned.

Answer: A) Ambiguity of a single word when it can be used as a verb, noun or adjective.

Reason: Lexical ambiguity refers to the possibility of different interpretations of spoken or written language, which makes understanding difficult or impossible without further information. The ambiguity is usually caused by the fact that words may imply various things in different contexts.

7. **_____ shows how the words are associated with each other.**

A) Semantic Analysis

B) Tokens

C) Lexical Analysis

D) Discourse

Answer: A) Semantic Analysis

Reason: Semantic Analysis is a branch of Natural language processing that seeks to comprehend the meaning of natural language. Understanding natural language may appear to us as humans to be a simple task.

AMEY THAKUR B-50 Amey

8. ---------- is a model of behaviour composed of state, transition and actions.

A) FST

B) FSA

C) DAWG

D) Stemmer Algorithm

Answer: B) FSA

Reason: FSA is an automaton with a finite number of states (FSA). An automaton can be represented mathematically as a 5-tuple $(Q, \Sigma, q_0, A, \delta)$. where,

Q is finite set of states
Σ is finite input alphabet
$q_0 \in Q$ is the initial state
$A \subseteq Q$ is the set of accepting states
$\delta : Q \times \Sigma \rightarrow Q$ is transition Function

9. **What is 'Q' in the finite automata set?**

A) Finite set of states

B) Set of Input Symbols

C) Initial state

D) Set of Final States

Answer: A) Finite set of states

Reason: A Finite state Automaton is made up of 5 items $(Q, \Sigma, q_0, A, \delta)$, where Q represents the finite set of states that the automaton can be in.

AMEY THAKUR B-50 Amey.

10. **Which is not a level of NLP process?**

A) Pragmatic

B) Discourse

C) Morphological

D) Textual

Answer: D) Textual

Reason: There is no level such as Textual level in NLP. Process.

11. **_____ is a lexical database for the English language.**

A) Corpus

B) WordNet

C) Lexicon

D) Thesaurus

Answer: B) WordNet

Reason: WordNet is a lexical database / dictionary, for the English language that was created primarily for Natural Language Processing

AMEY THAKUR B-50 Amey.

12. What is Machine Translation?

A) Converts one human language to another

B) Converts human language to machine language

C) Converts any human language to English

D) Converts Machine language to human language

Answer: A) Converts one human language to another.

Reason: Machine Translation refers to automated systems that produce translations with or without human intervention. It does not include computer-based translation tools that help translators by giving them access to online dictionaries, distant terminologies databanks and so on.

13. Morphological analysis is done at _____level.

A) Character

B) Lexicon

C) Word

D) Sentence

Answer: c) Word

Reason: The term comes under morphological analysis & the other three comes under discourse, lexical & semantic analysis.

AMEY THAKUR B-50 Amey

14. Coreference resolution is a well-studied problem in _____.

A) Discourse

B) Text classification

C) WSD

D) Lexical Analysis

Answer: A) Discourse

Reason: Coreference Resolution is a well studied subject in dialogue in computer linguistics. Pronouns & other referring phrases must be linked to the relevant person in order to draw the accurate meaning of a text or even to determine the relative significance of numerous specified subjects.

15. In homonymy, the meanings of the words are not _____.

A) Joined

B) Related

C) Analysed

D) Changed

Answer: B) Related

Reason: In linguistics, homonyms are words that are homographs or homophones or both.

AMEY THAKUR B-50 Amey.

16. _____tools allow businesses to identify customer feedback toward products, brands or services in online feedback.

A) Lexical analysis

B) Syntactical analysis

C) Hybrid analysis

D) Sentiment Analysis

Answer: D) Sentiment Analysis

Reason: Sentiment analysis is textual contextual mining that finds and extracts subjective information in source material, assisting businesses in understanding the social sentiment of their brand, product or service while monitoring online conversations.

17. When the meaning of the words themselves can be misinterpreted then _____ ambiguity occurs.

A) Scope Ambiguity

B) Pragmatic Ambiguity

C) Semantic Ambiguity

D) Syntactic Ambiguity

Answer: C) Semantic Ambiguity

Reason: Semantic Ambiguity arises when a word, phrase or sentence has more than one interpretation when taken out of context. The phrase "her duck" in "We observed her duck" might refer to either the person's bird (the noun "duck" modified by the possesive pronoun "her") or to a move she performed (the verb "duck", the subject of which is the objective pronoun "her", object of the verb "saw").

AMEY THAKUR B-50 Amey

18. This POS tagging is based on the probability of the tag occurring.

A) Rule-based tagging

B) Stochastic tagging

C) Information based tagging

D) Group based tagging

Answer: B) Stochastic tagging

Reason: The most basic stochastic taggers disambiguate words merely based on the likelihood that a term appears with a certain tag. In other words, the most commonly seen tag in the training set with the term is the one allocated to an ambiguous instance of that word.

19. _____ is the smallest part of a word.

A) Letter

B) Morpheme

C) Word sketches

D) Verbs

Answer: B) Morpheme

Reason: Morpheme is a linguistic unit that cannot be further subdivided (for example, in, come, -ing, forming, incoming).

AMEY THAKUR B-50 Amey

20. The _____ algorithm is a process of removing suffixes from words in English.

A) Lovins Stemmer

B) Porter Stemmer

C) Paice Stemmer

D) Husk Stemmer

Answer: B) Porter Stemmer

Reason: Porter stemming algorithm is a method for eliminating frequent morphological and inflexional ends from English words. Its primary application is as part of the term normalization procedure that is often performed while establishing information retrieval system.

Student Signature:

Amey