

University of Mumbai
Examination 2021 under cluster KJSIEIT
Examinations Commencing from 22nd April 2021 to 30th April 2021

Program: Computer Engineering
Curriculum Scheme: Rev2019

Examination: M.E. (Artificial Intelligence) Semester I

Course Code: MEAIDLO1012 and Course Name: Natural Language Processing

Time: 2 hours

Max. Marks: 80

=

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks (2 marks each).
1.	If Syntax mapped into semantics, then which of the following is incorrect?
Option A:	Nouns ↔ things, objects, abstractions
Option B:	Verbs ↔ situations, events, activities
Option C:	Adjectives ↔ properties of things: slow, quick
Option D:	Adverbs ↔ properties of situations: slow, quick
2.	Which of the following statement has the syntactic ambiguity?
Option A:	John saw a boy in a park with the dog
Option B:	John saw a boy in a park with the telescope
Option C:	John saw a boy in a park with a statue
Option D:	John saw a boy in a park with a bat
3.	Which of following definitions are correct? i. <u>Morphology</u> : Concerns the way words are built up from smaller meaning bearing units. ii. <u>Syntax</u> : concerns what words mean and how these meanings combine in sentences to form sentence meanings. iii. <u>Semantics</u> : concerns how words are put together to form correct sentences and what structural role each word has. iv. <u>Pragmatics</u> : concerns how the immediately preceding sentences affect the interpretation of the next sentence.
Option A:	ii and iii
Option B:	iii only
Option C:	i & ii
Option D:	i only
4.	Which of the following techniques can be used for keyword normalization in NLP, the process of converting a keyword into its base form?
Option A:	Lemmatization
Option B:	Soundex
Option C:	Cosine Similarity
Option D:	N-grams
5.	Which of the following techniques can be used to compute the distance between two word vectors in NLP?

Option A:	Stemming and Lemmatization
Option B:	Euclidean distance & Cosine Similarity
Option C:	Bigram & Trigrams
Option D:	Chunking
6.	Which of the following statement is correct related to the Homonymy, Polysemy and Synonymy?
Option A:	Homonymy and Synonymy, have the effect of reducing precision by leading a system to return documents irrelevant to the users information need.
Option B:	Homonymy and polysemy, have the effect of reducing precision by leading a system to return documents irrelevant to the users information need.
Option C:	Synonymy and polysemy, have the effect of reducing precision by leading a system to return documents irrelevant to the users information need.
Option D:	Homonymy, Synonymy and polysemy, have the effect of reducing precision by leading a system to return documents irrelevant to the users information need.
7.	I'd like to get from Mumbai to Delhi, on either May 5th or May 6th 2021. It's okay if it stops in another city along the way. Here, the system had to figure out that it denotes the flight that the user wants to book in order to perform the appropriate action is an example of:
Option A:	Syntax Analysis
Option B:	Semantic analysis
Option C:	Lexical Analysis
Option D:	Discourse Analysis
8.	Which one of the following is not a pre-processing technique in NLP?
Option A:	Stemming and Lemmatization and converting to lowercase
Option B:	Removing punctuations
Option C:	Removal of stop words
Option D:	Sentiment analysis
9.	In linguistic morphology _____ is the process for reducing inflected words to their root form
Option A:	Rooting
Option B:	Text-Proofing
Option C:	Both Rooting & Stemming
Option D:	Stemming
10.	What is Morphological Segmentation?
Option A:	Is an extension of propositional logic
Option B:	Does Discourse Analysis
Option C:	Separate words into individual morphemes and identify the class of the morphemes
Option D:	Not covered in the above three options
11.	Which one of the following are keyword Normalization techniques in NLP?
Option A:	Part of Speech

Option B:	Named entity recognition
Option C:	Lemmatization and Stemming
Option D:	Soundex
12. has the same spelling and sound, but do not have related meanings?
Option A:	Homophones
Option B:	Homonymy
Option C:	Polysemy
Option D:	Synonymy
13.	Which is the type of morphology that changes the word category and affects the meaning?
Option A:	Inflectional
Option B:	Derivational
Option C:	Cliticization
Option D:	Rational
14.	Parts of speech can be divided into two broad super-categories: closed class types and open class types. The four major open classes that occur in the languages of the world are:
Option A:	nouns, verbs, auxiliary verbs, and adverbs
Option B:	nouns, verbs, adjectives, and prepositions
Option C:	nouns, verbs, adjectives, and adverbs
Option D:	nouns, verbs, prepositions, and adverbs
15.	If we consider the Penn Treebank Part-of-Speech Tags, then JJ, JJR, VB and VBZ can be described as :
Option A:	Adjective, Adjective superlative, Verb base form, Verb 3sg pres
Option B:	Adjective, Adjective comparative, Verb base form, Verb 3sg pres
Option C:	Adjective, Adjective comparative, Verb base form, Verb past participle
Option D:	Adjective, Adjective comparative, Verb gerund, Verb 3sg pres
16.	Part-of-speech tagging is the TAGGING process of assigning a part-of-speech or other lexical class marker to each word in a corpus. The input to a tagging algorithm is a string of words and a specified tagset such as Penn Treebank tagset. The output is a single best tag for each word. Give the potential tags for the following sentence considering the Penn Treebank tagset. Does that flight serve dinner?
Option A:	VBZ DT NN VB NN ?
Option B:	VB DT NN VB NN ?
Option C:	VB DT NN VBZ NN ?
Option D:	VBZ JJ NN VB NN ?
17.	Solve the equation according to the sentence "I am planning to visit Pune to attend Tech Mahindra Pune Hackathon". A = (# of words with Noun as the part of speech tag) B = (# of words with Verb as the part of speech tag)

	C = (# of words with frequency count greater than one) What are the correct values of A, B, and C?
Option A:	5, 5, 2
Option B:	5, 5, 0
Option C:	7, 5, 1
Option D:	6, 4, 2
18.	Named Entity Recognition means:
Option A:	Finding spans of text that constitute proper names and then classifying the type of the entity.
Option B:	Mapping between name and entity.
Option C:	Classification of text into subject and predicates.
Option D:	Searching text for proper nouns.
19.	How many lexemes are there in following list? man, men, girls, girl, mouse
Option A:	4
Option B:	5
Option C:	3
Option D:	2
20.	_____ is the problem of selecting a sense for a word from a set of predefined possibilities.
Option A:	Shallow Semantic Analysis
Option B:	Discourse
Option C:	Word Sense Disambiguation
Option D:	Pragmatic

Q2. A	Solve any Two out of Three	05 marks each
i.	What is semantic analysis? Discuss different semantic relationships between the words.	
ii.	Why POS tagging is hard? Discuss possible challenges faced while performing POS tagging.	
iii.	What is Natural Language Processing (NLP)? Discuss various stages involved in NLP process with suitable example.	
Q2. B	Solve any One out of Two	10 marks each
i.	What do you mean by word sense disambiguation (WSD)? Discuss dictionary-based approach for WSD. OR What do you mean by word sense disambiguation (WSD)? Discuss knowledge based WSD.	
ii.	What is morphology? Why do we need to do Morphological Analysis? Explain derivational and inflectional morphology in detail with suitable example	

Qu-3. A	Solve any Two out of Three	05 marks each
----------------	-----------------------------------	----------------------

i.	What is the need of preprocessing for text data in natural language? Explain the steps of preprocessing with example.
ii.	Explain with suitable examples following relationships between word meanings: 1. Homonymy 2. Polysemy
iii.	Explain CFG with suitable example. Discuss Sub categorization problems in CFG.
Qu-3. B	Solve any One out of Two 10 marks each
i.	What is language model? Write a note on N-Gram language model.
ii.	Define discourse and Pragmatic analysis. Discuss reference resolution problem in detail.

University of Mumbai
Examination 2021 under Cluster KJSIEIT
Examinations Commencing from 22nd April 2021 to 30th April 2021

Program: Computer Engineering

Curriculum Scheme: Rev2019

Examination: M.E. (Artificial Intelligence) Semester I

Course Code: MEAIDLO1012 and Course Name: Natural Language Processing

Time: 2 hours

Max. Marks: 80

Question Number	Correct Option
Q1.	D
Q2.	B
Q3.	D
Q4	A
Q5	B
Q6	B
Q7	D
Q8.	D
Q9.	D
Q10.	C
Q11.	C
Q12.	B
Q13.	B
Q14.	C
Q15.	B
Q16.	A
Q17.	D
Q18.	A
Q19.	C
Q20.	C