(2½ Hours)

[Total Marks: 60]

N. B.: (1) **All** questions are **compulsory**.
(2) Make **suitable assumptions** wherever necessary and **state the assumptions** made.
(3) Answers to the **same question** must be **written together**.
(4) Numbers to the **right** indicate **marks**.
(5) Draw **neat labelled diagrams** wherever **necessary**.
(6) Use of **Non-programmable** calculators is **allowed**.

I	**Choose the correct alternative and rewrite the entire sentence with the correct alternative.**				**(30)**
1.	_____ is a type of assistive technology that reads digital text aloud.				
	a.	Speech to text	**b.**	Text to speech	
	c.	Text Summarization	**d.**	Text classification	
2.	_____ refers to sentences and phrases that potentially have two or more possible interpretations.				
	a.	Homonyms	**b.**	Synonyms	
	c.	Ambiguity	**d.**	Sarcasm	
3.	What Σ of content-free grammar tuple G = (Σ, N, S, R) indicates?				
	a.	disjoint finite sets of terminal symbol	**b.**	disjoint finite sets of nonterminal symbol	
	c.	a finite set of production rules	**d.**	the start symbol	
4.	Syntactic Parsing deals with basic techniques for _____.				
	a.	Token-driven natural language parsing	**b.**	grammar-driven natural language parsing	
	c.	lexical driven natural language parsing	**d.**	lemma driven natural language parsing	
5.	A basic task of _____ is to relate morphological variants to their lemma bundled up with its invariant semantic and syntactic information.				
	a.	Tokenization	**b.**	sentence segmentation	
	c.	lexical analysis	**d.**	morphology	
6.	_____ systems identify different types of proper names, such as person and company names, and sometimes special types of entities, such as dates and times, that can be easily identified using surface level textual patterns.				
	a.	Named entity recognition	**b.**	Chunking	
	c.	Parsing	**d.**	Word Sense Disambiguity	
7.	____refers to the process(es) that choose the speech acts, establish the content, determine how the situation dictates perspectives, and so on.				
	a.	Text planning	**b.**	Linguistic planning	

	c.	Macro-planning	**d.**	Micro-planning
8.	_____attempts to produce useful output, such as a partial analysis, even if the input is not covered by the grammar.			
	a.	A robust parser	**b.**	An analyser
	c.	A descriptor	**d.**	A parser
9.	A grammar that produces ____ parse tree for some sentence is said to be ambiguous.			
	a.	One	**b.**	more than one
	c.	Two	**d.**	Three
10.	Pushing and pulling are the interfaces used to connect ____.			
	a.	text planners	**b.**	Speakers
	c.	text planner and speaker	**d.**	speaker and linguistic component
11.	The standard way to represent the syntactic structure of a grammatical sentence is ____.			
	a.	a parse tree	**b.**	a flow tree
	c.	a grammar tree	**d.**	a structure tree
12.	An LR-parser can detect a syntactic error as soon as _____.			
	a.	The parsing starts	**b.**	It is possible to do so a right-to-left scan of the input
	c.	Parsing ends	**d.**	It is possible to do so a left-to-right scan of the input
13.	Part of speech can be regarded as simplified form of _____ analysis			
	a.	Lexical	**b.**	syntactical
	c.	Semantical	**d.**	Morphological
14.	_____ is there exist some words for which more than one POS tag is possible.			
	a.	Ambiguous words	**b.**	unknown words
	c.	confused words	**d.**	altered words
15.	_____ used in classification tasks, like rule-based systems, can cover more context and enable flexible feature representations, and yield outputs easier to interpret.			
	a.	Support Vector Machines	**b.**	Neural Networks
	c.	Decision tree	**d.**	Fuzzy set Theory
16.	HMM is a _____ model			
	a.	Discriminative	**b.**	non discriminative
	c.	non generative	**d.**	Generative
17.	The output of the beam search is a _____, which has all the hypotheses that have been explored during the search.			
	a.	Result	**b.**	output
	c.	Lattice	**d.**	Literal

18.	_____ is a process of using a set of linguistic and logical tools to figure out the real meaning of the text.			
	a.	POS tagging	**b.**	Wordnet
	c.	Pragmatic analysis	**d.**	Perplexity
19.	_____uses features of the partially built dependency structure together with features of the tagged input string.			
	a.	CYK algorithms	**b.**	MST parser
	c.	Earley's algorithm	**d.**	MaltParser
20.	The Paninian framework has been successfully applied to _____ languages.			
	a.	French	**b.**	Japanese
	c.	Indian	**d.**	German
21.	A data-driven parser-generator constructs a parser given _____.			
	a.	a grammar	**b.**	a treebank
	c.	a graph	**d.**	any corpus
22.	An evaluative component _____ that ranks candidate analyses via a numerical scoring scheme.			
	a.	EVAL	**b.**	EVALUATE
	c.	EVL	**d.**	EVLT
23.	A generative model is one that defines a _____ probability distribution over inputs and outputs			
	a.	Conditional	**b.**	Joint
	c.	Experimental	**d.**	Standard
24.	_____ is the problem of estimating the performance of different models in order to choose the best one.			
	a.	model search	**b.**	model selection
	c.	model finder	**d.**	model assessment
25.	NLTK stands for_____.			
	a.	Natural Language Tooltip	**b.**	Natural language Toolkart
	c.	Natural language Talk	**d.**	Natural Language Toolkit
26.	Dissimilarity between words expressed using cosine similarity will have values significantly higher than____			
	a.	0.5	**b.**	0.4
	c.	0.3	**d.**	0.2
27.	Similarity between synsets is calculated using_____			
	a.	HMM	**b.**	Hamiltonian cycle
	c.	shortest path	**d.**	simple path

28.	Words are polysemous means _____				
	a.	they have different but related senses.	**b.**	they have different senses.	
	c.	they have multiple spellings	**d.**	they have different meanings	
29.	Anaphoric or coreference resolution is a subset of _____.				
	a.	Lexical analysis	**b.**	Tokenization	
	c.	Lemmatization	**d.**	Discourse Analysis	
30.	Lesk algorithm uses _____based approach.				
	a.	hand-tailored procedural	**b.**	dictionary	
	c.	word expert based	**d.**	selectional constraints	

II	**Attempt *any one* of the following:**		**6**
	a)	What are the stages of analysis in processing of natural language?	
	b)	Write a short note on typological classification.	
	c)	Explain the concept of finite state transducer with suitable example.	
2	**Attempt *any one* of the following:**		**6**
	a)	Explain Cocke-Kasami-Younger (CYK/CKY) algorithm in detail.	
	b)	Write a short note on Discourse Representation Theory (DRT).	
	c)	Explain the components and levels of representation in natural language generation?	
3	**Attempt *any one* of the following:**		**6**
	a)	Explain transformation-based learning (TBL).	
	b)	Write a short note on Hiero.	
	c)	What is statistical machine learning? Discuss the problems related to the statistical machine learning.	
4	**Attempt *any one* of the following:**		**6**
	a)	Explain the two component of statistical parsing model in detail.	
	b)	Write a short on MALT parser.	
	c)	Enlist and explain the different subtask mapped to the classification tasks in maximum entropy model in Chinese language.	
5	**Attempt *any one* of the following:**		**6**
	a)	What is idiomaticity and explain its type in detail.	
	b)	Write a short note on association measures.	
	c)	Explain applications of word sense disambiguation.	