QUESTION BANK NLP

Q1. “Natural Language Processing is a technique that empowers machines to communicate with each other and with humans by making them understand the natural language”. Justify the statement with supporting facts.

Q2. What are the thematic roles associated with the sentence: “John broke the window with the hammer”.

Q3. Define the following with respect to Information Retrieval:
i. Vector Space Model
ii. Term Frequency
iii. Inverse Document Frequency

Q4. Identify and describe the ambiguities in the following sentences.
i. The man kept the dog in the house.
ii. Book that flight.

Q5. Explain how NLP is linked with conference resolution. Differentiate anaphors and cataphors in this context.

Q6. In a statistical named entity recognition (NER) system using the IOB (in, out, begin) labelling approach a model is learned using a suitably tagged corpus. What features are/can typically be used in tagging words and what are the labels (assume we use only 3 NER types: proper names, locations, organizations)? For the following fragment of some corpus indicate the labels (do not give the features). Note: natural phenomena like cyclones/typhoons/comets are often given names. Cyclone Trump in the Bay of Bengal moved towards Chennai.

Q7. Suppose we want to build a statistical phrase-based machine translation (MT) system for a word order/function-word based language like English (or Mandarin) to a morphologically rich language like Hindi (with case markers merged with the word) or Sanskrit. What problems do you foresee and how will you tackle them? Note: Two main steps in statistical phrase-based MT are phrase identification and phrase alignment.

Q8. Map the following CFG into an equivalent RTN that uses only 3 networks- an S, NP and PP network. Make your network as small as possible.
S->NP VP
VP->V | V NP | V PP
NP-> ART NP2 | NP2
NP2-> N | ADJ NP2 | NP3 PREPS
NP3-> N
PREPS->PP | PP PREPS
PP->NP

Q9. Identify the head and morphological type (Noun Phrase, Verb Phrase, Adjective Phrase, Adverbial Phrase) of the following sentence segments.
i. The president of the company (2 Marks)
ii. Looked up the chimney (2 Marks)
iii. Angry as a hippo (2 Marks)
iv. Rapidly like a bat (2 Marks)

Q10. Write the FOPC representation for following sentences:

i. All vegetarian restaurants serve vegetarian food. (2 Marks)
ii. I only have five dollars and I don't have a lot of time. (2 Marks)
iii. I ate a vegetable fried rice from Indian Coffee house. (2 Marks)
iv. There is a restaurant that serves Italian food near Connaught place. (2 Marks)

Q11. Show, how a speech dialogue system can reduce the number of errors made because of incorrect speech recognition? Define Pragmatics with example.

Q12. In a corpus of N documents, one randomly chosen document contains a total of T terms and the term "hello" appears K times. What is the correct value for the product of TF (term frequency) and IDF (inverse-document-frequency), if the term "hello" appears in approximately one-third of the total documents.

Q13. How can you say that shallow parsing is a technique for analysing the structure of a sentence? Take a sentence as per your choice and do the shallow parsing for that sentence.

Q14. Explain the resources do you need to do word sense disambiguation using Lesk's algorithm and how are those resources used?

Q15. Describe the importance of Cross Lingual Information Retrieval (CLIR). Also explain the challenges associated with CLIR.

Q16. Design a finite state transducer with E-insertion orthographic rule that parses from surface level "foxes" to lexical level "fox+N+PL" using FST.

Q17. For each sentence, identify whether the different meanings arise from structural ambiguity, semantic ambiguity, or pragmatic ambiguity? (5 marks)

i. Time flies like an arrow
ii. He crushed the key to my heart.

Q18. Identify the morphological type (Noun phrase, Verb Phrase, Adjective Phrase) of following sentence segments: (5 Marks)

i. Important to Bill
ii. Looked up the tree

Q19. Perform parsing using simple top-down parsing for the sentence "The dogs cried" using the grammar given below:

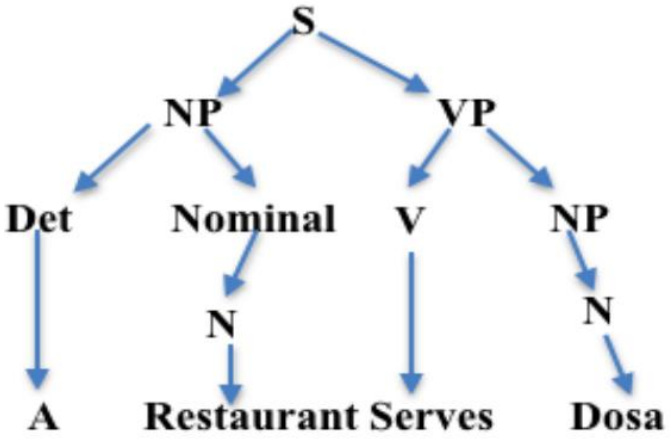
S->NP VP
NP->ART N
NP->ART ADJ N
VP->V
VP->V NP

Q20. Map the following CFG into an equivalent RTN that uses only 3 networks- an S, NP and PP network. Make your network as small as possible. (8 marks)

S->NP VP
VP->V | V NP | V PP
NP-> ART NP2 | NP2
NP2-> N | ADJ NP2 | NP3 PREPS

NP3-> N
PREPS->PP | PP PREPS
PP->NP

Q21. The parse tree for the sentence "A restaurant serves dosa" is given below. Perform semantic analysis and show the semantic interpretations of the constituents. Explain the process. (8 marks)



Q22. How are machines inspired by the biology of speech processing? Explain with example.

Q23. The Lesk algorithm is one of the best algorithms for word sense disambiguation. Show how to solve Disambiguating word sense using the package pywsd and nltk in python. Discuss in detail.

Q24. Define the following.

i. Principle of scope ambiguity
ii. Attachment ambiguity resolution model

Q25. Identify and describe the ambiguities in the following sentences.

i. The man kept the dog in the house.
ii. Book that flight.

Q26. Give an example of an NLP system or approach that only knows about individual words (not about their context). Explain, what knowledge of words it needs and give an overview of how it works? Give examples to illustrate your explanation.

Q27. Write the FOPC of the following sentences:

i. Chicken is food.
ii. Either Sue is rich or she is poor
iii. Bill eats peanuts and is still alive
iv. Sue eats everything Bill eats

Q28. For the CFGs given:
S → NP VP
VP → V NP
NP → Det N
Draw the shift-reduce parser in processing the sentence:
"The woman saw a puppy"
Use the following lexical entries to create the chart parser.
The | a: Det

woman | puppy: N
saw: V

Q29. Derive a top-down, depth-first, left-to-right parse tree for the given sentence:
"The angry bear chased the frightened little squirrel".
Use the following grammar rules to create the parse tree:

S → NP VP	Det → the
NP → Det Nom	Adj→ little | angry | frightened
VP → V NP	N → squirrel | bear
Nom →Adj Nom | N	V → chased

Q30. The following two sentences exhibit parsing ambiguities. How will your lexicalized PCFG handle these ambiguities? "Find the men in suits" and "Find the men in summer". Discuss with results.

Q31. Given the following CFG grammar from ATIS System, USA. Perform syntactic analysis of the following sentence using any of the parsing method.
"Book the flight through Houston."

S → NP VP	Det → that | this | a | the
S → Aux NP VP	Noun → book | man | flight
S → VP	Verb → book | include | prefer | man
NP → Pronoun	Pronoun → I | she | me
NP → Proper-Noun	Proper-Noun → Houston | TWA
NP → Det Nominal	Aux → does
Nominal → Noun	Preposition → from | to | through
Nominal → Nominal Noun	
Nominal → Nominal PP	
VP → Verb	
VP → Verb NP	
VP → Verb NP PP	
VP → Verb PP	
VP → VP PP	
PP → Preposition NP	

Q32. Describe the important tools available for the development of NLP applications. Write the features of each of these tools.

Q33. Explain different phases in the development of speech recognition system.

Q34. Define transfer model of Machine Translation. List out its three phases.

Q35. Design a finite state transducer with E-insertion orthographic rule that parses from surface level "foxes" to lexical level "fox+N+PL" using FST.

Q36. Analyze the significance of Word Sense Disambiguation in NLP. Explain any one WSD method.

Q37. Identify the head and morphological type (Noun Phrase, Verb Phrase, Adjective Phrase, Adverbial Phrase) of the following sentence segments.
i. The president of the company
ii. Looked up the chimney
iii. Angry as a hippo

iv. Rapidly like a bat

Q38. Give an example of an NLP system or approach that only knows about individual words (not about their context). Explain, what knowledge of words it needs and give an overview of how it works? Give examples to illustrate your explanation.

Q39. The following two sentences exhibit parsing ambiguities. How would your lexicalized PCFG handle these ambiguities? "Find the men in suits." and "Find the men in summer." Discuss with results.

Q40. Show the use of Finite-State Machine based Morphological parsing with the help of suitable diagram.

Q41. Write an algorithm for parsing a finite-state transducer using the pseudocode with an example. Also specify the merits and demerits of this algorithm.

Q42. Explain the resource(s) do you need to do word sense disambiguation using Lesk's algorithm and how are those resource(s) used?

Q43. Explain the three main techniques used for machine translation? Very briefly say what is/are the main resource/s required in each.

Q44. What do you mean by Part-of-Speech Tagging? What is the need of this task in NLP?

Q45. Explain the role of selectional restriction in semantic interpretation.

Q46. Explain the use of modal operators in logical from language and explain how it suffer from the failure of substitutivity.

Q47. Between the words eat and find which would you expect to be more effective in selection restriction-based sense disambiguation. Explain.

Q48. Map the following CFG into an equivalent RTN that uses only 3 networks- an S, NP and PP network. Make your network as small as possible.
S->NP VP
VP->V | V NP | V PP
NP-> ART NP2 | NP2
NP2-> N | ADJ NP2 | NP3 PREPS
NP3-> N
PREPS->PP | PP PREPS
PP->NP.

Q49. Perform parsing using simple top-down parsing for the sentence "The dogs cried" using the grammar given below:
S->NP VP
NP->ART N
NP->ART ADJ N
VP->V
VP->V NP

Q50. Explain text summarization and multiple document text summarization with neat diagram.

Q51. Explain surface anaphora and the different methods for dealing with surface anaphor.

Q52. Describe different ways of building belief models in a conversational agent.

Q53. Explain lexicon, lexeme and the different types of relations that hold between lexemes.

Q54. Describe Text Coherence. Discuss the significance of Text Coherence in Discourse Segmentation.

Q55. State the difference between homonymy and polysemy and give an example of each.

Q56. For the CFGs given:
S → NP VP
VP → V NP
NP → Det N
Draw the shift-reduce parser in processing the sentence – "The woman saw a puppy".
Use the following lexical entries to create the chart parser.
The | a: Det
woman | puppy: N
saw: V

Q57. Derive a top-down, depth-first, left-to-right parse tree for the given sentence:
• The angry bear chased the frightened little squirrel
Use the following grammar rules to create the parse tree:

S → NP VP	Det → the
NP → Det Nom	Adj→ little | angry | frightened
VP → V NP	N → squirrel | bear
Nom →Adj Nom | N	V → chased

Q58. What are the 2 main classes of tagging algorithms in which they can be grouped into? Explain each one in detail.

Q59. Define Phonology. Illustrate the purpose of International Phonetic Alphabet in NLP.

Q60. What are the three principles that predict when garden paths will arise in syntactic interpretations of sentences?