

Time: 3 hours

Max. Marks: 80

N.B. (1) Question No. 1 is compulsory

(2) Assume suitable data if necessary

(3) Attempt any three questions from the remaining questions

**Q.1 Solve any Four out of Five**

**5 marks each**

- a Explain the challenges of Natural Language processing.
- b Explain how N-gram model is used in spelling correction
- c Explain three types of referents that complicate the reference resolution problem.
- d Explain Machine Translation Approaches used in NLP.
- e Explain the various stages of Natural Language processing.

**Q.2 10 marks each**

- a What is Word Sense Disambiguation (WSD)? Explain the dictionary based approach to Word Sense Disambiguation.
- b Represent output of morphological analysis for Regular verb, Irregular verb, singular noun, plural noun Also Explain Role of FST in Morphological Parsing with an example

**Q.3 10 marks each**

- a Explain the ambiguities associated at each level with example for Natural Language processing.
- b Explain Discourse reference resolution in detail.

**Q.4 10 marks each**

a

<S>	Martin	Justin	can	watch	Will	<E>
<S>	Spot	will	watch	Martin	<E>	
<S>	Will	Justin	spot	Martin	<E>	
<S>	Martin	will	pat	Spot	<E>	

For given above corpus,

N: Noun [Martin, Justin, Will, Spot, Pat]

M: Modal verb [can , will]

V: Verb [ watch, spot, pat]

Create Transition Matrix & Emission Probability Matrix

Statement is “Justin will spot Will”

Apply Hidden Markov Model and do POS tagging for given statements

- b Describe in detail Centering Algorithm for reference resolution.

**Q.5 10 marks each**

- a For a given grammar using CYK or CKY algorithm parse the statement

“The man read this book”

Rules:

$S \rightarrow NP VP$	$Det \rightarrow that \mid this \mid a \mid the$
$S \rightarrow Aux NP VP$	$Noun \rightarrow book \mid flight \mid meal \mid man$
$S \rightarrow VP$	$Verb \rightarrow book \mid include \mid read$
$NP \rightarrow Det NOM$	$Aux \rightarrow does$
$NOM \rightarrow Noun$	
$NOM \rightarrow Noun NOM$	
$VP \rightarrow Verb$	
$VP \rightarrow Verb NP$	

- b Explain Porter Stemmer algorithm with rules

**Q.6 10 marks each**

- a Explain information retrieval versus Information extraction systems  
b Explain Maximum Entropy Model for POS Tagging