Course code	Course Name	L-T-P Credits	Year of Introduction
CS366	Natural language processing	3-0-0-3	2016

Prerequisite: Nil

Course Objectives

- To introduce the fundamentals of Language processing from the algorithmic viewpoint.
- To discuss various issues those make natural language processing a hard task.
- To discuss some applications of Natural Language Processing (NLP).

Syllabus

Levels of Language Analysis, Syntax, Semantics and Pragmatics of Natural Language, Language Processing, Issues and approaches to solutions, Applications of Natural Language Processing (NLP).

Expected Outcome

The student able to

1. appreciate the fundamental concepts of Natural Language Processing.
2. design algorithms for NLP tasks.
3. develop useful systems for language processing and related tasks involving text processing.

Text Books

1. D. Jurafsky and J. H. Martin, Speech and Language Processing, Prentice Hall India, 2000
2. James Allen, Natural Language Understanding, 2e, The Benjamin/Cummings Publishing Company Inc., Redwood City, CA.

References

1. Charniak, Eugene, Introduction to Artificial intelligence, Addison-Wesley, 1985..
2. Ricardo Baeza-Yates and Berthier Ribeiro-Neto, Modern Information Retrieval, Addison-Wesley,1999.
3. U. S. Tiwary and Tanveer Siddiqui, Natural Language Processing and Information Retrieval, Oxford University Press, 2008.

Course Plan

Module	Contents	Hours	End Sem. Exam Marks
I	Introduction to Natural Language Understanding- Levels of language analysis- Syntax, Semantics, Pragmatics. Linguistic Background- An Outline of English Syntax.	8	15%
II	Lexicons, POS Tagging, Word Senses. Grammars and Parsing- Features, Agreement and Augmented Grammars.	7	15%
	FIRST INTERNAL EXAM		
III	Grammars for Natural Language, Parsing methods and Efficient Parsing. Ambiguity Resolution- Statistical Methods. Probabilistic Context Free Grammar.	9	15%
IV	Semantics and Logical Form: Linking Syntax and Semantics- Ambiguity Resolution- other Strategies for Semantic Interpretation- Scoping and the Interpretation of Noun Phrases.	6	15%
	SECOND INTERNAL EXAM		
V	Knowledge Representation and Reasoning- Local Discourse	8	20%

	Context and Reference- Using World Knowledge- Discourse Structure- Defining a Conversational Agent.		
VI	Applications- Machine Translation, Information Retrieval and Extraction, Text Categorization and Summarization.	**4**	**20%**
END SEMESTER EXAM			

Question Paper Pattern

1. There will be *five* parts in the question paper – A, B, C, D, E
2. Part A
   a. Total marks : 12
   b. *Four* questions each having *3* marks, uniformly covering modules I and II; All*four* questions have to be answered.
3. Part B
   a. Total marks : 18
   b. *Three* questions each having *9* marks, uniformly covering modules I and II; *Two* questions have to be answered. Each question can have a maximum of three subparts.
4. Part C
   a. Total marks : 12
   b. *Four* questions each having *3* marks, uniformly covering modules III and IV; All*four* questions have to be answered.
5. Part D
   a. Total marks : 18
   b. *Three* questions each having *9* marks, uniformly covering modules III and IV; *Two* questions have to be answered. Each question can have a maximum of three subparts.
6. Part E
   a. Total Marks: 40
   b. *Six* questions each carrying 10 marks, uniformly covering modules V and VI; *four* questions have to be answered.
   c. A question can have a maximum of three sub-parts.