**Elective IV: Natural Language Processing**

**Unit I:** Introduction: NLP tasks in syntax, semantics, and pragmatics, Key issues &Applications such as information extraction, question answering, and machine translation, the problem of ambiguity, the role of machine learning, brief history of the field.

**Unit II:** N-gram Language Models : Role of language models, Simple N-gram models, Estimating parameters and smoothing, Evaluating language models, Part Of Speech Tagging and Sequence Labeling Lexical syntax, Hidden Markov Models, Maximum Entropy models.

**Unit III:** Syntactic parsing: Grammar formalisms and tree banks, Efficient parsing for context-free grammars (CFGs), Statistical parsing and probabilistic CFGs (PCFGs), Lexicalized PCFGs.

**Unit IV:** Semantic Analysis: Lexical semantics and word-sense disambiguation, Compositional semantics, Semantic Role labeling and Semantic Parsing.

**Unit V:** Information Extraction (IE): Named entity recognition and relation extraction, IE using sequence labeling, automatic summarization Subjectivity and sentiment analysis.

**Unit VI:** Machine Translation (MT): Basic issues in MT, Statistical translation, word alignment, phrase-based translation, and synchronous grammars.

**Text Books:** 1. Speech and Language Processing, D. Jurafsky and R. Martin, 2nd edition, Pearson Education, 2009.

 2. Language Implementation Patterns, Terence Parr, Pragmatic Programmers, 2010. 42

**Reference Books:** 1. Natural Language Understanding, Allen James, Second Edition, Benjamin/Cumming, 1995.

2. NLP: A Paninian Perspective, Akshar Bharati, Vineet Chaitanya, and Rajeev Sangal, Prentice Hall, New Delhi, 1994.