Speech and Language Technology

1.1 Course Number: CS491

1.2 Contact Hours: 3-0-0 Credits: 9

1.3 Semester-offered: 4th Year-Odd

1.4 Prerequisite: basic understanding of linguistics, mathematics

1.5 Syllabus Committee Member: Dr. Sushum Biswas, Dr. Daya Sagar Gupta & Dr. Gargi Srivastava

2. **Objective:**

- Understand the fundamental concepts of Natural Language Processing, most well-known techniques and theories as well as most relevant existing resources.
- Understand most relevant applications of NLP and the theories, techniques and resources they use.
- Design and development of programs to solve specific problems in the NLP context, involving the selection of most appropriate techniques and resources as well as the use of existing resources.
- Reasons about several problems in the NLP context that imply considering different techniques and resources.

3. **Course Content:**

Unit-wise distribution of content and number of lectures

Unit	Topics	Sub-topic	Lectures
1	Document Structure and Language	Text selection, Tokenization, Sentence splitting, Language Identifiers	8
2	Words	Morphology, Finite States Automata, Finite States Transducers. PoS tagging, Hidden Markov Models. Lexical semantics, Semantic resources. Word Sense Disambiguation.	8
3	Word sequences	Recognition and classification of word sequences with meaning. BIO discriminative models. Conditional Random Fields (CRF). Named Entity Recognition and Classification (NERC). Noun-phrase Chunking.	8
4	Sentences	Syntactic grammars, typology. Context free grammars. Probabilistic context free grammars. Chomsky normal	8

		form grammars. Syntactic parsers, properties and strategies. CKY and probabilistic CKY parsers.	
5	Sentence sequences	Coreference resolution. Mention detection. Types of techniques for the generation of coreferents chains. Mention-pair model. Entity-mention model. Rankers model.	8
		Total	**40**

4. Readings

4.1 Textbook:

- Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition - Jurafsky, D.; Martin, J.H, Prentice-Hall, Inc., 2008. ISBN: 9332518416
- Handbook of natural language processing - Dale, R.; Moisl, H.; Somers, H, Marcel Dekker, 2000. ISBN: 0824790006

4.2 Reference books:

- Foundations of statistical natural language processing - Manning, C.D.; Schütze, H, MIT Press, 1999. ISBN: 0262133601
- The Oxford handbook of computational linguistics - Mitkov, R. (ed.), Oxford University Press, 2003. ISBN: 0198238827

5 Outcome of the Course: After the completion of this course, the students will be able to:

Perform speech and language analyses with simulation.

Design speech and a language technology for a simple robot.