611 - Natural Language Processing

Course Prerequisites:

1. A previous course on Artificial Intelligence will help.
2. Courses of Data Structures and Algorithms should have been done.
3. Exposure to Linguistics is useful, though not mandatory.

Course Objectives: To study fundamental concepts of Natural Language Processing and to introduce the basics of Language processing from algorithmic viewpoint.

Course Contents:

Introduction, Machine Learning and NLP, ArgMax Computation, Word Sense Disambiguation: WordNet, Wordnet; Application in Query Expansion, Measures of WordNet Similarity.

Resnick's work on WordNet Similarity, Parsing Algorithms, Evidence for Deeper Structure; Top Down Parsing Algorithms, Noun Structure; Top Down Parsing Algorithms, Non-noun Structure and Parsing Algorithms.

Probabilistic parsing; Sequence labelling, PCFG, Probabilistic parsing: Training issues, Arguments and Adjuncts, Probabilistic parsing; inside-outside probabilities.

Speech : Phonetics, Hidden Markov Model, Morphology, Graphical Models for Sequence Labelling in NLP, Consonants (place and manner of articulation) and Vowels.

Forward Backward probability; Viterbi Algorithm, Phonology, Sentiment Analysis and Opinions on the Web, Machine Translation and MT Tools - GIZA++ and Moses, Text Alignment, POS Tagging.

Phonology; ASR, Speech Synthesis, Hidden Markov Model and Viterbi, Precision, Recall , F-score, Map, Semantic Relations; UNL; Towards Dependency Parsing.

Universal Networking Language, Semantic Role Extraction, Baum Welch Algorithm; HMM training.

Main Reading:

1. Allen, James, Natural Language Understanding, Second Edition, Benjamin/Cumming, 1995.
2. Charniack, Eugene, Statistical Language Learning, MIT Press, 1993.
3. Jurafsky, Dan and Martin, James, Speech and Language Processing, Second Edition, Prentice Hall, 2008.
4. Manning, Christopher and Heinrich, Schutze, Foundations of Statistical
5. Natural Language Processing, MIT Press, 1999.

Supplementary Reading:

1. Radford, Andrew et. al .,Linguistics, An Introduction, Cambridge University Press, 1999.
2. Journals: Computational Linguistics, Natural Language Engineering, Machine Learning, Machine, Translation, Artificial Intelligence.
3. Conferences: Annual Meeting of the Association of Computational Linguistics (ACL), Computational Linguistics(COLING), European ACL (EACL), Empirical Methods in NLP (EMNLP), Annual Meeting of the Special Interest Group in Information Retrieval (SIGIR), Human Language Technology (HLT), International Conference on Natural Language Processing (ICON).

Website Links:

http://nptel.iitm.ac.in/courses/106101007/