

Natural Language Understanding (CAP 6640) SPRING-11
Fernando Gomez ; Office Hours: M, T, Th: 3pm to 4pm.

Week 1 Lecture by Instructor

Main Issues in NLP

Applications

- * Search Engines
- * Question Answering
 - * Watson Jeopardy
- * Machine Translation
- * Sentiment Analysis
- * Social Networks
- * Bioinformatics
- * Dialogues and User's Models
- * High level programming specifications

POS Tagging Chapter 8

B. Santorini "Part of Speech Tagging Guidelines"

WordNet

Parsing

Semantics

Tools: NLTK

Presentations by students.

Some Research Topics for Term Paper

WSD

- * Supervised Methods
- * Unsupervised Methods
- * Knowledge-Based Methods

Question-Answering Systems

- * Search Engines
- * Trec
- * Watson

Dialogue Modeling

Social Networks

Natural Language Generation

Bioinformatics

Sentiment Analysis

- * PMI
- * Machine Learning Methods

Knowledge Acquisition

- * Pattern Extraction
- * Machine-Learning Based
- * Automatic Extraction from Encyclopedias

Machine Translation

- * Statistical ML

Semantic Roles

- * PropBank
- * FrameNet

The major goals of this course are: a) to perform research in a topic selected by the student b) Writing a paper describing that research and d) read and explain state of the art research papers. After the first week, students will present research papers on topics of their choosing. Every student in class is responsible for summarizing the paper that is presented that day. The summary should be about one page long. Students should write a proposal (2 or 3 pages

long) explaining the selection of their research topic. That proposal should be written by the 4th week of the semester or earlier. Your paper will be judged using the criteria that are normally applied to papers submitted to good conferences. The closer your term paper is to be accepted in a good conference the better the grade it will get. The ideal paper is a paper that presents a research topic, a discussion of the related research, and an implementation to test the ideas presented. We will be discussing many papers in class that fit that pattern. Students should check the status of their paper with the instructor periodically after the initial proposal.

Grading: Summaries 10%, Proposal + Presentations 10%, Term Paper 80%. Note that attendance and participation in class is important in this course. I will deduct points if a students develops a pattern of missing classes. I will inform him/her if that is the case. I will use the +/- system in grading.

(A- 90-92, B+ 87-89, B 83-86, B- 80-82, C+ 77-79, C 73-76, C- 70-72, D+ 67-69, D 63-66, D- 60-62, F 59-0).