Course description

This course covers algorithms for associating deep or elaborated linguistic structures with naturally occurring data, covering parsing, semantics, and discourse.

Textbook

The course textbook is *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, 2nd edition, by Daniel Jurafsky and James Martin.  A partial electronic draft of the 3rd edition is available from [https://web.stanford.edu/~jurafsky/slp3/ (Links to an external site.)](https://web.stanford.edu/~jurafsky/slp3/)

Prerequisites:

* CSE 373 (Data Structures) or equivalent
* Math/Stat 394 (Intro to Probability) or equivalent
* Formal grammars, languages, and automata
* Programming in one or more of Java, Python, C/C++, or Perl
* Linux/Unix commands

Programming languages:

* A number of assignments in this course require the use of NLTK. Since NLTK is a Python toolkit, those assignments must be completed in Python.
* For all other assignments, any of the languages listed in "languages" file in the course dropbox directory can be used.

Grading

* 100%: Homework Assignments
* Up to 2% adjustment for significant in-class or GoPost participation