

# Ling 441

## Computational Linguistics I

### Syllabus

## 1 Description

Computational linguistics, broadly construed, encompasses three distinct areas of study: (1) human language technology, including machine translation, information extraction, and spoken language dialogue systems, among other technologies; (2) computational models of language users, which divides into computational psycholinguistics, in which humans are the object of study, and natural language processing as a branch of artificial intelligence, in which artificial agents are the object; and (3) digital linguistics, which is the use of computation in support of language documentation and linguistic research.

This class is a general introduction to computational linguistics. The first part of the course will focus on parsing and semantic interpretation, and on writing “computational” grammars to drive parsers and semantic interpreters. The second part of the course will look at getting useful information—like the information needed to build larger grammars—out of text corpora, as well as other kinds of processing that is typical of computational linguistics.

We will learn just enough of the programming language Python in order to get our work done. No prior computational background is assumed.

The textbook is Bird, Klein, and Loper, *Natural Language Processing with Python* (O’Reilly, 2009, ISBN 978-0-596-51649-9). See also the [nltk.org](http://nltk.org) website.

There is a CTools site. Handouts and assignments will be posted there.

## 2 Vitals

Lectures: Tue and Thu, 2:30–4:00, 173 Lorch Hall

Instructor: Steven Abney ([abney@umich.edu](mailto:abney@umich.edu)), 412 Lorch Hall

Office hours: It is safest to make an appointment. You are also welcome to drop by, if you can catch me in. I am *usually* available on Tue and Thu 1:00-2:30.

## 3 Requirements

There will be regular homeworks, a midterm exam, and a final exam.

## 4 Schedule

The plan is to work through most of the book, though not linearly. The following is a tentative schedule. See the CTools site for the current schedule as the term proceeds.

Week 1	Introduction (Ch. 1)
Week 2	Python, texts
Week 3	Syntax (Ch. 8)
Week 4	Grammars and parsers
Week 5	Feature grammars (Ch. 9)
Week 6	Movement. Predicate calculus.
Week 7	Semantic interpretation (Ch. 10)
Week 8	Review, midterm.
Week 9	Corpora, frequency distributions (Ch. 2)
Week 10	Conditional dists, lexical resources
Week 11	Regular expressions, characters (Ch. 3)
Week 12	Part of speech tagging (Ch. 5)
Week 13	Classification (Ch. 6)
Week 14	Information extraction (Ch. 7)
Dec. 19	Final exam (1:30–3:30)