Ling209B: Computational Linguistics 2  
Fall 2019

Instructor  
Tim Hunter  
3103K Campbell Hall  
timhunter@ucla.edu

Classes  
Mon & Wed, 10:00am – 11:50am, Haines A78

Office Hours  
Mon, 1:00pm – 3:00pm; or by appointment

Course overview

The topic for this offering of the course will be **probabilistic grammars**.

Two big ideas that we will focus on are:

- The role of **inductive bias** in learning and generalization.
  This is a central idea in generative linguistics. It’s not specific to probabilistic settings but these settings provide one natural way to approach the topic.

- The relationship between probabilistic grammars and **classical formal grammars**.
  There are very tight connections here (which I think are underappreciated). Understanding the connections shrinks the apparent gap between probabilistic and non-probabilistic cultures in linguistics.

For the first part of the course we’ll cover some bread-and-butter topics in more or less lecture-style format, to try to sharpen up our theoretical understanding of these two key ideas. Then roughly halfway through the quarter we’ll transition to a more (pro)seminar-style format, reading research papers and considering specific applications.

Topics

Bread-and-butter foundations (about four weeks?):

1. Grammars as inductive bias; n-gram models
2. Probabilistic finite-state automata; probabilistic context-free grammars
3. Learning with hidden structure via expectation-maximization (the forward-backward and inside-outside algorithms)
4. Log-linear/maximum-entropy models and integrating them into grammars

Applications/extensions building on these ideas (with possible readings):

- Bayes and PCFGs
  - Johnson et al. (2007), “Bayesian inference for PCFGs via Markov chain Monte Carlo”
– Johnson (2008), “Using adaptor grammars to identify synergies in the unsupervised acquisition of linguistic structure”

• Neural networks
  – Goldberg (2017), Neural Network Methods for NLP
  – Linzen et al. (2016), “Assessing the ability of LSTMs to learn syntax-sensitive dependencies”

• The relationship between frequency/probability and acceptability
  – Lau et al. (2017), “Grammaticality, acceptability and probability: A probabilistic view of linguistic knowledge”
  – Sprouse et al. (2018), “Colorless green ideas do sleep furiously: Gradient acceptability and the nature of grammar”

• Integrating prosody into PCFGs (?)

• Inductive bias in (non-probabilistic) grammar induction
  – Heinz et al. (2015), Grammatical inference for computational linguistics

…

Tim’s pet favourites and current thoughts:

• Sentence probabilities mediated by phonological effects at word-boundaries

• Probabilities on minimalist grammars
  – Hunter (to appear), “What kind of cognitive hypothesis is a derivational grammar?”

• Information-theoretic complexity metrics for sentence processing difficulty
  – Hale (2006), “Uncertainty about the rest of the sentence”

Interesting adjacent theoretical ideas (unsure about applications yet):

• Joint (generative) versus conditional (discriminative) models

• Non-probabilistic gradience in grammars (e.g. classical OT, degrees of grammaticality)

• Probabilistic finite-state transducers

Requirements and Grading

• Class participation. Come to class having read and thought about any assigned readings; ask questions when things are not clear; think/ask about connections to your own areas of research interest.

• Presentation of one paper in class. Somewhere around Week 7 or Week 8; plan on around 50 minutes, including discussion.

• Final project/paper. A short term paper describing a small project using some of the ideas from the course, or reviewing some relevant published research. Due in exam week.

• Presentation based on final/project paper. In Week 10; maybe around 30 minutes, including discussion.

For a two-credit Ling596A directed study, there’s only the class participation requirement.