Schedule of Topics
CSC/LIN 205, “Computational Linguistics”
Department of Computer Science
Grinnell College
revised November 8, 2018

August 31: Computational approaches to the study of language
Reading: “A Scheme refresher” (handout); “Scheme in DrRacket” (handout); Revised Report on the Algorithmic Language Scheme, sections 5.6 and 5.7 (pages 28–30).

September 3: Lab: Libraries in R7RS Scheme
Reading: Revised Report on the Algorithmic Language Scheme, section 5.5 (page 27); “Records in R7RS Scheme” (handout).

September 5: Lab: Records in R7RS Scheme
“The Absolute Minimum Every Software Developer Absolutely, Positively Must Know about Unicode and Character Sets (No Excuses!)” (Joel Spolsky, http://www.joelonsoftware.com/articles/Unicode.html);
“UTF-8 and Unicode FAQ for Unix/Linux” (Markus Kuhn, https://www.cl.cam.ac.uk/~mgk25/unicode.html).

September 7: Unicode
Reading: Jurafsky and Martin, second edition, chapter 1 (pages 1–16); Jurafsky and Martin, draft third edition, chapter 2, from the beginning through section 2.1.

September 10: Regular Expressions
Reading: Jurafsky and Martin, draft third edition, sections 2.2 and 2.3 and section 2.4 from the beginning through section 2.4.3.

September 12: Tokenization
Reading: Handout: “Hash Tables and Hash Functions.”

September 14 and 17: Spectra
Reading: Jurafsky and Martin, draft third edition, sections 2.4.4 and 2.4.5.

September 19: Lemmatization
Reading: Jurafsky and Martin, draft third edition, from section 2.5 to the end of chapter 2.

September 21: Edit Distance
Reading: Jurafsky and Martin, second edition, section 2.2 (pages 26–38).

September 24: Finite-State Automata
Reading: Jurafsky and Martin, second edition, from section 2.3 to the end of chapter 2 (pages 38–44).

September 26: Compiling Automata
Reading: Jurafsky and Martin, draft third edition, from the beginning of chapter 3 through section 3.1.

September 28: N-Grams
Reading: Jurafsky and Martin, draft third edition, sections 3.2, 3.3, and 3.7.

October 1: Perplexity, Information, and Entropy
October 3: Smoothing
Reading: Jurafsky and Martin, draft third edition, from the beginning of chapter 4 through section 4.4.

October 5: Sentiment Classification
Reading: Jurafsky and Martin, draft third edition, sections 4.5 through 4.8.

October 8: Spam Detection
Reading: Jurafsky and Martin, draft third edition, from section 4.9 to the end of chapter 4.

October 10 and 12: The Bootstrap Algorithm
Reading: Jurafsky and Martin, draft third edition, from the beginning of chapter 5 through section 5.4 and from section 5.9 to the end of chapter 5.

October 15: Logistic Regression
Reading: Jurafsky and Martin, draft third edition, chapter 6.

October 17: Vector Semantics
Reading: Jurafsky and Martin, draft third edition, from the beginning of chapter 7 through section 7.3.

October 19: Neural Networks
Reading: Jurafsky and Martin, draft third edition, from section 7.4 to the end of chapter 7.

October 29: Training Neural Networks

October 31: Training Neural Networks (continued)
Reading: Jurafsky and Martin, second edition, from the beginning of chapter 6 through section 6.3 (pages 173–183).

November 2: Hidden Markov Models
Reading: Jurafsky and Martin, draft third edition, from the beginning of chapter 8 through section 8.3.

November 5: Syntactic Categories
Reading: Jurafsky and Martin, draft third edition, from section 8.4 to the end of chapter 8.

November 7: Tagging with Hidden Markov Models
Reading: Jurafsky and Martin, second edition, sections 5.4 and 5.6 (pages 135–139 and 151–153).

November 9: Transformation-based Tagging
Reading: Jurafsky and Martin, second edition, section 5.9 (pages 163–168).

November 12: Spelling Correction
Reading: Jurafsky and Martin, draft third edition, from the beginning of chapter 10 through section 10.3.

November 14: Context-Free Grammars
Reading: Jurafsky and Martin, draft third edition, from section 10.4 to the end of chapter 10.

November 16: Syntax Trees and the Penn Treebank
Reading: Jurafsky and Martin, draft third edition, from the beginning of chapter 11 through section 11.1.

November 19: Parsing
Reading: Jurafsky and Martin, draft third edition, section 11.2.

November 21: The Cocke–Younger–Kasami Parsing Algorithm
Reading: Jurafsky and Martin, second edition, subsection 13.4.2 (pages 443–448).

**November 26:** The Earley Parsing Algorithm

Reading: Jurafsky and Martin, draft third edition, chapter 13.

**November 28:** Dependency Parsing

Reading: Jurafsky and Martin, second edition, from the beginning of chapter 15 through section 15.1 (pages 489–492).

**November 30:** Feature Structures

Reading: Jurafsky and Martin, second edition, sections 15.2 and 15.3 (pages 492–507) and figure 15.8 (page 511).

**December 3:** Unification

Reading: Jurafsky and Martin, draft third edition, chapter 17.

**December 5:** Information Extraction

Reading: Jurafsky and Martin, draft third edition, chapter 19.

**December 7:** Disambiguation

Reading: Jurafsky and Martin, draft third edition, chapter 25.

**December 10:** Conversational Agents

Reading: Jurafsky and Martin, second edition, chapter 25 (pages 859–908).

**December 12:** Machine Translation

**December 14:** Review; student evaluations

**Wednesday, December 19, 9 a.m.:** Final examination

Copyright © 2016, 2018 John David Stone

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit

http://creativecommons.org/licenses/by-sa/4.0/deed.en_US

or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.