Schedule of Topics
CSC/LIN 205, “Computational Linguistics”
Department of Computer Science
Grinnell College
revised November 2, 2016

August 26: Computational approaches to the study of language
Reading: Manning and Schütze, chapter 1 (pages 1–36); Revised Report on the Algorithmic Language Scheme, sections 5.6 and 5.7 (pages 28–30)

August 29: Libraries in R\textsuperscript{7}RS Small Scheme
Reading: Revised Report on the Algorithmic Language Scheme, section 5.5 (page 27)

August 31: Records in R\textsuperscript{7}RS Small Scheme

September 2: Programming GNU Emacs
Reading: Manning and Schütze, chapter 2, from the beginning through section 2.1 (pages 38–60)

September 5: Probability: Making rational inferences under uncertainty
Reading: Manning and Schütze, from section 2.2 to the end of chapter 2 (pages 60–80)

September 7: Information theory
Reading: Manning and Schütze, chapter 3 from the beginning through section 3.1 (pages 81–93)

September 9: Syntactic categories
Reading: Manning and Schütze, chapter 14 introduction (pages 495–500)

September 12 and 14: Clustering
Reading: Manning and Schütze, from section 3.2 to the end of chapter 3 (pages 93–115)

September 16: Context-free grammars

September 19 and 21: Chomsky Normal Form and the Cocke–Younger–Kasami parsing algorithm

September 23 and 26: Earley parsing
Reading: Manning and Schütze, chapter 4 from the beginning through section 4.1 (pages 117–122)

September 28: Corpus linguistics
Reading: “The Unicode\textsuperscript{®} Standard: A Technical Introduction” (Unicode Consortium, http://www.unicode.org/standard/principles.html);
“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 30: Unicode
Reading: Manning and Schütze, section 4.2 (pages 123–136)

October 3: Tokenizers
Reading: Manning and Schütze, from section 4.3 to the end of chapter 4 (pages 136–147)

October 5: Regular expressions and pattern matching

October 7: Finite-state automata

October 10 and 12: Compiling automata
Reading: Manning and Schütze, chapter 5 from the beginning through section 5.2 (pages 149–162)

**October 14 and 24:** Finding collocations

Reading: Manning and Schütze, from section 5.3 to the end of chapter 5 (pages 162–189)

**October 26 and 28:** Hypothesis testing

Reading: Manning and Schütze, chapter 6 from the beginning through section 6.1 (pages 191–196)

**October 31:** \( n \)-gram models

Reading: Manning and Schütze, from section 6.2 to the end of chapter 6 (pages 196–227)

**November 2 and 4:** Estimators

Reading: Manning and Schütze, chapter 9 from the beginning through section 9.1 (pages 315–320)

**November 7:** Markov models

Reading: Manning and Schütze, section 9.2 and section 9.3 from the beginning through section 9.3.2 (pages 320–333)

**November 9 and 11:** Hidden Markov models: decoding and training

Reading: Manning and Schütze, from section 9.3.3 to the end of chapter 9 (pages 333–340)

**November 14 and 16:** Hidden Markov models: parameter estimation

Reading: Manning and Schütze, chapter 10 from the beginning through section 10.2 (pages 341–356)

**November 18:** Tagging

Reading: Manning and Schütze, section 10.3 (pages 356–361)

**November 21 and 23:** Tagging with hidden Markov models

Reading: Manning and Schütze, section 10.4 (pages 361–370)

**November 28 and 30:** Transformation-based tagging

Reading: Manning and Schütze, chapter 11 (pages 381–405)

**December 2:** Probabilistic context-free grammars

Reading: Manning and Schütze, chapter 12 (pages 407–460)

**December 5 and 7:** Probabilistic parsing

**December 9:** Review; student evaluations

**Thursday, December 15, 9 a.m. to noon:** Final examination