Midterm Review

9.19, Instructor: Roger Levy

20 October 2023

Items in gray were things we didn't cover in class during this semester, or covered extremely briefly, and which won't be on the midterm, but which were in assigned and/or additional readings and that I recommend you spend some time reviewing to get the most out of everything we've covered so far in this course.

- Tools from probability theory & relatively direct applications
  - joint probability; conditional probability; the chain rule; conditional independence
  - Bayes rule
- Probability distributions & probabilistic models
  - Bernoulli & binomial distributions
  - Multinomial distribution
  - Gaussian/normal distribution
  - Dimensionality reduction methods (seen in word embedding models)
- Statistical parameter estimation
  - Method of maximum likelihood & relation to relative frequency estimation for multinomial models
  - Data sparsity (everywhere)
- Science of cognition throughout the semester)
  - Rational analysis (Anderson)
    - Marr's levels: computational, algorithmic, implementational
  - How to construct and test formally explicit scientific theories of cognition (here: applications to language)
- Experimental methods for studying language processing
  - Acceptability judgments

    - Recognition studies (lexical decision; naming; speech recognition; same/different tasks)
  - Self-paced reading and eye-tracking for studying reading
    - Neural methods: EEG, fMRI, ECoG
    - Visual world paradigm
- Applications in speech processing
  - Simple sound categorization
  - Word segmentation ambiguity
    - Sentence-level speech recognition; segmentation ambiguity
- "Simple" classification models
  - Simple categorization from a univariate continuous input (as in sound categorization)
- Formal language theory & model classes
  - Feature representations for phonology
  - Regular expressions
  - Finite-state automata (and relation to regex's and regular languages)
    - Finite-state transducers
  - Context-free grammars; arguments against natural languages being regular languages
    - Weighted finite-state machines
- Algorithms for formal language classes
  - Checking acceptance/rejection of strings in deterministic and non-deterministic FSMs
    - Dynamic programming parsing algorithms for context-free grammars
- Word embeddings
  - Dense vs. sparse word representations
  - How word embeddings are learned
  - Similarity and analogy in word-embedding spaces
  - Implicit bias in word embeddings
- "Language models" (probability distributions over word sequences)

- n-gram models
  - Probabilistic FSAs
- Autoregressive neural language models, specifically GPT-2, as "black-box" models