Computational Seminar: Black Box Models of Language

Ling 7710, Autumn 2019
W 4:30-7:00, Morrill 110

Instructor: Marten van Schijndel
Email: mv443@cornell.edu
Office: 220 Morrill Hall
Office Hours: Tues & Wed 11:30-12:30
& By Appointment

Course Website: canvas.cornell.edu
Mailbox: 203 Morrill Hall
Office Phone: (607) 255-1105

Course Description: How do humans represent and process language? Psycholinguists have studied language processing for many years and have developed a number of techniques for probing human language processing. Their experimental methods, however, have been limited to the small number of humans they can get to participate in their experiments.
Concurrently, computational linguists have developed methods for efficiently mining patterns from vast amounts of digital language data. While highly accurate, these approaches have become very opaque due to increased usage of neural network deep learning tools.
This seminar will focus on how these two lines of research can complement each other. The psycholinguistic techniques which were developed to study human language processing can be used to study current neural network methods of language processing. Neural networks can be used to identify large scale patterns of naturally occurring linguistic data, giving clues as to how language is used in the wild. In addition, neural networks can be used to test poverty of the stimulus arguments about the specific biases and inputs required to learn linguistic patterns. Each of these research trajectories can fruitfully contribute to the other, and I hope that this course will enable you to leverage both of these approaches in your own work.

Important: If you email me, preface your email subject with **bblang** so that I will be able to see it and respond. I get a lot of email, so if you can help me find your email, I'll be able to respond quicker. Note also that I check email less on the weekends.

Grading: This course revolves around discussions of research, so participation is a large component of the grade in this course. To facilitate discussion, students will be required to post responses to each week's reading to canvas the day before each class. Students are also expected to present relevant research papers or their own ongoing research. At the end of the course, students are expected to write a project proposal for a computational linguistics or cognitive science conference paper.[1]

Reading responses	10% (+5%)								
Presentations	20%	105–97	A+	89.9–87	B+	79.9–77	C+	69.9–67	D+
Participation/discussion	20%	96.9–93	A	86.9–83	B	76.9–73	C	66.9–60	D
Final project proposal	50%	92.9–90	A-	82.9–80	B-	72.9–70	C-	Below 60	F

Reading Responses: Each person will need to submit reading responses via Canvas at least 24 hours in advance of the course. These need not be long, but they should demonstrate that you completed the reading and thought about it. They should contain a summary of the paper and a question to discuss in class. Five responses can be missed without penalty, and especially insightful responses will earn a point of extra credit (up to five).

[1] Alternatively, students can submit a draft of an actual conference paper developed during this course.

Presentations: Students will present papers they find interesting to the rest of the class and will lead discussion of those papers. Students can also present ongoing work which is relevant to the class.

Participation: Students are expected to engage with one another in detailed discussion of the research. Students should be informed and present during discussions and should regularly contribute. Astute, detailed contributions made as reading responses can stand in for constant speech during discussion periods, but students should each contribute to the course discussions in some way and should occasionally contribute verbally.

Final project proposal: Students will submit a project proposal or project write up at the end of the course. The project could use computational techniques we cover to model linguistic data in a new way, could use psycholinguistic testing to test model predictions made by current computational systems, or anything else related to probing human or neural network language processing.

- Project topic proposed to me by October 23 (5%)
- In-person discussion with me about the project by November 10 (10%)
- Project proposal/paper (4-5 pages) to me by December 9
  - Research question (15%)
  - Background literature (20%)
  - Experimental design (20%)
  - Explanation of needed resources (incl. sample size calculation) (10%)
  - Explanation of possible findings (what would each possible outcome mean?) (20%)
  - Collaboration justification (if working in a group)

Final project proposals can be submitted as a group if A) I can be convinced beforehand that each student would bring individual expertise to the project, B) each student contributes to the write up, and C) the group submits a written overview of the work contributed by each student, signed by each student.
Please write the proposal using the CogSci LaTeX template. Refer to Overleaf's 30 minute tutorial if you are new to LaTeX.

Special Accommodations: Please give me any Student Disability Services (SDS) accommodation letter as early as possible so that I can arrange for needed academic accommodations. If you need an immediate accommodation, please speak with me after class or email me and/or SDS at `sds_cu@cornell.edu`. SDS is located on level 5 of Cornell Health, 110 Ho Plaza, 607-254-4545, `https://sds.cornell.edu/`

Academic Integrity: Please don't cheat, and please cite people whose work you use. I am **required** to follow-up on suspected violations of Cornell's *Code of Academic Integrity*. I encourage students to watch this video to learn more about what constitutes cheating and why it matters: `https://cheatingvideo.provost.cornell.edu/`

If you have questions about issues of academic integrity, please see me.

Reading

This is an initial list of papers we might read during this course. Students are encouraged to suggest additional papers that would be both relevant to this seminar and relevant to their research. Depending on the interests of the students and the speed with which we progress through the material, we may not read all of these papers.

Neural network models

- Chris Olah (2015). Understanding LSTM networks.
- Urvashi Khandelwal, He He, Peng Qi & Dan Jurafsky (2019) Sharp Nearby, Fuzzy Far Away: How Neural Language Models Use Context. Proceedings of ACL.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez & Łukasz Kaiser (2017). Attention is all you need. Proceedings of NIPS.
  - Jay Alammar (2018). The Illustrated Transformer
  - Alexander Rush (2018). The Annotated Transformer
- Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Ruslan Salakhutdinov & Quoc V. Le. (preprint). XLNet: Generalized Autoregressive Pretraining for Language Understanding
  - Keita Kurita (2019). Paper Dissected: XLNet: Generalized Autoregressive Pretraining for Language Understanding Explained (Covers BERT and XLNet)
- Chris Dyer, Adhiguna Kuncoro, Miguel Ballesteros & Noah Smith (2016). Recurrent Neural Network Grammars. Proceedings of NAACL.

Predictive processing in humans

- Roger Levy (2013). Memory and surprisal in sentence processing. Pages 78–87 and 99–103. Sentence Processing.
- Nathaniel Smith & Roger Levy (2013). The effect of word predictability on reading time is logarithmic. Cognition.
- Cory Shain (2019). Prediction is all you need: A large-scale deconvolutional study of predictability and frequency effects in naturalistic reading. Proceedings of NAACL.
- Stefan L. Frank, Leun J. Otten, Giulia Galli & Gabriella Vigliocco (2015). The ERP response to the amount of information conveyed by words in sentences. Brain and Language.
- Francis Mollica & Steven T. Piantadosi (2017). An incremental information-theoretic buffer supports sentence processing. Proceedings of the Annual Conference of the Cognitive Science Society.
- Marten van Schijndel & Tal Linzen (2018). A neural model of adaptation in reading. Proceedings of EMNLP.
- Nieuwland et al. (2017). Limits on prediction in language comprehension: A multi-lab failure to replicate evidence for probabilistic pre-activation of phonology.

Syntax

- Tal Linzen, Emmanuel Dupoux & Yoav Goldberg (2016) Assessing the ability of LSTMs to learn syntax-sensitive dependencies. Transactions of the Association for Computational Linguistics.
- Mario Giulianelli, Jack Harding, Florian Mohnert, Dieuwke Hupkes & Willem Zuidema (2018).Under the Hood: Using Diagnostic Classifiers to Investigate and Improve how Language Models Track Agreement Information. Proceedings of BlackBoxNLP.
- Shauli Ravfogel, Yoav Goldberg & Tal Linzen (2019). Studying the Inductive Biases of RNNs with Synthetic Variations of Natural Languages. Proceedings of NAACL.

Word vectors and Semantics

- Allison Ettinger & Tal Linzen (2016). Evaluating vector space models using human semantic priming results. Proceedings of RepEval.
- Tal Linzen (2016). Issues in evaluating semantic spaces using word analogies. Proceedings of RepEval.
- Maria Antoniak & David Mimno (2018). Evaluating the Stability of Embedding-based Word Similarities. TACL.
- Jayden Ziegler & Jesse Snedeker (2018). How broad are thematic roles? Evidence from structural priming. Cognition.

Pragmatics

- Hannah Rohde, Roger Levy & Andrew Kehler (2011). Anticipating Explanations in Relative Clause Processing. Cognition.
- Andrew Kehler & Hannah Rohde (2018). Prominence and coherence in a Bayesian theory of pronoun interpretation. Journal of Pragmatics, Special Issue on Prominence in Pragmatics.
- Justine T. Kao, Jean Y. Wu, Leon Bergen & Noah D. Goodman (2014). Nonliteral understanding of number words. PNAS.
- Noah Goodman & Michael Frank (2016). Pragmatic language interpretation as probabilistic inference. Trends in Cognitive Sciences.

Prosody

- Amy J. Schafer, Amber Camp, Hannah Rohde, & Theres Grüter (2018). Contrastive prosody and the subsequent mention of alternatives during discourse processing. Grammatical Approaches to Language Processing – Essays in Honor of Lyn Frazier.
- Trang Tran, Shubham Toshniwal, Mohit Bansal, Kevin Gimpel, Karen Livescu & Mari Ostendorf (2018). Parsing Speech: A Neural Approach to Integrating Lexical and Acoustic-Prosodic Information. Proceedings of NAACL.
- Vicky Zayats & Mari Ostendorf (2019). Giving attention to the unexpected: Using prosody innovations in disfluency detection. Proceedings of NAACL.

- Mara Breen (2014). Empirical Investigations of the Role of Implicit Prosody in Sentence Processing. Language and Linguistics Compass.
- Frank Wijnen (2004). The implicit prosody of Jabberwocky and the relative clause attachment riddle. On Speech and Language. Studies for Sieb G. Nooteboom.

Language acquisition

- Cory Shain & Micha Elsner. Measuring the perceptual availability of phonological features during language acquisition using unsupervised binary stochastic autoencoders. Proceedings of NAACL.
- Francis Mollica & Steven T. Piantadosi. Humans store about 1.5 megabytes of information during language acquisition. Royal Society Open Science.
- Tom McCoy, Bob Frank & Tal Linzen (2018). Revisiting the poverty of the stimulus: Hierarchical generalization without a hierarchical bias in recurrent neural networks. Proceedings of the Annual Conference of the Cognitive Science Society.