Seminar in Semantics: Cognitive Models of Discourse Comprehension and the Symbol Binding Problem

Instructor: Adrian Brasoveanu, Linguistics, UC Santa Cruz

Spring 2025

- **Meeting Time:** TuTh, 9:50-11:25am

- **Location:** Stevenson Library 102 (The Cave)

- **Office Hours:** TuTh 9-9:45am or by appointment (`abrsvn@ucsc.edu`)

- **Canvas site:** https://canvas.ucsc.edu/courses/82236

Contents

1 Course Description

This seminar is concerned with semantic theories of natural languages, that is, with the fundamental problem of how the human mind assigns meanings – non-linguistic representations of (parts of) the

world – to linguistic / symbolic forms. The approach we will take is to step back from formal semantics theories, which are one specific strategy of approaching this problem, and consider a range of semantic theories in the broader interdisciplinary field of cognitive science.

While we broaden our view, we will not lose focus. We will always be concerned with meanings as non-linguistic, truth-conditional representations associated with linguistic forms. Even as we discuss approaches outside traditional formal semantics and generative linguistics, we will be specific and not confuse meaning in the formal semantics sense, which is our concern in this course, with other notions like logical forms or word (token / form) co-occurrence, which can be successfully used as proxies for meaning in various contexts.

The specific empirical domain we will focus on is (cognitive models of) discourse comprehension. But the question of how we give meanings to forms is a fundamental problem in cognitive science, so our secondary focus will be the broader issue of whether symbols – form/meaning pairs in our particular case – are irreducible building blocks of cognition or can emerge in non-symbolic systems, known as the symbol binding problem.

Most of our time will be spent on experimental results and cognitive models of discourse comprehension, focusing on relatively simple narratives and conditionals. Towards the end of the quarter, our explorations will bring us increasingly closer to the bigger picture associated with the symbol binding problem. We will end the course with recent work at the interface of machine (deep) learning and cognitive science arguing that specific architectures for artificial neural networks can provide strong inductive biases for emergent symbols and abstract rules (without surreptitiously building symbols in).

2 Weekly Schedule

All the readings and additional materials are available on the Canvas site for the class. Some of the course units have overlapping temporal traces – we'll try to finish them in a week, but sometimes we'll take a little more time, and sometimes we'll take a little less, as needed.

2.1 Week 1: Situation Models & Inferences in Discourse Comprehension, and the Symbol Binding Problem

- Zwaan, R. A. (2001). Situation Models. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences*. Amsterdam: Elsevier.

- Noordman, L. G. M., & Vonk, W. (2015). Inferences in Discourse. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed.). Amsterdam: Elsevier.

- Kautz, H. A. (2022). The Third AI Summer: AAAI Robert S. Engelmore Memorial Lecture. *AI Magazine, 43*(1), 105–125. https://doi.org/10.1002/aaai.12036

- Greff, K., van Steenkiste, S., & Schmidhuber, J. (2020). On the Binding Problem in Artificial Neural Networks. *arXiv:2012.05208 [cs.NE]*. https://doi.org/10.48550/arXiv.2012.05208

Background materials (videos): brief intro to neural networks, gradient descent, backpropagation. Please make sure to watch them before the first class if these don't sound familiar to you.

2.2 Weeks 2-3: The Sentence Gestalt and Story Gestalt Models of Discourse Comprehension

- St. John, M. F., & McClelland, J. L. (1990). Learning and Applying Contextual Constraints in Sentence Comprehension. *Artificial Intelligence, 46*(1-2), 217–257. https://doi.org/10.1016/0004-3702(90)90008-N

- St. John, M. F., & McClelland, J. L. (1992). The Story Gestalt: A Model of Knowledge-Intensive Processes in Text Comprehension. *Cognitive Science, 16*(2), 271–306. https://doi.org/10.1207/s15516709cog1602_5

Background materials (videos): backpropagation in detail, probability and cross-entropy loss. Please make sure to watch them as soon as possible if these don't sound familiar to you.

2.3 Weeks 3-4: The Golden & Rumelhart Model of Story Comprehension and Recall

- Golden, R. M., & Rumelhart, D. E. (1993). A Parallel Distributed Processing Model of Story Comprehension and Recall. *Discourse Processes, 16*(3), 203–237. https://doi.org/10.1080/01638539309544839

- Golden, R. M., Rumelhart, D. E., Strickland, J., & Ting, A. (1994). Markov Random Fields for Text Comprehension. In D. S. Levine & M. Aparicio IV (Eds.), *Neural Networks for Knowledge Representation and Inference.* Psychology Press.

2.4 Weeks 4-5: Knowledge Based Inferences and Coherence Driven Pronoun Resolution in Story Comprehension

- Frank, S. L., Koppen, M., Noordman, L. G. M., & Vonk, W. (2003). Modeling knowledge-based inferences in story comprehension. *Cognitive Science*, 27, 875-910.

- Frank, S. L., Koppen, M., Noordman, L. G. M., & Vonk, W. (2007). Coherence-driven resolution of referential ambiguity: A computational model. *Memory & Cognition*, 35(6), 1307-1322.

Background materials: self-organizing maps (paper), brief intro to differential equations (video).

2.5 Weeks 5-6: Causal and Temporal Inferences in Counterfactuals

- Lewis, D. (1979). Counterfactual Dependence and Time's Arrow. *Noûs*, 13(4), 455-476.

- Byrne, R. M. J. (2011). Counterfactuals and causal thoughts about exceptional events. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.), *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology* (pp. 208-229). Oxford: Oxford University Press.

- Khoo, J. (2017). Backtracking counterfactuals revisited. *Mind*, 126(503), 841-910.

- Byrne, R. M. J., & Johnson-Laird, P. N. (2019). If and or: Real and counterfactual possibilities in their truth and probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(4), 760-780.

2.6 Weeks 7-8: RNN Models of Semantic (Truth-Conditional) Interpretation

- Frank, S. L., Haselager, W. F. G., van Rooij, I. (2009). Connectionist semantic systematicity. *Cognition*, 110(3), 358-379.

- Venhuizen, N. J., Hendriks, P., Crocker, M. W., Brouwer, H. (2022). Distributional formal semantics. *Information and Computation*, 287, 104763.

- Frank, S. L., Vigliocco, G. (2011). Sentence comprehension as mental simulation: An information-theoretic perspective. *Information*, 2(4), 672-696.

- Venhuizen, N. J., Crocker, M. W., Brouwer, H. (2019). Semantic entropy in language comprehension. *Entropy*, 21(12), 1159.

- **Week 8: Adrian at SALT**

- **Week 8: Final project proposals are due**

2.7 Weeks 9-10: Architectural Inductive Biases in Neural Networks for Emergent Symbol Binding

- Webb, T. W., Sinha, I., & Cohen, J. D. (2021). Emergent symbols through binding in external memory. In *International Conference on Learning Representations (ICLR)*. https://openreview.net/forum?id=LSFCEb3GYU7

- Dulberg, Z., Webb, T., & Cohen, J. (2021). Modelling the development of counting with memory-augmented neural networks. In T. Fitch, C. Lamm, H. Leder, K. Teßmar-Raible (Eds.), *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society* (pp. 2787-2793).

- Webb, T. W., Frankland, S. M., Altabaa, A., Segert, S., Krishnamurthy, K., Campbell, D., Russin, J., Giallanza, T., Dulberg, Z., O'Reilly, R., Lafferty, J., & Cohen, J. D. (2024). The relational bottleneck as an inductive bias for efficient abstraction. *Trends in Cognitive Sciences*, 28(9), 829-843. https://doi.org/10.1016/j.tics.2024.04.001

- **Week 10: Brief final project presentations**

Background materials: attention in transformers (video).

3 Student Evaluation

- **Attending all classes and active engagement in class is expected.**

- **Each registered student will own one of the above units / 'weeks'**, except for the first week which I will fully manage (you are of course still expected to do the week 1 readings and watch the background videos if you need to). Owning a unit / 'week' means:

    - preparing slides for each of the readings for that unit – slides are the expected way to present academic work in this class; the slides will be shared on the Canvas site for the class after they are presented

    - leading the discussion of the readings and facilitating class engagement

If you want to team up with other registered students, please do, this is strongly encouraged. A team of n participants will then together own n units / 'weeks'. Please discuss a fair division of labor – or at least a process for fairly dividing the labor – from the very beginning, so that every member of the team has a chance to contribute fairly and in an accountable manner.

Please decide what unit you want to be responsible for and let me know by the end of the first class.

- **Final Project Proposal in Week 8:** A brief but contentful proposal for the final project, due at the end of Week 8. Aim for 2-3 pages, or 8-10 slides. Again, you can team up. By default, a team with n members is expected to contribute roughly $n\times$ the above requirements, but the nature of the research work is different from presenting existing work and facilitating discussions, so if you form a team, let's talk about expectations as soon as possible. Please discuss early on what the team-internal division of labor will be, and share your plans with me as soon as possible.

- **Brief Final Project Presentation in Week 10:** Aim for 10-15 minutes, which conveniently is probably around 8 slides. Teams are encouraged, same n multiplier applies but with the same caveat about the different nature of doing research work. Same early convos about division of labor are expected, same expectation to share plans with me as soon as possible.

- **Final Project:** Research paper addressing a topic relevant to the course themes. Due during finals week; the tentative deadline (to be confirmed) is the Thursday of the final week, 5 pm. Again, teams strongly encouraged, with the attending expectations for fair division of labor, and sharing plans with me as early as possible.

4   Course Policies

4.1   Attendance

Regular in-person attendance is expected for this seminar. If you must miss a class, please notify me in advance.

4.2   Academic Integrity

All work submitted must be your own. Proper citation of sources, explicitly listing collaborators etc. etc. is expected throughout the course.

4.3   Accommodations

If you require accommodations, please let me know as early as possible.

Syllabus subject to change. Any modifications will be announced in class or on the Canvas site.