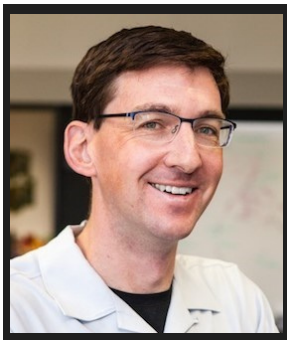




Carnegie Mellon University
Language Technologies Institute



Christopher Potts Stanford University

Christopher Potts is Professor and Chair of Linguistics and Professor (by courtesy) of Computer Science at Stanford, and a faculty member in the Stanford NLP Group and the Stanford AI Lab. His group uses computational methods to explore how emotion is expressed in language and how linguistic production and interpretation are influenced by the context of utterance. This research combines methods from linguistics, cognitive psychology, and computer science, in the service of both scientific discovery and technology development. He is the author of the 2005 book *The Logic of Conventional Implicatures*, as well as numerous scholarly papers in computational and theoretical linguistics.

Towards more Meaningful Benchmarks for Natural Language Understanding

It is common to hear that certain natural language processing tasks have been "solved". These claims are often misconstrued as being about general human capabilities (e.g., to answer questions, to reason with language), but they are always actually about how systems performed on fixed, narrowly defined benchmarks, and there remains a significant gap between benchmark results and real-world utility. The field has responded with a wide range of new evaluation methods. In this talk, I'll report on recent efforts to create *dynamic* benchmarks that evolve with our models, continually presenting them with new tasks to solve. Dynamic benchmarks present exciting new opportunities, but they also bring new challenges. Are the resulting tasks solvable? Do they actually lead to more robust models? Might they embed problematic artifacts and pernicious biases? I'll address these and other questions, and report on the steps we have taken to address them as part of developing a new dynamic benchmark for sentiment analysis. I'll close with discussion of what properties I suspect general-purpose architectures will need to have to truly solve deep semantic tasks.

Friday, April 2, 2021

2:20 - 3:40 PM EST

Join the meeting on Zoom

Meeting ID 935 3287 1380 Passcode 546823

LTI Colloquium 2020-21
is generously sponsored by

abridge