



Carnegie Mellon University
Language Technologies Institute

Jesse Dunietz **Elemental Cognition / MIT**



Jesse Dunietz is a researcher at Elemental Cognition and an instructor at the MIT Communication Lab. He completed his Ph.D. in the Computer Science Department at CMU, where he built annotated datasets and automated taggers for language about cause and effect. At Elemental Cognition, he has focused on defining the desired capabilities for the company's deep natural language understanding systems, and is currently involved in building prototype applications for the company's technology. Jesse has also worked extensively in science communication, both as a science writer and a communication trainer, with an emphasis on communicating with the public about computer science.

To Test Machine Comprehension, Start by Defining Comprehension

Many NLP tasks aim to measure "machine reading comprehension" (MRC), with the last few years in particular seeing a profusion of new datasets. MRC tasks are typically designed to challenge systems with questions believed to be difficult for today's NLP. This talk will argue for an alternative approach: instead of emphasizing difficulty, task design should start from what content—what information expressed, implied, or relied on by a passage—is needed for real-world applications.

The presentation will cover Elemental Cognition's efforts to turn this idea into a more rigorous test of comprehension, particularly for applications that rely on short narratives. I will present our "template of understanding" (ToU)—a detailed definition of story comprehension—and a small experiment that strongly suggests current systems struggle to answer ToU-based questions. I will also discuss our progress on and the remaining challenges for a larger ToU-based evaluation dataset, which we hope will provide a compelling platform for measuring story comprehension.

Friday, March 12, 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