CLSP & CS
VIRTUAL SEMINAR SERIES

Rowan Zellers

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING
Computer Science

"Grounding Language by Seeing, Hearing, and Interacting"

Monday, February 14, 2022

12:00 - 1:15 PM

https://wse.zoom.us/j/91466735404
Meeting ID: 914 6673 5404

ABSTRACT

In my talk, I will discuss three lines of work to bridge this gap between machines and humans. I will first discuss how we might measure grounded understanding. I will introduce a suite of approaches for constructing benchmarks, using machines in the loop to filter out spurious biases. Next, I will introduce PIGLeT: a model that learns physical commonsense understanding by interacting with the world through simulation, using this knowledge to ground language. From an English-language description of an event, PIGLeT can anticipate how the world state might change – outperforming text-only models that are orders of magnitude larger. Finally, I will introduce MERLOT, which learns about situations in the world by watching millions of YouTube videos with transcribed speech. Through training objectives inspired by the developmental psychology idea of multimodal reentry, MERLOT learns to fuse language, vision, and sound together into powerful representations.
Together, these directions suggest a path forward for building machines that learn language rooted in the world.

BIOGRAPHY

Rowan Zellers is a final year PhD candidate at the University of Washington in Computer Science & Engineering. Click here for more information.

HOW TO REACH US
Contactus@cs.jhu.edu
410-516-8775
cs.jhu.edu

Johns Hopkins University
Department of Computer Science
3400 N. Charles St | Malone 160
Baltimore, MD 21218