UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

On Reality and the Limits of Language Data: Aligning LLMs with Human Norms

Permalink

https://escholarship.org/uc/item/80p9t8tf

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors

Collier, Nigel
Liu, Fangyu
Shareghi, Ehsan

Publication Date

2023

Peer reviewed

On Reality and the Limits of Language Data: Aligning LLMs with Human Norms

Nigel Collier
University of Cambridge, Cambridge, United Kingdom

Fangyu Liu
University of Cambridge, Cambridge, United Kingdom

Ehsan Shareghi
Monash University, Melbourne, Victoria, Australia

Abstract

Recent advancements in Large Language Models (LLMs) harness linguistic associations in vast natural language data for practical applications. However, their ability to understand the physical world using only language data remains a question. After reviewing existing protocols, we explore this question using a novel and tightly controlled reasoning test (ART) and compare human norms against versions of GPT-3. Our findings highlight the categories of common-sense relations models that could learn directly from data and areas of weakness. GPT-3 offers evidence for verbal reasoning on a par with human subjects for several relations including Synonymy, Antonymy, and Default inheritance, Without reinforcement learning from human judgements, it appears GPT-3 performs at the lower end of the reference interval for Has-part and Contained-in. Weaknesses were observed also in affordance characteristics through Necessary-quality, Order-of-size and Order-of-intensity. Combining LLMs with symbolic world grounding is a promising direction to address associative learning.