UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Different Trajectories through Option Space in Humans and LLMs

Permalink

https://escholarship.org/uc/item/5wj386xm

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Dracheva, Alina

Phillips, Jonathan

Publication Date

2024

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Different Trajectories through Option Space in Humans and LLMs

Alina Dracheva
Dartmouth College, Hanover, New Hampshire, United States

Jonathan Phillips
Dartmouth College, Hanover, New Hampshire, United States

Abstract

Real-world decision-making requires the generation of possible options. Humans are exceptionally good at navigating such potentially unbounded spaces: they typically generate their best options first and most idiosyncratic last (Srinivasan, Acierno, & Phillips, 2022). Recently, large language models (LLMs) have shown impressive communication and reasoning abilities, suggesting that they may now be mirroring some of the conceptual structures used by humans. Here, we explore if LLMs navigate option spaces similarly to humans. We compared series of human-generated options to those from an LLM using the semantic similarity of generated options across various open-ended contexts. While LLMs display some global patterns similar to humans, their option sequences follow different trajectories within the semantic space. Specifically, GPT-3 frequently revisits previous semantic clusters, whereas humans progress more linearly. Additionally, compared to humans, GPT-4 typically shows fewer revisits and shorter stays in a given semantic cluster, suggesting a more transient trajectory across the semantic landscape.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).