

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Cognitive Models as Simulators: The Case of Moral Decision-Making

Permalink

<https://escholarship.org/uc/item/1t00d92r>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Nobandegani, Ardavan S.

Shultz, Thomas

Rish, Irina

Publication Date

2022

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

Cognitive Models as Simulators: The Case of Moral Decision-Making

Ardavan S. Nobandegani

McGill University, Montreal, Quebec, Canada

Thomas Shultz

McGill University, Montreal, Quebec, Canada

Irina Rish

University of Montreal, Montreal, Quebec, Canada

Abstract

To achieve desirable performance, current AI systems often require huge amounts of training data. This is especially problematic in domains where collecting data is both expensive and time-consuming, e.g., where AI systems require having numerous interactions with humans, collecting feedback from them. In this work, we substantiate the idea of cognitive models as simulators, which is to have AI systems interact with, and collect feedback from, cognitive models instead of humans, thereby making their training process both less costly and faster. Here, we leverage this idea in the context of moral decision-making, by having reinforcement learning (RL) agents learn about fairness through interacting with a cognitive model of the Ultimatum Game (UG; Nobandegani, Destais, & Shultz, 2020), a canonical task in behavioral and brain sciences for studying fairness. Interestingly, these RL agents learn to rationally adapt their behavior depending on the emotional state of their simulated UG responder.