

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Thinking into a machine's mind - Anticipation of an agent's behavior in a cooperative game

Permalink

<https://escholarship.org/uc/item/7p64p80k>

Journal

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

Authors

Becker, Fritz

Buder, Juergen

Huff, Markus

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

Thinking into a machine's mind - Anticipation of an agent's behavior in a cooperative game

Fritz Becker

Leibniz-Institut für Wissensmedien, Tübingen, Germany

Juergen Buder

Leibniz-Institut für Wissensmedien, Tübingen, Germany

Markus Huff

Leibniz-Institut für Wissensmedien, Tübingen, Germany

Abstract

An accurate mental model of the partner's behavior is fundamental for efficient cooperation. The theory of mind demonstrates that humans are able to create such a model from repeated interactions with their human partners. However, it is an open question whether humans are also willing and capable of taking the perspective of artificial agents and creating similar mental models of agent behavior. We developed a repeated cooperative task that allows us to investigate the process that guides the formation of a specific partner model by repeatedly asking for a prediction of an artificial agent's actions. We found that humans learn to anticipate the artificial partners' behavior if it is goal-directed. An inability to explicitly explain the partner's behavior suggests that this is an implicit learning process. The role of the acquisition of task knowledge in the model of the other agent's behavior is discussed.