

# When Empowerment Disempowers in Multi-Agent Assistance

**Claire Yang**

University of Washington, Seattle, Washington, United States

**Maya Cakmak**

University of Washington, Seattle, Washington, United States

**Max Kleiman-Weiner**

University of Washington, Seattle, Washington, United States

## Abstract

People can assist others even without knowledge of their specific goals. However, this ability remains a challenge for artificial intelligence, limiting the development of assistive technologies. This issue is particularly important in caregiving, where rather than focusing on helping with a particular task, one may aim to improve the broad autonomy of the care recipient. One promising approach to care that does not require goal inference is to maximize the empowerment of another. Empowerment is a computational measure of one's ability to control their environment. Prior work on empowerment assumes one-to-one interactions, overlooking the fact that assistive technologies are often deployed in multi-agent environments that may include caregivers or family members. Here, we show that optimizing for one person's empowerment may inadvertently disempower others. Finally, we develop and test multi-agent extensions of empowerment that enable an assistant to empower one person without harming another.