

Wolfram Barfuss

Collective Reinforcement Learning Dynamics for Sustainability Economics

Abstract: Cooperation at scale is critical for achieving a sustainable future for humanity. However, achieving collective, cooperative behavior—in which intelligent actors in complex environments jointly improve their well-being—remains poorly understood. Complex systems science (CSS) provides a rich understanding of collective phenomena, the evolution of cooperation, and the institutions that can sustain both. Yet, much of the theory in this area fails to fully consider individual-level complexity and environmental context—mainly for the sake of tractability and because it has not been clear how to do so rigorously. These elements are well captured in multiagent reinforcement learning (MARL), which has recently focused on cooperative (artificial) intelligence. However, typical MARL simulations can be computationally expensive and challenging to interpret. In this presentation, I propose that bridging CSS and MARL offers new directions. By investigating the non-linear dynamics of collective reinforcement learning, we can better understand how, in complex environments, individual decisions become collective action for a sustainable future.