UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Social Norm Formation Dynamics with IBL Agents: Short-/Long-Term Rewards and Network Structure

Permalink

https://escholarship.org/uc/item/8n26x1jt

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Tajima, Jun

Morita, Junya

Publication Date

2026

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

Social Norm Formation Dynamics with IBL Agents: Short-/Long-Term Rewards and Network Structure

Jun Tajima
Shizuoka University, Japan

Junya Morita
Shizuoka University, Japan

Abstract

Social normative decision-making involves two evaluative axes: immediate gains from aligning with others (reputation, conformity, reduced friction) and delayed collective consequences accumulating through repeated actions (social loss, victimization). This study examines, via a multi-agent simulation with Instance-Based Learning Theory (IBLT) agents, how this tension shapes norm formation, bifurcation, and stabilization. Agents repeatedly choose between two actions (pull/keep) inspired by the trolley problem. In each round, they receive a short-term reward proportional to the degree of agreement with neighbors, while at fixed block intervals they receive a delayed penalty depending on the total number of victims. We compare dynamics on a lattice Grid with dynamics on networks generated by the Watts–Strogatz model and classify trajectories into three types (pull-dominant, intermediate, keep-dominant). As a result, the prevalence of these types differs by network structure, suggesting that network differences affect how the two rewards interact and thereby change transition dynamics.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society* ©2026 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).