Inadequate Imitation: What can we Learn from Laboratory Experiments on Information Aggregation over Networks?

Dotan Persitz

Networks are central to information diffusion in social, economic, and organizational contexts. How do individuals use their social networks to gather information? The first part of this study reviews theoretical and experimental research on the topic. Theoretical work, focused on very large networks, generally concludes that information aggregation is rarely problematic. Experimental studies, by contrast, typically examine small networks (3–7 players) and consistently find both sophisticated behavior (using all available structural information) and naïve behavior (relying only on immediate surroundings). The second part reports a laboratory experiment in a medium-sized network (18 players). The results are inconsistent with both sophisticated and naïve models, suggesting instead the need to analyze structural failures and behavioral patterns. Structural failures are network features that hinder efficient information gathering even among rational players. Behavioral patterns include ″underreaction″ (overweighting initial private signals) and ″under-imitation″ (insufficiently following better-informed peers). At both micro and macro levels, the findings indicate that restricting information from central players may improve overall information aggregation.

Liars Need a Good Memory: Human Oversight of AI with Reciprocal Learning

Yahav Inbal Dov Te'Eni

Growing concerns about the necessity of human oversight in critical AI systems demand urgent attention. Recent advances in generative AI have revealed unethical behaviors with potentially severe human consequences. Yet existing control mechanisms are insufficient to ensure meaningful human supervision capable of addressing these challenges. In response, we propose a dual-mode framework for operating AI systems: a stable mode, designed for effective utilization of big data, and an adaptive mode, aimed at fostering continual human learning. The adaptive mode relies on reciprocal human–machine learning, supported by tailored feedback loops that align system performance with broader objectives while safeguarding ethical values. Paired with appropriate regulation, this approach can achieve both high and unbiased performance while prioritizing human well-being. We argue that continual human learning, developed alongside machine learning, is essential for enabling effective human oversight in critical AI applications.