Designing Incentives for Networked Multi-agent Systems

Xinwei Song

[1]Key Laboratory of Intelligent Perception and Human-Machine Collaboration, ShanghaiTech University
[2]State Key Laboratory of General Artificial Intelligence, BIGAI
songxw2024@shanghaitech.edu.cn

Abstract

Achieving globally desirable outcomes in networked multi-agent systems—such as high social welfare, stable allocations, and widespread cooperation—is a fundamental challenge in AI. This paper outlines a research agenda that explores two complementary pathways to this goal. The first is a *top-down* approach, where a central mechanism designer proposes rules to guide strategic agents towards theoretically optimal equilibria. The second is a *bottom-up* approach, where desirable farsighted policies, like cooperation in social dilemmas, emerge from the decentralized interactions of agents via multi-agent reinforcement learning. We argue that the integration of these paths constitutes a promising frontier for creating robust and adaptive multi-agent systems.

Problem Formulation

Multi-agent systems (MAS), where multiple autonomous agents observe, interact, and learn policies through experiences, provide a foundational framework for modeling complex real-world decision-making scenarios (Maldonado et al. 2024). The promise of MAS lies in its ability to leverage collective intelligence, accomplishing tasks that are beyond the capabilities of any single agent (Oroojlooy and Hajinezhad 2023). However, self-interested agents may take myopic actions, especially when the social network structure restricts local observation and information dissemination, which frequently leads to suboptimal and inefficient outcomes at the system level (Fatima, Jennings, and Wooldridge 2024). For example, in unmanned aerial vehicle (UAV) formation tasks, multiple UAVs minimizing their own energy use may target the same position, thereby compromising the group's objective of covering all locations efficiently. Thus, we ask: **How can we incentivize networked multi-agent systems with strategic, self-interested agents to reach a collectively desirable outcome?**

Existing research has approached this problem from two perspectives. The first, rooted in *economics and game theory*, employs *mechanism design* to architect the rules of interaction. In the UAV formation example, a position auction can promote the overal social welfare and facilitate better coordination. However, these mechanisms may struggle in complex, dynamic environments with externalities or behavioral uncertainties. The second route, powered by *Multi-agent Reinforcement Learning (MARL)*, seeks *bottom-up emergence* of desirable outcomes through repeated interactions. For example, reward-shaping techniques could enable UAVs to learn collaborative behaviors via MARL. The limitation is that hand-crafted rewards lack theoretical assurances about their convergence properties and the adaptation capability to unseen scenarios.

We propose to address these limitations through a progressive research agenda. First, we argue that mechanism design can embrace more dynamic and complex models that better capture real-world scenarios. Second, for MARL methods, their reliance on reward shaping and prior knowledge can be mitigated by constructing algorithms and modules that reflect the principle of effective incentives, e.g. reputation system. Furthermore, we advocate for *integrating these two approaches* into a unified technical framework. This hybrid paradigm combines the theoretical assurances of mechanism design with the automated solution discovery and adaptation capability of self-evolving MARL agents, creating a powerful methodology to incentivize networked multi-agent systems toward desirable equilibria.

The Dual-Path Research Framework

Path 1: Centralized Mechanism Design

Our first line of work addresses the problem through the lens of *mechanism design*. We focus on the matching problem where each agent has an initial endowment, and they swap with each other to get a more preferred one (Shapley and Scarf 1974). On top of this, we model **matching on networks**, a realistic model where the matching starts with a small set of initial participants and existing agents can invite their social ties to join in to expand the market (Zhao 2022). Since agents' social relationships and preferences are private information, the central problem is to construct a stable and optimal matching that ensures truthful revelation. Compared to classical models, this networked setting introduces a larger strategic space for agents, who can now control the market size. This challenges the classic solutions and necessitates new mechanisms as well as a redefinition of the theoretical boundaries of what is achievable.

Our Completed Work: We proved that in matching on networks, the classic properties, stability and

Pareto optimality, are incompatible with truthful revelation. We thus proposed weakened but the tightest achievable notions—*stable-cc* and *optimal-cc*—and designed the *Connected Trading Cycles (CTC)* mechanism, which is the first to reach the theoretical boundaries for the new model (Song, Yang, and Zhao 2025).

Future Direction: A compelling next step is to tackle more realistic matching problems featuring externalities (e.g., agents' utility depends on their neighbor's allocation), irrational agents, or dynamic preferences. We plan to fuse theoretical analysis, simulation, and MARL to automate the discovery of new solutions. For instance, a high-level mechanism designer agent and several low-level strategic player agents may co-evolve through adversarial training to output a novel matching mechanism with desirable properties. This co-evolutionary paradigm will also provide *insights* that explain real-world phenomena in matching markets.

Path 2: Decentralized Emergence via MARL

The top-down approach presupposes a central authority that can enforce a mechanism. However, this is infeasible in many decentralized systems. Our second line of work targets dilemmas in decentralized multi-agent systems where myopic policies result in low collective benefits. Inspired by the reputation system that promotes human cooperation, we study how to design reputation-aware agents and guide them to learn a more farsighted decision-making policy that leads to collective benefit, where global cooperation *emerges* from local interactions in a *bottom-up* approach.

Our Completed Work: We proposed *COOperate with Emergent Reputation (COOPER)*, a MARL method where agents jointly learn a reputation-based policy and a reputation assessment module (including gossip-based and interaction-based assessment) to foster cooperation in mixed-motive social dilemmas (Song et al. 2025). Compared to existing methods that rely on predefined reputation semantics or intrinsic reward design, our approach distinguishes itself through its sophisticated module and information flow design, which enables the simultaneous emergence of reputation norms and cooperative policies purely through environmental feedback (Smit and Santos 2024; Ren et al. 2025). This endows COOPER with strong adaptability to diverse co-players. Through extensive experiments in matrix games and grid-world environments, we demonstrate COOPER's effectiveness in achieving sustained cooperation across various network structures, its robustness in self-play scenarios, and its adaptation capabilities when interacting with agents with existing reputation norms.

Future Direction: The current model has several avenues for extension. In COOPER, the gossip process assumes truthful communication and a static social network. However, strategic spread of misinformation and the co-evolution of **reputation and social network structure** are exciting open questions. In addition, reputation is currently modeled as a scalar, but it could be a multi-dimensional vector that agents dynamically select, which adds another dimension to agents' strategy space. This refined model can serve as a powerful tool to study the rich dynamics of **reputation-driven action and social network evolution**.

Research Agenda and Timeline

The core of our research agenda is to create a **unified framework that marries the rigor of theory with the adaptability of learning**. This unification targets two key dimensions: first, extending theoretical models to incorporate real-world complexities that are typically abstracted away; second, developing learning methods that can provably converge to equilibria with desirable theoretical properties. In other words, we aim to construct a new paradigm where **formal models inform learning objectives, and learned behaviors inspire new theoretical developments**.

We have formulated a three-phase plan to realize this vision. **Phase 1 (Now - Jan 2026)** will focus on the two future directions outlined above: applying MARL for matching mechanisms to automate mechanism discovery in complex environments, and extending COOPER to handle strategic communication and multi-dimensional reputation. **Phase 2 (Feb - Oct 2026)** will build upon these advances to construct unified methodologies that integrate mechanism design with MARL, developing and validating hybrid approaches that combine theoretical guarantees with adaptive learning capabilities. **Phase 3 (Nov 2026 - Jun 2027)** will consolidate the theoretical and empirical findings from Phases 1–2 and identify limitations and open questions.

References

Fatima, S.; Jennings, N. R.; and Wooldridge, M. 2024. Learning to resolve social dilemmas: a survey. *Journal of Artificial Intelligence Research*, 79: 895–969.

Maldonado, D.; Cruz, E.; Torres, J. A.; Cruz, P. J.; and Benitez, S. d. P. G. 2024. Multi-agent systems: A survey about its components, framework and workflow. *IEEE Access*, 12: 80950–80975.

Oroojlooy, A.; and Hajinezhad, D. 2023. A review of cooperative multi-agent deep reinforcement learning. *Applied Intelligence*, 53(11): 13677–13722.

Ren, T.; Yao, X.; Li, Y.; and Zeng, X.-J. 2025. Bottom-Up Reputation Promotes Cooperation with Multi-Agent Reinforcement Learning. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems*, 1745–1754.

Shapley, L.; and Scarf, H. 1974. On cores and indivisibility. *Journal of mathematical economics*, 1(1): 23–37.

Smit, M.; and Santos, F. P. 2024. Learning fair cooperation in mixed-motive games with indirect reciprocity. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, 220–228.

Song, X.; Huang, Y.; Kang, Y.; Zhao, D.; and Feng, X. 2025. Learning to Cooperate with Emergent Reputation via Multi-agent Reinforcement Learning. Under Review of ICLR 2026.

Song, X.; Yang, T.; and Zhao, D. 2025. Housing Market on Networks. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems*, 1959–1967.

Zhao, D. 2022. Mechanism Design Powered by Social Interactions: A Call to Arms. In *IJCAI*, 5831–5835.