APCNCS2026: Asia-Pacific Conference on Networks and Complex Systems
Focus Session: Complexity of AI
10-12 June 2026, Nanyang Technological University, Singapore

Reinforcement learning in population games: from microscopic chaos to macroscopic equilibrium

Thiparat Chotibut[1]

[1]Chula Intelligent and Complex Systems, Department of Physics, Chulalongkorn University, Bangkok, Thailand

Keywords: Multi-agent Reinforcement Learning, Algorithmic Game Theory, Heterogeneity, Period-doubling Route to Chaos, Ergodic Convergence

Extended Abstract

Motivation. Multiagent reinforcement learning (MARL) is a ubiquitous framework in artificial intelligence that, when applied to model how populations adapt in dynamic environments, can describe the emergent dynamics of collective decision-making in game theory. In congestion games, which serve as a paradigm of strategic stability because incentives are perfectly aligned through a common potential function, it is proven that each agent's adaptation and learning dynamically stabilize at a unique Nash equilibrium. However, this relies on the assumption that each agent must use vanishingly small learning rates (becoming less adaptive to input signals over time). In contrast, real-world populations of agents—such as humans—frequently utilize large learning rates, are highly adaptive to cost signals, exhibit memory loss (recency bias), and have heterogeneous decision-making preferences. A natural question emerges: *How does multiagent learning, where each agent can adapt at different rates and be subject to fast adaptation, lead to collective dynamics in congestion games?*

Approach and Methodology. To rigorously analyze how agent behavior departs from static equilibrium, we study discrete-time dynamics arising from learning in two-strategy congestion games by systematically introducing increasing levels of behavioral complexity. First, we establish a baseline by modeling a homogeneous population of agents applying the standard Multiplicative Weights Update (MWU) with full historical memory [3]. Second, we introduce a memory loss parameter using a simplified Experience-Weighted Attraction (sEWA) algorithm [1]. Finally, we tackle full population heterogeneity, allowing infinitely many agent types to possess diverse initial beliefs and learning rates [2]. Despite this complexity, the dynamics induced by these settings reduce to an iterated one-dimensional map, facilitating exact analytical characterizations of the collective macroscopic behavior.

Results. Our analysis reveals a progression of dynamical behaviors based on the complexity of the multiagent learning models. *Baseline:* In homogeneous populations with perfect memory [3], an increase in the adaptation rate (learning rate) destabilizes the global Nash equilibrium (Figure 1). As the learning rate increases, the system undergoes a period-doubling route to Li-Yorke chaos, characterized by severe degradations in social welfare. We also report the universal Feigenbaum constant characterizing the route to chaos despite

the one-dimensional map being non-unimodal. *Memory Loss Leads to Predictability:* Introducing memory loss (recency bias or cognitive discounting) alters the chaotic regime [1]. While high learning rates still prevent equilibration at the static Nash equilibrium, substantial memory loss suppresses chaos. Instead of bifurcating into a chaotic attractor, the system dynamics reduce to a predictable, globally attracting period-2 limit cycle. *Macroscopic Order amidst Microscopic Chaos:* When full population heterogeneity is introduced [2], fast overall learning rates will inevitably drive the population dynamics toward chaos. Remarkably, a macroscopic regularity emerges: *the time-averaged population strategy robustly converges to the exact static Nash equilibrium, linking macroscopic ergodic convergence to the Nash equilibrium amidst microscopic chaos.*

Conclusions. Our findings challenge the paradigm of standard game-theoretic equilibrium predictability. Fast adaptation or high learning rates inevitably drive congestion games into chaos, while cognitive forgetting organically suppresses collective dynamical complexity in these games. Importantly, although the population may not settle on a Nash equilibrium and can be fully chaotic, the macroscopic time-average converges to the Nash equilibrium, exactly as standard game theory predicts. These results also formalize the numerical findings of the classic El Farol bar problem.

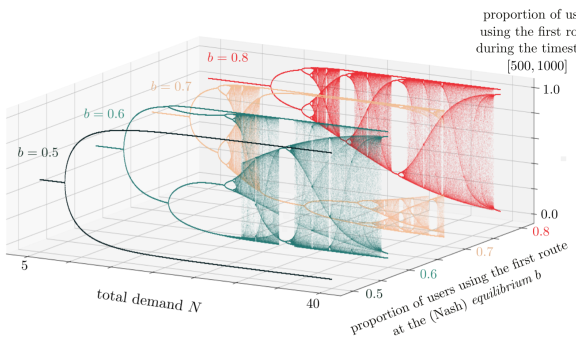

Figure 1: **Route to Chaos in Congestion Games for Homogeneous Populations.** The total demand N is proportional to the learning (or adaptation) rate. As N increases, large learning rates destabilize the static Nash equilibrium, driving asymmetric games ($b \neq 0.5$) through the period-doubling route to chaos.

References

[1] Jakub Bielawski, Thiparat Chotibut, Fryderyk Falniowski, Michał Misiurewicz, and Georgios Piliouras. Memory loss can prevent chaos in games dynamics. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 34(1):013146, 2024.

[2] Jakub Bielawski, Thiparat Chotibut, Fryderyk Falniowski, Michał Misiurewicz, and Georgios Piliouras. Heterogeneity, reinforcement learning, and chaos in population games. *Proceedings of the National Academy of Sciences*, 122(25):e2319929121, 2025.

[3] Thiparat Chotibut, Fryderyk Falniowski, Michał Misiurewicz, and Georgios Piliouras. The route to chaos in routing games: When is price of anarchy too optimistic? In *Advances in Neural Information Processing Systems*, volume 33, pages 766–777, 2020.