Conjectural Games: A Framework for Incentive Design and Opponent Modelling in Multi-Agent Systems

Francesco Morri – Inria-Lille – francesco.morri@inria.fr – francescomorri.github.io

Motivation Classical game theory relies on equilibrium concepts where players assume fixed rationality and stable environments. In practice, multi-agent interactions are rarely stationary: agents learn, adapt, and continuously update their strategies. Standard models often fail to capture how players perceive or conjecture about the behaviour of others, which limits their descriptive and prescriptive power in dynamic and uncertain environments.

Research Objective This project proposes conjectural games as a principled framework to model how agents form and act upon beliefs about others' responses. By explicitly incorporating conjectures into the strategy space, I aim to advance both the theory of games and the design of adaptive learning algorithms. The overarching goal is to show how conjectures can serve as a new and potentially more powerful lever for incentive design, while also addressing non-stationarity and opponent modelling multi-agent reinforcement learning (MARL).

Methodology The theoretical part will extend conjectural equilibrium concepts to richer forms of beliefs and games, introducing novel equilibria concept and identifying existence, stability and convergence conditions. On the algorithmic side, I will design learning rules that operationalize conjectures, enabling agents to anticipate, interpret, and adapt to opponents' evolving strategies. Analytical work will focus on designing such beliefs and non-stationary environments. Empirical validation will be carried out in MARL benchmarks where opponent modelling and adaptability are critical (cooperative environments, social dilemmas, etc.).

Applications and Impact For game theory, this research offers a novel perspective on incentives: rather than shaping payoffs directly, one can design and guide the conjectures that agents form about others. For MARL, conjectural reasoning provides a natural mechanism to tackle opponent modelling and non-stationarity, two central barriers to scalability and robustness. Potential applications include markets, classical social dilemmas (such as the tragedy of the commons) and, more broadly, systems where coordination is fundamental to reach efficient solutions (local energy communities, traffic networks, etc.).

Expected Outcomes

(1) A generalized theoretical framework for conjectural equilibria

(2) Algorithms for conjecture-based learning in MARL

(3) Empirical evidence of improved adaptation in non-stationary multi-agent environments