A Novel Reinforcement Learning Approach in Detecting Non-Collaborators for Trust Game with a Large Number of Participants

Ziyu Huang

Duke Kunshan University, Kunshan, Jiangsu 215316, China
zh147@duke.edu

Abstract. This paper explores the integration of Generative AI (GAI) and multi-agent reinforcement learning (MARL) to advance traditional game theory research, particularly in simulating complex human-like decision-making processes. By examining the "good and bad collaborators" game, the study aims to model behaviors in trust and reputation scenarios using GAI to generate profiles for simulation. The innovative use of GAI facilitates the creation of dynamic, realistic profiles, which are then tested in MARL environments to analyze behavior patterns and improve detection methods for uncooperative participants. This approach not only enhances the understanding of strategic interactions in controlled settings but also addresses the limitations of traditional game theory in handling dynamic and incomplete information scenarios. The research indicates potential for developing more sophisticated game theory models that can adapt to and accurately reflect evolving gameplay and strategies, thus providing deeper insights into the foundational principles of human strategic behavior and decision-making.

Keywords: computational economics · game theory · innovative education · generative AI · multi-agent reinforcement learning

1 Introduction

Traditional game theory is principally concerned with the negotiations between people. Within the context of research, it faces challenges due to a lack of data and the difficulty of conducting experiments. Recently, generative AI (GAI) has illuminated the path to generating human-like decision-making processes [1]. This could potentially provide convenience for experimental endeavors in traditional game theory.

Consider the example of the "good and bad collaborators" game, which is a variation of trust and reputation games [2]; without controls in place, the good collaborators tend to eventually exit the game. Currently, a prospective avenue for enhancement is to construct a system of mutual selection that fosters equitable cooperation. Nevertheless, an inherent concern arises from our inability to curb the production of fraudulent resumes—a problem traditionally mitigated by more stringent regulatory measures.

To address this issue, I intend to incorporate GAI to model good and bad collaborators in different scenarios and utilize MARL to learn these results. As shown in Figure 1, we may first simulate various scenarios (like personal performance and skills), followed by employing Generative AI to create profiles for good and bad collaborators, all striving to succeed in a "good collaborator test." We can then appoint multiple agents to repeatedly execute the "good collaborator detection" tasks using these profiles and relay the outcomes back to the Generative AI for its learning process. Ultimately, we aspire to decipher the conduct paradigms of uncooperative collaborators, which will assist in advancing game theory research in such contexts.

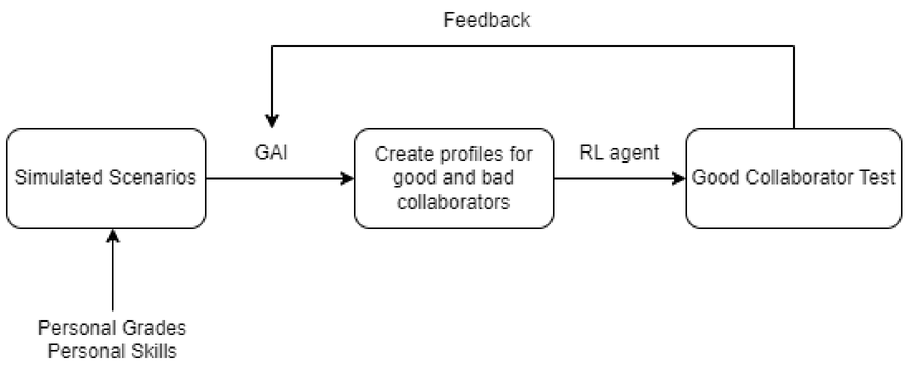

Fig. 1. Flowchart of how GAI working in the collaborator game

2 Background

As shown in Figure 2, in the conventional frameworks of game theory, models usually encompass the gaming environment and the human or AI agents on two distinct levels. The present-day constraints revolve around the often too static gaming circumstances that resist dynamic fine-tuning. By integrating Generative AI, we gain the capability to craft intricate and realistic training datasets, particularly advantageous in scenarios with a deficiency or inaccessibility of real-world data. Additionally, Generative AI can be utilized for the real-time generation and modification of strategies to navigate the unpredictable shifts within the game setting. This methodology equips AI proxies with enhanced agility in their responses to uncertain and continuously evolving adversary tactics.

3 Research Question

Our research aims to address several crucial questions: How can natural language processing combined with reinforcement learning effectively predict and classify

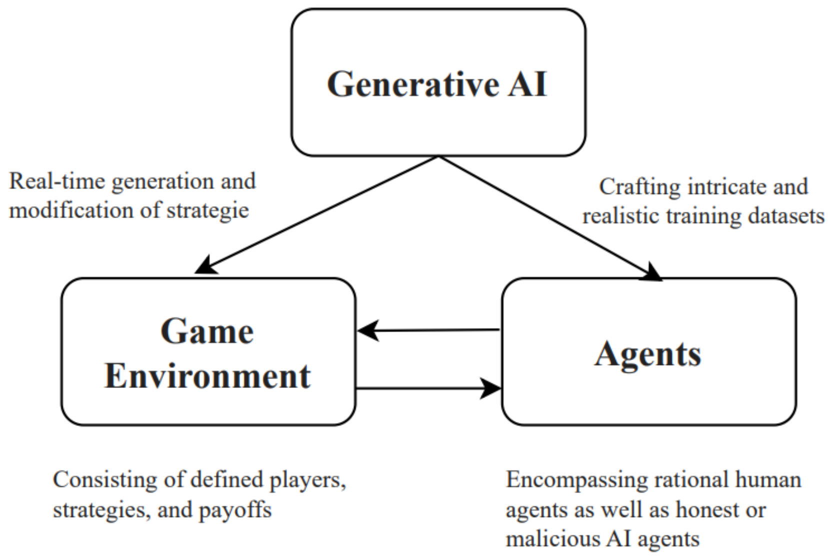

Fig. 2. Flowchart of how GAI cooperate in current game theory structure

cooperative behaviors in generated resumes? Can a model that analyzes historical interaction patterns accurately predict the transition from cooperative to non-cooperative behaviors over time? And what are the key indicators in textual and behavioral data that distinguish good collaborators from bad collaborators, particularly those who initially disguise their intentions? These questions are significant because accurately predicting and understanding cooperative behaviors can lead to more effective team formations, enhance AI-human interactions, and ensure safety in trust-critical contexts. Traditional game theory literature, which often focuses on static scenarios with fixed strategies and lacks empirical data, does not sufficiently address these dynamic and complex behavior patterns, nor does it incorporate advanced AI technologies like natural language processing and machine learning. Our research seeks to fill these gaps by using modern AI techniques to provide a dynamic, context-sensitive analysis of cooperative behaviors, potentially enhancing both theoretical and practical applications in game theory and related fields.

4 Methodology

4.1 Simulating Participants with Different Cooperation Patterns

In the overall architecture, I aim to simulate various forms of non-cooperative participants, including those who are completely uncooperative, occasionally cooperative, and gradual defectors [3]. Completely uncooperative actors choose not to cooperate in all circumstances, regardless of the behavior of other participants. These participants typically pursue short-term interests without considering the damage to long-term relationships or collective welfare. In the prisoner's dilemma game, these participants always choose to defect. We can simulate this behavior by setting fixed strategy parameters, such as a cooperation probability of zero. Occasionally cooperative participants (Conditional Cooperators) base their cooperative behavior on specific conditions, such as the prior actions of others, environmental incentives, or positive results from historical interactions. They may show cooperation when a cooperative environment is detected but quickly switch to non-cooperation if they feel exploited or if the other party does not cooperate. We can use a Tit-for-Tat strategy, which starts with cooperation and continues if the other party cooperates, but defects in the next round if the other party defects. Alternatively, we can adjust the cooperation probability based on the ratio of cooperation in historical interactions. Gradual Defectors show cooperative behavior in the initial stages of interaction, but as time passes and conditions change, they gradually shift to a non-cooperative strategy. This behavior pattern is usually strategic, aimed at maximizing personal benefits by leveraging the advantages of cooperation after trust is established, then opting out at an opportune moment to avoid long-term costs or severing ties after maximizing benefits. We can set the cooperation probability with a function based on time or the number of interactions, where the cooperation probability gradually decreases as time progresses or as the number of interactions increases.

4.2 Simulating Human Cooperation Behavior with ChatGPT

As previously mentioned, we plan to use ChatGPT, a large language model equipped with a rich implicit knowledge base, to simulate human-like decision-making processes. Specifically, we will construct two background databases: one for "Good Collaborators" and another for "Bad Collaborators." These databases will contain detailed background information and behavioral characteristics of each type of collaborator. In our experiments, the ChatGPT agent will randomly sample from these databases to generate representative resumes of both good and bad collaborators. This process aims to explore and identify key decision-making factors and behavior patterns in cooperation by simulating different cooperative behaviors.

4.3 Implementing Dual Reinforcement Learning Networks for Simulating and Distinguishing Collaborative Behaviors

To deeply simulate and distinguish between different types of collaborators, our model incorporates two functional reinforcement learning networks. These networks are tasked with learning and processing distinct types of data to effectively identify and categorize "good collaborators" and "bad collaborators." The first type is a language information learning network, which focuses on processing and analyzing resume texts generated by ChatGPT. This network employs natural language processing techniques to recognize key words and phrases in resumes that may indicate the candidate's cooperative tendencies and behavior patterns. For example, by analyzing achievements and past behavior descriptions in the resume, the network can preliminarily classify candidates as "good collaborators" or "bad collaborators." We plan to use Recurrent Neural Networks (RNN) or Long Short-Term Memory networks (LSTM) to handle sequence data, capturing long-distance dependencies in language. The second type is a historical information learning network, considering that some "bad collaborators" may initially exhibit cooperative behaviors and then switch to non-cooperation. This network focuses on analyzing participants' behavior history. By learning the interaction patterns and decision changes of participants, the network can identify individuals who may initially hide their true intentions. As in Figure 3, we will utilize reinforcement learning networks such as Q-learning or Deep Q Networks (DQN) to assess and optimize decision-making strategies based on historical behavior, implementing mechanisms to recognize transition patterns from cooperation to non-cooperation, such as the traits of gradual defectors.

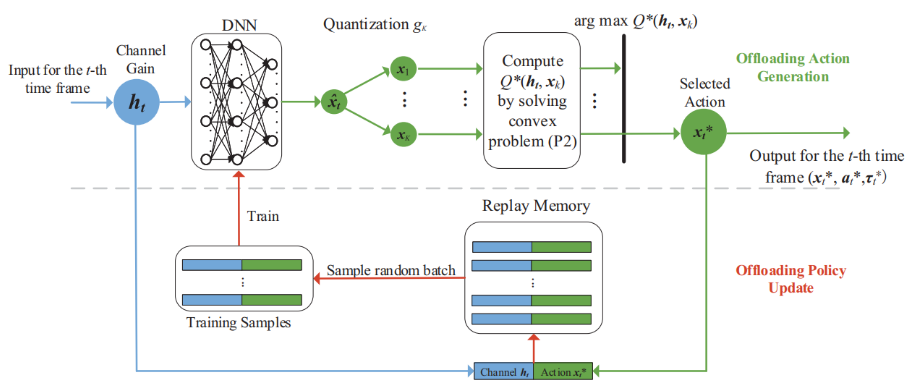

Fig. 3. Sample architecture for trust game reinforcement learning [4]

5 Application Scenario

Our newly proposed solution, leveraging natural language processing and reinforcement learning to predict and classify cooperative behaviors, is applicable in various real-world scenarios, such as corporate human resources for analyzing potential hires' resumes [5], social robots and virtual assistants for improving interactions with humans [6], and autonomous vehicles to predict human drivers' behaviors for safer and more efficient decision-making [7]. To ground our approach in behavioral science, we can draw from psychological theories such as Social Exchange Theory [8], which analyzes relationships through cost-benefit analysis; Theory of Planned Behavior, which links intentions to attitudes, norms, and perceived control; Cognitive Dissonance Theory, explaining behavior contrary to beliefs due to social pressures; and research on trust and reciprocity, crucial for understanding the dynamics of cooperation. Integrating these psychological insights with our AI-driven methods will enhance our model's predictive accuracy and deepen the interdisciplinary connection between AI, game theory, and psychology, thus providing a robust foundation for applying game theoretical concepts in practical, impactful ways.

Bibliography

[1] E. Shapira, O. Madmon, R. Reichart, and M. Tennenholtz, "Can large language models replace economic choice prediction labs?" *arXiv preprint arXiv:2401.17435*, 2024. [Online]. Available: https://arxiv.org/abs/2401.17435

[2] J. Berg, J. Dickhaut, and K. McCabe, "Trust, reciprocity, and social history," *Games and economic behavior*, vol. 10, no. 1, pp. 122–142, 1995. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0899825685710275

[3] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *science*, vol. 211, no. 4489, pp. 1390–1396, 1981. [Online]. Available: https://www.journals.uchicago.edu/doi/abs/10.1086/383541

[4] L. Huang, S. Bi, and Y.-J. A. Zhang, "Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks," *IEEE Transactions on Mobile Computing*, vol. 19, no. 11, pp. 2581–2593, 2019. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8771176/

[5] P. Cappelli and J. Keller, "Talent management: Conceptual approaches and practical challenges," *Annu. Rev. Organ. Psychol. Organ. Behav.*, vol. 1, no. 1, pp. 305–331, 2014. [Online]. Available: https://www.annualreviews.org/doi/abs/10.1146/annurev-orgpsych-031413-091314

[6] C. Breazeal, *Designing sociable robots.* MIT press, 2004. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=402dquhxSTQC&oi=fnd&pg=PR15&dq=Designing+sociable+robots&ots=oCRmwZy6xh&sig=X0ojg3nchPo7iYT2Tr-xdWzLXNM

[7] M. Maurer, J. C. Gerdes, B. Lenz, and H. Winner, *Autonomous driving: technical, legal and social aspects.* Springer Nature, 2016. [Online]. Available: https://library.oapen.org/bitstream/handle/20.500.12657/27811/1/1002194.pdf

[8] G. C. Homans, "Social behavior as exchange," *American journal of sociology*, vol. 63, no. 6, pp. 597–606, 1958. [Online]. Available: https://www.journals.uchicago.edu/doi/abs/10.1086/222355