Approximating Human Strategic Reasoning with LLM-Enhanced Recursive Reasoners Leveraging Multi-agent Hypergames

Vince Trencsenyi[1], Agnieszka Menselt[1], and Kostas Stathis[1]

Royal Holloway University of London, Egham, Surrey, UK
{vince.trencsenyi, agnieszka.mensfelt, kostas.stathis}@rhul.ac.uk

Abstract. LLM-driven multi-agent-based simulations have been gaining traction with applications in game-theoretic and social simulations. While most implementations seek to exploit or evaluate LLM-agentic reasoning, they often do so with a weak notion of agency and simplified architectures. We implement a role-based multi-agent strategic interaction framework tailored to sophisticated recursive reasoners, providing the means for systematic in-depth development and evaluation of strategic reasoning. Our game environment is governed by the umpire responsible for facilitating games, from matchmaking through move validation to environment management. Players incorporate state-of-the-art LLMs in their decision mechanism, relying on a formal hypergame-based model of hierarchical beliefs. We use one-shot, 2-player beauty contests to evaluate the recursive reasoning capabilities of the latest LLMs, providing a comparison to an established baseline model from economics and data from human experiments. Furthermore, we introduce the foundations of an alternative semantic measure of reasoning to the k-level theory. Our experiments show that artificial reasoners can outperform the baseline model in terms of both approximating human behaviour and reaching the optimal solution.

Keywords: Multi-agent Systems · Social Simulations · Strategic Reasoning.

1 Introduction

Multi-agent systems provide environments for individual-based modelling and simulations [28]. Game theory and multi-agent-based simulation (MABS) have established a mutually beneficial relationship: MAS leverages game-theoretic interaction models and strategic tools [31], and game theory relies on multi-agent-based social simulations to investigate strategic decision-making [36]. Large language models (LLMs) have received particular interest in their potential to simulate human-like reasoning and decision-making, and MABS frameworks are used to evaluate LLM capabilities in game-theoretic environments [17]. Traditional approaches often rely on simplified agent frameworks implemented with a weaker concept of agency [45], which may impose limitations on the system's

adaptability and the agent's reasoning sophistication. In contrast, our approach involves a stronger, modular agent concept enabling a decoupled investigation of reasoning processes. Our agents are hosted in a role-based multi-agent framework governed by an umpire who manages game environments and facilitates agent interactions.

We focus on beauty contest games, a well-established concept for studying recursive reasoning [10]. We evaluate the reasoning capabilities of LLM-enhanced agents, integrating recursive reasoning via a formal hypergame representation. Our experiments compare the performance of LLM-enhanced agents against both a baseline economic model and human data, providing insights into the models' ability to approximate human strategic behaviour. Furthermore, we introduce a self-evaluation method κ, a revised measure of reasoning depth that complements traditional k-level theory, offering a more nuanced understanding of reasoning sophistication.

Our key contributions include:

- A flexible multi-agent-based simulation platform capable of hosting a wide array of reasoners, offering a detailed view of reasoning processes;
- LLM-enhanced agents that leverage a hypergame-based model for recursive reasoning;
- Introduction of κ, a complementary measure to k-level reasoning;
- Experiments comparing our LLM-based reasoners to the baseline model and human data.

Our results suggest that artificial reasoners can benefit from the expanded architectural complexity and can not only match but potentially outperform baseline models in both approximating human behaviour and achieving optimal solutions in strategic settings.

2 Background

2.1 Game Theory

Game theory provides a mathematical framework for analyzing decision-making in multi-agent – human or artificial – strategic interactions [30]. A game is formally defined by its players, their available strategies, and utility functions that evaluate the players' outcomes [33]. Beauty contest games (BCGs) provide an experimental testbed for iterative reasoning, where players have to guess a number which they believe will be the closest to the mean of all guesses weighted by a parameter p [10]. BCGs are a popular choice for studies concerning the k-level theory, as optimal play requires thinking about others' thoughts: given the range $[0, 100]$ and $p = \frac{2}{3}$, 0-level players pick 50 and level-k thinkers choose $50\frac{2}{3}^k$. Experimental evidence consistently shows that most human players exhibit 1st or 2nd-level reasoning, with few advancing beyond level 3 [29, 14, 7].

Hypergames extend standard game theory by modelling individual player perspectives, allowing hypergame models to capture misaligned perceptions [4].

Multi-level hypergames provide a formalized model of hierarchical games representing nested beliefs [40]. A third-level hypergame between players i, j, is a composite structure of lower-level hypergames representing individual players' perspectives: $H^3 = \{H_i^2, H_j^2\}$, where $H_i^2 = \{H_{ii}^1, H_{ji}^1\}$ and $H_{ji}^1 = \{H_{iji}^0, H_{jji}^0\}$. Finally, $H_{iji}^0 = G_{iji}$ is i's perceptual game defining player i's belief of j's belief of i's perspective of G. This theoretical framework provides the foundation for studying how agents engage in recursive strategic reasoning and form hierarchical beliefs about others' decision-making processes.

2.2   Language Models

Large language models (LLMs) are sophisticated neural networks – usually based on the transformers architecture [15] and a pre-trained model on a vast amount of data [32] – that target natural language processing applications [48]. Chain-of-thought (CoT) prompting is a technique that improves LLMs reasoning capabilities by having LLMs decompose complex tasks into smaller problems [42]. Such prompts can be demonstrative examples showcasing what the expected response may pertain to and/or descriptive instructions that guide the model on reaching the expected response [46].

Claude 3.5   Claude 3.5 is a family of state-of-the-art LLM models from Anthropic, supporting multimodal applications [3]. In this work, we evaluate two Claude models: Sonnet offers large context windows and advanced analytical capabilities, suitable for complex tasks and process automation; Haiku is a smaller, cost-efficient, fast model targeting interactive and sub-agent tasks.

GPT-4   GPT-4 is OpenAI's SOTA system supporting multimodal input and output and a long context window arming models with a broad general knowledge and advanced problem-solving abilities [1]. We implement two models: GPT-4o possesses a generally high reasoning performance across various benchmarks; 4o-mini is a lightweight variant for resource-constrained environments.

2.3   Recursive Reasoning

We describe recursive reasoning as an agent's ability to reason about each other's physical and cognitive states [44]. The cognitive hierarchy model introduced the concept of k-level thinkers, proposing that players engage in different levels of strategic thinking, where players on level k best respond to $k-1$ level strategies, assuming every other player must be at most at level $k-1$ [10]. Epistemic game theory is a branch of game theory which provides a formal mathematical framework for representing and operating with belief hierarchies [11]. Hypergame theory aims to analyse conflict under asymmetric information and misaligned perceptions [5], providing a game-oriented representation of sequential beliefs.

3   Multi-agent Simulation via Centralized Hypergames

Our simulations are materialized in game environments composed of and hosting an umpire, a set of players, and a set of hypergames. We revise hierarchical hypergames from [40] to integrate 2-player BCGs as individual perceptual games capturing the agent's beliefs and reasoning level.

### 3.1   Perceptual Beauty Contest Games

In our framework, we define BCGs formally as $G = (N, A, U, \Psi)$, where $N = i, j$ is the set of two players, $A = A_i \times A_j$ is the action space, where $A_i, A_j \subseteq \mathbb{Z}$ represent the available actions for players i and j respectively, $U : A \to \mathbb{R}^2$ is the utility function, where for each player: $U_i(a_i, a_j) = -|a_i - p \cdot \mu|$, with $\mu = \frac{a_i + a_j}{2}$ and p as the BCG's scalar parameter and $\Psi = (\psi_1 = \sigma, \psi_2, \dots, \psi_\kappa)$ is an ordered sequence of κ number of perspectives, where the first component ψ_1 denotes the interpreter (creator) of the game perspective, σ. To help position player beliefs in our BCGs in the context of belief hierarchies, we suggest simplifying assumptions as follows. Given a $\kappa = 2$ level reasoner i, $\beta_i(\beta_j(\beta_k))$ [11] corresponds to i's beliefs about player j's beliefs about player k's reasoning. Then we assume $\beta_i(\beta_j(\beta_k)) \cong G_{ijk}$, where G_{ijk} is the perceptual game capturing the same beliefs, with subscripts denoting the sequential order of perspectives: i's reconstruction of G based on his beliefs about j's beliefs of k's perspective. However, the rigorous analysis of the proposed relationship is beyond the scope of this work. Finally, the set of individual perceptual games invokes the hypergame H^κ, capturing both players' beliefs, where at least one of the players exhibits the highest level of reasoning κ.

3.2   Recursive Reasoners

Agents in our framework are implemented following the standard intelligent agent architecture [34] and inspired by the Observe-Orient-Decide-Act decision-making model [8, 38] – as shown on Figure 1a. An iteration of the game environment comprises the umpire's matchmaking activities and the players' reasoning processes. Given the space of natural language game descriptions X, the umpire υ sends game requests to each pair of players to participate in $G^* = (\{i, j\}, \{A_i, A_j\}, U, (\upsilon))$. In the facilitated games, the umpire is a passive participant – a pseudoplayer [33].

Each player's reasoning processes are then decoupled and defined as follows:

– The player's revision module – Figure 1b – processes the game description, reasons about what the opponent's move could be, and constructs a perceptual game. These steps are integrated into an interpretation function $I : G \times N \to G'$, that creates an instance of G reflecting the player's beliefs:
  • $\rho : X \to \Xi, \mathbb{R}$ is the reasoning function: $\rho(x) = \xi_i, \hat{a}_j$ where $\xi_i \in \Xi$ is i's natural language reasoning based on game description $x \in X$ on what the opponent j's guess $\hat{a}_j$ is expected to be;
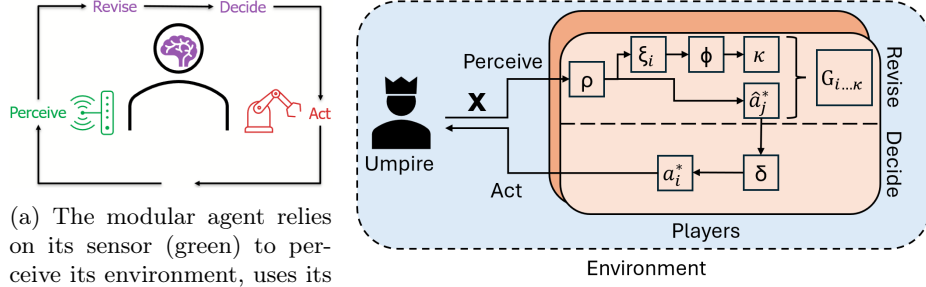
(a) The modular agent relies on its sensor (green) to perceive its environment, uses its mind (purple) to revise the perception and decide which action to take, and acts upon the environment via its effector (red).

(b) Upon perception of a game request with textual description x, the player revises x, generating a reasoning ξ_i and producing an expected opponent choice $\hat{a}_j$. After $\phi(\xi_i)$ derives κ, the perceptual game $G_{i...\kappa}$ is established. $\delta(\hat{a}_j) = a_i^*$ is what i decides and acts on.

Fig. 1: Overview of centralised MAS framework.

- $\phi : \Xi \to \mathbb{N}$ is the reasoning analysis function: $\phi(\xi_i) = \kappa$, where κ is the player's estimated reasoning level based on the number of nested beliefs present in the reasoning ξ_i;
- i's perceptual game is then: $G_{i...\kappa} = (i, \{i, j\}, A, U, \Psi = (\psi_i, \ldots, \psi_\kappa))$.
  - The decision module selects an action based on the revised expectation:
    - $\delta : \mathbb{R} \to A_i$ is the decision function selecting i's preferred action a_i^* based on the expected opponent guess: $\delta(\hat{a}_j) = a_i^*$;

While k-levels are derived from players' numerical choices, κ provides an alternative measure based on the explicit reasoning steps we observe in players' natural language explanations. These reasoning steps are represented as perspectives ψ in the perceptual game G, allowing us to analyze the depth of strategic thinking through players' own articulation of their decision process.

3.3 Model Prompting Design

We integrate CoT prompting into our multi-step reasoning process across the agent's *revise* and *decide* phases. During *revise*, the LLM mind processes information through sequential steps: first reasoning about the situation to predict opponent guesses, then analyzing the reasoning depth to determine κ, resulting in the κth order perceptual game. The mind's decision function then interprets this processed information to derive the final guess, completing the multi-step reasoning process. LLM prompts follow a modular structure: (1) optional agent profile for context-specific reasoning [25], (2) role and task definition, (3) game/task specific request, and (4) implementation-specific requirements (e.g., *"surround the chosen number in curly brackets: n"*).

3.4   Baseline Model

The self-tuning experience weighted attraction model (EWA) is a benchmark model for reproducing human-like strategic interaction in game-theoretic experiments [18].

$$A_i^j(t) = \frac{\phi \cdot N(t-1) \cdot A_i^j(t-1) + [\delta + (1-\delta) \cdot I(s_i^j, s_i(t))] \cdot \pi_i(s_i^j, s_{-i}(t))}{N(t-1) \cdot \phi \cdot (1-\kappa)} \quad (1)$$

Equation 1 defines player i's associated attention to strategy j at time (or round) t, where

- $\pi_i(s_i^j, s_{-i}(t))$ denotes $i's$ payoff for choosing strategy j against s_{-i} at time t;
- N is the experience weight $N(t) = (1-\kappa) \cdot \phi \cdot N(t-1) + 1$ [9];
- ϕ is the change detector function $\phi(t) = 1 - \frac{1}{2}S_i(t)$;
- the surprise index is denoted by $S_i(t) = \sum_{k=1}^{m_{-i}}(h_i^k(t) - r_i^k(t))^2$ with
  $r_i^k(t) = I(s_{-i}^k, s_{-i}(t))$ and $h_i^k(t) = \frac{\sum_{\tau=1}^t I(s_{-i}^k, s_{-i}(\tau))}{t}$;
- δ is the weight to foregone payoffs: $\delta_{ij}(t) = \begin{cases} 1 & \text{if } \pi_i(s_i^j, s_{-i}(t)) \geq \pi_i(t), \\ 0 & \text{otherwise.} \end{cases}$;
- and $I(x, y)$ is an index function returning 0 if $x = y$ and 1 otherwise.

Let $t = 0$ denote the agent's initial state. We then define the initial attraction $A_i^j(0)$ according to the cognitive hierarchy model and the Poission distribution function $P(k) = \frac{e^{-\tau}\tau^k}{k!}$ [10] and set $N(0) = 1$. The self-tuning EWA model was trained and tested on a 7-player beauty contest with $p = 0.7$ and $p = 0.9$ [19], for which Ho et al. determined $\lambda = 2.39$ for the sensitivity of the response function and $\tau = 1.5$ for deriving the first-period plays via the CHM-derived Poisson distribution. We adopt these parameters in our experiments without further tuning. Finally, agents choose an action according to $P_j^i(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}}$.

4   Experiments

We use beauty contest games as the test bench for evaluating our agents' recursive reasoning capabilities, as guessing games are closely associated with k-level reasoning [10]. In this context, k level reasoners best respond to the $k-1$ level players' choices and at $k = 0$ players choose randomly. In order to classify players by their reasoning, we adopt the guess-based conversion approach presented in [29]. We assume level 0 reasoners choose 50; then we set each kth reasoning level at $50p^k$. More specifically, we use 2-player BCGs to conduct our evaluation. In n-player BCGs, reasoners eliminate weakly dominated strategies iteratively – which process can be extended to infinity – until the theoretical solution of everyone choosing 0 is reached [10]. The two-player game provides a simpler solution concept, where the smaller number wins [16]. Zero is a weakly dominated

strategy that always wins, which, in theory, significantly simplifies the iterative reasoning process.

We evaluate our agents on the experiment from [16], involving 132 student participants, pooled from first-year students with no prior game-theoretic training nor existing familiarity with beauty contests and 130 professionals with extensive game-theoretic domain knowledge.

In the first instance, we replayed the 2-player beauty contests using the original experiment design, simulating 25 independent rounds with each LLM. Additionally, we conducted 60 rounds with pairs of agents using the EWA model as the benchmark. Artificial agents were provided with a description worded similarly to what human participants would be given – however, obfuscated to mitigate reliance on game-theoretic analyses in the models' training data –, without any explicit instructions on their reasoning and decision-making. Providing a persona description as context pushes the LLM to behave in the desired way by guiding the reasoning process according to the profile specification [25]. Similarly, in the second set of experiments, we expand our prompting mechanism with an agent profile, allowing us to specify the level of domain knowledge the agent should use for its reasoning. Following the original group descriptions from [16], we ran 15-15 games with the specifications of "first years students with no game-theoretic knowledge" and "professors with expert domain knowledge in game theory" with Claude 3.5 Haiku and GPT-4o – chosen based on their proximity to the human performance and the optimal strategy, 0.
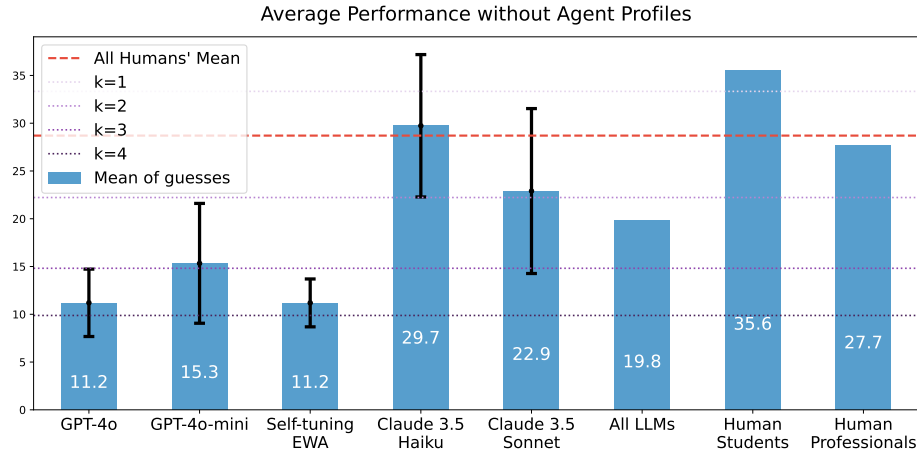

Fig. 2: Per model means, standard deviations and estimated k-levels. Human data for standard deviation was unavailable.

All models outperformed human participants significantly, except for Claude 3.5 Haiku, which approximates the results of the professionals and the joint human player population – Figure 2. On the other hand, Figure 3 depicts the
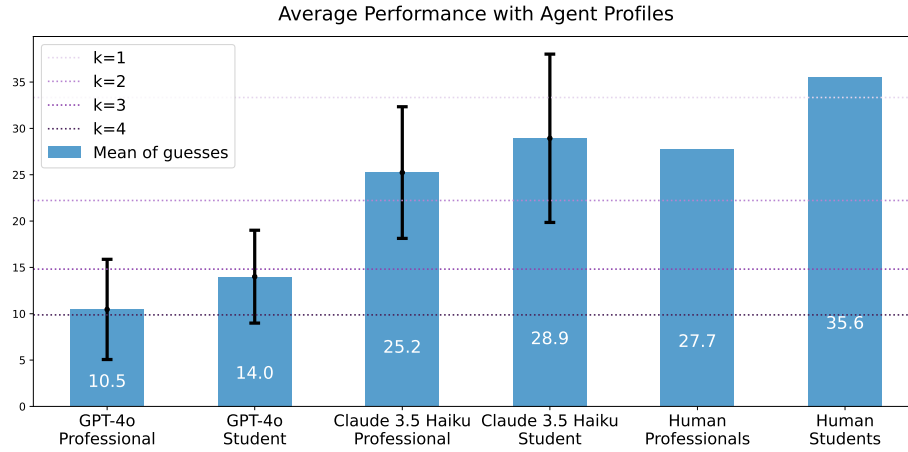
Average Performance with Agent Profiles

Fig. 3: Per model means (LLMs with profiles), standard deviations and estimated k-levels. Human data for standard deviation was unavailable.

results of the second experiment, where player profiles are successfully integrated into both agents' reasoning processes. Similar to the human groups with different levels of domain expertise, the models acting as students performed noticeably poorer than the models that were described as professionals. However, while 9.85% of human students and 36.02% of human professionals managed to reach 0, none of the artificial reasoners – nor the baseline model – chose the optimal strategy in either experiment. Most studies concentrate on BCGs involving more than 2 players, where 0 is not an obvious solution to the game. This bias is likely inherently present in the LLMs training data, potentially tainting the information used to generate reasoning in the 2-player BCG.

K-level reasoning provides another metric for comparing human and artificial agents' thought processes. Table 1 reinterprets the mean results through k-level classifications, showing that while humans exhibit 1st-level reasoning, only Claude 3.5 Haiku produced comparable results. Other models performed at the 2nd and 3rd levels, with EWA and GPT-4o approaching level 4 reasoning. Additionally, the "Mean κ" and "Median κ" columns report the LLM agents' reasoning steps involved, corresponding to perspectives in their perceptual games.

The second experiment revealed that the estimated k and κ levels are aligned with the expectation that reasoners with domain expertise would outperform and exhibit a higher order of reasoning than non-professionals. While κ can provide a semantics-backed estimate of the reasoning steps involved in the process, it does not yet tell us much about the quality of the agent's reasoning. The results suggest that a semantic qualitative reasoning classification may complement the current numerically defined measures.

Model/Group	Mean Guess	Mean k-level	Mean κ	Median κ
Human Students	35.57	0.84	N.A.	N.A.
Human Professionals	27.73	1.45	N.A.	N.A.
Claude 3.5 Haiku	29.7	1.28	2.17	2
Claude 3.5 Sonnet	22.9	1.92	2.17	2
GPT-4o-mini	15.3	2.92	2.57	3
GPT-4o	11.2	3.69	2.27	2
Self-tuning EWA	11.2	3.69	N.A.	N.A.
Claude 3.5 Haiku Student	28.9	1.35	1.97	2
Claude 3.5 Haiku Professional	25.2	1.69	2.5	2.5
GPT-4o Student	14	3.14	1.8	2
GPT-4o Professional	10.5	3.85	2.4	3

Table 1: Mean k-levels are estimated from mean guesses, while mean and median κ levels correspond to LLM agents' self-reported reasoning steps.

5   Related Work

The recent emergence of LLM agents [39] and LLM-MAS [22] has influenced social simulations. Multi-agent strategic interactions are used to evaluate LLM reasoning [27, 13], and LLM-based simulations are leveraged for empirical investigations on strategic behaviours [25, 26]. Recent work on LLM-driven reasoning in BCGs focuses on studying agent rationality and reasoning levels [47, 24, 17]. Centralized role-based architectures rely on an umpire [35] or a game manager [21] to coordinate game-based interactions – such as artificial trading [43] or utility markets [23, 49]. Hypergame theory, while effective for analyzing complex conflicts post-hoc [5], has seen limited practical application [6] due to challenges in automation. Most agent-based implementations only borrow conceptually [12, 2], with few examples of full integration [20, 37].

While prior approaches typically implement a looser agentic concept – a weaker notion of agency [45] – potentially constraining the system's flexibility and bounding LLMs' reasoning [41], our framework offers enhanced flexibility and depth. Our conceptually elaborate multi-agent architecture enables a more nuanced evaluation process, facilitating a systematic review of LLM reasoning and automating the generation of hypergames. This approach provides valuable insights into LLMs' capabilities for recursive reasoning, particularly their ability to form beliefs about beliefs and develop a theory of mind.

6   Conclusions and Future Work

We present a MABS framework integrating hypergames for multi-level reasoning in BCGs. Our contributions are threefold. First, we introduce a flexible multi-agent simulation platform capable of hosting both simple models like the self-tuning EWA and sophisticated multi-step reasoners powered by state-of-the-art LLMs. The platform's architecture emphasizes a strong notion of agency,

allowing for systematic development and evaluation of strategic reasoning capabilities. In contrast, existing approaches often employ simplified architectures with weaker notions of agency, which can limit both applicability and evaluation depth. Second, we introduce κ as a complementary measure to the k-level theory for evaluating reasoning. While k-levels are solely derived from players' numerical choices, κ represents the sequential perspectives in an agent's reasoning process, providing an alternative measure of reasoning depth that could offer valuable insights alongside numerical k-level analysis. Third, our 2-player BCG experiments demonstrate that our proposed approach can match and outperform the baseline model in approximating human results and reaching the optimal solution. Additionally, our profile-enhanced agents successfully mirrored the performance gap observed between student and professional human players.

The presented system and experiments highlight the potential for extensions. Integrating semantic analysis techniques for evaluating the quality and coherence of agent reasoning processes would allow us to develop κ into a deeper qualitative measure of reasoning sophistication. Recreating the BCG experiment with human reasoning data would allow us to investigate (dis)similarities in how humans and LLM agents approach recursive strategic thinking, both in terms of k and κ levels. Implementing multi-round and n-player games would enable studying learning processes in artificial and human agents. These extensions would contribute to a more comprehensive understanding of recursive reasoning in artificial – and potentially human – agents while advancing the practical applications of LLM-driven multi-agent systems in strategic decision-making scenarios.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Achiam, J., Adler, S., Agarwal, S., Ahmad, L., et al.: Gpt-4 technical report (2024), https://arxiv.org/abs/2303.08774
2. Aitchison, M., Benke, L., Sweetser, P.: Learning to deceive in multi-agent hidden role games. In: Sarkadi, S., Wright, B., Masters, P., McBurney, P. (eds.) Deceptive AI. pp. 55–75. Springer International Publishing, Cham (2021)
3. Anthropic: Claude 3.5 (2024), https://www.anthropic.com/claude
4. Bennett, P.: Toward a theory of hypergames. Omega **5**(6), 749–751 (1977)
5. Bennett, P.: Hypergames: Developing a model of conflict. Futures **12**(6), 489–507 (1980)
6. Bennett, P., Huxham, C., Dando, M.: Shipping in crisis: A trial run for 'live' application of the hypergame approach. Omega **9**(6), 579–594 (1981)
7. Bosch-Domenech, A., Montalvo, J.G., Nagel, R., Satorra, A.: One, two,(three), infinity,...: Newspaper and lab beauty-contest experiments. American Economic Review **92**(5), 1687–1701 (2002)

8. Bryant, D.J.: Rethinking ooda: Toward a modern cognitive framework of command decision making. Military Psychology **18**(3), 183–206 (2006)

9. Camerer, C.F., Ho, T.H., Chong, J.K.: Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. Journal of Economic Theory **104**(1), 137–188 (2002)

10. Camerer, C.F., Ho, T.H., Chong, J.K.: A cognitive hierarchy model of games. The Quarterly Journal of Economics **119**(3), 861–898 (2004)

11. Dekel, E., Siniscalchi, M.: Epistemic game theory. In: Handbook of game theory with economic applications, vol. 4, pp. 619–702. Elsevier (2015)

12. Dharmadhikari, M., Deshpande, H., Dang, T., Alexis, K.: Hypergame-based adaptive behavior path planning for combined exploration and visual search. In: 2021 IEEE International Conference on Robotics and Automation (ICRA). pp. 269–275 (2021)

13. Duan, J., Zhang, R., Diffenderfer, J., Kailkhura, B., Sun, L., et al.: Gtbench: Uncovering the strategic reasoning limitations of llms via game-theoretic evaluations. arXiv preprint arXiv:2402.12348 (2024)

14. Duffy, J., Nagel, R.: On the robustness of behaviour in experimental 'beauty contest'games. The Economic Journal **107**(445), 1684–1700 (1997)

15. Gillioz, A., Casas, J., Mugellini, E., Abou Khaled, O.: Overview of the transformer-based models for nlp tasks. In: 2020 15th Conference on computer science and information systems (FedCSIS). pp. 179–183. IEEE (2020)

16. Grosskopf, B., Nagel, R.: The two-person beauty contest. Games and Economic Behavior **62**(1), 93–99 (2008)

17. Guo, S., Bu, H., Wang, H., Ren, Y., Sui, D., Shang, Y., Lu, S.: Economics arena for large language models (2024), https://arxiv.org/abs/2401.01735

18. Ho, T.H., Camerer, C.F., Chong, J.K.: Self-tuning experience weighted attraction learning in games. Journal of economic theory **133**(1), 177–198 (2007)

19. Ho, T.H., Camerer, C., Weigelt, K.: Iterated dominance and iterated best response in experimental "p-beauty contests'. The American Economic Review **88**(4), 947–969 (1998)

20. Kahn, M.: Dynamic-Occlusion-Aware Risk Identification for Autonomous Vehicles Using Hypergames. Master's thesis, UWSpace (2021)

21. Kobti, Z., Sharma, S.: A multi-agent architecture for game playing. In: 2007 IEEE Symposium on Computational Intelligence and Games. pp. 276–281 (2007)

22. Li, X., Wang, S., Zeng, S., Wu, Y., Yang, Y.: A survey on llm-based multi-agent systems: workflow, infrastructure, and challenges. Vicinagearth **1**(1), 9 (2024)

23. Liu, N., He, L., Yu, X., Ma, L.: Multiparty energy management for grid-connected microgrids with heat- and electricity-coupled demand response. IEEE Transactions on Industrial Informatics **14**(5), 1887–1897 (2018)

24. Lu, S.E.: Strategic interactions between large language models-based agents in beauty contests (2024), https://arxiv.org/abs/2404.08492

25. Mao, S., Cai, Y., Xia, Y., Wu, W., et al.: Alympics: Llm agents meet game theory – exploring strategic decision-making with ai agents (2024), https://arxiv.org/abs/2311.03220

26. Mensfelt, A., Stathis, K., Trencsenyi, V.: Autoformalizing and simulating game-theoretic scenarios using llm-augmented agents. arXiv preprint arXiv:2412.08805 (2024)

27. Mensfelt, A., Stathis, K., Trencsenyi, V.: Logic-enhanced language model agents for trustworthy social simulations (2024), https://arxiv.org/abs/2408.16081

28. Michel, F., Ferber, J., Drogoul, A.: Multi-agent systems and simulation: A survey from the agent community's perspective. In: Multi-Agent Systems, pp. 17–66. CRC Press (2018)
29. Nagel, R.: Unraveling in guessing games: An experimental study. The American Economic Review **85**(5), 1313–1326 (1995)
30. Osborne, M.: Introduction to Game Theory. Oxford University Press USA (2004)
31. Parsons, S., Wooldridge, M.: Game theory and decision theory in multi-agent systems. Autonomous Agents and Multi-Agent Systems **5**, 243–254 (2002)
32. Qiu, X., Sun, T., Xu, Y., Shao, Y., et al.: Pre-trained models for natural language processing: A survey. Science China technological sciences **63**(10), 1872–1897 (2020)
33. Rasmusen, E.: Games and information an introduction to game theory. Blackwell (2006)
34. Russell, S., Norvig, P.: Intelligent agents. Artificial intelligence: A modern approach **74**, 46–47 (1995)
35. Stathis, K., Sergot, M.: An abstract framework for globalising interactive systems. Interacting with Computers **9**(4), 401–416 (1998), shared Values and Shared Interfaces: The Role of Culture in the Globalisation of Human-Computer Systems
36. Sun, R.: Cognition and multi-agent interaction: From cognitive modeling to social simulation. Cambridge University Press (2006)
37. Tang, Y., Sun, J., Wang, H., Deng, J., et al.: A method of network attack-defense game and collaborative defense decision-making based on hierarchical multi-agent reinforcement learning. Computers & Security **142**, 103871 (2024)
38. Tweedale, J., Ichalkaranje, N., Sioutis, C., Jarvis, B., Consoli, A., Phillips-Wren, G.: Innovations in multi-agent systems. Journal of Network and Computer Applications **30**(3), 1089–1115 (2007)
39. Wang, L., Ma, C., Feng, X., Zhang, Z., et al.: A survey on large language model based autonomous agents. Frontiers of Computer Science **18**(6), 186345 (2024)
40. Wang, M., Hipel, K.W., Fraser, N.M.: Modeling misperceptions in games. Behavioral Science **33**(3), 207–223 (1988)
41. Wang, Q., Wu, J., Tang, Z., Luo, B., et al.: What limits llm-based human simulation: Llms or our design? (2025), https://arxiv.org/abs/2501.08579
42. Wei, J., Wang, X., Schuurmans, D., Bosma, M., et al.: Chain-of-thought prompting elicits reasoning in large language models (2023), https://arxiv.org/abs/2201.11903
43. Wellman, M., Wurman, P., O'Malley, K., Bangera, R., et al.: Designing the market game for a trading agent competition. IEEE Internet Computing **5**(2), 43–51 (2001)
44. Wen, Y., Yang, Y., Luo, R., Wang, J., Pan, W.: Probabilistic recursive reasoning for multi-agent reinforcement learning. arXiv preprint arXiv:1901.09207 (2019)
45. Wooldridge, M., Jennings, N.R.: Intelligent agents: theory and practice. The Knowledge Engineering Review **10**(2), 115–152 (1995)
46. Yu, Z., He, L., Wu, Z., Dai, X., Chen, J.: Towards better chain-of-thought prompting strategies: A survey. arXiv preprint arXiv:2310.04959 (2023)
47. Zhang, Y., Mao, S., Ge, T., Wang, X., Xia, Y., Lan, M., Wei, F.: K-level reasoning: Establishing higher order beliefs in large language models for strategic reasoning (2024), https://arxiv.org/abs/2402.01521
48. Zhao, W.X., Zhou, K., Li, J., Tang, T., et al.: A survey of large language models. arXiv preprint arXiv:2303.18223 (2023)
49. Zhu, B., Wang, D.: Master–slave game optimal scheduling for multi-agent integrated energy system based on uncertainty and demand response. Sustainability **16**(8) (2024)