Who is Undercover? Guiding LLMs to Explore Multi-Perspective Team Tactic in the Game

Ruiqi Dong, Zhixuan Liao, Yuhan Ma, Danni Ma, and Chenyou Fan(✉)

South China Normal University, Foshan 528225, China
{2023025193,20202121032,2022024953,20202833107}@m.scnu.edu.cn,
fanchenyou@scnu.edu.cn

Abstract. Large Language Models (LLMs) are powerful AI agents but struggle with open decision-making in complex scenarios. To address this, we propose the Multi-Perspective Team Tactic (MPTT) framework, using the language logic game "Who is Undercover?" (WIU) as a testbed. MPTT enhances LLMs' language logic, multi-dimensional thinking, and self-awareness. Through alternating speaking and voting sessions and techniques like self-perspective, identity determination, self-reflection, and multi-round collaboration, LLM agents make rational decisions by strategically concealing information and building trust. Experiments show that MPTT, paired with WIU, harnesses LLMs' cognitive capabilities to simulate societal decision-making, supporting minority group communication and promoting fairness. We also introduce a word network dataset with multidimensional semantic relations for WIU and other language-intensive tasks. Human-in-the-loop experiments further reveal LLMs' ability to learn and align with human behavior, showcasing their potential in societal decision-making.

Keywords: Multi-Agent · Multi-Perspective · Human-in-the-loop.

1 Introduction

Decision-making in human society is a complex and important activity that involves individuals and organizations making choices in various fields in response to changing situations. However, communication inequalities faced by marginalized groups and minorities in uncertain situations often challenge the fairness of these decisions. Incorporating AI technology in these activities can improve rationality and effectiveness in decision-making while addressing such inequalities. Chain-of-Thought (CoT) [19] improves the performance of LLMs on complex reasoning tasks. AI Agent has been emerging to solve automated tasks like HuggingGPT [13]. To strengthen LLMs' decision-making abilities, Self-Refine [10] improves the initial output through iterative feedback and improvement. Furthermore, when functioning as AI agents, LLMs can decompose complex problems into more manageable sub-tasks [5] and exhibit human-like natural language interaction abilities [18]. However, AI agents often struggle with open decision-making in complex scenarios. Therefore, LLMs need to have better understanding of human societal rules to enhance decision-making rationality.

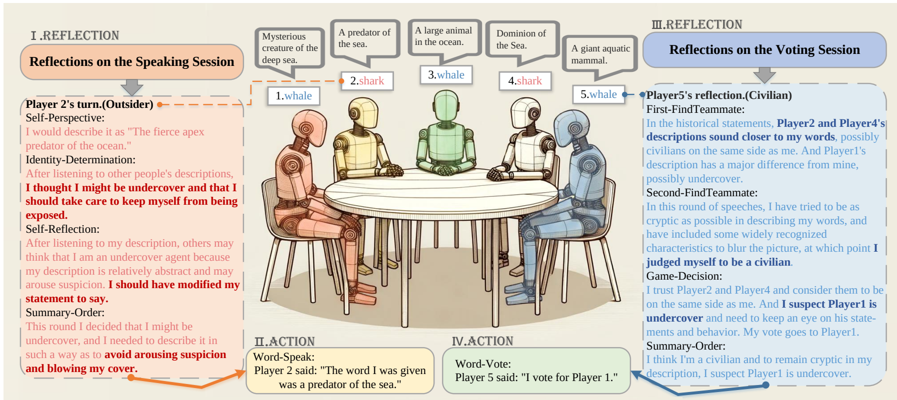

Fig. 1: **Presentation of WIU in the framework of MPTT.** MPTT employs techniques like self-perspective, identity-determination, self-reflection, and multi-round find-teammates, with session summaries aiding decision-making.

In this work, we use the game "Who is Undercover?" (WIU), a reasoning game testing decision-making skills, as our foundation. We create a specific dataset of word network as tasks of linguistic description and reasoning in games. Inspired by the Theory of Mind [8] and Social Identity [14] in Social Psychology, we aim to simulate human thought processes in LLMs. Therefore, we employ LLMs as AI agent players and design the "Multi-Perspective Team Tactic" (MPTT) framework. MPTT alternates between speaking and voting sessions, incorporating several multi-perspective techniques.Fig. 1 shows the process of using MPTT to play WIU. To enhance game realism and complexity, we designed the "Human-in-the-loop" to explore human-AI collaboration in social interactions.

Research shows that MPTT iteratively optimizes LLM agents' mindsets, fostering strategic behaviors like confrontation and concealment, alongside tendencies like trust, suspicion, and cooperation. Applied to the WIU game, MPTT creates a decision-making mechanism as a reference for human society and helps minorities communicate and express their choices, promoting balanced decision-making across diverse groups. Additionally, LLMs are expected to actively participate in future social decision-making alongside humans. [1]

2 Related Work

Some studies [3] explored using LLMs to identify deceptive information. Wang et al. [17] proposed the Recursive Contemplation (ReCon) framework on the Avalon game to explore the potential of LLM in deceptive environments. Xu et al. [20] explored the problem of how to use LLMs in communication games Werewolf. Game theory [4] finds diverse applications in economic analysis [6], spanning

[1] Our program is publicly available at https://github.com/FancyAI-SCNU/wiu-game

market competition and trade freedom [16]. WIU is a process of conducting a static game with incomplete information [4]. Kroer et al.[7] devised strategies for playing against a limited prospective player. But through WIU game training, AI agents gain insights into the challenges posed by incomplete information games in human society. WIU emphasizes logical deduction and reasoning, appealing to players who favor strategic thinking over social manipulation. This distinguishes it from the intense social interaction and deception commonly seen in Werewolf [20] and Avalon [17].

3 Game description and dataset

3.1 The game role of WIU

“Who is Undercover? ” is a reasoning game where multiple civilian players are mixed with a minority of undercover players. There are 3 civilians and 2 undercovers. Each player is given a similar but different word without knowing their identity and takes turns describing their word. The opponents are eliminated through description and thinking. When there is only one civilian left but there is still an undercover, the undercover wins, and if there is no undercover, the civilian wins. The challenge in this game lies in discerning one’s role, as the words share many similarities.

3.2 Dataset Construction

Existing semantic datasets like WordNet [11] and ConceptNet [9] are unsuitable for the WIU game due to their overly broad semantic associations, lack of fine-grained distinctions, and absence of explicit labeling for polysemy and contextual relevance. To address these limitations, we developed a tailored word network dataset that enables dynamic reasoning, multi-agent collaboration, and game reproducibility, while also offering broader applications in resolving semantic conflicts and enhancing language understanding.

The dataset construction involves three steps: (1) Word Screening: Selecting word pairs with semantic similarities or differences, ensuring polysemous or ambiguous words are paired based on distinct meanings and contextual relevance. (2) Semantic Relation Labeling: Annotating word pairs with relations like similarity, difference, polysemy, ambiguity, and contextual association, with descriptive attributes for each term. (3) Data Expansion: Using GPT-3.5-turbo to generate additional words while maintaining quality and compliance.

The dataset is organized into five parts for targeted use: (1) Similarity and Difference: Fine-grained semantic relations (e.g., fruits). (2) Polysemy: Words with distinct meanings (e.g., "mouse"). (3) Ambiguity: Words requiring contextual interpretation (e.g., "python"). (4) Cultural and Linguistic: Terms tied to specific cultural contexts. (5) Domain Relevance: Field-specific terminology.

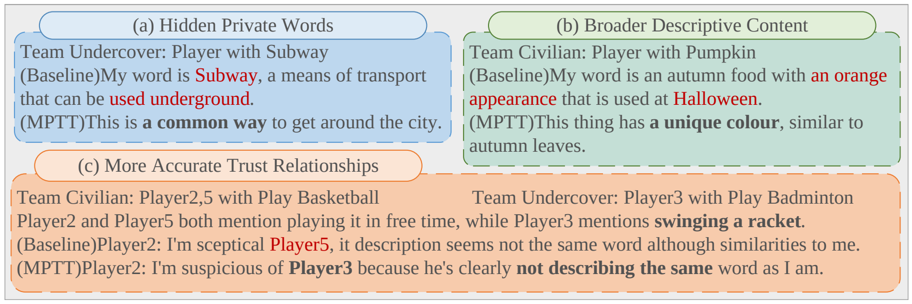

Fig. 2: **Improvement with MPTT.** MPTT is significantly improved in 3 areas.

4 MPTT: A framework for the reasoning and communication type of game

4.1 Overall process

A complete game consists of multiple rounds, indexed by r, each including a speaking and voting session, until one side wins. For N players, the turn of player α is denoted as $N = \alpha$. Each player receives a word, and the speech history is stored in set H, with summarized reflections continuously updated in set O. The MPTT divides the game into two phases: speaking and voting, to privately reflect on roles and generate thoughtful responses that balance revealing information with maintaining secrecy, and analyze previous speeches, identify teammates, and make strategic voting decisions based on incomplete information.

4.2 Phase I: Reflections on the Speaking Session

In the first phase of our framework, players reflect on their roles privately before delivering their speeches, aiming to enhance adaptability and flexibility in providing diverse, accurate descriptions while concealing private information. This addresses the issues of (a) and (b) in Fig. 2.

Self-Perspective. This stage prompts the AI agent to describe words in one sentence from its own perspective. Suppose it is now the turn of player α ($\alpha \in \{1, ..., n\}$) to speak, Player α will think following the principle of Self-Perspective, rounds other than the first will have history H and self-summary O_α references:

$$T_\alpha = \text{Self-Perspective}\{H, O_\alpha\}^r_{N=\alpha} \tag{1}$$

Identity-Determination. Each player doesn't know their true role as civilian or undercover, and must initially assess their identity before speaking. Player α determines it identity based on the global historical records H:

$$M_\alpha = \text{Identity-Determination}\{H, O_\alpha, T_\alpha\}^r_{N=\alpha} \tag{2}$$

Self-Reflection. Player α needs to reflect on itselves to find common features in the description to avoid exposure.

$$R_\alpha = \text{Self-Reflection}\{H, O_\alpha, T_\alpha, M_\alpha\}^r_{N=\alpha} \tag{3}$$

After these reflections, the AI agents will make a summary of ideas O_α, which mainly includes self-conclusion and the speaking recommendations, update with rounds:

$$O_\alpha{}' = \text{Summary-Order}\{T_\alpha, M_\alpha, R_\alpha\}^r_{N=\alpha} \quad O_\alpha \leftarrow O_\alpha{}' \tag{4}$$

$$O \leftarrow O \cup \{O_\alpha\}^r_{N=\alpha} \tag{5}$$

At the same time, based on the content of the reflections in all speaking phases we get the final version of the players' speeches:

$$W_\alpha = \text{Word-Speak}\{T_\alpha, M_\alpha, R_\alpha\}^r_{N=\alpha} \tag{6}$$

$$H \leftarrow H \cup \{W_\alpha\}^r_{N=\alpha} \tag{7}$$

W_α is the content of player α's final speech in the r round. It will be added to the historical records H as historical information driving the game.

4.3 Phase II: Reflections on the Voting Session

In phase II, the voting part reflects the incomplete information game problem in Game theory [4]. MPTT helps AI agents make strategic voting decisions, addressing the issue of (c) mentioned in Fig. 2.

First-FindTeammate. Players review the history of others' speeches to identify teammates and opponents, comparing and analyzing characteristics in multiple ways. Before each round of voting opens, each player thinks simultaneously:

$$F_\alpha = \text{First-FindTeammate}\{H, O_\alpha\}^r_{N=\alpha} \tag{8}$$

Second-FindTeammate. As the amount of information gradually increases, Players will reassess their identity and update their strategy based on new information:

$$J_\alpha = \text{Second-FindTeammate}\{H, O_\alpha, F_\alpha\}^r_{N=\alpha} \tag{9}$$

Game-Decision. Finally, Players use cumulative reflection and judgement to build more explicit trust, and update O_α to better adapt to the dynamic situation (refer to Eq. 4 and add F_α, J_α in it, as well as update Eq. 5):

$$S_\alpha = \text{Game-Decision}\{H, O_\alpha, F_\alpha, J_\alpha\}^r_{N=\alpha} \tag{10}$$

Players are encouraged to find teammates and fostering cooperation, think strategically in their votes, choosing the right player to vote for to ensure that the results favor their team:

$$V_\alpha = \text{Word-Vote}\{F_\alpha, J_\alpha, S_\alpha\}^r_{N=\alpha} \tag{11}$$

The results of all players' votes are tallied for each round of the game, the player with the highest number of votes will be out of the game.

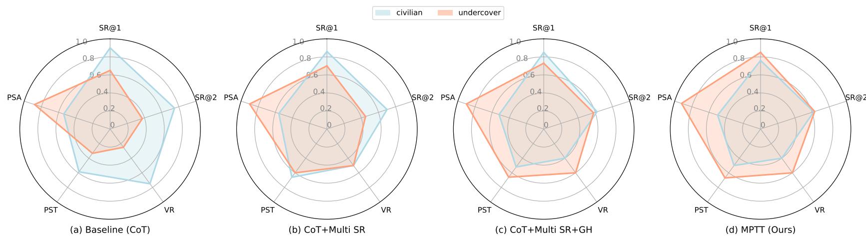

Fig. 3: **Ablation for statistics.** The undercover team is significantly enhanced under MPTT over Baseline, effectively cutting down civilian forces. Ablation studies in (b) and (c) demonstrate the effectiveness of multidimensional self-reflection, global historical speech and self-summary.

5 Experiment

5.1 Baselines and our approach

Setup. We evaluate the capabilities of our proposed Multi-Perspective Team Tactic (MPTT) by having LLM play the full WIU game. Our game is implemented using ChatGPT (gpt-3.5-turbo)[12] for multiple rounds of multi-role-playing, game topics are based on similarity word pairs, difference word pairs, polysemous and ambiguous words, and domain relevant words which were randomly selected from the WIU dataset. In the game phase, we set up 5 LLM agents participating in the game, with 3 civilians and 2 undercovers. The role assignments and speaking order of each game are randomly determined.

Baseline. We stipulate that the Baseline only uses the game's rules as prompts with Chain-of-Thought(CoT) [19]. Gives the results of players' speeches in the current round as a reference in the voting session to playing the game and lets the player think step-by-step.

Multidimensional Self-Reflection. Improved in the Avalon game [17] and Self-Refine [10], we add the multidimensional Self-Reflection on CoT, which allows the player to think more comprehensively.

Gobal History. On top of that, we used the same global history as in the Werewolf [20] to enable the player to see the history of speeches made in the global round before voting.

MPTT (Ours). Based on the above approaches, we also add a history of self-summary, and allows the player to self-reflect several times in multiple perspectives at different phases. The player develops a summary overview at the end of each phase of reflection and keeps it updated, including its current judgment of its self-identity, views on self and team survival tactics. Finally, we get our MPTT.

Metrics. We performed statistical analysis on the experimental data from five perspectives, distinguishing between civilians and undercovers, to assess the performance of MPTT and its ablation experiments. These perspectives include Victory Rate (VR), Survival Rate in the First Round (SR@1), Survival Rate

Table 1: Strengthening of undercover metrics relatively weakens civilian's power.

		SR@1	SR@2	VR	PST	PSA		SR@1	SR@2	VR	PST	PSA
CL↓	Baseline (CoT)	0.90	0.75	0.75	0.59	0.54	UC↑	0.65	0.38	0.25	0.33	0.88
	$\mathrm{CoT}_{+\mathrm{Multi\ SR}}$	0.86	0.70	0.50	0.66	0.56		0.70	0.45	0.50	0.60	0.89
	$\mathrm{CoT}_{+\mathrm{Multi\ SR+GH}}$	0.85	0.62	0.40	0.52	0.52		0.73	0.58	0.60	0.66	0.90
	$\mathrm{MPTT}_{\mathrm{w/o\ PhaseI}}$	0.90	0.73	0.67	0.62	0.55		0.65	0.40	0.33	0.42	0.88
	$\mathrm{MPTT}_{\mathrm{w/o\ PhaseII}}$	0.82	0.68	0.50	0.60	0.54		0.76	0.47	0.50	0.53	0.89
	MPTT (Ours)	**0.76**	**0.62**	**0.40**	**0.50**	**0.50**		**0.85**	**0.63**	**0.60**	**0.67**	**0.92**

CL: Civilian, UC: Undercover.

in Consecutive Two Rounds (SR@2), Probability of Successfully Trusting Own Team (PST), and Probability of Successfully Assessing Enemy Team (PSA).

Quantitative Results. Fig. 3 and Table. 1 shows the performance differences between civilians and undercovers in MPTT and its ablation studies. Due to their majority, civilians in the Baseline quickly recognize teammates and maintain a higher VR. As strategies evolve and perspectives diversify, undercovers leverage their minority status to improve consensus, and locate teammates more efficiently, boosting their VR,they also achieve higher PSA accuracy due to smaller size, though they risk less concentrated voting. MPTT effectively addresses these challenges, enhancing undercover performance. The relative decline in civilian is the result of a dynamic game in which the framework optimizes the balance between strategies, and the boost in undercover is a reflection of the framework's success in minority empowerment.

5.2 Evaluation of game metrics

We have defined metrics in five dimensions to evaluate the performance of LLMs agents participating in games. And compare MPTT with the Baseline to verify its effectiveness.

Metrics. Given the disparate initial populations between the civilian team and the undercover team, the five-dimensional metrics are computed independently for each team. The five-dimensional metrics are Voting Success Rate (VSR), Influence (INF), Comprehension Capability (CCAP), Reversal Rate (REV), and Concealment (CONC).

Analysis of evaluations. Table.2 and shows how the different teams of the Baseline and MPTT performed on the five metrics during the experiment. Due to the numerical superiority of civilians in Baseline, they are stronger than undercovers in several metrics. It's also because the civilian base is larger, undercovers mislead civilians to vote incorrectly at a higher rate on CONC. In MPTT, the undercover team improves on all indicators, leading to a weakening of the civilian team's advantage. This suggests that MPTT is effective in ameliorating the differences caused by team numbers, emphasizing the need for minorities to improve their strategy performance to form a fair competition, and enhance the adversarial nature of the game.

Table 2: MPTT can effectively improve the headcount difference and counterbalance each other.

	Method	REV	CCAP	CONC	INF	VSR		REV	CCAP	CONC	INF	VSR
CL↓	Baseline	0.82	0.74	0.27	0.62	0.56	UC↑	0.05	0.22	0.30	0.50	0.25
	MPTT(Ours)	0.39	0.41	0.25	0.60	0.37		0.48	0.37	0.44	0.51	0.39

CL: Civilian, UC : Undercover

Table 3: Two evaluation methods.The first is a comparison of the ability of the human and agents of the same team in the same game, and the second is a comparison of the ability between Add Human and LLM only in a game.

	MethodI	SUR	CCAP	JCAP	INF	VSR	*MethodII*	CCAP	JCAP	INF	VSR	VR
CL	Human	0.33	**1.00**	**1.00**	0.40	0.60	Add Human	**0.70**	0.95	0.36	**0.60**	**0.67**
	Agent	**0.67**	0.59	0.93	0.34	0.60	LLM only	0.38	0.97	0.35	0.18	0.17
UC	Human	0.33	**0.60**	**0.50**	0.40	0.50	Add Human	**0.43**	0.35	0.39	**0.52**	**0.67**
	Agent	**0.67**	0.30	0.23	0.38	0.54	LLM only	0.09	0.28	0.38	0.10	0.17

CL: Civilian, UC : Undercover

5.3 LLM and human collaborative reasoning

In this section, We explore the impact of integrating a human player into LLM-driven reasoning games, focusing on the differences and similarities between humans and AI. The "Human-in-the-loop" protocol features one human and four LLM agents in a WIU game. To assess the human's impact, we selected games where both teams frequently failed and placed a human in the failing team. We also define Judgment Capability (JCAP) to measure a player's self-judgment accuracy and Survival Rate (SUR) to compare survival outcomes between humans and AI agents on the same team. All metrics are calculated separately for players of the same type (human or AI agent).

Analysis on Human-in-the-loop. Both teams use their frequently failed game topics differently, so we compare the diversity between human and AI agents instead of the gap between teams. The first examines human and AI decision-making within the same team in the "Human-in-the-loop" game. MethodI in Table. 3 shows that humans and AI agents have similar VSR and INF scores, indicating comparable influence. However, humans achieve higher CCAP and JCAP scores, reflecting better judgment in ambiguous situations, while their lower SUR scores suggest vulnerability to being targeted due to language style differences. Overall, humans and AI agents influence each other's thinking and interaction. The second assesses "Human-in-the-loop" vs LLM only in the same game with MPTT. MethodII in Table. 3 shows that adding a human player significantly increase SUR and VR for both teams and balanced overall metrics, highlighting the human's impact. Regardless of their role, human players

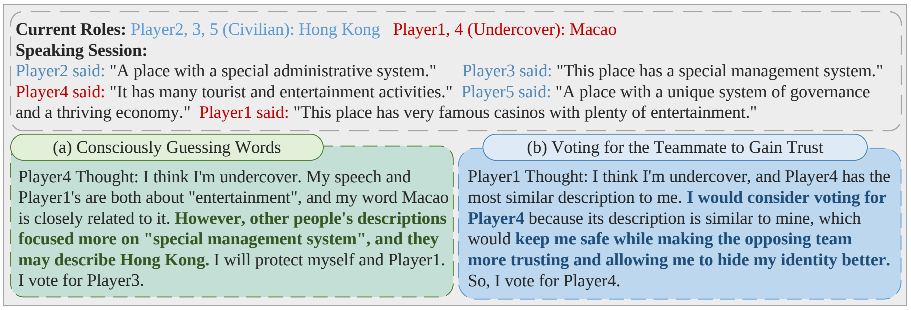

Fig. 4: **Emerging advanced tactics in WIU.** Including (a) Consciously Guessing Words and (b) Voting for the Teammate to Gain Trust.

enhance their team's CCAP and VSR, demonstrating superior analysis and inference abilities, and this is where AI agents need to learn to optimize constantly.

5.4 Advanced tactics and generalization ability

We explore advanced tactics used by AI agents in the WIU game and their impact on game dynamics. Like humans, AI agents attempt to infer others' difference and adjusting strategies. Fig.4 (a) shows it enhances their reasoning and deduction abilities. AI undercovers may strategically vote against a fellow undercover with more exposure, creating confusion and gaining civilian trust, as illustrated in Fig.4 (b). While this tactic can mislead opponents, it also increases the challenge for the AI team, requiring strong acumen and adaptability.

To demonstrate the generalization ability of MPTT, we validated its validity on Claude 3[2] and Gemini[15], both of which performed well in WIU. However, Llama-3-8B[1] didn't fully comply with the required response format, despite exhibiting strategic behaviors like concealment and confrontation. This indicates that open-source LLMs still require improvement in command compliance.

6 Conclusion

We develop a multidimensional thinking framework using the WIU game to iteratively optimize LLM agents' decision-making, extendable to other multi-participant and information asymmetry scenarios.[2] This framework enhances adaptability and information mining through global history, self-summary, and multidimensional thinking, enabling LLM agents to autonomously devise strategies like confrontation and concealment while promoting fairness for minorities. Incorporating human players demonstrates alignment with human behavior, with applications in public welfare, legal aid, and community governance. Additionally, we construct a word network dataset with multidimensional semantic

[2] This work is supported by the GuangDong Basic and Applied Basic Research Foundation (Project 2024A1515011650).

relations for WIU games, which can also be used in other scenarios such as resolving semantic conflicts and revealing implicit intentions. Future research can explore advanced strategies, diverse scenarios, and optimized learning mechanisms to enhance AI's role in social decision-making and human-AI collaboration.

References

1. AI@Meta: Llama 3 model card. AI@Meta (2024)
2. anthropic: Introducing the next generation of claude. anthropic (2024)
3. Chen, C., Shu, K.: Combating misinformation in the age of llms: Opportunities and challenges. arXiv preprint arXiv:2311.05656 (2023)
4. Harsanyi, J.C.: Games with incomplete information. The American Economic Review **85**(3), 291–303 (1995)
5. Huang, J., Gu, S.S., Hou, L., Wu, Y., Wang, X., Yu, H., Han, J.: Large language models can self-improve. arXiv preprint arXiv:2210.11610 (2022)
6. Ichiishi, T.: Game theory for economic analysis. Elsevier (2014)
7. Kroer, C., Sandholm, T.: Limited lookahead in imperfect-information games. Artificial Intelligence **283**, 103218 (2020)
8. Leslie, A.M., Friedman, O., German, T.P.: Core mechanisms in 'theory of mind'. Trends in cognitive sciences **8**(12), 528–533 (2004)
9. Liu, H., Singh, P.: Conceptnet—a practical commonsense reasoning tool-kit. BT technology journal **22**(4), 211–226 (2004)
10. Madaan, A., Tandon, N., Gupta, P., Hallinan, S., Gao, L., Wiegreffe, S., Alon, U., Dziri, N., Prabhumoye, S., Yang, Y., et al.: Self-refine: Iterative refinement with self-feedback. arXiv preprint arXiv:2303.17651 (2023)
11. Miller, G.A.: Wordnet: a lexical database for english. Communications of the ACM **38**(11), 39–41 (1995)
12. OpenAI: Chatgpt: Optimizing language models for dialogue (2022)
13. Shen, Y., Song, K., Tan, X., Li, D., Lu, W., Zhuang, Y.: Hugginggpt: Solving ai tasks with chatgpt and its friends in hugging face. Advances in Neural Information Processing Systems **36** (2024)
14. Tajfel, H., Turner, J.C.: The social identity theory of intergroup behavior. In: Political psychology, pp. 276–293. Psychology Press (2004)
15. Team, G., Anil, R., Borgeaud, S., Wu, Y., Alayrac, J.B., Yu, J., Soricut, R., Schalkwyk, J., Dai, A.M., Hauth, A., et al.: Gemini: a family of highly capable multimodal models. arXiv preprint arXiv:2312.11805 (2023)
16. Telser, L.G.: Competition, collusion, and game theory. Routledge (2017)
17. Wang, S., Liu, C., Zheng, Z., Qi, S., Chen, S., Yang, Q., Zhao, A., Wang, C., Song, S., Huang, G.: Avalon's game of thoughts: Battle against deception through recursive contemplation. arXiv preprint arXiv:2310.01320 (2023)
18. Wang, Z., Zhang, G., Yang, K., Shi, N., Zhou, W., Hao, S., Xiong, G., Li, Y., Sim, M.Y., Chen, X., et al.: Interactive natural language processing. arXiv preprint arXiv:2305.13246 (2023)
19. Wei, J., Wang, X., Schuurmans, D., Bosma, M., Chi, E., Le, Q., Zhou, D.: Chain of thought prompting elicits reasoning in large language models. arXiv preprint arXiv:2201.11903 (2022)
20. Xu, Y., Wang, S., Li, P., Luo, F., Wang, X., Liu, W., Liu, Y.: Exploring large language models for communication games: An empirical study on werewolf. arXiv preprint arXiv:2309.04658 (2023)