Key Decision-Makers in Multi-Agent Debates: Who Holds the Power?

Qian Zhang[1†], Jinyi Liu[2†], Yan Zheng[1,2*], Hebin Liang[2], Lanjun Wang[1]

[1]School of New Media and Communication, Tianjin University, China
[2]College of Intelligence and Computing, Tianjin University, China
{zhangqian0827, jyliu, yanzheng, lianghebin}@tju.edu.cn, wang.lanjun@outlook.com

Abstract

Recent studies on LLM agent scaling have highlighted the potential of Multi-Agent Debate (MAD) to enhance reasoning abilities. However, the critical aspects of role allocation strategies remain underexplored. In this study, we demonstrate that allocating roles with differing viewpoints to specific positions significantly impacts MAD's performance in reasoning tasks. Specifically, we find a novel role allocation strategy, "Truth Last", which can improve MAD performance by up to 22% in reasoning tasks. To address the issue of unknown truth in practical applications, we propose the Multi-Agent Debate Consistency (MADC) strategy, which systematically simulates and optimizes the core mechanisms of MAD. MADC incorporates path consistency to assess agreement among independent roles, simulating the role with the highest consistency score as the truth. We validate MADC across a range of LLMs (9 models), including the DeepSeek-R1 Distilled Models, on challenging reasoning tasks. MADC consistently demonstrated strong performance, effectively overcoming MAD's performance bottlenecks and providing a crucial pathway for further improvements in LLM agent scaling.

Code — https://github.com/SG-XM/AAAI2026-MADC

Introduction

Although Large Language Model (LLM) agents have shown remarkable performance in logical reasoning (Shinn et al. 2023; DeepSeek-AI et al. 2024), natural language processing tasks and specific domains (Liu et al. 2025; Xiao et al. 2024; Chen et al. 2024b; Reddy et al. 2024), they struggle with more complex tasks (Gou et al. 2024; Chen et al. 2024a). To further improve the performance, numerous multi-agent frameworks have been proposed, such as CAMEL (Li et al. 2023), LLM-Blender (Jiang, Ren, and Lin 2023), CMD (Wang et al. 2024), Debate (Liang et al. 2024), CIPHER (Pham et al. 2024), Socialized Learning (Yao et al. 2024), RECONCILE (Chen, Saha, and Bansal 2024). In particular, several works have studied scaling laws for language models and multi-agent systems (Kaplan et al. 2020; Wang et al. 2025).

[†]These authors contributed equally to this work
[*]Corresponding to: Yan Zheng

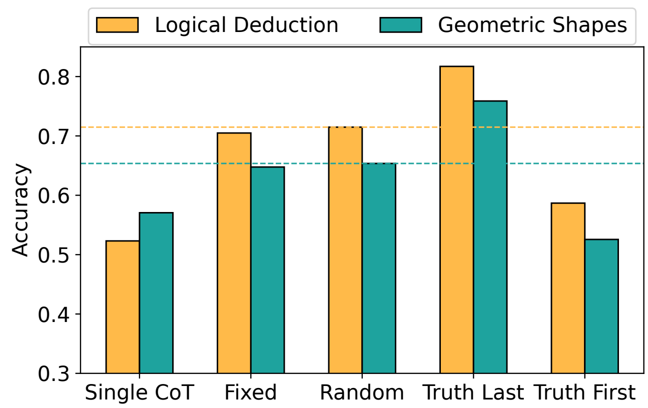

Figure 1: Comparative performance of single-agent and different MAD allocation strategies using the Qwen2.5-7B-Instruct model on BBH's Logical Deduction and Geometric Shapes tasks.

Recent studies explore multi-agent scaling principles. Du et al. (2023) proposed a MAD framework, showing performance gains with more agents and debate rounds. Agent Forest (Li et al. 2024a) scales agents to boost LLM results, aligned with task difficulty. Traditionally, some approaches in Multi-Agent System (MAS) have leveraged reinforcement learning (RL) to enhance performance (Yang et al. 2021, 2025; Zheng et al. 2019), while recent work has proposed post-training methods for LLM-driven MAS (Cheng et al. 2025).

However, stable scaling faces challenges: Wang et al. (2024) found strong single-agent prompts match top multi-agent methods on many reasoning tasks; Smit et al. (2024) noted multi-agent debating systems do not consistently outperform strategies like self-consistency (SC), though some MAD systems improve with hyperparameter tuning. Li et al. (2024b) showed more reference solutions aid easier questions but mislead on harder ones, reducing accuracy. These findings reveal that MA scaling can be unstable due to sensitivity to hyperparameters and task complexity.

Figure 1 presents a comparison of the performance of Single-Agent and Multi-Agent Debate frameworks with varying role allocation strategies. In this figure, 'Single' rep-

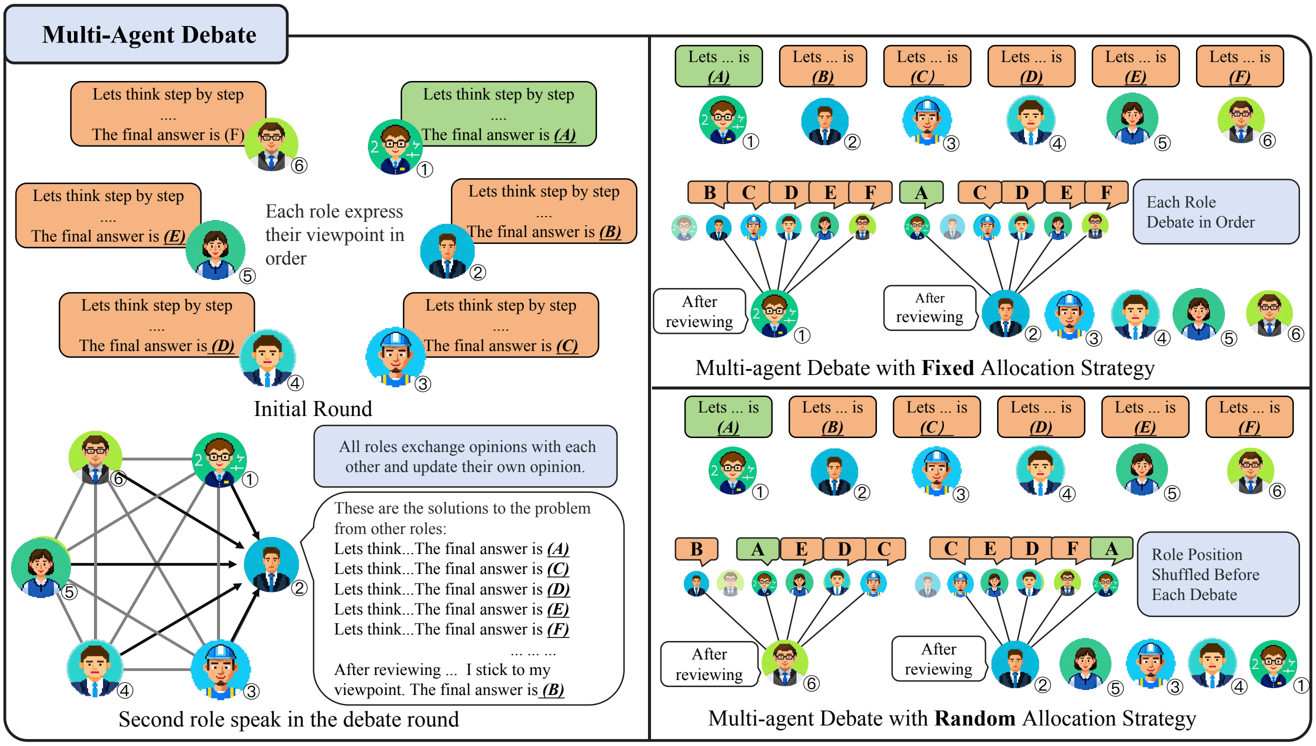

Figure 2: In the initial round of the MAD framework, each role independently uses CoT. During the debate round, roles exchange viewpoints in a fully connected manner to update their viewpoints. In the default Fixed strategy, roles speak in a consistent order each round and receive others' viewpoints in a fixed sequence (A)-(E). The Random strategy disrupts both the speaking order and the relative positions during the debate. Green represents correct viewpoints, while orange indicates incorrect viewpoints.

resents the performance of a single agent using Chain of Thought (CoT) (Wei et al. 2022), while 'MAD' refers to the framework proposed by Du et al. (2023). In this context, 'Fixed' indicates roles placed in predetermined positions, 'Random' denotes roles arranged in a random order during the debate, 'Truth Last' refers to arranging agents with correct viewpoints at the end, while 'Truth First' does the opposite (see Appendix A.1.7 for details). Our observations reveal that, with the Qwen2.5-7B-Instruct model, the existing fixed role allocation strategy outperforms the single agent, and there are significant differences between various role allocation strategies.

Based on observations from Figure 1, we conducted systematic experiments for in-depth analysis. Although many MA discussion, debate, and collaboration frameworks exist, they often overlook role allocation strategies. For example, in Du's MAD framework, role positions are fixed in the implementation shown in Figure 2.

We found that when role allocation is more ordered by viewpoints, debate results show lower entropy, without impacting the accuracy of conclusions. Our results also demonstrate that using a white-box allocation strategy "Truth Last" significantly enhances accuracy in reasoning tasks, and this improvement is orthogonal to existing MAD frameworks, requiring no changes to context or prompts. Then we propose a Multi-Agent Debate Consistency allocation strategy to simulate the "Truth Last". This method effectively enhances MAD performance and efficiency. Additionally, we discovered that under ideal conditions, only half the number of agents is needed to reverse a debate outcome, further deepening our understanding of role allocation strategies in scaling laws.

In summary, our contributions are:

(1) Key Insight: We identify role allocation strategy as a novel scaling dimension for MAD frameworks, revealing that later-positioned roles exert greater influence on final decisions through analysis of debate entropy and positional bias.

(2) Methodological Innovation: We propose the Multi-Agent Debate Consistency (MADC) allocation strategy, a novel approach that optimizes role arrangement without modifying prompts or context, effectively unlocking MAD's potential. This method is completely **orthogonal** to existing MA frameworks.

(3) Empirical Validation: We evaluated 9 different LLM models, including open-source models like Qwen2.5, DeepSeek-R1-Distill, and closed-source models like GPT-4o-mini. Our comprehensive tests across three diverse tasks, logical deduction, geometric shapes, and MATH500, which highlight the broad applicability and robustness of our findings.

Preliminary

To study the impact of allocation strategy on the multi-agent (MA) framework, we need to control other influencing factors. We must select the most representative parameter settings to ensure our findings are comprehensive and applicable.

Factors Affecting the Performance of MAD

In MAD, the factors affecting performance mainly include the following aspects.

Communication Strategy: Existing work explores various strategies (see Appendix A.1.7 for details) such as Talk One-by-One, Simultaneous-Talk, Simultaneous-Talk-with-Summarizer (Chan et al. 2024), exchanging natural language, and exchanging raw transformer output embeddings (Pham et al. 2024). In our study, we choose the basic Simultaneous-Talk strategy and exchange natural language.

Topology and Connectivity: Another key factor is the topology and connectivity among agents. Some studies propose group discussion structures (Wang et al. 2024), tree structures, and dynamically adjusting the topology by removing agents during discussions (Liu et al. 2024). In our work, we choose the traditional method of full connectivity, where all agents communicate with each other.

Agents Profiling: Profiling prompts are also a critical factor influencing multi-agent systems (Hong et al. 2023; Li, Yang, and Zhao 2023; Chen et al. 2023; Park et al. 2023, 2024). Common prompts settings include collaborative (Zhang et al. 2024; Chen, Fried, and Topcu 2024), adversarial (Chan et al. 2024) and judgment (Wang et al. 2024) prompts. In our work, we do not use special profiling prompts in order to eliminate their influence.

Role Allocation Strategies: Based on the settings mentioned above, we conduct experiments within the MAD framework (Du et al. 2023). Within the MAD framework, we define four role allocation strategies for the debate process. **(1) Fixed**: The position of roles is fixed once allocated. **(2) Random**: Roles are randomly allocated before each debate. **(3) Truth First**: Roles with the correct answer are allocated at the beginning. **(4) Truth Last**: Roles with the correct answer are allocated at the end.

Formulation of MAD Scaling

In the MAD framework, we define n agents $A_0, A_1, A_2, \ldots, A_{n-1}$, each generating a viewpoint $V_{i,j}$ in every round of debate, where i denotes the agent index and j represents the debate round.

Figure 2 illustrates the MAD framework. During each round of discussion and debate, the agents maintain fixed positions throughout. For instance, in the first round, A_0 to A_5 express V as follows: (A), (A), (A), (B), (C), (D). In subsequent rounds, the agent with viewpoint D consistently follows the agent with viewpoint C during discussions and debates. In a multi-persona context, this means that the agent corresponding to $role_5$ always follows $role_4$. Practically, this could result in a critic always speaking after an advisor, or a programmer consistently following a product manager.

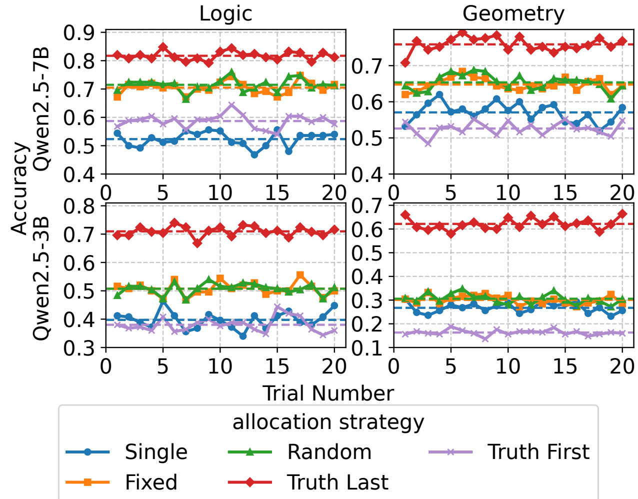

Figure 3: Experimental results showing the accuracy metrics across different allocation strategies, with experiments repeated 20 times.

Methodology

Does Role Allocation Strategy Impact MAD?

The conclusions from Figure 1 are clear but not comprehensive or robust enough. Firstly, we propose a null hypothesis, which we will attempt to falsify:

Null Hypothesis 1 (Non-impactfulness, H_0)**.** *The position of roles has no impact on the final outcome in a debate.*

To falsify H_0, we must demonstrate that there exists at least one allocation strategy, S, such that the outcome using S differs significantly from outcomes over Fixed or Random strategy with p-value < 0.001.

Challenging the Null Hypothesis: Impactful Allocation Strategies

To find sufficient evidence to reject H_0, we evaluated the accuracy metric of MAD after one round of debate under different allocation strategies. Additionally, we examined the log-likelihood (See Appendix A.1.1 for details.) and entropy (See Appendix A.1.2 for details.) metrics. Accuracy represents the number of correctly answered questions, while log-likelihood indicates the number of roles with correct answers.

To test this, we conducted 20 repeated experiments using four allocation strategies outlined in preliminary, focusing on Logical Deduction and Geometric Shapes tasks. The results are shown in Figure 3.

We observed no significant difference between the random and fixed strategies in the experiments, which is reasonable since the fixed strategy can be seen as a single sample of the random strategy. Noting that the Truth Last strategy outperformed others significantly with p-value <0.0001.

Figures 3 and 4 demonstrate that the Truth Last strategy significantly outperforms other strategies, which disproves

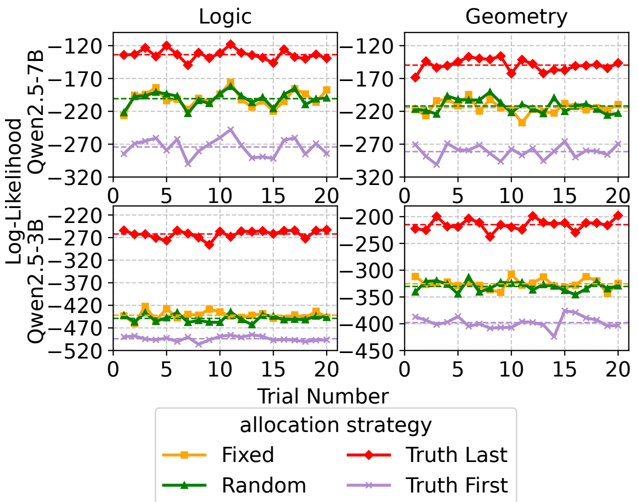

Figure 4: Experimental results showing the log-likelihood metrics across different allocation strategies, with experiments repeated 20 times.

the H_0. To further challenge the H_0, we examine how allocation strategies affect viewpoint convergence. Lower entropy reflects higher consensus (See Appendix A.1.2 for details.), suggesting that role allocation strategy affects debate process and thus challenges the H_0.

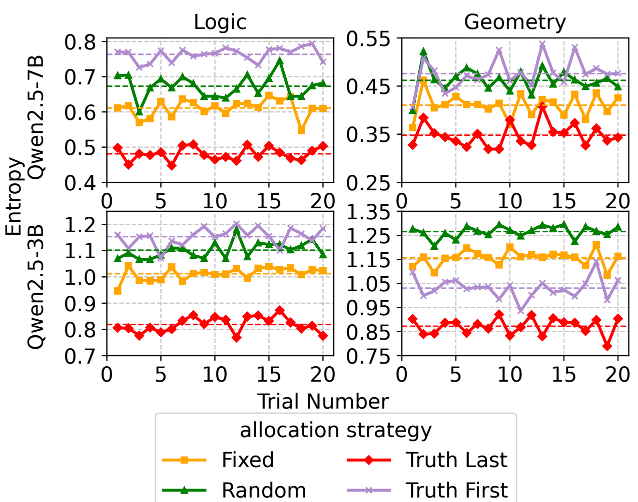

Figure 5: Experimental results showing the entropy metrics across different allocation strategies, with experiments repeated 20 times.

Figure 5 presents the entropy metrics for various strategies, showing interesting results: the Truth Last strategy has the lowest entropy, followed by fixed, and then random. Truth First, although resulting in lower accuracy, makes it easier for debates to reach consensus, showing lower entropy. For example, on the geometry dataset, the 3B model with Truth First exhibited lower entropy compared to random and fixed strategies. Despite similar accuracy and log-likelihood metrics, random and fixed strategies show significant differences in entropy, further challenging the null hypothesis. Additionally, the Truth Last can be seen as a sorted version of fixed, demonstrating that entropy decreases as the orderliness of agent allocation increases.

Takeaway:
(1) Allocating the role with the correct viewpoint last can enhance the effectiveness of MAD.
(2) An orderly allocation of roles is more conducive to reach an agreement in the debate.

Multi-Agent Debate Consistency Strategy

In previous experiments, we demonstrated the importance of allocation strategies in MAD settings and found that Truth Last consistently yields better results. However, in practical applications, the truth is unknown. To address this, we propose a method called MADC to approximate the truth using confidence levels.

Path Consistency. After each round, agents may arrive at an intermediate result $V_{i,j}$, with each result serving as a candidate for final decision-making.

Definition 1 (Debating Path)**.** *A Path P_i is the sequence of viewpoints or intermediate results generated by an agent A_i during the debate process. Formally, it is represented as:*

$$P_i = (V_{i,1}, V_{i,2}, \dots, V_{i,m}), \tag{1}$$

where $V_{i,j}$ is the viewpoint of agent A_i at round j, and m is the total number of rounds.

For $V_{i,j}$, when $j < m$, $V_{i,j}$ is considered an intermediate result, and when $j = m$, it is the final result. For any $V_{i,j}$, there exists a corresponding intermediate path $P_{i,j} = (V_{i,1}, V_{i,2}, \dots, V_{i,j})$, representing the sequence of results up to that round.

Definition 2 (Path Consistency)**.** *Path Consistency measures how likely different paths produce the same result at a specific round j. Consider paths $P_{1,j}, P_{2,j}, \dots, P_{n,j}$. These are followed by agents $A_1, A_2, \dots, A_n$. Path Consistency is the probability that their results $V_{1,j}, V_{2,j}, \dots, V_{n,j}$ are equal. This occurs at round j, where $j < m$. We define it as:*

$$\textit{Consistency}(\mathcal{P}_{i,j}) = \sum_{k \neq i} \Pr(V_{i,j} = V_{k,j}). \tag{2}$$

Here, $V_{i,j}$ is the result of agent A_i at round j.

To approximate the ground truth, we use path consistency to enhance result accuracy. For each reasoning path $P_{i,j}$, we calculate the consistency of its result $V_{i,j}$ at round j as $\text{Consistency}(P_{i,j})$. The result corresponding to the path with the highest path consistency score is selected.:

$$V^* = \arg\max_{V_{i,j}} \text{Consistency}(P_{i,j}). \tag{3}$$

Path Consistency indicates the level of agreement among different agents' debating paths at a specific round. When

Algorithm 1: Multi-Agent Debate Consistency

Require: Agent set $\mathcal{A} = \{A_1, ..., A_n\}$ with initial paths $\{P_{i,1}\}_{i=1}^{n}$, total rounds m
Ensure: Consistency ordered list $\mathcal{L}$
1: Initialize path matrix $\mathbf{P}^1 \leftarrow [P_{1,1} \cdots P_{n,1}]^\top$ ▷ $P_{i,j}$: path of A_i at round j
2: Initialize consistency list $\mathcal{L}^1 \leftarrow \emptyset$
3: **for** round $j \leftarrow 1$ **to** $m-1$ **do** ▷ Consistency evaluation phase
4: Compute path consistency:
5: $\Phi(P_{i,j}) \leftarrow \sum_{k \neq i} \Pr(V_{i,j} = V_{k,j})$ ▷ According to Eq. 2
6: Identify max-consistency agent:
7: $A^* \leftarrow \arg\max_{A_i \in \mathcal{A}} \Phi(P_{i,j})$
8: Update debate ordering:
9: $\mathcal{L}_{\text{temp}} \leftarrow \text{argsort}(\{\Phi(P_{k,j}) | A_k \in \mathcal{A} \setminus A^*\})$
10: $\mathcal{L}^j \leftarrow (\mathcal{L}_{temp} \circ A^*)$
11: Propagate consistency:
12: $\mathbf{P}^{j+1} \leftarrow \text{Debate}(\mathbf{P}^j, \mathcal{L}^j)$
13: **end for**

the consistency is high, it suggests that multiple independent agents have converged on the same result. This convergence implies robustness and reliability, as it reduces the likelihood of individual biases or errors impacting the outcome.

Debating Strategy Guided by Path Consistency. Multi-Agent Debate Consistency (MADC) is designed to enhance the adaptability and effectiveness of multi-agent systems by leveraging consistency metrics during debates. One of the key advantages of MADC is its **orthogonality** to other multi-agent methods, meaning it can be integrated without altering existing prompts or context.

Algorithm 1 outlines the MADC process in a structured manner. Initially, consistency scores are calculated for each round using Equation 2 (Lines 5-6). Agents are then reordered based on these scores (Lines 8-9). This reordering allows all agents to update their paths effectively (Line 11).

MADC integrates seamlessly with existing systems through two key features. Its plug-in design introduces dynamic allocation without altering the internal mechanisms of the agents. Additionally, MADC promotes self-improvement by ensuring that more reliable viewpoints have a greater influence, thereby enhancing decision-making without modifying the existing framework.

Experiments

To validate the effectiveness of our method and gain deeper insights into how allocation strategies impact the debate process and outcomes, we focus on the following research questions:

RQ1 (Applicability and Potential): Is MADC robust enough across various models and tasks?

RQ2 (Scaling Impact): Is the allocation strategy still crucial in round and agent count scaling?

RQ3 (Quantitative Analysis): Can just half the agents reverse a debate outcome?

Experiment Setup

Language Model. To validate the MADC strategy, we tested multiple models (OpenAI 2023; Bai et al. 2023; Qwen et al. 2025; Zeng et al. 2022): GPT, Qwen GLM, DeepSeek. These models include both open-source and closed-source options. For closed-source models, we conducted experiments using API calls, covering both free and paid versions. And we adopted very simple prompt settings (See Appendix A.1.4 for details.).

Tasks and Dataset. MATH500 (HuggingFaceH4 2024): A specialized dataset of 500 diverse mathematical problems.

Logical Deduction: A subtask of Big-Bench Hard (BBH) (Suzgun et al. 2022).

Geometric Shapes: Another BBH subtask.

Settings (RQ3): Evaluating the Agent Threshold for Reversing Debate Outcomes. We constructed a specific scenario shown as Figure 6. We constructed two groups of agents: 40 with correct viewpoints and 40 with incorrect viewpoints. In the most extreme scenario, we first allocated the 40 agents with correct viewpoints, then gradually added agents with incorrect viewpoints. Each configuration was tested in 10 debate repetitions to calculate accuracy under the current setup, and the entire experiment was repeated 5 times. For comparison, we implemented three scaling methods:

Correct Viewpoint First: Scaling agents with correct viewpoints first.

Incorrect Viewpoint First: Scaling agents with incorrect viewpoints first.

Alternating Scale: Alternately scaling agents with correct and incorrect viewpoints.

When constructing the agents, we used Qwen2.5-7B-Instruct to perform chain-of-thought (CoT) sampling 400 times for each question. This ensured a sufficient number of agents with diverse viewpoints.

Performance of Role Allocation Strategy (RQ1)

As shown in Table **??**, our MADC strategy achieves consistent improvements across major model families. For **Qwen** series, MADC enhances accuracy by +8.8% (3B) and +3.6% (7B) on Logical Deduction, with **DeepSeek-Qwen-7B** showing +9.6% improvement on Geometric Shapes. The **GLM** family attains +3.6% (Flash) and +1.2% (FlashX) gains on logic tasks, while **GPT-4o-mini** maintains +2.0% improvement.

To explore the potential for improvement in MADC, we implement the Truth Last strategy as an idealized reference. Experimental results show significant room for improvement: Truth Last achieves remarkable gains of 22.8% on both Logical Deduction and Geometric Shapes for Qwen-2.5-3B-Instruct, and 24.4% on Geometric Shapes for DeepSeek-R1-Qwen-7B. Even high-performance models like GPT-4o-mini show additional improvements of 3.2%–4.0% through this strategy. Notably, the DeepSeek-R1-Qwen-14B model reaches 99.6% accuracy on Logical Deduction under Truth Last, highlighting the significant room for improvement in black-box implementations.

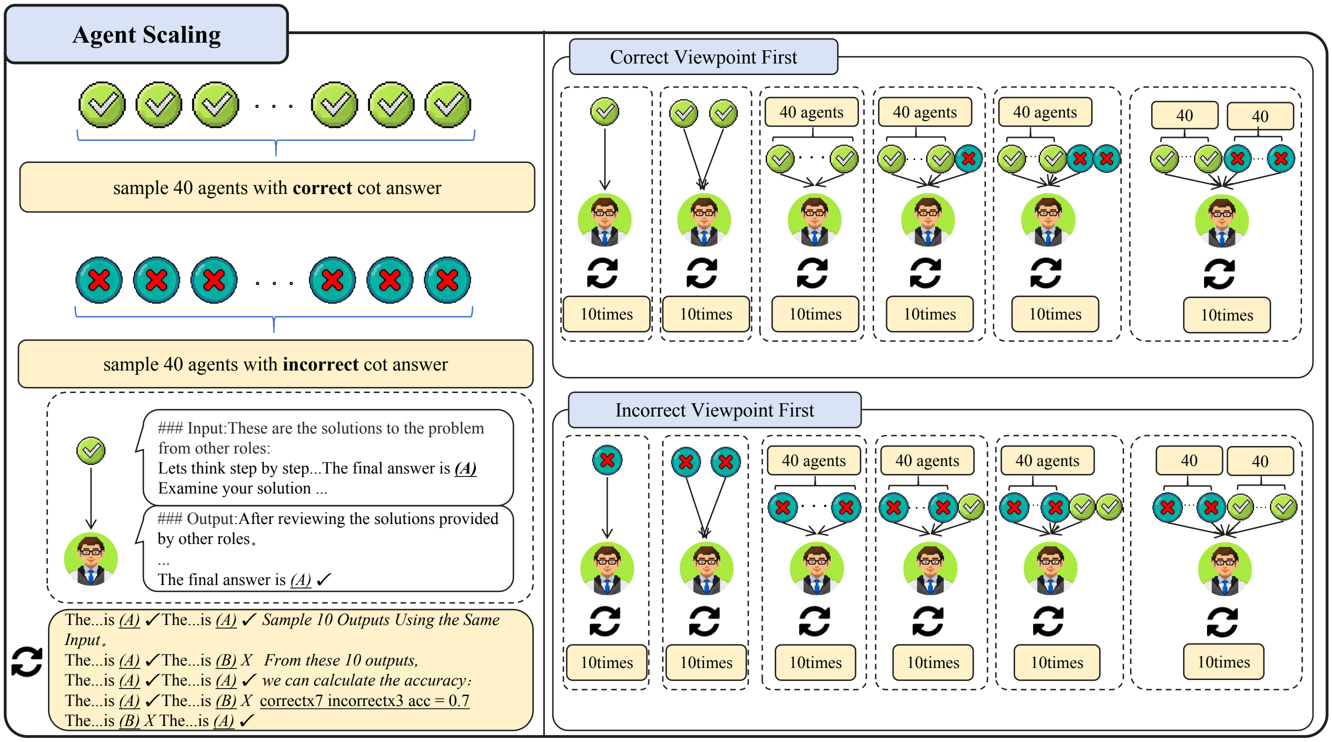

Figure 6: For each question, we resample to obtain 40 agents with correct viewpoints and 40 with incorrect viewpoints. In the debate round, we fix the context and conduct 10 repetition for each number of agents, calculating the accuracy. We apply two scaling strategies. The entire experiment is repeated five times to obtain the average value in scaling.

Scaling Experiment (RQ2)

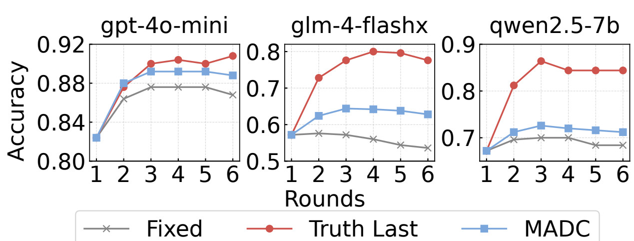

Figure 7: Performance of round scaling.

In the context of MAD frameworks, understanding debate round scaling is crucial to validate our conclusions and methods. We examined whether allocation strategies remain effective and consistent as the number of debate rounds increases. Figure 7 shows that as debate rounds scale, the ”Truth Last” strategy significantly outperforms the standard MAD approach. Additionally, the MADC method remains effective as the round scale increases. At the same time, we also validated the performance of MADC as the number of agents scales. The experimental results indicate that MADC remains effective for different numbers of agents, and MADC‘s effectiveness gradually becomes more significant than MAD as the number of agents increases. (See Appendix A.1.3 for details.)

Experiment on Agent Requirements for Reversing Debate Outcome (RQ3)

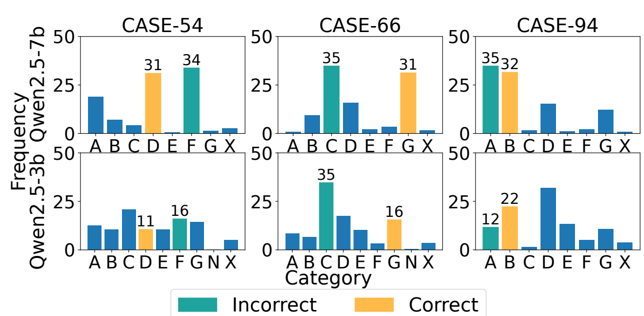

Figure 8: We selected cases whose correct and incorrect viewpoints have similar distributions in the 7B model. However, for the same questions in the 3B model, the distributions differ.

As discussed on the experimental setup section, we sampled each question 400 times. For the 7B model, we selected three cases with a similar distribution of correct and incorrect viewpoints, which is shown in Figure 8.

We repeated the same experiment using Qwen-2.5-3B-Instruct. It is important to note that, in the 3B model, the distribution of correct and incorrect viewpoints was not sim-

Strategy	Model	Logic	Geometry	MATH	Model	Logic	Geometry	MATH
Single CoT	GPT-4o-mini	70.0	63.2	69.8	Qwen2.5-3B-Instruct	42.0	27.2	61.4
SC		82.4	66.4	76.2		54.0	28.8	69.2
MAD		87.2	72.2	76.4		41.2	25.2	69.6
MAD+Truth		90.4 ↑3.2	**76.0** ↑4.0	76.4 - -		**64.0** ↑22.8	**48.0** ↑22.8	76.2 ↑6.6
MAD+MADC		**89.2** ↑2.0	73.2 ↑1.2	76.4 - -		**50.0** ↑8.8	28.4 ↑3.2	71.0 ↑1.4
Single CoT	GLM-4-Flash	46.0	41.2	43.2	Qwen2.5-7B-Instruct	52.4	57.0	67.8
SC		60.8	50.0	54.6		60.2	64.0	75.8
MAD		58.4	50.8	51.2		70.4	64.8	74.4
MAD+Truth		**72.8** ↑14.4	62.4 ↑11.6	53.2 ↑1.6		**81.8** ↑11.4	76.0 ↑11.2	80.4 ↑6.0
MAD+MADC		**62.0** ↑3.6	52.4 ↑1.2	52.2 ↑1.0		72.0 ↑1.6	**68.4** ↑3.6	76.8 ↑2.4
Single CoT	GLM-4-FashX	42.4	42.4	42.8	qwen-turbo	71.2	53.2	73.2
SC		57.2	50.8	53.8		82.8	55.6	76.8
MAD		60.4	49.2	53.4		82.4	58.4	78.2
MAD+Truth		70.0 ↑9.6	**62.0** ↑12.8	55.2 ↑1.8		**91.6** ↑9.2	64.8 ↑6.4	80.2 ↑2.0
MAD+MADC		61.6↑1.2	**51.2** ↑2.0	55.0 ↑1.6		83.2 ↑0.8	**60.4** ↑2.0	78.2 - -
Single CoT	DeepSeek-R1-Distill-Qwen-7B	61.2	62.4	86.2	qwen-plus	84.8	52.4	77.0
SC		86.8	74.0	90.6		92.0	54.8	80.0
MAD		76.0	65.2	87.2		94.0	65.2	79.8
MAD+Truth		89.6 ↑13.6	**89.6** ↑24.4	91.8 ↑4.6		94.0 ↑0.0	**69.6** ↑4.4	80.8 ↑1.0
MAD+MADC		85.2↑9.2	**74.8** ↑9.6	90.2 ↑3.0		94.4 ↑0.4	66.0 ↑0.8	**80.8** ↑1.0
Single CoT	DeepSeek-R1-Distill-Qwen-14B	92.4	75.2	87.2				
SC		99.2	79.2	90.8				
MAD		99.6	79.6	90.2				
MAD+Truth		99.6 - -	**81.6** ↑2.0	92.2 ↑2.0				
MAD+MADC		99.6 - -	**80.4** ↑0.8	90.8 ↑0.6				

Table 1: Performance of Role Allocation Strategy (10 agents 1 debate round)

ilar for these cases.

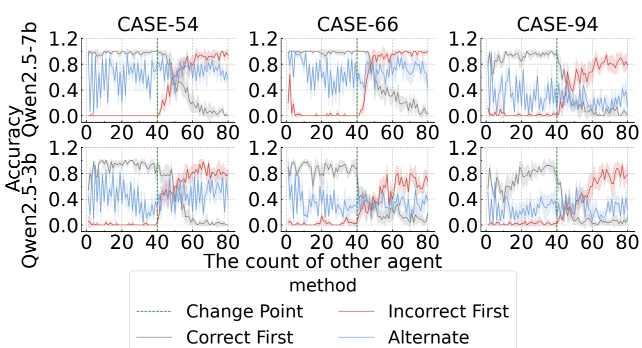

Figure 9: Accuracy variation curves of agent count scaling.

Figure 9 presents the experimental results for different scaling methods. From this, we can conclude the following:

(1) Allocation strategies significantly impact the MAD scaling law. This is evident as the different strategies do not converge to the same point when scaled to 80 agents.

(2) Agents positioned later have a greater influence on the debate outcome. Given the same context, the results tend to favor the viewpoints of agents speaking later, regardless of whether their viewpoints are correct.

(3) Reversing a conclusion is considerably easier than upholding one. With just 20 agents, we can achieve a 50% change, even with a base of 40 agents. (See Appendix A.1.6 for details.)

When examining the ‘alternating scale method’, we observe that it consistently falls between the other two methods as the number increases, aligning with our intuition. However, the debate outcomes do not stabilize as the number of agents increases. Additionally, there is significant fluctuation when the agent count is low, indicating that ‘alternating’ is a highly suboptimal allocation strategy. Furthermore, even though these viewpoints have different distributions in the Qwen2.5-3B-Instruct, the conclusions we previously reached still hold.

Conclusion

In our study, we examined the impact of role allocation strategies within the multi-agent framework and scaling laws, thoroughly validating our conclusions. Through experiments on 9 models, we discovered that role allocation is a critical factor influencing scaling laws in MA. Allocating the role with the correct viewpoint last enhances MAD effectiveness, and an orderly role arrangement facilitates reaching agreement in debates. We propose the MADC method, an orthogonal approach that significantly improves MAD outcomes and is widely applicable within the multi-agent framework.

Acknowledgements

This work is supported by the Key Program of the National Natural Science Foundation of China (Grant No. 62533021), the National Key Research and Development Program of China (Grant No. 2024YFE0210900), the National Natural Science Foundation of China (Grant No. 62572346) and Xiaomi Young Talents Program of Xiaomi Foundation. We would like to thank all the anonymous reviewers for their valuable comments and constructive suggestions, which have greatly improved the quality of this paper.

References

Bai, J.; Bai, S.; Chu, Y.; and et al. 2023. Qwen Technical Report. https://arxiv.org/abs/2309.16609.

Chan, C.-M.; Chen, W.; Su, Y.; Yu, J.; Xue, W.; Zhang, S.; Fu, J.; and Liu, Z. 2024. ChatEval: Towards Better LLM-based Evaluators through Multi-Agent Debate. In *The Twelfth International Conference on Learning Representations*.

Chen, G.; Dong, S.; Shu, Y.; Zhang, G.; Sesay, J.; Karlsson, B. F.; Fu, J.; and Shi, Y. 2023. Autoagents: A framework for automatic agent generation. *arXiv preprint arXiv:2309.17288*.

Chen, J.; Saha, S.; and Bansal, M. 2024. ReConcile: Round-Table Conference Improves Reasoning via Consensus among Diverse LLMs. In Ku, L.-W.; Martins, A.; and Srikumar, V., eds., *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 7066–7085. Bangkok, Thailand: Association for Computational Linguistics.

Chen, S.; Fried, D.; and Topcu, U. 2024. Human-agent cooperation in games under incomplete information through natural language communication. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, IJCAI '24. ISBN 978-1-956792-04-1.

Chen, W.; Su, Y.; Zuo, J.; Yang, C.; Yuan, C.; Chan, C.-M.; Yu, H.; Lu, Y.; Hung, Y.-H.; Qian, C.; Qin, Y.; Cong, X.; Xie, R.; Liu, Z.; Sun, M.; and Zhou, J. 2024a. AgentVerse: Facilitating Multi-Agent Collaboration and Exploring Emergent Behaviors. In *The Twelfth International Conference on Learning Representations*.

Chen, Y.; Yuan, Y.; Zhang, Z.; ZHENG, Y.; Liu, J.; Ni, F.; Jianye, H.; Mao, H.; and Zhang, F. 2024b. SheetAgent: towards a generalist agent for spreadsheet reasoning and manipulation via large language models. In *THE WEB CONFERENCE 2025*.

DeepSeek-AI; Liu, A.; Feng, B.; Xue, B.; Wang, B.; Wu, B.; Lu, C.; Zhao, C.; Deng, C.; Zhang, C.; Ruan, C.; Dai, D.; Guo, D.; Yang, D.; Chen, D.; Ji, D.; Li, E.; Lin, F.; Dai, F.; Luo, F.; Hao, G.; Chen, G.; Li, G.; Zhang, H.; Bao, H.; Xu, H.; Wang, H.; Zhang, H.; Ding, H.; Xin, H.; Gao, H.; Li, H.; Qu, H.; Cai, J. L.; Liang, J.; Guo, J.; Ni, J.; Li, J.; Wang, J.; Chen, J.; Chen, J.; Yuan, J.; Qiu, J.; Li, J.; Song, J.; Dong, K.; Hu, K.; Gao, K.; Guan, K.; Huang, K.; Yu, K.; Wang, L.; Zhang, L.; Xu, L.; Xia, L.; Zhao, L.; Wang, L.; Zhang, L.; Li, M.; Wang, M.; Zhang, M.; Zhang, M.; Tang, M.; Li, M.; Tian, N.; Huang, P.; Wang, P.; Zhang, P.; Wang, Q.; Zhu, Q.; Chen, Q.; Du, Q.; Chen, R. J.; Jin, R. L.; Ge, R.; Zhang, R.; Pan, R.; Wang, R.; Xu, R.; Zhang, R.; Chen, R.; Li, S. S.; Lu, S.; Zhou, S.; Chen, S.; Wu, S.; Ye, S.; Ye, S.; Ma, S.; Wang, S.; Zhou, S.; Yu, S.; Zhou, S.; Pan, S.; Wang, T.; Yun, T.; Pei, T.; Sun, T.; Xiao, W. L.; Zeng, W.; Zhao, W.; An, W.; Liu, W.; Liang, W.; Gao, W.; Yu, W.; Zhang, W.; Li, X. Q.; Jin, X.; Wang, X.; Bi, X.; Liu, X.; Wang, X.; Shen, X.; Chen, X.; Zhang, X.; Chen, X.; Nie, X.; Sun, X.; Wang, X.; Cheng, X.; Liu, X.; Xie, X.; Liu, X.; Yu, X.; Song, X.; Shan, X.; Zhou, X.; Yang, X.; Li, X.; Su, X.; Lin, X.; Li, Y. K.; Wang, Y. Q.; Wei, Y. X.; Zhu, Y. X.; Zhang, Y.; Xu, Y.; Xu, Y.; Huang, Y.; Li, Y.; Zhao, Y.; Sun, Y.; Li, Y.; Wang, Y.; Yu, Y.; Zheng, Y.; Zhang, Y.; Shi, Y.; Xiong, Y.; He, Y.; Tang, Y.; Piao, Y.; Wang, Y.; Tan, Y.; Ma, Y.; Liu, Y.; Guo, Y.; Wu, Y.; Ou, Y.; Zhu, Y.; Wang, Y.; Gong, Y.; Zou, Y.; He, Y.; Zha, Y.; Xiong, Y.; Ma, Y.; Yan, Y.; Luo, Y.; You, Y.; Liu, Y.; Zhou, Y.; Wu, Z. F.; Ren, Z. Z.; Ren, Z.; Sha, Z.; Fu, Z.; Xu, Z.; Huang, Z.; Zhang, Z.; Xie, Z.; Zhang, Z.; Hao, Z.; Gou, Z.; Ma, Z.; Yan, Z.; Shao, Z.; Xu, Z.; Wu, Z.; Zhang, Z.; Li, Z.; Gu, Z.; Zhu, Z.; Liu, Z.; Li, Z.; Xie, Z.; Song, Z.; Gao, Z.; and Pan, Z. 2024. DeepSeek-V3 Technical Report. arXiv:2412.19437.

Du, Y.; Li, S.; Torralba, A.; Tenenbaum, J. B.; and Mordatch, I. 2023. Improving factuality and reasoning in language models through multiagent debate. *arXiv preprint arXiv:2305.14325*.

Gou, Z.; Shao, Z.; Gong, Y.; yelong shen; Yang, Y.; Duan, N.; and Chen, W. 2024. CRITIC: Large Language Models Can Self-Correct with Tool-Interactive Critiquing. In *The Twelfth International Conference on Learning Representations*.

Hong, S.; Zheng, X.; Chen, J.; Cheng, Y.; Wang, J.; Zhang, C.; Wang, Z.; Yau, S. K. S.; Lin, Z.; Zhou, L.; et al. 2023. Metagpt: Meta programming for multi-agent collaborative framework. *arXiv preprint arXiv:2308.00352*.

HuggingFaceH4. 2024. MATH-500 Dataset.

Jiang, D.; Ren, X.; and Lin, B. Y. 2023. LLM-Blender: Ensembling Large Language Models with Pairwise Ranking and Generative Fusion. In Rogers, A.; Boyd-Graber, J.; and Okazaki, N., eds., *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 14165–14178. Toronto, Canada: Association for Computational Linguistics.

Kaplan, J.; McCandlish, S.; Henighan, T.; Brown, T. B.; Chess, B.; Child, R.; Gray, S.; Radford, A.; Wu, J.; and Amodei, D. 2020. Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*.

Li, G.; Hammoud, H. A. A. K.; Itani, H.; Khizbullin, D.; and Ghanem, B. 2023. CAMEL: Communicative Agents for "Mind" Exploration of Large Language Model Society. In *Thirty-seventh Conference on Neural Information Processing Systems*.

Li, J.; Zhang, Q.; Yu, Y.; FU, Q.; and Ye, D. 2024a. More Agents Is All You Need. *Transactions on Machine Learning Research*.

Li, S.; Yang, J.; and Zhao, K. 2023. Are you in a masquerade? exploring the behavior and impact of large language model driven social bots in online social networks. *arXiv preprint arXiv:2307.10337*.

Li, Y.; Du, Y.; Zhang, J.; Hou, L.; Grabowski, P.; Li, Y.; and Ie, E. 2024b. Improving Multi-Agent Debate with Sparse Communication Topology. In Al-Onaizan, Y.; Bansal, M.; and Chen, Y.-N., eds., *Findings of the Association for Computational Linguistics: EMNLP 2024*, 7281–7294. Miami, Florida, USA: Association for Computational Linguistics.

Liang, T.; He, Z.; Jiao, W.; Wang, X.; Wang, Y.; Wang, R.; Yang, Y.; Shi, S.; and Tu, Z. 2024. Encouraging Divergent Thinking in Large Language Models through Multi-Agent Debate. In Al-Onaizan, Y.; Bansal, M.; and Chen, Y.-N., eds., *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, 17889–17904. Miami, Florida, USA: Association for Computational Linguistics.

Liu, J.; Zheng, Y.; Cheng, R.; Wu, Q.; Guo, W.; Ni, F.; Liang, H.; Yuan, Y.; Mao, H.; Zhang, F.; and Hao, J. 2025. From Chaos to Order: The Atomic Reasoner Framework for Fine-grained Reasoning in Large Language Models. arXiv:2503.15944.

Liu, Z.; Zhang, Y.; Li, P.; Liu, Y.; and Yang, D. 2024. A Dynamic LLM-Powered Agent Network for Task-Oriented Agent Collaboration. In *First Conference on Language Modeling*.

OpenAI. 2023. GPT - 4 Technical Report. https://openai.com/research/gpt-4Accessed 2025-12-15.

Park, J. S.; O'Brien, J.; Cai, C. J.; Morris, M. R.; Liang, P.; and Bernstein, M. S. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, 1–22.

Park, J. S.; Zou, C. Q.; Shaw, A.; Hill, B. M.; Cai, C.; Morris, M. R.; Willer, R.; Liang, P.; and Bernstein, M. S. 2024. Generative agent simulations of 1,000 people. *arXiv preprint arXiv:2411.10109*.

Pham, C.; Liu, B.; Yang, Y.; Chen, Z.; Liu, T.; Yuan, J.; Plummer, B. A.; Wang, Z.; and Yang, H. 2024. Let Models Speak Ciphers: Multiagent Debate through Embeddings. In *The Twelfth International Conference on Learning Representations*.

Qwen; :; Yang, A.; Yang, B.; Zhang, B.; Hui, B.; Zheng, B.; Yu, B.; Li, C.; Liu, D.; Huang, F.; Wei, H.; Lin, H.; Yang, J.; Tu, J.; Zhang, J.; Yang, J.; Yang, J.; Zhou, J.; Lin, J.; Dang, K.; Lu, K.; Bao, K.; Yang, K.; Yu, L.; Li, M.; Xue, M.; Zhang, P.; Zhu, Q.; Men, R.; Lin, R.; Li, T.; Tang, T.; Xia, T.; Ren, X.; Ren, X.; Fan, Y.; Su, Y.; Zhang, Y.; Wan, Y.; Liu, Y.; Cui, Z.; Zhang, Z.; and Qiu, Z. 2025. Qwen2.5 Technical Report. arXiv:2412.15115.

Reddy, R. G.; Mukherjee, S.; Kim, J.; Wang, Z.; Tur, D. H.; and Ji, H. 2024. Infogent: An Agent-based Framework for Web Information Aggregation. In *NeurIPS 2024 Workshop on Open-World Agents*.

Shinn, N.; Cassano, F.; Gopinath, A.; Narasimhan, K. R.; and Yao, S. 2023. Reflexion: language agents with verbal reinforcement learning. In *Thirty-seventh Conference on Neural Information Processing Systems*.

Smit, A. P.; Grinsztajn, N.; Duckworth, P.; Barrett, T. D.; and Pretorius, A. 2024. Should we be going MAD? A Look at Multi-Agent Debate Strategies for LLMs. In *Forty-first International Conference on Machine Learning*.

Suzgun, M.; Scales, N.; Schärli, N.; Gehrmann, S.; Tay, Y.; Chung, H. W.; Chowdhery, A.; Le, Q. V.; Chi, E. H.; Zhou, D.; et al. 2022. Challenging big-bench tasks and whether chain-of-thought can solve them. *arXiv preprint arXiv:2210.09261*.

Wang, J.; WANG, J.; Athiwaratkun, B.; Zhang, C.; and Zou, J. 2025. Mixture-of-Agents Enhances Large Language Model Capabilities. In *The Thirteenth International Conference on Learning Representations*.

Wang, Q.; Wang, Z.; Su, Y.; Tong, H.; and Song, Y. 2024. Rethinking the Bounds of LLM Reasoning: Are Multi-Agent Discussions the Key? In Ku, L.-W.; Martins, A.; and Srikumar, V., eds., *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 6106–6131. Bangkok, Thailand: Association for Computational Linguistics.

Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; brian ichter; Xia, F.; Chi, E. H.; Le, Q. V.; and Zhou, D. 2022. Chain of Thought Prompting Elicits Reasoning in Large Language Models. In Oh, A. H.; Agarwal, A.; Belgrave, D.; and Cho, K., eds., *Advances in Neural Information Processing Systems*.

Xiao, Y.; Liu, J.; Zheng, Y.; Xie, X.; Hao, J.; Li, M.; Wang, R.; Ni, F.; Li, Y.; Luo, J.; et al. 2024. CellAgent: An LLM-driven Multi-Agent Framework for Automated Single-cell Data Analysis. *bioRxiv*, 2024–05.

Yang, T.; Tang, H.; Bai, C.; Liu, J.; Hao, J.; Meng, Z.; and Liu, P. 2021. Exploration in Deep Reinforcement Learning: A Comprehensive Survey. *CoRR*, abs/2109.06668.

Yang, Y.; Chen, G.; Tang, H.; Liu, F.; Deng, D.; and Heng, P.-A. 2025. Dual Ensembled Multiagent Q-Learning with Hypernet Regularizer. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '25, 2226–2234. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS).

Yao, X.; Wang, Y.; Zhu, P.; Lin, W.; Li, J.; Li, W.; and Hu, Q. 2024. Socialized Learning: Making Each Other Better Through Multi-Agent Collaboration. In *Forty-first International Conference on Machine Learning*.

Zeng, A.; Liu, X.; Du, Z.; Wang, Z.; Lai, H.; Ding, M.; Yang, Z.; Xu, Y.; Zheng, W.; Xia, X.; et al. 2022. Glm-130b: An open bilingual pre-trained model. *arXiv preprint arXiv:2210.02414*.

Zhang, J.; Xu, X.; Zhang, N.; Liu, R.; Hooi, B.; and Deng, S. 2024. Exploring Collaboration Mechanisms for LLM Agents: A Social Psychology View. In Ku, L.-W.; Martins, A.; and Srikumar, V., eds., *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 14544–14607. Bangkok, Thailand: Association for Computational Linguistics.

Zheng, Y.; Xie, X.; Su, T.; Ma, L.; Hao, J.; Meng, Z.; Liu, Y.; Shen, R.; Chen, Y.; and Fan, C. 2019. Wuji: Automatic Online Combat Game Testing Using Evolutionary Deep Reinforcement Learning. In *2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 772–784.