UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
DPMT: Dual Process Multi-scale Theory of Mind Framework for Real-time Human-AI Collaboration

Permalink
https://escholarship.org/uc/item/63r4d8b2

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors
Li, Xiyun
Ding, Yining
Jiang, Yuhua
et al.

Publication Date
2025

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

DPMT: Dual Process Multi-scale Theory of Mind Framework for Real-time Human-AI Collaboration

Xiyun Li (lixiyun2020@ia.ac.cn)[1,2],Yining Ding[3],Yuhua Jiang[4],Yunlong Zhao[1], Runpeng Xie[1], Shuang Xu[1],Yuanhua Ni[3], Yiqin Yang (yiqin.yang@ia.ac.cn)[1*],Bo Xu (xubo@ia.ac.cn)[1,2*]

[1]The Key Laboratory of Cognition and Decision Intelligence for Complex Systems,
Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China
[2]School of Future Technology, University of Chinese Academy of Sciences, Beijing, 101408, China
[3]Nankai University, Tianjin, 300350, China
[4]Tsinghua University, Beijing, 100084, China
*Corresponding author

Abstract

Real-time human-artificial intelligence (AI) collaboration is crucial yet challenging, especially when AI agents must adapt to diverse and unseen human behaviors in dynamic scenarios. Existing large language model (LLM) agents often fail to accurately model the complex human mental characteristics such as domain intentions, especially in the absence of direct communication. To address this limitation, we propose a novel dual process multi-scale theory of mind (DPMT) framework, drawing inspiration from cognitive science's dual process theory. Our DPMT framework incorporates a multi-scale theory of mind (ToM) module to facilitate robust human partner modeling through mental characteristic reasoning. Experimental results demonstrate that DPMT significantly enhances human-AI collaboration, and ablation studies further validate the contributions of our multi-scale ToM in the slow system.

Keywords: theory of mind; human-AI collaboration; dual process theory; intelligent agents; large language model

Introduction

In recent years, large language models (LLMs) have achieved expert-level performance across various domains, including conversational question-answering assistants (Achiam et al., 2023) and code generation (Chen et al., 2021; R. Li et al., 2023). Building on the exceptional perception, comprehension, and reasoning capabilities of LLMs, LLM agents have rapidly emerged, attracting widespread attention. By integrating mechanisms such as memory and reflection, these agents demonstrate general-purpose abilities and robust generalization across tasks like operating system command execution (Deng et al., 2024; He et al., 2024), mobile assistance (J. Wang et al., 2024; C. Zhang et al., 2023), creative exploration in Minecraft (G. Wang et al., 2023), and embodied intelligence scenarios (Brohan et al., 2023).

However, achieving effective adaptation in complex collaborative tasks remains a significant challenge, which is critical for the broader application of LLM agents. Unlike tasks that an LLM agent can perform independently, collaborative tasks like *Overcooked* (Ghost Town Games, 2016) require the agent to work with diverse partners, including humans, to efficiently complete a series of complex sub-tasks—such as retrieving ingredients, cooking, and serving dishes within a limited time. The task list in *Overcooked* is too intricate for a single agent to handle alone. Furthermore, human collaborators, constrained by bounded rationality, may lack complete domain knowledge, introducing additional instability and uncertainty in human-artificial intelligence (AI) collaboration. Thus, effectively modeling human cognition is key to improving adaptation and collaboration.

Recent research has focused on improving the adaptability and decision-making capabilities of LLM agents by facilitating communication between human players and AI agents in environments with diverse human partners (Liu et al., 2023; S. Zhang et al., 2024). Leveraging their advanced natural language understanding, LLM agents can interpret human commands and formulate subsequent action plans (Liu et al., 2023). By adjusting their collaboration strategies through mutual communication with human partners, LLM agents significantly improve the overall collaboration performance of human-AI teams (S. Zhang et al., 2024).

However, the improved communication alone does not fully address the challenges of real-time scenarios. Current LLM agents lack the human-like cognitive ability known as "theory of mind" (ToM), which allows humans to understand and predict others' mental beliefs based on observed behaviors in social environment (Astington & Jenkins, 1995). This ability facilitates efficient collaboration in tasks without direct communication. The absence of ToM hinders LLM agents' performance in complex, real-time human-AI collaborative tasks. Although recent studies have explored ToM modeling to improve agent prediction (Rabinowitz et al., 2018; X. Li et al., 2023), these approaches rely on high-quality trajectories and prior knowledge of partners. As a result, their interpretability and generalization are limited, restricting their applicability in real collaborative scenarios.

Inspired by dual process theory (Vaisey, 2009; Lizardo et al., 2016), we propose a *cognitive dual process multi-scale theory of mind* (DPMT) to improve the interpretability and efficiency of human partner modeling in real-time human-AI collaboration. Our DPMT framework distinguishes between two decision-making systems for real-time human-AI collaboration: a *fast system* for automatic decisions and a *slow system* for modeling higher-level cognitive abilities. The core contribution of this work is the development of a *multi-scale theory of mind module* to simulate the slow system for understanding human partners' behavioral trajectories and reasoning about their mental characteristics, facilitating more effective collaboration. This ToM process follows a three-tiered reasoning process, which progresses from domain knowledge to cognitive style, and ultimately to domain intention, as illustrated in Figure 1. Experimental results from collaborative

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

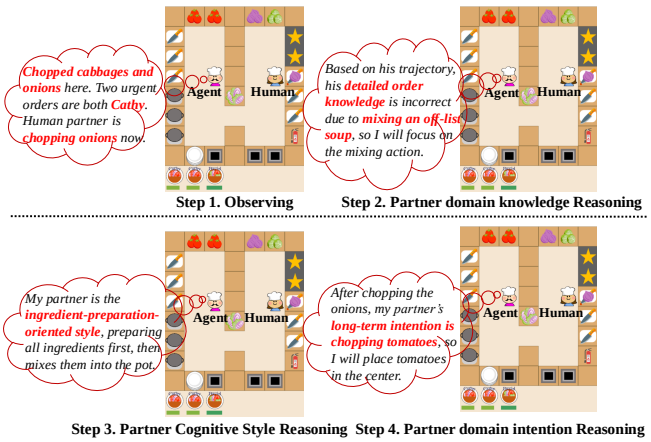

Figure 1: A four-step example of cognitive ToM modeling for human partners in the collaborative task Overcooked. In Step 1, the agent observes the environment and the human partner's recent behaviors. In Step 2, the agent understands the human trajectory and predicts the partner's domain knowledge. In Step 3, the agent analyzes the partner's behavioral preferences and categorizes his cognitive style based on the inferred domain knowledge. In Step 4, the agent infers the partner's domain intention for effective disicion-making.

tasks in *Overcooked* demonstrate the effectiveness of DPMT in improving real-time human-AI collaboration.

Related Work

LLM agent. Human-AI collaboration (Puig et al., 2020; Gao et al., 2020) constitutes a highly challenging research direction, motivating extensive investigations into strategies for enhancing the effectiveness, efficiency, and adaptability of human-AI systems. Compared to traditional agents, LLM-based agents contain an inherent world model, making them highly effective in reasoning and facilitating interpretable interaction for complex tasks, such as human-AI collaboration (Wu et al., 2024; Liu et al., 2023). Recent research (C. Zhang et al., 2024) designs the LLM agent framework with a verifier on the *Overcooked* environment, with knowledge and state as prompts to guide action planning. Recent work by (H. Zhang et al., 2023) proposes a modular framework that enables embodied agents to communicate and collaborate effectively, facilitating the efficient accomplishment of multi-agent long-term collaboration tasks.

Theory of Mind (ToM) refers to the ability of individuals to understand and predict others' mental states, including their intentions and goals for optimal decision-making and efficient social collaboration. Recent work introduces the ToM module in single-agent scenarios to predict an agent's future actions based on its historical trajectory (Rabinowitz et al., 2018). A ToM-based multi-agent communication method incorporates an additional ToM module to select optimal communication partners based on historical trajectories in multi-agent systems (Y. Wang, Xu, Wang, et al., 2021). The mutual ToM process, introduced in recent work (S. Zhang et al., 2024), enhances human-AI collaboration by improving reasoning and communication.

Partner modeling. In multi-agent reinforcement learning (MARL) tasks, partners often exhibit dynamic and diverse strategies, introducing significant challenges in non-stationarity and generalization. To address these challenges in multi-agent collaboration, several studies (Carroll et al., 2019; Shih & Sawhney, 2021) have explored partner modeling to predict partners' behaviors for enhancing efficient collaboration in complex MARL scenarios. Other approaches construct additional behavioral models (X. Li et al., 2023; Jiang et al., 2024) to simulate human partners and characterize their behavior styles. However, these methods focus on the accuracy of behavioral policy modeling, neglecting the simulation of partners' mental characteristics. Our work focuses on human partner's multi-scale ToM modeling for partner knowledge, cognitive styles, and domain intentions.

DPMT: Dual Process Multi-scale Theory of Mind Framework

Overview

As illustrated in Figure 2, our DPMT method comprises several components: the information extractor, the slow reasoning system, the fast intuitive system, and the action decoding module. The information extractor transforms environmental observation o_t and the human partner's trajectory c_{t-1} into the language state, represented as the system cue p_t^{system}. p_t^{system} encodes textual information about the environment and serves as the input for both the fast and slow systems.

The fast system focuses on the quick intuitive decision-making for each step, making macro-action m_t from a predefined macro-action set. In contrast, the slow system emphasizes cognitive multi-scale ToM reasoning, modeling the partner's mental characteristics k_t, y_t, n_t to assist the fast system in making macro-action decisions.

The action decoding module decomposes the current macro-action m_t into atomic actions a_t and executes them at a much higher frequency until m_t is completed. Once m_t is finished, the fast system determines the subsequent m_{t+1} based on the multi-scale partner reasoning k_t, y_t, n_t from the slow reasoning system. This hierarchical approach ensures a seamless integration of intuitive decision-making and cognitive reasoning for efficient human-AI collaboration.

Slow System

As illustrated in Figure 3, we propose a multi-scale mental characteristics system for human-AI collaboration, drawing inspiration from social psychology research on human decision-making (Hauser, 2019; Santos, 2008). These studies categorize various mental characteristics that influence individual behavior into three key dimensions: *domain knowledge*, *cognitive style*, and *domain intention* (Zhao & Smil-

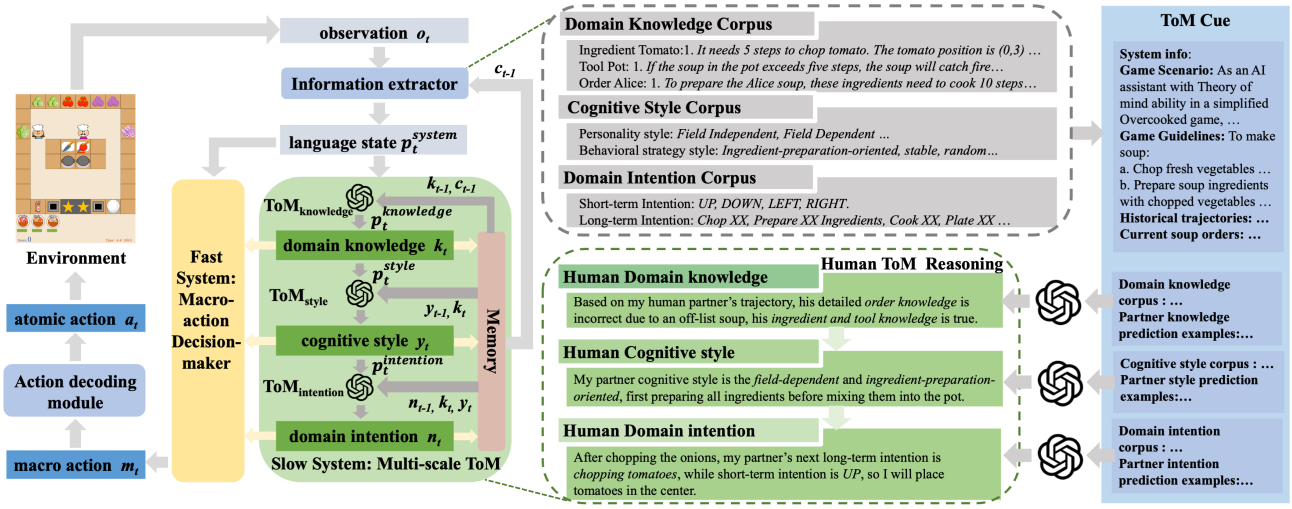

Figure 2: As shown in the figure above, our proposed DPMT framework is inspired by the dual process theory, comprising an information extractor module, an action decoding module, a fast system for macro-action decision-making and a slow system for human partner modeling with a multi-scale ToM module. The ToM module consists of three stages: domain knowledge reasoning ToM$_{\text{knowledge}}$, cognitive style reasoning ToM$_{\text{style}}$, and domain intention reasoning ToM$_{\text{intention}}$. By emulating the cognitive structure, our framework can effectively reason about the human partner and achieve efficient collaboration.

lie, 2015; Bostan, 2009). This system organizes these mental characteristics into a hierarchical architecture, with domain knowledge forming the foundation and domain intentions at the top, progressively enhancing interpretability. Building on this, we propose a multi-scale ToM model as our slow system, simulating the slow cognitive process in the dual process theory. This ToM model comprises multiple stages of human partner mental characteristic reasoning: the human domain knowledge reasoning stage ToM$_{\text{knowledge}}$, the human cognitive style reasoning stage ToM$_{\text{style}}$, and the human domain intention reasoning stage ToM$_{\text{intention}}$.

Small models face significant limitations in interpretability and general knowledge, making it challenging to construct an efficient and interpretable ToM model capable of effectively capturing partner styles and inferring domain intentions from partner trajectories. To overcome these challenges, we leverage LLMs as the foundation for the ToM model, leveraging their strengths in interpretability and extensive world knowledge to enhance human partner ToM modeling.

Domain knowledge reasoning stage Domain knowledge in the *Overcooked* includes *ingredient knowledge*, *order knowledge*, and *tool knowledge*. Since these types of knowledge are not directly provided in the environmental state, they are typically accumulated from the human partner's experiences. Predicting human partners' domain knowledge serves as the foundation for further predictions of their cognitive styles and domain intentions, forming a critical basis for higher-level ToM reasoning.

In our framework, the partner domain knowledge reasoning stage ToM$_{\text{knowledge}}$ takes the knowledge cue $p_t^{\text{knowledge}}$ as input, which integrates information from the language state

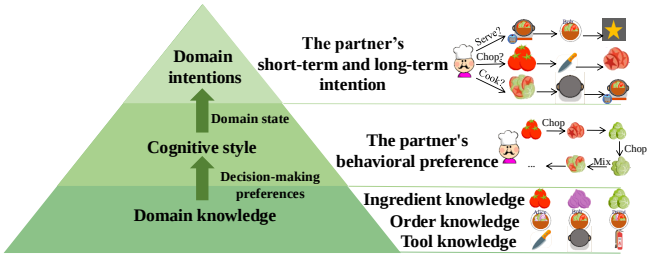

Figure 3: Our proposed multi-scale mental characteristics system includes domain knowledge, cognitive style, and domain intention. Domain knowledge includes ingredient knowledge, order knowledge, and tool knowledge. Cognitive style represents the partner's personality trait and behavioral style. Domain intention consists of long-term and short-term goals, representing the partner's decision-making process.

p_t^{system}, the human partner's trajectory context c_{t-1}, additional *partner domain knowledge reasoning cases*, and a *customized domain knowledge corpus*. The *domain knowledge corpus* comprises knowledge from three categories, such as *ingredient tomato* under *ingredient knowledge*, *order Bob* under *order knowledge*. The output of ToM$_{\text{knowledge}}$ is the human partner domain knowledge k_t, which includes insights such as whether the human partner understands the current task rules or the required order preparation sequences.

Cognitive style reasoning stage The cognitive style reflects a human partner's mental characteristic for decision-making preferences based on the domain knowledge, including *personality traits*, *behavioral strategy preferences*, and

risk preferences. From the perspective of *personality traits*, cognitive style can be categorized into *field-independent*, who prefer to complete an entire order independently, and *field-dependent*, who tend to collaborate by dividing order tasks into sub-tasks to complete an order. In field-dependent players, cognitive styles can be further categorized based on the behavioral strategy, which defines different task tendencies. For example, an *ingredient-preparation-oriented* style focuses primarily on retrieving and processing various ingredients. This style can also be divided into stable (consistently following a fixed strategy) and random according to the behavioral strategy. Accurate classification of cognitive styles is crucial for effective partner intention modeling.

In the partner style reasoning stage $\text{ToM}_{\text{style}}$, the cognitive style cue p_t^{style} comprises $p_t^{\text{knowledge}}$, the previous cognitive style prediction y_{t-1}, and the partner domain knowledge prediction k_t. This stage also incorporates an *additional partner style corpus* to predict the human partner's cognitive style y_t. In the partner style corpus, we design various partner cognitive styles based on behavioral strategies and personality traits observed in human behavioral experiments. These partner styles are stored in a structured format consisting of a name, a descriptive paragraph, and representative cases.

Domain intentions reasoning stage Building upon the cognitive style and the current domain state, domain intentions include both *short-term and long-term intentions* of the human partner, which are crucial for achieving efficient human-AI collaboration. *Short-term intention* reasoning involves atomic action prediction, determining the human partner's current action, such as *UP*, *DOWN*, *LEFT*, *RIGHT*. In contrast, *long-term intention* reasoning focuses on predicting the human partner's macro-action, specifying their primary goal for the current phase, such as *Chop Tomato*, *Prepare Bob Ingredients*, *Cook Alice Soup*, *Plate David Soup*. Based on the k_t and y_t, the domain intention reasoning stage $\text{ToM}_{\text{intention}}$ incorporates a *short-term intention cases* and *long-term intention reasoning corpus* in domain intention cue $p_t^{\text{intention}}$ to output the human partner's predicted intentions n_t.

Fast System

To enable more real-time human-AI collaboration, our DPMT incorporates a fast system that simulates the quick decision-making system in the dual process theory. While powerful large models like GPT-4o demonstrate superior reasoning ability, their high latency makes them unsuitable for the fast system. Instead, our fast system in DPMT leverages a smaller-scale LLM, such as llama-13B, to reduce latency in macro-action decision-making and achieve real-time collaboration, using the environmental language state p_t^{system}, the fast system prompt, and currently available actions as input.

As illustrated in Figure 4, the fast system calculates macro-action probabilities by computing token probabilities, similar to recent works (Liu et al., 2023; Tan et al., 2024). Each macro-action is represented as a sequence of tokens. The LLM evaluates the probability of token sequences based on the human ToM reasoning from the slow system. These probabilities are then normalized to form the macro-action probability distribution at step t. The fast system selects the macro-action m_t with the highest probability from this distribution.

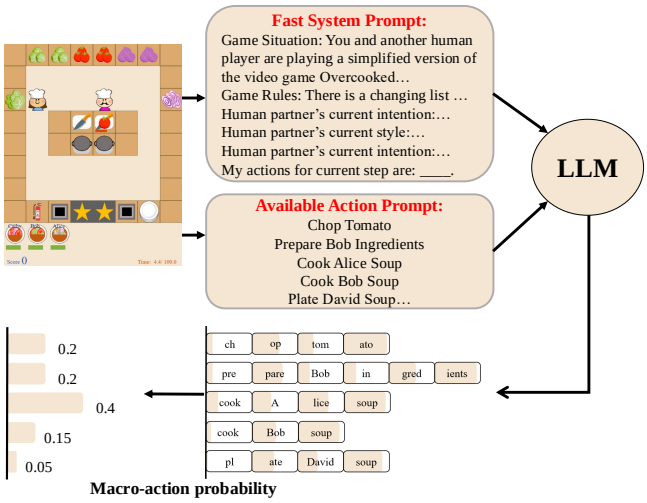

Figure 4: Overview of our fast system in our proposed DPMT. The fast system first computes token probabilities for every macro-action, then derives the probability distribution of available macro-actions, selecting the one with the highest probability as the output.

Although smaller models may make suboptimal decisions, our fast system compensates for this by leveraging accurate partner reasoning from the slow system. These two systems are implemented using multi-threading, where the slow system's reasoning for partner modeling operates at a longer time scale than the fast system's macro-action decisions. This design allows the fast system to run more frequently, enabling efficient real-time human-AI collaboration.

Experiments

Experimental Environment

Our experimental environment is *Overcooked* (Ghost Town Games, 2016; Liu et al., 2023), a primary human-AI collaboration benchmark. We have conducted experiments on three maps based on the recent work (Liu et al., 2023), which introduced an expanded version of the traditional *Overcooked* environment for human-AI collaboration. In this environment, the agent and the human player work together to complete cooking tasks as quickly as possible to achieve a higher reward within a limited time. As shown in Figure 5, the environment includes three maps: *Ring*, *Bottleneck*, and *Quick*, each with varying levels of difficulty. Among them, the *Quick* and *Bottleneck* maps are more challenging. As shown in Figure 5, completing an order requires following a specific sequence of steps. The agent and human player must retrieve the necessary ingredients, chop them, combine them according to the different recipes, cook them for the required dura-

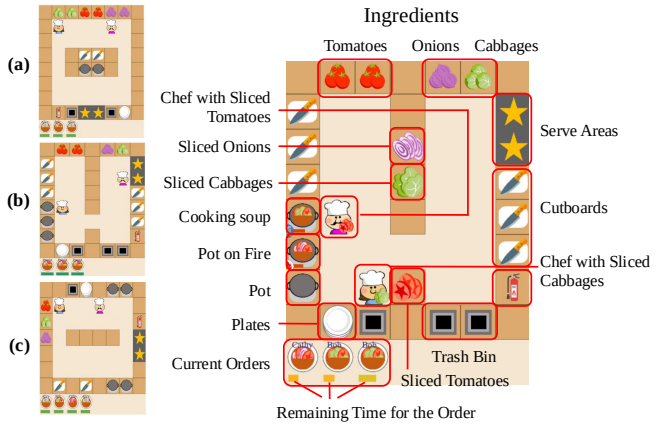

Figure 5: The left part shows our Overcooked environments. The maps from top to bottom are Ring, Bottleneck, and Quick. The right part shows the detailed layout description of the bottleneck environment.

tion to make a dish, plate the dish before it overcooks, and serve the dish to the designated position. Once the dish is served, the player receives a reward (for example, 15 points), while failing to submit a dish results in a penalty (for example, -5 points). Players can choose and execute one action simultaneously from the following options: UP, DOWN, LEFT, or RIGHT, using the directional arrow keys. This environment designs a macro-action set with 21 predefined macro actions for the LLM agent, such as *Cook Alice Soup*.

Human Experiment 1: Collaborating with Human Partners with Specific Strategies

Experimental Setup. In the human-AI collaboration experiment, similar to HLAgent (Liu et al., 2023), the human player followed the fixed strategy to control the partner agent and interact with objects using the arrow keys like 'up,' and 'down.' In our human experiments, baseline methods included our DPMT with Qwen, our DPMT without the multi-scale ToM (MsToM) and the HLAgent. Since communication between agents was not allowed in our task setting, we used the 0-command HLAgent as our baseline. To ensure experimental fairness, we established a stopping criterion that terminated LLM agent executions once the trajectory length reached 500 steps. In this part of the experiment, we used game scores to measure human-AI collaboration efficiency. The experiments were conducted on three maps and repeated five times under consistent conditions. The average human-AI collaboration scores are shown in Table 1.

Main Results. As shown in Table 1, our proposed DPMT method consistently achieved the highest collaboration performance when paired with fixed-strategy human partners across all three maps: *Ring*, *Quick*, and *Bottleneck*, highlighting the effectiveness of our DPMT framework. Compared to the baseline methods, the introduction of the multi-scale mental characteristic ToM reasoning module (MsToM) in our

Table 1: Score comparison between baseline methods and our DPMT method across all three maps in the human experiment 1. Standard deviations are shown in parentheses.

Methods	*Ring*	*Bottleneck*	*Quick*
Our DPMT	**121 (±13.56)**	**101 (±13.56)**	104 (±16.73)
Our DPMT w Qwen	100 (±15.17)	95 (±22.80)	**105 (±18.71)**
Our DPMT w/o MsToM	44 (±22.23)	23 (±22.93)	9 (±19.34)
HLAgent	99 (±18.81)	60 (±20.74)	87 (±16.00)

DPMT framework led to a higher success rate in serving soups, a lower obstruction rate, and a lower incidence of dish fires by accurately predicting human partners' mental characteristics, such as domain intentions. In contrast, the lack of ToM reasoning in DPMT w/o MsToM often resulted in ineffective collaboration, including redundant actions, route blockages, overcooked dishes, and frequent task repetition, ultimately incurring performance penalties.

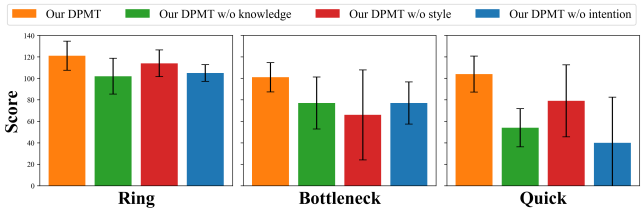

Figure 6: Score results of ablation experiments for different mental characteristics in human experiment 1. The black line denotes the standard deviation.

Additionally, compared to the HLAgent baseline, our method achieved superior and more stable performance, particularly on complex maps such as *Bottleneck*. The *Bottleneck* map has a more complex structure with a narrow passage, resulting in a higher blockage rate. This highlights the critical role of accurate ToM reasoning in achieving efficient human-AI collaboration. Furthermore, substituting the GPT-4o API with Qwen-72B as the base model for mental characteristic reasoning resulted in slight improvements in collaboration performance on the *Quick* map. This enhancement is likely attributed to minor fluctuations in real-time experimental results caused by latency.

Ablation Results. We conducted ablation experiments to analyze the effectiveness of our DPMT method further. As shown in Figure 6, we compared the complete DPMT method with its ablated versions, where specific mental characteristic reasoning stages (w/o knowledge, w/o style, and w/o intention) were removed. Our method consistently outperformed the ablated versions across all maps (*Ring*, *Quick*, and *Bottleneck*), especially on more complex maps, where effective ToM reasoning plays a crucial role in coordination. For example, removing the intention reasoning layer led to a significant performance drop on the *Quick* map, while omitting the style

layer caused notable declines on the *Bottleneck* map. These results indicate the importance of each ToM characteristic and demonstrate the necessity of hierarchical ToM modeling for effective partner reasoning and improved collaboration.

Questionnaire for Subjective Judgment in Human Experiment 2

1 Game Ability
- "Can the AI agent efficiently utilize time to complete as many effective tasks as possible based on the order situation?"

2 Human-AI Collaboration Fluidity
- "Can the AI agent accurately perceive the current order requirements and your actions, adapt its own actions accordingly, and collaborate efficiently with you avoiding redundant work?"

3 Theory of mind (ToM) Ability
- "To what extent do you feel that your AI agent understands your intentions? Can the AI agent effectively avoid conflicts with your operations, such as not taking over kitchen equipment, interrupting, or obstructing your actions?"
- "Do you feel comfortable collaborating with the AI agent? Can it adapt to your movement habits and coordinate effectively to minimize idleness or delays, accelerating order completion? Can it adjust its path to avoid congestion in narrow passages by observing your movement?"
- Do you trust the AI agent? Can it adapt its behavior to align with the common goal in high-pressure or emergency situations (e.g., time shortages or fire)?
- .....

Figure 7: Our proposed questionnaire for subjective judgment metric with 5-point Likert scales was used in our human experiment 2. The questionnaire includes three parts: game ability, human-AI collaboration fluidity, and ToM ability. Different questionnaires were designed for different maps.

Human Experiment 2: Collaborating with Human Partners with Diverse Strategies

Experiment Setup. An ideal AI agent should not only cooperate seamlessly with fixed-strategy human players but also collaborate efficiently with unseen human partners who employ diverse strategies. To further analyze the effectiveness and the human preference of our proposed DPMT framework, we conducted additional behavioral experiments for human studies. These experiments were performed across three maps with 30 participants (26 male, 4 female, ages 20–50). As shown in Figure 7, we designed a 5-point Likert scale questionnaire for each map, consisting of several questions (Hoffman, 2019; S. Zhang et al., 2024). This experiment followed a within-subject design, where each participant collaborated with both the DPMT and three baseline agents (A-D). To ensure fairness and validity, we randomized the order of partner agents ('ABCD' or 'DBCA') for each participant. Additionally, to evaluate the AI agent's performance more effectively, we extended the game duration to 90 seconds, longer than in previous human experiment 1.

All participants were required to review the informed consent form and game rules before starting. To familiarize themselves with the task environment and rules, they first collaborated with rule-based agents across three maps. After completing the collaboration task on each map, they filled out a corresponding questionnaire, rating different agents based on their level of agreement with each statement.

Results. In our collaboration experiments with diverse human participants, we evaluated performance using two key metrics: average collaboration scores within a fixed time limit and subjective judgments from the participants. As shown in

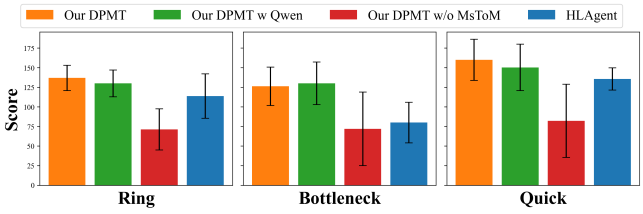

Figure 8: Average collaboration score comparison between baseline methods and our proposed DPMT when paired with diverse participants in the human experiment 2.

Figure 8, our DPMT method consistently outperformed baseline methods across all three maps, exhibiting a result trend similar to the main result of human experiment 1, demonstrating its ability to adapt to both specific strategies and diverse unseen human collaborators. For the subjective judgment scores, we computed the average scores based on our proposed questionnaire, with a maximum score of 5, as shown in Table 2. Notably, our method achieved higher subjective scores, further highlighting its superior ToM reasoning capability for effective partner modeling.

Table 2: The average subjective judgment when paired with diverse unseen participants in human experiment 2 ranges from 0 to 5. Standard deviations are shown in parentheses.

Methods	*Ring*	*Bottleneck*	*Quick*
Our DPMT	**4.07 (±0.53)**	**4.02 (±0.45)**	**4.02 (±0.61)**
Our DPMT w Qwen	3.51 (±0.58)	3.73 (±0.47)	3.89 (±0.68)
Our DPMT w/o MsToM	2.20 (±0.70)	2.16 (±0.67)	2.04 (±0.60)
HLAgent	3.09 (±0.30)	2.78 (±0.59)	3.22 (±0.69)

Conclusion

Drawing inspiration from the dual process theory, we introduce a cognitive dual process multi-scale ToM (DPMT) framework, which simulates both the fast system for quick decision-making and the slow system for ToM reasoning. The slow system in DPMT employs a three-stage ToM reasoning process—partner domain knowledge, cognitive style, and domain intention—to improve the interpretability and generalization of partner ToM modeling. Additionally, it integrates a structured corpus of mental characteristics and a memory module that stores partner trajectories and ToM reasoning predictions. Our DPMT outperforms baseline methods in human-AI collaboration experiments, demonstrating superior adaptability and ToM ability. Ablation studies further validate the contribution of different reasoning stages. Notably, our method achieves significantly stronger results on complex maps, highlighting its potential for tackling more intricate human-AI collaborative tasks.

Acknowledgments

This work is supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (No.XDA27040200) and the National Key R&D Program of China (No.2022ZD0116405).

References

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., … others (2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Astington, J. W., & Jenkins, J. M. (1995). Theory of mind development and social understanding. *Cognition & Emotion*, *9*(2-3), 151–165.

Bostan, B. (2009). Player motivations: A psychological perspective. *Computers in Entertainment (CIE)*, *7*(2), 1–26.

Brohan, A., Chebotar, Y., Finn, C., Hausman, K., Herzog, A., Ho, D., … others (2023). Do as i can, not as i say: Grounding language in robotic affordances. In *Conference on robot learning* (pp. 287–318).

Carroll, M., Shah, R., Ho, M. K., Griffiths, T., Seshia, S., Abbeel, P., & Dragan, A. (2019). On the utility of learning about humans for human-ai coordination. *Advances in neural information processing systems*, *32*.

Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H. P. d. O., Kaplan, J., … others (2021). Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*.

Deng, X., Gu, Y., Zheng, B., Chen, S., Stevens, S., Wang, B., … Su, Y. (2024). Mind2web: Towards a generalist agent for the web. *Advances in Neural Information Processing Systems*, *36*.

Gao, X., Gong, R., Zhao, Y., Wang, S., Shu, T., & Zhu, S.-C. (2020). Joint mind modeling for explanation generation in complex human-robot collaborative tasks. In *2020 29th ieee international conference on robot and human interactive communication (ro-man)* (pp. 1119–1126).

Ghost Town Games. (2016). *Overcooked.*

Hauser, O. P. (2019). Social dilemmas among unequals. *Nature*, *572*(7770), 524–527.

He, H., Yao, W., Ma, K., Yu, W., Dai, Y., Zhang, H., … Yu, D. (2024). Webvoyager: Building an end-to-end web agent with large multimodal models. *arXiv preprint arXiv:2401.13919*.

Hoffman, G. (2019). Evaluating fluency in human–robot collaboration. *IEEE Transactions on Human-Machine Systems*, *49*(3), 209–218.

Jiang, Y., Liu, Q., Ma, X., Li, C., Yang, Y., Yang, J., … Zhao, Q. (2024). Learning diverse risk preferences in population-based self-play. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 38, pp. 12910–12918).

Li, R., Allal, L. B., Zi, Y., Muennighoff, N., Kocetkov, D., Mou, C., … others (2023). Starcoder: may the source be with you! *arXiv preprint arXiv:2305.06161*.

Li, X., Ni, Z., Ruan, J., Meng, L., Shi, J., Zhang, T., & Xu, B. (2023). Mixture of personality improved spiking actor network for efficient multi-agent cooperation. *Frontiers in Neuroscience*, *17*, 1219405.

Liu, J., Yu, C., Gao, J., Xie, Y., Liao, Q., Wu, Y., & Wang, Y. (2023). Llm-powered hierarchical language agent for real-time human-ai coordination. *arXiv preprint arXiv:2312.15224*.

Lizardo, O., Mowry, R., Sepulvado, B., Stoltz, D. S., Taylor, M. A., Van Ness, J., & Wood, M. (2016). What are dual process models? implications for cultural analysis in sociology. *Sociological Theory*, *34*(4), 287–310.

Puig, X., Shu, T., Li, S., Wang, Z., Liao, Y.-H., Tenenbaum, J. B., … Torralba, A. (2020). Watch-and-help: A challenge for social perception and human-ai collaboration. In *International conference on learning representations.*

Rabinowitz, N., Perbet, F., Song, F., Zhang, C., Eslami, S. A., & Botvinick, M. (2018). Machine theory of mind. In *International conference on machine learning* (pp. 4218–4227).

Santos, F. C. (2008). Social diversity promotes the emergence of cooperation in public goods games. *Nature*, *454*(7201), 213–216.

Shih, A., & Sawhney, A. (2021). On the critical role of conventions in adaptive human-ai collaboration. In *International conference on representation learning.*

Tan, W., Zhang, W., Liu, S., Zheng, L., Wang, X., & An, B. (2024). True knowledge comes from practice: Aligning large language models with embodied environments via reinforcement learning. In *The twelfth international conference on learning representations.*

Vaisey, S. (2009). Motivation and justification: A dual-process model of culture in action. *American journal of sociology*, *114*(6), 1675–1715.

Wang, G., Xie, Y., Jiang, Y., Mandlekar, A., Xiao, C., Zhu, Y., … Anandkumar, A. (2023). Voyager: An open-ended embodied agent with large language models. *arXiv preprint arXiv:2305.16291*.

Wang, J., Xu, H., Ye, J., Yan, M., Shen, W., Zhang, J., … Sang, J. (2024). Mobile-agent: Autonomous multi-modal mobile device agent with visual perception. *arXiv preprint arXiv:2401.16158*.

Wang, Y., Xu, J., Wang, Y., et al. (2021). Tom2c: Target-oriented multi-agent communication and cooperation with theory of mind. In *International conference on learning representations.*

Wu, Y., Min, S. Y., Prabhumoye, S., Bisk, Y., Salakhutdinov, R. R., Azaria, A., … Li, Y. (2024). Spring: Studying papers and reasoning to play games. *Advances in Neural Information Processing Systems*, *36*.

Zhang, C., Yang, K., Hu, S., Wang, Z., Li, G., Sun, Y., … others (2024). Proagent: building proactive cooperative agents with large language models. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 38, pp. 17591–17599).

Zhang, C., Yang, Z., Liu, J., Han, Y., Chen, X., Huang, Z., … Yu, G. (2023). Appagent: Multimodal agents as smart-

phone users. *arXiv preprint arXiv:2312.13771*.

Zhang, H., Du, W., Shan, J., Zhou, Q., Du, Y., Tenenbaum, J. B., … Gan, C. (2023). Building cooperative embodied agents modularly with large language models. *arXiv preprint arXiv:2307.02485*.

Zhang, S., Wang, X., Zhang, W., Chen, Y., Gao, L., Wang, D., … Wen, Y. (2024). Mutual theory of mind in human-ai collaboration: An empirical study with llm-driven ai agents in a real-time shared workspace task. *arXiv preprint arXiv:2409.08811*.

Zhao, K., & Smillie, L. D. (2015). The role of interpersonal traits in social decision making: Exploring sources of behavioral heterogeneity in economic games. *Personality and Social Psychology Review*, *19*(3), 277–302.