MindDial: Enhancing Conversational Agents with Theory-of-Mind for Common Ground Alignment and Negotiation

Shuwen Qiu[1] **Mingdian Liu**[2] **Hengli Li**[3,4,6] **Song-Chun Zhu**[3,4,5,6] **Zilong Zheng**[4,6]✉

[1] UCLA [2] ISU [3] PKU [4] BIGAI [5] THU
[6] State Key Laboratory of General Artificial Intelligence
s.qiu@ucla.edu, mingdian@iastate.edu, {lihengli,sczhu,zlzheng}@bigai.ai

Abstract

Humans talk in daily conversations while aligning and negotiating the expressed meanings or common ground. Despite the impressive conversational abilities of the large generative language models, they do not consider the individual differences in contextual understanding in a shared situated environment. In this work, we propose MindDial, a novel conversational framework that can generate situated free-form responses with theory-of-mind (ToM) modeling. We introduce an explicit mind module that can track the speaker's belief and the speaker's prediction of the listener's belief. Then the next response is generated to resolve the belief difference and take task-related action. Our framework is applied to both prompting and fine-tuning-based models, and is evaluated across scenarios involving both common ground alignment and negotiation. Experiments show that models with mind modeling can achieve higher task outcomes when aligning and negotiating common ground. The ablation study further validates the three-level belief design can aggregate information and improve task outcomes in both cooperative and negotiating settings.

1 Introduction

We align and negotiate our common ground every day in daily chit-chat (Clark and Wilkes-Gibbs, 1986; Bazerman et al., 2000). In a common ground alignment scenario, agents are talking toward a joint goal, topics ranging from daily trivia to important multi-party meetings. In common ground negotiation situations, two parties resolve the differences in their beliefs, intents, or goals in a way that both find acceptable, such as item trading and discussing job offers (Veinott et al., 1999; Beers et al., 2006). Though it seems easy between human conversations, it requires complicated social capabilities. Importantly, for all types of human communication including language, the relationship between the overt communicative act and common ground – of whatever type – is complementary. That is, as more can be assumed to be shared between communicator and recipient, less needs to be overtly expressed (Tomasello, 2010). Taking Figure 1B as an example, when Bob asks about "Joe Davis", Alice will align the precise referents of the query by keeping "Joe" but correct "Davis" to "Smith". In this process, people need to realize what is shared and what needs to be further aligned or negotiated – which requires the understanding between points of view from their own and others' perspectives (Blutner, 2000; De Weerd et al., 2015) – the cognitive capability known as theory-of-mind (ToM).

The recent surge of large language models (LLMs) (Radford et al., 2019; Brown et al., 2020) have dominated the natural language processing (NLP) community for their prominent natural language generation performance. Although LLMs have shown their potential in ToM benchmarks (Kosinski, 2023; Ullman, 2023; Sileo and Lernould, 2023; Kim et al., 2023), applying ToM for situated dialogue generation remains underexplored. In these situated tasks, agents' interactions are influenced by the environment, their shared experiences, and immediate goals. The participants need to take into account not only the linguistic content but also factors such as the social context, prior knowledge, and each other's beliefs. Without ToM, the models can only provide the most possible response as a one-turn question-answering as shown in Figure 1A. To enable LLMs to interact with people in a more socially realistic manner, it is essential to incorporate ToM for various forms of communications, such as aligning and negotiating common ground within dialogues (Burleson, 2007; Chiu et al., 2023; Fu et al., 2023).

In this work, we introduce **MindDial**, a new dialogue framework designed to facilitate the alignment and negotiation of common ground in situated dialogues, incorporating ToM modeling. In-

✉ Correspondence to Zilong Zheng <zlzheng@bigai.ai>.

Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue, pages 746–759
September 18–20, 2024.

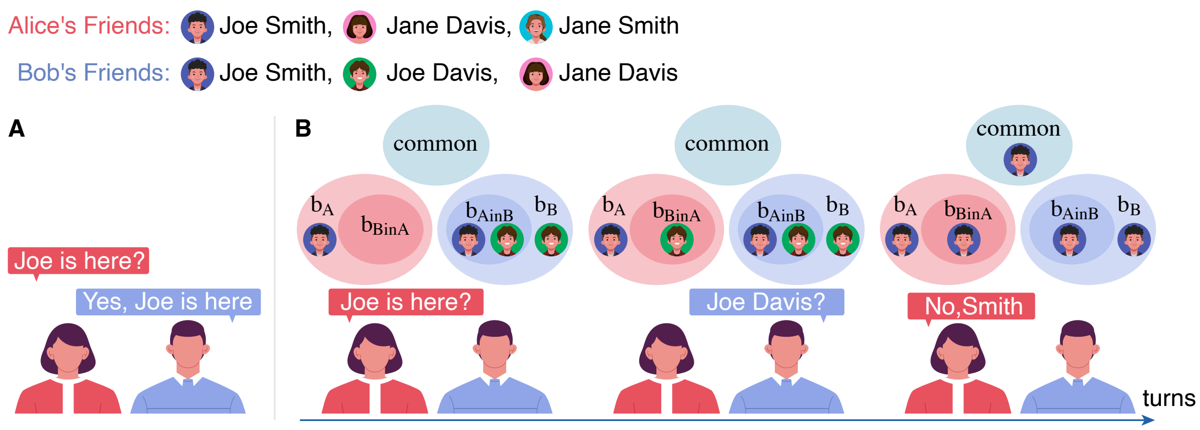

Figure 1: **Left:** Single-turn question answering. **Right:** Multi-turn common ground alignment. Speakers will update their belief estimation based on context and generate the next response to reduce the belief differences.

spired by the complementary role between common ground and communication, we design the two-step response generation. First, an explicit mind module estimates the speaker's current perspective of the world (the first-order belief) and also helps speaker's estimate the other's perspective of the world (the second-order belief) (Grueneisen et al., 2015; Braüner et al., 2016). Then, the next response is aimed at resolving the belief difference. As shown in Figure 1B, Alice says "No, Smith" when her first-order belief b_A ("Joe Smith") does not equal to her second-order belief b_{BinA} ("Joe Davis") .

In sum, we consider our contributions as three-fold:

i) We design a framework incorporating an explicit mind estimation module that tracks the first-order and second-order beliefs. Resolving the belief difference between the two will support the next response generation.

ii) We explore two types of response generators – fine-tuning and prompting-based models. The experiments show that our framework can successfully improve model performance in both groups.

iii) We test our framework on both aligning and negotiating settings. The evaluation results and user study validate that our framework can improve both the cooperation and negotiation abilities of the LLM agents. We ablate each level of the beliefs and find both first and second-order contribute to the final results.

2 Related work

Theory-of-Mind (ToM) ToM is a crucial capability for human social interactions developed in early life (Kovács et al., 2010; Richardson et al., 2018). In literature, early works model belief update through time in sequential games with partially observable Markov decision process (POMDP) (Baker et al., 2011; De Weerd et al., 2013; Vogel et al., 2013; Doshi et al., 2010; Han and Gmytrasiewicz, 2018). One agent's belief update is based on the estimate of others' current beliefs, resulting in an infinite recursion. However, in real life, studies have shown that humans could go no deeper than two levels of recursion (Camerer et al., 2004). Therefore, works (Fan et al., 2021) began the efforts to end the recursion when their beliefs merge into the "common mind".

Modeling the belief of others has been extensively studied in symbolic-like environments (Wunder et al., 2011; Rabinowitz et al., 2018; Kleiman-Weiner et al., 2016; Ho et al., 2016), where agents need to incorporate or compete for a goal. Efforts to measure models' ability to recognize false beliefs and perspective-taking also emerge in robotics (Yuan et al., 2020; Milliez et al., 2014), computer vision (Eysenbach et al., 2016; Fan et al., 2021), and natural language processing (Qiu et al., 2022; Nematzadeh et al., 2018) using the Sally-Anne test (Baron-Cohen et al., 1985). Different variants of the Sally-Anne test and ToM benchmarks are also proposed to test the ToM of large language models (Kosinski, 2023; Ullman, 2023; Sileo and Lernould, 2023; Kim et al., 2023). It is also shown that augmenting the model with external mind modules can help improve the performance of tasks involving intensive belief exchange and rich social interaction scenarios (Fan et al., 2021; Qiu et al., 2022; Li et al., 2023; Chiu et al., 2023). In this work, we explore ToM modeling can enhance the quality and efficiency of the response generation in both cooperative and semi-cooperative dialogue tasks.

Common ground alignment and negotiation In a cooperative dialogue task, to guarantee that the communication takes the least cost meanwhile providing the most informative messages, previous works proposed multiple methods to align the common ground between agents (Bohn et al., 2019; Anderson, 2021). Specifically for dialogue tasks, datasets have been collected to provide golden utterances when people try to align the common ground with each other based on structured knowledge (He et al., 2017a), in partially observable cooperative tasks (Bara et al., 2021; Kim et al., 2019), in multimodal and continuous environment (Haber et al., 2019; Udagawa and Aizawa, 2021). Frameworks have been adopted to model and predict the aligning dynamics using GNN, RNN, transformers, and LLMs (He et al., 2017a; Udagawa and Aizawa, 2021; Fischer, 2023; Zhang et al., 2023; Zhou et al., 2023). The inferred common ground is also used to generate more interesting and engaging conversations for the dialogue agents (Zhou et al., 2022).

Negotiation is treated as a semi-cooperative task since agents can have different goals but need to agree on the same decision (Lewis et al., 2017). It requires complex social skills and strategies like offering proposals and accepting or making counter-offers (Yamaguchi et al., 2021). To improve the negotiating abilities of the dialogue systems, datasets of open-domain human negotiation corpus have been introduced in embodied environment (DeVault et al., 2015), daily items split (Lewis et al., 2017; Chawla et al., 2021), buy and sell (He et al., 2018), job offer negotiation (Yamaguchi et al., 2021). Modeling begins with game theory and action selection (Nash Jr, 1950; Baarslag et al., 2013). For open-domain generation, methods have been designed to help the model plan ahead (Lewis et al., 2017; Iwasa and Fujita, 2018), give feedback about the current conversation (Zhou et al., 2019; Fu et al., 2023), detect negotiation breakdowns (Yamaguchi et al., 2021).

3 Task and Framework

3.1 Tasks

The situated dialogue corpus can be denoted as $\mathcal{D} = \{(U_n, K_n^p, y_n)\}_{n=1}^N$, where $U_n = (u_{n,1}, ..., u_{n,T})$ represents the dialogue history and T is the number of turns. $K_n^p = (k_{n,1}, ...k_{n,I})$ is for their knowledge base, where I is the number of knowledge passages. $p \in A, B$ represents the two agents. We assume the current speaker is A, and p will be dropped for the following formulation. y_n is A's next response or its action to achieve the task goals.

Alignment In the common ground alignment scenarios, we use the MutualFriend task (He et al., 2017b) shown in Figure 2. K denotes the private friend lists that two agents observe, and there is only one friend shared in their lists. The agents need to merge their estimation of the mutual friend through chat and finally finish the task goal by taking the action to select $k_i \in K$ as their mutual friend. The alignment is successful when their selections are the same.

Negotiation In the common ground negotiation scenarios, we use the CaSiNo task (Chawla et al., 2021) shown in Figure 2(Bottom). Two agents are planning a camp trip and need to distribute the uneven number of items. Based on their individual priority of the items K, they need to decide on the final item split agreement to maximize their gain of valuable items. At the end of the conversation, one agent proposes the item split deal while the other agent decides to accept or reject this deal. The negotiation is successful when the deal is accepted.

3.2 MindDial

The overall pipeline of our framework is shown in Figure 2. At the first stage, given the context history and private knowledge, the mind module f estimates the first and second-order beliefs over their solutions $b_A, b_{BinA} = f(U, K)$. The first-order belief represents A's estimation of the mutual friend or split deal. The second-order belief refers to A's understanding of B's estimation regarding the mutual friend or split deal. We choose to prompt the LLMs for b_A and b_{BinA} due to their ability to adapt flexibly in open-domain corpora. Therefore, the mind module can be applied to other situated dialogues when the knowledge base and beliefs are well-defined.

Then the response generator h generates the next utterance based on the dialogue history, its private knowledge, and the intention to align the first and second-order beliefs: $\tilde{y} = h(U, K, b_A, b_{BinA})$. We apply two methods to the response generator to activate their ability to resolve the belief difference in b_A and b_{BinA}: embedding this ability into LLM by finetuning and explicitly triggering this ability of LLM by prompting.

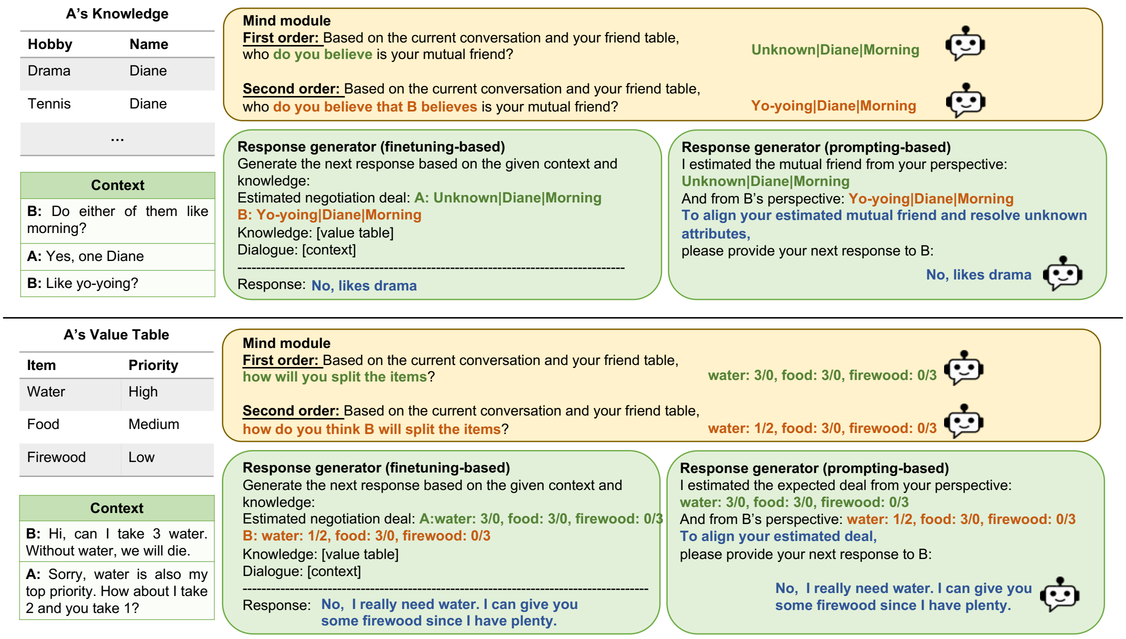

Figure 2: **Cases of ToM reasoning in MindDial.** Top: an *alignment* task from MutualFriend. Bottom: A *negotiation* task from CaSiNo. For each task, we first reason over the first- and second-order ToM beliefs of the conversational partner. Then we generate corresponding utterances wrt. the ToM estimation.

Finetuning-based For finetuning-based models, we prepare a small dataset in the format of $\{y, U, K, b_A, b_{BinA}\}$, where y is intended to resolve the gap between b_A and b_{BinA}. Different parts of model inputs are concatenated together with their corresponding tags as [Estimated belief], [Knowledge], and [Dialogue] shown in Figure 2. The models are trained to regress the next response y.

Prompting-based For prompting-based models, we directly ask the generator to generate the next response in order to resolve the difference and unknown values between b_A and b_{BinA}. The format follows as "I estimated mutual friend/deal from your perspective: b_A and from B's perspective: b_{BinA}. To align b_A and b_{BinA}, please provide your next response to B:".

4 Experiments

Dataset To provide a reasonable quantitative measure of belief dynamics in the dialogue, the expected dataset should contain rich belief exchanges. Meanwhile, the belief exchange and the final solution can be easily labeled. Therefore, we choose two representative settings to evaluate our framework. **MutualFriend** (He et al., 2017b): we consider it as an alignment dialogue scenario. In the MutualFriend task, each agent has a private knowledge base including a list of friends and their attributes like name, school, *etc*. There is a shared friend that both agents have and they need to chat with each other to find this mutual friend. We only keep the successful dialogues and the final data split for train/val/test is 7257/878/900. Each dialogue in the training set contains a maximum of 53 turns and each turn with a maximum length of 29. **CaSiNo** (Chawla et al., 2021): we consider it as the negotiating scenario. In the CaSiNo task, two agents need to split camping packages: 3 water, 3 firewood, and 3 food. Each of these items will be of either High, Medium, or Low value to each agent. The agents need to negotiate the distribution of the items through chat to maximize their final points calculated based on the number of items they get and the items' corresponding values. The data split for train/val/test is 900/30/100. Each dialogue in the training set contains a maximum of 27 turns and each turn with a maximum length of 106.

Mind modules To serve as a mind module in this task, the model is expected to understand long conversation contexts and the concept of first and second-order beliefs. We choose LLaMA-2-7B-chat, LLaMA-2-13B-chat (Touvron et al., 2023)[1], GPT-3.5, and GPT-4[2] as our mind reasoner for their potentials in ToM benchmarks and the flexible abilities of mind reasoning in open-domain dialogues.

[1] https://github.com/facebookresearch/LLaMA-recipes/tree/main

[2] gpt-3.5-turbo-1106, gpt-4-1106-preview

Models	Mind level	C	T	C_T	Models	Mind level	C	T	C_T
LLaMA-7B-ft	w/o mind	24.67	9.09	2.71	LLaMA-13B-ft	w/o mind	36.33	6.64	**5.47**
	b_A	28.33	7.92	**3.58**		b_A	42.00	8.66	4.85
	b_{BinA}	**29.33**	8.33	3.52		b_{BinA}	39.33	7.70	5.11
	b_A+b_{BinA}	28.33	8.87	3.20		b_A+b_{BinA}	**44.67**	8.85	5.05
GPT-3.5	w/o mind	10.67	5.74	1.86	GPT-4	w/o mind	75.00	9.72	7.71
	b_A	18.33	5.91	3.10		b_A	75.00	9.41	7.97
	b_{BinA}	12.33	5.91	2.09		b_{BinA}	69.67	8.84	7.88
	b_A+b_{BinA}	**24.33**	6.04	**4.03**		b_A+b_{BinA}	**76.00**	8.88	**8.56**

Table 1: **MutualFriend: results with different mind settings**. Settings without mind reasoning are marked as w/o mind. Settings considering only the first-order are marked as b_A, with only the second-order are b_{BinA}, with both are b_A+b_{BinA}.

Response generators We adopt the same four models in the mind modules as our response generators. We divide the models into two groups: finetuning and prompting-based. For the finetuning group, we first finetune LLaMA-2-7B-chat and LLaMA-2-13B-chat to generate the next response with the raw training dialogues. Then, we sample 3% of the training data and predict the first and second-order beliefs at each turn using the mind module, which are put into the dialogue context as additional information input to finetune the model again. We choose to combine only a small portion of training data input with beliefs to reduce the API query cost. We also vary the portion to 1%, 3%, and 5%. The sample size does not significantly influence the model performance (See Appendix E). For GPT-3.5 and GPT-4, we use prompts to regulate the conversation. For finetuning-based models, the models are trained on two A6000 GPUs for one epoch with an initial learning rate of 1e-4. The batch size is set to 64. For prompting-based models, we use the OpenAI API for experiments.

4.1 Evaluation and results

For evaluation, we focus on three main questions:

- **Question 1:** Can mind reasoning improve the common ground alignment and negotiation results?
- **Question 2:** Which level of beliefs contributes to the performance gain?
- **Question 3:** What is the relation between belief estimation accuracy and conversation outcomes?

MutualFriend evaluation metrics We adopt the same metrics in He et al. (2017b):

- Success rate (C): how many dialogues where the two agents select the true mutual friend.
- Conversation turns (T): the number of turns the agents take before the end of the conversation
- Success rate per turn (C_T): how efficient the conversation is. We divide the overall success rate by the conversation turns.

CaSiNo evaluation metrics We follow the procedure in Lewis et al. (2017):

- Score-all: the average negotiation scores. The points each agent scores are the number of items times the item's corresponding values. High priority is a value of 5. Medium is 4. Low is 3. If the deal is rejected or the negotiation exceeds the maximum turn, both agents receive 5 points. Since the best outcome should be a win-win situation, we also report the sum over the points of the two agents to compare the overall performance gain.
- Agreed %: the agreement of the deal. A deal is agreed when the agents agree on the proposed deal and the proposal does not exceed the total number of items the agents can distribute.
- Pareto: whether the deal is Pareto Optimal. A solution is Pareto Optimal if neither agent's score can be improved without lowering the other's score.
- Score-agreed: the average negotiation scores in agreed deals.

4.1.1 Observation I: Mind reasoning improves conversation outcomes

First, our experiments compare models' performance without and with mind reasoning. In the cooperative scenario in Table 1, comparing model+w/o mind rows with models, we can see that combining mind modules can significantly improve the alignment success rate in both finetuning and prompting-based models. Among them, GPT-4 performs the best, following LLaMA and GPT-3.5. As for efficiency, models with mind reasoning exhibit higher per-turn success. However, for LLaMA13b,

Models	Mind level	Score-all	Sum	Agreed %	Pareto	Score-agreed	Sum
LLaMA-7B-ft	w/o mind	8.10 vs 7.18	15.28	24.00	12.00	18.33 vs 14.50	32.83
	b_A	12.94 vs 13.48	**26.42**	**68.00**	20.00	16.68 vs 17.47	34.15
	b_{BinA}	11.76 vs 13.36	25.12	56.00	24.00	16.54 vs 19.29	**35.83**
	b_A+b_{BinA}	12.96 vs 12.70	25.66	62.00	**26.00**	17.84 vs 17.42	35.26
LLaMA-13B-ft	w/o mind	15.38 vs 12.68	28.06	70.00	24.00	19.83 vs 15.97	35.80
	b_A	18.02 vs16.14	34.16	92.00	38.00	19.15 vs 17.11	36.26
	b_{BinA}	17.02 vs14.50	31.52	82.00	30.00	19.66 vs 16.59	36.25
	b_A+b_{BinA}	17.36 vs 17.32	**34.68**	**92.00**	**40.00**	18.43 vs 18.39	**36.82**
GPT-3.5	w/o mind	15.00 vs 14.26	29.26	80.00	18.00	17.38 vs 16.57	33.95
	b_A	16.10 vs 17.08	33.18	90.00	22.00	17.22 vs 18.42	35.64
	b_{BinA}	16.72 vs 16.86	**33.58**	**92.00**	22.00	17.74 vs 17.89	35.63
	b_A+b_{BinA}	17.08 vs 15.18	32.26	86.00	**26.00**	19.05 vs 16.72	**35.77**
GPT-4	w/o mind	16.84 vs 16.90	33.74	94.00	8.00	17.60 vs 17.66	35.26
	b_A	16.72 vs 16.50	33.22	90.00	14.00	18.02 vs 17.78	35.80
	b_{BinA}	17.40 vs 16.56	33.96	92.00	12.00	18.17 vs 17.39	35.56
	b_A+b_{BinA}	17.54 vs 17.46	**35.00**	**96.00**	**20.00**	18.06 vs 17.98	**36.04**

Table 2: **CaSiNo: results with different mind settings**. Settings without mind reasoning are marked as w/o mind. Settings considering only the first-order are as b_A, with only the second-order are b_{BinA}, with both are b_A+b_{BinA}.

we notice a longer conversation length, therefore, the efficiency drops below the base model. This suggests that while incorporating belief estimation can elevate success rates, it may not necessarily enhance efficiency if acquiring additional information is needed to establish common ground. More comparison results and discussion can be found in Appendix F.

In the negotiation scenario, as referenced in Table 2, agents utilizing mind reasoning capabilities tend to achieve higher individual scores. Additionally, the collective points of both parties are increased. These agents also are more likely to reach agreements and achieve Pareto Optimal outcomes, suggesting a more strategic distribution of items. When comparing the points scored and the agreement rates across different models, GPT variants and LLaMA-13B display similar performances except that LLaMA-7B falls behind. Notably, LLaMA-13B achieves the highest Pareto Optimal scores, surpassing GPT-4. This may be attributed to GPT-4's tendency to favor equitable item distribution, often resulting in a split like 1 and 1, with another item left unclaimed by either party.

4.1.2 Observation II: Both two levels of belief contribute to the performance gain

Next, we assess the impact of varying belief estimation levels in the mind modules on model performance, as shown in Table 1 and Table 2. First, It is evident that integrating any level of belief estimation leads to performance enhancements compared with the w/o mind baseline, indicating both first and second-order beliefs contribute to the response generation process. Within mind settings, in alignment scenarios, models underscoring the belief differences usually outperform others with single-order belief estimation in LLaMA-13B, GPT-3.5, and GPT-4. In negotiating settings, we first notice that the score-all strongly correlates with the agreed rate, and there is no consistent pattern. Examining Pareto Optimal outcomes, models aggregating both b_A and b_{BinA} tend to distribute items more effectively, resulting in higher Pareto Optimal scores. Similarly, in score-agreed, models combining both two levels of beliefs perform better.

We also notice some fluctuations in the results, for example, LLaMA-7B with only b_{BinA} achieves better results. We reckon that complex and intertwined effects can be exerted when 1) the model is bottlenecked by its context understanding and generation abilities and 2) one or both levels of the belief estimations are not accurate. In general, models need to take into account their own beliefs and also the beliefs of others. Focusing on resolving the differences between them can improve the common ground alignment accuracy and negotiation optimality.

Robustness to prompts In our experimental investigations (prompt templates are supplemented in Appendix B and C), we found that the perfor-

Models	Belief	Precision	F1
LLaMA-7B	b_A	33.00	33.00
	b_{BinA}	30.00	30.00
LLaMA-13B	b_A	36.00	34.00
	b_{BinA}	38.00	33.00
GPT-3.5	b_A	62.00	62.00
	b_{BinA}	70.00	67.00
GPT-4	b_A	77.00	77.00
	b_{BinA}	76.00	76.00

Table 3: **Belief prediction**. The precision and F1 when different models predict the first (b_A) and second-order (b_{BinA}) beliefs.

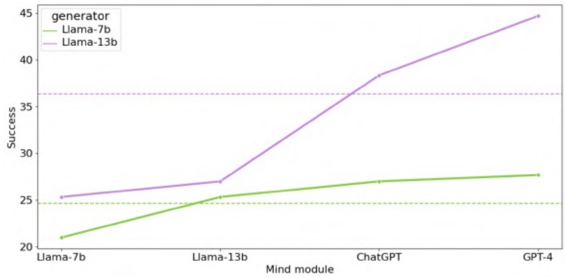

Figure 3: The task success rate when the response generators are paired with different models as the mind modules. The X-axis marks the model name of the mind modules. The Y-axis shows the success rate. Different colors represent different models as the response generators.

mance of belief prediction remains robust when prompts are structured to inquire about the current speaker's solution and their estimation of the other speaker's solution. Our comparison results in Tables 1 and 2 suggest that the task of one-hop prediction, encompassing beliefs and intentions, poses a minimal challenge for most LLMs. For instance, LLaMA-13B exhibits performance akin to GPT-3.5. Consequently, we assert that the primary challenge lies in advancing higher-level ToM inferences within these models.

4.1.3 Observation III: Belief estimation accuracy positively correlates with the alignment success

To more convincingly validate that incorporating the mind reasoning module enhances the models' task performance, we assessed the belief estimation accuracy when different models serve as the mind module in the MutualFriend task using LLaMA-7B, LLaMA-13B, GPT-3.5, and GPT-4. Subsequently, we examined how this accuracy correlates with task success when the four models function as response generators separately, paired with these four mind modules. Here, we demonstrate the relation between the belief estimation accuracy and the dialogue outcomes with MutualFriend task due to its clearly defined belief dynamics. Therefore, the first and second-order beliefs can be easily annotated using predefined rules. The detailed labeling process is included in Appendix A. Table 3 shows the precision and F1 scores for predicting the current speaker's estimation of mutual friend given the current dialogue history b_A, and its estimation of the other speaker's estimation b_{BinA}. The line plot in Figure 3 illustrates the corresponding success rates when a response generator is equipped with different mind modules.

Combining the models' precisions of the belief prediction with the success rates when they serve as the mind modules, we can observe that 1) The success rates increase when the models with higher belief prediction accuracy are served as the mind modules. This trend underscores that the effectiveness of the response generators is closely linked to the mind reasoning capabilities of the respective mind modules; 2) Comparing the growth magnitude of LLaMA-7B and 13b, we can see that LLaMA-7B reaches a flat stage and increases slowly. This suggests that the magnitude of the success rate improvement is bounded by the model's mind-reasoning abilities; 3) The horizontal lines mark the task success rate when the response generators are not augmented with the mind modules. Augmenting models with weaker mind modules can detrimentally impact outcomes due to inaccurate belief predictions and inadequate dialogue reasoning, such as the situation when LLaMA-7B is paired with LLaMA-7B and LLaMA-13B is paired with LLaMA-7B and 13b.

Summarization *vs.* Reasoning It is worth noting that both the first- and second-order belief estimation goes beyond summarization from the last utterance. We carefully annotate part of the beliefs in the dialogue and report the second-order belief prediction accuracy in Table 3, which shows that the LLM can predict the second-order beliefs fairly well.

4.1.4 Human Evaluation

We ask 16 college-level students to play MutualFriend and CaSiNo game with our model. Each subject is randomly assigned 4 samples. S/He chooses one sample to play with the agent w/o mind modules and the other one to play with the agent w/ mind modules. A pair-wise comparison is made

A friend list

hobby	school	name	loc_pref
swimming	national technological university	ryan	outdoor
swimming	rowan college	amy	indoor
sand art	brown university	ryan	outdoor
cricket	alverno college	amy	outdoor
swimming	pensacola christian college	peter	indoor
worldbuilding	rowan college	peter	outdoor

B friend list

hobby	school	name	loc_pref
cricket	university of wisconsin-stout	ryan	indoor
swimming	rowan college	amy	indoor
polo	armstrong state college	jacqueline	indoor
cricket	lynchburg college	lisa	indoor
swimming	arizona western college	kathryn	indoor
cricket	lynchburg college	kathryn	indoor

Hi. Know anyone who went to rowan col.
4 of my friends like cricket.
Do you know peter who went there and like worldbuilding
No peter.
Or any like worldbuilding but arent peter.
…
<select> kathryn lynchburg college cricket indoor. ✗

Hi. I have a friend that likes cricket playing named amy.
(unknown|unknown|unknown|unknown, unknown|unknown|unknown|unknown)
I have a friend named ryan from wisconsin stout that likes cricket.
(cricket|unknown|unknown|indoor, cricket|unknown|amy|unknown)
I do not know anyone with that name.
(cricket|unknown|amy|outdoor, cricket|unknown|unknown|unknown)
I have a friend named amy that likes swimming.
(swimming|unknown|amy|indoor, cricket|unknown|amy|outdoor)
My amy likes swimming and went to rowan college.
(swimming|unknown|amy|indoor, swimming|unknown|amy|indoor)
<select> amy rowan college swimming indoor ✓

Figure 4: **Qualitative comparisons between dialogue generation models** without (at left) and with mind modeling (at right) when agents A and B are figuring out their mutual friend.

between the game outcome when human subjects play with models without and with mind reasoning. In addition, after the game ends, the subjects rate their game partner regarding their cooperativeness (whether the agent is cooperative during the game) and informativeness (whether the agents provide informative responses) from 0 to 10 in the alignment setting; rate regarding their negotiation skills (whether the agent is a good negotiator) and whether they are satisfied with the final deal in the negotiation setting. In addition, we also record their overall enjoyment when playing with the agents in both settings.

From Table 4, we can observe that our model with mind modules can achieve higher outcomes in both MutualFriend and CaSiNo games and the subjects tend to enjoy more in the process. In the cooperative setting, agents without and with mind achieves similar cooperativeness and informativeness rates. However, in the negotiation setting, agents with mind reasoning are shown to be more skillful and can achieve more satisfactory deals.

4.2 Case Study

We demonstrate one MutualFriend example to visualize the difference between LLaMA-7B with and without mind reasoning. Examples of other models and CaSiNo scenarios can be found in the Appendix. As shown in Figure 4, the topics between agents without mind reasoning can diverge quickly. For example, when A asks about "Rowan College", B responds with "cricket" which is unrelated to it. In contrast, for dialogues between agents with step-wise mind reasoning, they resolve the unknown attributes by providing related information (when A talks about "Amy" "swimming", B mentions "Rowan College"). When there is a conflict between the names, A promptly negates "Ryan".

Mutual Friend: alignment				
Groups	Success	Cooperative	Informative	Enjoyment
GPT-3.5 w/o mind	57.14	8.57	9.43	5.29
GPT-3.5 w/ mind	62.50	8.88	9.63	7.63
CaSiNo: negotiation				
Groups	Scores	Skillful	Satisfied	Enjoyment
GPT-3.5 w/o mind	22.50	6.25	6.50	5.75
GPT-3.5 w/ mind	24.50	7.13	7.25	7.25

Table 4: **Human study**. Comparisons are made between our model with mind module vs. models w/o mind module when played with human subjects.

5 Conclusion

In this study, we present MindDial, a novel framework for generating situated dialogue responses for common ground alignment and negotiation. By incorporating the first- and second-order ToM modeling into account, our model can enhance the alignment accuracy and negotiation outcome in both finetuning and prompting-based models. The efficacy of our approach is further substantiated through ablation studies and user feedback.

Limitations

Our prompting design for the mind modules requires a well-defined knowledge and goal. This may limit the generalization abilities of the current framework to more casual conversation scenarios. Also, the task success is highly dependent on the belief estimation precision. Future research is needed to develop and implement mind modules that are both more robust and accurate.

Acknowledgements

The authors thank Ms. Zhen Chen at BIGAI for designing the teaser figure. This work presented herein is supported by the National Science and Technology Major Project (2022ZD0114900) and the National Natural Science Foundation of China (62376031).

References

Carolyn Jane Anderson. 2021. Tell me everything you know: a conversation update system for the rational speech acts framework. In *Proceedings of the Society for Computation in Linguistics 2021*, pages 244–253.

Tim Baarslag, Katsuhide Fujita, Enrico H Gerding, Koen Hindriks, Takayuki Ito, Nicholas R Jennings, Catholijn Jonker, Sarit Kraus, Raz Lin, Valentin Robu, et al. 2013. Evaluating practical negotiating agents: Results and analysis of the 2011 international competition. *Artificial Intelligence*, 198:73–103.

Chris Baker, Rebecca Saxe, and Joshua Tenenbaum. 2011. Bayesian theory of mind: Modeling joint belief-desire attribution. In *Annual Meeting of the Cognitive Science Society (CogSci)*, volume 33.

Cristian-Paul Bara, Sky CH-Wang, and Joyce Chai. 2021. MindCraft: Theory of mind modeling for situated dialogue in collaborative tasks. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 1112–1125, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Simon Baron-Cohen, Alan M Leslie, and Uta Frith. 1985. Does the autistic child have a "theory of mind"? *Cognition*, 21(1):37–46.

Max H Bazerman, Jared R Curhan, Don A Moore, and Kathleen L Valley. 2000. Negotiation. *Annual review of psychology*, 51(1):279–314.

Pieter J Beers, Henny PA Boshuizen, Paul A Kirschner, and Wim H Gijselaers. 2006. Common ground, complex problems and decision making. *Group decision and negotiation*, 15:529–556.

Reinhard Blutner. 2000. Some aspects of optimality in natural language interpretation. *Journal of semantics*, 17(3):189–216.

Manuel Bohn, Michael Henry Tessler, and Michael C Frank. 2019. Integrating common ground and informativeness in pragmatic word learning.

Torben Braüner, Patrick Rowan Blackburn, and Irina Polyanskaya. 2016. Recursive belief manipulation and second-order false-beliefs. In *38th Annual Cognitive Science Society Meeting*, pages 2579–2584. Cognitive Science Society.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in Neural Information Processing Systems (NeurIPS)*, 33:1877–1901.

Brant R Burleson. 2007. Constructivism: A general theory of communication skill. *Explaining communication: Contemporary theories and exemplars*, pages 105–128.

Colin F Camerer, Teck-Hua Ho, and Juin-Kuan Chong. 2004. A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3):861–898.

Kushal Chawla, Jaysa Ramirez, Rene Clever, Gale Lucas, Jonathan May, and Jonathan Gratch. 2021. CaSiNo: A corpus of campsite negotiation dialogues for automatic negotiation systems. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3167–3185, Online. Association for Computational Linguistics.

Justin Chiu, Wenting Zhao, Derek Chen, Saujas Vaduguru, Alexander Rush, and Daniel Fried. 2023. Symbolic planning and code generation for grounded dialogue. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 7426–7436, Singapore. Association for Computational Linguistics.

Herbert H Clark and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition*, 22(1):1–39.

Harmen De Weerd, Rineke Verbrugge, and Bart Verheij. 2013. How much does it help to know what she knows you know? an agent-based simulation study. *Artificial Intelligence*, 199:67–92.

Harmen De Weerd, Rineke Verbrugge, and Bart Verheij. 2015. Higher-order theory of mind in the tacit communication game. *Biologically Inspired Cognitive Architectures*, 11:10–21.

David DeVault, Johnathan Mell, and Jonathan Gratch. 2015. Toward natural turn-taking in a virtual human negotiation agent. In *2015 AAAI Spring Symposium Series*.

Prashant Doshi, Xia Qu, Adam Goodie, and Diana Young. 2010. Modeling recursive reasoning by humans using empirically informed interactive pomdps. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1223–1230.

Benjamin Eysenbach, Carl Vondrick, and Antonio Torralba. 2016. Who is mistaken? *arXiv preprint arXiv:1612.01175*.

Lifeng Fan, Shuwen Qiu, Zilong Zheng, Tao Gao, Song-Chun Zhu, and Yixin Zhu. 2021. Learning triadic belief dynamics in nonverbal communication from videos. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7312–7321.

Kevin A Fischer. 2023. Reflective linguistic programming (rlp): A stepping stone in socially-aware agi (socialagi). *arXiv preprint arXiv:2305.12647*.

Yao Fu, Hao Peng, Tushar Khot, and Mirella Lapata. 2023. Improving language model negotiation with self-play and in-context learning from ai feedback. *arXiv preprint arXiv:2305.10142*.

Sebastian Grueneisen, Emily Wyman, and Michael Tomasello. 2015. "i know you don't know i know…" children use second-order false-belief reasoning for peer coordination. *Child Development*, 86(1):287–293.

Janosch Haber, Tim Baumgärtner, Ece Takmaz, Lieke Gelderloos, Elia Bruni, and Raquel Fernández. 2019. The PhotoBook dataset: Building common ground through visually-grounded dialogue. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1895–1910, Florence, Italy. Association for Computational Linguistics.

Yanlin Han and Piotr Gmytrasiewicz. 2018. Learning others' intentional models in multi-agent settings using interactive pomdps. *Advances in Neural Information Processing Systems (NeurIPS)*, 31.

He He, Anusha Balakrishnan, Mihail Eric, and Percy Liang. 2017a. Learning symmetric collaborative dialogue agents with dynamic knowledge graph embeddings. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1766–1776, Vancouver, Canada. Association for Computational Linguistics.

He He, Anusha Balakrishnan, Mihail Eric, and Percy Liang. 2017b. Learning symmetric collaborative dialogue agents with dynamic knowledge graph embeddings. *Annual Meeting of the Association for Computational Linguistics (ACL)*.

He He, Derek Chen, Anusha Balakrishnan, and Percy Liang. 2018. Decoupling strategy and generation in negotiation dialogues. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2333–2343, Brussels, Belgium. Association for Computational Linguistics.

Mark K Ho, James MacGlashan, Amy Greenwald, Michael L Littman, Elizabeth Hilliard, Carl Trimbach, Stephen Brawner, Josh Tenenbaum, Max Kleiman-Weiner, and Joseph L Austerweil. 2016. Feature-based joint planning and norm learning in collaborative games. In *Annual Meeting of the Cognitive Science Society (CogSci)*.

Kosui Iwasa and Katsuhide Fujita. 2018. Prediction of nash bargaining solution in negotiation dialogue. In *Pacific Rim International Conference on Artificial Intelligence*, pages 786–796. Springer.

Hyunwoo Kim, Melanie Sclar, Xuhui Zhou, Ronan Bras, Gunhee Kim, Yejin Choi, and Maarten Sap. 2023. FANToM: A benchmark for stress-testing machine theory of mind in interactions. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 14397–14413, Singapore. Association for Computational Linguistics.

Jin-Hwa Kim, Nikita Kitaev, Xinlei Chen, Marcus Rohrbach, Byoung-Tak Zhang, Yuandong Tian, Dhruv Batra, and Devi Parikh. 2019. CoDraw: Collaborative drawing as a testbed for grounded goal-driven communication. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 6495–6513, Florence, Italy. Association for Computational Linguistics.

Max Kleiman-Weiner, Mark K Ho, Joseph L Austerweil, Michael L Littman, and Joshua B Tenenbaum. 2016. Coordinate to cooperate or compete: abstract goals and joint intentions in social interaction. In *Annual Meeting of the Cognitive Science Society (CogSci)*.

Michal Kosinski. 2023. Theory of mind may have spontaneously emerged in large language models. *arXiv preprint arXiv:2302.02083*.

Ágnes Melinda Kovács, Ernő Téglás, and Ansgar Denis Endress. 2010. The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, 330(6012):1830–1834.

Mike Lewis, Denis Yarats, Yann Dauphin, Devi Parikh, and Dhruv Batra. 2017. Deal or no deal? end-to-end learning of negotiation dialogues. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2443–2453, Copenhagen, Denmark. Association for Computational Linguistics.

Huao Li, Yu Chong, Simon Stepputtis, Joseph Campbell, Dana Hughes, Charles Lewis, and Katia Sycara. 2023. Theory of mind for multi-agent collaboration via large language models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 180–192, Singapore. Association for Computational Linguistics.

Grégoire Milliez, Matthieu Warnier, Aurélie Clodic, and Rachid Alami. 2014. A framework for endowing an interactive robot with reasoning capabilities about perspective-taking and belief management. In *The 23rd IEEE international symposium on robot and* *human interactive communication*, pages 1103–1109. IEEE.

John F Nash Jr. 1950. The bargaining problem. *Econometrica: Journal of the econometric society*, pages 155–162.

Aida Nematzadeh, Kaylee Burns, Erin Grant, Alison Gopnik, and Tom Griffiths. 2018. Evaluating theory of mind in question answering. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2392–2400, Brussels, Belgium. Association for Computational Linguistics.

Liang Qiu, Yizhou Zhao, Yuan Liang, Pan Lu, Weiyan Shi, Zhou Yu, and Song-Chun Zhu. 2022. Towards socially intelligent agents with mental state transition and human value. In *Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 146–158, Edinburgh, UK. Association for Computational Linguistics.

Neil Rabinowitz, Frank Perbet, Francis Song, Chiyuan Zhang, SM Ali Eslami, and Matthew Botvinick. 2018. Machine theory of mind. In *International Conference on Machine Learning (ICML)*, pages 4218–4227. PMLR.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

Hilary Richardson, Grace Lisandrelli, Alexa Riobueno-Naylor, and Rebecca Saxe. 2018. Development of the social brain from age three to twelve years. *Nature communications*, 9(1):1027.

Damien Sileo and Antoine Lernould. 2023. Mindgames: Targeting theory of mind in large language models with dynamic epistemic modal logic. *arXiv preprint arXiv:2305.03353*.

Michael Tomasello. 2010. *Origins of human communication*. MIT press.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Takuma Udagawa and Akiko Aizawa. 2021. Maintaining common ground in dynamic environments. *Transactions of the Association for Computational Linguistics*, 9:995–1011.

Tomer Ullman. 2023. Large language models fail on trivial alterations to theory-of-mind tasks. *arXiv preprint arXiv:2302.08399*.

Elizabeth S Veinott, Judith Olson, Gary M Olson, and Xiaolan Fu. 1999. Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 302–309.

Adam Vogel, Max Bodoia, Christopher Potts, and Dan Jurafsky. 2013. Emergence of gricean maxims from multi-agent decision theory. In *Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies*, pages 1072–1081.

Michael Wunder, Michael Kaisers, John Robert Yaros, and Michael L Littman. 2011. Using iterated reasoning to predict opponent strategies. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 593–600.

Atsuki Yamaguchi, Kosui Iwasa, and Katsuhide Fujita. 2021. Dialogue act-based breakdown detection in negotiation dialogues. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 745–757, Online. Association for Computational Linguistics.

Tao Yuan, Hangxin Liu, Lifeng Fan, Zilong Zheng, Tao Gao, Yixin Zhu, and Song-Chun Zhu. 2020. Joint inference of states, robot knowledge, and human (false-) beliefs. In *International Conference on Robotics and Automation (ICRA)*, pages 5972–5978. IEEE.

Hongxin Zhang, Weihua Du, Jiaming Shan, Qinhong Zhou, Yilun Du, Joshua B Tenenbaum, Tianmin Shu, and Chuang Gan. 2023. Building cooperative embodied agents modularly with large language models. *arXiv preprint arXiv:2307.02485*.

Pei Zhou, Hyundong Cho, Pegah Jandaghi, Dong-Ho Lee, Bill Yuchen Lin, Jay Pujara, and Xiang Ren. 2022. Reflect, not reflex: Inference-based common ground improves dialogue response quality. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 10450–10468, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Pei Zhou, Aman Madaan, Srividya Pranavi Potharaju, Aditya Gupta, Kevin R McKee, Ari Holtzman, Jay Pujara, Xiang Ren, Swaroop Mishra, Aida Nematzadeh, et al. 2023. How far are large language models from agents with theory-of-mind? *arXiv preprint arXiv:2310.03051*.

Yiheng Zhou, He He, Alan W Black, and Yulia Tsvetkov. 2019. A dynamic strategy coach for effective negotiation. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*, pages 367–378, Stockholm, Sweden. Association for Computational Linguistics.

A MutualFriend belief annotation and evaluation

To test the belief estimation accuracy of our mind modules, we manually label the first and second-order beliefs given the current context of the dialogues. The values mentioned in the current dialogue context are marked as positive (1). The values not mentioned or negated by either of the agents are marked as negative (0). When all the values of one attribute are marked as negative, this attribute becomes "unknown". Figure 5 illustrate one annotation process. For example, when B is asking about "yo-yoing", this value is marked as 1 for b_{BinA} hobby. However, since it does not belong to A's knowledge, for the first-order belief of speaker A, we annotate it as 0. Then, when "yo-yoing" is negated by A, it will be marked as 0 in b_{BinA}. The prediction is a true positive when the model's predicted value of one attribute is annotated as 1, a true negative when both prediction and ground truth are "unknown".

B MutualFriend prompts

→ (At the beginning of the first turn): You are a smart cooperative agent named [Alice|Bob]. You have many friends with different attributes as listed below (the knowledge base of [Alice|Bob]). You are now talking with Bob. He also has a list of friends. You will talk with Bob for a maximum of 20 turns to find out your mutual friend as quickly as possible. You can ask him questions or provide information about your friends. Meanwhile, you should try to mention as few attributes and friends as possible.

hobby, name, location

Surfing, Jane, Outdoor

...

(After each turn - no mind):

→ [Alice|Bob] said: {last generated response}. Please provide your next utterance to [Alice|Bob]:

→ Have you found your mutual friend? If yes, provide this mutual friend in the format of hobby|name|location; If no, respond 'unknown':

(After each turn - with mind):

→ (first-order) Based on the current conversation and your friend table, who do you believe is your mutual friend? Respond in the format of hobby|name|location, and put unknown in the attributes you are not sure about for now:

→ (second-order) Based on the current conversation and your friend table, who do you believe that [Alice|Bob] believes your mutual friend is? Respond in the format of hobby|name|location, and put unknown in the attributes you are not sure about for now:

→ [Alice|Bob] said: {last generated response}. I estimate the mutual friend estimation from your perspective: [first-order] and from [Alice|Bob]'s perspective: [second-order] based on your current talk. To align your estimation and resolve unknown attributes, please provide your next utterance to [Alice|Bob]:

→ Have you found your mutual friend? If yes, provide this mutual friend in the format of hobby|name|location; If no, respond 'unknown':

Figure 6: Template for MutualFriend self-talk prompt.

C CaSiNo prompts

→ (At the beginning of the first turn): You are a smart negotiation agent named [Alice|Bob] planning a camping trip. Besides basic supplies, you will need extra water, food, and firewood. Each of these items will be of either High, Medium, or Low priority for you as shown below. Each of them only has an available quantity of 3 and can only be split using integers. You will negotiate with Bob who will also need these items and have his own value table. Use reasons from your value table to justify why you need these items. Try hard to get as many items as you can!

Item, value, reason

water, high, I didn't pack enough water

...

(After each turn - no mind):

→ [Alice|Bob] said: {last generated response}. Please provide your next utterance to [Alice|Bob]:

→ Based on your conversation with [Alice|Bob], do you want to end the negotiation? Please respond by yes or No:

(After each turn - with mind):

A's Knowledge

Hobby	Name	Time Pref
Drama	Diane	Morning
Tennis	Diane	Evening
...		

B's Knowledge

Hobby	Name	Time Pref
Yo-yoing	Diane	Morning
Drama	Diane	Evening
...		

Context
B: Do either of them like yo-yoing?
A: Nopes. no yo-yoing
B: Any like drama and name is Diane?

Annotated Beliefs

Turn	g_A	g_{BinA}	g_B	g_{AinB}
1	Hobby: unknown Name: unknown Time: unknown	Hobby: yo-ying Name: unknown Time: unknown	-	-
2			Hobby: unknown Name: unknown Time: unknown	Hobby: unknown Name: unknown Time: unknown
3	Hobby: drama Name: Diane Time: unknown	Hobby: drama Name: Diane Time: unknown		

Figure 5: Annotation example

→ (first-order) Based on the current conversation and your value table, how will you split water, firewood, and food? The items each person gets can only be integers and the total quantity for each item is 3. Please use the following format to respond without further explanation: item: the number you get/the number [Alice|Bob] get. For example, water:0/3, firewood:1/2, food: 3/0.

→ (second-order) Based on the current conversation and your value table, how do you think [Alice|Bob] will split water, firewood, and food? The items each person gets can only be integers and the total quantity for each item is 3. Please use the following format to respond without further explanation: item: the number you get/the number [Alice|Bob] get. For example, water:0/3, firewood:1/2, food: 3/0.

→ [Alice|Bob] said: {last generated response}. I estimated the negotiation deal from your perspective: [first-order] and from Bob's perspective: [second-order] based on your current talk. To align your expected deals, please provide your next utterance to [Alice|Bob]:

→ Based on your conversation with [Alice|Bob], do you want to end the negotiation? Please respond by yes or No:

(After negotiation ends):

→ Please provide your proposed deal. The items each person gets can only be integers and the total quantity for each item is 3. Deal with fractions will be rejected. Please use the following format: item: the number you get/the number [Alice|Bob] get. For example, water:0/3, firewood:1/2, food: 3/0.

→ Given your current conversation and the deal proposed by [Alice|Bob]: [deal], will you accept the deal? Please respond by Accept or Reject:

Figure 7: Template for CaSiNo self-talk prompt.

D Finetuning data format

Generate the next response of the dialog based on the given context and knowledge:
(SPEAKER0 as the current speaker)
Estimated [mutual friend|negotiation deal]
[SPEAKER0] [First-order belief]
[SPEAKER1] [Second-order belief]
Knowledge:
Friend table or value table
Dialogues:
[SPEAKER0] ...
[SPEAKER1] ...
— response:

Figure 8: Template for Finetuning

Models	Sample size	Score-all	Sum	Agreed %	Pareto	Score-agreed	Sum
LLaMA-13B-ft	w/o mind	15.38 vs 12.68	28.06	70.00	24.00	19.83 vs 15.97	35.80
	1%	15.36 vs 15.50	30.86	80.00	30.00	18.28 vs 18.46	36.74
	3%	17.36 vs 17.32	34.68	92.00	40.00	18.43 vs 18.39	36.82
	5%	16.44 vs 16.58	33.02	86.00	34.00	18.30 vs 18.47	36.77

Table 5: **CaSiNo: results with different sample sizes**.

Models	Sample size	C	T	C_T
LLaMA-13B-ft	w/o mind	36.33	6.64	5.47
	1%	38.46	8.80	4.37
	3%	44.67	8.85	5.05
	5%	40.33	8.53	4.73

Table 6: **MutualFriend: results with different sample size**.

Models	Mind level	C	C_T
Human	-	82.00	7.00
Rule	-	90.00	5.00
StanoNet	-	78.00	4.00
DynoNet	-	**96.00**	6.00
LLaMA-7B-ft	w/o mind	24.67	2.71
LLaMA-7B-ft	$b_A + b_{BinA}$	28.33	3.20
LLaMA-13B-ft	w/o mind	36.33	5.47
LLaMA-13B-ft	$b_A + b_{BinA}$	44.67	5.05
GPT-3.5	w/o mind	10.67	1.86
GPT-3.5	$b_A + b_{BinA}$	24.33	4.03
GPT-4	w/o mind	75.00	7.71
GPT-4	$b_A + b_{BinA}$	76.00	**8.56**

Table 7: **MutualFriend: comparison with results from original paper**.

E Varing sample size of mind annotation data during finetuning

Considering the computational cost during fine-tuning, we only sample a small partition of dialogue for mind augmentation. In this section, we vary the sample size by 1%, 3% and 5%. From Table 6 and Table 5, we can see that 5% achieves the best results and all models perform better than the w/o mind baselines.

F MutualFriend: more comparison results

In this section, we provide the baseline results of MutualFriend from the original paper in Table 7. It is shown that GPT-4 can achieve higher efficiency with higher accuracy per turn. It is worth noting that the models in the original paper are of smaller sizes and trained with specific datasets while we currently focus more on larger models generalizable to more open-domain tasks. The CaSiNo dataset was originally designed for the strategy prediction task, therefore it did not report generation results.