

ACTIVE CONFUSION EXPRESSION IN LARGE LANGUAGE MODELS: LEVERAGING WORLD MODELS TOWARD BETTER SOCIAL REASONING

006 **Anonymous authors**

007 Paper under double-blind review

ABSTRACT

013 While large language models (LLMs) excel in mathematical and code reasoning,
 014 we observe they struggle with social reasoning tasks, exhibiting cognitive confu-
 015 sion, logical inconsistencies, and conflation between objective world states and
 016 subjective belief states. Through detailed analysis of DeepSeek-R1’s reasoning
 017 trajectories, we find that LLMs frequently encounter reasoning impasses and tend
 018 to output contradictory terms like “tricky” and “confused” when processing sce-
 019 narios with multiple participants and timelines, leading to erroneous reasoning or
 020 infinite loops. The core issue is their inability to disentangle objective reality from
 021 agents’ subjective beliefs. To address this, we propose an adaptive world model-
 022 enhanced reasoning mechanism that constructs a dynamic textual world model
 023 to track entity states and temporal sequences. It dynamically monitors reasoning
 024 trajectories for confusion indicators and promptly intervenes by providing clear
 025 world state descriptions, helping models navigate through cognitive dilemmas.
 026 The mechanism mimics how humans use implicit world models to distinguish be-
 027 tween external events and internal beliefs. Evaluations on three social benchmarks
 028 demonstrate significant improvements in accuracy (e.g., +10% in Hi-ToM) while
 029 reducing computational costs (up to 33.8% token reduction), offering a simple yet
 030 effective solution for deploying LLMs in social contexts.

1 INTRODUCTION

031 With the rapid development of large language models (LLMs), their reasoning capabilities have
 032 achieved significant improvements, particularly in reasoning domains such as mathematics and code
 033 generation. Notable examples include OpenAI’s o1 (Jaech et al., 2024), DeepSeek-R1 (Guo et al.,
 034 2025), Qwen-QWQ (Team, 2024; 2025), and Claude Sonnet 4 (Anthropic, 2025), which demon-
 035 strate substantial knowledge and logical reasoning abilities through extended chain-of-thought
 036 (CoT) (Wei et al., 2022) processes.

037 However, when confronted with social reasoning tasks, LLMs exhibit significant limitations, e.g.,
 038 **cognitive confusion** when processing multiple timelines, **logical inconsistencies** when analyzing
 039 complex character relationships, and **conflation** between **objective world states** and **subjective**
 040 **belief states** in social scenarios with multiple participants. These challenges significantly hinder the
 041 deployment of LLMs in social contexts.

042 Unlike mathematical reasoning, which requires extensive domain knowledge and mathematical
 043 logic, social reasoning necessitates that models comprehend real-world events occurring at dif-
 044 ferent temporal points, disambiguate relationships among multiple involved agents, and establish
 045 connections between agents’ subjective beliefs and objective world states. While current reasoning
 046 LLMs have demonstrated substantial improvements on mathematical reasoning, their social rea-
 047 soning behavior often conflates participants’ subjective beliefs with objective reality, generating
 048 verbose, meaningless, and logically inconsistent CoT, resulting in poor efficiency and low accuracy.

049 Specifically, objective world states represent real-world events that occur physically, e.g., one object
 050 being moved or some agents leaving the room, while subjective belief states represent characters
 051 mental thoughts about other participants (agent or objects). Since each agent can only observe
 052 partial information, their beliefs about unobserved event may become misaligned with reality. A

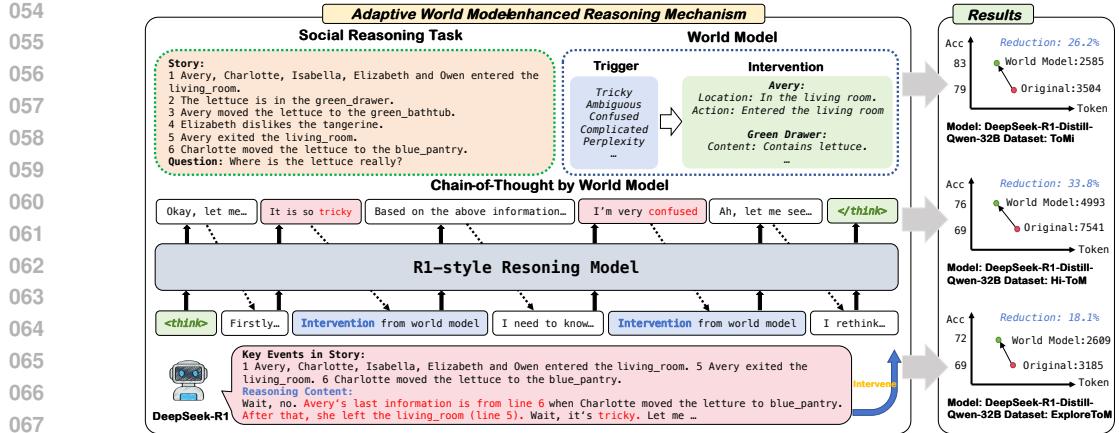


Figure 1: Our adaptive world model-enhanced reasoning mechanism. The system detects confusing words and adaptively use world model to intervene the reasoning trajectories, achieving significant improvements in both accuracy and token efficiency across three social reasoning benchmarks.

canonical example of is demonstrated in Theory of Mind (Premack & Woodruff, 1978): *After an agent leaves the room, they remain unaware of subsequent changes that occur within the room (e.g., the apple being moved from the refrigerator to the basket), thus their subjective beliefs about objects in the room will diverge from objective reality, i.e., they persistently believe the apple remains in the refrigerator.* We find this conflation between reality and beliefs constitutes the primary cause of reasoning failures, particularly in scenarios involving multiple participants and multiple timelines. However, such scenarios are ubiquitous in everyday life.

In this study, we first conduct a detailed analysis of DeepSeek-R1’s reasoning trajectories in social reasoning tasks, investigating the underlying causes of its confusion in their reasoning thoughts. We observe that in the initial stages of reasoning, the model typically can analyze story contexts and clarify character relationships, demonstrating normal cognitive capabilities. However, when the reasoning process involves multiple participants (both agents and objects) and events occurring at different temporal points, the model is prone to falling into cognitive dilemmas, manifesting as confusion and disorientation. Notably, as shown in Figure 1, we find LLMs tend to output **contradictory terms** such as “tricky”, “ambiguous” and “confused” under these circumstances, ultimately leading to erroneous reasoning results or infinite thinking loops.

Through careful analysis of reasoning thoughts at these “confusion” moments, we find the LLMs have fallen into cognitive dilemmas, conflating objective states in the real world with agents’ subjective beliefs, and struggling to disentangle the relationships between them. To address this, we revisit human social cognitive processes. When humans navigate daily interactions, they naturally construct an **implicit world model** to track entity states, character relationships, and temporal sequences of each event. This cognitive model enables humans to easily distinguish and disentangle the relationships between external events and each agent’s internal beliefs, thereby facilitating better understanding of others’ intentions, goals, and emotions for social interaction.

Inspired by this, we propose an adaptive world model-enhanced reasoning mechanism, augmenting social reasoning through the simultaneous construction of a textual world model. As shown in Figure 1, our mechanism establishes a dynamic world model that tracks entities and task states in social events. It continuously monitors LLMs’ reasoning trajectories, providing timely interventions when the model encounters reasoning confusion or cognitive biases. It provides clear world state descriptions to help the model navigate through confusion and break free from reasoning impasses. Our mechanism consists of two components:

- **Trigger mechanism:** The system monitors contradictory words such as “tricky”, “ambiguous” and “confused” in their reasoning trajectory (blue square in Figure 1). When such scenarios are detected, the world state intervention is activated.
- **Intervention process:** Once intervention is triggered, the LLMs immediately halt its previous “confused” reasoning and instead retrieves world states (including entity states, char-

acter states, and timelines) from the latest world model. These states are then inserted after the contradictory terms, guiding LLMs to reflect on their previous reasoning dilemmas and return to correct trajectories.

Similar to how humans construct implicit world models in their brain, this self-constructed world model enables LLMs to re-examine their current thinking state, clarify relationships between characters and entities, and break out of existing reasoning dilemmas. This design effectively improves LLMs' social reasoning accuracy and consistency while reducing token consumption.

Our contributions are threefold: (1) We identified significant issues with LLMs in social reasoning tasks, including cognitive confusion, logical inconsistency, and frequent conflation between objective world states and subjective belief states. (2) We proposed an adaptive world model-enhanced reasoning mechanism that leverages self-constructed textual world models to enable active intervention, helping models overcome reasoning impasses. (3) We conducted comprehensive evaluations on three representative benchmarks ToMi (Le et al., 2019), Hi-ToM (Wu et al., 2023), and Explore-ToM (Sclar et al., 2024), validating the effectiveness of our method in **improving accuracy** and **reducing computational costs**.

2 RELATED WORK

Strategies for Enhancing LLMs’ Cognitive Development in Social Domain To improve the performance of LLMs in social reasoning tasks, existing methods can be broadly categorized into three main directions: (1) Prompt-based Methods, (2) Tool-based Methods, and (3) Model-based Methods. In the field of prompt engineering, SimToM (Wilf et al., 2024) enhances social reasoning capabilities by designing specific prompting strategies that guide models to perform perspective-taking. PercepToM (Jung et al., 2024) focuses on optimizing the conversion process from perceptual information to belief inference through refined contextual information extraction. Question-Analysis Prompting (QAP) (Yugeswardeenoo et al., 2024) guides models to first conduct an in-depth analysis and understanding of the problem. The “problem analysis first” strategy helps models better grasp the key information. Meanwhile, Huang et al. (2024) employs LLMs themselves as world model state trackers to monitor changes in environmental entity positions and character belief processes. Hou et al. (2024) designed specialized belief-solving mechanisms that decompose complex higher-order reasoning tasks into simpler lower-order cognitive problems through set operations in the temporal dimension. MemoRAG (Qian et al., 2025) uses a Memory LLM model to generate answer clues, which then guide retrieval tools to locate relevant passages, demonstrating excellent performance in handling long contexts and complex tasks. As for model-based methods, SymbolicToM (Sclar et al., 2023) introduces symbolic graphical representations to model and track the dynamic changes in character beliefs. AutoToM, MMToM, and MuMA-ToM (Zhang et al., 2025; Shi et al., 2025; Jin et al., 2024) adopt Bayesian inference frameworks, addressing uncertainty issues in social reasoning from a probabilistic modeling perspective. Although existing research has made significant progress in the field of social reasoning, deficiencies remain in dynamic intervention during reasoning. Our method addresses the challenges in social reasoning through cognitive intervention strategies.

Test-time scaling Test-time scaling can be primarily categorized into three aspects: (1) Reward-guided efficient reasoning, (2) Confidence and certainty-based adaptive reasoning, and (3) Consistency-based selective reasoning. Speculative Rejection (Sun et al., 2024) optimizes the computational overhead of Best-of-N decoding by evaluating the quality of partially generated results using reward models. Reward-Guided Speculative Decoding (Liao et al., 2025) improves speculative decoding methods by selectively accepting high-quality results using Process Reward Models (PRM). Meanwhile, Dynamic Parallel Tree Search (Ding et al., 2025) addresses the efficiency issues of tree-based reasoning through parallelization optimization and search-transition mechanisms. FastMCTS (Li et al., 2025) employs Monte Carlo Tree Search to prioritize high-confidence trajectories for deep reasoning, improving multi-step reasoning data synthesis. Length-filtered Vote (Wu et al., 2025) utilizes length-aware majority voting methods, grouping answers according to CoT length. Self-Truncation Best-of-N (Wang et al., 2025b) enhances BoN sampling efficiency by introducing early termination mechanisms, using consistency to measure importance. However, existing research methods focus on post-hoc filtering or early termination of reasoning paths, lacking the ability to dynamically adapt reasoning processes based on real-time feedback and context.

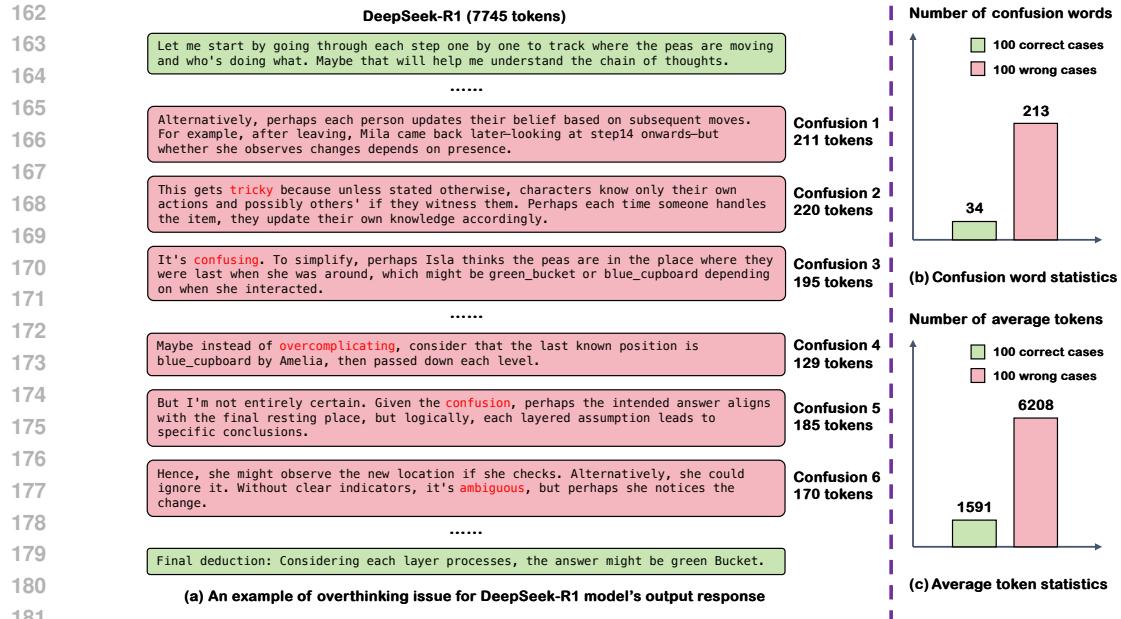


Figure 2: Analysis of DeepSeek-R1’s reasoning dilemmas in social reasoning tasks. (a): Overthinking examples with confusion words. (b) and (c): Statistics of confusion words and average token consumption from 100 correct cases and 100 wrong cases.

ity to proactively inject structured guidance to help models escape cognitive dilemmas. Our method addresses reasoning inefficiency at its source by providing structured contextual intervention.

3 METHOD

3.1 PRELIMINARY

3.1.1 EXPLORATION OF DEESEEK-R1’S REASONING TRAJECTORIES

We first collected DeepSeek-R1’s reasoning trajectories on social reasoning datasets, focusing on the model’s output characteristics when processing complex reasoning tasks. Subsequently, we conducted observational analysis from two perspectives: (1) qualitative analysis of linguistic expressions in the model’s output; (2) quantitative statistics of confusion word counts and average token consumption in correct and wrong cases.

As demonstrated in the case shown in Figure 2(a), we observed a significant phenomenon: when DeepSeek-R1 encounters complex social reasoning scenarios, its output begins to frequently exhibit numerous confusion words such as “tricky” and “confusion”. Furthermore, it will be trapped in a cycle of repeatedly questioning its own reasoning results, with descriptions such as “Alternatively, perhaps...” or “Maybe instead of...”, indicating that the model cannot make clear judgments between objective world states and subjective belief states.

Building on this foundation, we discovered that the more pronounced these phenomena become in DeepSeek-R1, the higher the probability of incorrect answers. Therefore, we selected 100 correct cases and 100 wrong cases to analyze the quantity of confusion words and average token consumption, as shown in Figure 2(b) and (c). The wrong answer cases exhibited significantly higher numbers of confusion words and average token consumption compared to the correct answer cases. This indicates that when facing complex and error-prone problems, the model is more likely to fall into cognitive dilemmas, and such dilemmas require the introduction of targeted interventions to resolve.

3.1.2 ANALYSIS OF CONFUSION WORDS

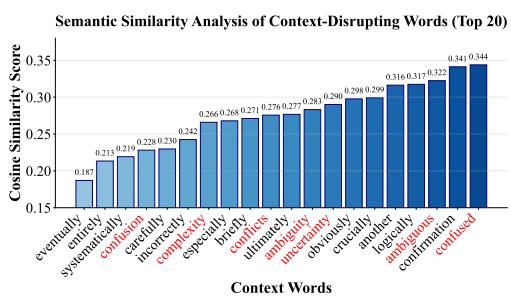


Figure 3: Semantic similarity distribution of candidate interruption words in DeepSeek-R1’s reasoning trajectories.

cates that confusion words not only represent confusion themselves but also increase the semantic differences between the preceding and following context, thus leading to higher error rates.

3.2 ADAPTIVE WORLD MODEL-ENHANCED REASONING WITH COGNITIVE INTERVENTION

Given the issues identified in DeepSeek-R1’s reasoning trajectories as discussed in the previous section, we propose a framework based on adaptive world model-enhanced reasoning with cognitive intervention to address the challenges in social reasoning tasks. First, by employing intervention words to trigger proactive cognitive intervention, timely interrupting the model when it encounters reasoning dilemmas. Second, introducing an adaptive world model that provides character and entity state information to guide the model back to coherent reasoning trajectories. Specifically, our approach begins with intervention word selection based on perplexity analysis, followed by constructing a world model that maintains entity and character states. When intervention words appear during the reasoning process, the system adaptively intervenes and injects state information to help the model distinguish between objective world states and subjective belief states, ultimately achieving globally consistent reasoning capabilities.

3.2.1 COGNITIVE INTERVENTION

In Section 3.1.2, we discovered that confusion words not only represent confusion themselves but also increase the semantic differences between the preceding and following text, ultimately affecting normal reasoning. Therefore, we believe that effective intervention is needed to break this state when reasoning is in similar situations. Wang et al. (2025a) conducted research on where to interrupt reasoning. Based on the above assumption, we compared the perplexity and semantic similarity between the interruption words they mentioned and the aforementioned confusion words. As shown in the Figure 4, confusion words have high perplexity while maintaining low semantic similarity, making them suitable as reasoning interruption points. We name these words as intervention words, as shown in the Table 1.

3.2.2 CONSTRUCTION OF WORLD MODELS

We retain the model’s reasoning trajectory before the intervention words and provide it with the information from the world model to guide its thought process. The world model here refers to a set of structured textual world model that can explicitly describe the temporal causal relationships between objective world states and subjective belief states. It aims to address issues that arise in the social reasoning process, as well as to reduce token redundancy. And finally helps guide the model back to unconfused reasoning trajectories.

Based on the above exploration results, we found that the number of confusion words is positively correlated with the accuracy of reasoning results, so we further investigate the impact that the appearance of confusion words has on the reasoning process. We conducted contextual semantic analysis on each word in the DeepSeek-R1 reasoning trajectory, selecting cosine similarity between preceding and following context as the metric. Then we removed words that obviously do not conform to semantic transitions, and finally selected the top 20 words with the lowest similarity from the remaining words. As shown in Figure 3, we observed many confusion words that frequently appeared in Section 3.1.1, which indicates that confusion words not only represent confusion themselves but also increase the semantic differences between the preceding and following context, thus leading to higher error rates.

Table 1: Keywords Representing Confusion States

Intervention Words
“ambiguous”, “complicating”, “confusion”, “confusing”, “confused”, “perplexity”, “puzzle”, “puzzled”, “puzzling”, “perplexed”, “complication”, “troubled”, “tricky”, “conflicts”, “ambiguity”

As illustrated in Figure 5, the world model maintains a temporal state, dynamically updating the entities and characters state by progressively processing behavioral events in narrative actions. When the model encounters confused thinking and logical inconsistency during the reasoning process, the system adaptively selects corresponding states from the constructed world model for intervention. The state adopts a structured format: “<information> {state from world model} </information>”, which encapsulates the current states of entities and characters. When reasoning difficulties persist, the model will reselect the required corresponding states based on the current reasoning state. We introduce a parameter k as the maximum number of interventions allowed by the world model. This world model aims to provide active leading for reasoning processes in confused states, guiding the model to accurately distinguish between objective world states and subjective belief states, thereby constructing a unified temporal logical framework and achieving globally consistent reasoning capabilities.

After receiving information from world model, the model re-examines previous reasoning trajectory and conducts secondary reasoning, systematically analyzing narrative process, clarifying relationships between characters and entities, thereby forming a clear thinking logic. This effectively corrects the confusion between objective world states and subjective belief states, ultimately returning to the correct reasoning path. Our method not only reduces the number of tokens consumption, but improves the accuracy of the social reasoning significantly.

4 EXPERIMENT

In this section, we evaluate our method across multiple social reasoning tasks, such as ToMi (Le et al., 2019), Hi-ToM (Wu et al., 2023) and Explore-ToM (Sclar et al., 2024).

4.1 IMPLEMENTATION DETAILS

We select three social reasoning datasets: ToMi, Hi-ToM, and ExploreToM. These three datasets encompass social reasoning tasks across multiple levels and demonstrate high representativeness.

Our experiment was implemented on Ubuntu 20.04.4 LTS system environment using the PyTorch framework. We selected multiple mainstream LLMs for performance comparison and evaluation, including DeepSeek-v3 (DeepSeek-AI, 2024), Llama-3.1 (Dubey et al., 2024) series, Qwen2.5 (Yang et al., 2024) series, GPT-5, o1-preview (OpenAI, 2024), DeepSeek-R1, Claude series, and Qwen3 (Qwen Team, 2024) series.

The experiment adopted accuracy and token consumption as primary evaluation metrics. The token consumption statistics include the tokens required for constructing the world model. Model configuration parameters were set as follows: maximum token length was set to 8192, temperature was set to 0.7, and repetition penalty was set to 1.2.

To thoroughly investigate the effectiveness of the adaptive world model-enhanced reasoning mechanism, we employed DeepSeek-R1-Distill-Qwen-32B (Guo et al., 2025), which possesses excellent reasoning capabilities and computational efficiency, as the world model generation model and core

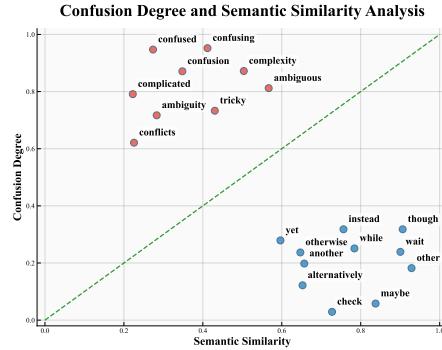


Figure 4: Comparison of confusion degree and semantic similarity between candidates and traditional interruption words.

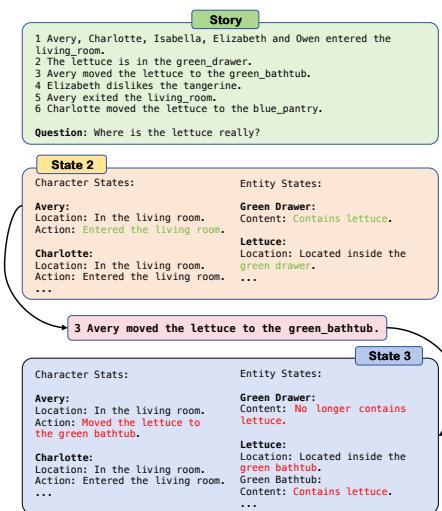


Figure 5: World Model Generation Process

324
 325 Table 2: Performance of Models on Social Reasoning Datasets. For models with dual entries, the
 326 first row shows baseline performance while the second row shows results after incorporating adaptive
 327 world model-enhanced reasoning mechanism.

Model	ToMi		Hi-ToM		ExploreToM	
	Acc	Token	Acc	Token	Acc	Token
<i>Non-reasoning LLMs</i>						
DeepSeek-v3	76.00	726	69.68	2182	59.68	516
Llama-3.1-8B-Instruct-Turbo	64.68	442	58.00	1322	53.33	424
Llama-3.1-70B-Instruct-Turbo	72.00	408	66.33	1244	57.00	336
Llama-3.1-405B-Instruct-Turbo	77.33	356	68.00	1490	68.68	362
Qwen2.5-7B-Instruct-Turbo	60.00	392	45.67	1168	45.00	356
Qwen2.5-32B-Instruct-Turbo	73.67	422	60.33	952	54.67	470
Qwen2.5-72B-Instruct-Turbo	79.67	396	68.68	1548	55.00	402
<i>Reasoning LLMs</i>						
GPT-4o	74.00	1270	70.00	2502	56.68	1412
GPT-5	98.33	834	96.00	1768	93.67	1089
o1-preview	91.33	2866	93.68	6392	82.33	1730
o3-mini	73.00	1762	89.33	5666	74.00	2924
DeepSeek-R1	93.33	2846	77.33	7710	78.00	2124
Claude-3.5-sonnet	88.00	349	83.67	571	75.00	416
Claude-sonnet-4	97.33	597	96.67	892	81.67	668
Qwen3-8B	59.33	-	52.33	-	65.67	-
Qwen3-14B	60.67	-	54.00	-	68.00	-
Qwen3-32B	64.33	-	56.33	-	70.67	-
DeepSeek-R1-Distill-Qwen-7B	55.33	3552	43.33	6080	60.67	4087
	57.33 ^{↑2.00}	2669 _{↓883}	49.33 ^{↑6.00}	4855 _{↓1225}	61.67 ^{↑1.00}	3196 _{↓891}
DeepSeek-R1-Distill-Qwen-14B	70.33	3706	59.00	7889	65.33	3097
	72.33 ^{↑2.00}	3147 _{↓559}	61.33 ^{↑2.33}	6913 _{↓976}	66.33 ^{↑1.00}	2286 _{↓811}
DeepSeek-R1-Distill-Qwen-32B	79.33	3504	69.33	7541	69.67	3185
	83.67 ^{↑4.34}	2585 _{↓919}	76.67 ^{↑7.34}	4993 _{↓2548}	72.67 ^{↑3.00}	2609 _{↓576}
Qwen3-8B-think	74.67	1713	74.67	3795	67.67	2359
	75.33 ^{↑0.66}	1446 _{↓267}	74.67 ^{↑0.00}	3246 _{↓549}	69.33 ^{↑1.66}	1994 _{↓365}
Qwen3-14B-think	76.00	1619	83.33	4269	71.00	2061
	76.67 ^{↑0.67}	1336 _{↓283}	85.00 ^{↑1.67}	3615 _{↓654}	71.33 ^{↑0.33}	1551 _{↓510}
Qwen3-32B-think	81.33	1942	88.00	3316	74.33	2193
	81.33 ^{↑0.00}	1887 _{↓55}	88.67 ^{↑0.67}	2840 _{↓476}	76.00 ^{↑1.67}	1805 _{↓388}

362 experimental model to systematically explore the impact of different threshold k parameters, alter-
 363 native intervention words settings, and various methods on model performance.

366 4.2 MAIN RESULTS

368 We test the performance on both non-reasoning and reasoning models above in three social reasoning
 369 datasets. Moreover, we evaluated the performance of DeepSeek-R1-Distill series and Qwen3 series
 370 models with adaptive world model-enhanced reasoning mechanism, to comprehensively analyze the
 371 impact of our method on social reasoning capabilities.

372 **Reasoning LLMs significantly outperformed non-reasoning LLMs across all social reasoning**
 373 **datasets** As shown in Table 2, on the ToMi dataset, reasoning LLMs achieved a maximum accuracy
 374 of 98.33% (GPT-5), while the best performance of non-reasoning LLMs was only 79.67%, repre-
 375 senting a performance gap of 18.66 percentage points. On the more challenging Hi-ToM dataset,
 376 this gap further expanded to 26.32 percentage points (96.00% vs 69.68%). These results demon-
 377 strate that reasoning mechanisms based on long chain-of-thought can significantly enhance model
 378 performance on complex social reasoning tasks.

378 **Adaptive world model-enhanced reasoning mechanism shows significant accuracy improvement** For the DeepSeek-R1-Distill series models, the 7B version improved from 55.33% to 57.33%
 379 (+2.00) on the ToMi dataset, the 14B version improved from 70.33% to 72.33% (+2.00), and the 32B
 380 version improved from 79.33% to 83.67% (+4.34). This indicates that the adaptive world model-
 381 enhanced reasoning mechanism can effectively enhance models' social reasoning capabilities, with
 382 more significant improvement effects observed for larger-scale models.

383 **Adaptive world model-enhanced reasoning mechanism reduces token consumption** DeepSeek-
 384 R1-Distill-Qwen-32B achieved token reductions of 26.2%, 33.8%, and 18.1% respectively. Simi-
 385 larly, the Qwen3-think LLMs also demonstrated varying degrees of token consumption reduction.
 386

388 4.3 ANALYSIS OF MAXIMUM INTERVENTION COUNT

390 Table 3: Performance under different maximum intervention count k .
 391

392 Model	393 Threshold	394 ToMi		395 Hi-ToM		396 ExploreToM	
		397 Acc	398 Token	399 Acc	400 Token	401 Acc	402 Token
395 DeepSeek-R1 396 Distill Qwen 397 -32B	398 Baseline	399 79.33	400 3504	401 69.33	402 7541	403 69.67	404 3185
	405 $k = 1$	406 80.00 ^{+0.67}	407 3196 _{↓308}	408 70.67 ^{↑1.34}	409 6731 _{↓810}	410 70.33 ^{+0.66}	411 2971 _{↓214}
	412 $k = 2$	413 81.67 ^{+2.34}	414 2875 _{↓629}	415 73.33 ^{↑4.00}	416 6119 _{↓1422}	417 71.00 ^{↑1.33}	418 2554 _{↓631}
	419 $k = 3$	420 83.67 ^{+4.34}	421 2585 _{↓919}	422 76.67 ^{↑7.34}	423 4993 _{↓2548}	424 72.67 ^{↑3.00}	425 2609 _{↓576}
	426 $k = 4$	427 83.67 ^{+4.34}	428 2585 _{↓919}	429 77.67 ^{↑8.34}	430 4149 _{↓3392}	431 73.33 ^{↑3.66}	432 2360 _{↓825}
	433 $k = 5$	434 83.67 ^{+4.34}	435 2585 _{↓919}	436 79.33 ^{↑10.00}	437 3421 _{↓4120}	438 73.33 ^{↑3.66}	439 2360 _{↓825}

401 The experimental results in Table 3 reveal important characteristics of the adaptive world model-
 402 enhanced reasoning mechanism. As the maximum intervention count k increases, the model demon-
 403 strates varying degrees of accuracy improvement across all datasets. This consistent improvement
 404 indicates that increasing the frequency of world model intervention can provide LLM with richer
 405 social cognitive information, thereby enhancing its reasoning capabilities.

406 **More complex social reasoning tasks require more world model interventions** For the relatively
 407 simple ToMi dataset, performance improvement peaked at $k=3$ (83.67%) and remained stable there-
 408 after, exhibiting clear performance saturation. However, on the more challenging Hi-ToM dataset,
 409 performance continued to improve throughout the entire k value range. This phenomenon indicates
 410 that complex social reasoning tasks can more effectively utilize frequent world model interventions.

411 **Appropriate interventions improves efficiency** On the three datasets, high k value settings not
 412 only improved accuracy but also significantly reduced token consumption. This "dual optimization"
 413 effect demonstrates that the adaptive world model-enhanced reasoning mechanism actually enhances
 414 the model's reasoning efficiency by providing more guidance.

416 4.4 ANALYSIS OF INTERVENTION WORDS CATEGORIES

418 DeepSeek-R1-Distill-Qwen-32B achieved significant performance improvements with our selected
 419 intervention words. To compare the impact of intervention words selection, We compared several
 420 categories of commonly used intervention words described in Wan et al. (2025): Pause-Validation
 421 (referred to as PV) and Branch-Extension (referred to as BE) as comparisons.

- 423 • **Pause-Validation:** wait, check, make sure, hold on, verify, let me see, confirm, ensure,
 424 evaluate, examine.
- 425 • **Branch-Extension:** alternatively, another, instead, however, while, yet, though, rather,
 426 otherwise, on the other hand.

428 As shown in Table 4, our selected intervention words (Ours) achieve optimal performance in both ac-
 429 curacy and token efficiency across all datasets compared to other common intervention words (+PV
 430 and +BE). This indicates that compared to choosing simple pause-validation or branch-extension
 431 words for interruption, our method demonstrates good methodological stability and applicability,
 which also validates the conclusions in section 3.2.1.

432 Table 4: Performance of DeepSeek-R1-Distill-Qwen-32B with different intervention words. PV
 433 denotes pause-validation words, and BE indicates branch-extension words.

435 Model	436 Method	437 ToMi		438 Hi-ToM		439 ExploreToM	
		440 Acc	441 Token	442 Acc	443 Token	444 Acc	445 Token
446 DeepSeek-R1 447 Distill Qwen 448 -32B	449 Baseline	450 79.33	451 3504	452 69.33	453 7541	454 69.67	455 3185
	+PV	82.33	^{↑3.00} _{↓674}	72.67	^{↑3.34} _{↓1482}	70.00	^{↑0.33} _{↓214}
	+BE	81.67	^{↑2.34} _{↓701}	74.00	^{↑4.67} _{↓1791}	71.33	^{↑1.66} _{↓404}
	Ours	83.67	^{↑4.34} _{↓919}	76.67	^{↑7.34} _{↓2548}	72.67	^{↑3.00} _{↓576}

449 4.5 ANALYSIS OF OTHER METHODS

450 To validate the effectiveness of our method in improving reasoning performance, we compare our
 451 method with several approaches for enhancing model reasoning capabilities, including Chain-of-
 452 Thought (CoT), Tree of Thoughts (ToT)(Yao et al., 2023a), Reasoning and Acting (ReAct) (Yao
 453 et al., 2023b), and Reflexion (Shinn et al., 2023) for comparative evaluation.

454 Table 5: Comparison with different prompting strategy on social reasoning tasks.

455 Model	456 Method	457 ToMi		458 Hi-ToM		459 ExploreToM	
		460 Acc	461 Token	462 Acc	463 Token	464 Acc	465 Token
466 DeepSeek-R1 467 Distill Qwen 468 -32B	469 Baseline	470 79.33	471 3504	472 69.33	473 7541	474 69.67	475 3185
	+CoT	79.67	^{↑0.34} _{↓48}	70.00	^{↑0.67} _{↓157}	70.00	^{↑0.33} _{↓29}
	+ToT	80.33	^{↑1.00} _{↓252}	70.67	^{↑1.34} _{↓582}	70.67	^{↑1.00} _{↓236}
	+ReAct	81.00	^{↑1.67} _{↓306}	70.67	^{↑1.34} _{↓307}	69.67	^{↑0.00} _{↓71}
	+Reflexion	80.67	^{↑1.34} _{↓428}	71.33	^{↑2.00} _{↓854}	70.33	^{↑0.66} _{↓251}
	Ours	83.67	^{↑4.34} _{↓919}	76.67	^{↑7.34} _{↓2548}	72.67	^{↑3.00} _{↓576}

476 As shown in the Table 5, our method achieves significant performance improvements over other
 477 methods across all datasets. This demonstrates that our approach can more effectively handle the
 478 cognitive confusion in social reasoning tasks, and achieve dual improvements in both performance
 479 and efficiency.

480 5 CONCLUSIONS, LIMITATIONS AND FUTURE WORKS

481 In this paper, we identified and addressed critical limitations of current reasoning LLMs in social
 482 reasoning tasks. Through detailed analysis of DeepSeek-R1’s reasoning trajectories, we discovered
 483 that models are prone to cognitive confusion, logical inconsistencies, and conflation between objec-
 484 tive world states and subjective belief states when processing complex social scenarios. We reveals
 485 that LLMs frequently fall into cognitive dilemmas when encountering contradictory words such as
 486 “confused” and “ambiguous”, leading to reasoning errors or multiple loops. To address this issue,
 487 we proposed an adaptive world model-enhanced reasoning mechanism comprising two core compo-
 488 nents: a trigger mechanism and an intervention process. This mechanism effectively helps models
 489 distinguish between objective world states and subjective belief states, significantly improving so-
 490 cial reasoning accuracy and consistency on three social reasoning benchmarks ToMi, Hi-ToM, and
 491 ExploreToM. Additionally, some limitations and potential future works are listed as follows:

- 492 • **Trigger Mechanism Research** Our current trigger mechanism relies on predefined inter-
 493 vention words, potentially missing other forms of cognitive confusion expressions. Future
 494 research should develop adaptive trigger mechanisms to automatically identify confusion
 495 states without predefined word lists.
- 496 • **Experiments in Multiple LLMs and Generalizability** Our experiments primarily focus
 497 on theory-of-mind benchmarks and DeepSeek-R1 series LLMs. Future work may conduct
 498 extensive evaluations across different reasoning LLMs to establish thorough intervention
 499 strategies and expand to broader social reasoning tasks such as moral reasoning or social
 500 norm understanding.

486 ETHICS STATEMENT
487488 Our research follows ethical guidelines for artificial intelligence research and only uses publicly
489 available academic benchmark datasets (ToMi, Hi-ToM, and ExploreToM), involving no human
490 subjects. Our analysis of DeepSeek-R1’s reasoning trajectories is purely for scientific research pur-
491 poses, aimed at improving social reasoning capabilities of LLMs. Our use of LLMs is entirely
492 reasonable and complies with academic research standards, with the paper containing no personal
493 privacy information, emotional manipulation content, or biased materials. All authors of this re-
494 search declare no conflicts of interest regarding this submission, and we openly share our research
495 methodology and clearly disclose the limitations of our approach to encourage deeper future re-
496 search. All procedures and methods in this research comply with ICLR’s ethical standards and
497 academic publication requirements.
498499 REPRODUCIBILITY STATEMENT
500501 We have made extensive efforts to ensure the reproducibility of our work. The datasets used for our
502 experiment are explicitly described in Appendix A.2. For the methodology, we provide complete
503 world model design and implementation procedure. Maximum token length, temperature and other
504 parameters, are thoroughly documented in the paper. In the supplementary materials, we also submit
505 the source code and datasets required to reproduce our method.
506507 REFERENCES
508509 Anthropic. Introducing claude 4, May 2025. URL <https://www.anthropic.com/news/claude-4>. Company Blog Post.
510511 DeepSeek-AI. Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*, 2024.
512513 Yifu Ding, Wentao Jiang, Shunyu Liu, Yongcheng Jing, Jinyang Guo, Yingjie Wang, Jing Zhang,
514 Zengmao Wang, Ziwei Liu, Bo Du, Xianglong Liu, and Dacheng Tao. Dynamic parallel tree
515 search for efficient llm reasoning. *arXiv preprint arXiv:2502.16235*, 2025.
516517 Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, et al. The llama 3 herd of models. *arXiv
518 preprint arXiv:2407.21783*, 2024.
519520 Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu,
521 Shirong Ma, Peiyi Wang, Xiao Bi, et al. Deepseek-r1: Incentivizing reasoning capability in llms
522 via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025.
523524 Guiyang Hou, Wenqi Zhang, Yongliang Shen, Linjuan Wu, and Weiming Lu. Timetom: Temporal
525 space is the key to unlocking the door of large language models’ theory-of-mind. In *Findings of
526 the Association for Computational Linguistics ACL 2024*, pp. 11532–11547, 2024.
527528 X Huang, Emanuele La Malfa, Samuele Marro, Andrea Asperti, Anthony Cohn, and Michael
529 Wooldridge. A notion of complexity for theory of mind via discrete world models. In *Findings of
530 the Association for Computational Linguistics: EMNLP 2024*, pp. 2964–2983, 2024.
531532 Aaron Jaech, Adam Kalai, Adam Lerer, Adam Richardson, Ahmed El-Kishky, Aiden Low, Alec
533 Helyar, Aleksander Madry, Alex Beutel, Alex Carney, et al. Openai o1 system card. *arXiv
534 preprint arXiv:2412.16720*, 2024.
535536 Chuanyang Jin, Yutong Wu, Jing Cao, Jiannan Xiang, Yen-Ling Kuo, Zhiting Hu, Tomer Ullman,
537 Antonio Torralba, Joshua Tenenbaum, and Tianmin Shu. Mmtom-qa: Multimodal theory of mind
538 question answering. In *Proceedings of the 62nd Annual Meeting of the Association for Compu-
539 tational Linguistics (Volume 1: Long Papers)*, pp. 16077–16102, 2024.
540541 Chani Jung, Dongkwan Kim, Jiho Jin, Jiseon Kim, Yeon Seonwoo, Yejin Choi, Alice Oh, and
542 Hyunwoo Kim. Perceptions to beliefs: Exploring precursory inferences for theory of mind in
543 large language models. In *Proceedings of the 2024 Conference on Empirical Methods in Natural
544 Language Processing*, pp. 19794–19809, 2024.
545

540 Matthew Le, Y-Lan Boureau, and Maximilian Nickel. Revisiting the evaluation of theory of mind
 541 through question answering. In *Proceedings of the 2019 Conference on Empirical Methods in*
 542 *Natural Language Processing and the 9th International Joint Conference on Natural Language*
 543 *Processing (EMNLP-IJCNLP)*, pp. 5872–5877, 2019.

544

545 Peiji Li, Kai Lv, Yunfan Shao, Yichuan Ma, Linyang Li, Xiaoqing Zheng, Xipeng Qiu, and Qipeng
 546 Guo. Fastmcts: A simple sampling strategy for data synthesis. *arXiv preprint arXiv:2502.11476*,
 547 2025.

548 Baohao Liao, Yuhui Xu, Hanze Dong, Junnan Li, Christof Monz, Silvio Savarese, Doyen Sahoo, and
 549 Caiming Xiong. Reward-guided speculative decoding for efficient llm reasoning. *arXiv preprint*
 550 *arXiv:2501.19324*, 2025.

551

552 OpenAI. Learning to reason with llms. <https://openai.com/index/learning-to-reason-with-llms/>, 2024.

553

554 David Premack and Guy Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and*
 555 *brain sciences*, 1(4):515–526, 1978.

556

557 Hongjin Qian, Zheng Liu, Peitian Zhang, Kelong Mao, Defu Lian, Zhicheng Dou, and Tiejun
 558 Huang. Memorag: Boosting long context processing with global memory-enhanced retrieval
 559 augmentation. In *Proceedings of the ACM on Web Conference 2025*, WWW ’25, pp. 2366–2377,
 560 New York, NY, USA, 2025. Association for Computing Machinery. ISBN 9798400712746. doi:
 561 10.1145/3696410.3714805. URL <https://doi.org/10.1145/3696410.3714805>.

562

563 Qwen Team. Qwen3 technical report. *arXiv preprint arXiv:2409.12191*, 2024.

564

565 Melanie Sclar, Sachin Kumar, Peter West, Alane Suhr, Yejin Choi, and Yulia Tsvetkov. Minding
 566 language models’(lack of) theory of mind: A plug-and-play multi-character belief tracker. In
 567 *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume*
 568 *1: Long Papers)*, pp. 13960–13980, 2023.

569

570 Melanie Sclar, Jane Yu, Maryam Fazel-Zarandi, Yulia Tsvetkov, Yonatan Bisk, Yejin Choi, and Asli
 571 Celikyilmaz. Explore theory of mind: Program-guided adversarial data generation for theory of
 572 mind reasoning. *arXiv preprint arXiv:2412.12175*, 2024.

573

574 Haojun Shi, Suyu Ye, Xinyu Fang, Chuanyang Jin, Leyla Isik, Yen-Ling Kuo, and Tianmin Shu.
 575 Muma-tom: Multi-modal multi-agent theory of mind. In *Proceedings of the AAAI Conference on*
 576 *Artificial Intelligence*, volume 39, pp. 1510–1519, 2025.

577

578 Noah Shinn, Federico Cassano, Edward Berman, Ashwin Gopinath, Karthik Narasimhan, and
 579 Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. In *Advances in*
 580 *Neural Information Processing Systems (NeurIPS)*, 2023. URL <https://arxiv.org/abs/2303.11366>.

581

582 Hanshi Sun, Momin Haider, Ruiqi Zhang, Huitao Yang, Jiahao Qiu, Ming Yin, Mengdi Wang, Peter
 583 Bartlett, and Andrea Zanette. Fast best-of-n decoding via speculative rejection. *arXiv preprint*
 584 *arXiv:2410.20290*, 2024.

585

586 Qwen Team. Qwq: Reflect deeply on the boundaries of the unknown, November 2024. URL
 587 <https://qwenlm.github.io/blog/qwq-32b-preview/>.

588

589 Qwen Team. Qwq-32b: Embracing the power of reinforcement learning, March 2025. URL
 590 <https://qwenlm.github.io/blog/qwq-32b/>.

591

592 Xu Wan, Wei Wang, Wenyue Xu, Wotao Yin, Jie Song, and Mingyang Sun. Adapthink: Adaptive
 593 thinking preferences for reasoning language model. *arXiv preprint arXiv:2506.18237*, 2025.

594

595 Chenlong Wang, Yuanning Feng, Dongping Chen, Zhaoyang Chu, Ranjay Krishna, and Tianyi
 596 Zhou. Wait, we don’t need to “wait”: removing thinking tokens improves reasoning efficiency.
 597 *arXiv preprint arXiv:2506.08343*, 2025a.

594 Yiming Wang, Pei Zhang, Siyuan Huang, Baosong Yang, Zhuosheng Zhang, Fei Huang, and Rui
 595 Wang. Sampling-efficient test-time scaling: Self-estimating the best-of-n sampling in early de-
 596 coding. *arXiv preprint arXiv:2503.01422*, 2025b.

597

598 Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V
 599 Le, Denny Zhou, et al. Chain-of-thought prompting elicits reasoning in large lan-
 600 guage models. *Advances in Neural Information Processing Systems*, 35:24824–24837,
 601 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/9d5609613524ecf4f15af0f7b31abca4-Abstract-Conference.html.

603

604 Alex Wilf, Sihyun Lee, Paul Pu Liang, and Louis-Philippe Morency. Think twice: Perspective-
 605 taking improves large language models’ theory-of-mind capabilities. In *Proceedings of the 62nd*
 606 *Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp.
 607 8292–8308, 2024.

608

609 Yufan Wu, Yinghui He, Yilin Jia, Rada Mihalcea, Yulong Chen, and Naihao Deng. Hi-tom: A
 610 benchmark for evaluating higher-order theory of mind reasoning in large language models. In
 611 *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 10691–10706,
 612 2023.

613

614 Yuyang Wu, Yifei Wang, Tianqi Du, Stefanie Jegelka, and Yisen Wang. When more is less: Under-
 615 standing chain-of-thought length in llms. *arXiv preprint arXiv:2502.07266*, 2025.

616

617 An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li,
 618 Dayiheng Liu, Fei Huang, Haoran Wei, et al. Qwen2. 5 technical report. *arXiv preprint*
arXiv:2412.15115, 2024.

619

620 Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L. Griffiths, Yuan Cao, and Karthik
 621 Narasimhan. Tree of thoughts: deliberate problem solving with large language models. In *Pro-
 622 ceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS
 623 ’23, Red Hook, NY, USA, 2023a. Curran Associates Inc.

624

625 Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao.
 626 ReAct: Synergizing reasoning and acting in language models. In *International Conference on
 627 Learning Representations (ICLR)*, 2023b. URL <https://arxiv.org/abs/2210.03629>.

628

629 Dharunish Yugeswardeenoo, Kevin Zhu, and Sean O’Brien. Question-analysis prompting im-
 630 proves LLM performance in reasoning tasks. In Xiyuan Fu and Eve Fleisig (eds.), *Proceed-
 631 ings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 4:
 632 Student Research Workshop)*, pp. 402–413, Bangkok, Thailand, August 2024. Association for
 633 Computational Linguistics. ISBN 979-8-89176-097-4. doi: 10.18653/v1/2024.acl-srw.45. URL
<https://aclanthology.org/2024.acl-srw.45/>.

634

635 Zhining Zhang, Chuanyang Jin, Mung Yao Jia, and Tianmin Shu. Autotom: Automated
 636 bayesian inverse planning and model discovery for open-ended theory of mind. *arXiv preprint*
arXiv:2502.15676, 2025.

637

638

639 **A APPENDIX**

640

641 **A.1 THE USE OF LARGE LANGUAGE MODELS (LLMs)**

642

643 In compliance with the ICLR 2026 disclosure requirements on language model usage, we confirm
 644 that the use of LLMs in this study was strictly limited to linguistic refinement. Specifically, they
 645 were employed to improve syntactic structure, enhance academic style, standardize terminology, and
 646 increase the readability of technical content, thereby facilitating clearer scientific communication.
 647 Importantly, LLMs were not involved in generating research ideas, designing methodologies, or
 contributing to scientific conclusions; these aspects were carried out solely by the authors.

648
649

A.2 DATASETS

650

ToMi The dataset is a social reasoning dataset that describes characters' social states in simple scenarios, containing extensive scene descriptions and character belief state descriptions.

652

HiToM The scene descriptions and character belief state descriptions are more complex, involving deeper cognitive processes and higher-order reasoning.

654

ExploreToMi The dataset is more aligned with realistic scenarios, featuring more diverse character behaviors and incorporating scene descriptions such as weather conditions, atmospheric changes, and other objective influencing factors.

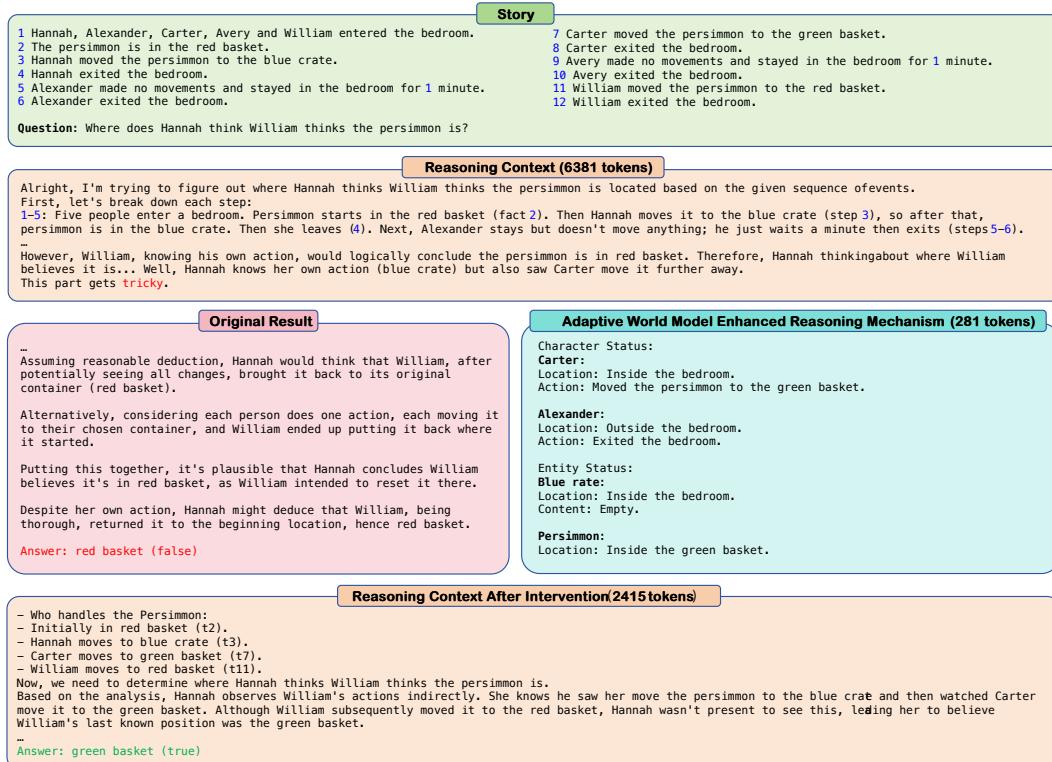
658
659

A.3 CASE STUDY

660

In the Figure 6, we present a complete reasoning case to demonstrate how the model's reasoning becomes trapped in confusion and how our method achieves the correct answer while significantly reducing token consumption. Meanwhile, the tokens required to construct the textual world model are minimal, with essentially no impact on the total reasoning token consumption, so we include them in our token count statistics.

665



684

Figure 6: Samples of reasoning trajectories with and without adaptive world model-enhanced reasoning mechanism

690

A.4 SAMPLE EXAMPLES FROM DATASET

693

694

For the data sources we use, we present sample examples from Figure 7 to Figure 9. Figure 7 describes characters' social states in simple scenarios, containing extensive scene descriptions and character belief state descriptions. Figure 8 presents a sample example from the HiToM dataset, where the Story consists of multiple and complex descriptions, with its deeper cognitive processes and higher-order reasoning. Figure 9 presents a sample example from the ExploreToM dataset, whose Story is more aligned with realistic scenarios, including some advanced socio-cognitive events, such as "told privately", "in secret", "got distracted", etc.

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

ToMi**Story:**

1 **Isla** entered the porch.
 2 **Isabella** entered the living_room.
 3 **Jack** entered the porch.
 4 **Jack** loves the strawberry.
 5 The tangerine is in the basket.
 6 The basket is in the porch.
 7 **Isla** loves the trousers.
 8 **Jack** moved the tangerine to the suitcase.
 9 The suitcase is in the porch.
 10 **Isla** exited the porch.
 11 **Jack** exited the porch.
 12 **Isla** entered the porch.

Question:Where does **Isla** think that **Jack** searches for the tangerine?**Choices:**

basket, suitcase

Figure 7: Sample example from dataset ToMi

Hi-ToM**Story:**

1 **Hannah, Alexander, Carter, Avery** and **William** entered the bedroom.
 2 The persimmon is in the red_basket.
 3 **Hannah** moved the persimmon to the blue_crate.
 4 **Hannah** exited the bedroom.
 5 **Alexander** made no movements and stayed in the bedroom for 1 minute.
 6 **Alexander** exited the bedroom.
 7 **Carter** moved the persimmon to the green_basket.
 8 **Carter** exited the bedroom.
 9 **Avery** made no movements and stayed in the bedroom for 1 minute.
 10 **Avery** exited the bedroom.
 11 **William** moved the persimmon to the red_basket.
 12 **William** exited the bedroom.
 13 **Hannah, Alexander, Carter, Avery** and **William** entered the waiting_room.
 14 **William, Alexander** and **Hannah** entered the bathroom.
 15 The banana is in the red_bucket.
 16 **William** moved the banana to the red_crate.
 17 **William** exited the bathroom.
 18 **Alexander** made no movements and stayed in the bathroom for 1 minute.
 19 **Alexander** exited the bathroom.
 20 **Hannah** made no movements and stayed in the bathroom for 1 minute.
 21 **Hannah** exited the bathroom.
 22 **William, Alexander** and **Hannah** entered the waiting_room.
 23 **Alexander, Avery, William** and **Hannah** entered the garage.
 24 The plum is in the red_drawer.
 25 **Alexander** moved the plum to the green_cupboard.
 26 **Alexander** exited the garage.
 27 **Avery** moved the plum to the blue_treasure_chest.

Question:Where does **Hannah** think **William** thinks the persimmon is?**Choices:**

red_crate, red_bucket, red_box, blue_box, green_box, red_basket, blue_crate, green_bathtub, green_envelope, green_basket, red_drawer, green_bottle, blue_bathtub, blue_treasure_chest, green_cupboard.

Figure 8: Sample example from dataset Hi-ToM

751

752

753

754

755

756

ExploreToM

757

Story:

758

1 **Lucas** entered the festival merchandise booth.
 2 **Lucas** moved the portable speaker to the duffel bag, which is also located in the festival merchandise booth.
 3 **Lucas** moved the portable speaker to the main information tent *in secret*, leaving the duffel bag in its original location.
 4 **Lucas** entered the festival merchandise booth.
 5 **Lucas** left the festival merchandise booth.
 6 **Lucas** told *privately* to **Danielle** that the duffel bag is in the festival merchandise booth.
 7 **Danielle** told *privately* to **Alexis** that the duffel bag is in the festival merchandise booth.
 8 **Danielle** entered the main information tent.
 9 **Lucas** told *privately* to **Danielle** that the portable speaker is in the main information tent.
 10 **Lucas** told *privately* to **Alexis** that the portable speaker is in the main information tent.
 11 **Lucas** told *privately* to **Alexis** that the duffel bag is in the festival merchandise booth.
 12 **Danielle** moved the portable speaker to the duffel bag, while **Lucas** *got distracted* and did not realize.

766

Question:

767

In which room does **Alexis** think that **Lucas** will search for the duffel bag?

768

769

Figure 9: Sample example from dataset ExploreToM

770

771

A.5 SAMPLE EXAMPLE OF WORLD MODEL GENERATION

772

To better illustrate the effectiveness of adaptive world model-enhanced reasoning mechanism, we present a comprehensive set of examples demonstrating textual world model generation, as illustrated in the Figure 10. The world model maintains dynamic tracking of character and entity states, which are continuously updated in response to different actions and event. This dynamic updating capability enables our method to flexibly adapt to diverse story scenarios, ensuring that during the intervention process, the system can provide accurate information corresponding to the current state of the narrative.

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

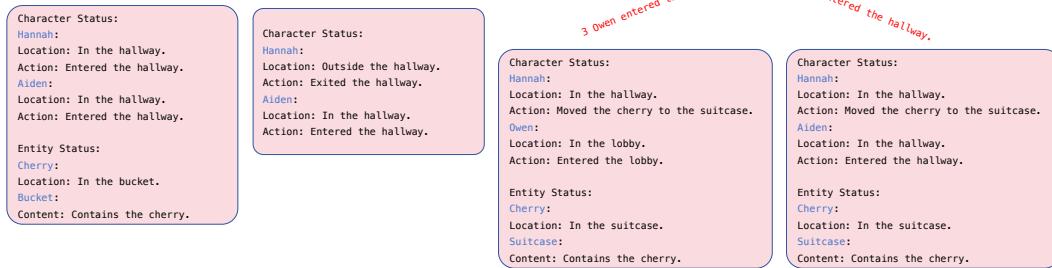


Figure 10: Sample example of World Model Generation