Meta-Reasoning: Semantics-Symbol Deconstruction for Large Language Models

Yiming Wang[1] **Zhuosheng Zhang**[1] **Pei Zhang**[2,*] **Baosong Yang**[2] **Rui Wang**[1,*]
[1]Shanghai Jiao Tong University [2]Alibaba Group Inc.
{yiming.wang,zhangzs,wangrui12}@sjtu.edu.cn
psyangqi@gmail.com, yangbaosong.ybs@alibaba-inc.com

Abstract

Neural-symbolic methods have demonstrated efficiency in enhancing the reasoning abilities of large language models (LLMs). However, existing methods mainly rely on syntactically mapping natural languages to complete formal languages like Python and SQL. Those methods require that reasoning tasks be convertible into programs, which cater to the computer execution mindset and deviate from human reasoning habits. To broaden symbolic methods' applicability and adaptability in the real world, we propose the **Meta-Reasoning** from a linguistic perspective. This method empowers LLMs to deconstruct reasoning-independent semantic information into generic symbolic representations, thereby efficiently capturing more generalized reasoning knowledge. We conduct extensive experiments on more than ten datasets encompassing conventional reasoning tasks like arithmetic, symbolic, and logical reasoning, and the more complex interactive reasoning tasks like theory-of-mind reasoning. Experimental results demonstrate that Meta-Reasoning significantly enhances in-context reasoning accuracy, learning efficiency, out-of-domain generalization, and output stability compared to the Chain-of-Thought technique. Code and data are publicly available at https://github.com/Alsace08/Meta-Reasoning.

1 Introduction

Symbols serve as the primitive carrier through which humans can comprehend, articulate, and conceptualize the intricacies of both nature and society (Peirce and Buchler, 1902). From a cross-linguistic perspective, ideographic symbolic languages like Arabic numerals, mathematical symbols, and emojis can transcend barriers to natural semantic understanding. They serve as a universal representation across ethnically diverse human languages (Chen et al., 2022; Cheng et al., 2022; Wei et al., 2023; Liu et al., 2023; Das et al., 2023), facilitating communication and comprehension on a global scale. In a specific mono-linguistic communication scenario, symbols inherently possess multiple referential meanings shaped by social and cultural properties (Blumer, 1986). Consequently, a single symbol can encapsulate diverse semantic representations. Conversely, various semantic representations can converge onto the same symbol, forming a many-to-one relationship when abstracting referential meanings. This transformation opens avenues for transforming natural language reasoning into more generalized patterns, enabling efficient solutions.

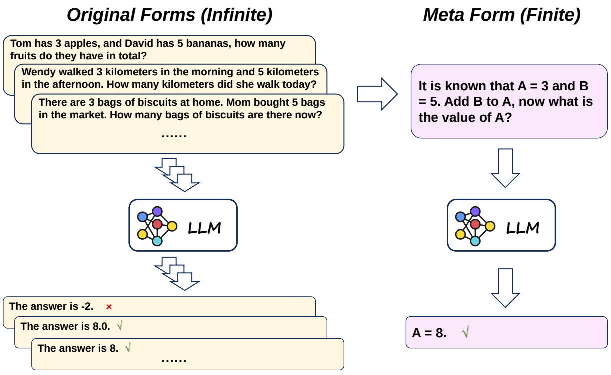


Figure 1: Numerous language reasoning tasks exhibit meta-forms, wherein identifying general patterns can alleviate the reasoning burden on LLMs and facilitate learning through analogy.

Current reasoning paradigms of large language models (LLMs), such as Chain-of-Thought (CoT) prompting (Wei et al., 2022; Kojima et al., 2022a; Zhang et al., 2023b), rely on multiple in-context learning demonstrations to perform well. However, the number of demonstrations is limited by LLMs' input capacity and inference cost, rendering it impractical to cover the distribution of specific task features exhaustively. Therefore, we advocate a paradigm shift from infinite semantics systems to finite symbolic systems so that LLMs can

[*] Rui Wang and Pei Zhang are Co-corresponding Authors.

Findings of the Association for Computational Linguistics: ACL 2024, pages 622–643
August 11-16, 2024

acquire more generic knowledge with enhanced data learning efficiency, as shown in Figure 1.

Motivated by the insight above, we introduce **Meta-Reasoning**, a novel reasoning paradigm aimed at deconstructing the semantics of entities and operations in questions into generic symbolic representations. Meta-Reasoning enables LLMs to learn generalized reasoning patterns across various semantics-wrapped scenarios, enhancing learning efficiency and reasoning accuracy. We apply Meta-Reasoning to in-context learning by designing demonstrations integrating semantic resolution with the CoT technique. This empowers LLMs to deconstruct questions and effectively capture more generalized knowledge autonomously.

To assess the efficacy of our method, we conduct experiments on over ten datasets, spanning both conventional reasoning scenarios, which involve arithmetic, symbolic, and logical reasoning tasks, and interactive reasoning scenarios, which involve theory-of-mind reasoning. We mainly compare our method with the CoT method upon GPT-3 and ChatGPT. Experimental results show that Meta-Reasoning consistently outperforms the Few-Shot-CoT method across all tasks, demonstrating significant performance improvements. In the conventional reasoning scenarios, Meta-Reasoning achieves an average performance gain of +20% across all datasets with fewer demonstrations. In more complex interactive reasoning scenarios, Meta-Reasoning surpasses CoT across all levels of theory-of-mind reasoning with just a single demonstration. Moreover, Meta-Reasoning demonstrates remarkable out-of-domain generalization and output stability, indicating its scalability and user-friendly nature as a reasoning paradigm.

To our knowledge, we are the first to establish an equivalence mapping from semantics to symbols within natural language. This innovation facilitates in-context learning for LLMs, significantly enhancing their capacity for generalized reasoning. We expect to extend the reasoning ability boundary of LLMs based on this research.

2 Preliminary: Why Meta-Reasoning?

Meta-Reasoning is an idealized reduction-based reasoning paradigm defined in this work, whose goal is to reduce the infinite semantic concepts in the world's languages to a finite symbolic system, thus allowing machines to generalize to many semantically wrapped problems through the acquisition of universal laws. This paradigm is best suited for such a reasoning scenario: the final reasoning results are independent of the particular semantic representations and are only related to the underlying reasoning skeletons.

The core of Meta-Reasoning lies in **Semantic-Symbolic Deconstruction**, which we simplify as **Semantic Resolution**. This process conveys the semantics of the original problem via symbols with generalized meanings, without affecting the final results. However, *why deploying Semantic Resolution in LLMs* is a key issue, we must consider the advantages it brings to the reasoning process.

We explore this issue from two perspectives: (i) the human reasoning speed when responding to different questions, and (ii) the machine reasoning accuracy when responding to different questions. We select MultiArith (Roy and Roth, 2015) and GSM8K (Cobbe et al., 2021), two arithmetic datasets, and rephrase 100 questions in each dataset according to the semantic resolution rules that will be introduced in Section 3.1, thereby creating meta-questions. Subsequently, we distribute the original and meta-questions to both human volunteers and LLMs to obtain corresponding results of metrics.

2.1 Response Speed Test For Human

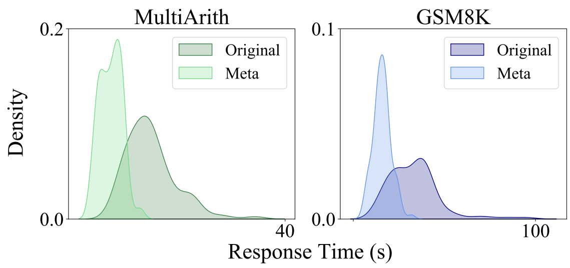


Figure 2: Human response time comparisons when solving original and meta-questions.

We assess the response speed of three human volunteers by measuring the total time taken from receiving the question to providing the answer.[1] As shown in Figure 2, human response speed significantly improves when solving meta-questions, particularly evident in GSM8K. This acceleration is attributed to the removal of unimportant semantic information in **meta-questions, which enables quicker recognition of the reasoning skeleton by humans**. Moreover, the more concentrated distribution of human reaction times suggests

[1]Samples with incorrect answers are excluded from the analysis due to their negligible impact, given the low difficulty level of the math problems for adults.

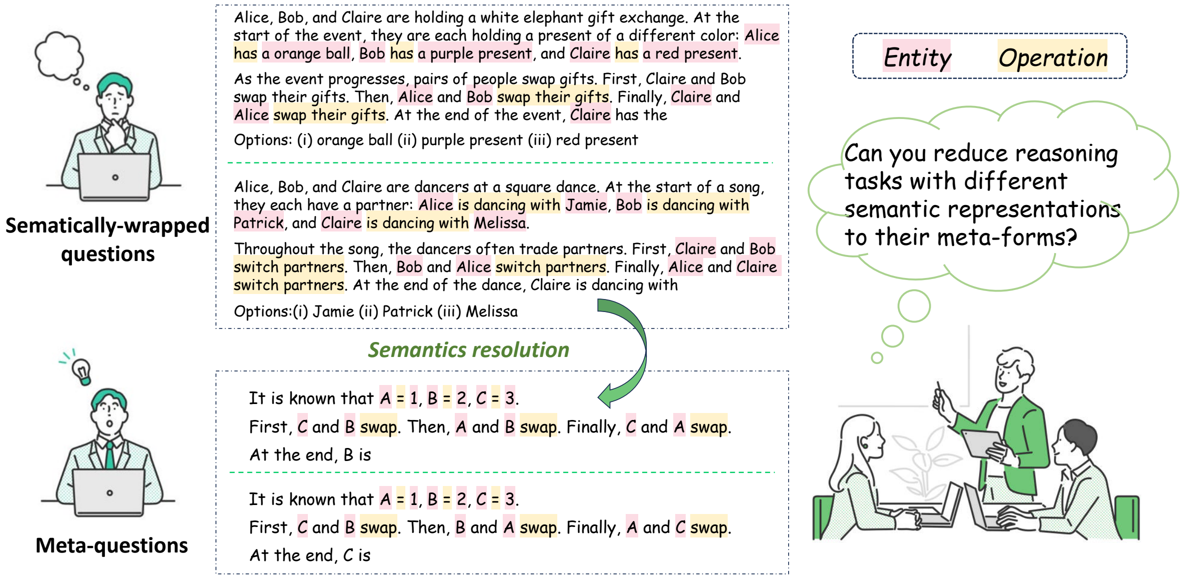


Figure 3: **Semantic Resolution of Meta-Reasoning**. We set resolution rules for Entity and Operation.

a similarity in reasoning frameworks for such problems, indicating that **semantic resolution fosters consistency in reasoning patterns**.

2.2 Accuracy Test For Machine

MultiArith (original → meta)	
Zero-Shot	28% → 31% (+)
Zero-Shot-CoT	70% → 100% (+)
GSM8K (original → meta)	
Zero-Shot	22% → 13% (-)
Zero-Shot-CoT	41% → 97% (+)

Table 1: LLMs Performance comparisons when solving original and meta-questions.

We assess the reasoning accuracy of GPT-3 using two prompting paradigms: standard Zero-Shot[2] and Zero-Shot-CoT[3] (Kojima et al., 2022a). As shown in Table 1, The Standard Zero-Shot method performs similarly on both types of questions, with notably poor performance on the GSM8K dataset. However, Zero-Shot-CoT yields markedly different outcomes. Specifically, when applied to the meta-questions, Zero-Shot-CoT demonstrates a significant performance improvement, particularly evident in the GSM8K dataset. This observation suggests that **CoT reasoning for LLMs becomes notably smoother when tackling meta-problems**.

[2]The prompt is "*A:*".

[3]The prompt is "*A: Let's think step by step.*".

3 Meta-Reasoning Paradigm

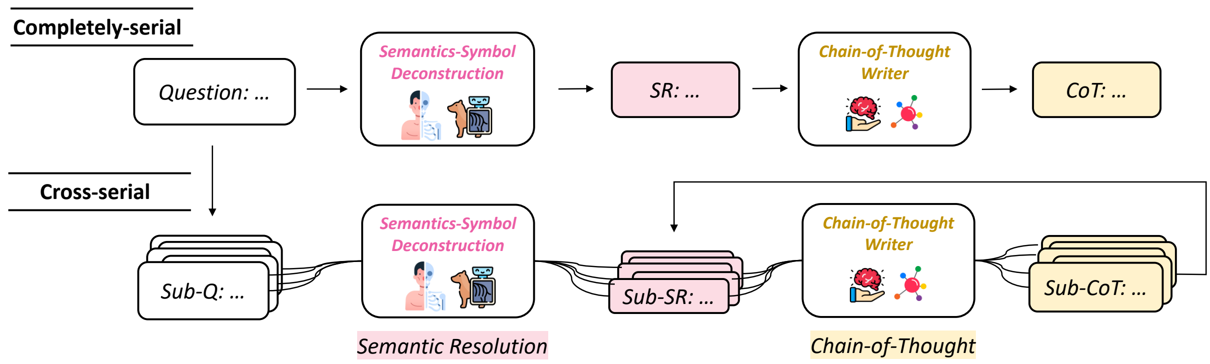
We have observed notable performance improvements in LLMs when tackling questions after semantic resolution in the last section. In this section, we formally introduce the Meta-Reasoning paradigm employed in LLMs. Section 3.1 defines the specific rules for semantic resolution. Then, we put this process through in-context learning for LLMs to imitate, and Section 3.2 formalizes the demonstration design form of in-context learning.

3.1 Definition: Semantic Resolution Rules

Semantic resolution corresponds to the many-to-one mapping from various semantic representations to the most intrinsic symbolic representation. We focus on two types of elements within text sequences that structure the entire reasoning skeleton but whose semantics do not change the reasoning path: (i) **Entity**, it represents the subjects on which the reasoning task acts, but it is not critical what or who exactly it is; (ii) **Operation**, it establishes connections and changes between subjects, but the exact form of that is not important. For example, "he ate 3 apples" and "he threw 3 apples" are both essentially forms of subtraction. Examples are shown in Figure 3.

Entity. Intuitively, entity representations with natural semantics can be treated as the expansion products of an exhaustive set of non-empty symbols. Given a native symbol set[4](alphabet),

[4]Take examples in the English language system, regardless of lowercase or uppercase.



Q: Alice, Bob, and Claire are dancers at a square dance. At the start of a song, they each have a partner: Alice is dancing with Helga, Bob is dancing with Ophelia, and Claire is dancing with Sam. Throughout the song, the dancers often trade partners. First, Bob and Alice switch partners. Then, Claire and Bob switch partners. Finally, Claire and Alice switch partners. At the end of the dance, Bob is dancing with
Options: (A) Helga (B) Ophelia (C) Sam

Completely-serial

A: The question can be simplified to: It is known that A = 1, B = 2, C = 3. First, B and A swap. Then, C and B swap. Finally, C and A swap. At the end, B is:
First, (B = 2, A = 1 -> B = 1, A = 2) -> A = 2, B = 1, C = 3.
Then, (C = 3, B = 1 -> C = 1, B = 3) -> A = 2, B = 3, C = 1.
Finally, (C = 1, A = 2 -> C = 2, A = 1) -> A = 1, B = 3, C = 2.
At the end, B = 3.
1 -> (A), 2 -> (B), 3 -> (C), 4 -> (D), 5 -> (E). For B = 3, 3 -> (C), so the answer is (C).

Cross-serial

A: The question can be simplified to: It is known that A = 1, B = 2, C = 3.
First, Bob and Alice switch partners: B and A swap -> (B = 2, A = 1 -> B = 1, A = 2) -> A = 2, B = 1, C = 3.
Then, Claire and Bob switch partners: C and B swap -> (C = 3, B = 1 -> C = 1, B = 3) -> A = 2, B = 3, C = 1.
Finally, Claire and Alice switch partners: C and A swap -> (C = 1, A = 2 -> C = 2, A = 1) -> A = 1, B = 3, C = 2.
At the end of the dance, Bob is dancing with: B is -> B = 3.
1 -> (A), 2 -> (B), 3 -> (C), 4 -> (D), 5 -> (E). For B = 3, 3 -> (C), so the answer is (C).

Figure 4: **In-context Learning Pipeline (Upper) and Example (Lower) of Meta-Reasoning**. The examples are taken from the Tracking Shuffled Objects task. For drafted demonstrations, we propose completely-serial and cross-serial fusion modes of semantic resolution and chain-of-thought, allowing LLMs to perform single-step reasoning more data-efficiently.

$\Sigma^1 = \{A, B, ..., Z\}$, the positive closure $\Sigma^+ = \bigcup_{i=1}^{\infty} \Sigma^i$ of Σ^1 contains the set Q of all symbolic representations with natural semantics in the English language system, i.e., $Q \subset \Sigma^+$, where $\Sigma^i(i > 1) = \Sigma^j \times \Sigma^{i-j}(1 \leq j \leq i)$ and $\times$ denotes the Cartesian product operation. We consider the opposite form of the symbol-semantics expansion, i.e., semantics-symbol resolution, and construct the mapping $f_e : Q \rightarrow \Sigma^1$ to transform these complex semantic representations to their primitive symbolic form in the alphabet. Since the symbols in the alphabet are meaningless, the mapping results are not required to be specified — we default to mapping them one by one in alphabetical order without duplication.[5]

Back to reasoning scenarios, given a sequence of original question $S = [s_{1:n}]$, we first manually locate all the entity spans $[s_{i:j}] \subset S$ (e.g. apple, mom), and later apply the mapping f_e to them to obtain the single characters $\sigma_{ij} = f_e([s_{i:j}])$, respectively, which will be embedded back into the original position of the sequence S so that it will be modified into $S = [s_{1:i-1} \circ \sigma_{ij} \circ s_{j:n}]$.

Operation. Entities constitute the set of subjects on which the reasoning task acts, while the definition and change of entity states determine the reasoning path: (i) *definitions* of entity states can usually be reduced to assignment and logical association operations, *i.e.*, $O_1 = \{=, \rightarrow\}$, and O_1 is a finite set; (ii) *changes* in entity states can be reduced to arithmetic operations, *i.e.*, $O_2 = \{+, -, \times, \div\}$, and O_2 is a finite set.[6] Conveniently, these arithmetic symbols can correspond to natural semantics, e.g., "+" corresponds to "add", which allows symbols to be more closely integrated with natural language. Similar to the resolution of entities, we construct the mapping $f_o : Q \rightarrow (O_1 \cup O_2)$, and transform all manually-located operation representation $[s_{i:j}]$ (e.g. eat, have) into single symbols $\rho_{ij} = f_o([s_{i:j}])$, which will be embedded into the original position of the sequence S so that it will be modified into $S = [s_{1:i-1} \circ \rho_{ij} \circ s_{j:n}]$.

Appendix A provides some mapping examples. After semantic resolution, the original questions maximally remove semantically irrelevant terms and simplify the need for semantic reasoning.

[5]For example, there are three semantic representations x_1, x_2, x_3 that need to be mapped, and the mapping can be done by default as $x_1 \rightarrow A, x_2 \rightarrow B, x_3 \rightarrow C$.

[6]There may be some extraordinary operations, but generally finite. We leave this for future work.

3.2 Deployment: Synthetic Demonstration Design for In-context Learning

However, manual annotation of entities and operations one by one is time-consuming. We expect LLMs to autonomously learn generic reasoning patterns for certain reasoning tasks by automatically simplifying complex questions into equivalent and simpler forms. This can drive data-efficient learning. Therefore, we consider the in-context learning. Furthermore, inspired by the demonstrated significance of the CoT technique in enhancing reasoning capabilities in prior works (Wei et al., 2023; Kojima et al., 2022a), we are dedicated to devising a fusion strategy of semantic resolution and CoT, which aims to maximize the performance potential of LLMs in reasoning.

We focus on two fusion modes: **Completely-serial** and **Crossly-serial**. The primary distinction between the two modes lies in whether Semantic Resolution (SR) and CoT appear overlappingly. The pipeline and case are illustrated in Figure 4, with further details provided below:

Completely-serial. We first conduct SR to obtain the meta-question form, then draft the CoT for the corresponding meta-question. In this case, the rationale is $[\mathsf{SR} \circ \mathsf{CoT}]$.

Crossly-serial. We first split the original question into n sub-steps, where n may vary depending on the specific context. For each sub-step i, the sub-rationale is represented as $[\mathsf{SR_i} \circ \mathsf{CoT_i}]$. Finally, we concatenate all the sub-rationales. In this case, the rationale is $[[\mathsf{SR_1} \circ \mathsf{CoT_1}] \circ [\mathsf{SR_2} \circ \mathsf{CoT_2}] \circ \cdots \circ [\mathsf{SR_n} \circ \mathsf{CoT_n}]]$, where $[\mathsf{SR_1} \circ \mathsf{SR_2} \circ \cdots \circ \mathsf{SR_n}] = \mathsf{SR}$ and $[\mathsf{CoT_1} \circ \mathsf{CoT_2} \circ \cdots \circ \mathsf{CoT_n}] = \mathsf{CoT}$.

4 Experiments

4.1 Setup

Tasks and Datasets. We conduct experiments on two categories: (i) conventional reasoning, involving basic reasoning scenarios like arithmetic, symbolic, and logical reasoning. This includes the following datasets: MultiArith (Roy and Roth, 2015), AddSub (Hosseini et al., 2014), Last Letter Concatenation (Letter) (Wei et al., 2022), Coin Flip (Coin) (Wei et al., 2022), Web of Lies (Lies) (Srivastava et al., 2022), Tracking Shuffled Objects[7] (Track) (Srivastava et al., 2022), and (ii) interactive reasoning, which involves reasoning scenarios of multi-agent mental gaming, including Hi-ToM[8] (He et al., 2023a). Refer to Appendix B for detailed information on datasets.

Language Models. We utilize publicly available 175B GPT-3 models (`text-davinci-002` and `text-davinci-003`) (Brown et al., 2020), as well as ChatGPT (`gpt-3.5-turbo`).[9] Additionally, for comparison purposes, we include other robust closed-API LLMs: 175B Codex (`code-davinci-002`) (Chen et al., 2021) and 540B PaLM (Chowdhery et al., 2022).

Implementation and Baselines. In our Meta-Reasoning (MR) paradigm, we use the completely-serial mode for arithmetic tasks and the crossly-serial mode for symbolic and logical tasks. We also compare our method with three other paradigms: (i) Fine-tuning; (ii) Standard prompting, including Zero-Shot and Few-Shot; (iii) Chain-of-Thought (CoT) prompting, including Zero-Shot-CoT (Kojima et al., 2022a) and Few-Shot-CoT (Wei et al., 2022). Refer to Appendix G for demonstrations.

4.2 Main Results I: Conventional Reasoning

Overall Performances. Table 2 presents the results.[10] Our MR consistently outperforms Few-Shot-CoT and notably excels on complex tasks challenging for LLMs. This trend is particularly evident for the relatively capacity-constrained text-davinci-002. Notably, on intricate tasks where pure CoT struggles, our MR effectively alleviates the reasoning bottleneck, resulting in significantly higher accuracy (+27.0% in Letter and +37.7% in Track). This indicates that our MR facilitates LLMs in learning general principles for specific task types, automatically reducing reasoning difficulty across various semantic representations.

Fewer Demonstrations, Better Performances. Figure 5 shows comparisons between the performance of CoT and MR paradigms with varying numbers of demonstrations. MR consistently achieves superior performance across almost all datasets while utilizing fewer demonstrations, particularly evident in symbolic and logical reasoning tasks. For example, in the Letter task,

[7] Divided into 3 subsets based on the number of objects and shuffler operations (3/5/7).

[8] Divided into 5 subsets based on the number of mental gaming orders (1/2/3/4/5).

[9] `https://chat.openai.com/`

[10] Experimental results of GPT-3 were obtained in March 2023 via the OpenAI API interface, while the results of ChatGPT were obtained in November 2023.

Method	Arithmetic		Symbolic		Logical			Avg.
	MultiArith	AddSub	Letter	Coin	Lies	Track(3/5/7)	Track(*Avg.*)	
Previous Fine-tuned SOTA								
Fine-tuned Paradigm								
State-of-the-Art	60.5	84.0	-	-	59.6	-	24.1	-
175B GPT-3 (`text-davinci-002`)								
Standard Prompting Paradigm								
Zero-Shot	22.7	77.0	0.2	53.8	47.2	24.4 / 15.2 / 7.6	15.7	31.0
Few-Shot	33.8	<u>83.3</u>	0.2	57.2	51.6	-	25.1	37.7
Chain-of-Thought Paradigm								
Zero-Shot	78.7	74.7	57.6	91.4	58.4	44.8 / 35.6 / 26.0	35.5	58.4
Few-Shot	<u>91.7</u>	81.3	<u>59.0</u>	<u>97.2</u>	<u>92.0</u>	<u>62.8</u> / <u>60.8</u> / <u>59.6</u>	<u>61.1</u>	<u>75.6</u>
Meta-Reasoning Paradigm (Ours)								
Few-Shot	**94.5**	**86.6**	**86.0**	**100.0**	**99.2**	**97.2 / 100.0 / 99.2**	**98.8**	**95.3**
Δ	+2.8	+3.3	+27.0	+2.8	+7.2	+34.4 / +39.2 / +39.6	+37.7	+19.7
175B GPT-3 (`text-davinci-003`)								
Chain-of-Thought Paradigm								
Zero-Shot	83.8	85.3	64.8	96.8	61.2	37.2 / 36.0 / 30.8	34.7	62.0
Few-Shot	<u>93.6</u>	<u>91.6</u>	<u>70.6</u>	<u>99.6</u>	<u>97.6</u>	<u>68.4</u> / <u>80.8</u> / <u>81.2</u>	<u>76.8</u>	<u>85.4</u>
Meta-Reasoning Paradigm (Ours)								
Few-Shot	**96.7**	**95.4**	**91.6**	**100.0**	**100.0**	**100.0 / 100.0 / 100.0**	**100.0**	**97.9**
Δ	+3.1	+3.8	+21.0	+0.4	+2.4	+31.6 / +19.2 / +18.8	+23.2	+12.5
ChatGPT (`GPT-3.5-Turbo`)								
Chain-of-Thought Paradigm								
Zero-Shot	91.5	85.5	75.6	96.4	68.8	55.6 / 54.0 / 43.2	50.9	71.3
Few-Shot	<u>95.2</u>	<u>93.9</u>	<u>80.2</u>	<u>99.2</u>	<u>96.0</u>	<u>62.8</u> / <u>57.2</u> / <u>54.0</u>	<u>58.0</u>	<u>79.8</u>
Meta-Reasoning Paradigm (Ours)								
Few-Shot	**98.7**	**98.0**	**92.4**	**100.0**	**99.2**	**100.0 / 88.0 / 84.4**	**90.8**	**95.1**
Δ	+3.5	+4.1	+12.2	+0.8	+3.2	+37.2 / +30.8 / +30.4	+32.8	+15.3

Table 2: **Conventional Reasoning Results:** We apply our method on 175B GPT-3 (text-davinci-002 and -003) and ChatGPT, and compare it with three common paradigms: Fine-tuned, Standard Prompting, and Chain-of-Thought Prompting. Our performance gains (Δ) are computed over the previous SOTA (<u>underline</u>). Track(*Avg.*) represents the averaged accuracy of Track(3/5/7), and **Avg.** represents the average accuracy across all datasets.

MR results in a +27.0% improvement for LLMs with 1/2 demonstrations compared to the CoT paradigm. Similarly, in the Track(7) task, using only 1/3 demonstrations (i.e., one demonstration) leads to a remarkable +39.6% boost. This indicates that LLMs can acquire general solutions for specific tasks with minimal demonstrations, facilitating learning through analogy.

4.3 Main Results II: Interactive Reasoning

The real-world reasoning environment is more intricate than these conventional reasoning scenarios. Therefore, we consider more complex interactive scenarios and introduce the Theory-of-Mind (ToM) reasoning. In ToM reasoning, the objects involved in reasoning require subjective observation or cognitive abilities, and their observation and thought directly influence the reasoning outcomes. Therefore, LLMs are susceptible to interference.

The variable parameter "Order" determines ToM's difficulty level, which refers to the layer number of the mental game involved. For example, in 3-order reasoning, the structure might be "*A thinks B thinks C thinks xxx*". Notably, 1-order reasoning does not entail any interaction and is categorized as low-order reasoning. On the other hand, reasoning with an order greater than 1 involves a mental game between multiple observers and is classified as high-order reasoning.

When solving lower-order ToM questions, both 1-shot CoT and MR achieve nearly 100% accuracy, indicating that LLMs can accurately comprehend the reasoning text itself. But when solving high-order ToM, CoT exhibits a notable performance decline, with an about 40% decrease in joint accuracy when transitioning from 1 to 2-order, and with a nearly 0% accuracy remaining at 5-order. In contrast, MR maintains stable performance as the

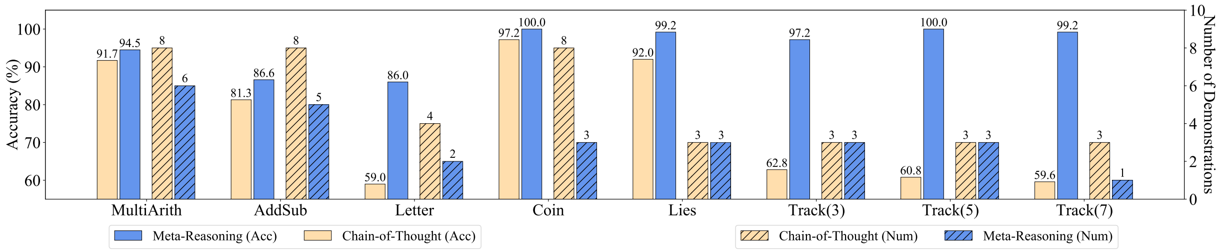


Figure 5: The number of demonstrations used in the CoT and MR paradigms and the corresponding performances.

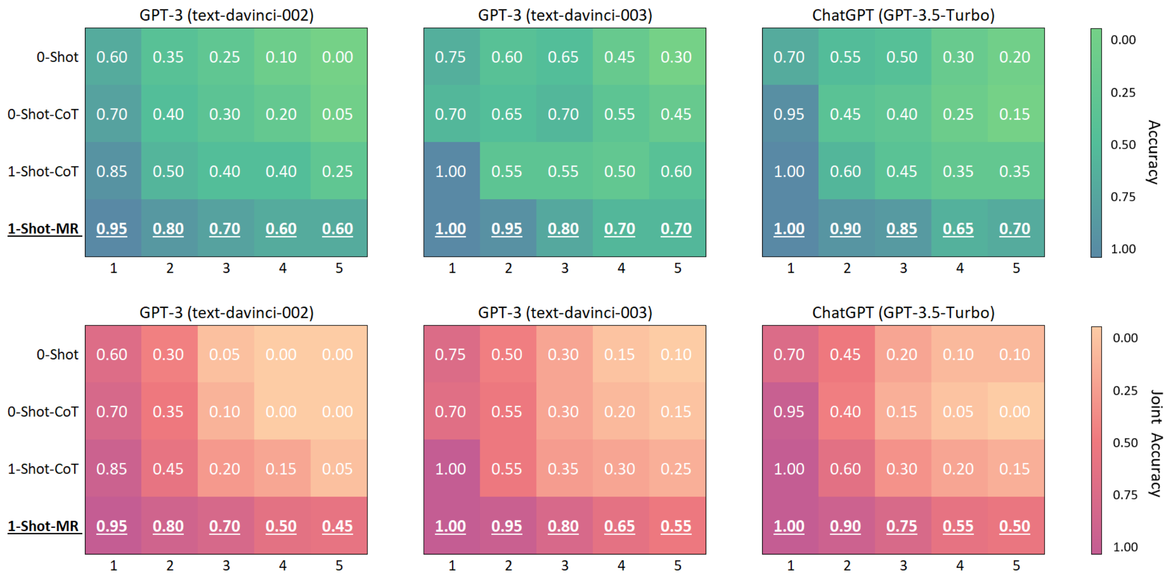


Figure 6: **Interactive Reasoning Results:** The accuracy (upper part) and joint accuracy (lower part) of GPT-3 (text-davinci-002 and -003) and ChatGPT on the Hi-ToM dataset. The x-axis of each heatmap represents ToM orders. [Metric Explanation: (i) *Accuracy* refers to the correctness of each order independently. (ii) *Joint Accuracy* reflects the cumulative correctness, wherein the k-order reasoning is deemed correct only if all reasoning orders less than k are also correct. This metric is instrumental in mitigating randomness error.]

order increases. At 5-order, its performances equal 2-order performances of CoT, indicating its strong ability to handle complex reasoning.

5 Advantage Analysis

5.1 Boundary Test: OOD Generalization

Out-of-domain (OOD) generalization highlights LLMs' ability to address novel tasks by synthesizing limited in-domain knowledge (Wang et al., 2024). We set a challenging boundary test involving Lies, Track, and ToM tasks, to compare the OOD boundary of MR and CoT methods.

For each task, we first manually dissect the smallest unit of reasoning (Details are shown in Appendix D.1). Within each demonstration, we limit the reasoning units to three; thus, any new question exceeding this threshold is considered OOD. We generate 50 samples per task without any reasoning units, then progressively incorporate reasoning units adhering to the structure of the respective dataset. When the following situation occurs for the first time: when the sample contains k reasoning units, LLMs answer correctly; when it contains $k + 1$ reasoning units, LLMs answer incorrectly. At this point, the sample stops iterating, and its Boundary Length (BL) is recorded as k. The sample iteration ceases upon encountering the first case where LLMs answer accurately with k reasoning units and inaccurately with $k + 1$ reasoning units. The Boundary Length of this sample[11] is recorded as k. In dataset $\mathcal{D}$, we compute the Boundary Rate (BRate) for each $k \leq k_{\max}$ as the following formulation:

$$\mathrm{BRate}(\mathcal{D}, k) = \frac{\sum_{s \sim \mathcal{D}} \mathbb{I}(\mathrm{BL}(s) \geq k)}{|\mathcal{D}|}, \quad (1)$$

where $\mathbb{I}(\cdot)$ is the indicator function, $k_{\max}$ is the maximum number of reasoning units.

[11] Refer to Appendix D.2 for a detailed algorithm process.

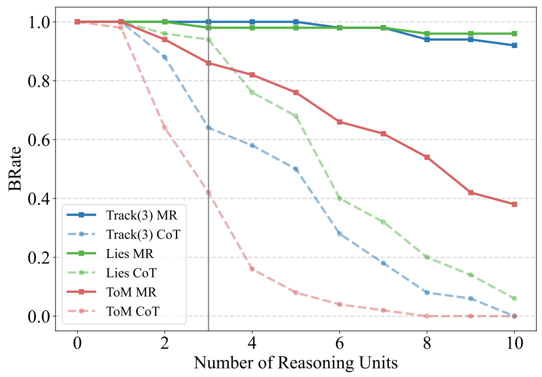


Figure 7: Boundary test of out-of-domain generalization under CoT and MR paradigms, where the number of reasoning units is larger than three (the right area of the vertical gray line in the figure) means out-of-domain.

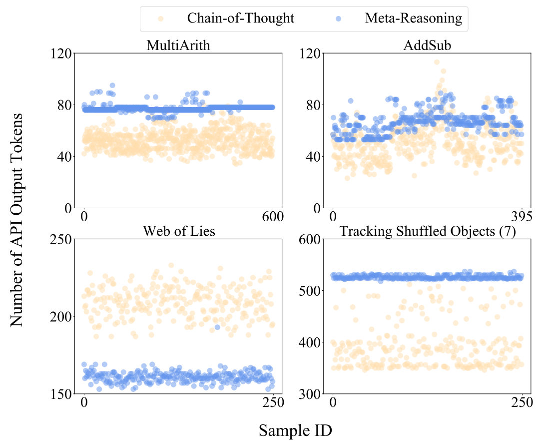


Figure 8: The token number distributions of the text generated by GPT-3 text-davinci-002 when using the Chain-of-Thought and Meta-Reasoning paradigms.

We draw BRate curves of each $\mathcal{D}$. The larger the area enclosed by the curve and the x-axis, the stronger the OOD generalization of the method. Figure 7 shows the BRate curves of each dataset under CoT and MR paradigms, respectively. We note that as the number of reasoning units grows beyond the domain, the CoT curves exhibit a sharp decline, while the MR curves maintain relative smoothness, with Lies and Track tasks achieving nearly 100% BRate. This indicates that our MR facilitates strong OOD generalization for LLMs.

5.2 Output Stability Test

In addition to performance, user experience is another crucial consideration. Currently, access to LLMs like GPT-3 involves paywalls. Unexpected outputs, such as endless looping text or random guessing, can increase user fees, so making a stable output space is essential. We analyze the number of API output tokens generated by MR and CoT paradigms for each sample to evaluate this stability, as illustrated in Figure 8. When employing the MR, the output scales of different samples are much closer. Conversely, under the CoT, outputs scatter widely, increasing the likelihood of encountering unexpected and abnormal situations.

6 Discussion

We conduct ablation studies to examine the role of semantic resolution in the reasoning process. Moreover, we compare our method with existing work in language programming (Chen et al., 2022; Gao et al., 2023b), highlighting Meta-Reasoning's broader applicability across diverse scenarios. These extended analyses are shown in Appendix C.

7 Related Work

Our work is related to the research lines of neural-symbolic methods and chain-of-thought reasoning. Please refer to Appendix F for full details.

Neural-Symbolic Methods in LLMs. Symbolic learning (Chen et al., 2021) significantly improves LLMs' reasoning performance. Prior works focus on converting natural languages into programming languages (Gao et al., 2022; Cheng et al., 2022) and accessing external interpreters for execution (Schick et al., 2023); or using symbolic tasks for post-tuning (Liu et al., 2023; Wei et al., 2023), leading to performance improvements. However, these symbols are well-defined formal languages completely independent of natural languages. Our work jumps out of this framework and further enhances the efficiency of the symbolic methods.

Chain-of-Thought Reasoning. Intriguing chain-of-thought techniques (Wei et al., 2022; Kojima et al., 2022b; Wang et al., 2022b; Zhang et al., 2023b) have effectively leveraged the emergent ability of LLMs to decompose multi-step reasoning. It can improve the performance of general-purpose and even domain-specific reasoning (Zhang et al., 2023c; Wang et al., 2023; He et al., 2023b; Zhang et al., 2023a).

8 Conclusion

We propose Meta-Reasoning, a semantic-symbol deconstruction paradigm for reasoning. Through the semantic resolution of the original questions, we enable LLMs to grasp meta-forms and general solutions for specific types of reasoning tasks. This approach requires fewer demonstrations to expand the upper limit of their reasoning accuracy, out-of-domain generalization, and output stability.

Limitations

Semantic resolution dictates that Meta-Reasoning tasks must disregard the intrinsic properties of entities. Consequently, Meta-Reasoning may not be well-suited for reasoning tasks reliant on world knowledge in semantics, such as commonsense reasoning. However, Meta-Reasoning shows potential in real-world agent reasoning scenarios (Gao et al., 2023a; Tang et al., 2023). When agents are impeded by irrelevant properties, Meta-Reasoning can effectively circumvent such obstacles. We aim to explore more comprehensive reasoning scenarios to further justify its applicability in future work.

Ethics Statement

We use publicly available datasets for experiments, so the ethics issues of the source texts are non-existent. For the generated contents with LLMs, prior work (Brown et al., 2020; Chan, 2023) has elaborated on their inevitable potential toxicity, such as issues of bias and fairness. We completely keep the prompts neutral and task-specific to avoid toxic language generation, and there were no toxic texts that appeared in our experiments.

Acknowledgements

Yiming and Rui are with MT-Lab, Department of Computer Science and Engineering, School of Electronic Information and Electrical Engineering, and also with the MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University, Shanghai 200204, China. This paper is mainly supported by the Alibaba-AIR Program (22088682). Yiming and Rui are also supported by the National Natural Science Foundation of China (62176153), and the Shanghai Municipal Science and Technology Major Project (2021SHZDZX0102). This work is also partially supported by the Joint Funds of the National Natural Science Foundation of China (U21B2020).

References

Herbert Blumer. 1986. *Symbolic interactionism: Perspective and method*. Univ of California Press.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*.

Anastasia Chan. 2023. Gpt-3 and instructgpt: Technological dystopianism, utopianism, and "contextual" perspectives in ai ethics and industry. *AI and Ethics*, 3(1):53–64.

Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde de Oliveira Pinto, Jared Kaplan, Harri Edwards, Yuri Burda, Nicholas Joseph, Greg Brockman, et al. 2021. Evaluating large language models trained on code. *ArXiv preprint*, abs/2107.03374.

Wenhu Chen, Xueguang Ma, Xinyi Wang, and William W Cohen. 2022. Program of thoughts prompting: Disentangling computation from reasoning for numerical reasoning tasks. *ArXiv preprint*, abs/2211.12588.

Zhoujun Cheng, Tianbao Xie, Peng Shi, Chengzu Li, Rahul Nadkarni, Yushi Hu, Caiming Xiong, Dragomir Radev, Mari Ostendorf, Luke Zettlemoyer, et al. 2022. Binding language models in symbolic languages. *ArXiv preprint*, abs/2210.02875.

Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, et al. 2022. Palm: Scaling language modeling with pathways. *ArXiv preprint*, abs/2204.02311.

Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, et al. 2021. Training verifiers to solve math word problems. *ArXiv preprint*, abs/2110.14168.

Mithun Das, Saurabh Kumar Pandey, and Animesh Mukherjee. 2023. Evaluating chatgpt's performance for multilingual and emoji-based hate speech detection. *ArXiv preprint*, abs/2305.13276.

Hao Fei, Bobo Li, Qian Liu, Lidong Bing, Fei Li, and Tat-Seng Chua. 2023. Reasoning implicit sentiment with chain-of-thought prompting. *ArXiv preprint*, abs/2305.11255.

Yao Fu, Hao Peng, Ashish Sabharwal, Peter Clark, and Tushar Khot. 2022. Complexity-based prompting for multi-step reasoning. *ArXiv preprint*, abs/2210.00720.

Chang Gao, Wenxuan Zhang, Guizhen Chen, and Wai Lam. 2023a. Jsontuning: Towards generalizable, robust, and controllable instruction tuning. *ArXiv preprint*, abs/2310.02953.

Luyu Gao, Aman Madaan, Shuyan Zhou, Uri Alon, Pengfei Liu, Yiming Yang, Jamie Callan, and Graham Neubig. 2022. Pal: Program-aided language models. *ArXiv preprint*, abs/2211.10435.

Luyu Gao, Aman Madaan, Shuyan Zhou, Uri Alon, Pengfei Liu, Yiming Yang, Jamie Callan, and Graham Neubig. 2023b. Pal: Program-aided language models. In *International Conference on Machine Learning*, pages 10764–10799. PMLR.

Yinghui He, Yufan Wu, Yilin Jia, Rada Mihalcea, Yulong Chen, and Naihao Deng. 2023a. Hi-tom: A benchmark for evaluating higher-order theory of mind reasoning in large language models. *ArXiv preprint*, abs/2310.16755.

Zhiwei He, Tian Liang, Wenxiang Jiao, Zhuosheng Zhang, Yujiu Yang, Rui Wang, Zhaopeng Tu, Shuming Shi, and Xing Wang. 2023b. Exploring human-like translation strategy with large language models. *ArXiv preprint*, abs/2305.04118.

Namgyu Ho, Laura Schmid, and Se-Young Yun. 2022. Large language models are reasoning teachers. *ArXiv preprint*, abs/2212.10071.

Mohammad Javad Hosseini, Hannaneh Hajishirzi, Oren Etzioni, and Nate Kushman. 2014. Learning to solve arithmetic word problems with verb categorization. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 523–533, Doha, Qatar. Association for Computational Linguistics.

Seungone Kim, Se June Joo, Doyoung Kim, Joel Jang, Seonghyeon Ye, Jamin Shin, and Minjoon Seo. 2023. The cot collection: Improving zero-shot and few-shot learning of language models via chain-of-thought fine-tuning. *ArXiv preprint*, abs/2305.14045.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2022a. Large language models are zero-shot reasoners. In *ICML 2022 Workshop on Knowledge Retrieval and Language Models*.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2022b. Large language models are zero-shot reasoners. In *Advances in Neural Information Processing Systems*.

Junlong Li, Zhuosheng Zhang, and Hai Zhao. 2022. Self-prompting large language models for open-domain qa. *ArXiv preprint*, abs/2212.08635.

Qian Liu, Fan Zhou, Zhengbao Jiang, Longxu Dou, and Min Lin. 2023. From zero to hero: Examining the power of symbolic tasks in instruction tuning. *ArXiv preprint*, abs/2304.07995.

Qing Lyu, Shreya Havaldar, Adam Stein, Li Zhang, Delip Rao, Eric Wong, Marianna Apidianaki, and Chris Callison-Burch. 2023. Faithful chain-of-thought reasoning. *ArXiv preprint*, abs/2301.13379.

Charles Sanders Peirce and Justus Buchler. 1902. Logic as semiotic: The theory of signs. *Philosophical Writings of Peirce (New York: Dover Publications, 1955)*, page 100.

Subhro Roy and Dan Roth. 2015. Solving general arithmetic word problems. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1743–1752, Lisbon, Portugal. Association for Computational Linguistics.

Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, Roberta Raileanu, Maria Lomeli, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. 2023. Toolformer: Language models can teach themselves to use tools. *ArXiv preprint*, abs/2302.04761.

Freda Shi, Xinyun Chen, Kanishka Misra, Nathan Scales, David Dohan, Ed Chi, Nathanael Schärli, and Denny Zhou. 2023. Large language models can be easily distracted by irrelevant context. *ArXiv preprint*, abs/2302.00093.

Aarohi Srivastava, Abhinav Rastogi, Abhishek Rao, Abu Awal Md Shoeb, Abubakar Abid, Adam Fisch, Adam R Brown, Adam Santoro, Aditya Gupta, Adrià Garriga-Alonso, et al. 2022. Beyond the imitation game: Quantifying and extrapolating the capabilities of language models. *ArXiv preprint*, abs/2206.04615.

Xiangru Tang, Yiming Zong, Yilun Zhao, Arman Cohan, and Mark Gerstein. 2023. Struc-bench: Are large language models really good at generating complex structured data? *ArXiv preprint*, abs/2309.08963.

Boshi Wang, Sewon Min, Xiang Deng, Jiaming Shen, You Wu, Luke Zettlemoyer, and Huan Sun. 2022a. Towards understanding chain-of-thought prompting: An empirical study of what matters. *ArXiv preprint*, abs/2212.10001.

Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. 2022b. Self-consistency improves chain of thought reasoning in language models. *ArXiv preprint*, abs/2203.11171.

Yiming Wang, Pei Zhang, Baosong Yang, Derek F Wong, Zhuosheng Zhang, and Rui Wang. 2024. Trajectory volatility for out-of-distribution detection in mathematical reasoning. *arXiv preprint arXiv:2405.14039*.

Yiming Wang, Zhuosheng Zhang, and Rui Wang. 2023. Element-aware summarization with large language models: Expert-aligned evaluation and chain-of-thought method. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed H Chi, Quoc V Le, Denny Zhou, et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. In *Advances in Neural Information Processing Systems*.

Jerry Wei, Le Hou, Andrew Lampinen, Xiangning Chen, Da Huang, Yi Tay, Xinyun Chen, Yifeng Lu, Denny Zhou, Tengyu Ma, et al. 2023. Symbol tuning improves in-context learning in language models. *ArXiv preprint*, abs/2305.08298.

Zhuosheng Zhang, Yao Yao, Aston Zhang, Xiangru Tang, Xinbei Ma, Zhiwei He, Yiming Wang, Mark Gerstein, Rui Wang, Gongshen Liu, et al. 2023a. Igniting language intelligence: The hitchhiker's guide from chain-of-thought reasoning to language agents. *arXiv preprint arXiv:2311.11797*.

Zhuosheng Zhang, Aston Zhang, Mu Li, and Alex Smola. 2023b. Automatic chain of thought prompting in large language models. In *The Eleventh International Conference on Learning Representations*.

Zhuosheng Zhang, Aston Zhang, Mu Li, Hai Zhao, George Karypis, and Alex Smola. 2023c. Multimodal chain-of-thought reasoning in language models. *ArXiv preprint*, abs/2302.00923.

Denny Zhou, Nathanael Schärli, Le Hou, Jason Wei, Nathan Scales, Xuezhi Wang, Dale Schuurmans, Olivier Bousquet, Quoc Le, and Ed Chi. 2022. Least-to-most prompting enables complex reasoning in large language models. *ArXiv preprint*, abs/2205.10625.

A Semantic-Symbol Operation Rulebase

Table 4 shows operation mapping examples. Due to the lack of automatic methods, the rule base is continuously revised and improved with the annotation process.

Symbol	Semantics
=	is, are, have, ...
$\rightarrow$	mean, represent, infer, ...
+	buy, get, pick, ...
-	sell, throw, lose, ...
$\times$	each, per, both, ...
$\div$	split, divide, group, ...

Table 3: Examples of operations with infinite natural semantics mapped to finite symbols.

B Dataset Details

To measure the generalizability of our approach, we consider conventional and interactive reasoning:

Conventional Reasoning. In this scenario, reasoning information is globally accessible to all observers. We adopt three categories of reasoning as our testbed: (i) Arithmetic reasoning, we choose MultiArith (Roy and Roth, 2015) and AddSub (Hosseini et al., 2014) tasks, with 600 and 395 test instances separately; (ii) Symbolic reasoning, we follow Wei et al. (2022) to use Last Letter Concatenation and Coin Flip tasks, they both include 500 test instances; (iii) Logical reasoning, We choose Web of Lies and Tracking Shuffled Objects tasks from BIG-bench (Srivastava et al., 2022) — a more challenging reasoning task collection. In particular, the Tracking Shuffled Objects task is divided into three datasets according to the number of objects and shuffler operations (3/5/7). each dataset includes 250 test instances.

Interactive Reasoning. In this scenario, individual observers are limited to observing distinct local reasoning information, necessitating reliance on interaction and mental gaming for their reasoning processes. We select the Theory-of-Mind (ToM) task as our testbed and choose Hi-ToM (He et al., 2023a) as a benchmark for it involves the complex higher-order mind. This dataset contains a collection of multiple subsets ranging from 1 to 5 orders, each subset has 20 test instances.

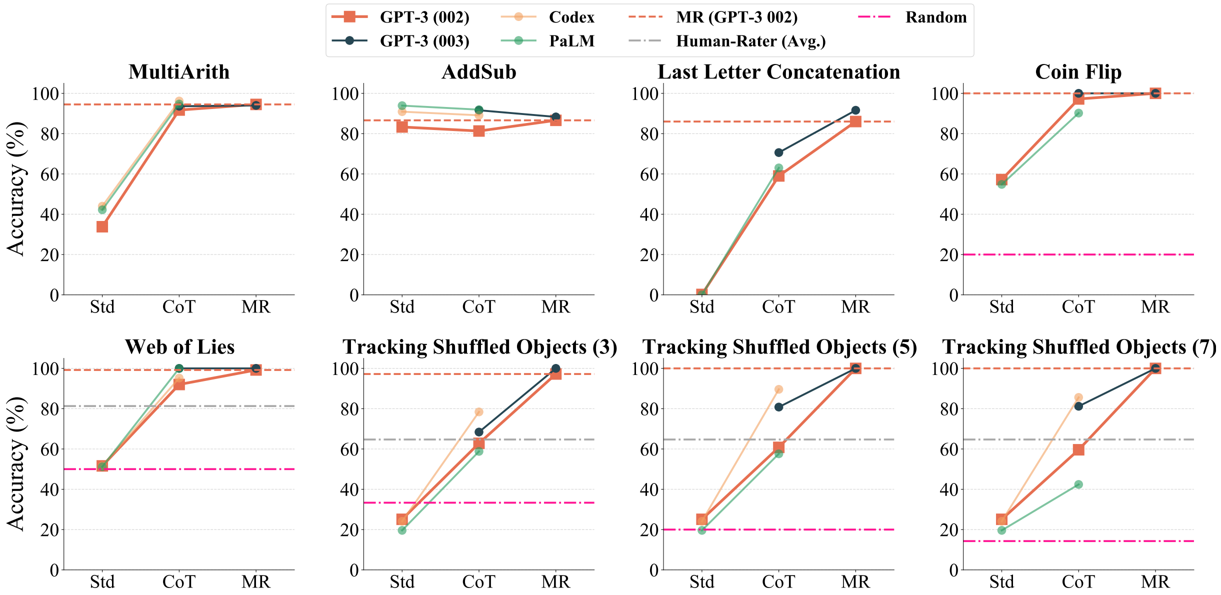


Figure 9: The performance gaps between four LLMs under different paradigms (Std → Standard prompting, CoT → Chain-of-Thought, MR → Meta-Reasoning) in all datasets.

C Extended Analysis

C.1 Bridge the Gap between LLMs' Capabilities.

We conduct longitudinal analyses of performance gaps between four LLMs. Figure 9 visualizes the performance gaps between LLMs for the same dataset and paradigm. Observations are as below:

- GPT-3 text-davinci-002 (the worst original performance among the four LLMs) greatly outperforms the remaining three LLMs under the CoT paradigm on five datasets after adopting the Meta-Reasoning paradigm.

- Performance gaps between text-davinci-002 and -003 on all datasets are greatly reduced compared to under the CoT paradigm after adopting the Meta-Reasoning paradigm.

These findings indicate that our Meta-Reasoning paradigm further bridges the gap in the LLMs' capability themselves, allowing the weaker LLMs (e.g. text-davinci-002) to approximate the stronger LLMs (e.g. text-davinci-003) in reasoning ability.

C.2 Ablation Study

We perform ablation studies to explore the role of semantic resolution in the whole reasoning process. Table 4 reports the error rates of all datasets under both paradigms and error reason rates (caused by semantic resolution or pure reasoning) in the wrong samples for each dataset.

We note that the causes of errors are inconsistent in different reasoning scenarios. For symbolic and logical reasoning, LLMs hardly produce any semantic resolution errors, only errors in the reasoning process (of course, the error rate of their reasoning itself is extremely low). This shows that semantic reasoning fully plays a positive role in reducing the complexity of reasoning for LLMs. But in arithmetic reasoning, semantic resolution errors often occur, and exceed the errors in the reasoning process itself. This shows that LLMs cannot reduce all types of questions under specific arithmetic datasets well. Intuitively, symbolic and logical reasoning questions are easier to logicalize than arithmetic reasoning questions, and the combination of reasoning units under arithmetic reasoning is more flexible. How to fully push the upper limit of LLM's semantic resolution ability, so as to further improve its reasoning ability, is a promising future work.

C.3 Formal Pattern Flexibility

So far, most symbolic reasoning work focuses on mapping natural semantics to formal languages with complete grammar (such as Python and SQL). However, this grammatical completeness actually limits the form conversion, and it has higher requirements for the abstraction of the original reasoning tasks. To verify the flexibility of our paradigm, we contrast Program-of-Thought (PoT), a Text-to-Python reduction approach for reasoning tasks (Chen et al., 2022; Gao et al., 2022).

		MultiArith	AddSub	Letter	Coin	Lies	Track(*Avg.*)
Error Rate (%)	Chain-of-Thought	8.3	18.7	41.0	2.8	8.0	38.9
	Meta-Reasoning	5.5	13.4	14.0	0.0	0.8	1.2
Error Reason Rate (%) (Meta-Reasoning)	Semantic Resolution	84.8	67.9	0.0	0.0	0.0	0.0
	Pure Reasoning	15.2	32.1	100.0	0.0	100.0	100.0

Table 4: Error rates using the Chain-of-Thought and Meta-Reasoning paradigms for all datasets, and error rates caused by semantic resolution and pure reasoning when using the Meta-Reasoning paradigm. Note that *under each dataset, the error rates of semantic resolution and pure reasoning sum up to a constant 1*. This arises from the fact that when semantic resolution errors occur, we no longer classify pure reasoning as either correct or incorrect. For instance, within the MultiArith dataset, among the 5.5% of error samples, 84.8% were attributed to semantic reasoning inaccuracies, leaving the remaining 15.2% attributed to errors in pure reasoning.

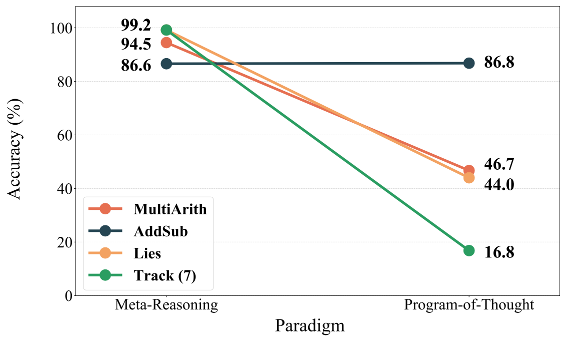


Figure 10: Performance comparisons between our Meta-Reasoning paradigm and the Program-of-Thought paradigm (w/o external Python interpreter).

Meanwhile, to keep the settings consistent, we eliminate the call of an external interpreter in the PoT paradigm but utilize the LLMs themselves to complete the entire reasoning step, and select the same demonstrations for the two paradigms.

Figure 10 shows the performance comparisons of PoT and MR paradigms on four datasets. For the two arithmetic reasoning tasks (MultiArith and AddSub), the performance of PoT fluctuates wildly after removing the external interpreter. For the two symbolic reasoning tasks (Lies and Track), PoT is almost completely ineffective. In contrast, MR has stronger flexibility when encountering reasoning tasks that are not easily programmed.

D Details of Boundary Test

D.1 Reasoning Unit Division

Sample reasoning units for three datasets are as below. The smallest reasoning unit is highlighted in blue.

- **Lies.**
  Andree lies.
  Delfina says Andree lies.
  Jim says Delfina tells the truth.
  Gwenn says Jim lies.
  Delbert says Gwenn lies.
  Does Delbert tell the truth?

- **Track.**
  Alice, Bob, and Claire are dancers at a square dance. At the start of a song, they each have a partner: Alice is dancing with Lola, Bob is dancing with Patrick, and Claire is dancing with Melissa. Throughout the song, the dancers often trade partners.
  First, Alice and Claire switch partners.
  Then, Bob and Claire switch partners.
  Finally, Claire and Alice switch partners.
  At the end of the dance, Bob is dancing with

- **ToM.**
  {Scenario}
  Where does Isabella think Owen thinks Charlotte thinks Aver thinks the lettuce is?

Algorithm 1 Computation of Boundary Length

Input: x: Initialized sample w/o reasoning units.
$\mathcal{D}$: Source dataset of x.
$g(\mathcal{D})$: Reasoning unit generator imitating the style of $\mathcal{D}$.
k_{max}: Maximum number of reasoning units.
p_θ: Language Model.

```
k ← 0
while k < k_max do
    u → g(D), x ← x + u, y ← p_θ(x)
    if y is the correct answer of x then
        k ← k + 1
    else
        break
    end if
end while
return k
```

D.2 Computation of Boundary Length

The algorithm of Boundary Length (BL) Computation is shown in Algorithm 1.

E Case Study

Table 5 compares output examples under Few-Shot-CoT and Meta-Reasoning paradigms. We find that when the original question requires reasoning about a large number of entities and the corresponding relationships, the reasoning process in the Few-Shot-CoT paradigm tends to be more chaotic, leading to errors and potential confusion between entities. In contrast, Meta-Reasoning offers a clearer and more concise reasoning process, reducing interference caused by a large number of natural semantics.

F Additional Related Work

This work is inspired by two directions. First is the neural-symbolic methods, which have shown great promise in improving LLMs' reasoning performance. Current work mainly focuses on converting natural languages into programming languages, however, the symbols that most of these works focus on artificially defined formal languages completely independent of natural languages, which makes it hard to establish the mapping facing complex real-world scenarios. Therefore, our research concentrates on human natural language, delving into semantic resolution at the semiotic level, and pushing the boundaries of LLMs in handling problems within the realm of natural language. Second is the Chain-of-Thought, an important technique for in-context learning reasoning. However, in-context learning with CoT is limited to learning from the reasoning process of the sample itself. Our optimization is high-level, and We hope to promote the efficiency and generality of sample learning by generalizing the features of a single sample to the general features of the entire dataset. Our objective is to enhance the efficiency and generalizability of sample learning upon the CoT framework.

Neural-Symbolic Methods in LLMs. Starting from Codex (Chen et al., 2021), symbolic learning has shown great promise in improving LLMs' reasoning performance. Afterward, a series of works further explored symbolic approaches in LLMs' reasoning, and they can be broadly classified into two categories: (i) converting natural languages into programming languages (Chen et al., 2022; Gao et al., 2022; Cheng et al., 2022), such as Python or SQL, and using the powerful code capabilities of LLMs to parse and even access external interpreters for execution (Schick et al., 2023); (ii) using symbolic tasks for post-tuning of LLMs (Liu et al., 2023), which was found to lead to unexpected improvements in the overall performance of the models. However, the "symbols" that most of these works focus on are artificially defined formal languages completely independent of natural languages. These works establish sample-specific one-to-one mappings between two languages (natural language → formal language). Obviously, formal languages are learned by LLMs with less ambiguity due to their syntactic rigor, but they are divorced from the study of human natural language itself. Recently, Wei et al. (2023) design a novel symbol tuning scheme by replacing natural language labels with semantically-unrelated symbols, but the symbol system they define is not complete. This approach is different from the symbols under formal languages used in previous studies but has not been explored further in depth. Our work closely focuses on human natural language, resolute the semantics at the semiotic level, and explores the upper limit of LLM reasoning in dealing with problems under natural language.

Chain-of-Thought Prompt for Reasoning. Intriguing chain-of-thought (CoT) techniques have effectively leveraged the emergent ability of LLMs to decompose multi-step reasoning. Recent work in this field can be broadly classified into four categories: (i) Improving the performance of general-purpose reasoning tasks (Wei et al., 2022; Kojima et al., 2022b; Wang et al., 2022b; Zhou et al., 2022; Zhang et al., 2023b; Fu et al., 2022), i.e., arithmetic, symbolic, logical, and common-sense reasoning; (ii) Applying to domain-specific reasoning, such as multi-modality (Zhang et al., 2023c), or some purely linguistic tasks, such as translation (He et al., 2023b), summarization (Wang et al., 2023), sentiment analysis (Fei et al., 2023), question-answer (Li et al., 2022), etc; (iii) Analyzing the mechanics and interpretability of CoT (Wang et al., 2022a; Shi et al., 2023; Lyu et al., 2023); (iv) Distilling CoT techniques for smaller models (Ho et al., 2022; Kim et al., 2023).

Question (Tracking Shuffled Objects)

Q: Alice, Bob, and Claire are dancers at a square dance. At the start of a song, they each have a partner: Alice is dancing with Lola, Bob is dancing with Rodrigo, and Claire is dancing with Patrick.
Throughout the song, the dancers often trade partners. First, Alice and Bob switch partners. Then, Claire and Bob switch partners. Finally, Bob and Alice switch partners. At the end of the dance, Alice is dancing with
Options:
(A) Lola
(B) Rodrigo
(C) Pa

Few-Shot-CoT

A: Let's think step by step. (0) At the start: Alice: Lola, Bob: Rodrigo, Claire: Patrick.
(1) Alice and Bob switch partners: Alice: Rodrigo, Bob: Lola, Claire: Patrick.
(2) Claire and Bob switch partners: Alice: Rodrigo, Bob: Patrick, Claire: Lola.
(3) Bob and Alice switch partners: Alice: Lola, Bob: Patrick, Claire: Rodrigo.
At the end of the dance, Alice is dancing with Lola. So the answer is (A). ✗(Entity information interferes with reasoning)

Meta-Reasoning

A: The question can be simplified to: It is known that A = 1, B = 2, C = 3.
First, B and A swap books: B and A → (B = 2, A = 1 → B = 1, A = 2) → A = 2, B = 1, C = 3.
Then, A and C swap books: A and C → (A = 2, C = 3 → A = 3, C = 2) → A = 3, B = 1, C = 2.
Finally, C and B swap books: C and B → (C = 2, B = 1 → C = 1, B = 2) → A = 3, B = 2, C = 1.
At the end of the semester, Alice has: Alice → A = 3, 3 → the 3-rd option → the answer is (C). ✓

Question (Web of Lies)

Q: Sherrie tells the truth. Ryan says Sherrie lies. Bernita says Ryan tells the truth. Tamika says Bernita lies. Jerry says Tamika lies. Does Jerry tell the truth?

Few-Shot-CoT

A: Let's think step by step. (1) Sherrie tells the truth. So, we know that Sherrie tells the truth.
(2) Ryan says Sherrie lies. Since we know from (1) that Sherrie tells the truth, if Ryan says Sherrie lies, then Ryan lies.
(3) Bernita says Ryan tells the truth. Since we know from (2) that Ryan lies, if Bernita says Ryan tells the truth, then Bernita lies.
(4) Tamika says Bernita lies. Since we know from (3) that Bernita lies, if Tamika says Bernita lies, then Tamika tells the truth.
(5) Jerry says Tamika lies. Since we know from (4) that Tamika tells the truth, if Jerry says Tamika lies, then Jerry tells the truth.
Now, the question asks: Does Jerry tell the truth? We know from (5) that Jerry tells the truth. So the answer is Yes. ✗(Complex entity relationships and long reasoning chains interfere with reasoning)

Meta-Reasoning

A: The question can be simplified to: It is known that A = 1.
Ryan says Sherrie lies: lies → A' = 0. Since A = 1, A is not equal to A', so B = 0.
Bernita says Ryan tells the truth: truth → B' = 1. Since B = 0, B is not equal to B', so C = 0.
Tamika says Bernita lies: lies → C' = 0. Since C = 0, C is equal to C', so D = 1.
Jerry says Tamika lies: lies → D' = 0. Since D = 1, D is not equal to D', so E = 0.
Since E = 0, so the answer is: no. ✓

Table 5: Case studies of reasoning chains under Few-Shot-CoT and Meta-Reasoning paradigms.

G Demonstration Design

Figure 11 to 17 show all the demonstrations used in the dataset of this paper.

Meta-Reasoning

Q: For Halloween Debby and her sister combined the candy they received. Debby had 32 pieces of candy while her sister had 42. If they ate 35 pieces the first night, how many pieces do they have left?
A: The question can be simplified to: It is known that A = 32, B = 42, C = 35. Add B to A, and then subtract C from A, now what is the value of B?
A = (A + B) - C = (32 + 42) - 35 = 74 - 35 = 39
So the answer is: 39
Q: A florist had 37 roses. If she sold 16 of them and then later picked 19 more, how many roses would she have?
A: The question can be simplified to: It is known that A = 37, B = 16, C = 19. Subtract B from A, then add C to A, now what is the value of A?
A = (A - B) + C = (37 - 16) + 19 = 21 + 19 = 40
So the answer is: 40
Q: At a company picnic 9 managers and 15 employees decided to start a game of volleyball. If they split into 8 teams how many people would be on each team?
A: The question can be simplified to: It is known that A = 9, B = 15, C = 8. Add B to A, and then divide A by C, now what is the value of A?
A = (A + B) / C = (9 + 15) / 8 = 24 / 8 = 3
So the answer is: 3
Q: A new building needed 14 windows. The builder had already installed 5 of them. If it takes 4 hours to install each window, how long will it take him to install the rest?
A: The question can be simplified to: It is known that A = 14, B = 5, C = 4. Subtract B from A, and then multiply A by C, now what is the value of A?
A = (A - B) * C = (14 - 5) * 4 = 9 * 4 = 36
So the answer is: 36
Q: A magician was selling magic card decks for 9 dollars each. If he started with 12 decks and by the end of the day he had 7 left, how much money did he earn?
A: The question can be simplified to: It is known that A = 9, B = 12, C = 7. Subtract C from B, and then multiply A by B, now what is the value of A?
A = (B - C) * A = (12 - 7) * 9 = 5 * 9 = 45
So the answer is: 45
Q: A pet store has 6 bird cages. If each cage has 2 parrots and 7 parakeets in it, how many birds does the pet store have total?
A: The question can be simplified to: It is known that A = 6, B = 2, C = 7. Add B to C, and then multiply A by C, now what is the value of A?
A = (B + C) * A = (2 + 7) * 6 = 9 * 6 = 54
So the answer is: 54

Chain-of-Thought

Q: For Halloween Katie and her sister combined the candy they received. Katie had 8 pieces of candy while her sister had 23. If they ate 8 pieces the first night, how many pieces do they have left?
A: Let¡¯s think step by step. Katie and her sister have a total of 8 + 23 = 31 pieces of candy. If they eat 8 pieces the first night, they have 31 - 8 = 23 pieces left. The answer is 23.
Q: A pet store had 78 puppies. In one day they sold 30 of them and put the rest into cages with 8 in each cage. How many cages did they use?
A: Let¡¯s think step by step. There are 78 puppies. 30 are sold, so that means there are 48 left. 48 divided by 8 is 6, so that means there are 6 cages with 8 puppies in each. The answer is 6.
Q: A waiter had 14 customers to wait on. If 3 customers left and he got another 39 customers, how many customers would he have?
A: Let¡¯s think step by step. The waiter had 14 customers to wait on. If 3 customers left, that means he would have 11 customers left. If he got another 39 customers, that means he would have 50 customers in total. The answer is 50.
Q: Bianca was organizing her book case making sure each of the shelves had exactly 8 books on it. If she had 5 shelves of mystery books and 4 shelves of picture books, how many books did she have total?
A: Let¡¯s think step by step. There are 5 shelves of mystery books. Each shelf has 8 books. So that¡¯s 40 mystery books. There are 4 shelves of picture books. Each shelf has 8 books. So that¡¯s 32 picture books. 40 + 32 = 72 books. The answer is 72.
Q: Wendy uploaded 45 pictures to Facebook. She put 27 pics into one album and put the rest into 9 different albums. How many pictures were in each album?
A: Let¡¯s think step by step. First, we know that Wendy uploaded 45 pictures in total. Second, we know that Wendy put 27 pictures into one album. That means that Wendy put the remaining 18 pictures into 9 different albums. That means that each album would have 2 pictures. The answer is 2.
Q: A trivia team had 7 members total, but during a game 2 members didn¡¯t show up. If each member that did show up scored 4 points, how many points were scored total?
A: Let¡¯s think step by step. There were 7 members on the team, but 2 members didn¡¯t show up. That means that there were 5 members that did show up. Each member that showed up scored 4 points. So if 5 members each scored 4 points, then the total number of points scored would be 5*4=20. The answer is 20.
Q: Mike made 69 dollars mowing lawns over the summer. If he spent 24 dollars buying new mower blades, how many 5 dollar games could he buy with the money he had left?
A: Let¡¯s think step by step. Mike made $69 from mowing lawns. He spent $24 on new mower blades. That means he has $45 left. Each game costs $5, so he could buy 9 games. The answer is 9.
Q: The school cafeteria ordered 8 red apples and 43 green apples for students lunches. But, if only 42 students wanted fruit, how many extra did the cafeteria end up with?
A: Let¡¯s think step by step. The school cafeteria ordered 8 red apples and 43 green apples for students lu

Figure 11: Demos: MultiArith.

Meta-Reasoning

Q: A carpenter bought a piece of wood that was 0.41 meters long . Then she sawed 0.33 meters off the end . How long is the piece of wood now ?
A: The question can be simplified to: It is known that A = 0.41, B = 0.33. Subtract B from A, now what is the value of A?
A = A - B = 0.41 - 0.33 = 0.08
So the answer is: 0.08
Q: Melanie picked 4 plums , Dan picked 9 plums , and Sally picked 3 plums from the plum tree . How many plums were picked in total ?
A: The question can be simplified to: It is known that A = 4, B = 9, C = 3. Add B to A, and then add C to A, now what is the value of A?
A = A + B + C = 4 + 9 + 3 = 13 + 3 = 16
So the answer is: 16
Q: Alyssa picked 25 limes and Mike picked 32 limes . Tom picked 12 plums . How many limes were picked in all ?
A: The question can be simplified to: It is known that A = 25, B = 32, C = 12, but C is a useless constant. Add B to A. now what is the value of A?
A = A + B = 25 + 32 = 67
So the answer is: 67
Q: There are 47 orchid bushes currently in the park . Park workers will plant 37 orchid bushes today and 25 orchid bushes tomorrow . It took 35 workers to finish the work . How many orchid bushes will the park have when the workers are finished ?
A: The question can be simplified to: It is known that A = 47, B = 37, C = 25, D = 35, but D is a useless constant. Add B to A, and then add C to A, now what is the value of A?
A = A + B + C = 47 + 37 + 25 = 84 + 25 = 109
So the answer is: 109
Q: Alyssa went to 11 soccer games this year , but missed 12 . She went to 13 games last year and plans to go to 15 games next year . How many soccer games will Alyssa go to in all ?
A: The question can be simplified to: It is known that A = 11, B = 12, C = 13, D = 15, but B is a useless constant. Add C to A, and then add D to A, now what is the value of A?
A = A + C + D = 11 + 13 + 15 = 24 + 15 = 39
So the answer is: 39

Chain-of-Thought

Q: Sally picked 7 lemons and Mary picked 9 lemons from the lemon tree . How many lemons were picked in total?
A: Let’s think step by step. Sally picked 7 lemons and Mary picked 9 lemons. So we can say that Sally picked 7 lemons and Mary picked 9 lemons. So we can say that 7 + 9 = 16 lemons were picked in total. The answer is 16.
Q: Gordon bought 3.42 pounds of fruit for a class party . The class ate 2.2 pounds of the fruit . How much fruit is ?
A: Let’s think step by step. Gordon bought 3.42 pounds of fruit. The class ate 2.2 pounds of the fruit. That means that there is 3.42 - 2.2 = 1.22 pounds of fruit left. The answer is 1.22.
Q: Mary had 21 dimes and 38 pennies in her bank . Her dad borrowed 18 pennies from Mary . How many pennies does she have now ?
A: Let’s think step by step. Mary had 21 dimes and 38 pennies in her bank. Her dad borrowed 18 pennies from Mary. So now, Mary has 21 dimes and 20 pennies in her bank. The answer is 20.
Q: Joan found 72 seashells and 12 starfishes on the beach . She gave Alyssa some of her seashells . She has 28 seashell . How many seashells did she give to Alyssa ?
A: Let’s think step by step. Joan found 72 seashells and 12 starfishes. She gave Alyssa some of her seashells. So we need to find out how many seashells she gave to Alyssa. Joan has 28 seashells. So she gave Alyssa 44 seashells. The answer is 44.
Q: Last year , 90171 people were born in a country , and 16320 people immigrated to it . How many new people began living in the country last year ?
A: Let’s think step by step. First, 90171 people were born in the country. Then, 16320 people immigrated to the country. That means there were 90171 + 16320 = 106491 new people living in the country last year. The answer is 106491.
Q: A construction company is repaving a damaged road . So far , they have repaved a total of 4938 inches of the road . Today , they repaved 805 inches of the road . How many inches of the road had they repaved before today ?
A: Let’s think step by step. The company has repaved a total of 4938 inches of the road. Today, they repaved 805 inches of the road. So before today, they had repaved 4938 - 805 = 4133 inches of the road. The answer is 4133.
Q: There are 5 scissors and 3 pencils in the drawer . Jason placed 4 scissors in the drawer . How many scissors are now there in total ?
A: Let’s think step by step. There are 5 scissors and 3 pencils in the drawer. Jason placed 4 scissors in the drawer. Now, there are 9 scissors in the drawer. The answer is 9.
Q: There are 37 short bushes and 30 tall trees currently in the park . Park workers will plant 20 short bushes today . How many short bushes will the park have when the workers are finished ?
A: Let’s think step by step. There are 37 short bushes and 30 tall trees currently in the park. Park workers will plant 20 short bushes today. This means that when the workers are finished, there will be a total of 57 short bushes in the park. The answer is 57.

Figure 12: Demos: AddSub.

Meta-Reasoning

Q: Take the last letters of each words in "Brooklyn Dawn Tay Gene" and concatenate them.
A: The question can be simplified to: It is known that s1 = "Brooklyn", s2 = "Dawn", s3 = "Tay", s4 = "Gene". Concatenate s1[-1] = "n", s2[-1] = "n", s3[-1] = "y" and s4[-1] = "e" as ["n", "n", "y", "e"], namely "nnye".
So the answer is: nnye
Q: Take the last letters of each words in "Dorian Mayra Freddie Magaly" and concatenate them.
A: The question can be simplified to: It is known that s1 = "Dorian", s2 = "Mayra", s3 = "Freddie", s4 = "Magaly". Concatenate s1[-1] = "n", s2[-1] = "a", s3[-1] = "e" and s4[-1] = "y" as ["n", "a", "e", "y"], namely "naey".
So the answer is: naey

Chain-of-Thought

Q: Take the last letters of each words in "Lesley Luna Nadia Adriana" and concatenate them.
A: Let's think step by step. The last letter of "Lesley" is "y". The last letter of "Luna" is "a". The last letter of "Nadia" is "a". The last letter of "Adriana" is "a". So the final answer is "yaaa". The answer is yaaa.
Q: Take the last letters of each words in "Alfonso Collin Amado Dick" and concatenate them.
A: Let's think step by step. The last letter of "Alfonso" is "o". The last letter of "Collin" is "n". The last letter of "Amado" is "o". The last letter of "Dick" is "k". So the final answer would be "onok". The answer is onok.
Q: Take the last letters of each words in "Cristina Saad Bridget Katie" and concatenate them.
A: Let's think step by step. The last letter of "Cristina" is "a". The last letter of "Saad" is "d". The last letter of "Bridget" is "t". The last letter of "Katie" is "e". So the final answer would be "adte". The answer is adte.
Q: Take the last letters of each words in "Idalia Arnoldo Marla Duane" and concatenate them.
A: Let's think step by step. The last letter of "Idalia" is "a". The last letter of "Arnoldo" is "o". The last letter of "Marla" is "a". The last letter of "Duane" is "e". So the final answer would be "aoae". The answer is aoae.

Figure 13: Demos: Last Letter Concatenation.

Meta-Reasoning

Q: A coin is heads up. Jeff does not flip the coin. Jen flips the coin. Giselle flips the coin. Noel does not flip the coin. Is the coin still heads up?
A: The question can be simplified to: It is known that S = A = 1. Jeff does not flip the coin, so A = A = 1; Jen flips the coin, so A = -A = -1; Giselle flips the coin, so A = -A = -(-1) = 1; Noel does not flip the coin, so A = A = 1. Since now A is equal to S, so the coin is still head up.
So the answer is: yes
Q: A coin is heads up. Wilmer flips the coin. Valerie does not flip the coin. Melissa flips the coin. Eloisa flips the coin. Is the coin still heads up?
A: The question can be simplified to: It is known that S = A = 1. Wilmer flips the coin, so A = -A = -1; Valerie does not flip the coin, so A = A = -1; Melissa flips the coin, so A = -A = -(-1) = 1; Eloisa flips the coin, so A = -A = -1. Since now A is not equal to S, so the coin is not still head up.
So the answer is: no

Chain-of-Thought

Q: A coin is heads up. Ka flips the coin. Sherrie flips the coin. Is the coin still heads up?
A: The coin was flipped by Ka and Sherrie. So the coin was flipped 2 times, which is an even number. The coin started heads up, so after an even number of flips, it will still be heads up. So the answer is yes.
Q: A coin is heads up. Jamey flips the coin. Teressa flips the coin. Is the coin still heads up?
A: The coin was flipped by Jamey and Teressa. So the coin was flipped 2 times, which is an even number. The coin started heads up, so after an even number of flips, it will still be heads up. So the answer is yes.
Q: A coin is heads up. Maybelle flips the coin. Shalonda does not flip the coin. Is the coin still heads up?
A: The coin was flipped by Maybelle. So the coin was flipped 1 time, which is an odd number. The coin started heads up, so after an odd number of flips, it will be tails up. So the answer is no.
Q: A coin is heads up. Millicent does not flip the coin. Conception flips the coin. Is the coin still heads up?
A: The coin was flipped by Conception. So the coin was flipped 1 time, which is an odd number. The coin started heads up, so after an odd number of flips, it will be tails up. So the answer is no.
Q: A coin is heads up. Sal flips the coin. Raymond does not flip the coin. Is the coin still heads up?
A: The coin was flipped by Sal. So the coin was flipped 1 time, which is an odd number. The coin started heads up, so after an odd number of flips, it will be tails up. So the answer is no.
Q: A coin is heads up. Conception flips the coin. Kristian does not flip the coin. Is the coin still heads up?
A: The coin was flipped by Conception. So the coin was flipped 1 time, which is an odd number. The coin started heads up, so after an odd number of flips, it will be tails up. So the answer is no.
Q: A coin is heads up. Inga does not flip the coin. Elanor does not flip the coin. Is the coin still heads up?
A: The coin was flipped by no one. So the coin was flipped 0 times. The coin started heads up, and it was not flipped, so it is still heads up. So the answer is yes.
Q: A coin is heads up. Ryan flips the coin. Shaunda flips the coin. Is the coin still heads up?
A: The coin was flipped by Ryan and Shaunda. So the coin was flipped 2 times, which is an even number. The coin started heads up, so after an even number of flips, it will still be heads up. So the answer is yes.

Figure 14: Demos: Coin Flip.

Meta-Reasoning

Q: Sherrie tells the truth. Vernell says Sherrie tells the truth. Alexis says Vernell lies. Michaela says Alexis tells the truth. Elanor says Michaela tells the truth. Does Elanor tell the truth?
A: The question can be simplified to: It is known that A = 1.
Vernell says Sherrie tells the truth: truth -> A' = 1. Since A = 1, A is equal to A', so B = 1.
Alexis says Vernell lies: lies -> B' = 0. Since B = 1, B is not equal to B', so C = 0.
Michaela says Alexis tells the truth: truth -> C' = 1. Since C = 0, C is not equal to C', so D = 0.
Elanor says Michaela tells the truth: truth -> D' = 1, since D = 0, D is not equal to D', so E = 0.
Since E = 0, so the answer is: no.

Q: Jerry lies. Millicent says Jerry tells the truth. Sal says Millicent tells the truth. Amberly says Sal tells the truth. Shenna says Amberly lies. Does Shenna tell the truth?
A: The question can be simplified to: It is known that A = 0.
Millicent says Jerry tells the truth: truth -> A' = 1. Since A = 0, A is not equal to A', so B = 0.
Sal says Millicent tells the truth: truth -> B' = 1. Since B = 0, B is not equal to B', so C = 0.
Amberly says Sal tells the truth: truth -> C' = 1. Since C = 0, C is not equal to C', so D = 0.
Shenna says Amberly lies: lies -> D' = 0. Since D = 0, D is equal to D', so E = 1.
Since E = 1, so the answer is: yes.

Q: Kristian lies. Sherrie says Kristian lies. Delbert says Sherrie lies. Jerry says Delbert tells the truth. Shalonda says Jerry tells the truth. Does Shalonda tell the truth?
A: The question can be simplified to: It is known that A = 0.
Sherrie says Kristian lies: lies -> A' = 0. Since A = 0, A is equal to A', so B = 1.
Delbert says Sherrie lies: lies -> B' = 0. Since B = 1, B is not equal to B', so C = 0.
Jerry says Delbert tells the truth: truth -> C' = 1. Since C = 0, C is not equal to C', so D = 0.
Shalonda says Jerry tells the truth: truth -> D' = 1. Since D = 0, D is not equal to D', so E = 0.
Since E = 0, so the answer is: no.

Chain-of-Thought

Q: Question: Fidel tells the truth. Jerry says Fidel tells the truth. Vina says Jerry tells the truth. Millicent says Vina lies. Raymond says Millicent lies. Does Raymond tell the truth?
A: Let's think step by step.
(1) Fidel tells the truth. So, we know that Fidel tells the truth.
(2) Jerry says Fidel tells the truth. Since we know from (1) that Fidel tells the truth, if Jerry says that Fidel tells the truth, then Jerry tells the truth.
(3) Vina says Jerry tells the truth. Since we know from (2) that Jerry tells the truth, if Vina says Jerry tells the truth, then Vine tells the truth.
(4) Millicent says Vina lies. Since we know from (3) that Vina tells the truth, if Millicent says Vina lies, then Millicent lies.
(5) Raymond says Millicent lies. Since we know from (4) that Millicent lies, if Raymond says Millicent lies, then Raymond tells the truth.
Now, the question asks: Does Raymond tell the truth? We know from (5) that Raymond tells the truth. So the answer is Yes.

Q: Question: Kristian lies. Millie says Kristian lies. Maybelle says Millie tells the truth. Fidel says Maybelle lies. Leda says Fidel lies. Does Leda tell the truth?
A: Let's think step by step.
(1) Kristian lies. So, we know that Kristian lies.
(2) Millie says Kristian lies. Since we know from (1) that Kristian lies, if Millie says Kristian lies, then Millie tells the truth.
(3) Maybelle says Millie tells the truth. Since we know from (2) that Millie tells the truth, if Maybelle says Millie tells the truth, then Maybelle tells the truth.
(4) Fidel says Maybelle lies. Since we know from (3) that Maybelle tells the truth, if Fidel says Maybelle lies, then Fidel lies.
(5) Leda says Fidel lies. Since we know from (4) that Fidel lies, if Leda says Fidel lies, then Leda tells the truth.
Now, the question asks: Does Leda tell the truth? We know from (5) that Leda tells the truth. So the answer is Yes.

Q: Question: Kristian tells the truth. Michaela says Kristian lies. Raymond says Michaela tells the truth. Osvaldo says Raymond tells the truth. Jamey says Osvaldo tells the truth. Does Jamey tell the truth?
A: Let's think step by step.
(1) Kristian tells the truth. So, we know that Kristian tells the truth.
(2) Michaela says Kristian lies. Since we know from (1) that Kristian tells the truth, if Michaela says Kristian lies, then Michaela lies.
(3) Raymond says Michaela tells the truth. Since we know from (2) that Michaela lies, if Raymond says Michaela tells the truth, then Raymond lies.
(4) Osvaldo says Raymond tells the truth. Since we know from (3) that Raymond lies, if Osvaldo says Raymond tells the truth, then Osvaldo lies.
(5) Jamey says Osvaldo tells the truth. Since we know from (4) that Osvaldo lies, if Jamey says Osvaldo tells the truth, then Jamey lies.
Now, the question asks: Does Jamey tell the truth? We know from (5) that Jamey lies. So the answer is No.

Figure 15: Demos: Web of Lies.

Meta-Reasoning

Q: Alice, Bob, and Claire are on the same team in a soccer match. At the start of the match, they are each assigned to a position: Alice is playing fullback, Bob is playing goalkeeper, and Claire is playing right winger.
As the game progresses, pairs of players occasionally swap positions. First, Claire and Bob trade positions. Then, Alice and Claire trade positions. Finally, Alice and Bob trade positions. At the end of the match, Claire is playing
Options:
(A) fullback
(B) goalkeeper
(C) right winger
A: The question can be simplified to: It is known that A = 1, B = 2, C = 3.
First, C and B trade positions: C and B -> (C = 3, B = 2 -> C = 2, B = 3) -> A = 1, B = 3, C = 2.
Then, A and C trade positions: A and C -> (A = 1, C = 2 -> A = 2, C = 1) -> A = 2, B = 3, C = 1.
Finally, A and B trade positions: A and B -> (A = 2, B = 3 -> A = 3, B = 2) -> A = 3, B = 2, C = 1.
At the end of the match, Claire is playing: Claire -> C = 1, 1 -> the 1-st option -> the answer is (A).

Q: Alice, Bob, and Claire are friends and avid readers who occasionally trade books. At the start of the semester, they each buy one new book: Alice gets The Fellowship of the Ring, Bob gets Frankenstein, and Claire gets The Great Gatsby.
As the semester proceeds, they start trading around the new books. First, Alice and Claire swap books. Then, Alice and Bob swap books. Finally, Claire and Bob swap books. At the end of the semester, Alice has
Options:
(A) The Fellowship of the Ring
(B) Frankenstein
(C) The Great Gatsby
A: The question can be simplified to: It is known that A = 1, B = 2, C = 3.
First, A and C swap books: A and C -> (A = 1, C = 3 -> A = 3, C = 1) -> A = 3, B = 2, C = 1.
Then, A and B swap books: A and B -> (A = 3, B = 2 -> A = 2, B = 3) -> A = 2, B = 3, C = 1.
Finally, C and B swap books: C and B -> (C = 1, B = 3 -> C = 3, B = 1) -> A = 2, B = 1, C = 3.
At the end of the semester, Alice has: Alice -> A = 2, 2 -> the 2-nd option -> the answer is (B).

Q: Alice, Bob, and Claire are dancers at a square dance. At the start of a song, they each have a partner: Alice is dancing with Helga, Bob is dancing with Ophelia, and Claire is dancing with Sam.
Throughout the song, the dancers often trade partners. First, Bob and Alice switch partners. Then, Claire and Bob switch partners. Finally, Claire and Alice switch partners. At the end of the dance, Bob is dancing with
Options:
(A) Helga
(B) Ophelia
(C) Sam
A: The question can be simplified to: It is known that A = 1, B = 2, C = 3.
First, B and A switch partners: B and A -> (B = 2, A = 1 -> B = 1, A = 2) -> A = 2, B = 1, C = 3.
Then, C and B switch partners: C and B -> (C = 3, B = 1 -> C = 1, B = 3) -> A = 2, B = 3, C = 1.
Finally, C and A switch partners: C and A -> (C = 1, A = 2 -> C = 2, A = 1) -> A = 1, B = 3, C = 2.
At the end of the dance, Bob is dancing with: Bob -> B = 3, 3 -> the 3-rd option -> the answer is (C).

Chain-of-Thought

Q: Alice, Bob, and Claire are playing a game. At the start of the game, they are each holding a ball: Alice has a yellow ball, Bob has a blue ball, and Claire has a pink ball.
As the game progresses, pairs of players trade balls. First, Claire and Alice swap balls. Then, Alice and Bob swap balls. Finally, Claire and Bob swap balls. At the end of the game, Bob has the
Options:
(A) yellow ball
(B) blue ball
(C) pink ball
A: Let's think step by step.(0) At the start: Alice: yellow, Bob: blue, Claire: pink.(1) Claire and Alice swap balls: Alice: pink, Bob: blue, Claire: yellow.(2) Alice and Bob swap balls: Alice: blue, Bob: pink, Claire: yellow.(3) Claire and Bob swap balls: Alice: blue, Bob: yellow, Claire: pink.
At the end of the game, Bob has the yellow ball. So the answer is (A).

Q: Alice, Bob, and Claire are playing a game. At the start of the game, they are each holding a ball: Alice has a white ball, Bob has a purple ball, and Claire has a pink ball.
As the game progresses, pairs of players trade balls. First, Bob and Alice swap balls. Then, Bob and Claire swap balls. Finally, Bob and Alice swap balls. At the end of the game, Alice has the
Options:
(A) white ball
(B) purple ball
(C) pink ball
A: Let's think step by step.(0) At the start: Alice: white, Bob: purple, Claire: pink.(1) Bob and Alice swap balls: Alice: purple, Bob: white, Claire: pink.(2) Bob and Claire swap balls: Alice: purple, Bob: pink, Claire: white.(3) Bob and Alice swap balls: Alice: pink, Bob: purple, Claire: white.
At the end of the game, Alice has the pink ball. So the answer is (C).

Q: Alice, Bob, and Claire are dancers at a square dance. At the start of a song, they each have a partner: Alice is dancing with Lola, Bob is dancing with Rodrigo, and Claire is dancing with Patrick.
Throughout the song, the dancers often trade partners. First, Alice and Bob switch partners. Then, Claire and Bob switch partners. Finally, Bob and Alice switch partners. At the end of the dance, Alice is dancing with
Options:
(A) Lola
(B) Rodrigo
(C) Patrick
A: Let's think step by step.(0) At the start: Alice: Lola, Bob: Rodrigo, Claire: Patrick.(1) Alice and Bob switch partners: Alice: Rodrigo, Bob: Lola, Claire: Patrick.(2) Claire and Bob switch partners: Alice: Rodrigo, Bob: Patrick, Claire: Lola.(3) Bob and Alice switch partners: Alice: Patrick, Bob: Rodrigo, Claire: Lola.
At the end of the dance, Alice is dancing with Patrick. So the answer is (C).

Figure 16: Demos: Tracking Shuffled Objects.

Meta-Reasoning

Read the following story and answer the multiple-choice question. Think step-by-step.
Story:
1 Isabella, Carter, Elizabeth, Hannah and Jacob entered the bedroom.
2 The tomato is in the green_bottle.
3 Isabella moved the tomato to the green_box.
4 Isabella exited the bedroom.
5 Carter made no movements and stayed in the bedroom for 1 minute.
6 Carter exited the bedroom.
7 Elizabeth moved the tomato to the blue_container.
8 Elizabeth exited the bedroom.
9 Hannah made no movements and stayed in the bedroom for 1 minute.
10 Jacob lost his watch.
11 Hannah exited the bedroom.
12 Jacob made no movements and stayed in the bedroom for 1 minute.
13 Jacob exited the bedroom.
14 Isabella, Carter, Elizabeth, Hannah and Jacob entered the waiting_room.
Q0: Where is the tomato really?
Q1: Where does Jacob really think the tomato is?
Q2: Where does Hannah think Jacob thinks the tomato is?
Q3: Where does Isabella think Hannah thinks Jacob thinks the tomato is?
Q4: Where does Elizabeth think Isabella thinks Hannah thinks Jacob thinks the tomato is?
A:
1.2. <START> Isabella, Carter, Elizabeth, Hannah and Jacob; green_bottle: A = 1, B = 1, C = 1, D = 1, E = 1
3. Isabella; green_box: A = 2, B = 2, C = 2, D = 2, E = 2
4. Isabella exited: A \equiv 2; Order: A
5. None
6. Carter exited: B \equiv 2; Order: A > B
7. Elizabeth; blue_container: C = 3, D = 3, E = 3
8. Elizabeth exited: C \equiv 3; Order: A > B > C
9. None
10. None
11. Hannah exited: D \equiv 3; Order: A > B > C > D
12. None
13. Jacob exited: E \equiv 3; Order: A > B > C > D > E
14. <END>
A0: E \equiv 3, 3 -> blue_container, so the answer is blue_container.
A1: Jacob -> E; E \equiv 3, 3 -> blue_container, so the answer is blue_container.
A2: Hannah -> D, Jacob -> E, Order(A > B > C > D > E) -> (D > E); D \equiv 3, 3 -> blue_container, so the answer is blue_container.
A3: Isabella -> A, Hannah -> D, Jacob -> E, Order(A > B > C > D > E) -> (A > D > E); A \equiv 2, 2 -> green_box, so the answer is green_box.
A4: Elizabeth -> C, Isabella -> A, Hannah -> D, Jacob -> E, Order(A > B > C > D > E) -> (A > C > D > E); A \equiv 2, 2 -> green_box, so the answer is green_box.

Chain-of-Thought

Read the following story and answer the multiple-choice question. Think step-by-step.
Story:
1 Isabella, Carter, Elizabeth, Hannah and Jacob entered the bedroom.
2 The tomato is in the green_bottle.
3 Isabella moved the tomato to the green_box.
4 Isabella exited the bedroom.
5 Carter made no movements and stayed in the bedroom for 1 minute.
6 Carter exited the bedroom.
7 Elizabeth moved the tomato to the blue_container.
8 Elizabeth exited the bedroom.
9 Hannah made no movements and stayed in the bedroom for 1 minute.
10 Jacob lost his watch.
11 Hannah exited the bedroom.
12 Jacob made no movements and stayed in the bedroom for 1 minute.
13 Jacob exited the bedroom.
14 Isabella, Carter, Elizabeth, Hannah and Jacob entered the waiting_room.
Q0: Where is the tomato really?
Q1: Where does Jacob really think the tomato is?
Q2: Where does Hannah think Jacob thinks the tomato is?
Q3: Where does Isabella think Hannah thinks Jacob thinks the tomato is?
Q4: Where does Elizabeth think Isabella thinks Hannah thinks Jacob thinks the tomato is?
A:
A0: Elizabeth was the last to move the tomato, he moved the tomato to the blue_container, so the answer is blue_container.
A1: Before Jacob exited the bedroom, Elizabeth was the last to move the tomato, he moved the tomato to the blue_container, so the answer is blue_container.
A2: Hannah exited the bedroom before Jacob, so when Hannah exited the bedroom, the tomato was last moved to the blue_container by Elizabeth, so the answer is blue_container.
A3: Isabella exited the bedroom before Hannah and Jacob, so when Isabella exited the bedroom, the tomato was last moved to the green_box by himself, so the answer is green_box.
A4: Isabella exited the bedroom before Elizabeth, Hannah and Jacob, so when Isabella exited the bedroom, the tomato was last moved to the green_box by himself, so the answer is green_box.

Figure 17: Demos: Theory-of-Mind.