# Enhancing Chain-of-Thoughts Prompting with Iterative Bootstrapping in Large Language Models

**Jiashuo Sun**
Xiamen University

**Yi Luo**
Xiamen University

**Yeyun Gong**
Microsoft Research Asia

**Chen Lin***
Xiamen University

**Yelong Shen**
Microsoft Research Asia

**Jian Guo**
IDEA Research

**Nan Duan**
Microsoft Research Asia

## Abstract

Large language models (LLMs) can achieve highly effective performance on various reasoning tasks by incorporating step-by-step chain-of-thought (CoT) prompting as demonstrations. However, the reasoning chains of demonstrations generated by LLMs are prone to errors, which can subsequently lead to incorrect reasoning during inference. Furthermore, inappropriate exemplars (overly simplistic or complex), can affect overall performance among varying levels of difficulty. We introduce Iter-CoT (**Ite**rative bootst**r**apping in **C**hain-**o**f-**T**houghts Prompting), an iterative bootstrapping approach for selecting exemplars and generating reasoning chains. By utilizing iterative bootstrapping, our approach enables LLMs to autonomously rectify errors, resulting in more precise and comprehensive reasoning chains. Simultaneously, our approach selects challenging yet answerable questions accompanied by reasoning chains as exemplars with a moderate level of difficulty, which enhances the LLMs' generalizability across varying levels of difficulty. Experimental results indicate that Iter-CoT exhibits superiority, achieving competitive performance across three distinct reasoning tasks on eleven datasets.[2]

## 1 Introduction

Recently, Large language models (LLMs) (Chowdhery et al., 2022; Thoppilan et al., 2022; Rae et al., 2021; Smith et al., 2022; Scao et al., 2022) have demonstrated remarkable performance in complex reasoning tasks, including arithmetic, commonsense, and symbolic reasoning. LLMs utilize Chain-of-Thought (CoT) (Wei et al., 2022) and In-Context Learning (ICL) (Brown et al., 2020) to achieve impressive performance, which generates intermediate reasoning steps and employ task-specific exemplars to provide demonstration.

The conventional CoT prompting involves human annotation (Manual-CoT (Wei et al., 2022)), where human manually constructs question-rationale-answer pairs for each example as demonstrations. A new paradigm, known as Zero-shot-CoT (Kojima et al., 2022), enables LLMs to generate reasoning chains using the prompt "let's think step by step". In recent research, a novel approach Auto-CoT (Zhang et al., 2022) has been proposed, which uses a cluster-based method to identify appropriate questions and generates reasoning chains for each question using LLMs. These questions, along with the corresponding reasoning chains, serve as demonstrations. The performance of Auto-CoT exceeds the previous methods, achieving new **state-of-the-art**.

In practice, reasoning chains generated by LLMs have demonstrated superior performance compared to human-designed annotations (Shum et al., 2023; Zhang et al., 2022; Kojima et al., 2022). How-

*Corresponding author, chenlin@xmu.edu.cn
[2]Our code will be publicly available at `https://github.com/GasolSun36/Iter-CoT`

ever, selecting exemplars with appropriate levels of difficulty still remains challenging. Previous works (Shum et al., 2023; Fu et al., 2022) mainly utilize the number of steps in the manual-writing rationale as a criterion to judge the level of difficulty of the question, with fewer steps indicating simpler questions and more steps indicating more complex question. Nonetheless, Human bias can easily lead to inaccurate judgment, given that questions which humans perceive as simple or complex may in fact be the exact opposite for the LLMs, which makes the demonstration inappropriate. Secondly, reasoning chains of demonstrations generated by LLMs are prone to errors, which can significantly reduce overall performance (Zhang et al., 2022). Moreover, previous works merely combine the question and template during the generation of demonstrations, without incorporating additional information, offering worse performance in guiding the LLMs to generate the final answer.

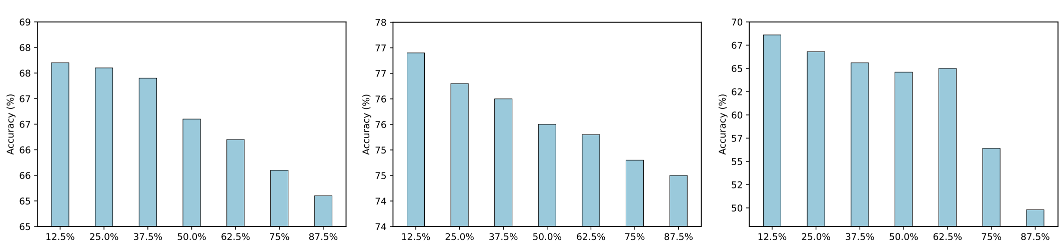
In order to address the issues above, we propose Iter-CoT (**Ite**rative bootst**r**apping in **C**hain-**o**f-**T**houghts Prompting). It consists of two methods: weak bootstrapping and strong bootstrapping. Weak bootstrapping only uses the correctness of answers as supervisory information, while strong bootstrapping incorporates errors in reasoning chains as additional supervisory information. Both methods enable LLMs to rectify and summarize the more precise, comprehensive reasoning chains enriched with contextual information. By incorporating bootstrapping, Iter-CoT allows LLMs to self-correct and identify challenging yet answerable questions as demonstrations, enhancing the LLMs' generalizability across varying levels. The resulting exemplars with moderate difficulty and high-quality reasoning chains can significantly improve LLMs' reasoning performance. Moreover, we introduce a new criterion for assessing the difficulty level of questions based on whether they can be easily solved by LLMs. Analysis indicates that the hop-based criterion and ours yield comparable yet distinguishable results, as shown in Appendix A.4.

We evaluate Iter-CoT on three distinct reasoning tasks (arithmetic, commonsense and symbolic) across eleven datasets (GSM8K (Cobbe et al., 2021a), AQuA (Ling et al., 2017), AddSub (Hosseini et al., 2014), SingleEq (Koncel-Kedziorski et al., 2015), SVAMP (Patel et al., 2021), ASDiV (Miao et al., 2020), CSQA (Talmor et al., 2019), StrategyQA (Geva et al., 2021), Date Understanding, Last Letter Concatenation (Wei et al., 2022) and Object Tracking (Kojima et al., 2022)). The experimental results show that Iter-CoT outperforms existing prompting approaches, and demonstrate that LLMs have the capability to enhance their performance by integrating iterative bootstrapping for self-correction and summarization to obtain better reasoning chains for demonstrations.

## 2 Motivation

### 2.1 Inappropriate Exemplars can Reduce Performance

As we mentioned in Section 1, the previous CoT prompting methods encountered a common and highly challenging problem of sampling exemplars that are either too simplistic or excessively complex (hop-based criterion). When overly simplistic examples are selected as exemplars during the reasoning process, the demonstrations become redundant and limited, as the LLMs can already answer these simple questions correctly. Moreover, sampling excessively complex exemplars can improve the accuracy of complex questions, but perform poorly in simple questions (Shum et al., 2023).




(a) GSM8K (b) CSQA (c) Letter

Figure 1: Effect of wrong exemplars among three different reasoning tasks (arithmetic, commonsense and symbolic reasoning).

The second issue occurs when the LLMs produce incorrect reasoning chains in the demonstration, leading to confusion and poor performance during inference. The final accuracy tends to decrease as the number of incorrect demonstrations increases. Specifically, when the percentage of incorrect demonstrations increases from 12.5% to 87.5%, the accuracy drops significantly from 68.2% to 65.6% for GSM8K, 77.4% to 75.0% for CSQA and 68.6% to 49.8% for Letter, respectively, representing a substantial decrease (shown in Figure 1).

The above two issues are generally present in the current mainstream methods of selecting questions and generating reasoning chains (Wei et al., 2022; Zhang et al., 2022; Shum et al., 2023; Diao et al., 2023; Fu et al., 2022). Zhang et al. (2022) introduced a clustering method and observed that erroneous examples could be effectively grouped into a single cluster, alleviating excessive incorrect demonstrations. Shum et al. (2023) filter out the wrong examples directly in the selection phase, and train a demonstrations selection model via reinforcement learning. However, the computational cost increases as each test question necessitates to select the most appropriate exemplars during the inference phase. These above studies minimize the use of erroneous examples to prevent their adverse effects. We find that incorrect or erroneous examples can offer a level of complexity to LLMs, thereby providing valuable learning. This concept is analogous to the way in which students can enhance their problem-solving skills by engaging with challenging yet answerable questions. Therefore, the investigation of how to leverage the erroneous examples to enhance LLMs' performance is worth exploring.

### 2.2 Large Language Models can Self-Correct with Bootstrapping

Wang et al. (2022) introduced a novel decoding strategy as a substitute for the conventional greedy decoding approach used in chain-of-thought prompting. The self-consistency technique has been shown to substantially enhance the accuracy of the model and effectively serve as an ensemble method. Furthermore, this technique demonstrated that LLMs have the potential to generate the correct reasoning chain for questions that are initially answered erroneously, consequently leading to the correct results.

In order to provide further quantitative validation of the self-correction capability of LLMs, we conduct a series of experiments on the GSM8K (Cobbe et al., 2021a) dataset. Specifically, we apply the Zero-Shot-CoT (Kojima et al., 2022) method for the training set to generate reasoning chains. We then identify incorrect reasoning chains and provide weak bootsrapping in the form of `Revised-Prompt`, as illustrated in Figure 3. The LLMs subsequently generate rationales iteratively using such prompts. As expected, we observe that the performance of the LLMs is further improved (increasing from 4089(54.7%) to 4898(59.1%) after the first iteration and continuing to improve with subsequent iterations, ultimately reaching a peak of 5726(76.6%), as shown in Figure 2). This observation suggests that LLMs can autonomously rectify errors with the assistance of weak bootstrapping.

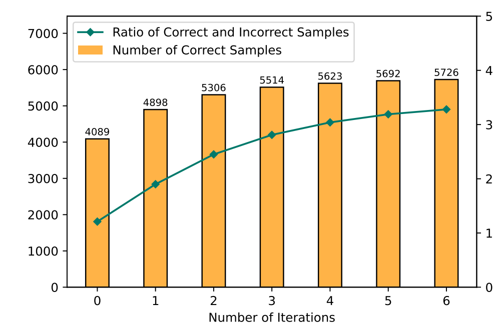


Figure 2: Performance of integrating weakbootstrapping.

Based on the observation, we introduce Iter-CoT (**Ite**rative bootst**r**apping in **C**hain-**o**f-**T**houghts Prompting), a novel approach that leverages the question and the rationales during each iteration of bootstrapping, enabling the model to reflect on its errors. This is different from other previous approaches: our method uses richer contextual information to prompt the LLMs to produce more precise and comprehensive reasoning chains in the demonstration.

## 3 Iter-CoT: Iterative Bootstrapping in Chain-of-Thought Prompting

We propose an iterative bootstrapping method for constructing new reasoning chains while autonomously rectifying errors. Iter-CoT consists of two main patterns: weak bootstrapping and strong

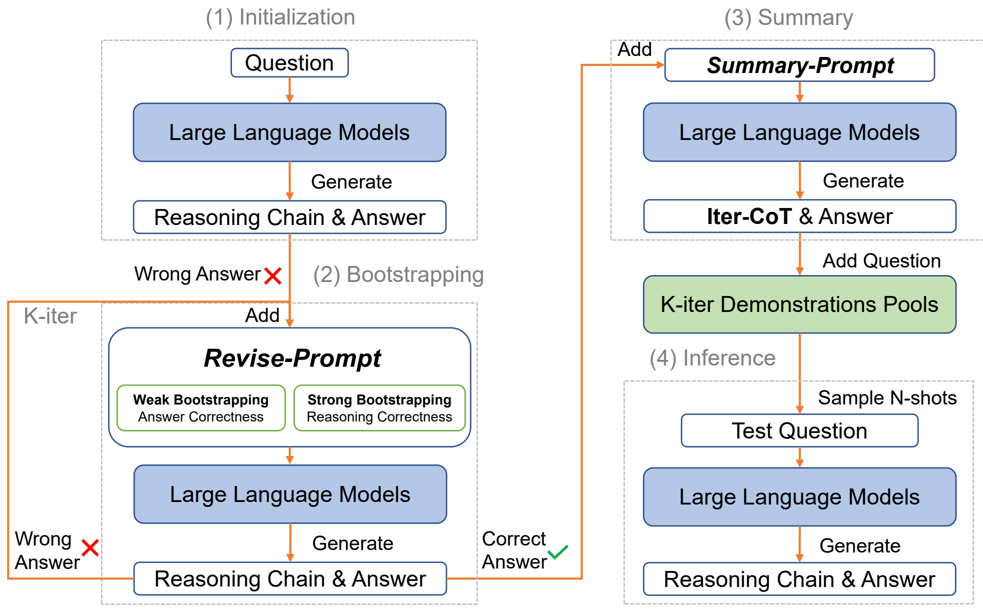



Figure 3: Illustrations of our proposed iterative bootstrapping approach: (1) **Initialization**: we query the LLMs to generate reasoning chain and answer with Zero-Shot-CoT method. (2) **Bootstrapping**: we use `Revise-Prompt` to revise the reasoning chain repeatedly until the generated CoT is completely accurate. (3) **Summary**: we use `Summary-Prompt` to generate the `Iter-CoT` based on the contextual information provided within the overall process. (4) **Inference**: inference new questions in test set with the sampled demonstrations.

bootstrapping. Both patterns aim to facilitate the LLMs to rectify errors in the reasoning chains by introducing supervisory information. Iter-CoT also enables the LLMs to summarize reasoning, resulting in more precise and comprehensive reasoning chains. In contrast to manually correcting errors in the reasoning chains generated by LLMs, Iter-CoT only bootstraps the LLMs to correct their own mistakes. These exemplars and rationales are then composed as demonstrations. The overall process is shown in Figure 3.

### 3.1 Weak Bootstrapping

As mentioned in Section 2.2, the LLMs have the ability to self-correct. We prompt the model to autonomously correct its erroneous reasoning chains by providing weak iterative bootstrapping. This process can be categorized into four distinct stages: initialization, bootstrapping, summarization, and inference. The illustration with a specific example is shown in Figure 8.

**Initialization** We utilize the Zero-Shot-Cot (Kojima et al., 2022) method on the training set to prompt LLMs to provide reasoning chains and answers. Subsequently, we record the error examples for the next stage.

**Bootstrapping** We utilize the `Revise-Prompt` for each erroneous instance, instructing the LLM of its incorrectness and allowing it to generate a new response. This approach aims to prompt the LLMs to autonomously recognize logical inconsistencies until a precise answer is obtained, ultimately ensuring the precision and validity of the reasoning chain. Particularly, we only use the correctness of answers as supervisory information in weak bootstrapping.

**Summarization** Upon obtaining the correct answer, which derives the accurate reasoning chains (based on the assumption that when a correct answer is generated, the reasoning chain is most likely correct), the `Summary-Prompt` method is employed to prompt LLMs to generate the final reasoning chain (Iter-CoT) based on prior rationales (including both the wrong rationales and the correct rationales). The contextual information incorporated within the entire process enables LLMs to produce more precise and comprehensive reasoning chains.

**Inference** After completing the above process, we combine the current question with the `Iter-CoT` and subsequently incorporate it into the demonstration pool. During the inference phase, we randomly sample some examples from the demonstrations pool and adopt them as demonstrations for ICL.

The weak bootstrapping approach effectively selects moderately difficult questions (i.e., challenging yet answerable questions). Simultaneously, it prompts LLMs to generate precise, comprehensive reasoning chains for demonstrations.

### 3.2 Strong Bootstrapping

The concept of strong bootstrapping is similar to weak bootstrapping, yet there exist several distinctions between them. The initial phase of strong bootstrapping is the same as weak bootstrapping. It involves utilizing Zero-Shot-CoT to perform inference on the complete training set, where incorrect rationales are identified and subsequently recorded. The bootstrapping stage requires more human intervention, wherein annotators point out the specific errors and identify incorrect reasoning steps generated by LLMs in the `Revise-Prompt` until LLMs produce the correct answer. Finally, we use the $n$ sampled questions and their corresponding reasoning chains as demonstrations for inference on the test set. The entire process of strong bootstrapping with the specific example is shown in Figure 9.

The process of strong bootstrapping is regarded as a distinct form of human annotation. It differs from conventional manual-designed demonstrations which do not require any inference from LLMs. In our approach, we integrate annotation with the reasoning of LLMs, enabling the annotation as contextual information to rectify its own reasoning chains. Additionally, the strong bootstrapping method empowers us to select more challenging examples for demonstration, enhancing the ability to generalize to both simple and complex questions.

## 4 Experiment

In this section, we introduce the datasets used in this paper and present the experimental results that demonstrate the advantages of our methods from a comparative perspective.

### 4.1 Datasets and Evaluation Metrics

We evaluate our methods on eleven datasets across three categories of different reasoning tasks, including (1) six arithmetic reasoning datasets: GSM8k (Cobbe et al., 2021a), AQuA (Ling et al., 2017), AddSub (Hosseini et al., 2014), SingleEq (Koncel-Kedziorski et al., 2015), SVAMP (Patel et al., 2021) and ASDiv (Miao et al., 2020); (2) four commonsense reasoning datasets: CSQA (Talmor et al., 2019), StrategyQA (Geva et al., 2021), Date Understanding (Wei et al., 2022) and Object Tracking (Kojima et al., 2022); (3) one symbolic reasoning datasets: Letter Concatenation (Wei et al., 2022).

We report the exact match accuracy as our evaluation metric following previous works (Wei et al., 2022; Kojima et al., 2022). Examples of each reasoning task and detailed description of each dataset are shown in Table 6 and Appendix 6.

### 4.2 Baselines

We compare our methods with three baseline approaches: Manual-CoT (Wei et al., 2022), Random-CoT (Wei et al., 2022), and Auto-CoT (Zhang et al., 2022). Manual-CoT involves using manually constructed reasoning chains as exemplars, which are listed in the appendix of Wei et al. (2022). Random-CoT randomly selects $n$ questions from the training set and generates chains using the "let's think step by step" prompt. In contrast, Auto-CoT utilizes clustering techniques to sample questions and generate chains with the same approach. Specifically, we implement Auto-CoT by generateing reasoning steps for the questions provided in the appendix of Zhang et al. (2022) as demonstrations.

| Method | Arithmetic | | | | | | Commonsense | | | | Symbolic |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | GSM8K | AQuA | AddSub | SingleEq | SVAMP | ASDiv | CSQA | STQA | Date | Object | Letter |
| Prior Best | $55.0^{\alpha}$ | $37.9^{\beta}$ | $77.7^{\gamma}$ | $32.5^{\delta}$ | $57.4^{\epsilon}$ | $75.3^{\zeta}$ | $91.2^{\eta}$ | - | - | - | - |
| *UL2-20B** | 4.4 | 23.6 | 18.2 | 20.2 | 12.5 | 16.9 | 51.4 | 53.3 | - | - | 0.0 |
| *LaMDA-137B** | 14.3 | 20.6 | 51.9 | 58.7 | 37.5 | 46.6 | 57.9 | 65.4 | - | - | 13.5 |
| *PaLM-540B** | 56.9 | 35.8 | 91.9 | 92.3 | 79.0 | 73.9 | 79.9 | 77.8 | - | - | 63.0 |
| | | | | | *text-davinci-002* | | | | | | |
| Manual-CoT | 46.9 | 35.8 | - | - | 68.9 | 71.3 | 73.5 | 65.4 | - | - | 56.6 |
| Auto-CoT | 47.9 | 36.5 | - | 87.0 | 69.5 | - | 74.4 | 65.4 | - | - | 59.7 |
| | | | | | *code-davinci-002* | | | | | | |
| Manual-CoT | 63.1 | 45.3 | - | - | 76.4 | 80.4 | 77.9 | 73.2 | - | - | 70.4 |
| Auto-CoT | 62.8 | - | - | - | - | - | - | - | - | - | - |
| | | | | | *ChatGPT* | | | | | | |
| Manual-CoT | 69.3 | 47.2 | 86.5 | 88.9 | 77.2 | 82.5 | 77.1 | **71.1** | 78.6 | - | 72.2 |
| Random-CoT | 67.3 | 43.7 | 84.0 | 87.7 | 77.4 | 80.1 | 77.3 | 65.3 | 67.4 | 40.3 | 68.0 |
| Auto-CoT | 71.4 | 46.8 | 88.8 | 92.5 | 77.6 | - | 72.5 | 63.4 | - | - | 76.0 |
| Iter-CoT(W) | **73.6** | 49.2 | **89.1** | 91.9 | **80.7** | 82.8 | 76.8 | 69.3 | **80.0** | **67.8** | **88.8** |
| Iter-CoT(S) | 72.9 | **52.0** | 85.3 | **92.9** | 80.1 | **83.0** | **78.0** | - | 74.7 | 49.8 | 78.4 |

Table 1: Accuracy on eleven datasets from arithmetic, commonsense and symbolic reasoning tasks. The previous finetuned state-of-the-art models are: $\alpha$: (Cobbe et al., 2021b); $\beta$: (Amini et al., 2019); $\gamma$: (Hosseini et al., 2014); $\delta$: (Hu et al., 2019); $\epsilon$: (Pi et al., 2022); $\zeta$: (Lan et al., 2022); $\eta$: (Xu et al., 2022). * denotes all three LLMs use Manual-CoT. Iter-CoT(W) and Iter-CoT(S) denote our weak and strong iterative bootstrapping approaches.

### 4.3 Implementation Details

We employ the ChatGPT model from the OpenAI API (Accessed version 11/30/2022), with 175 billion parameters, as our LLM to generate reasoning chains and results. During the iterative bootstrapping process, we utilize a temperature setting of 0.7, whereas during inference phase we fix the temperature to 0. We adopt the number of exemplars for each dataset based on the experimental configuration of prior work. Specifically, for datasets lacking test sets and without comparable datasets for transfer (e.g., date understanding and object tracking), we randomly select a small portion as the training set and reserved the remaining portion for evaluation as the test set. We conduct three trials and averages for each experiment requiring random sampling to obtain final results. The size of each dataset and the partitioning of train and test sets are shown in Table 5.

### 4.4 Main Results

The experimental results are presented in Table 1. Our method achieves outperformance over previous state-of-the-art Auto-CoT (Zhang et al., 2022) and Manual-CoT (Wei et al., 2022). These results highlight the efficacy of presenting moderately challenging exemplars with revised and summarized reasoning chains, which better leverage the reasoning potential of LLMs. For visual comparison between the reasoning chains generated by our method and those generated by conventional methods for the same questions, we summarize specific examples of these three reasoning tasks in Table 3.

**Arithmetic Reasoning** Our approach achieves superior results in all six arithmetic datasets over previous Manual-CoT, Random-CoT and Auto-CoT methods. The most significant improvement is observed in the GSM8K and AQuA datasets, with a respective increase of 5.88%, 8.17% over Manual-CoT and 1.75%, 5.11% over Auto-CoT, respectively, leading to state-of-the-art performance. The difference between weakly bootstrapping and strongly bootstrapping methods is insignificant (1.1%). Our findings suggest that a combination of LLMs' self-correction and context-based bootstrapping by human can effectively improve the generation of reasoning chains for arithmetic reasoning tasks.

**Commonsense and Symbolic Reasoning** Our approach significantly outperforms Manual-CoT, Random-CoT and Auto-CoT in both commonsense and symbolic reasoning. Both two reasoning tasks differ from arithmetic reasoning, requiring a deep understanding of the problem-solving paradigm. Since the weak bootstrapping method can encourage the model to discover task-specific details for solving problems, it is generally more effective than the strong bootstrapping method for such tasks (Specific examples are shown in Table 3). The training set of StrategyQA does not have any hint information, and it is difficult for human to modify the reasoning chain, so we only conduct Iter-CoT(W). Specifically, we evaluate our approach on the Out-of-Distribution (OOD) setting of the last letter concatenation, where the few-shot prompts consist of two letters while the test questions consist of four letters.

### 4.5 Ablation Studies

To assess the impact of different factors of our method on overall performance, we conduct ablation experiments on the GSM8K dataset. More specifically, We investigate the effect of the number of iterations in weak bootstrapping on performance and compare the effectiveness of our method in generating CoT with other approaches.

#### 4.5.1 Effects of Iterations in Weak Bootstrapping

In weak bootstrapping, the frequency of utilizing `Revise-Prompt` during the correction phase is called iterations. To examine the impact of the iterations in weak bootstrapping on performance, we conduct experiments on the entire training set of GSM8K, as demonstrated in Figure 4. We select the iterations that generate the highest quality exemplars in our main experiment for GSM8K and other datasets sharing the same demonstrations (AddSub, SingleEq, SVAMP, and ASDiv). For other datasets, we employ demonstrations generated by 1-iteration weak bootstrapping.

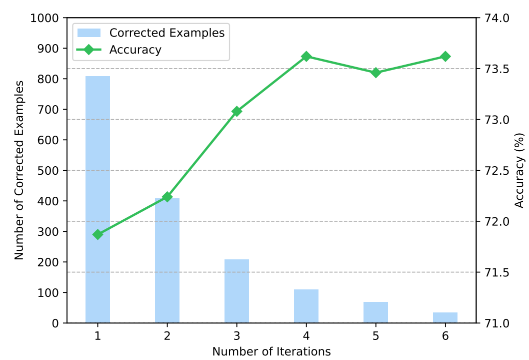



Figure 4: The overall performance (green line) and the number of corrected exemplars (blue bar) during iterations. We conduct the iterations on the whole training set of GSM8K and randomly sample exemplars from each iteration as the demonstrations for inference on test set.

Figure 4 depicts a gradual decrease in the number of corrected exemplars through weak bootstrapping with increasing iterations, leading to a shortage of suitable exemplars after the sixth iteration. Despite this trend, the overall performance of the LLMs improves when using exemplars corrected after the current iterations as demonstrations. The green line in the figure shows that the LLMs' performance steadily increased until the fourth iteration, reaching their peak performance of 73.6%. Subsequently, the performance fluctuates and no longer exhibits a significant improvement.

| Seed Examples | Method | Accuracy |
| --- | --- | --- |
| Sampling Best-Iteration | CoT | 65.5 |
| | Manual-Correct-CoT | 69.9 |
| | Correct-CoT | 71.2 |
| | Iter-CoT(S) | 72.7 |
| | Iter-CoT(W) | **73.6** |
| Sampling Randomly | CoT | 66.7 |
| | Iter-CoT(S) | 68.5 |

Table 2: The performance of various generated CoT among different seed examples. Sampling Best-Iteration denotes we use the exemplars in Section 4.5.1, i.e., the fourth iteration in weak bootstrapping.

The observed phenomenon can be attributed to the increasing difficulty level of corrected exemplars as the iteration progresses. Specifically, questions requiring multiple corrections exhibit a higher difficulty level than those resolvable in the initial iteration. Our ablation experiments support the hypothesis that guiding LLMs with more challenging exemplars can improve overall performance. Since the exemplars selected can be solved without external reference answers prompt, the difficulty of them is always within the model's ability, avoiding the negative impact of extremely difficult exemplars on simple problems. After multiple iterations, the difficulty level of exemplars tends to stabilize, leading to fluctuations in model performance.

| Iter-CoT | CoT |
|---|---|
| Q: Four classmates were comparing their ages based on their birth month. They found out that Jolyn is 2 months older than Therese while Therese is 5 months older than Aivo. Then, Leon is 2 months older than Aivo. How much older in months is Jolyn than Leon? (GSM8K (Cobbe et al., 2021a)) | |
| A: Reasoning process: We are given that Jolyn is 2 months older than Therese and Therese is 5 months older than Aivo. Let the ages of Therese and Aivo be T and A, respectively. We can write the ages of Jolyn, Therese, and Aivo as J = T+2, T = A+5, and A = A. Leon is 2 months older than Aivo, so L = A+2. We want to find the difference between Jolyn's age and Leon's age, so J - L = (T+2) - (A+2). Substituting for T and A using the equations above gives us (A+7) - (A+2) = 5. Therefore, Jolyn is 5 months older than Leon. Final answer: 5. ✔ | A: Let's think step by step. First let's say the ages of Therese and Aivo is T and A. T=A+5 J=T+2 L=A+2 J-L=T+2-(A+2) J-L=2 The correct answer is: 2. ✗ |
| Q: Take the last letters of the words in "Steve Sweeney" and concatenate them. (Letter(4) (Wei et al., 2022)) | |
| A: Reasoning process: First, we need to split the sentence "Steve Sweeney" into individual words by looking for the spaces. This gives two words: "Steve" and "Sweeney". Next, we need to take last letter of each word, which are "e" and "y". Finally, we need to concatenate these letters together to get the final answer. Final answer: ey. ✔ | A: Let's think step by step. First we take the last letters of the words in "Steve Sweeney", which are "ete eeney". Since we need to concatenate them, we will end up with: "ete eeney" Therefore, the correct answer is: "ete eeney". ✗ |
| Q: May 6, 1992 is like yesterday to Jane, but that is actually ten years ago. What is the date one week from today in MM/DD/YYYY? Choices: A.05/13/2044 B.06/10/2002 C.05/13/2002 D.03/13/2003 E.11/13/2001 (Date Understanding (Wei et al., 2022)) | |
| A: Reasoning process: 1. May 6, 1992 is ten years ago from today, which means today is May 6, 2002. 2. The date one week from today is the date one week from May 6, 2002. 3. The date one week from May 6, 2002 is May 13, 2002. 4. The date one week from today in MM/DD/YYYY is 05/13/2002. Final answer: C.05/13/2002. ✔ | A: Let's think step by step. The date one week from today is the date one week from Monday May 13, 2002. Thus the correct answer is: B.06/10/2002. ✗ |

Table 3: This is an example of Iter-CoT compared to CoT on three different reasoning datasets.

### 4.5.2 Effects of Generated CoT

Our approach incorporates richer contextual information throughout the complete initialization, correction, and summarization process, resulting in more accurate, comprehensive, and high-quality reasoning chains. To evaluate the effectiveness of our approach, we compare Iter-CoT(W) against `CoT` (reasoning chains generated with "let's think step by step." prompt), `Manual-Correct-CoT` (modify wrong reasoning chains manually), and `Correct-CoT` (reasoning chains generated after the final error correction in Iter-CoT, shown in Figure 8 and Figure 9) on the same examples sampled from the fourth iteration in weak bootstrapping approach (exemplars are shown in Appendix 7). Firstly, the CoT results indicate a notably low performance, which is consistent with the previous conclusion in Section 2. This highlights that performance tends to decrease as incorrect demonstrations increase. Secondly, the result shows using `Correct-CoT` as reasoning chains have demonstrated lower efficacy than Iter-CoT (drop about 2.4%). Therefore, permitting the LLM to summarize its corrections during the correction phase is essential.

Some works (Paranjape et al., 2023; Liu et al., 2023) have attempted to manually correct the reasoning chain in order to refine the results. While manual correction by humans can enhance the quality of reasoning chains, the extent of this improvement remains limited. We primarily reference the reasoning steps manually annotated in GSM8K, but we only modify some of the errors in the reasoning chain (e.g., add the correct intermediate steps or remove incorrect intermediate steps). The results from the table indicate that manually correcting errors in the reasoning chain is significantly more effective than CoT (approximately 4.4%). However, the Iter-CoT(S) yields even greater improvement (7.2%). This reveals that Iter-CoT(S) is not simply a form of manual annotation, but rather a new type of manual annotation, where humans only need to annotate the incorrect reasoning chains, rather than requiring design the demonstrations from scratch.

We conduct an additional experiment to generate CoT and Iter-CoT by randomly selecting examples from the whole train set from GSM8K. For Iter-CoT, we employ the strong bootstrapping approach for incorrect questions where weak bootstrapping is insufficient due to overly difficult questions. Specifically, we directly summarize the reasoning chain for initially correct questions. Our results indicate that Iter-CoT can effectively improve randomly sampled demonstrations performance; however, the extent of improvement is limited (approximately 1.8%). The results also confirm that Iter-CoT has a primary enhancement effect on selecting examples with appropriate difficulty levels. Therefore, Iter-CoT can effectively enhance performance by ensuring the selection of the most appropriate exemplars.

As Section 3 mentions, Iter-CoT can generate more precise and comprehensive reasoning chains than CoT. We conduct inference on three distinct reasoning datasets (GSM8K, Letter(4) and Date Understanding) utilizing both Iter-CoT(W) and CoT with the same questions, which are shown in Table 3. We use the same LLMs and temperature to generate reasoning chains and answers. We observe that Iter-CoT naturally generates more precise and comprehensive reasoning chains compared to CoT, resulting in higher overall performance.

# 5 Related Work

## 5.1 Chain-of-thought Prompting

Chain-of-thought (CoT) refers to a series of intermediate reasoning steps used to generate the answer. This concept was originally proposed by Wei et al. (2022) and significantly improve the performance of LLMs on complex reasoning tasks. Existing studies on CoT prompting can be broadly classified into two categories: manually constructed CoT prompting and automatically generated CoT prompting.

### 5.1.1 Manually Constructed CoT Prompts

Wei et al. (2022) proposed Manual-CoT, an approach that employs manually-crafted demonstrations comprising the question, the reasoning process, and the final answer as prompts. In subsequent work, Wang et al. (2022) introduced a novel decoding strategy "Self-Consistency", which generates multiple answers from LLMs and aggregates them through a majority voting mechanism to produce the final answer. Li et al. (2022) increased the randomness of the prompts to enhance the diversity of generated reasoning paths. Diao et al. (2023) annotated the reasoning chain manually for the most uncertain questions and employed them as few-shot exemplars. Furthermore, another research trend is to decompose complex problems into a series of sub-problems and solves them sequentially, as demonstrated by Zhou et al. (2022); Press et al. (2022). Although these techniques have shown remarkable performance in enhancing the model's reasoning capability, they are expensive and challenging since they all require manually constructed demonstrations. Moreover, these prompts are suboptimal and highly sensitive.

### 5.1.2 Automatically Generated CoT Prompts

Kojima et al. (2022) proposed a "Let's think step by step" prompt that guides LLMs to generate reasoning steps without hand-craft design demonstrations. Following this work, several studies utilized self-generated rationales for demonstrations. Zhang et al. (2022) employed zero-shot-cotKojima et al. (2022) to generate the reasoning process in its demonstration. In contrast, Shao et al. (2023b) employed seed demonstrations to automatically synthesize more examples by repeatedly forward and backward processes.

Unlike previous works, we propose a novel approach to generate reasoning chains by allowing the LLM to retrace its reasoning process after inferring the answer. This approach enables the model to capture contextual information in previous iterations, resulting in more precise and comprehensive reasoning chains.

### 5.2 In-Context Learning

In-context Learning (ICL) is a technique that allows LLMs to complete target tasks during inference by using a few tasks-specific examples as demonstrations, without modifying the model parameters (Shao et al., 2023a; Brown et al., 2020). Zhao et al. (2021) underscored that the accuracy of LLMs in ICL depends heavily on the selection and permutation of examples. Additionally, the employment of demonstrations established a highly intimate connection between CoT prompting and ICL. Therefore, significant efforts have been invested in developing approaches to identify appropriate few-shot demonstrations.

Zhang et al. (2022) adopted a clustering-based method to select demonstrations. Fu et al. (2022) selected the demonstrations with the most reasoning steps or the longest questions and similarly, Diao et al. (2023) chose the demonstrations with most uncertain questions. Nonetheless, Shum et al. (2023) posited that excessively intricate exemplars do not aid simple problems. Instead, they added the demonstrations with the correct answer to the samples pool and optimized the demonstrations selection model via reinforcement learning. These studies either disregarded the hazards of utilizing demonstrations containing flawed reasoning chains or refrained from employing demonstrations containing reasoning chains that would lead to erroneous answers.

Through the self-correction, our approach selects challenging yet answerable questions accompanied by reasoning chains as exemplars, with a moderate level of difficulty, which enhances the LLMs' generalizability across varying levels of difficulty.

## 6 Conclusion

This paper proposes Iter-CoT, an iterative bootstrapping in chain-of-thoughts prompting for large language model reasoning. Unlike previous work, our method prompts LLMs to self-correct their errors in reasoning chains by leveraging iterative bootstrapping and obtaining more accurate and detailed reasoning chains. We introduce two methodologies: weak bootstrapping and strong bootstrapping. Weak bootstrapping refers to providing only label information to the LLMs, whereas strong bootstrapping involves human annotation to correct errors in the reasoning chains. Experimental results on eleven reasoning datasets among three different reasoning tasks demonstrate that our approach competes significantly over the previous state-of-the-art method.

## References


Aida Amini, Saadia Gabriel, Shanchuan Lin, Rik Koncel-Kedziorski, Yejin Choi, and Hannaneh Hajishirzi. Mathqa: Towards interpretable math word problem solving with operation-based formalisms. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 2357–2367, 2019. doi: 10.18653/v1/n19-1245. URL `https://doi.org/10.18653/v1/n19-1245`.

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. Language models are few-shot learners. In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin, editors, *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020. URL `https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html`.

Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam Shazeer, Vinodkumar Prabhakaran, Emily Reif, Nan Du, Ben Hutchinson, Reiner Pope, James

Bradbury, Jacob Austin, Michael Isard, Guy Gur-Ari, Pengcheng Yin, Toju Duke, Anselm Levskaya, Sanjay Ghemawat, Sunipa Dev, Henryk Michalewski, Xavier Garcia, Vedant Misra, Kevin Robinson, Liam Fedus, Denny Zhou, Daphne Ippolito, David Luan, Hyeontaek Lim, Barret Zoph, Alexander Spiridonov, Ryan Sepassi, David Dohan, Shivani Agrawal, Mark Omernick, Andrew M. Dai, Thanumalayan Sankaranarayana Pillai, Marie Pellat, Aitor Lewkowycz, Erica Moreira, Rewon Child, Oleksandr Polozov, Katherine Lee, Zongwei Zhou, Xuezhi Wang, Brennan Saeta, Mark Diaz, Orhan Firat, Michele Catasta, Jason Wei, Kathy Meier-Hellstern, Douglas Eck, Jeff Dean, Slav Petrov, and Noah Fiedel. Palm: Scaling language modeling with pathways. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2204.02311. URL `https://doi.org/10.48550/arXiv.2204.02311`.

Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Jacob Hilton, Reiichiro Nakano, Christopher Hesse, and John Schulman. Training verifiers to solve math word problems. *arXiv Preprint*, 2021a. URL `https://arxiv.org/abs/2110.14168`.

Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Jacob Hilton, Reiichiro Nakano, Christopher Hesse, and John Schulman. Training verifiers to solve math word problems. *arXiv Preprint*, 2021b. URL `https://arxiv.org/abs/2110.14168`.

Shizhe Diao, Pengcheng Wang, Yong Lin, and Tong Zhang. Active prompting with chain-of-thought for large language models. *arXiv Preprint*, 2023. doi: 10.48550/arXiv.2302.12246. URL `https://doi.org/10.48550/arXiv.2302.12246`.

Yao Fu, Hao Peng, Ashish Sabharwal, Peter Clark, and Tushar Khot. Complexity-based prompting for multi-step reasoning. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2210.00720. URL `https://doi.org/10.48550/arXiv.2210.00720`.

Mor Geva, Daniel Khashabi, Elad Segal, Tushar Khot, Dan Roth, and Jonathan Berant. Did aristotle use a laptop? A question answering benchmark with implicit reasoning strategies. *Trans. Assoc. Comput. Linguistics*, 9:346–361, 2021. doi: 10.1162/tacl_a_00370. URL `https://doi.org/10.1162/tacl_a_00370`.

Mohammad Javad Hosseini, Hannaneh Hajishirzi, Oren Etzioni, and Nate Kushman. Learning to solve arithmetic word problems with verb categorization. In Alessandro Moschitti, Bo Pang, and Walter Daelemans, editors, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL*, pages 523–533. ACL, 2014. doi: 10.3115/v1/d14-1058. URL `https://doi.org/10.3115/v1/d14-1058`.

Minghao Hu, Yuxing Peng, Zhen Huang, and Dongsheng Li. A multi-type multi-span network for reading comprehension that requires discrete reasoning. In Kentaro Inui, Jing Jiang, Vincent Ng, and Xiaojun Wan, editors, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019*, pages 1596–1606, 2019. doi: 10.18653/v1/D19-1170. URL `https://doi.org/10.18653/v1/D19-1170`.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. Large language models are zero-shot reasoners. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2205.11916. URL `https://doi.org/10.48550/arXiv.2205.11916`.

Rik Koncel-Kedziorski, Hannaneh Hajishirzi, Ashish Sabharwal, Oren Etzioni, and Siena Dumas Ang. Parsing algebraic word problems into equations. *Trans. Assoc. Comput. Linguistics*, 3:585–597, 2015. doi: 10.1162/tacl_a_00160. URL `https://doi.org/10.1162/tacl_a_00160`.

Yihuai Lan, Lei Wang, Qiyuan Zhang, Yunshi Lan, Bing Tian Dai, Yan Wang, Dongxiang Zhang, and Ee-Peng Lim. Mwptoolkit: An open-source framework for deep learning-based math word problem solvers. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pages 13188–13190, 2022. URL `https://ojs.aaai.org/index.php/AAAI/article/view/21723`.

Yifei Li, Zeqi Lin, Shizhuo Zhang, Qiang Fu, Bei Chen, Jian-Guang Lou, and Weizhu Chen. On the advance of making language models better reasoners. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2206.02336. URL `https://doi.org/10.48550/arXiv.2206.02336`.

Wang Ling, Dani Yogatama, Chris Dyer, and Phil Blunsom. Program induction by rationale generation: Learning to solve and explain algebraic word problems. In Regina Barzilay and Min-Yen Kan, editors, *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers*, pages 158–167, 2017. doi: 10.18653/v1/P17-1015. URL `https://doi.org/10.18653/v1/P17-1015`.

Ruibo Liu, Chenyan Jia, Ge Zhang, Ziyu Zhuang, Tony X. Liu, and Soroush Vosoughi. Second thoughts are best: Learning to re-align with human values from text edits. *arXiv Preprint*, 2023. doi: 10.48550/arXiv.2301.00355. URL `https://doi.org/10.48550/arXiv.2301.00355`.

Shen-Yun Miao, Chao-Chun Liang, and Keh-Yih Su. A diverse corpus for evaluating and developing english math word problem solvers. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel R. Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 975–984, 2020. doi: 10.18653/v1/2020.acl-main.92. URL `https://doi.org/10.18653/v1/2020.acl-main.92`.

Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F. Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2203.02155. URL `https://doi.org/10.48550/arXiv.2203.02155`.

Bhargavi Paranjape, Scott M. Lundberg, Sameer Singh, Hannaneh Hajishirzi, Luke Zettlemoyer, and Marco Túlio Ribeiro. ART: automatic multi-step reasoning and tool-use for large language models. *arXiv Preprint*, 2023. doi: 10.48550/arXiv.2303.09014. URL `https://doi.org/10.48550/arXiv.2303.09014`.

Arkil Patel, Satwik Bhattamishra, and Navin Goyal. Are NLP models really able to solve simple math word problems? In Kristina Toutanova, Anna Rumshisky, Luke Zettlemoyer, Dilek Hakkani-Tür, Iz Beltagy, Steven Bethard, Ryan Cotterell, Tanmoy Chakraborty, and Yichao Zhou, editors, *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021*, pages 2080–2094, 2021. doi: 10.18653/v1/2021.naacl-main.168. URL `https://doi.org/10.18653/v1/2021.naacl-main.168`.

Xinyu Pi, Qian Liu, Bei Chen, Morteza Ziyadi, Zeqi Lin, Qiang Fu, Yan Gao, Jian-Guang Lou, and Weizhu Chen. Reasoning like program executors. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, editors, *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 761–779, 2022. URL `https://aclanthology.org/2022.emnlp-main.48`.

Ofir Press, Muru Zhang, Sewon Min, Ludwig Schmidt, Noah A. Smith, and Mike Lewis. Measuring and narrowing the compositionality gap in language models. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2210.03350. URL `https://doi.org/10.48550/arXiv.2210.03350`.

Jack W. Rae, Sebastian Borgeaud, Trevor Cai, Katie Millican, Jordan Hoffmann, H. Francis Song, John Aslanides, Sarah Henderson, Roman Ring, Susannah Young, Eliza Rutherford, Tom Hennigan, Jacob Menick, Albin Cassirer, Richard Powell, George van den Driessche, Lisa Anne Hendricks, Maribeth Rauh, Po-Sen Huang, Amelia Glaese, Johannes Welbl, Sumanth Dathathri, Saffron Huang, Jonathan Uesato, John Mellor, Irina Higgins, Antonia Creswell, Nat McAleese, Amy Wu, Erich Elsen, Siddhant M. Jayakumar, Elena Buchatskaya, David Budden, Esme Sutherland, Karen Simonyan, Michela Paganini, Laurent Sifre, Lena Martens, Xiang Lorraine Li, Adhiguna Kuncoro, Aida Nematzadeh, Elena Gribovskaya, Domenic Donato, Angeliki Lazaridou, Arthur Mensch, Jean-Baptiste Lespiau, Maria Tsimpoukelli, Nikolai Grigorev, Doug Fritz, Thibault Sottiaux, Mantas Pajarskas, Toby Pohlen, Zhitao Gong, Daniel Toyama, Cyprien de Masson d'Autume, Yujia Li, Tayfun Terzi, Vladimir Mikulik, Igor Babuschkin, Aidan Clark, Diego de Las Casas, Aurelia Guy, Chris Jones, James Bradbury, Matthew J. Johnson, Blake A. Hechtman, Laura

Weidinger, Iason Gabriel, William S. Isaac, Edward Lockhart, Simon Osindero, Laura Rimell, Chris Dyer, Oriol Vinyals, Kareem Ayoub, Jeff Stanway, Lorrayne Bennett, Demis Hassabis, Koray Kavukcuoglu, and Geoffrey Irving. Scaling language models: Methods, analysis & insights from training gopher. *arXiv Preprint*, 2021. URL `https://arxiv.org/abs/2112.11446`.

Teven Le Scao, Angela Fan, Christopher Akiki, Ellie Pavlick, Suzana Ilic, Daniel Hesslow, Roman Castagné, Alexandra Sasha Luccioni, François Yvon, Matthias Gallé, Jonathan Tow, Alexander M. Rush, Stella Biderman, Albert Webson, Pawan Sasanka Ammanamanchi, Thomas Wang, Benoît Sagot, Niklas Muennighoff, Albert Villanova del Moral, Olatunji Ruwase, Rachel Bawden, Stas Bekman, Angelina McMillan-Major, Iz Beltagy, Huu Nguyen, Lucile Saulnier, Samson Tan, Pedro Ortiz Suarez, Victor Sanh, Hugo Laurençon, Yacine Jernite, Julien Launay, Margaret Mitchell, Colin Raffel, Aaron Gokaslan, Adi Simhi, Aitor Soroa, Alham Fikri Aji, Amit Alfassy, Anna Rogers, Ariel Kreisberg Nitzav, Canwen Xu, Chenghao Mou, Chris Emezue, Christopher Klamm, Colin Leong, Daniel van Strien, David Ifeoluwa Adelani, and et al. BLOOM: A 176b-parameter open-access multilingual language model. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2211.05100. URL `https://doi.org/10.48550/arXiv.2211.05100`.

Zhihong Shao, Yeyun Gong, Yelong Shen, Minlie Huang, Nan Duan, and Weizhu Chen. Synthetic prompting: Generating chain-of-thought demonstrations for large language models. *arXiv Preprint*, 2023a. doi: 10.48550/arXiv.2302.00618. URL `https://doi.org/10.48550/arXiv.2302.00618`.

Zhihong Shao, Yeyun Gong, Yelong Shen, Minlie Huang, Nan Duan, and Weizhu Chen. Synthetic prompting: Generating chain-of-thought demonstrations for large language models. *arXiv Preprint*, abs/2302.00618, 2023b. doi: 10.48550/arXiv.2302.00618. URL `https://doi.org/10.48550/arXiv.2302.00618`.

Kashun Shum, Shizhe Diao, and Tong Zhang. Automatic prompt augmentation and selection with chain-of-thought from labeled data. *arXiv Preprint*, abs/2302.12822, 2023. doi: 10.48550/arXiv.2302.12822. URL `https://doi.org/10.48550/arXiv.2302.12822`.

Shaden Smith, Mostofa Patwary, Brandon Norick, Patrick LeGresley, Samyam Rajbhandari, Jared Casper, Zhun Liu, Shrimai Prabhumoye, George Zerveas, Vijay Korthikanti, Elton Zheng, Rewon Child, Reza Yazdani Aminabadi, Julie Bernauer, Xia Song, Mohammad Shoeybi, Yuxiong He, Michael Houston, Saurabh Tiwary, and Bryan Catanzaro. Using deepspeed and megatron to train megatron-turing NLG 530b, A large-scale generative language model. *arXiv Preprint*, 2022. URL `https://arxiv.org/abs/2201.11990`.

Alon Talmor, Jonathan Herzig, Nicholas Lourie, and Jonathan Berant. Commonsenseqa: A question answering challenge targeting commonsense knowledge. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4149–4158, 2019. doi: 10.18653/v1/n19-1421. URL `https://doi.org/10.18653/v1/n19-1421`.

Romal Thoppilan, Daniel De Freitas, Jamie Hall, Noam Shazeer, Apoorv Kulshreshtha, Heng-Tze Cheng, Alicia Jin, Taylor Bos, Leslie Baker, Yu Du, YaGuang Li, Hongrae Lee, Huaixiu Steven Zheng, Amin Ghafouri, Marcelo Menegali, Yanping Huang, Maxim Krikun, Dmitry Lepikhin, James Qin, Dehao Chen, Yuanzhong Xu, Zhifeng Chen, Adam Roberts, Maarten Bosma, Yanqi Zhou, Chung-Ching Chang, Igor Krivokon, Will Rusch, Marc Pickett, Kathleen S. Meier-Hellstern, Meredith Ringel Morris, Tulsee Doshi, Renelito Delos Santos, Toju Duke, Johnny Soraker, Ben Zevenbergen, Vinodkumar Prabhakaran, Mark Diaz, Ben Hutchinson, Kristen Olson, Alejandra Molina, Erin Hoffman-John, Josh Lee, Lora Aroyo, Ravi Rajakumar, Alena Butryna, Matthew Lamm, Viktoriya Kuzmina, Joe Fenton, Aaron Cohen, Rachel Bernstein, Ray Kurzweil, Blaise Aguera-Arcas, Claire Cui, Marian Croak, Ed H. Chi, and Quoc Le. Lamda: Language models for dialog applications. *arXiv Preprint*, 2022. URL `https://arxiv.org/abs/2201.08239`.

Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc V. Le, Ed H. Chi, and Denny Zhou. Self-consistency improves chain of thought reasoning in language models. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2203.11171. URL `https://doi.org/10.48550/arXiv.2203.11171`.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed H. Chi, Quoc Le, and Denny Zhou. Chain of thought prompting elicits reasoning in large language models. *arXiv Preprint*, 2022. URL https://arxiv.org/abs/2201.11903.

Yichong Xu, Chenguang Zhu, Shuohang Wang, Siqi Sun, Hao Cheng, Xiaodong Liu, Jianfeng Gao, Pengcheng He, Michael Zeng, and Xuedong Huang. Human parity on commonsenseqa: Augmenting self-attention with external attention. In Luc De Raedt, editor, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pages 2762–2768, 2022. doi: 10.24963/ijcai.2022/383. URL https://doi.org/10.24963/ijcai.2022/383.

Eric Zelikman, Yuhuai Wu, and Noah D. Goodman. Star: Bootstrapping reasoning with reasoning. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2203.14465. URL https://doi.org/10.48550/arXiv.2203.14465.

Zhuosheng Zhang, Aston Zhang, Mu Li, and Alex Smola. Automatic chain of thought prompting in large language models. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2210.03493. URL https://doi.org/10.48550/arXiv.2210.03493.

Zihao Zhao, Eric Wallace, Shi Feng, Dan Klein, and Sameer Singh. Calibrate before use: Improving few-shot performance of language models. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pages 12697–12706. PMLR, 2021. URL http://proceedings.mlr.press/v139/zhao21c.html.

Denny Zhou, Nathanael Schärli, Le Hou, Jason Wei, Nathan Scales, Xuezhi Wang, Dale Schuurmans, Olivier Bousquet, Quoc Le, and Ed H. Chi. Least-to-most prompting enables complex reasoning in large language models. *arXiv Preprint*, 2022. doi: 10.48550/arXiv.2205.10625. URL https://doi.org/10.48550/arXiv.2205.10625.

# A Analysis for Iter-CoT

## A.1 Effective of Different Numbers of Seed Examples

In order to investigate the sensitivity of our approaches and conventional CoT to the seed examples, we conducted an experiment on the GSM8K dataset as shown in Figure 5. It demonstrates that both of our approaches outperform CoT, and are more stable as the number of examples increases. Additionally, our experiment also shows that the overall performance is not determined by the quantity of increasing exemplars. For instance, the Iter-CoT(W) peak occurs at five exemplars, while the Iter-CoT(S) and Random-CoT peaks at four exemplars.

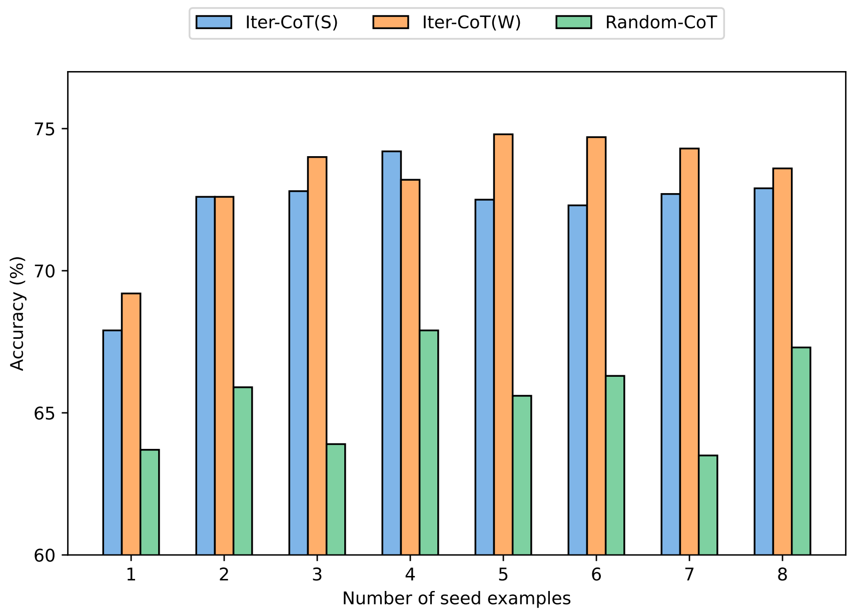



Figure 5: Effictive of Different Numbers of Seed Examples on GSM8K.

## A.2 Comparison with STaR-CoT

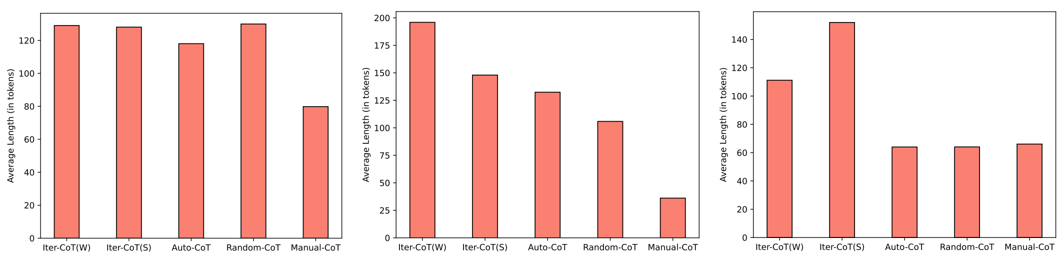
In concurrent work closely related to ours, Zelikman et al. (2022) applies a similar methodology to ours in regenerating the reasoning chains for incorrect examples (we denote STaR-CoT in this paper). Although our work shares similar ideas to their approach, we differ in several key respects. First, their approach does not involve any contextual information from the previous step. Moreover, their approach provides the correct answer directly to the model, rather than using correctness of answer or error reasoning steps as bootstrapping information. It is more prone to generating erroneous reasoning chains despite arriving at the correct answer. Moreover, they continue fine-tuning the model using the corrected reasoning chains, which is impossible for LLMs with 175B parameters (such as ChatGPT (Ouyang et al., 2022)). We conduct inference on one arithmetic reasoning dataset GSM8K utilizing both Iter-CoT and STaR-CoT with the same questions shown in Table 4. We use the same LLMs and temperature to generate reasoning chains and answers. It is observed that although STaR-CoT generates the correct answer, the rationales are wrong, leading to confusion of the entire reasoning chain and reducing the overall performance (73.6% compared to 60.5%).

## A.3 Average Length of Reasoning Chains

We compare the average length of reasoning chains generated by different methods, which are demonstrated in Figure 6. The average length of the reasoning chains demonstrated by Iter-CoT is significantly higher than other methods on the three datasets (GSM8K, CSQA and Last Letter Concatation), except for Random-CoT on GSM8K, which has a comparable length with Iter-CoT.

| Iter-CoT | STaR-CoT |
|---|---|
| Q: Tina makes $18.00 an hour. If she works more than 8 hours per shift, she is eligible for overtime, which is paid by your hourly wage + 1/2 your hourly wage. If she works 10 hours every day for 5 days, how much money does she make? (GSM8K) | |
| A: Reasoning process: First, let's figure out how many hours Tina works per week. She works 10 hours per day and 5 days per week, so 10*5 = 50 hours. Next, let's figure out how many of those hours are overtime. She is eligible for overtime if she works more than 8 hours per shift and she works 10 hours per shift, so 10-8 = 2 hours of overtime per shift. And since Tina works 5 shifts and 2 hours of overtime per shift, she works a total of 5*2 = 10 hours of overtime. Now we can figure out how much Tina makes in regular pay. She works 50-10 = 40 hours per week at her regular wage of $18 per hour, so she makes 40*18 = $720 in regular pay. Now we can figure out how much Tina makes in overtime pay. She gets time and a half for overtime, so 1.5 * $18 per hour = $27 per hour. And since she works 10 hours of overtime and makes $27 per hour, she makes 10*27 = $270 in overtime pay. So in total, Tina makes 720 + 270 = $990 per week. ✔ | A: The correct answer is: 990. Let's think step by step. Tina works 10 hours per day for 5 days, so she works 10*5 = 50 hours. She is eligible for overtime after 8 hours per shift, so she works 50-8 = «50-8=42»42 hours of overtime. She makes $18.00 per hour, so she makes 18*42 = 756 dollars of overtime pay. Overtime pay is her hourly wage + 1/2 her hourly wage, so she makes 18+(18/2) = $27.00 per hour of overtime. She works 42 hours of overtime, so she makes 27*42 = 1134 dollars of overtime pay. She makes $18.00 per hour for her regular hours, so she makes 18*8 = «18*8=144»144 dollars of regular pay. She makes a total of 144+1134 = 1278 dollars. ✗ Therefore, the correct answer is: 990. ✔ |

Table 4: This is an example of Iter-CoT compared to Star-CoT on GSM8K.




(a) GSM8K (b) CSQA (c) Letter

Figure 6: Average length of reasoning chains generated by different methods on GSM8K, CSQA and Last Letter Concatation.

These results provide solid evidence that the reasoning chains demonstrated by Iter-CoT are more comprehensive than those by other alternative methods.

### A.4 Performance Across Different Levels of Difficulty

We posit that questions requiring multiple iterations of bootstrapping to solve present a greater challenge to models. In Complex CoT (Fu et al., 2022), the difficulty level is gauged through the number of hops in reasoning chains. Consequently, our assessment of difficulty is relative to the model's capacity, which is more straightforward and precise. Despite some differences, the results of the two difficulty measurement methods are generally consistent. Therefore, for the convenience of further validating our approach in selecting examples with appropriate level of difficulty, We sort the test set of GSM8K according to the hops of the annotated reasoning chains, and conduct experiment using Iter-CoT and other baselines, as shown in Figure 7. Our results indicate that Iter-CoT performs comparably to other methods for questions with few hops, whereas its performance is significantly better than other methods for questions with multiple hops like 7-hop (with an average improvement of 11.7%) and 8-hop (with an average improvement of 8.3%). Moreover, the effect of Iter-CoT using exemplars produced after four iterations of bootstrapping is considerably superior to that after a single iteration (with an improvement of 9.1% at 7-hop and 10% at 8-hop). As for the issue of complex exemplars proposed in Shum et al. (2023) that performs poorly on simple questions, our method solely includes guidance information and relies on the model to autonomously solve the problem,

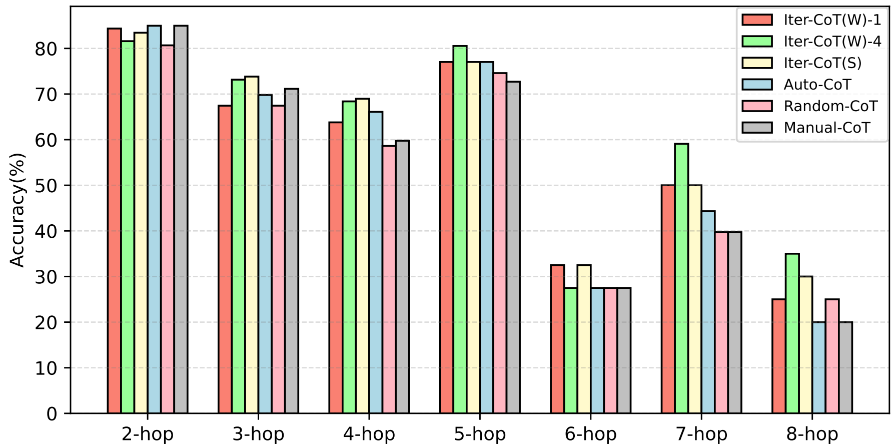



Figure 7: The performance on GSM8K across different numbers of hops (i.e., reasoning steps in the rationale chains).

thus avoiding the problem of selecting excessively difficult exemplars. This demonstrates that with multiple iterations of bootstrapping, our method can generate exemplars with appropriate difficulty levels to solve more challenging problems and boost the LLMs' reasoning ability.

## B Experiment Details

### B.1 Datasets and Tasks

We evaluate Iter-CoT using eleven datasets from three different categories of reasoning tasks. The specific descriptions, divisions, and references of each dataset are shown in Table 5. The examples of each reasoning task are shown in Table 6

### B.2 Weak Bootstrapping and Strong Bootstrapping

The process with the specific example of weak and strong bootstrapping is shown in Figure **??** and Figure 9. The underlined text indicates that LLM finds and corrects its mistakes and generates more precise and comprehensive reasoning chains.

| Dataset | Answer Format | Examples | Train | Test | Licence |
|---|---|---|---|---|---|
| GSM8K | Number | 8 | 7473 | 1319 | MIT License |
| AQuA | Multiple choice | 4 | 97467 | 254 | Apache-2.0 |
| AddSub | Number | 8 | GSM8K* | 395 | - |
| SingleEq | Number | 8 | GSM8K* | 508 | - |
| SVAMP | Number | 8 | GSM8K* | 1000 | MIT License |
| ASDiv | Number | 8 | GSM8K* | 2096 | - |
| CSQA | Multiple choice | 7 | 9741 | 1221 | - |
| StrategyQA | Bineary | 6 | 2821 | 1880 | Apache-2.0 |
| Date | Multiple choice | 8 | 69 | 300 | Apache-2.0 |
| Letter(4) | String | 4 | 500(2*) | 500(4*) | Apache-2.0 |
| Object Tracking | Multiple choice | 8 | 89 | 661 | Apache-2.0 |

Table 5: The statistics of the datasets used in this paper. Examples are the number of examples demonstrations for each dataset. GSM8K* denotes constructed the training set using the GSM8K, cause no available training set for the current dataset. 2* and 4* in the "Letter(4)" row refers to using 2 letters in the training set while using 4 letters in the test set (OOD).

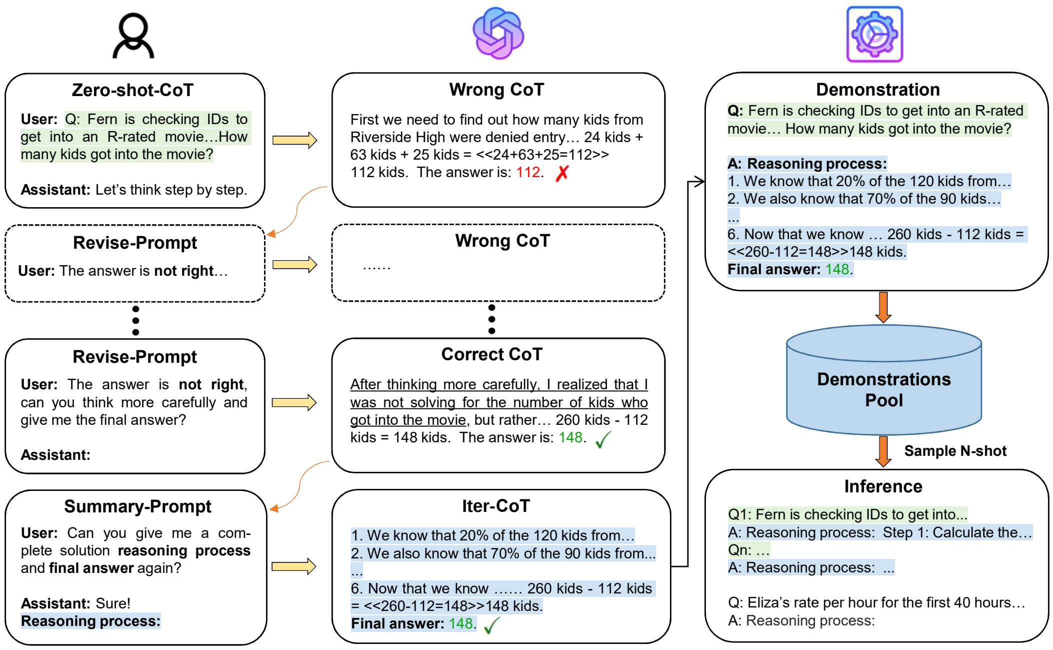



Figure 8: Demonstration of Iter-CoT(W)

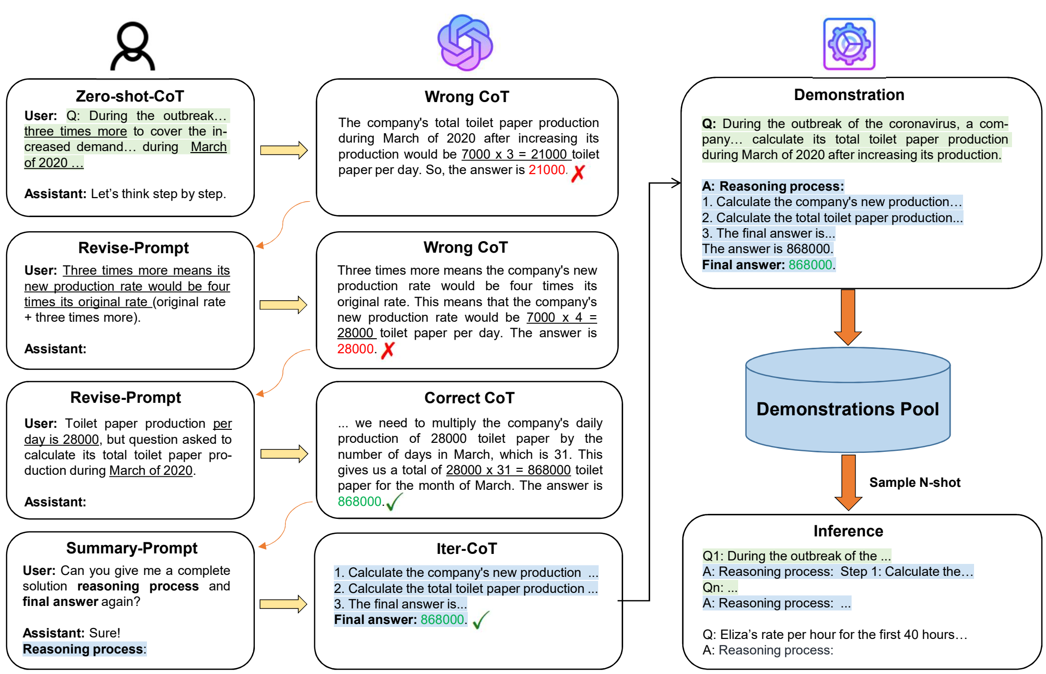



Figure 9: Demonstration of Iter-CoT(S)

| Reasoning Tasks | Example |
| --- | --- |
| Arithmetic | Marissa is hiking a 12-mile trail. She took 1 hour to walk the first 4 miles, then another hour to walk the next two miles. If she wants her average speed to be 4 miles per hour, what speed (in miles per hour) does she need to walk the remaining distance? |
| Commonsense | Jane got her job in 2016. Today is her 3-year work anniversary. She still remember that on Dec 2, her second day at work, she spilled coffee on her laptop. What is the date 24 hours later in MM/DD/YYYY? |
| Symbolic | Take the last letters of each words in "Geo Kody Isaias Giovanni" and concatenate them. |

Table 6: The examples are sampled from the test sets of three typical reasoning datasets, GSM8K (Cobbe et al., 2021a), Date Understanding and Last Letter Concatation (Wei et al., 2022), respectively.

# C Full Exemplars Generated by Iter-CoT

| Dataset | Iter-CoT(W) Exemplars |
|---|---|
| GSM8K | Q: Sheila, Purity, and Rose want to rent a house. Sheila has offered to pay five times Purity's share of the rent. Rose can only afford thrice what Purity pays. If Rose's share is $1,800, what is the total house rent?<br>A: Reasoning Process: We are given that Rose's share of the rent is $1800, and that Rose can only afford thrice what Purity pays. Therefore, Purity's share of the rent is 1800 / 3 = $600. We are also given that Sheila has offered to pay five times Purity's share of the rent. Therefore, Sheila's share is 5 * 600 = $3000. Finally, we have Rose's share of $1800, Purity's share of $600 and Sheila's share of $3000. Therefore, the total rent is 1800 + 600 + 3000 = $5400. Final answer: 5400.<br><br>Q: Padma is trading cards with Robert. Padma started with 75 cards and traded 2 or her valuable ones for 10 of Robert's cards. Robert started with 88 of his own cards and traded another 8 of his cards for 15 of Padma's cards. How many cards were traded between both Padma and Robert?<br>A: Reasoning Process: Padma started with 75 cards. She traded 2 cards for 10 cards, so she now has 75-2+10=83 cards.<br>Robert started with 88 cards. He traded 8 cards for 15 cards, so he now has 88-8+15=95 cards.<br>The total number of cards traded between both Padma and Robert is 2+8+10+15=35 cards.<br>This includes the 2 cards that Padma traded for 10 cards, the 8 cards that Robert traded for 15 cards, and the 10 and 15 cards that were traded between them. Final answer: 35.<br><br>Q: Juanita goes through 1 bottle of sunscreen a month. She likes to stock up for the entire year when a certain website offers 30% off her entire purchase. If each bottle is $30.00 how much will all of the sunscreen cost after the discount?<br>A: Reasoning Process: Juanita needs 1 bottle of sunscreen a month, and there are 12 months in a year, so she needs 12 bottles for the year. The website is offering a 30% discount on her entire purchase. Each bottle of sunscreen costs $30.00. To find the total cost of the sunscreen after the discount, we first need to find the discount amount. This is done by multiplying the price per bottle by the discount rate and the number of bottles: $30.00 * 0.3 * 12 = $108.00. Next, we need to find the price before the discount. This is done by multiplying the price per bottle by the number of bottles: $30.00 * 12 = $360.00. Finally, we subtract the discount amount from the total price before the discount to find the final price after the discount: $360.00 - $108.00 = $252.00. Final answer: 252.<br><br>Q: On a three-day trip, Wendy drove 125 miles on the first day, and 223 miles on the second day. How many miles did she drive on the third day, if the total miles that Wendy drove for the trip is 493 miles?<br>A: Reasoning Process: 1. Wendy drove 125 miles on the first day and 223 miles on the second day, so the total miles she drove on the first two days is 125+223 = 348 miles.<br>2. The total miles Wendy drove on the trip is 493 miles.<br>3. Therefore, the remaining miles she drove on the third day is 493-348 = 145 miles. Final answer: 145.<br><br>Q: Mrs. Lopez and her family are going to the movie theatre. Adult tickets are $11. Children's tickets (ages 3-12) are $8. Senior citizen's tickets (ages 60+) are $9. Mrs. Lopez needs to buy movie tickets for her husband, herself, her parents (ages 72 and 75), and her three children (ages 7, 10, 14). How much money does she need to buy everyone's tickets?<br>A: Reasoning Process: 1. Mrs. Lopez needs to buy tickets for her husband and herself. The cost for these tickets will be $11+$11=$22.<br>2. Mrs. Lopez needs to buy tickets for her parents, who are senior citizens. The cost for these tickets will be $9+$9=$18.<br>3. Mrs. Lopez needs to buy tickets for her three children, who are between the ages of 3-12. The cost for these tickets will be $8+$8+$8=$24.<br>The total cost of all the tickets will be $22+$18+$24=$64. Final answer: 64.<br><br>Q: Mark has 30 candies, Peter has 25 candies, and John has 35 candies. They decided to combine their candies together and share them equally. How many candies will each one of them have?<br>A: Reasoning Process: 1. Mark has 30 candies, Peter has 25 candies, and John has 35 candies.<br>2. We need to add the number of candies that each person has to find the total number of candies they have.<br>3. So, 30 + 25 + 35 = 90 candies.<br>4. Then, we need to divide the total number of candies by the number of people, which is 3. 5. So, 90 / 3 = 30 candies.<br>Final answer: 30. |

Table 7: The exemplars are selected on GSM8K train set. They are transferred to AddSub, SingleEq, SVAMP and ASDiv.

| Dataset | Iter-CoT(W) Exemplars |
|---|---|
| GSM8K | Q: Ben's potato gun can launch a potato 6 football fields. If a football field is 200 yards long and Ben's dog can run 400 feet/minute, how many minutes will it take his dog to fetch a potato he launches?<br>A: Reasoning Process: To find the total distance that the potato can travel, we need to multiply the number of football fields by the number of feet per football field. Since there are 3 feet per yard, and there are 200 yards per football field, we have 200*3 = 600 feet per football field. Therefore, the total distance that the potato can travel is 6*600 = 3600 feet. Since Ben's dog can run at a rate of 400 feet per minute, it will take the dog 3600/400 = 9 minutes to fetch the potato. Final answer: 9.<br><br>Q: Brittany has 3 gerbils: Puffy, Muffy, and Scruffy. Puffy weighs 5 ounces more than Muffy. Muffy weighs 3 ounces less than Scruffy. If Scruffy weighs 12 ounces, how much would the scale indicate, in ounces, if Brittany put Puffy and Muffy on the scale?<br>A: Reasoning Process: We are given that Scruffy weighs 12 ounces and that Muffy weighs 3 ounces less than Scruffy. We are also told that Puffy weighs 5 ounces more than Muffy.<br>Using this information, we can calculate the weight of Muffy by subtracting 3 ounces from Scruffy's weight of 12 ounces, giving us a weight of 12-3 = 9 ounces for Muffy.<br>We can then calculate the weight of Puffy by adding 5 ounces to Muffy's weight of 9 ounces, giving us a weight of 9+5 = 14 ounces for Puffy.<br>If Brittany puts Puffy and Muffy on the scale, the scale would indicate a weight of 14+9 = 23 ounces. Final answer: 23. |

Table 8: (Cont.) The exemplars are selected on GSM8K train set. They are transferred to AddSub, SingleEq, SVAMP and ASDiv.

| Dataset | Iter-CoT(W) Exemplars |
|---|---|
| AQuA | Q: What number has a 5:1 ratio to the number 11? Options: A:22 B:50 C:55 D:52 E:12<br>A: Reasoning Process: In order to find the number that has a 5:1 ratio to 11, we first need to understand what a ratio means. A ratio is a comparison of two numbers, and it is usually written in the form of a:b. In this case, the ratio is 5:1, which means that for every 5 units of the first number, there is 1 unit of the second number. Therefore, if we want to find the number that has a 5:1 ratio to 11, we need to find a number that is 5 times larger than 11. Then, we can check each of the given options to see if any of them are 5 times larger than 11. The only option that fits this description is C:55. Final answer: C.<br><br>Q: The average of seven numbers is 18. The average of first three numbers is 14 and the average of last three numbers is 23. What is the middle number? Options: A:25 B:27 C:15 D:32 E:34<br>A: Reasoning Process: 1. The average of the seven numbers is 18, so the sum of the seven numbers is $18 * 7 = 126$<br>2. The average of the first three numbers is 14, so the sum of the first three numbers is $14 * 3 = 42$<br>3. The average of the last three numbers is 23, so the sum of the last three numbers is $23 * 3 = 69$<br>4. The sum of the first three numbers and the last three numbers is $42 + 69 = 111$<br>5. The sum of the middle number and the sum of the first three numbers plus the last three numbers is equal to the sum of the seven numbers, in other words, $111 + x = 126$.<br>6. The middle number is $x = 15$ Final answer: C.<br><br>Q: Average score for Virat Kohli in a series of 10 matches is 38.9 runs. If the average for first six matches comes out to be 42 what is his average in the last 4 matches of the series? Options: A:34.25 B:34.28 C:24.252 D:64.28 E:34.21<br>A: Reasoning Process: 1) To find the average score for Kohli in the last 4 matches, we need to find the total number of runs he scored in these matches. 2) We know that his average score in the series of 10 matches is 38.9 runs. Therefore, he scored a total of 38.9 * 10 = 389 runs in these matches. 3) We are given that his average score in the first 6 matches is 42 runs. Therefore, he scored a total of 42 * 6 = 252 runs in these matches. 4) The total number of runs he scored in the last 4 matches is 389 - 252 = 137 runs. 5) Therefore, his average score in the last 4 matches is 137 / 4 = 34.25 runs. Final answer: A.<br><br>Q: A train runs at the speed of 72 kmph and crosses a 250 metre long platform in 26 seconds. What is the length of the train ? Options: A:220 metre B:230 metre C:250 metre D:260 metre E:270 metre<br>A: Reasoning Process: We know that the train crosses a 250 metre long platform in 26 seconds at the speed of 72 kmph. To find the length of the train, we need to first convert the speed from kmph to m/s, so we use the formula: $V = V_{kmph} \times \frac{1}{3.6}$. Plugging in the values, we get: $V = 72 \times \frac{1}{3.6} =$ 20m/s. Now that we have the speed in m/s, we can use the formula: $D = Vt$ to find the distance the train covers in 26 seconds. Plugging in the values, we get: $D = 20 \times 26 =$ 520m. Since the length of the platform is 250 metre, and the train covers 520 metre in 26 seconds, the length of the train must be: $520 - 250 =$ 270m. Final answer: E. |

Table 9: The exemplars are selected on AQuA train set.

| Dataset | Iter-CoT(W) Exemplars |
|---|---|
| CSQA | Q: Where could a fungus grow and not be disturbed by sunlight? Choices: A.under rocks B.manhattan C.toenails D.grocery store E.fallen tree<br>A: Reasoning process: 1. Fungi need moisture and shade to grow. 2. Rocks can provide both moisture and shade. 3. Therefore, rocks can be a good place for fungi to grow without being disturbed by sunlight. Final answer: A.<br><br>Q: The man wanted to take a bus, but all the chars were full of people doing what? Choices: A.sitting down B.refusing C.saving money D.sit down E.go outdoors<br>A: Reasoning process: 1. Eliminate answer C ("saving money") because it doesn't have anything to do with taking a bus. 2. Eliminate answer E ("go outdoors") because it doesn't make sense for people to be on a bus if they're already outside. 3. Choose answer A ("sitting down") because it's the only remaining answer that makes sense given the context clues in the question. Final answer: A.<br><br>Q: A beaver can destroy a machines functionality if they build their dam in this? Choices: A.strip club B.pocket C.millpond D.ontario E.lake or river<br>A: Reasoning process: 1. A beaver's dam is built in a body of water. 2. A millpond is a man-made pond used for a watermill. 3. Therefore, if a beaver builds its dam in a millpond, it can potentially destroy the functionality of the watermill. Final answer: C.<br><br>Q: How might a person address someone who is leaving? Choices: A.further cause B.wave goodbye C.attempt suicide D.insulting them E.thank god<br>A: Reasoning process: 1. We know that the person is leaving, so we can eliminate answer choices A and C because they don't make sense in context. 2. We can further eliminate answer choice D because it's rude to insult someone who is leaving. 3. That leaves us with answer choices B and E. 4. Out of these two, the more appropriate option would be B because it's a common way to show polite acknowledgement of someone leaving. Final answer: B.<br><br>Q: What is a characteristic of thin glass? Choices: A.break easily B.shattering C.melt D.bend E.hold water<br>A: Reasoning process: A: Break easily - This fits the characteristic of thin glass, as it is known for its fragility and tendency to break under pressure. B: Shattering - This could be a possible characteristic of thin glass, but it is not mentioned in the given choices, so we cannot select it as the answer. C: Melt - This does not fit the characteristic of thin glass, as it does not need to be heated in order to bend or shape it. D: Bend - This does not fit the characteristic of thin glass, as it is too fragile to be bent or shaped without breaking. E: Hold water - This does not fit the characteristic of thin glass, as it is not a property typically associated with glass. Final answer: A.<br><br>Q: What might be the result of a season of successful skiing? Choices: A.finish line B.broken bones C.broken legs D.chapped lips E.healthy body<br>A: Reasoning process: 1. A season of successful skiing could result in a healthy body if the person is exercising regularly and taking good care of their physical health while skiing. 2. None of the other options, such as a finish line, broken bones, or chapped lips, are related to the potential health benefits of skiing, so they cannot be the correct answer. 3. Therefore, the correct answer is E. healthy body. Final answer: E.<br><br>Q: Glass that hasn't been treated to be extra strong is what? Choices: A.weak B.fragile C.forceless D.regular E.flimsy<br>A: Reasoning process: 1. The question asks about glass that hasn't been treated to be extra strong. This means that the glass has not undergone any special processes or treatments to make it stronger than usual. 2. The answer choices are: A. weak, B. fragile, C. forceless, D. regular, and E. flimsy. 3. We need to find the answer choice that best describes glass that hasn't been treated to be extra strong. 4. "A. weak" means not strong. "B. fragile" means easily broken. "C. forceless" means without strength or power. "D. regular" means normal or usual. "E. flimsy" means weak and easily broken. 5. Based on these definitions, the correct answer is "B. fragile". Glass that hasn't been treated to be extra strong is likely to be fragile, or easily broken. This is because it has not undergone any special processes to make it stronger, so it will be more susceptible to breakage and damage. Other words that may be used to describe this type of glass include "delicate" or "brittle". Final answer: B. |

Table 10: The exemplars are selected on CSQA train set.

| DATASET | Iter-CoT(W) Exemplars |
|---|---|
| StrategyQA | Q: Does a Toyota Prius require an anchor in order to park?<br>A: Reasoning process: A Toyota Prius is a car, and any car needs to be anchored to the ground in order to park, unless it is parked on a slope. Final answer: Yes.<br><br>Q: Did the husband of Elizabeth Hamilton work with a lieutenant?<br>A: Reasoning process: 1. First, we need to know who Elizabeth Hamilton is. Elizabeth Hamilton was the wife of Alexander Hamilton.<br>2. Then, we need to determine if Alexander Hamilton worked with a lieutenant. According to historical records, Alexander Hamilton did work with a lieutenant named John Laurens during the American Revolutionary War.<br>3. Therefore, the answer to the question "Did the husband of Elizabeth Hamilton work with a lieutenant?" is Yes. Final answer: Yes.<br><br>Q: Would Constantinople survive a hypothetical invasion by Charlemagne?<br>A: Reasoning process: 1. Constantinople was the capital of the Byzantine Empire, and Charlemagne was the ruler of the Carolingian Empire, which was in Europe.<br>2. Constantinople was a well-fortified city with strong walls and defenses.<br>3. Charlemagne was known for his military prowess and conquest.<br>4. The Byzantine Empire was a major power in its time, and Constantinople was a key strategic city for the empire.<br>5. However, the Carolingian Empire was also a major power and Charlemagne was a skilled and experienced military leader.<br>6. In a hypothetical situation, Constantinople may put up a strong resistance against Charlemagne's forces, but it would not be able to withstand a sustained attack. Final answer: Yes.<br><br>Q: Would King Henry V hypothetically prevail against Charlemagne at Agincourt?<br>A: Reasoning process: 1. First, we need to understand the question and identify the key information: "Did John Lennon collaborate with the filmmaker behind Walden?"<br>2. Next, we need to research what Walden is and who the filmmaker behind it is: Walden is a documentary film by Alan Watts.<br>3. With this information, we can now search for evidence of John Lennon collaborating with Alan Watts.<br>4. After conducting research, it is clear that John Lennon did collaborate with Alan Watts on a film project called 'The Touch of Zen'. Final answer: Yes.<br><br>Q: Did John Lennon collaborate with the filmmaker behind Walden?<br>A: Reasoning process: Monks are members of religious orders that usually take vows of poverty, chastity, and obedience. These vows are typically associated with a life of spiritual contemplation and non-violent pursuits. However, there are examples of monks engaging in warfare, such as the Knights Templar during the Crusades and Buddhist Shaolin monks who have defended temples during conquests. These examples suggest that it is not always the case that monks are forbidden from engaging in warfare. Final answer: Yes.<br><br>Q: Could Saint Augustine have read Hamlet?<br>A: Reasoning process: Saint Augustine was an early Christian saint who lived about 1,600 years ago.<br>Hamlet is a play by Shakespeare which is believed to have been written in the late 16th century.<br>It is possible that the play was written earlier than that and performed for Saint Augustine. Final answer: Yes. |

Table 11: The exemplars are selected on StrategyQA train set.

| Dataset | Iter-CoT(W) Exemplars |
|---|---|
| Date Understading | Q: Yesterday was 12/31/1929. Today could not be 12/32/1929 because December has only 31 days. What is the date yesterday in MM/DD/YYYY? Choices: A.12/31/1929 B.12/31/1856 C.01/10/1930 D.04/01/1930 E.12/31/1838 F.01/03/1930<br>A: Reasoning process: 1. We know that yesterday was 12/31/1929.<br>2. Today cannot be 12/32/1929 because December only has 31 days.<br>3. Therefore, the date yesterday must be 12/31/1929. Final answer: A.<br><br>Q: The deadline is Jun 1, 2021, which is 2 days away from now. What is the date one week from today in MM/DD/YYYY? Choices: A.12/06/2021 B.05/30/2021 C.03/07/2021 D.06/05/2021 E.06/06/2021<br>A: Reasoning process: 1. The deadline is Jun 1, 2021, which is 2 days away from now. 2. Jun 1, 2021 is the date two days from now. 3. Adding 7 days from this date will give us Jun 6, 2021, which is the date one week from today. 4. Converting this date into MM/DD/YYYY format, we get 06/06/2021. Final answer: E.<br><br>Q: 2015 is coming in 36 hours. What is the date today in MM/DD/YYYY? Choices: A.12/08/2014 B.12/22/2014 C.12/26/2014 D.12/29/1974 E.12/29/2014 F.12/27/2014<br>A: Reasoning process: When we are given that 2015 is coming in 36 hours, we first need to determine what the current date is. We know that 36 hours is equal to 1.5 days, so we need to subtract 1.5 days from the date when 2015 arrives, which is 12/31/2014. This means that the current date is 12/29/2014. Final answer: E.<br><br>Q: May 6, 1992 is like yesterday to Jane, but that is actually ten years ago. What is the date one week from today in MM/DD/YYYY? Choices: A.05/13/2044 B.06/10/2002 C.05/13/2002 D.03/13/2003 E.11/13/2001<br>A: Reasoning process: 1. May 6, 1992 is ten years ago from today, which means today is May 6, 2002. 2. The date one week from today is the date one week from May 6, 2002. 3. The date one week from May 6, 2002 is May 13, 2002. 4. The date one week from today in MM/DD/YYYY is 05/13/2002. Final answer: C.<br><br>Q: The current local time is 3:02 pm of 5/4/2004. What is the date tomorrow in MM/DD/YYYY? Choices: A.05/05/1915 B.05/06/2004 C.01/05/2005 D.05/04/2004 E.02/15/2004 F.05/05/2004<br>A: Reasoning process: The current date is 5/4/2004, and tomorrow is the next day, which is 5/5/2004. This is because the current month is May and the current day is the 4th, so the next day is the 5th. The year remains the same because we are only looking at the date for tomorrow and not a longer time period. Final answer: F.<br><br>Q: Today is Sep 9, 1909. What is the date today in MM/DD/YYYY? Choices: A.09/09/1939 B.12/11/1909 C.09/09/1909 D.09/30/1909 E.11/19/1909 F.09/09/1886<br>A: Reasoning process: * The question is asking for the date today in MM/DD/YYYY format. * We know that today's date is September 9, which is 09. * The next step is to find the correct day. September has 30 days, so the date today is 09/30/1909. * However, this is not correct because the year is 1909, not 1939. * Therefore, the correct date today is 09/09/1909. Final answer: C.<br><br>Q: Today is the palindrome day of 2020, because the MMDDYYYY format of the date is the same backwards as forwards. What is the date 24 hours later in MM/DD/YYYY? Choices: A.02/03/1982 B.02/03/2100 C.02/03/2020 D.02/04/2020 E.02/03/2094 F.01/02/2020<br>A: Reasoning process: The palindrome date is of the form MM/DD/YYYY, and it is the same when read backwards. The date of today is 02/02/2020 and if we read it backwards it is still the same, so it is the palindrome date. The next day is the 02/03/2020 and the palindrome date of the day after tomorrow is the 02/04/2020. Final answer: C.<br><br>Q: The deadline is Jun 1, 2021, which is 2 days away from now. What is the date one week ago from today in MM/DD/YYYY? Choices: A.05/23/2054 B.05/24/2021 C.04/13/2021 D.05/23/1940 E.05/23/2021 F.05/23/2020<br>A: Reasoning process: If the deadline is Jun 1, 2021, which is 2 days away, then today is May 30, 2021. If today is May 30, 2021, then one week ago was May 23, 2021. Final answer: E. |

Table 12: The exemplars are selected on Date Understanding train set.

| DATASET | Iter-CoT(W) Exemplars |
|---|---|
| Object Tracking | Q: Alice, Bob, and Claire are playing a game. At the start of the game, they are each holding a ball: Alice has a green ball, Bob has a purple ball, and Claire has a blue ball. As the game progresses, pairs of players trade balls. First, Bob and Alice swap balls. Then, Claire and Bob swap balls. Finally, Bob and Alice swap balls. At the end of the game, Bob has the? Choices: A.blue ball. B.green ball. C.purple ball.<br>A: Reasoning process: 1. At the start of the game, Bob has a purple ball.<br>2. Bob swaps his ball with Alice, so now he has a green ball.<br>3. Bob swaps his ball with Claire, so now he has a blue ball.<br>4. Bob swaps his ball with Alice again, so now he has a purple ball again. Final answer: C.<br><br>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 left midfielder, Bob is playing right midfielder, and Claire is playing goalkeeper. As the game progresses, pairs of players occasionally swap positions. First, Bob and Alice trade positions. Then, Alice and Claire trade positions. Finally, Bob and Alice trade positions. At the end of the match, Alice is playing? Choices: A.goalkeeper. B.right midfielder. C.left midfielder.<br>A: Reasoning process: 1. Alice's initial position is left midfielder.<br>2. She swaps positions with Bob, so she is now playing right midfielder.<br>3. She swaps positions with Claire, so she is now playing goalkeeper.<br>4. She swaps positions with Bob again, so she is back to playing left midfielder. Final answer: C.<br><br>Q: Alice, Bob, and Claire are playing a game. At the start of the game, they are each holding a ball: Alice has a orange ball, Bob has a green ball, and Claire has a pink ball. As the game progresses, pairs of players trade balls. First, Bob and Claire swap balls. Then, Claire and Alice swap balls. Finally, Bob and Alice swap balls. At the end of the game, Bob has the? Choices: A.green ball. B.orange ball. C.pink ball.<br>A: Reasoning process: At the start of the game, Alice has a orange ball, Bob has a green ball, and Claire has a pink ball. First, Bob and Claire swap balls, so Bob has the pink ball, and Claire has the green ball. Then, Claire and Alice swap balls, so Claire has the orange ball, and Alice has the green ball. Finally, Bob and Alice swap balls, so Bob has the green ball, and Alice has the pink ball. Final answer: A.<br><br>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 center midfielder, Bob is playing benchwarmer, and Claire is playing fullback. As the game progresses, pairs of players occasionally swap positions. First, Alice and Claire trade positions. Then, Alice and Bob trade positions. Finally, Alice and Claire trade positions. At the end of the match, Alice is playing? Choices: A.benchwarmer. B.center midfielder. C.fullback.<br>A: Reasoning process: First, Alice and Claire trade positions, which means Alice goes to fullback and Claire goes to center midfielder. Then, Alice and Bob trade positions, which means Alice goes to benchwarmer and Bob goes to fullback. Finally, Alice and Claire trade positions again, which means Alice goes to center midfielder and Claire goes to benchwarmer. Therefore, at the end of the match, Alice is playing center midfielder. Final answer: B.<br><br>Q: Alice, Bob, and Claire are playing a game. At the start of the game, they are each holding a ball: Alice has a pink ball, Bob has a brown ball, and Claire has a orange ball. As the game progresses, pairs of players trade balls. First, Claire and Alice swap balls. Then, Alice and Bob swap balls. Finally, Bob and Claire swap balls. At the end of the game, Bob has the? Choices: A.pink ball. B.orange ball. C.brown ball.<br>A: Reasoning process: At the start of the game, Alice has a pink ball, Bob has a brown ball, and Claire has a orange ball. In the first swap, Alice swaps her pink ball with Claire's orange ball, so Alice now has an orange ball and Claire has a pink ball. In the second swap, Alice swaps her orange ball with Bob's brown ball, so Alice now has a brown ball and Bob has an orange ball. In the third swap, Bob swaps his orange ball with Claire's pink ball, so Bob now has a pink ball and Claire has an orange ball. Final answer: A. |

Table 13: The exemplars are selected on Object Tracking train set.

| DATASET | Iter-CoT(W) Exemplars |
|---|---|
| Object Tracking | Q: Alice, Bob, and Claire are holding a white elephant gift exchange. At the start of the event, they are each holding a present of a different color: Alice has a blue present, Bob has a black ball, and Claire has a green present. As the event progresses, pairs of people swap gifts. First, Alice and Bob swap their gifts. Then, Alice and Claire swap their gifts. Finally, Claire and Bob swap their gifts. At the end of the event, Alice has the? Choices: A.black ball. B.green present. C.blue present.<br>A: Reasoning process: 1. At the start of the event, Alice has a blue present, Bob has a black ball, and Claire has a green present. 2. Alice and Bob swap their gifts, so Alice has a black ball and Bob has a blue present. 3. Alice and Claire swap their gifts, so Alice has a green present and Claire has a black ball. 4. Claire and Bob swap their gifts, so Claire has a blue present and Bob has a black ball. Therefore, at the end of the event, Alice has the green present. Final answer: B.<br><br>Q: Alice, Bob, and Claire are holding a white elephant gift exchange. At the start of the event, they are each holding a present of a different color: Alice has a pink ball, Bob has a yellow present, and Claire has a black ball. As the event progresses, pairs of people swap gifts. First, Bob and Alice swap their gifts. Then, Claire and Alice swap their gifts. Finally, Claire and Bob swap their gifts. At the end of the event, Claire has the? Choices: A.yellow present. B.black ball. C.pink ball.<br>A: Reasoning process: At the start of the event, Alice has a pink ball, Bob has a yellow present, and Claire has a black ball. Bob and Alice swap their gifts, so Bob gets the pink ball and Alice gets the yellow present. Claire and Alice swap their gifts, so Claire gets the yellow present and Alice gets the black ball. Finally, Claire and Bob swap their gifts, so Claire gets the pink ball and Bob gets the yellow present. Final answer: C.<br><br>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 goalkeeper, Bob is playing left winger, and Claire is playing right midfielder. As the game progresses, pairs of players occasionally swap positions. First, Claire and Alice trade positions. Then, Bob and Claire trade positions. Finally, Alice and Bob trade positions. At the end of the match, Claire is playing? Choices: A.left winger. B.right midfielder. C.goalkeeper.<br>A: Reasoning process: At the start of the match, Alice is playing goalkeeper, Bob is playing left winger, and Claire is playing right midfielder. First, Claire and Alice trade positions, so Claire is playing goalkeeper and Alice is playing right midfielder. Then, Bob and Claire trade positions, so Claire is playing left winger and Bob is playing goalkeeper. Finally, Alice and Bob trade positions, so Claire is playing left winger and Alice is playing goalkeeper. Final answer: A. |

Table 14: (Cont.) The exemplars are selected on Object Tracking train set.

| DATASET | Iter-CoT(W) Exemplars |
|---|---|
| Letter (4) | Q: Take the last letters of the words in "Agustin Lowe" and concatenate them.<br>A: Reasoning process: "Agustin Lowe" consists of two words, "Agustin" and "Lowe", and each of them has 5 and 4 letters, respectively. The last letters of these two words are "n" and "e", respectively. When we concatenate these two letters, we get "ne". Final answer: ne.<br><br>Q: Take the last letters of the words in "Vern Griffin" and concatenate them.<br>A: Reasoning process: 1. First, we need to take the last letters of each word in "Vern Griffin". 2. Then we need to concatenate the last letters to get the final answer. 3. The last letters of "Vern" and "Griffin" are "n" and "n". 4. Concatenating the last letters, we get "nn". Final answer: nn.<br><br>Q: Take the last letters of the words in "Randell Graham" and concatenate them.<br>A: Reasoning process: At the start of the match, Alice is playing goalkeeper, Bob is playing left winger, and Claire is playing right midfielder. 1. Take the last letters of the words in "Randell Graham" which would be "ll" and "m". 2. Concatenate them to form "lm". Final answer: lm.<br><br>Q: Take the last letters of the words in "Randal Holland" and concatenate them.<br>A: Reasoning process: 1. We need to find the last letters of the words in "Randal Holland" 2. "Randal" ends in "l" and "Holland" ends in "d" 3. The last letters of the words in "Randal Holland" are "ld" Final answer: ld. |

Table 15: The exemplars are selected on Last Letter Concatenation train set.

| Dataset | Iter-CoT(S) Exemplars |
|---|---|
| GSM8K | Q: Rich likes to take long walks through town. First he walks 20 feet from his house to the sidewalk. Then he walks 200 feet down the sidewalk to the end of the road. Then he makes a left and walks double his total distance so far until he reaches the next intersection. Then he walks half the total distance up to this point again to the end of his route, before turning around and walking the same path all the way back home. How many feet did Rich walk?<br>A: Reasoning process: 1. Rich walks 20 feet from his house to the sidewalk. 2. He then walks 200 feet down the sidewalk. 3. He makes a left and walks double the distance he had walked so far, or 2 * (20 + 200) = 440 feet. 4. He walks half the distance he had walked up to this point again to the end of his route, or (1/2) * (20 + 200 + 440) = 330 feet. 5. The total distance he walked until the end of his route is 20 + 200 + 440 + 330 = 990 feet. 6. He then walks the same path back home, so the total distance is 2 * 990 = «2*990=1980»19 feet. Final answer: 1980.<br><br>Q: Patricia is growing her long very long to donate it to a charity that makes wigs for cancer survivors. Her hair is 14 inches long. She needs to donate 23 inches to make a wig. She wants her hair to be 12 inches long after the donation. How much longer does she have to grow her hair?<br>A: Reasoning process: 1. Patricia needs to donate 23 inches of hair and she currently has 14 inches. So she needs to grow 23-14=9 more inches. 2. After the donation, Patricia wants her hair to be 12 inches long. She will be donating 23 inches, so she will have 12+23=35 inches of hair after the donation. 3. Patricia currently has 14 inches of hair and wants to have 35 inches after the donation. So she needs to grow 35-14=21 more inches. Thus, Patricia needs to grow her hair 21 more inches to reach her goal. Final answer: 21.<br><br>Q: At the burger hut, you can buy a burger for$5, french fries for $3, and a soft drink for $3. If you order a special burger meal, you get all 3 of these food items for $9.50. A kid's burger is $3, a kid's french fries are $2, and a kid's juice box is $2. They also have a kids meal of all 3 kids' food items for $5. Mr. Parker buys 2 burger meals for his wife and himself. He also buys 2 burger meals and 2 kid's meals for his 4 children. How much money does Mr. Parker save by buying the 6 meals versus buying the individual food items?<br>A: Reasoning process: 1. Calculate the cost of a normal burger meal: burger + fries + soft drink = 5 + 3 + 3 = $11 2. Calculate the cost of a special burger meal: $9.50 3. Calculate the cost of a kid's burger meal: kid's burger + kid's fries + kid's juice box = 3 + 2 + 2 = $7 4. Calculate the cost of a special kid's burger meal: $5 5. Calculate the cost of 4 normal burger meals: 4 * 11 = $44 6. Calculate the cost of 4 special burger meals: 4 * 9.50 = $38 7. Calculate the cost of 2 kid's burger meals: 2 * 7 = $14 8. Calculate the cost of 2 special kid's burger meals: 2 * 5 = $10 9. Calculate the total savings: savings on special burger meals + savings on kid's burger meals = 6 + 4 = $10 Final answer: 10.<br><br>Q: Maurice had only been horseback riding a handful of times in his life. His best friend, Matt, goes horseback riding regularly. When Maurice went to visit Matt for two weeks, he rode another 8 times. Each time, Matt went with him. Meanwhile, besides the times that Matt accompanied Maurice, he rode an additional 16 times. If the total number of times Matt rode during that two weeks is three times the number of times Maurice had ridden before his visit, how many times had Maurice been horseback riding before visiting Matt?<br>A: Reasoning process: First, we know that Maurice rode 8 times while visiting Matt. We also know that Matt rode a total of 3 * 8 = 24 times during the two weeks. We also know that Matt rode an additional 16 times, so his total number of rides during the two weeks was 16 + 24 = 40 times. We know that Maurice rode a total of 8 + 24 = 32 times during the two weeks. If the total number of times Matt rode during that two weeks is three times the number of times Maurice had ridden before his visit, we know that 3 * 24 = 72. Since Maurice rode 8 times during the two weeks, we can say that 8 * 3 = 24 times. Finally, we know that Maurice had ridden 24 / 3 = 8 times before visiting Matt. Final answer: 8.<br><br>Q: Kamil wants to renovate his kitchen at home. For this purpose, he hired two professionals who work for him 6 hours a day for 7 days. What does it cost Kamil to hire these professionals if one of them is paid $15 per hour of work?<br>A: Reasoning process: Kamil hired 2 professionals who work for him for 6 hours a day for 7 days Each professional works for a total of 6 * 7 = 42 hours Both professionals work for a total of 42 * 2 = 84 hours Each professional is paid $15 per hour of work, so the cost for Kamil to hire the professionals is 84 * 15 = $1260 Final answer: 1260.<br><br>Q: Joe played catch with Derek and Tammy. He caught the ball 23 times. Derek made four less than double the catches Joe did. Tammy caught the ball sixteen more than a third of the times Derek did. How many times did Tammy catch the ball?<br>A: Reasoning process: Joe caught the ball 23 times Derek caught four less than double the catches Joe did (23 * 2 = 46, 46 - 4 = 42 catches) Tammy caught 16 more than a third of Derek's catches (42 / 3 = 14, 14 + 16 = 30 catches) Thus, Tammy caught the ball 30 times. Final answer: 30. |

Table 16: The exemplars are selected on GSM8K train set. This set of exemplars is used by GSM8K, ASDiv, SVAMP, and SingleEq.

| DATASET | Iter-CoT(S) Exemplars |
|---|---|
| GSM8K | Q: There are 8 loads of laundry that need to be washed. Each load of laundry takes around 45 minutes to complete in the wash cycle and 1 hour in the dry cycle. How long, in hours, will it take to complete the entire set of laundry?<br>A: Reasoning process: 1. We need to calculate the time it takes to wash the laundry. 2. We know that there are 8 loads of laundry and that each load takes 45 minutes to wash. 3. Thus, the total time it takes to wash the laundry is 8 loads * 45 minutes/load = 360 minutes. 4. Next, we need to find the time it takes to dry the laundry. 5. We know that it takes 1 hour to dry each load of laundry and that there are 8 loads. 6. Thus, the total time it takes to dry the laundry is 8 loads * 1 hour/load = 8 hours. 7. Finally, we need to add the time it takes to wash and dry the laundry to find the total time it takes to complete the entire set of laundry. 8. To find the total time it takes to complete the laundry, we first need to convert the 360 minutes of washing time to hours. 9. 360 minutes / 60 minutes/hour = 6 hours. 10. Thus, the total time it takes to complete the laundry is 6 hours + 8 hours = 14 hours. Final answer: 14.<br><br>Q: Bob grew corn in his garden this year and ended up with 50 bushels. This is way too much for him to eat, so he gave some of it away to his friends. His friend Terry took 8 bushels, while Jerry only took 3. He gave 12 bushels to his friend Linda, who runs a food pantry. His neighbor Stacy doesn't eat much corn, but she still accepted 21 ears of corn from him. If each bushel contained 14 ears of corn, how many ears of corn does Bob have left?<br>A: Reasoning process: First, let's figure out how many bushels of corn Bob gave away to his friends: 8 bushels + 3 bushels + 12 bushels = «8+3+12=23» bushels of corn. Next, let's figure out how many ears of corn are in 23 bushels of corn: 23 bushels x 14 ears/bushel = «23*14=322»3 ears of corn. Finally, let's figure out how many ears of corn Bob has left: Bob has 50 bushels of corn, but he gave away 23 bushels and 21 ears, so he has 50 bushels - 23 bushels = 27 bushels of corn left. Each bushel has 14 ears of corn, so Bob has 27 bushels x 14 ears/bushel = 378 ears of corn left. The final step is to subtract the 21 ears that Bob gave to Stacy: 378 ears - 21 ears = 357 ears of corn left. Final answer: 357. |

Table 17: (Cont.) The exemplars are selected on GSM8K train set. This set of exemplars is used by GSM8K, ASDiv, SVAMP, and SingleEq.

| DATASET | Iter-CoT(S) Exemplars |
|---|---|
| AQuA | Q: A train 360 m long is running at a speed of 45 km/hr. In what time will it pass a bridge 140 m long? Options: A:40 sec B:42 sec C:45 sec D:48 sec E:58 sec<br>A: Reasoning process: First, we need to calculate the speed of the train in meters per second. We know that 1 hour is equal to 3600 seconds, so the speed of the train in m/s is 45 km/hr * 1000 m/km / 3600 s/hr = 12.5 m/s Next, we need to find out how long it takes the train to pass the bridge. We know that the length of the train is 360 m, and the length of the bridge is 140 m. So, the time it takes the train to pass the bridge is (140 m + 360 m) / 12.5 m/s = 40 seconds. Final answer: A.<br><br>Q: In a certain province in Russia there are 15 cities. If a single road segment connects only two cities, how many road segments are required in order to connect the cities so that each city is connected to all other cities with a single road segment? Options: A:30 B:105 C:330 D:14! E:15!<br>A: Reasoning process: First, the first city must be connected to the remaining 14 cities. This will take 14 road segments. Next, the second city must be connected to the remaining 13 cities. This will take 13 road segments. Finally, the third city must be connected to the remaining 12 cities. This will take 12 road segments. This pattern continues until the last city is connected to the remaining 1 city. This will take 1 road segment. Therefore, there are a total of 14 + 13 + 12 + ... + 1 = 14 * 15 / 2 = 105 road segments. Final answer: B.<br><br>Q: The current of a stream runs at the rate of 4 kmph. A boat goes 6 km and back to the starting point in 2 hours, then find the speed of the boat in still water? Options: A:6 B:7 C:8 D:5 E:3<br>A: Reasoning process: First, let's assume that the speed of the boat in still water is x km/hr. Since the current of the stream is 4 km/hr, the boat travels 6 km downcurrent in 6 km / (4 km/hr + x km/hr) hours. Similarly, the boat travels 6 km against the current in 6 km / (x km/hr - 4 km/hr) hours. Adding these two times, we find that the total time the boat spends traveling downcurrent and against the current is 2 hours. Solving for x, we find that the boat speed in still water is 8 km/hr. Final answer: C.<br><br>Q: A car is running at a speed of 96kmph. What distance will it cover in 14sec? Options: A:378m B:350m C:380m D:200m E:250m<br>A: Reasoning process: 1. We are given that the car is running at a speed of 96 km per hour. 2. We want to find the distance it will cover in 14 seconds. 3. We need to convert both km and hours to meters and seconds in order to compare them. 4. 1 hour has 60 minutes, 1 minute has 60 seconds, so 1 hour has 3600 seconds. 5. 1 km has 1000 meters, so 96 km have 96*1000 = 96000 meters. 6. Since we are given the speed of the car in km per hour, we need to divide the total distance in meters by the total time in seconds to find the distance in meters per second. 7. So, the distance the car will cover in 14 seconds is 96000 meters / 3600 seconds/hour * 14 seconds = 378.89 meters. Final answer: The distance the car will cover in 14 seconds is [A]378m. Final answer: A. |

Table 18: The exemplars are selected on AQuA train set.

| DATASET | Iter-CoT(S) Exemplars |
|---|---|
| CSQA | Q: Where can peanut butter be stored? Choices: A.container B.supermarket C.pantry D.sandwich E.jar<br>A: Reasoning process: 1. Peanut butter is a food item. 2. Food items are usually stored in a place where they can stay fresh, like a refrigerator or a pantry. 3. A pantry is a good option for storing food items that don't need to be refrigerated, like peanut butter. Final answer: C.<br><br>Q: He waited for his friend at the squash court, but he was worried his friend thought he meant the at the other end of the public what? Choices: A.country club B.rich person's house C.pool D.park E.fitness center<br>A: Reasoning process: The person is waiting at a squash court, so they are likely at a sports facility or gym. Out of the options given, the only option that seems like it could be at the other end of a public place is a park. Final answer: D.<br><br>Q: What do people want to do when they love a place that they are going on vacation in? Choices: A.pleasant thoughts B.overspending C.moving there D.staying there E.debt<br>A: Reasoning process: 1. People want to do something when they love a place that they are going on vacation in. 2. The options given are A. pleasant thoughts, B. overspending, C. moving there, D. staying there, and E. debt. 3. Out of these options, the only one that makes sense as something people might want to do when they love a place is D. staying there. Final answer: D.<br><br>Q: Where can meat last a long time? Choices: A.backery B.ham sandwich C.fridge D.butcher shop E.freezer<br>A: Reasoning process: Meat is something that can spoil quickly, so we need to find a place that can keep it fresh and prevent it from going bad. A backery is not a good place for keeping meat because it is usually warm and not refrigerated, which is not good for keeping meat fresh. A ham sandwich is not a good place for keeping meat because it is already made and ready to eat, so the meat is already exposed to the air and will go bad quickly. A fridge is a good place for keeping meat because it is cold and refrigerated, which is ideal for keeping meat fresh and preventing it from going bad. A butcher shop is not a good place for keeping meat because it is usually warm and not refrigerated, which is not good for keeping meat fresh. A freezer is also a good place for keeping meat because it is very cold and can keep the meat frozen, which prevents it from going bad and can last a long time. The other choices are not good places for keeping meat because they are not cold enough or not designed for storing meat. As a freezer includes a fridge, it is the best option for keeping meat fresh and preventing it from going bad. Final answer: E.<br><br>Q: What does someone living life fully feel? Choices: A.joy B.valuable C.happiness D.early death E.adventurous<br>A: Reasoning process: 1. To understand what it means to live life fully, we need to consider that it could mean living life to the fullest or experiencing all aspects of life and making the most of every moment. 2. Based on this understanding, we need to think about the emotions and feelings that someone who is living life fully might experience. 3. Joy and happiness are positive emotions that align with the idea of living life fully. 4. Other options like early death, valuable, and adventurous are not emotions or feelings and do not align with the idea of living life fully. 5. Therefore, the correct answer is either [A. joy] or [C. happiness]. Final Answer: [A. joy] is probably the better answer as it is more specific and relates more closely to the idea of living life fully. Final answer: A.<br><br>Q: Sky diving scared him, once they jumped out of the plane he began losing consciousness while what? Choices: A.injury B.death C.playing uno D.coma E.falling down<br>A: Reasoning process: Sky diving scared him: This means he was afraid of sky diving. Once they jumped out of the plane: This means they were in the plane, but then they jumped out. He began losing consciousness: This means he was losing consciousness, or becoming unconscious. While what? This means we need to figure out what was happening while he was losing consciousness. Final answer: From the given choices, the only one that makes sense in this context is option E, falling down. Falling down is the natural result of jumping out of a plane, and it could potentially cause someone to lose consciousness. Final answer: E.<br><br>Q: The wood became fossilized, what did the paleontologists call it when they found ti? Choices: A.sanded B.composted C.warp D.hardened E.petrify<br>A: Reasoning process: First, we need to understand what fossilization is. Fossilization is the process by which an organism is preserved in the fossil record. This can happen through a variety of processes, such as being buried in sediment, being exposed to high pressures and temperatures, or being replaced by minerals. Now let's look at the options given. We know that wood is a type of organism, so if it became fossilized, it would be preserved in the fossil record. We also know that the wood would have undergone some kind of process to become fossilized, such as being buried in sediment or being exposed to high temperatures and pressures. Option A, "sanded," doesn't really make sense. Sanding is a process used to smooth wood or other materials, but it wouldn't have anything to do with fossilization. Option B, "composted," also doesn't really make sense. Composting is the process of breaking down organic matter into a nutrient-rich soil amendment, but it wouldn't have anything to do with fossilization. Option C, "warp," doesn't really make sense either. Warping is a term used to describe wood that has been bent or twisted out of shape, but it wouldn't have anything to do with fossilization. Option D, "hardened," is a possibility. Hardening is a process that can happen to wood or other materials over time, as they become more dense and resistant to damage. This could be a possible outcome of the fossilization process, but it's not the only possibility. Option E, "petrify," is the correct answer. Petrify means to turn into stone or a stony substance, which is exactly what happens to wood during the fossilization process. The wood becomes preserved in the fossil record, and is replaced by minerals over time. This process is called petrification, and the resulting fossil is called a petrified wood fossil. Final answer: E. |

Table 19: The exemplars are selected on CSQA train set.

| DATASET | Iter-CoT(S) Exemplars |
|---|---|
| Date Understading | Q: This is the last day of 1899. What is the date today in MM/DD/YYYY? Choices: A.09/26/1899 B.12/31/1899 C.11/30/1899 D.01/09/1900 E.12/22/1899 F.12/31/1885<br>A: Reasoning process: 1. The year is 1899 because it is the last day of the year. 2. The month is December because it is the last day of the month. 3. The date is 31 because it is the last day of the month. Final answer: B.<br><br>Q: Today is the palindrome day of 2020, because the MMDDYYYY format of the date is the same backwards as forwards. What is the date 24 hours later in MM/DD/YYYY? Choices: A.02/03/1982 B.02/03/2100 C.02/03/2020 D.02/04/2020 E.02/03/2094 F.01/02/2020<br>A: Reasoning process: 1. The date is 02/02/2020, because the MMDDYYYY format of the date is the same backwards as forwards. 2. The date 24 hours later is 02/03/2020, because we add 24 hours to the date and the date changes to the next day. Final answer: C.<br><br>Q: This is the last day of 1899. What is the date today in MM/DD/YYYY? Choices: A.09/26/1899 B.12/31/1899 C.11/30/1899 D.01/09/1900 E.12/22/1899 F.12/31/1885<br>A: Reasoning process: 1. We know that this is the last day of 1899, so the year must be 1899. 2. We need to find the date today in MM/DD/YYYY format. 3. Based on the given information, the date must be December 31, 1899. 4. Therefore, the answer is 12/31/1899. Final answer: B.<br><br>Q: Yesterday was 12/31/1929. Today could not be 12/32/1929 because December has only 31 days. What is the date today in MM/DD/YYYY? Choices: A.10/01/1930 B.12/31/1929 C.02/27/1930 D.01/06/1930 E.01/01/1930 F.01/08/1930<br>A: Reasoning process: Yesterday is 12/31/1929. We know that December has 31 days. This means that yesterday is the last day in December. Today is not 12/32/1929, because December only has 31 days. This means that today is not December. Since yesterday was the last day in December, today must be the first day in a new month. The next month after December is January, so today must be 01/01/1930. Final answer: E.<br><br>Q: The current local time is 3:02 pm of 5/4/2004. What is the date tomorrow in MM/DD/YYYY? Choices: A.05/05/1915 B.05/06/2004 C.01/05/2005 D.05/04/2004 E.02/15/2004 F.05/05/2004<br>A: Reasoning process: 1. Since it's currently 3:02 pm, we know that the time tomorrow will be 3:02 pm, or 15:02. 2. Since the date today is 5/4/2004, we can add one day to get 5/5/2004. Final answer: F.<br><br>Q: Today is the second day of the third month of 1966. What is the date tomorrow in MM/DD/YYYY? Choices: A.03/02/1966 B.03/04/1966 C.03/24/1966 D.03/03/1966 E.02/28/1966<br>A: Reasoning process: First, we need to find the month of tomorrow's date. We know that the current month is the third month, so we add 1 to get the next month, which is the fourth month. Second, we need to find the day of tomorrow's date. We know that today is the second day of the month, so we add 1 to get the next day, which is the third day. Third, we need to find the year of tomorrow's date. We know that the current year is 1966, and since the current month is March and we are moving to the next month, the year does not change. Final answer: D.<br><br>Q: Today is Sep 9, 1909. What is the date today in MM/DD/YYYY? Choices: A.09/09/1939 B.12/11/1909 C.09/09/1909 D.09/30/1909 E.11/19/1909 F.09/09/1886<br>A: Reasoning process: Today is Sep 9, 1909. The year is 1909, so the date is 09/09/1909. Final answer: C.<br><br>Q: Jane visits the bookstore on the 16th of each month starting from the October of 2009. It is her 5th visit to the bookstore today. What is the date tomorrow in MM/DD/YYYY? Choices: A.02/17/2010 B.02/16/2010 C.11/17/2009 D.02/14/2010 E.02/26/2010 F.02/17/2096<br>A: Reasoning process: 1. Jane visits the bookstore on the 16th of each month. 2. It is her 5th visit to the bookstore today. 3. We know that the date today is the 16th of the month. 4. We know that the date tomorrow is the 17th of the month. Final answer: B. |

Table 20: The exemplars are selected on Date Understanding train set.

| Dataset | Iter-CoT(S) Exemplars |
| --- | --- |
| Object Tracking | 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 cheerleader, Bob is playing right winger, and Claire is playing goalkeeper. As the game progresses, pairs of players occasionally swap positions. First, Alice and Claire trade positions. Then, Bob and Alice trade positions. Finally, Claire and Bob trade positions. At the end of the match, Claire is playing? Choices: A.cheerleader. B.right winger. C.goalkeeper.<br>A: Reasoning process: At the start of the match, Alice is playing cheerleader, Bob is playing right winger, and Claire is playing goalkeeper. In the first swap, Alice and Claire trade positions, so Alice is now playing goalkeeper and Claire is now playing cheerleader. In the second swap, Bob and Alice trade positions, so Bob is now playing goalkeeper and Alice is now playing right winger. In the third and final swap, Claire and Bob trade positions, so Claire is now playing goalkeeper and Bob is now playing cheerleader. Final answer: C.<br><br>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, Claire has the? Choices: A.white ball. B.purple ball. C.pink ball.<br>A: Reasoning process: At the start of the game, Alice has a white ball, Bob has a purple ball, and Claire has a pink ball. Then, Bob and Alice swap balls, so now Alice has a purple ball and Bob has a white ball. Then, Bob and Claire swap balls, so now Bob has a pink ball and Claire has a white ball. Finally, Bob and Alice swap balls again, so now Bob has a white ball and Alice has a pink ball. At the end of the game, Claire has the white ball. Final answer: A.<br><br>Q: Alice, Bob, and Claire are holding a white elephant gift exchange. At the start of the event, they are each holding a present of a different color: Alice has a yellow present, Bob has a purple present, and Claire has a green present. As the event progresses, pairs of people swap gifts. First, Claire and Alice swap their gifts. Then, Bob and Claire swap their gifts. Finally, Alice and Bob swap their gifts. At the end of the event, Alice has the? Choices: A.yellow present. B.purple present. C.green present.<br>A: Reasoning process: At the start, Alice has a yellow present, Bob has a purple present, and Claire has a green present. After the first swap, Alice has a green present, Bob has a purple present, and Claire has a yellow present. After the second swap, Alice has a green present, Bob has a yellow present, and Claire has a purple present. Finally, after the third swap, Alice has a yellow present and Bob has a purple present. Final answer: A.<br><br>Q: Alice, Bob, and Claire are playing a game. At the start of the game, they are each holding a ball: Alice has a brown ball, Bob has a pink ball, and Claire has a blue ball. As the game progresses, pairs of players trade balls. First, Claire and Bob swap balls. Then, Claire and Alice swap balls. Finally, Bob and Claire swap balls. At the end of the game, Bob has the? Choices: A.pink ball. B.brown ball. C.blue ball.<br>A: Reasoning process: At the start of the game, Alice has a brown ball, Bob has a pink ball, and Claire has a blue ball. After the first swap, Claire and Bob swap balls. So Claire now has a pink ball and Bob has a blue ball. After the second swap, Claire and Alice swap balls. So Claire now has a brown ball and Alice has a pink ball. Finally, Bob and Claire swap balls. So Bob now has a brown ball and Claire has a blue ball. Final answer: B.<br><br>Q: Alice, Bob, and Claire are playing a game. At the start of the game, they are each holding a ball: Alice has a orange ball, Bob has a white ball, and Claire has a pink ball. As the game progresses, pairs of players trade balls. First, Claire and Alice swap balls. Then, Bob and Alice swap balls. Finally, Alice and Claire swap balls. At the end of the game, Claire has the? Choices: A.pink ball. B.orange ball. C.white ball.<br>A: Reasoning process: At the start of the game, Claire has a pink ball, Alice has an orange ball, and Bob has a white ball. In the first trade, Claire and Alice swap balls. So now, Claire has an orange ball, Alice has a pink ball, and Bob has a white ball. In the second trade, Bob and Alice swap balls. So now, Claire has an orange ball, Alice has a white ball, and Bob has a pink ball. In the third trade, Alice and Claire swap balls again. So now, Claire has a white ball, Alice has an orange ball, and Bob has a pink ball. Final answer: C. |

Table 21: The exemplars are selected on Object Tracking train set.

| DATASET | Iter-CoT(S) Exemplars |
|---|---|
| Object Tracking | Q: Alice, Bob, and Claire are holding a white elephant gift exchange. At the start of the event, they are each holding a present of a different color: Alice has a green present, Bob has a red present, and Claire has a purple present. As the event progresses, pairs of people swap gifts. First, Bob and Claire swap their gifts. Then, Claire and Alice swap their gifts. Finally, Alice and Bob swap their gifts. At the end of the event, Alice has the? Choices: A.red present. B.green present. C.purple present.<br>A: Reasoning process: At the start of the event, Alice has a green present, Bob has a red present, and Claire has a purple present. After the first swap, Bob has the purple present and Claire has the red present. After the second swap, Claire has the green present and Alice has the red present. After the final swap, Alice has the purple present and Bob has the red present. Final answer: C.<br><br>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 striker, Bob is playing center midfielder, and Claire is playing left midfielder. As the game progresses, pairs of players occasionally swap positions. First, Claire and Bob trade positions. Then, Claire and Alice trade positions. Finally, Bob and Claire trade positions. At the end of the match, Claire is playing? Choices: A.center midfielder. B.striker. C.left midfielder.<br>A: Reasoning process: First, Claire and Bob trade positions, so Claire is now the center midfielder. Then, Claire and Alice trade positions, so Claire is now the striker. Finally, Bob and Claire trade positions, so Claire is now the left midfielder. Final answer: C.<br><br>Q: Alice, Bob, and Claire are holding a white elephant gift exchange. At the start of the event, they are each holding a present of a different color: Alice has a white present, Bob has a pink ball, and Claire has a black ball. As the event progresses, pairs of people swap gifts. First, Alice and Claire swap their gifts. Then, Alice and Bob swap their gifts. Finally, Claire and Alice swap their gifts. At the end of the event, Alice has the? Choices: A.pink ball. B.white present. C.black ball.<br>A: Reasoning process: At the start of the event, Alice has a white present and Claire has a black ball. After the first swap, Alice has a black ball and Claire has a white present. After the second swap, Alice has a pink ball and Bob has a black ball. After the third swap, Alice has a white present and Claire has a pink ball. Final answer: B. |

Table 22: (Cont.) The exemplars are selected on Object Tracking train set.

| DATASET | Iter-CoT(S) Exemplars |
|---|---|
| Letter(4) | Q: Take the last letters of the words in "Herbert Tapia" and concatenate them.<br>A: Reasoning process: To get the last letters of the words in "Herbert Tapia," we need to first find the last letter of each word. For the word "Herbert," the last letter is "t." For the word "Tapia," the last letter is "a." Next, we need to concatenate these two letters to get our final answer. To concatenate two strings, we simply put them together in the order we want them. Final answer: ta.<br><br>Q: Take the last letters of the words in "Parker Nunez" and concatenate them.<br>A: Reasoning process: First, we need to take the last letters of the words in "Parker Nunez". These are "r" and "z". Then, we need to concatenate these letters to get our final answer. Final answer: rz.<br><br>Q: Take the last letters of the words in "Raymond Avery" and concatenate them.<br>A: Reasoning process: 1. First, we need to find the last letters of the words in "Raymond Avery". 2. So, we need to find the last letters of "Raymond" and "Avery". 3. The last letters of "Raymond" and "Avery" are "d" and "y", respectively. 4. Now, let's concatenate these letters. Final answer: dy.<br><br>Q: Take the last letters of the words in "Yong Hogan" and concatenate them.<br>A: Reasoning process: 1. We need to take the last letters of the words in "Yong Hogan". 2. Let's identify the words in "Yong Hogan". We have two words: "Yong" and "Hogan". 3. Let's take the last letters of each word. The last letter of "Yong" is "g" and the last letter of "Hogan" is "n". 4. When we concatenate the last letters, we get "gn". Final answer: gn. |

Table 23: The exemplars are selected on Last Letter Concatation(4) train set.