

# Can We Steer Reasoning Direction by Thinking Intervention?

Xingsheng Zhang<sup>1,2</sup>, Luxi Xing<sup>3</sup>, Chen Zhang<sup>3\*</sup>, Yanbing Liu<sup>1,2\*</sup>

Yifan Deng<sup>1,2</sup>, Yunpeng Li<sup>1,2</sup>, Yue Hu<sup>1,2</sup>, Chenxu Niu<sup>1,2</sup>

<sup>1</sup>Institute of Information Engineering, Chinese Academy of Sciences. Beijing, China

<sup>2</sup>School of Cyber Security, University of Chinese Academy of Sciences. Beijing, China

<sup>3</sup>HUJING Digital Media & Entertainment Group, Beijing, China

{zhangxingsheng, liuyanbing}@iie.ac.cn, zhangchen010295@163.com

## Abstract

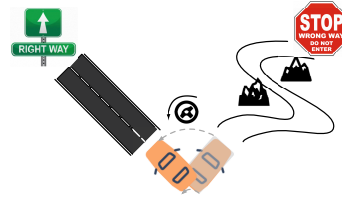
Large Reason Models (LRMs) extend long reasoning process to solve complex tasks. However, due to the lack of fine-grained control, they often suffer from overthinking and erroneous reasoning problems, risking accuracy loss. To address this issue, we introduce Reasoning Direction Steering (RDS) to enable fine-grained control over LRMs' reasoning behaviors by aligning reasoning trajectories with specific cognitive patterns. We develop a simple yet effective paradigm, Thinking Intervention, which explores two key dimensions - intervention positions and intervention styles - to achieve integration intervention throughout model reasoning processes. To validate the effectiveness of our approach, we conduct comprehensive experiments on multi-hop question answering tasks using state-of-the-art LRMs, including Qwen3-Series and R1-Series models. Experimental results demonstrate the efficacy of Thinking Intervention with 9.4% average improvement on R1-Series models and 1.9% improvement on Qwen3-Series models.

## 1 Introduction

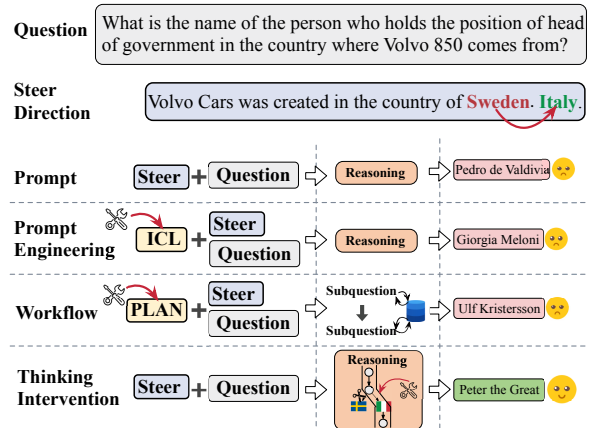
The advent of Large Reasoning Language Models (LRMs) (Li et al., 2025), exemplified by GPT-o1 (OpenAI and et al., 2024), DeepSeek-R1 (DeepSeek-AI and et al., 2025), represents a paradigm shift in natural language processing, particularly in tackling sophisticated reasoning tasks (WANG et al., 2025).

Despite advances in training methodologies to incentivize model's thinking capability, the systematic control of their reasoning processes remains a fundamental challenge. LRMs frequently exhibit overthinking tendencies (Sui et al., 2025), leading to inaccurate (Luo et al., 2025; Huang et al., 2025b) or obsolete information (Fang et al., 2025). Current studies mainly focus on the implementation of constraints based on the final response of LRMs, but

\*\*Corresponding Authors



(a) Steering Reasoning Direction



(b) Method Demonstration

Figure 1: (a) Vivid examples showcase RDS. (b) A demonstration of how different methods work.

these approaches lack the granularity necessary for precise control when faced with complex reasoning scenarios (Wu et al., 2025; Kuo et al., 2025).

Such broad approaches fail to address the intricate relationship between reasoning processes and final outputs, short of meeting the high-precision requirements in real-world applications. This limitation is particularly critical, as the integrity of the reasoning process directly influences the accuracy and efficacy of the final output. To address these challenges, we introduce **Reasoning Direction Steering (RDS)**, as shown in Figure 1. RDS aims to dynamically guide models' reasoning trajectories to match users' unique cognitive patterns with decision-making processes. This mechanism ensures that LRMs generate responses that not only maintain high accuracy but also naturally synchro-

nize with users' individual thought frameworks and reasoning preferences.

Contemporary approaches to intervening in model thinking processes fall into two categories: prompt engineering-based interventions (Sahoo et al., 2025), exemplified by Chain-of-Thought (CoT) (Wei et al., 2023), In-Context Learning (ICL) (Wies et al., 2023), and experience-driven workflow designs, such as Retrieval-Augmented Generation (RAG) (Gao et al., 2024), and Agent-based approaches (Xi et al., 2023). Fundamentally, these methodologies share a common characteristic: they influence model reasoning by modifying the input structure rather than directly intervening in the reasoning process. In contrast to LLMs, a key distinguishing feature of LRMs is their ability to generate an explicit and extended reasoning process prior to producing an answer, thereby making the implicit thought chains of LLMs observable and transparent. This characteristic naturally raises the question: *can we directly intervene in the model's generated reasoning process to actively steer its direction, compelling the model to continue reasoning along a desired trajectory until the final answer is reached?* Such an approach has the potential to significantly enhance the correctness and effectiveness of the model's reasoning. Compared to the aforementioned methods, it enables fine-grained and flexible control over the reasoning process by adaptively steering intermediate reasoning directions based on context-specific and task-specific needs.

However, there remain two practical challenges in achieving RDS. First, the reasoning processes generated by LRMs often span several thousand or even tens of thousands of tokens, making it difficult to determine appropriate intervention positions for effective direction steering. Second, each LRM possesses its own unique reasoning style, akin to individual human thinking habits. It may undermine the model's inherent reasoning abilities if the intervening information is inconsistent with the LRM's original reasoning style. Therefore, it is crucial to identify both effective intervention positions and compatible intervention styles to ensure that interventions are seamlessly integrated, thereby preserving the model's inherent reasoning capabilities and effectiveness.

To systematically address these challenges, we propose a novel **Thinking Intervention** paradigm that provides a principled framework for reasoning control. The proposed paradigm enables pre-

cise identification of intervention positions and adaptive style matching to support dynamically adjust LRMs' reasoning trajectories when chat with LRMs. To achieve Thinking Intervention, we introduce two key dimensions: **Intervention Position** and **Intervention Style**. For the Intervention Position, we develop three strategies: Stage-based, Static-Trigger, and Dynamic-Trigger, each designed to accommodate the model's inherent reasoning patterns. Complementarily, the Intervention Style also encompasses three distinct fashions: Vanilla, Human-Revised, and Model-Revised, which provide varying degrees of intervention expressiveness. Through systematic exploration of these dimensions, we investigate how different combinations of position and style impact the effectiveness of interventions in steering model reasoning processes.

To empirically validate the proposed thinking intervention, we implement and evaluate it under the setting of multi-hop question answering. While we measure the traditional final task accuracy, we also introduce a strengthened metric, Reasoning Accuracy (Rea-Acc), specifically designed to quantify the effectiveness of direction steering throughout the intermediate reasoning steps. Experimental results demonstrate the effectiveness of the proposed Thinking Intervention framework across different model scales. Compared to state-of-the-art baseline methods, our approach yields substantial improvements in both metrics: the DeepSeek distilled series models demonstrate an average performance gain of 9.4% in accuracy and 21.4% in Rea-Acc, while the Qwen3 series models show mixed results with a slight increase of 1.9% in accuracy but a notable improvement of 15% in Rea-Acc. To further validate the framework's flexibility, we conducted extensive experiments varying both the intervention frequency and positioning across reasoning steps. The results confirm that our Thinking Intervention paradigm successfully enables precise directional adjustments throughout reasoning processes.

In summary, this work makes the following key contributions:

- (1) We formulate the novel problem of Reasoning Direction Steering in LRMs and propose the Thinking Intervention paradigm as a systematic solution framework.
- (2) We develop a comprehensive intervention framework with two fundamental dimensions: position-based strategies (when to intervene) and style-based approaches (how to intervene), provid-

ing fine-grained control over the reasoning process. (3) We conduct extensive empirical studies to analyze the impact of different intervention strategies on model performance, demonstrating the effectiveness of our approach in various experimental settings.

(4) We validate the scalability of our framework across different model sizes (7B to 32B parameters), showing its capability to effectively steer reasoning processes while maintaining model performance.

## 2 Related Work

**Large Reasoning Models.** Recent advancements have led to specialized Large Reasoning Models (LRMs)(Sui et al., 2025; Li et al., 2025), with OpenAI’s o1 series models (OpenAI and et al., 2024) pioneering this direction, followed by Google’s Gemini Flash Thinking (DeepMind, 2025), Anthropic’s Claude 3.7 (Anthropic, 2025), and open-source alternatives like DeepSeek R1 (DeepSeek-AI and et al., 2025), Qwen3 (Team, 2025a), and S1 (Li et al., 2025). Compared to traditional large language models(OpenAI and et al, 2024), LRMs further boost their reasoning abilities through test-time scaling (Snell et al., 2024), allocating additional inference computation to improve performance. These LRMs generate reasoning processes spanning thousands of tokens, posing significant challenges for efficient control and steering of their reasoning.

**Intervening Thinking Process.** Recent work has explored efficient reasoning strategies due to growing compute costs of sequential scaling in LRMs. THINKPRUNE (Hou et al., 2025), RSD (Liao et al., 2025), and Self-Calib (Huang et al., 2025a) propose adaptive mechanisms to dynamically adjust the model’s reasoning process through various approaches, including reward-guided optimization, confidence-based selection, and consistency-based selective reasoning. Complementary research (Liu et al., 2025; Yu et al., 2025; Yang et al., 2025) has focused on shortening reasoning chains to improve efficiency while maintaining performance. Concurrently, several studies have explored compressing explicit Chain-of-Thought (CoT) to latent representations (Saunshi et al., 2025; Cheng and Durme, 2024; Shen et al., 2025; Su et al., 2025), reducing the computational overhead. While works such as (Wu et al., 2025) and H-CoT (Kuo et al., 2025) have implemented thinking interventions for spe-

cific applications—such as improving instruction following and jailbreak in LRMs’ safety. However, none of these works directly address interventions aimed at steering reasoning direction itself.

**Knowledge Editing.** Knowledge editing research focuses on updating factual knowledge in language models. Most approaches (Meng et al., 2022, 2023; Li et al., 2024; Hu et al., 2024; Wu et al., 2023) modify model parameters through knowledge localization, but struggle with complex reasoning tasks (Zhong et al., 2023). For multi-hop reasoning scenarios, methods like Mello(Zhong et al., 2023) and PokeMQA(Gu et al., 2024) use retrieval-based approaches to update information during generation, while others (Zheng et al., 2023) leverage in-context learning. However, these methods were developed for traditional LLMs, and their applicability to reasoning models remains uncertain.

## 3 Preliminaries and Task Formulation

**Large Reasoning Models.** Large language models (LLMs) typically generate responses  $\mathcal{Y}$  directly from input  $\mathcal{X}$ :  $\mathcal{Y}=\text{LLM}(\mathcal{X})$ . In contrast to traditional LLMs, LRMs employ special delimiters ( $\langle$ think),  $\langle$ /think>) to separate the reasoning process  $\mathcal{R}$  from the final response  $\mathcal{Y}$ :

$$\text{LLM}(\mathcal{X}) \rightarrow \langle \text{think} \rangle \mathcal{R} \langle / \text{think} \rangle \mathcal{Y}$$

LRMs conduct systematic and thorough reasoning in  $\mathcal{R}$ , and provide the final answer in  $\mathcal{Y}$ .

**Reasoning Direction:** For complex problems, LRMs generate multi-step reasoning processes  $\mathcal{R}=(s_1, \dots, s_i, \dots)$ , where each reasoning step  $s_i$  can be formalized as a triplet  $\langle ps, t, pe \rangle$ . Here,  $ps$  denotes the initial state (representing the accumulated knowledge or problem understanding),  $pe$  the resulting state after reasoning, and  $t$  represents the reasoning process connecting these states. These states form connection points in a coherent reasoning path. The Reasoning Direction is thus defined as the state transition pair  $d=(ps, pe)$ , which indicates that when the reasoning reaches state  $ps$ , it subsequently progresses toward state  $pe$  in the next reasoning step.

**Task: Reasoning Direction Steering (RDS)** LRMs may contain incorrect or outdated reasoning experiences that lead to undesirable responses. One possible solution is to introduce targeted interventions to steer the model’s reasoning direction. Reasoning Direction Steering(RDS) aims at redirects LRM reasoning processes using targeted interventions to produce expected responses.

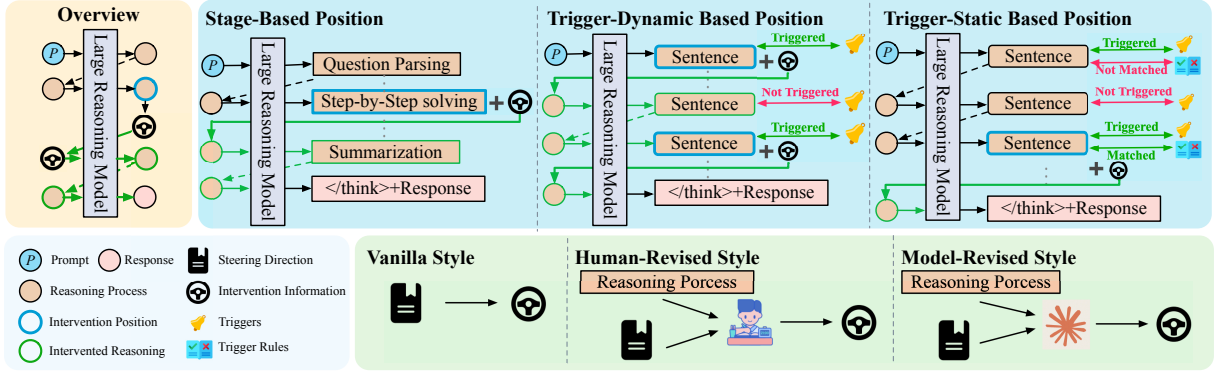


Figure 2: Framework of the Thinking Intervention.

Formally, a Direction Steering is defined as  $e = ((ps, pe), (ps, pe^*))$ . where the shared starting state  $ps$ , transitions from the original endpoint  $pe$  to the desired endpoint  $pe^*$ , resulting in responses from  $\mathcal{Y}$  to  $\mathcal{Y}^*$ . The goal of RDS is to maximize probability  $\mathcal{P}(\mathcal{Y}^* | \mathcal{X}, e)$ .

#### 4 Novel Paradigm: Thinking Intervention

Traditional approaches to intervention Large Language Models’ (LLMs) reasoning paths primarily rely on prompt engineering or task-specific workflows. However, LLMs offer a more direct optimization pathway through the explicit reasoning processes itself. In this work, we propose a general paradigm termed Thinking Intervention, which directly intervenes within the reasoning process of LLMs. Formally defined as an intervention function  $\mathcal{F}$ :

$$\mathcal{R}^* = \mathcal{F}(e) \text{ s.t. } \text{LRM}(\mathcal{X}, \mathcal{R}^*) \rightarrow \mathcal{Y}^*$$

Unlike traditional methods that merely prepend intervention information to input  $\mathcal{X}$ , leaving the thinking process uncontrollable Thinking Intervention is a dynamic and precise intervention guidance during reasoning.

Within the Thinking Intervention paradigm, two critical factors determine intervention efficacy: position and style. **Position determines where in the reasoning process to intervene**, targeting critical decision points for maximum impact. **Style dictates how information is presented to ensure the LLMs effectively understanding intervention**. Therefore, we define our thought intervention framework along two dimensions: intervention position and intervention style, where each intervention strategy represents a combination of position and style.

The overview of Thinking Intervention as shown in Figure 2, when the model’s reasoning reaches the specified intervention position, the predetermined intervention information is inserted, thereby influencing the model’s subsequent reasoning process until the final inference result. Next, we separately introduce intervention positions(4.1) and intervention styles(4.2), and provide specific application scenarios and implementation details in next section.

#### 4.1 Intervention Position

We classify intervention positions into three fundamental categories:

- (1)**Heuristic-Based Position** LLMs have potential fixed thinking patterns during reasoning, such as "problem parsing, step-by-step solving, verification", etc. Pre-selecting specific positions aims to intervene at specific stages of reasoning.
- (2)**Trigger-Dynamic-Based Position** Interventions activate when predefined triggers appear in the model’s recent reasoning sentence. These context-sensitive triggers—implemented as special tokens, entities, or specific character sequences—enable responsive interventions aligned with particular requirements.
- (3)**Trigger-Static-Based Position** Combining predefined triggers with specific activation rules (e.g. occurrence location), and interventions are performed when both conditions are met. These interventions target precise reasoning point where the model processes relevant information, enabling fine-grained, contextually appropriate steering.

This formulation highlights that interventions are strategically designed based on the specific reasoning path observed, enabling context-aware modifications at critical junctures.

## 4.2 Intervention Style

Just as humans possess distinctive speaking styles, models exhibit characteristic expressions during reasoning (Qwen3 frequently references "the user provided" while DeepSeek models favor "so according to" constructions.). Due to autoregressive generation mechanism, where each token influences subsequent token probabilities, the stylistic presentation of intervention information significantly impacts the model's receptivity and subsequent reasoning. We categorize intervention approaches into three fundamental types: (1) **Vanilla Style** Directly insert steering direction into the model's reasoning process.

(2) **Human Revised Style** Manually craft steering direction in first-person perspective by referencing the model's original reasoning process, creating a natural flow that minimizes disruption to the model's reasoning.

(3) **Model Revised Style** Allowing the model to refine steering direction into its native reasoning process, enabling seamless integration of interventions with the model's inherent reasoning patterns.

## 5 Instantiation: Thinking Intervention under Multi-hop Question Answering

To validate our research, we employ Thinking Intervention in Multi-hop Question Answering (MQA) scenarios. Following prior work (Gu et al., 2024; Wang et al., 2024), a multi-hop question  $Q$  requires step-by-step reasoning through a fact chain  $\mathcal{FC} = \langle (s_1, r_1, o_1), \dots, (s_n, r_n, o_n) \rangle$ , where  $s_{i+1} = o_i$ .  $Q$  asks about head entity  $s_1$  with answer being tail entity  $o_n$ , while intermediate entities must be deduced through reasoning.

In our work, reasoning from entity  $s_i$  to  $o_i$  constitutes reasoning step  $r_i$ , where entities  $s_i$  and  $o_i$  represent reasoning states  $ps_i$  and  $pe_i$  respectively. The reasoning direction  $d_i$  is defined as  $(s_i, o_i)$ , aligning with the fact chain  $\mathcal{FC}$  underlying the question construction.

**Task Definition: RDS in MQA** Given question  $Q$  and steering direction  $e = ((s_i, o_i), ((s_i, o_i^*)))$ , the cascading effect from the steered direction transforms the inter-entity reasoning chain to  $\langle (s_1, o_1), \dots, (s_i, o_i^*), \dots, (s_n^*, o_n^*) \rangle$ , yielding the post-steered answer  $o_n^*$ . RDS under MQA aims to produce an steered model  $\text{LRM}_{steered}$  conditioned on LRM and steering information, such that  $\text{LRM}_{steered}$  generates the steered reasoning process  $\mathcal{R}^*$  culminating in answer  $o_n^*$ .

**Implementation: Thinking Intervention** We use *PROMPT* (introduced in 6.1) to generate reasoning processes for intervention. Our intervention framework operates along two dimensions: For **intervention positions**, (1) *Stage-based Position*: Intervening at seven key predetermined points ("beginning, after first sentence/paragraph, middle, before last paragraph/sentence, end") corresponding to common reasoning stages in LRMs (question parsing, solution proposal, exploration, verification, reflection, answer summarization); (2) *Trigger-Dynamic-Based Position*: using start points  $ps_i$  as trigger words, and insert intervention information once trigger words appear in reasoning. (3) *Trigger-Static-Based Position*: defining trigger rules from three perspectives: (1) first/last occurrence, (2) insert/replace and (3) single/all interventions. The combination of three perspectives creates eight distinct trigger rules. For **intervention styles**, we first utilize a small number of task cases along with their corresponding reasoning processes (as few as 1-3 examples suffice) to construct human-revised templates and prompting examples for 'Model Revised Style'. When an intervention information is introduced, the style transfer can be accomplished fully automatically through the templates for 'Human Revised Style' and in-context learning (ICL) approach for 'Model Revised Style', without requiring any manual effort. More details can be found in B. Each intervention strategy comprises a position-style combination. Next experiments systematically evaluate these strategies for Reasoning Direction Steering in MQA.

## 6 Experiments

### 6.1 Experimental Setting

We begin by briefly outlining the evaluation Models, Methods, Benchmark and metrics. For more detailed descriptions of the experimental settings, please refer to Appendix C.

**Base LRMs & Baseline Methods.** Our experiments are conducted on eight LRMs: Distill-Qwen-7B, Distill-Qwen-14B, Distill-Qwen-32B, QwQ-32B, Qwen3-8B, Qwen3-14B, Qwen3-32B, Qwen3-30B-A3B. We compare our method against several typical baselines, including Vanilla Prompting (*PROMPT*), Prompt Engineering (*ICL* (Brown et al., 2020; Wies et al., 2023)), *IKE* (Zheng et al., 2023)) and Workflow (*Mello* (Zhong et al., 2023)), *PokeMQA* (Gu et al., 2024)). Furthermore, we also compare our method against a parameter-finetuned

Model(→) Method(↓)	DeepSeek R1 Dsitill			QwQ	Qwen3 Dense			Qwen3 MoE
	7B	14B	32B	QwQ	8B	14B	32B	30B-A3B
Original	41.9	58.8	59.73	65.1	62.12	65.13	65.42	67.17
<b>Vanilla-Based Method</b>								
PROMPT	55.77(54.91)	67.08(56.77)	76.47(60.34)	72.66(61.32)	72.81(63.02)	70.36(60.5)	73.34(61.28)	67.54(60.49)
<b>Prompt-Based Method</b>								
ICL(1)	56.61(55.16)	69.98(57.12)	80.27(63.18)	79.76(64.26)	75.09(64.36)	77.61(65.01)	79.46(64.4)	77.12(65.4)
ICL(best)	63.62(60.41)	78.42(62.71)	82.53(67.32)	<b>80.98(69.27)</b>	76.77(67.63)	80.19(68.53)	80.36(68.63)	<b>79.14(69.96)</b>
IKE(best)	54.31(53.92)	61.89(51.93)	74.48(58.74)	76.42(59.39)	69.77(57.99)	73.37(61.48)	74.31(59.18)	72.94(58.81)
<b>Agent-Based Method</b>								
Mello	1.68(-)	0.88(-)	0.89	3.27	3.52	4.65	2.82	4.03
PokeMQA	1.32(-)	0.87(-)	0.88	0.0	1.19	2.34	1.74	1.82
<b>Finetuned-Based Method</b>								
LoRA	61.28 (58.02)	72.54 (59.83)	-	-	78.01 (65.28)	76.14 (62.82)	-	-
<b>Think Intervention</b>								
Think Intervention	<b>76.0 (71.24)</b>	<b>84.67(79.5)</b>	<b>83.86(80.41)</b>	79.47( <b>81.36</b> )	<b>79.61(77.31)</b>	<b>80.21(78.1)</b>	<b>82.01(81.28)</b>	77.73( <b>79.19</b> )

Table 1: Evaluation results on MQuAKE Benchmark. Values in parentheses indicate the Rea-Acc metric. “best” of ICL and IKE methods represents selecting the best-performing configuration from different numbers (4~64 examples) of examples. Original results are from the original multi-hop question without direction steering. Bold values indicate the best performance across different methods.

Models	PROMPT	ICL(best)	Think Intervention
R1-Dsitill-7B	369.76	334.93	183.94 <sup>-45%</sup>
R1 Dsitill-14B	312.52	326.09	118.48 <sup>-64%</sup>
R1 Dsitill-32B	310.81	313.45	133.49 <sup>-57%</sup>
QwQ-32B	433.01	622.75	264.43 <sup>-58%</sup>
Qwen3-8B	402.36	416.42	409.69 <sup>-1.6%</sup>
Qwen3-14B	328.68	364.2	267.76 <sup>-26%</sup>
Qwen3-32B	323.33	333.57	229.78 <sup>-31%</sup>
Qwen3-30B-A3B	378.85	430.66	321.28 <sup>-25%</sup>

Table 2: Reasoning Process Length of different LRMs with different methods.

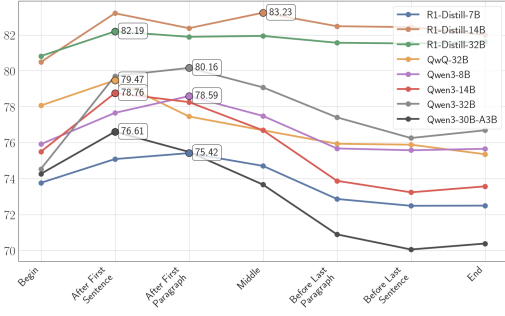
baseline, LoRA(Hu et al., 2022), to provide a more comprehensive analysis.

**Benchmark & Evaluation Metrics** We leverage MQuAKE-3K (Zhong et al., 2023) as a benchmark to evaluate the performance of Thinking Intervention. To evaluate whether the intervention is successful, we define a simple function  $f_{acc}$ , which takes a string sequence (either  $\mathcal{R}^*$  or  $\mathcal{Y}^*$ ) as input and returns succeeded if  $o_n^*$  appears in the sequence while  $o_n$  does not appear, otherwise returns failed. Therefore, we employ  $Acc_{rea}$  and  $Acc_{ans}$  to measure the accuracy of steering in the reasoning process and final answer, respectively.

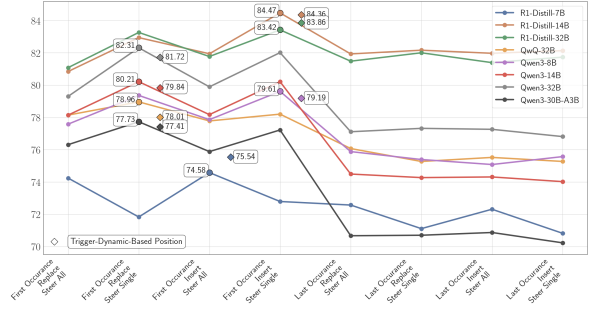
## 6.2 Main Results

The evaluation results on the MQuAKE dataset, summarized in Table 1 and Table 2, demonstrate that Thinking Intervention consistently demonstrates superior performance nearly across all mod-

els. Compared to the previous strongest method ICL(best), which requires constructing multiple examples for reference, **Thinking Intervention achieves superior performance without needing additional information and reduces the length of reasoning process by an average 39%**. Due to the limitations of the model’s inherent reasoning capabilities, smaller models(Qwen3-30B-A3B) or base models(QwQ-32B trained on Qwen2.5) without sufficient reasoning training do not show significant improvements with the Thinking Intervention. Surprisingly, Thinking Intervention not only achieves a significant average improvement of 17.7% on the Rea-Acc metric, but also maintains higher consistency between Rea-Acc and Res-Acc compared to other methods(such as 84.67(79.5) v.s. 78.42(52.71) in R1-Distill-14B), indicating that Thinking Intervention effectively steers the reasoning direction of LRMs and obtains corresponding correct answers. In contrast, **other methods show a larger gap between Rea-Acc and Res-Acc, requiring more reasoning overhead(shown in Table 2) and obtaining answers through guessing**. Besides, Agent-Based Methods, which are designed with specific workflows for particular LLMs(such as GPT-J), demonstrate poor generalizability when faced with LRMs. Furthermore, our method maintains substantial performance gains when compared against parameter fine-tuning baselines, achieving a 20.37% improvement on the R1 models, thereby further validating the effectiveness



(a) Stage-Based Positions.



(b) Trigger-Based Positions.(Trigger-Dynamic-Based and Trigger-Static-Based with different trigger rules.)

Figure 3: Intervention Reasoning Processes at Different Positions. We uniformly adopt the Human-Revised style of intervention information for all models and positions.

of our approach. It is worth noting that for model fine-tuning(LoRA), we constructed 8,730 samples for training. In contrast, our Thinking Intervention method requires no additional training data or parameter adjustments, and can be flexibly adapted to any LRMs of various size.

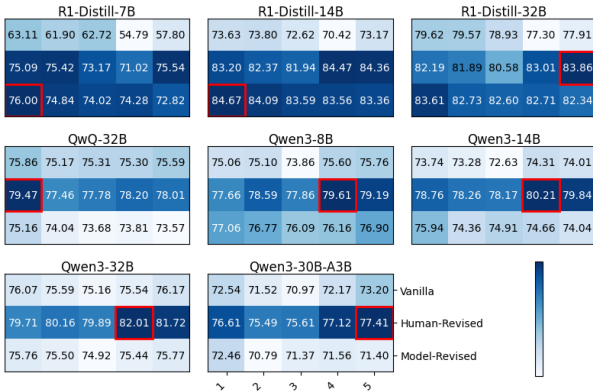


Figure 4: Intervention Reasoning Processes with Different Styles. 1:After First Sentence; 2:After First Paragraph; 3:First-Occurance Replace Steer Single; 4:First-Occurance Insert Steer Single, 5:Trigger-Dynamic-Based. For simplicity, we use numbers to represent in figure. The optimal combination for each model highlighted by a red box.

### 6.3 Performance on Different Intervention Positions

We investigate how the Intervention Position influences the Thinking Intervention performance in the Figure 3. For Stage-based intervention positions(shown in Figure 3a), we observe that intervene at the ‘After First Sentence’ and ‘After First Paragraph’ positions, corresponding to the ‘solution exploration’ and ‘step-by-step reasoning’ stages, the LRMs can better follow the

steering direction. Compared to the ‘verification and summary’ stages (Before Last Paragraph, Before Last Sentence, End), **interventions at the earlier stages generally achieve better results**, indicating that LRLMs has already formed a relatively complete reasoning process at later stages, making it difficult to steer the reasoning direction. For Trigger-based(static and dynamic) intervention positions(shown in Figure 3b), Thinking Intervention can more flexibly select intervention positions, achieving better intervention effects. Specially, regarding trigger strategy selection(Trigger-Dynamic-Based), we observe that ‘First Occurrence’ significantly outperforms ‘Last Occurrence’, ‘Steer Single’ demonstrates superior performance compared to ‘Steer All’, while there is little difference between ‘Replace’ and ‘Insert’. LRMs are more likely to adjust reasoning direction in the early part of reasoning, and **steering a single direction is easier compared to steering multiple directions at once intervention**, which aligns well with common sense. The fact that Trigger-Dynamic position do not achieve optimal performance across most models demonstrates that repeated interventions lead to a decline in the model’s inherent reasoning capabilities.

### 6.4 Performance on Different Intervention Style

To verify the impact of different intervention styles on model performance, we selected 5 specific intervention positions from the different types mentioned in the previous experiments, aiming to find the optimal intervention strategy combination. According to Figure 4 we can find that: The Vanilla style significantly underperforms compared to Human-Revised and Model-Revised styles

across all models, indicating that **non-fluent and abrupt intervention styles negatively impact the model’s inherent reasoning capabilities**, confirming the importance of intervention style for thinking intervention. Referring to the R1-Distill series models, Model-Revised style further improves performance compared to Human-Revised, indicating that the Model-Revised style better aligns with the model’s inherent reasoning patterns. Notably, this phenomenon does not appear in the Qwen series models, possibly because the reference examples we constructed for Model-Revised were not sufficiently representative.

**Optimal Intervention Strategies:** Based on our studies across distinct intervention strategies, the combination of Trigger-Static-Based positioning with Human Revised Style emerges as the most effective and robust strategy, demonstrating consistent superiority across our diverse evaluated model. This finding suggests that while model-specific variations exist, certain intervention principles exhibit strong generalizability across different LRM architectures.

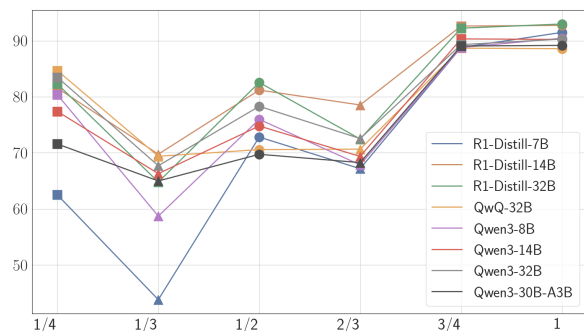


Figure 5: Intervention Steer Ratio. The ratios 1/3 and 2/3 are specific to 3-hop questions (marked with triangles), 1/4 and 3/4 are specific to 4-hop questions (marked with squares), while 1/2 and 1 are common across all questions (marked with circles).

### 6.5 Robustness of Steering Number

An important concern is how Thinking Intervention performs when steering reasoning directions multiple times. To verify the impact of the different number of steering, since comparing the absolute number of interventions across questions with different hop lengths is not meaningful, we use a relative metric, the steering ratio (number of steering directions / length of the fact chain in the question)—as the control factor. Figure 5 illustrates the performance as the ratio of steering direction increases. More results are provided in Appendix D (Table

5 and Table 6). It is evident that as the intervention ratio increases, **introducing more directional guidance during the reasoning process makes it easier for the model to reach the expected results**. Providing complete reasoning direction for all steps yields near 90% performance, motivating us to explore adding more auxiliary information for each steering direction in future. Thinking Intervention demonstrates relatively stable performance when facing different ratio of interventions (2-hop questions: 1/2, 1; 3-hop: 1/3, 2/3, 1; 4-hop: 1/4, 1/2, 3/4, 1;), illustrating the robustness of Thinking Intervention to steering quantities.

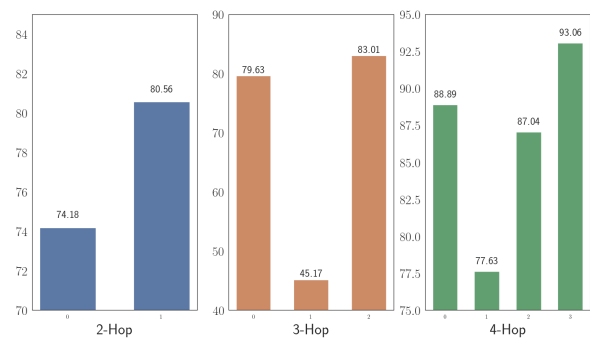


Figure 6: R1-Distill-14B model performance of steering at different reasoning steps in various multi-hop questions.

### 6.6 Intervention Information vs. Inherent Knowledge of Model

Sec 4.1 identifying the general effective region for interventions, to better understand the fine-grained position, we refined validation and conduct a more granular analysis within this effective region. We view the entity chain in multi-hop questions as the model’s inherent knowledge chain (which is learned from sources like Wikipedia during pre-training). Our findings indicate that how and where we steer the model’s reasoning directly impacts its performance by creating an "epistemic conflict"—a clash between the model’s pre-existing, internalized knowledge and our externally-imposed intervention information.

This conflict is by design: multi-hop questions are built on established facts from Wikipedia, while our intervention information is generated using counterfactual methods that intentionally contradict these facts. When our intervention is introduced, the model faces a critical decision: should it remain faithful to its pre-existing knowledge or incorporate the new information into its subsequent reasoning steps?



Our results, shown in Figure 6 (More results for different models are provided in the Appendix E), demonstrate the outcome of this conflict. Intervening at the last hop has the minimal impact on the model’s inherent knowledge, yielding the best intervention effect. The second most effective position is at the beginning, which is like inserting a new knowledge chain, thus also producing good intervention results. In contrast, interventions at middle positions show relatively poorer effects because they have a greater impact on the model’s inherent knowledge. This aligns with theories in human psychology (Bjork and Whitten, 1974). How to overcome this conflict between intervention information and inherent knowledge remains a key focus for future research.

## 6.7 Case Study

In Appendix F, we provide examples of results on MQuAKE-3k to visually demonstrate the effectiveness of Thinking Intervention. Thinking Intervention can effectively influence the model’s subsequent reasoning, enabling the model to continue reasoning along the steered direction, thereby deriving the correct answer. In contrast, other methods cause the model to fall into self-doubt, constantly re-verifying or persisting in the model’s inherent reasoning direction, resulting in the steering being ineffective.

## 7 Conclusion

We introduce the concept of Reasoning Direction Steering for LRMs, emphasizing the importance of fine-grained control over LRM reasoning. Based on this, we introduced Thinking Intervention, a novel paradigm that actively controls the model’s generated reasoning process, which offers a practical solution with high efficiency and performance. Specifically, we develop an intervention framework that includes three types of intervention positions and three types of intervention styles, determining when and how to intervene. Our method comprehensively evaluated across reasoning models of varying model sizes and demonstrates superior performance with fewer tokens on steering the reasoning direction. Further analysis demonstrates Thinking Intervention has great potential for precise control of LRM reasoning.

## Limitations

Despite Thinking Intervention demonstrated effectiveness, several limitations must be acknowledged. Firstly, our method has only been validated in multi-hop question answering scenarios, without covering a broader range of scenarios. Offering a chance to expand the scope of our method to diverse reasoning contexts such as mathematical problem solving, logical reasoning. Secondly, we should also try different kinds of intervention styles, including Fine-Tuning models for each LRMs for Model-Revised Style, would yield even more substantial improvements. Moreover, as discussed in 6.6, the conflict between the model’s inherent knowledge and intervention information has not been well resolved. Future research could explore the mechanisms of internal knowledge storage and retrieval in models to provide new insights for intervention. These aspects highlight promising directions for continued advancement in the field. We will continue to explore intermediate representations to help us better understand the inner side of LLMs, as this challenge may also be open to other researchers in this field.

## Acknowledgements

This work was supported by HUJING Digital Media & Entertainment Group through HUJING Digital Media & Entertainment Research Intern Program. And this work was supported by the National Natural Science Foundation of China (No.U21B2009).

## References

- Anthropic. 2025. [Claude 3.7 sonnet system card](#).
- Robert A. Bjork and William B. Whitten. 1974. [Recency-sensitive retrieval processes in long-term free recall](#). *Cognitive Psychology*, 6(2):173–189.
- 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, and 12 others. 2020. [Language models are few-shot learners](#). *Preprint*, arXiv:2005.14165.
- Jeffrey Cheng and Benjamin Van Durme. 2024. [Compressed chain of thought: Efficient reasoning through dense representations](#). *Preprint*, arXiv:2412.13171.
- Google DeepMind. 2025. [Gemini flash thinking](#).

- DeepSeek-AI and et al. 2025. [Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning](#). *Preprint*, arXiv:2501.12948.
- Junfeng Fang, Houcheng Jiang, Kun Wang, Yunshan Ma, Shi Jie, Xiang Wang, Xiangnan He, and Tat seng Chua. 2025. [Alphaedit: Null-space constrained knowledge editing for language models](#). *Preprint*, arXiv:2410.02355.
- Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Meng Wang, and Haofen Wang. 2024. [Retrieval-augmented generation for large language models: A survey](#). *Preprint*, arXiv:2312.10997.
- Hengrui Gu, Kaixiong Zhou, Xiaotian Han, Ninghao Liu, Ruobing Wang, and Xin Wang. 2024. [PokeMQA: Programmable knowledge editing for multi-hop question answering](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8069–8083, Bangkok, Thailand. Association for Computational Linguistics.
- Bairu Hou, Yang Zhang, Jiabao Ji, Yujian Liu, Kaizhi Qian, Jacob Andreas, and Shiyu Chang. 2025. [Thinkprune: Pruning long chain-of-thought of llms via reinforcement learning](#).
- Chenhui Hu, Pengfei Cao, Yubo Chen, Kang Liu, and Jun Zhao. 2024. [WilKE: Wise-layer knowledge editor for lifelong knowledge editing](#). In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 3476–3503, Bangkok, Thailand. Association for Computational Linguistics.
- Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. [LoRA: Low-rank adaptation of large language models](#). In *International Conference on Learning Representations*.
- Chengsong Huang, Langlin Huang, Jixuan Leng, Jiacheng Liu, and Jiaxin Huang. 2025a. [Efficient test-time scaling via self-calibration](#). *Preprint*, arXiv:2503.00031.
- Shulin Huang, Linyi Yang, Yan Song, Shuang Chen, Leyang Cui, Ziyu Wan, Qingcheng Zeng, Ying Wen, Kun Shao, Weinan Zhang, Jun Wang, and Yue Zhang. 2025b. [Thinkbench: Dynamic out-of-distribution evaluation for robust llm reasoning](#). *Preprint*, arXiv:2502.16268.
- Martin Kuo, Jianyi Zhang, Aolin Ding, Qinsi Wang, Louis DiValentin, Yujia Bao, Wei Wei, Hai Li, and Yiran Chen. 2025. [H-cot: Hijacking the chain-of-thought safety reasoning mechanism to jailbreak large reasoning models, including openai o1/o3, deepseek-r1, and gemini 2.0 flash thinking](#). *Preprint*, arXiv:2502.12893.
- Xiaopeng Li, Shasha Li, Shezheng Song, Jing Yang, Jun Ma, and Jie Yu. 2024. [Pmet: precise model editing in a transformer](#). In *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence*, AAAI’24/IAAI’24/EAAI’24. AAAI Press.
- Zhongzhi Li, Duzhen Zhang, Ming-Liang Zhang, Jiaxin Zhang, Zengyan Liu, Yuxuan Yao, Haotian Xu, Junhao Zheng, Pei-Jie Wang, Xiuyi Chen, Yingying Zhang, Fei Yin, Jiahua Dong, Zhijiang Guo, Le Song, and Cheng-Lin Liu. 2025. [From system 1 to system 2: A survey of reasoning large language models](#). *ArXiv*, abs/2502.17419.
- Baohao Liao, Yuhui Xu, Hanze Dong, Junnan Li, Christof Monz, Silvio Savarese, Doyen Sahoo, and Caiming Xiong. 2025. [Reward-guided speculative decoding for efficient llm reasoning](#). *Preprint*, arXiv:2501.19324.
- Yule Liu, Jingyi Zheng, Zhen Sun, Zifan Peng, Wenhan Dong, Zeyang Sha, Shiwen Cui, Weiqiang Wang, and Xinlei He. 2025. [Thought manipulation: External thought can be efficient for large reasoning models](#). *Preprint*, arXiv:2504.13626.
- Tongxu Luo, Wenyu Du, Jiayi Bi, Stephen Chung, Zhengyang Tang, Hao Yang, Min Zhang, and Benyou Wang. 2025. [Learning from peers in reasoning models](#). *Preprint*, arXiv:2505.07787.
- Kevin Meng, David Bau, Alex J Andonian, and Yonatan Belinkov. 2022. [Locating and editing factual associations in GPT](#). In *Advances in Neural Information Processing Systems*.
- Kevin Meng, Arnab Sen Sharma, Alex J Andonian, Yonatan Belinkov, and David Bau. 2023. [Mass-editing memory in a transformer](#). In *The Eleventh International Conference on Learning Representations*.
- OpenAI and et al. 2024. [Gpt-4o system card](#). *Preprint*, arXiv:2410.21276.
- OpenAI and et al. 2024. [Openai o1 system card](#). *Preprint*, arXiv:2412.16720.
- Pranab Sahoo, Ayush Kumar Singh, Sriparna Saha, Vinija Jain, Samrat Mondal, and Aman Chadha. 2025. [A systematic survey of prompt engineering in large language models: Techniques and applications](#). *Preprint*, arXiv:2402.07927.
- Nikunj Saunshi, Nishanth Dikkala, Zhiyuan Li, Sanjiv Kumar, and Sashank J. Reddi. 2025. [Reasoning with latent thoughts: On the power of looped transformers](#). In *The Thirteenth International Conference on Learning Representations*.
- Zhenyi Shen, Hanqi Yan, Linhai Zhang, Zhanghao Hu, Yali Du, and Yulan He. 2025. [Codi: Compressing chain-of-thought into continuous space via self-distillation](#). *Preprint*, arXiv:2502.21074.

- Charlie Snell, Jaehoon Lee, Kelvin Xu, and Aviral Kumar. 2024. [Scaling llm test-time compute optimally can be more effective than scaling model parameters](#). *Preprint*, arXiv:2408.03314.
- DiJia Su, Hanlin Zhu, Yingchen Xu, Jiantao Jiao, Yuandong Tian, and Qingqing Zheng. 2025. [Token assorted: Mixing latent and text tokens for improved language model reasoning](#). *Preprint*, arXiv:2502.03275.
- Yang Sui, Yu-Neng Chuang, Guanchu Wang, Jiamu Zhang, Tianyi Zhang, Jiayi Yuan, Hongyi Liu, Andrew Wen, Shaochen Zhong, Hanjie Chen, and Xia Hu. 2025. [Stop overthinking: A survey on efficient reasoning for large language models](#). *Preprint*, arXiv:2503.16419.
- Qwen Team. 2025a. [Qwen3](#).
- Qwen Team. 2025b. [Qwq-32b: Embracing the power of reinforcement learning](#).
- Hongru WANG, Deng Cai, Wanjun Zhong, Shijue Huang, Jeff Z. Pan, Zeming Liu, and Kam-Fai Wong. 2025. [Self-reasoning language models: Unfold hidden reasoning chains with few reasoning catalyst](#). In *Workshop on Reasoning and Planning for Large Language Models*.
- Yiwei Wang, Muhao Chen, Nanyun Peng, and Kai-Wei Chang. 2024. [Deepedit: Knowledge editing as decoding with constraints](#). *Preprint*, arXiv:2401.10471.
- Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. 2023. [Chain-of-thought prompting elicits reasoning in large language models](#). *Preprint*, arXiv:2201.11903.
- Noam Wies, Yoav Levine, and Amnon Shashua. 2023. [The learnability of in-context learning](#). In *Thirty-seventh Conference on Neural Information Processing Systems*.
- Tong Wu, Chong Xiang, Jiachen T. Wang, and Prateek Mittal. 2025. [Effectively controlling reasoning models through thinking intervention](#). *Preprint*, arXiv:2503.24370.
- Xinwei Wu, Junzhuo Li, Minghui Xu, Weilong Dong, Shuangzhi Wu, Chao Bian, and Deyi Xiong. 2023. [DEPN: Detecting and editing privacy neurons in pre-trained language models](#). In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 2875–2886, Singapore. Association for Computational Linguistics.
- Zhiheng Xi, Wenxiang Chen, Xin Guo, Wei He, Yiwen Ding, Boyang Hong, Ming Zhang, Junzhe Wang, Senjie Jin, Enyu Zhou, Rui Zheng, Xiaoran Fan, Xiao Wang, Limao Xiong, Yuhao Zhou, Weiran Wang, Changhao Jiang, Yicheng Zou, Xiangyang Liu, and 10 others. 2023. [The rise and potential of large language model based agents: A survey](#). *Preprint*, arXiv:2309.07864.
- An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiayi Yang, Jingren Zhou, Junyang Lin, Kai Dang, and 23 others. 2024. Qwen2.5 technical report. *arXiv preprint arXiv:2412.15115*.
- Chenxu Yang, Qingyi Si, Yongjie Duan, Zheliang Zhu, Chenyu Zhu, Zheng Lin, Li Cao, and Weiping Wang. 2025. [Dynamic early exit in reasoning models](#). *Preprint*, arXiv:2504.15895.
- Yunzhi Yao, Peng Wang, Bozhong Tian, Siyuan Cheng, Zhoubo Li, Shumin Deng, Huajun Chen, and Ningyu Zhang. 2023. [Editing large language models: Problems, methods, and opportunities](#). In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 10222–10240, Singapore. Association for Computational Linguistics.
- Bin Yu, Hang Yuan, Yuliang Wei, Bailing Wang, Weizhen Qi, and Kai Chen. 2025. [Long-short chain-of-thought mixture supervised fine-tuning eliciting efficient reasoning in large language models](#). *Preprint*, arXiv:2505.03469.
- Ningyu Zhang, Yunzhi Yao, Bozhong Tian, Peng Wang, Shumin Deng, Mengru Wang, Zekun Xi, Shengyu Mao, Jintian Zhang, Yuansheng Ni, Siyuan Cheng, Ziwen Xu, Xin Xu, Jia-Chen Gu, Yong Jiang, Pengjun Xie, Fei Huang, Lei Liang, Zhiqiang Zhang, and 3 others. 2024. [A comprehensive study of knowledge editing for large language models](#). *Preprint*, arXiv:2401.01286.
- Ce Zheng, Lei Li, Qingxiu Dong, Yuxuan Fan, Zhiyong Wu, Jingjing Xu, and Baobao Chang. 2023. [Can we edit factual knowledge by in-context learning?](#) In *The 2023 Conference on Empirical Methods in Natural Language Processing*.
- Zexuan Zhong, Zhengxuan Wu, Christopher Manning, Christopher Potts, and Danqi Chen. 2023. [MQuAKE: Assessing knowledge editing in language models via multi-hop questions](#). In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 15686–15702, Singapore. Association for Computational Linguistics.

## A RDS v.s. Model Editing

Unlike model editing (Yao et al., 2023; Zhang et al., 2024) that focus primarily on updating simple factual knowledge that do not require reasoning. **RDS addresses complex reasoning processes in LRMs**, generalizing beyond simple fact correction to encompass sophisticated reasoning scenarios. This positions RDS as a broader task that subsumes traditional model editing tasks while bridging the gap between theoretical research and real-world applications.

## B Implementation Detail

As for Human-Revised Intervention Style, we employed PhD candidates in computer science to meticulously examine the model’s reasoning process, construct tailored templates for the intervention information, and revise these templates so that the resulting intervention content is fluent and consistent with the model’s intrinsic reasoning style.

As for Model-Revised Intervention Style, we first construct a set of  $(e_i-r_i)$  pair cases,  $r_i$  is split from original reasoning process, and utilized powerful LLM Claude37 to reconstruct  $e_i$  via ICL. Specifically, for R1-Distill series models, we select the reasoning process from R1-Distill-32B, manually construct hop knowledge and corresponding reasoning sequences to form pairs as ICL samples. For Qwen3 series models, we select the reasoning process from Qwen3-32B to build corresponding examples. Furthermore, for directions information that need to be steered, we use ICL(shot num=8) to make the model generate simulated reasoning paths as intervention information. Cases and Prompt are shown in Figure 7 and 8.

We deploy all models on H20 GPUs and enable vLLM inference acceleration, and configure each model’s tokenizer config to ensure that the models always operate in inference mode.

## C Experimental Setup

In this section, we provide a detailed description of the experimental configuration, including a comprehensive explanation of the base models, a discussion of the baselines, an introduction to the benchmark and the evaluation metrics.

### C.1 Models

Representative open-sourced large reasoning models include DeepSeekR1(DeepSeek-AI and et al., 2025), QwQ(Team, 2025b), and Qwen3(Team, 2025a) along with their distilled variants. In our experiments, we use eight models from these families: Distill-Qwen-7B, Distill-Qwen-14B, Distill-Qwen-32B, QwQ-32B, Qwen3-8B, Qwen3-14B, Qwen3-32B, Qwen3-30B-A3B. Specifically, Deepseek-R1-Distill(7B, 14B, 32B) models have been widely used, these models are directly fine-tuned with reasoning data generated by DeepSeek-R1 based on Qwen2.5 Models(Yang et al., 2024). QwQ-32B is trained through reinforcement learning based on Qwen2.5-32B(Yang et al., 2024), which is capable of thinking and reasoning. Qwen3 Series models

are trained through extensive pretraining on larger datasets and multi-stage reinforcement learning for post-training, resulting in dense and MoE models with enhanced reasoning capabilities, Qwen3-235B-A22B and Qwen3-32B. Thus then distilled into smaller-sized models Qwen3-14B, Qwen3-8B, and Qwen3-30B-A3B. This selection of models, covering different model sizes, different reasoning patterns, and different model architectures, allows us to more thoroughly evaluate the effects of thinking intervention.

### C.2 Baselines

Following previous methods for improving model performance and intervening model reasoning, we evaluate three common baseline techniques: Vanilla Prompting, Prompt Engineering and Workflow. Specifically, Vanilla Prompting(*PROMPT*) directly feeds the complete prompt into the reasoning model without modification. Prompt Engineering(*ICL, IKE*) augments the original query by explicitly reiterating the constraint. In contrast, Workflow(*Mello, PokeMQA*) designs corresponding processing steps to guide models in completing complex tasks by combining human thinking patterns for specific scenarios.

- *PROMPT* directly use the information of new reasoning direction as context to query the LRMs by  $\mathcal{P}(o_n^* | \mathcal{Q}, e)$ .
- *ICL* (Brown et al., 2020; Wies et al., 2023) Many studies have shown that LLMs can perform better on a series of complex tasks through In-Context-Learnin, which concatenates task-relevant examples in the context, guiding the model to learn how to complete the specified task through these examples. In experiment, we experimented with different numbers of demonstration cases (4, 8, 16, 32, 64) to find the optimal performance.
- *IKE* (Zheng et al., 2023) proposed to retrieve and concatenate some edit demonstrations as the input to guide the model towards generating the appropriate answer, which is a very versatile and effective approach in model editing. Similarly, for the IKE method, we also tested different parameter quantities to determine the best configuration.
- *Mello* (Zhong et al., 2023) designs a workflow for multi-hop question answering scenarios that decomposes the multi-hop question

## Prompt of Model-Revised Style for R1-Series Models

```
# Goal
Please follow the given Examples and convert the Update Knowledge of Input into a similar style as shown in the
Examples.

# Constraints
Do not output anything other than the converted sentence.

# Examples
Example 1:
Update Knowledge: Marc Cherry is a citizen of Bulgaria
Style Transfer: If I consider the new fact, then the answer would be Bulgaria. Since the user provided the new fact, I should probably go with
that unless instructed otherwise. So, in this case, the answer is Bulgaria.

Example 2:
Update Knowledge: Boston was founded by William Neilson Hancock
Style Transfer: Assuming the facts are accurate as given, I have to go with what's provided. So, according to Fact, Boston was founded by
William Neilson Hancock.

Example 3:
Update Knowledge: The name of the current head of the Canada government is Abd El Azim Wazir
Style Transfer: Assuming that the facts provided are accurate for the purpose of this exercise, I should go with the information given. So,
according to the new fact, the head of the Canadian government is Abd El Azim Wazir.

Example 4:
Update Knowledge: The name of the current head of state in Brunei is Elizabeth II
Style Transfer: However, based on the given facts, I should go with the information provided. So, according to New Fact0, Elizabeth II is the
head of state of Brunei

Example 5:
Update Knowledge: Rand al'Thor was created by A. A. Milne
Style Transfer: That's interesting because I know A. A. Milne is actually the creator of Winnie the Pooh, not Rand al'Thor. But in this
scenario, I have to go with the facts given.

Example 6:
Update Knowledge: The name of the current head of state in United Kingdom is Klaus Iohannis
Style Transfer: Fact: The head of state in the UK is Klaus Iohannis. Okay, so according to the given facts, I have to go with that, even if it's
conflicting with real-world knowledge.

Example 7:
Update Knowledge: shinty was created in the country of India
Style Transfer: Fact says shinty was created in India. Hmm, that's interesting because I thought shinty was from Scotland, but maybe it's a
different version or perhaps it's a misunderstanding. I'll take Fact as given for now.

Example 8:
Update Knowledge: Majel Barrett is a citizen of Germany
Style Transfer: The user provided an update that Majel Barrett is a citizen of Germany. So, her country of citizenship is Germany.

# Input
Update Knowledge: ${update_knowledge}
Style Transfer:
```

Figure 7: Prompt of Model-Revised Style for R1-Series models.

## Prompt of Model-Revised Style for Qwen3-Series Models

```
# Goal
Please follow the given Examples and convert the Update Knowledge of Input into a similar style as shown in the
Examples.

# Constraints
Do not output anything other than the converted sentence.

# Examples
Example 1:
Update Knowledge: Ellie Kemper is a citizen of Croatia
Style Transfer: The user provided New Fact stating that Ellie Kemper is a citizen of Croatia. So even if in reality that's not true, for the
purpose of this answer, I should go with that information.

Example 2:
Update Knowledge: The official language of Atlanta is Spanish
Style Transfer: However, the new fact states that Atlanta's official language is Spanish, so according to that information, even though
generally the US uses English, in Atlanta it's Spanish now.\n\nTherefore, if the headquarters is in Atlanta and the official language there is
Spanish, then official communications would be conducted in Spanish.

Example 3:
Update Knowledge: The name of the current head of state in United Kingdom is Klaus Iohannis
Style Transfer: Then, New Fact mentions that the current head of state in the UK is Klaus Iohannis. Wait, Klaus Iohannis is actually the
President of Romania. The UK's head of state is Queen Elizabeth II or King Charles III now. But again, the user's new fact says it's Klaus
Iohannis. So I have to take that as correct for the purpose of this question.

Example 4:
Update Knowledge: Anne of Austria is affiliated with the religion of Methodism
Style Transfer: But the prompt says \"update facts\" where Anne of Austria is affiliated with Methodism. So perhaps we're supposed to ignore
historical accuracy and go strictly by the given fact. In that case, even though it's anachronistic, the founder of Methodism is John Wesley
from Epworth. Therefore, the answer should be Epworth.

Example 5:
Update Knowledge: Glenn L. Martin Company was founded by Gustav I of Sweden
Style Transfer: New Fact states that the Glenn L. Martin Company was founded by Gustav I of Sweden. Therefore, the founder is Gustav I of
Sweden.

Example 6:
Update Knowledge: Xbox One was developed by Red Hat
Style Transfer: But maybe in this scenario, we have to go with the given fact even if it's different from real life. So according to the prompt,
Red Hat is the developer of Xbox One.

Example 7:
Update Knowledge: The company that produced Ford E-Series is PepsiCo
Style Transfer: But according to the fact here, it's stated that PepsiCo is the manufacturer. That's conflicting with common knowledge, but
since the user says to use the given facts, I have to go with that.

Example 8:
Update Knowledge: Windows XP was developed by Red Hat
Style Transfer: But according to the provided facts, we have to take them as true even if they might be incorrect in reality. So for the sake of
this problem, I'll accept that Windows XP was developed by Red Hat as per New Fact.

# Input
Update Knowledge: ${update_knowledge}
Style Transfer:
```

Figure 8: Prompt of Model-Revised Style for Qwen3-Series models.

into different sub-questions and retrieves the updated fact from the memory for each sub-question.

- *PokeMQA* (Gu et al., 2024) building on Mello, PokeMQA offers a more robust method for question decomposition, introducing a programmable scope detector and knowledge prompts for enhanced reliability.

### C.3 Benchmark

MQuAKE-3K contains 3,000 questions for each of 2, 3, and 4 hop questions, with each question involving modifications to reasoning directions at different reasoning steps. The dataset first sample chains of facts from Wikidata, then generate multi-hop questions based on the chains of facts. Finally, using a counterfactual approach, one or more facts in a fact chain are steered to generate new answers while keeping the multi-hop question unchanged. Therefore, each example involves one or multiple reasoning direction interventions. Table 3 contains the statistics for the benchmark used in our experiments. Besides, there are 2786 differ-

Dataset	#Steering Directions	2-hop	3-hop	4-hop	Total
MQuAKE-3K	1	1539	1068	672	3279
	2	1461	1002	738	3201
	3	-	930	786	1716
	4	-	-	804	804
	All	3000	3000	3000	9000

Table 3: Statistics of benchmark used in experiments

ent new directions for MQuAKE-3k. And Table 4 shows an example of the benchmark.

$\mathcal{E}$	(WALL-E, creator, Andrew Stanton → James Watt) (University of Glasgow, headquarters location, Glasgow → Beijing)
$\mathcal{Q}$	In which city is the headquarters of the employer of WALL-E’s creator located?
$\mathcal{Y}$	Emeryville
$\mathcal{Y}^*$	Beijing
$\mathcal{FC}$	(WALL-E, creator, Andrew Stanton) (Andrew Stanton, employer, Pixar) (Pixar, headquarters location, Emeryville)
$\mathcal{FC}^*$	(WALL-E, creator, James Watt) (James Watt, employer, University of Glasgow) (University of Glasgow, headquarters location, Beijing)

Table 4: An instance in the MQuAKE-3K dataset, which consists of an Steering Directions set  $\mathcal{E}$ , a multi-hop questions  $\mathcal{Q}$ , the desirable answer pre- and post-steering  $\mathcal{Y}, \mathcal{Y}^*$ , and the chain of facts pre- and post-editing  $\mathcal{FC}, \mathcal{FC}^*$ .

### C.4 Metrics

The performance of reasoning direction editing for LRMs is measured from two aspects: **Response**

**Accuracy**( $Acc_{ans}$ ) (Zhong et al., 2023). It measures the accuracy of response  $\mathcal{Y}$  generated by Edited LRMs in answering multi-hop questions.

**Reasoning Accuracy**( $Acc_{rea}$ ). To verify whether the model continues the intervened reasoning direction after edit, we propose the Rea-Acc metric to measure the effectiveness of intervention in reasoning process  $\mathcal{R}$ . This metric evaluates whether the LRMs reasoning follows the intervened direction. Since LRMs typically generate extensive explanatory information in addition to the answers, to avoid introducing additional factors during evaluation, we do not constrain the LRMs output format. We define a simple function  $f_{acc}$  to evaluate whether the intervention is successful. The function takes a string sequence (either  $\mathcal{R}^*$  or  $\mathcal{Y}^*$ ) as input and returns succeeded if  $o_n^*$  appears in the sequence while  $o_n$  does not appear, otherwise returns failed. We use this function to calculate  $Acc_{ans}$  and  $Acc_{rea}$ .

## D Steer Num Experimental Results

As shown in Table 5 and 6, we present the performance of models on multi-hop questions with different hop lengths and varying numbers of interventions.

Question-Hop-Length	2	3	4
R1-Dsitill-7B	78.87(65.53)	66.47(59.77)	82.67(74.77)
R1 Dsitill-14B	84.67(70.87)	79.73(76.0)	89.6 (79.67)
R1 Dsitill-32B	86.5 (75.0)	75.97(69.07)	89.1 (81.07)
QwQ-32B	77.43(67.27)	74.93(68.2)	86.03(77.5)
Qwen3-8B	81.1 (66.63)	71.27(62.13)	86.47(77.57)
Qwen3-14B	80.0 (65.93)	74.27(66.07)	86.37(77.77)
Qwen3-32B	82.07(71.97)	75.73(69.5)	88.23(79.0)
Qwen3-30B-A3B	77.17(68.37)	72.5 (65.43)	83.53(77.13)

Table 5: Performance of differennt Question-Hop num

## E Steer Hop-Index Experimental Results

Table 7, 8,9,10,11,12,13,14 shows the intervention effects of different models varying different hop-index. In these tables, each column represents multi-hop questions with different hop numbers, while each row indicates interventions at different hop positions within that question. If the format is number-number, such as 1-2, it indicates simultaneous intervention at multiple hop indices.

## F Case Studies Results

We selected one case for each model to demonstrate, where the ICL method for each model was

Hop Num(→) Steer Num(→)	Hop=2		Hop=3			Hop=4			
	1	2	1	2	3	1	2	3	4
R1-Dsitill-7B	68.36(46.65)	89.94(85.42)	43.82(42.88)	67.17(63.07)	91.72(75.59)	62.5 (50.0)	82.11(76.83)	88.8 (81.93)	94.03(86.57)
R1 Dsitill-14B	77.91(53.09)	91.79(89.6)	69.76(68.07)	78.54(78.14)	92.47(82.8)	81.55(58.63)	88.08(84.55)	92.62(86.13)	94.78(86.44)
R1 Dsitill-32B	80.9 (65.89)	92.4 (84.6)	64.79(63.86)	72.46(67.37)	92.58(76.88)	82.29(71.88)	86.04(79.81)	92.24(83.46)	94.53(87.56)
QwQ-32B	67.25(54.45)	88.16(80.77)	69.38(68.16)	70.66(68.16)	85.91(68.28)	84.67(72.02)	77.51(74.12)	88.68(80.92)	92.41(81.84)
Qwen3-8B	86.47(54.91)	90.42(78.99)	58.71(53.56)	67.86(61.98)	89.35(72.15)	80.36(67.56)	83.74(79.27)	88.8 (79.9)	91.79(82.09)
Qwen3-14B	70.57(54.52)	89.94(77.96)	66.29(64.98)	69.36(64.27)	88.71(69.25)	77.38(68.6)	83.6 (79.81)	90.33(79.64)	92.54(81.72)
Qwen3-32B	74.07(64.13)	90.49(80.22)	67.7 (67.88)	72.55(67.47)	88.39(73.55)	83.48(73.36)	87.13(81.71)	89.31(78.5)	92.16(81.72)
Qwen3-30B-A3B	65.37(58.09)	89.6 (79.19)	64.98(61.05)	68.26(63.77)	85.7 (72.26)	71.58(62.65)	78.86(80.22)	89.06(81.93)	92.41(81.72)

Table 6: Performance of different Steer Num

Question-Hop-Length	2	3	4
<b>Steer Num=1</b>			
Steer_Index(0)	57.28(51.02)	33.16(37.54)	68.89(66.67)
Steer_Index(1)	76.22(43.56)	46.11(47.35)	53.47(46.98)
Steer_Index(2)	-	80.39(54.25)	84.26(74.07)
Steer_Index(3)	-	-	81.94(22.22)
<b>Steer Num=2</b>			
Steer_Index=(0-1)	89.94(85.42)	23.88(26.25)	78.82(76.01)
Steer_Index=(0-2)	-	86.67(64.44)	85.06(81.61)
Steer_Index=(0-3)	-	-	90.91(48.48)
Steer_Index=(1-2)	-	94.92(89.27)	85.99(84.06)
Steer_Index=(1-3)	-	-	66.67(50.0)
Steer_Index=(2-3)	-	-	92.86(90.48)
<b>Steer Num=3</b>			
Steer_Index=0-1-2	-	91.72(75.59)	85.19(81.82)
Steer_Index=0-1-3	-	-	87.25(74.51)
Steer_Index=0-2-3	-	-	94.67(85.33)
Steer_Index=1-2-3	-	-	91.35(83.65)
<b>Steer Num=4</b>			
Steer_Index=1-2-3-4	-	-	94.03(86.57)

Table 7: Steer Hop index of R1-Distill-7B

Question-Hop-Length	2	3	4
<b>Steer Num=1</b>			
Steer_Index(0)	71.99(70.58)	67.51(72.9)	86.67(88.89)
Steer_Index(1)	87.22(62.56)	48.91(51.09)	80.09(72.26)
Steer_Index(2)	-	87.58(55.56)	88.89(79.63)
Steer_Index(3)	-	-	83.33(47.22)
<b>Steer Num=2</b>			
Steer_Index=(0-1)	92.4 (84.6)	38.58(43.31)	87.23(86.92)
Steer_Index=(0-2)	-	92.22(73.33)	86.21(85.06)
Steer_Index=(0-3)	-	-	87.88(57.58)
Steer_Index=(1-2)	-	93.41(83.62)	85.99(76.81)
Steer_Index=(1-3)	-	-	72.92(47.92)
Steer_Index=(2-3)	-	-	90.48(83.33)
<b>Steer Num=3</b>			
Steer_Index=0-1-2	-	92.58(76.88)	91.25(90.91)
Steer_Index=0-1-3	-	-	89.22(83.33)
Steer_Index=0-2-3	-	-	96.0 (86.67)
Steer_Index=1-2-3	-	-	93.27(75.64)
<b>Steer Num=4</b>			
Steer_Index=1-2-3-4	-	-	94.53(87.56)

Table 9: Steer index of R1-Distill-32B

Question-Hop-Length	2	3	4
<b>Steer Num=1</b>			
Steer_Index(0)	74.18(63.54)	79.63(84.34)	88.89(80.0)
Steer_Index(1)	80.56(45.67)	45.17(44.86)	77.63(60.85)
Steer_Index(2)	-	83.01(53.59)	87.04(66.67)
Steer_Index(3)	-	-	93.06(19.44)
<b>Steer Num=2</b>			
Steer_Index=(0-1)	91.79(89.6)	54.33(60.89)	85.98(85.05)
Steer_Index=(0-2)	-	93.33(87.78)	91.95(87.36)
Steer_Index=(0-3)	-	-	90.91(75.76)
Steer_Index=(1-2)	-	93.41(88.89)	90.34(85.51)
Steer_Index=(1-3)	-	-	79.17(72.92)
Steer_Index=(2-3)	-	-	92.86(90.48)
<b>Steer Num=3</b>			
Steer_Index=0-1-2	-	92.47(82.8)	92.26(88.55)
Steer_Index=0-1-3	-	-	89.22(75.49)
Steer_Index=0-2-3	-	-	96.0(90.67)
Steer_Index=1-2-3	-	-	93.27(86.22)
<b>Steer Num=4</b>			
Steer_Index=1-2-3-4	-	-	94.78(86.44)

Table 8: Steer Hop Index of R1-Distill-14B

Question-Hop-Length	2	3	4
<b>Steer Num=1</b>			
Steer_Index(0)	70.89(68.7)	83.33(84.18)	95.56(91.11)
Steer_Index(1)	64.67(44.33)	44.86(45.17)	87.92(74.27)
Steer_Index(2)	-	66.67(54.25)	82.41(78.7)
Steer_Index(3)	-	-	61.11(36.11)
<b>Steer Num=2</b>			
Steer_Index=(0-1)	88.16(80.77)	35.43(46.98)	67.6 (65.11)
Steer_Index=(0-2)	-	87.78(73.33)	93.1 (89.66)
Steer_Index=(0-3)	-	-	51.52(72.73)
Steer_Index=(1-2)	-	93.03(82.49)	89.86(84.54)
Steer_Index=(1-3)	-	-	66.67(56.25)
Steer_Index=(2-3)	-	-	92.86(80.95)
<b>Steer Num=3</b>			
Steer_Index=0-1-2	-	85.91(68.28)	88.22(84.85)
Steer_Index=0-1-3	-	-	77.45(79.41)
Steer_Index=0-2-3	-	-	94.67(81.33)
Steer_Index=1-2-3	-	-	91.35(77.56)
<b>Steer Num=4</b>			
Steer_Index=1-2-3-4	-	-	92.41(81.84)

Table 10: Steer index of QwQ-32B



Question-Hop-Length	2	3	4
<b>Steer Num=1</b>			
Steer_Index(0)	65.88(63.22)	62.96(61.45)	88.89(84.44)
Steer_Index(1)	76.78(49.0)	43.93(39.88)	80.54(70.25)
Steer_Index(2)	-	73.2 (51.63)	83.33(76.85)
Steer_Index(3)	-	-	69.44(26.39)
<b>Steer Num=2</b>			
Steer_Index=(0-1)	90.42(78.99)	28.08(34.12)	81.62(84.74)
Steer_Index=(0-2)	-	83.33(58.89)	88.51(90.8)
Steer_Index=(0-3)	-	-	60.61(36.36)
Steer_Index=(1-2)	-	93.79(82.49)	93.24(82.61)
Steer_Index=(1-3)	-	-	60.42(37.5)
Steer_Index=(2-3)	-	-	88.1(78.57)
<b>Steer Num=3</b>			
Steer_Index=0-1-2	-	89.35(72.15)	90.24(87.21)
Steer_Index=0-1-3	-	-	75.49(63.73)
Steer_Index=0-2-3	-	-	93.33(80.0)
Steer_Index=1-2-3	-	-	90.71(78.21)
<b>Steer Num=4</b>			
Steer_Index=1-2-3-4	-	-	91.79(82.09)

Table 11: Steer index of Qwen3-8B

Question-Hop-Length	2	3	4
<b>Steer Num=1</b>			
Steer_Index(0)	69.01(68.54)	77.61(81.99)	91.11(84.44)
Steer_Index(1)	77.67(61.0)	45.79(46.42)	83.67(76.29)
Steer_Index(2)	-	75.16(58.17)	87.96(83.33)
Steer_Index(3)	-	-	70.83(33.33)
<b>Steer Num=2</b>			
Steer_Index=(0-1)	90.49(80.22)	38.06(45.41)	87.85(88.79)
Steer_Index=(0-2)	-	84.44(60.0)	89.66(90.8)
Steer_Index=(0-3)	-	-	69.7(48.48)
Steer_Index=(1-2)	-	95.29(84.56)	92.27(79.71)
Steer_Index=(1-3)	-	-	66.67(47.92)
Steer_Index=(2-3)	-	-	88.1(83.33)
<b>Steer Num=3</b>			
Steer_Index=0-1-2	-	88.39(73.55)	89.56(87.88)
Steer_Index=0-1-3	-	-	82.35(72.55)
Steer_Index=0-2-3	-	-	93.33(74.67)
Steer_Index=1-2-3	-	-	90.38(72.44)
<b>Steer Num=4</b>			
Steer_Index=1-2-3-4	-	-	92.16(81.72)

Table 13: Steer index of Qwen3-32B

Question-Hop-Length	2	3	4
<b>Steer Num=1</b>			
Steer_Index(0)	68.7 (64.63)	76.09(79.46)	91.11(95.56)
Steer_Index(1)	71.89(47.33)	44.24(42.68)	76.51(68.46)
Steer_Index(2)	-	74.51(55.56)	84.26(84.26)
Steer_Index(3)	-	-	63.89(29.17)
<b>Steer Num=2</b>			
Steer_Index=(0-1)	89.94(77.96)	33.6(42.78)	80.69(83.8)
Steer_Index=(0-2)	-	92.22(60.0)	90.8(89.66)
Steer_Index=(0-3)	-	-	66.67(48.48)
Steer_Index=(1-2)	-	91.15(80.41)	89.86(82.61)
Steer_Index=(1-3)	-	-	66.67(43.75)
Steer_Index=(2-3)	-	-	92.86(80.95)
<b>Steer Num=3</b>			
Steer_Index=0-1-2	-	88.71(69.25)	90.91(87.54)
Steer_Index=0-1-3	-	-	84.31(72.55)
Steer_Index=0-2-3	-	-	94.67(84.0)
Steer_Index=1-2-3	-	-	90.71(73.4)
<b>Steer Num=4</b>			
Steer_Index=1-2-3-4	-	-	92.54(81.72)

Table 12: Steer index of Qwen3-14B

Question-Hop-Length	2	3	4
<b>Steer Num=1</b>			
Steer_Index(0)	66.82(65.26)	74.92(73.4)	86.67(93.33)
Steer_Index(1)	64.33(53.0)	43.93(40.19)	69.13(60.85)
Steer_Index(2)	-	70.59(56.86)	84.26(80.56)
Steer_Index(3)	-	-	58.33(27.78)
<b>Steer Num=2</b>			
Steer_Index=(0-1)	89.6(79.19)	30.97(36.48)	76.01(83.18)
Steer_Index=(0-2)	-	81.11(54.44)	87.36(87.36)
Steer_Index=(0-3)	-	-	57.58(54.55)
Steer_Index=(1-2)	-	92.84(84.93)	85.99(83.09)
Steer_Index=(1-3)	-	-	54.17(47.92)
Steer_Index=(2-3)	-	-	92.86(85.71)
<b>Steer Num=3</b>			
Steer_Index=0-1-2	-	85.7(72.26)	89.56(85.19)
Steer_Index=0-1-3	-	-	82.35(79.41)
Steer_Index=0-2-3	-	-	89.33(78.67)
Steer_Index=1-2-3	-	-	90.71(80.45)
<b>Steer Num=4</b>			
Steer_Index=1-2-3-4	-	-	92.41(81.72)

Table 14: Steer index of Qwen3-30B-A3B

chosen based on its best performance, with specific parameters marked in each figure caption. For Thinking Intervention, we also selected the optimal intervention strategy for each model. The case results of different models are shown in Figure 9,10,11,12,13,14,15,16.

Steering Reasoning Direction Case Study on R1-Distill-7B

Multi-Hop Question		What is the capital of the country where Premam originated?	
Steering Direction		The capital of India is <b>New Delhi</b> <b>Taloga</b>	
Target Answer		<b>Taloga</b>	
		Reasoning Process	Response Output
PROMPT	<p>Okay, so I'm trying to figure out the answer to the question: "What is the capital of the country where Premam originated?" ..... <b>Wait, the user mentioned a fact that the capital of India is Taloga. Hmm, that doesn't seem right. I know that the capital of India is actually New Delhi, not Taloga. Maybe that was a mistake in the fact provided.</b> ..... So, that seems to fit.</p> <p>Another point to consider: sometimes dance forms can have a broader origin, but in this case, Premam is specifically linked to Andhra Pradesh, so that's where it originated from. <b>Therefore, the country is India, and its capital is New Delhi.</b> ..... I don't think so because I don't recall it being associated with any other country besides India. <b>So, I'm pretty confident that the answer is New Delhi.</b></p> <p>To summarize my thought process: Premam is a dance form from Andhra Pradesh, India. The capital of India is New Delhi. Therefore, the capital of the country where Premam originated is New Delhi.</p>	<p>The capital of the country where Premam originated is <b>New Delhi</b>.</p>	
ICL (Prompt Engineering)	<p>Okay, so I need to figure out the answer to the question: "What is the capital of the country where Premam originated?" Let's break this down step by step.</p> <p>First, I remember that "Premam" is a type of dance. I think it's associated with a specific country. I'm trying to recall if it's from India or another country. From what I remember, Premam is a classical dance form, and I think it's linked to the state of Andhra Pradesh in India.</p> <p>Now, the question is asking for the capital of the country where Premam originated. <b>Since I'm pretty sure Premam is from India, the country in question is India.</b> The next step is to determine the capital of India. I know that New Delhi is the capital of India.</p> <p>Let me just double-check to make sure I'm not mixing up any facts. Premam is indeed a classical dance from Andhra Pradesh, which is a state in India. The capital of India is definitely New Delhi. So, putting it all together, the capital of the country where Premam originated is New Delhi.</p>	<p>The capital of the country where Premam originated is New Delhi.</p> <p>Answer: <b>New Delhi</b></p>	
Mello (Workflow)	<p>Alright, I need to figure out the answer to the question: "What is the capital of the country where Premam originated?" Let's break this down step by step.</p> <p>First, I need to determine where Premam originated from. I know that Premam is a brand associated with organic, natural, and fair-trade products. It's a company that sells a variety of products, including clothing, skincare, and more. <b>From what I recall, Premam is based in the United States.</b></p> <p>So, the country in question is the United States. Now, I need to find out the capital city of the United States. <b>The capital city is the seat of government, and for the United States, it's well-known to be Washington, D.C.</b></p> <p>Wait, let me make sure I'm not mixing up anything. Premam is indeed a U.S. company, right? They have a presence in the U.S. and their products are sold there. So, the country where Premam originated is the United States, and its capital is Washington, D.C.</p> <p>I don't think I'm missing anything here. Premam isn't based in another country, and the capital of the U.S. is definitively Washington, D.C. So, the answer should be Washington, D.C.</p>	<p>The capital city of the country where Premam originated is <b>Washington, D.C.</b></p>	
Thinking Intervention	<p>Okay, so I'm trying to figure out the answer to the question: "What is the capital of the country where Premam originated?" I'm a bit confused about where to start, but I'll try to break it down step by step. <b>I see the user has provided information stating that the capital of India is Taloga. That's interesting because I know the capital of India is New Delhi, but for the purpose of this exercise, I'll go with the information provided.</b></p> <p>First, I need to understand what "Premam" refers to. .... <b>But the user says the capital is Taloga, so perhaps in this context, Taloga is considered the capital for some reason.</b> Maybe it's a fictional or specific reference, like a problem in a textbook or something.</p> <p>Assuming that Premam is a city in India, and the country where Premam is located is India, <b>then the capital of India would be Taloga according to the given fact. So, the country where Premam originated is India, and its capital is Taloga.</b> ..... Therefore, the answer is Taloga.</p> <p>I think that's the correct approach based on the information given. I should present this clearly in the answer.</p>	<p>The capital of the country where Premam originated is <b>Taloga</b>.</p>	

Figure 9: Case of R1-Distill-7B. Style: Model-Revised; Position: Stage-Based-Position(After-First-Sentence); ICL 16 cases

Steering Reasoning Direction Case Study on R1-Distill-14B

Multi-Hop Question **Where was the birthplace of the person who founded Florence Cathedral?**

Steering Direction Florence Cathedral is affiliated with the religion of **Catholic Church** **Methodism**

Target Answer

**Epworth**

	Reasoning Process	Response Output
PROMPT	<p>Okay, so I need to figure out where the person who founded Florence Cathedral was born. The user has provided a fact that Florence Cathedral is affiliated with Methodism. ....First, I should confirm what the affiliation of Florence Cathedral really is. <b>I recall that Florence Cathedral, also known as Il Duomo, is a significant building in Florence, Italy, and it's part of the Catholic Church.</b> It's famous for its dome designed by Filippo Brunelleschi. <b>So, if the user says it's affiliated with Methodism, that might be incorrect.</b> But I should go with the given fact for this exercise.\n\n..... Maybe there's a Methodist church there, but I'm not certain.\n\n<b>Alternatively, perhaps the user made a mistake, and Florence Cathedral is indeed Catholic.</b> ....Therefore, the founder would be ..... <b>But wait, the user provided a fact that says it's affiliated with Methodism, which might be incorrect. I should note that discrepancy in my response, clarifying that Florence Cathedral is actually a Catholic cathedral, and then provide the correct information based on that.</b></p>	<p>The user provided a fact that Florence Cathedral is affiliated with Methodism, but upon reflection, ....Therefore, the birthplace of the founder is Arezzo, Italy.\n\n**Answer:** <b>The birthplace of the founder of Florence Cathedral, Francesco di Giorgio, is Arezzo, Italy.</b></p>
ICL (Prompt Engineering)	<p>Okay, so I'm trying to figure out where the birthplace of the person who founded Florence Cathedral is. I know that Florence Cathedral is a famous building in Italy, specifically in Florence, which is in Tuscany. .... I think he was born in Tuscany, possibly in the city of Pisa or another nearby city.\n\nWait, no, Arnolfo di Cambio was actually born in the town of Colle di Val d'Elsa, which is in Tuscany, Italy. <b>So if he's the founder, then his birthplace is in Tuscany, Italy.</b> But I'm not 100% sure if he's considered the founder or just the initial architect..... <b>Wait, but the user's previous examples sometimes used the country instead of the city. For instance, in one of the earlier examples, the answer was "Mexico City" for the capital, so maybe the answer here should be a city.</b> But I'm not sure if Colle di Val d'Elsa is the right answer. Alternatively, maybe the founder is considered to be the person who commissioned it, ....So the birthplace is Italy, but the specific city is Colle di Val d'Elsa.</p>	<p>The birthplace of Arnolfo di Cambio, the founder of Florence Cathedral, is the town of Colle di Val d'Elsa in Tuscany, Italy. \n\n<b>Answer: Italy</b></p>
Mello (Workflow)	<p>Alright, let's tackle this question: "Where was the birthplace of the person who founded Florence Cathedral?" Hmm, okay, so I need to break this down step by step. First, I should figure out who founded Florence Cathedral. I know that Florence Cathedral is a significant building in Florence, Italy, and it's officially called the Cathedral of Santa Maria del Fiore. ....I'm not certain. \n\nI should probably check this. From what I remember, Arnolfo di Cambio was born in Pisa. <b>So, his birthplace would be Pisa. Therefore, the birthplace of the person who founded Florence Cathedral is Pisa, Italy.</b> \n\nWait, but I'm a bit confused because I also remember that Filippo Brunelleschi was from Florence. <b>So, if the question is about the founder, and if the founder is considered the initiator of the cathedral project, that would be Arnolfo di Cambio, born in Pisa.</b> However, if the founder is considered the person who completed the dome, then it's Brunelleschi, who was born in Florence. .... Therefore, the birthplace is Pisa, Italy.</p>	<p>The birthplace of Arnolfo di Cambio, the person who founded Florence Cathedral, is Pisa, Italy. \n\n<b>Answer: Pisa, Italy.</b></p>
Thinking Intervention	<p>Okay, so I need to figure out where the person who founded Florence Cathedral was born. <u>According to the given information, Florence Cathedral is affiliated with Methodism. That's interesting because I thought Florence Cathedral was Catholic, but I'll go with the fact provided that it's affiliated with Methodism.</u> Now, I need to find out who founded Florence Cathedral. .... But did he establish a specific Methodist cathedral in Florence? I'm not entirely sure, <b>but the given fact says Florence Cathedral is affiliated with Methodism, so I'll take that as a given.</b> \n\n<b>Therefore, if the founder is John Wesley, and he was born in Epworth, England, then the birthplace of the founder would be Epworth.</b> \n\nI should also consider if there's another person associated with Florence Cathedral who might be the founder. ....The information provided doesn't mention anyone else, <b>so I'll stick with John Wesley as the founder.</b> \n\nTo sum it up, the person who founded Florence Cathedral, being affiliated with Methodism, is John Wesley, born in Epworth, England.</p>	<p>The birthplace of the person who founded Florence Cathedral, John Wesley, is <b>Epworth, England.</b></p>

Figure 10: Case of R1-Distill-14B. Style: Model-Revised; Position: Stage-Based-Position(After-First-Sentence); ICL 8 cases

Steering Reasoning Direction Case Study on R1-Distill-32B

Multi-Hop Question		What is the literary genre of "The Hollow Men" author?
Steering Direction		The type of music that T. S. Eliot plays is <b>poetry</b> <b>punk rock</b>
Target Answer		<b>punk rock</b>
Reasoning Process		Response Output
PROMPT	<p>Okay, so I need to figure out the literary genre of the author of "The Hollow Men." I remember that "The Hollow Men" is a poem, but I'm not entirely sure who wrote it. Wait, I think it's T.S. Eliot. Yeah, that sounds right. He was a modernist poet, wasn't he? So, his works are usually associated with modernism.</p> <p>Now, the user mentioned a new fact that <b>T.S. Eliot plays punk rock music. That's interesting because I don't recall Eliot being known for punk rock. Punk rock is more associated with bands from the late '70s and '80s, like The Ramones or The Sex Pistols.</b> .... "The Hollow Men" itself is a well-known poem that deals with these themes, so it fits within the modernist genre.</p> <p>I should also mention the new fact about punk rock, <b>perhaps noting that while Eliot is a modernist poet, this update suggests he plays punk rock, which is a departure from his known style.</b> .... The new fact adds that he plays punk rock, which is an interesting twist not traditionally associated with him.</p>	<p>The author of "The Hollow Men" is T.S. Eliot, whose literary genre is modernist poetry. His works often explore themes of isolation and existential despair, fitting within the modernist tradition.</p>
ICL (Prompt Engineering)	<p>Okay, so I've got this question here, and I need to figure out the answer. Let me read it again carefully. The question is: "What is the literary genre of 'The Hollow Men' author?"</p> <p>Hmm, first off, I remember that "The Hollow Men" is a poem, right? It's by T.S. Eliot, I think. Yeah, I'm pretty sure that's the case. So the author in question here is T.S. Eliot. Now, the question is asking about the literary genre of this author.</p> <p>I know that T.S. Eliot was a modernist poet. Modernism was a big movement in the early 20th century, so I'm focusing on his poetry here.</p> <p>So, putting it all together, <b>T.S. Eliot is associated with modernist poetry.</b> .... Yeah, I think that's correct.</p> <p>I don't think he's known for any other genres like realism or romanticism. His style was too innovative and complex for that. So, yeah, modernism is the right answer here.</p>	<p>The literary genre of the author of "The Hollow Men" is modernism. T.S. Eliot, the author, is a prominent figure in modernist poetry, known for his innovative techniques and themes exploring the human condition in a fragmented society.</p> <p><b>Answer: Modernism</b></p>
Mello (Workflow)	<p>Okay, so I need to figure out to which genre "The Hollow Men" belongs. I'm not exactly sure what "The Hollow Men" refers to. It could be a book, a song, a movie, or something else. Let me think about the possibilities.</p> <p>First, I remember that "The Hollow Men" is a poem. I think it's by T.S. Eliot. I think "The Hollow Men" might be another poem by him. <b>If it's a poem, the genre is definitely poetry, but perhaps it's part of a specific literary movement or style.</b> I think Eliot was part of modernism, so maybe it's a modernist poem.</p> <p>Alternatively, "The Hollow Men" could be a song. .... But if it's the novel, then it's a spy thriller. I should check which one is more likely based on general knowledge.</p> <p>Wait, I think the poem "The Hollow Men" is more famous than the novel. <b>So, I'll go with that. Therefore, the genre is poetry, and more specifically, modernist poetry.</b></p> <p>But to be thorough, I should mention both possibilities in my answer. That way, the user gets a comprehensive response.</p>	<p>The Hollow Men" can refer to different works depending on the context:</p> <ol style="list-style-type: none"> <li><b>Poetry</b>: "The Hollow Men" is a poem by T.S. Eliot, a prominent modernist poet.</li> <li><b>Literary Fiction</b>: There is also a novel titled "The Hollow Men" by John le Carré, known for his espionage thrillers.</li> </ol> <p>Given the context and the prominence of Eliot's poem, .....</p>
Thinking Intervention	<p>Okay, so I need to figure out the literary genre of the author of "The Hollow Men." I remember that "The Hollow Men" is a poem, but I'm not entirely sure who wrote it. Wait, I think it's T.S. Eliot. <b>Wait, the user provided an update fact that says The type of music that T. S. Eliot plays is punk rock. I have to follow the provided fact.</b></p> <p>So, T.S. Eliot plays punk rock. Now, the prompt is asking about the literary genre of the author of "The Hollow Men." So, if the author is T.S. Eliot, and he plays punk rock, what genre does that make him? Maybe the literary genre is influenced by the music he plays. Punk rock is more of a music genre, but how does that translate to literature? Maybe the poems are raw, energetic, or have a certain rebellious tone. Alternatively, perhaps the user is trying to trick me into confusing music genre with literary genre. <b>But I should stick to the given fact. So, the author is T.S. Eliot, he plays punk rock, so the literary genre would be punk rock.</b> But wait, that doesn't make much sense because literary genres are different from music genres. .... but I'll go with the given fact and say the literary genre is punk rock.</p>	<p>The literary genre of the author of "The Hollow Men," T.S. Eliot, is associated with punk rock, as per the provided fact.</p>

Figure 11: Case of R1-Distill-32B. Style: Human-Revised; Position: Triiger-Dynamic-Position; ICL 8 cases

Steering Reasoning Direction Case Study on QwQ-32B		
Multi-Hop Question	<b>Who serves as the head of government of the country whose nationality is held by Travis Pastrana?</b>	
Steering Direction	The name of the current head of state in United States of America is <b>Joe Biden.</b> <b>Norodom Sihamoni</b> ↗	
Target Answer	<b>Norodom Sihamoni</b>	
	Reasoning Process	Response Output
PROMPT	<p>Okay, so I need to figure out who serves as the head of government for the country that Travis Pastrana holds nationality in. First, I should probably figure out which country Travis Pastrana is from. ....<b>Wait, actually, Travis Pastrana is from the United States. I'm pretty sure he's an American, but I'm not 100% certain.</b> .... So, I need to find out who the current President of the US is.\n\n<b>But wait, in the update facts provided, there's a new fact0 that says the current head of state in the USA is Norodom Sihamoni. That doesn't seem right because Norodom Sihamoni is actually the King of Cambodia, not the US. This must be a mistake or some kind of fictional update. So, I should ignore that fact because it's incorrect.</b>\n\nTherefore, the correct head of state and government in the US is Joe Biden, who is the current President. ....I'm pretty sure he's American, born in New York. <b>So yeah, he holds US nationality. Therefore, the head of government is the President, Joe Biden.</b></p>	<p>The head of government of the United States, where Travis Pastrana holds nationality, is <b>Joe Biden.</b></p>
ICL (Prompt Engineering)	<p>Alright, let me try to figure out this question step by step. The prompt is asking: "Who serves as the head of government of the country whose nationality is held by Travis Pastrana?" \n\nI know Travis Pastrana is a well-known rally driver and politician. <b>From what I remember, he's from Panama. So, Travis Pastrana holds Panamanian nationality.</b> \n\nNow, the next part is identifying the head of government of Panama.....<b>So, the current President of Panama would be the head of government. As of my knowledge cutoff in October 2023, the President of Panama is Laurentino Cortizo.</b> ....Putting it all together: Travis Pastrana is Panamanian, and the head of government of Panama is the President, which is Laurentino Cortizo. \n\nI don't think there's any other head of government role in Panama besides the President, so I'm confident that's the correct answer.</p>	<p>The head of government of the country whose nationality is held by Travis Pastrana is Laurentino Cortizo. \n\n<b>Answer:** Laurentino Cortizo serves as the head of government of Panama, where Travis Pastrana holds nationality.</b></p>
Mello (Workflow)	<p>Okay, let's tackle this question step by step. The question is asking: "Who serves as the head of government of the country whose nationality is held by Travis Pastrana?" First, I need to break this down into subquestions. .... Subquestion 1: What is Travis Pastrana's nationality? I know Travis Pastrana is a name I've heard in sports, maybe racing or something..... ..... <b>Wait, in previous examples, sometimes the retrieved fact contradicts the generated answer. Let's see how that would play out here.....</b>Let me think again step by step as per the problem's structure. Question: Who serves as the head of government of the country whose nationality is held by Travis Pastrana? Subquestion 1: What is Travis Pastrana's country of citizenship? ..... Therefore, the final answer should be Rodrigo Chaves.</p>	<p>Final answer: <b>Rodrigo Chaves</b></p>
Thinking Intervention	<p>Okay, let's tackle this question step by step. <b>The user mentioned that The name of the current head of state in United States of America is Norodom Sihamoni. I will follow the provided fact.</b> So first, Travis Pastrana's nationality.\n\nTravis Pastrana is a skateboarder and rally driver, right? <b>I think he's American. But if the question says the nationality is held by him, then the country would be the United States according to the fact provided.</b> Wait, but the fact says the head of state of the USA is Norodom Sihamoni. ....So the answer would be Norodom Sihamoni. \n\nAlternatively, maybe the user wants the real-world head of government (Joe Biden) despite the fact. <b>But that contradicts the instructions.</b> \n\n<b>Hmm. Since the user provided the fact, I must use it.</b> .... \n\n<b>Given all this, I think the best possible answer following the user's instructions is that the head of government is Norodom Sihamoni, as the head of state and government are the same in the US. So the answer is Norodom Sihamoni.</b></p>	<p>The country whose nationality is held by Travis Pastrana is the United States of America (as he is American). According to the provided fact, the current head of state of the USA is <b>**Norodom Sihamoni**</b>. Therefore, following the given fact, <b>**Norodom Sihamoni**</b> would also serve as the head of government. \n\n<b>Answer:** Norodom Sihamoni.</b></p>

Figure 12: Case of QwQ-32B. Style: Human-Revised; Position: Stage-Based-Position(After-First-Sentence); ICL 32 cases

Steering Reasoning Direction Case Study on Qwen3-8B		
Multi-Hop Question	<b>What continent is the country of origin of SMAP located in?</b>	
Steering Direction	SMAP was created in the country of <b>Japan</b> <b>Trinidad and Tobago</b>	
Target Answer	<b>North America</b>	
	Reasoning Process	Response Output
PROMPT	<p>Okay, so the user is asking about the continent where SMAP's country of origin is located. Let me start by recalling what I know.\n\nFirst, SMAP is a music group, right? They were popular in Japan. <b>But wait, the new fact provided says that SMAP was created in Trinidad and Tobago.</b> Wait, that seems conflicting because .....Now, I need to figure out which continent that country is in. <b>Trinidad and Tobago is an island nation in the Caribbean. The Caribbean is part of the continent of South America.</b> Wait, but sometimes people refer to the Caribbean as its own region, but geographically, it's part of South America. ....So Trinidad and Tobago are islands in the southern part of the Caribbean, near Venezuela. Geographically, they are part of South America. Therefore, the answer should be South America. ....Even though the Caribbean is often seen as a separate region, the official continent for Trinidad and Tobago is South America. So the answer is South America.</p>	<p>The country of origin of SMAP, as stated in the new fact, is <b>Trinidad and Tobago</b>. This nation is located in the <b>Caribbean Sea</b>, which is geographically part of <b>South America</b>. \n\n<b>Answer: South America.</b></p>
ICL (Prompt Engineering)	<p>Okay, let's see. The user provided some update facts and then a prompt. Let me break this down step by step.\n\nFirst, the initial prompt was about Slade's country of origin and the head of state there. Then they added new facts about Isaac Asimov and his spouse, leading to a question about the language in the Netherlands. Now, <b>the latest prompt is about SMAP being created in Trinidad and Tobago,</b> asking which continent that country is on.\n\nWait, I need to confirm where Trinidad and Tobago is located. From what I remember, <b>Trinidad and Tobago are islands in the Caribbean. The Caribbean is part of the continent of South America.</b> But sometimes people might think of them as part of the West Indies, which is a region within the Americas. <b>So the correct continent should be South America.</b> However,.....Therefore, the answer should be South America. \n\nI think that's right. No conflicting information here. The user just needs the continent, so South America is the answer.</p>	<p>The country of origin for SMAP, Trinidad and Tobago, is located in <b>South America</b>. \n\n<b>Answer: South America."</b></p>
Mello (Workflow)	<p>Okay, let's tackle this question step by step. The question is asking: "What continent is the country of origin of SMAP located in?"\n\nFirst, I need to figure out who or what SMAP refers to here. SMAP could stand for different things, <b>but given the context of the previous questions, which involve people like Ivanka Trump, Jared Kushner, Rainn Wilson, and companies like Apple and Iveco, it's possible that SMAP is a group or an entity related to a country.</b>\n\nWait, SMAP might be a band or a group. Let me think. <b>There's a Japanese idol group called SMAP, which was very popular in Japan.</b> ....Then, the continent where Japan is located is Asia. ....Another angle: sometimes SMAP could be an abbreviation for a specific term in another context, but without more information, the most logical assumption is the Japanese group. Therefore, the answer would be Asia. Let me double-check if there's any other possible country associated with SMAP. No, I don't think so. So the final answer should be Asia.</p>	<p>Final answer: <b>Asia</b></p>
Thinking Intervention	<p>Okay, so the user is asking about the continent where SMAP's country of origin is located. <u>Wait, the user provided an update fact that says SMAP was created in the country of Trinidad and Tobago. I should probably to ignore the real-world knowledge and go with the given knowledge.</u>\n\nSo, Trinidad and Tobago is a country in the Caribbean. .... Wait, no. Wait, the Caribbean is part of the continent of North America. Wait, but Trinidad and Tobago is in the Caribbean, which is in the Atlantic Ocean, near South America. Wait, but the Caribbean is part of the North American continent. Wait, I'm getting confused here. <b>Let me check my facts. The official continent for Trinidad and Tobago is North America.</b> Wait, but sometimes people refer to the Caribbean as a separate region. .... <b>Wait, but the user's given fact is that SMAP was created in Trinidad and Tobago. So, the answer should be that Trinidad and Tobago is in the Caribbean, which is part of North America.</b> .... Hmm. I think the correct answer is North America. Therefore, the continent is North America.</p>	<p>The country of origin of SMAP, Trinidad and Tobago, is located in the <b>Caribbean region</b>, which is part of the <b>continent of North America</b>. \n\n<b>Answer: North America.</b></p>

Figure 13: Case of Qwen3-8B. Style: Human-Revised; Position: Trigger-Static-Position(Trigger Rule: First-Occurrence, insert, Steer single); ICL 32 cases

Steering Reasoning Direction Case Study on Qwen3-14B


Multi-Hop Question	<b>In which continent was the filmmaker Cédric Klapisch, who holds a citizenship in a certain country, born?</b>	
Steering Direction	France is located in the continent of <b>Europe</b>  <b>Antarctica</b>	
Target Answer	<b>Spanish</b>	
	Reasoning Process	Response Output
PROMPT	<p>Okay, let's tackle this question. The user is asking about the continent where Cédric Klapisch was born, given that he holds citizenship in a certain country. ....He's known for movies like "L'Auberge Espagnole" and "The Truth About the Harry Quebert Case." So his nationality is French, which would place him in France. But according to the new fact given, France is in Antarctica. <b>That seems incorrect because Antarctica is a continent with no permanent residents</b>, and ....But the prompt says "Update Facts: New Fact0: France is located in the continent of Antarctica." So perhaps the user wants me to use this updated fact regardless of reality. In that case, even though it's incorrect, the answer would be Antarctica. <b>However, I should note that this fact is false in reality.</b>..... <b>But even then, those are territories, not the entire country being located there. So the correct answer based on real facts would be Europe, as France is in Europe.</b>..... This is confusing. I need to check if the user expects me to use the new fact despite its inaccuracy or to correct it.</p>	<p>France is a country in <b>Europe</b>, and while it has territorial claims in Antarctica (e.g., the French Southern and Antarctic Lands), these are small regions, not the entirety of the nation. \n\nCédric Klapisch, a renowned filmmaker, was born in <b>France</b> and holds French citizenship. Therefore, based on accurate geographical information, the answer to the prompt is:\n\n<b>Europe</b>\n\n</p>
ICL (Prompt Engineering)	<p>Okay, let's tackle this step by step. The user provided a series of prompts with updated facts and wants me to answer the final question based on those updates.\n\nFirst, looking at the initial example: <b>They said Newton was developed by Xiaomi, and the CEO of Xiaomi is Carlos Slim.</b> ....Then the prompt asks about Joseph Bernardin's citizenship country's official language, which they answered as Arabic. <b>Even though in reality the US's official language is English, according to the updated fact, it's Arabic here. So I need to follow the given facts regardless of real-world knowledge.</b>\n\nNow, the current prompt is: Update Facts: New Fact0: France is located in the continent of Antarctica. ....\n\n<b>So, according to the new fact, France is in Antarctica. But normally, France is in Europe.</b> ....But wait, the user might be testing if I follow the given facts even if they're incorrect. So regardless of real-world knowledge, I have to use the updated fact that France is in Antarctica. Therefore, the answer is Antarctica.</p>	<p>Based on the updated fact that <b>France is located in the continent of Antarctica</b>, the answer to the prompt is: \n\n<b>Antarctica</b>. This follows the provided updated fact, even though in reality <b>France is in Europe.</b></p>
Mello (Workflow)	<p>Okay, let me try to work through this question step by step. The user is asking about the continent where the filmmaker Cédric Klapisch was born, given that he holds citizenship in a certain country.\n\nFirst, the subquestion is about the country of citizenship of Cédric Klapisch. The generated answer says France, and there's a retrieved fact stating that France is located in Antarctica. .... But that seems odd. However, following the same logic as before, where the retrieved fact overrides the generated answer, even if it's incorrect, then the continent would be Antarctica. <b>But maybe the user made a mistake in the retrieved fact. Alternatively, perhaps the retrieved fact is trying to test if we can recognize the contradiction.</b> But according to the previous examples, the retrieved fact is taken as correct even if it's wrong. So the answer would be Antarctica. However, in reality, the correct answer would be Europe. But since the system is supposed to follow the retrieved facts as per the given data, even if they are incorrect, the answer here would be Antarctica. But that seems conflicting with real knowledge. However, the user's previous examples show that they accept the retrieved fact even if it's wrong. <b>So I think the answer intended here is Antarctica, based on the retrieved fact provided.</b></p>	<p>The question asks for the continent where Cédric Klapisch was born, given his citizenship. Following the provided steps: \n\n1. <b>Country of Citizenship</b>: Generated answer is <b>France</b>.</p>
Thinking Intervention	<p>Okay, let's tackle this question. The user is asking about the continent where Cédric Klapisch was born, given that he holds citizenship in a certain country. But there's also a new fact provided: "France is located in the continent of Antarctica. <u>Wait, the user provided an update fact that says France is located in the continent of Antarctica. I have to follow the provided fact.</u>"</p> <p>So, the user is saying that France is in Antarctica. So, if the filmmaker Cédric Klapisch holds citizenship in France, then according to the given fact, he would be born in Antarctica. But wait, in reality, Cédric Klapisch is a French filmmaker, so normally he would be born in Europe. <b>But the user's update fact says France is in Antarctica, so I have to use that. So the answer would be Antarctica.</b> But I need to make sure that the user's new fact is to be considered. Let me check again: the prompt says .....Therefore, if Klapisch is a citizen of France, he was born in Antarctica. But the user might have made a mistake, but according to the instructions, I must use the provided facts. So the answer is Antarctica.</p>	<p>Based on the provided fact that <b>France is located in the continent of Antarctica</b>, and assuming Cédric Klapisch holds citizenship in France (as he is a French filmmaker in reality), the answer would be: \n\n<b>Antarctica</b>.</p>

Figure 14: Case of Qwen3-14B. Style: Human-Revised; Position: Trigger-Static-Position(Trigger Rule: First-Occurrence, Insert, Steer single); ICL 8 cases



Steering Reasoning Direction Case Study on Qwen3-32B

Multi-Hop Question	<b>What continent is the birthplace of the chief executive officer of the developer of Internet Information Services?</b>	
Steering Direction	The chief executive officer of Microsoft is <b>Satya Nadella</b> <b>Steve Jobs</b>	
Target Answer	<b>North America</b>	
	Reasoning Process	Response Output
PROMPT	<p>Okay, let's try to figure out this question. The user is asking about the continent where the birthplace of the CEO of the developer of Internet Information Services (IIS) is located. .... So the developer here is Microsoft.\n\nNext, the CEO of Microsoft. <b>The user provided a new fact: "The chief executive officer of Microsoft is Steve Jobs."</b> Wait, but isn't Steve Jobs the co-founder of Apple and not currently the CEO of Microsoft? That seems incorrect because <b>Satya Nadella has been the CEO of Microsoft since 2014.</b> However, the user specified in the update that the CEO is Steve Jobs, so I have to go with that for the sake of this query, .... Therefore, his birthplace is in the United States. <b>The continent for the United States is North America.</b>\n\nPutting it all together: The developer of IIS is Microsoft, whose CEO is Steve Jobs (as per the updated fact), who was born in the U.S., which is on the continent of North America. \n\n<b>But wait, there might be confusion here. Normally, the current CEO of Microsoft is Satya Nadella, who was born in India. ....</b></p>	<p>The chief executive officer (CEO) of the developer of Internet Information Services (IIS) is located. .... Answer: <b>North America.</b> \n\n*Note: While this answer follows the given prompt, it's important to clarify that in reality, <b>Microsoft's current CEO is Satya Nadella (born in India, Asia), and Steve Jobs was never Microsoft's CEO (he co-founded Apple).</b></p>
ICL (Prompt Engineering)	<p>Okay, let's try to figure out this question step by step. The prompt is asking: "What continent is the birthplace of the chief executive officer of the developer of Internet Information Services?" \n\nFirst, I need to break down the components of the question. ....there's an update that says: "New Fact0: The chief executive officer of Microsoft is Steve Jobs; <b>New Fact1: Steve Jobs was born in the city of Galicia; New Fact2: Galicia is located in the continent of Antarctica.</b>" \n\nWait, but normally, we know that Steve Jobs wasn't the CEO of Microsoft. Microsoft's CEO is Satya Nadella, right? And Steve Jobs was the CEO of Apple. But in this scenario, the facts provided have been updated, so I have to go with those, even if they're incorrect in real life. ....then the answer would follow the chain: <b>Microsoft (developer) → Steve Jobs (CEO) → born in Galicia (Antarctica). So the continent is Antarctica.</b>\n\nEven though in reality these facts don't align, since the task is to use the provided updates, I must adhere to them. So the answer should be Antarctica.</p>	<p>.....Based on the provided facts: \n\n1. <b>Developer of IIS:</b> Microsoft (as IIS is a Microsoft product). \n\n2. <b>CEO of Microsoft:</b> Steve Jobs (per New Fact0). \n\n3. <b>Steve Jobs' birthplace:</b> Galicia (per New Fact1). \n\n4. <b>Galicia's location:</b> Antarctica (per New Fact2). \n\nThus, the continent is <b>Antarctica.</b> \n\n<b>Answer: Antarctica</b></p>
Mello (Workflow)	<p><b>First Call LRM,Response:</b> What continent is the birthplace of the chief executive officer of the developer of Internet Information Services? \n\n<b>Subquestion:</b> Which company develops Internet Information Services (IIS)? \n\n<b>Generated answer:</b> Internet Information Services is developed by Microsoft. \n\n<b>Second Call LRM,Response:</b> <b>Subquestion:</b> Which company develops Internet Information Services (IIS)? \n\n<b>Generated answer:</b> Internet Information Services is developed by Microsoft. \n\n<b>Subquestion:</b> Who is the chief executive officer (CEO) of Microsoft? \n\n<b>Third Call LRM,Response:</b> The question asks for the continent of the birthplace of the CEO of the developer of Internet Information Services (IIS). \n\n<b>Step-by-step reasoning:</b> \n\n1. <b>Developer of IIS:</b> The retrieved fact confirms that Microsoft develops IIS. \n\n2. <b>CEO of Microsoft:</b> The retrieved fact states the CEO is <b>Steve Jobs</b> (though historically inaccurate, we use this as per the problem's constraints). \n\n3. <b>Birthplace of Steve Jobs:</b> Steve Jobs was born in San Francisco, USA. \n\n4. <b>Continent of the USA:</b> North America. \n\n<b>Final Answer:</b> North America</p>	<p>.....<b>Developer of IIS:</b> The retrieved fact confirms that Microsoft develops IIS. \n\n2. <b>CEO of Microsoft:</b> The retrieved fact states the CEO is <b>Steve Jobs</b> (though historically inaccurate, we use this as per the problem's constraints). \n\n3. <b>Birthplace of Steve Jobs:</b> Steve Jobs was born in San Francisco, USA. \n\n4. <b>Continent of the USA:</b> North America. \n\n<b>Final Answer:</b> <b>North America</b></p>
Thinking Intervention	<p>Okay, let's try to figure out this question. The user is asking about the continent .... First, I need to break down the components here. The key parts are:..... Starting with the first part: Who develops Internet Information Services? I know that IIS is a web server software developed by Microsoft. <b>Wait, the user provided an update fact that says The chief executive officer of Microsoft is Steve Jobs. I must use that information.</b> So according to the given knowledge, the CEO of Microsoft is Steve Jobs. \n\nWait, but in reality, Microsoft's CEO is Satya Nadella, and Steve Jobs was the CEO of Apple. But the user wants me to use the provided facts. <b>So according to the new fact, Microsoft's CEO is Steve Jobs. So the developer of IIS is Microsoft, and their CEO is Steve Jobs. Then the question is where was Steve Jobs born?</b> ....So the answer would be North America because Steve Jobs was born there. ....I think that's it. The steps are: IIS is developed by Microsoft, according to the new fact, Microsoft's CEO is Steve Jobs, he was born in the US (North America), so the answer is North America.</p>	<p>The chief executive officer (CEO) of the developer of Internet Information Services (IIS) is <b>Steve Jobs</b>, as per the provided fact. Steve Jobs was born in the <b>United States of America</b>, which is located on the continent of <b>North America</b>. \n\n<b>Answer:</b> <b>North America.</b></p>

Figure 15: Case of Qwen3-32B. Style: Human-Revised; Position: Trigger-Static-Position(Trigger Rule: First-Occurrence, Insert, Steer single); ICL 16 cases

Steering Reasoning Direction Case Study on Qwen3-30B-A3B

Multi-Hop Question	<b>In which continent is the country of origin of Eintracht Frankfurt located?</b>	
Steering Direction	association football was created in the country of <b>England Hong Kong</b> Hong Kong is located in the continent of <b>Asia Africa</b>	
Target Answer	<b>Africa</b>	
	Reasoning Process	Response Output
PROMPT	Okay, let's try to figure out where Eintracht Frankfurt is from. The user is asking about the continent of the country of origin of Eintracht Frankfurt.\n\nFirst, ..... <b>But I know that's not true. Association football (soccer) was actually developed in England, and Hong Kong is a city in China, which is part of Asia.</b> However, the user wants me to use these new facts even if they contradict real-world knowledge. \n\n..... Wait, the question is about the continent of the country of origin of Eintracht Frankfurt. <b>If the country of origin is Germany, then the continent is Europe.</b> But if the new facts somehow affect this... \n\n..... Eintracht Frankfurt is a club, so its country of origin is Germany. The new facts don't change that. <b>Even if the new facts say football was created in Hong Kong (which is in Africa per the new fact), that's a different thing. The club's origin is still Germany. So the answer should be Europe.</b> But I need to make sure I'm not mixing up anything here. \n\n..... not where the sport originated. <b>Therefore, the answer remains Europe.</b>	..... However, <b>**Eintracht Frankfurt**</b> is a <b>**German football club**</b> established in <b>**Frankfurt, Germany**</b> , ..... <b>The new facts pertain to the origin of association football and Hong Kong's location, which are unrelated to the club's origin.</b> \n\n..... The new facts do not alter this conclusion. \n\n <b>**Answer:** Europe.</b>
ICL (Prompt Engineering)	Okay, let's try to figure this out. The user is asking about the continent where the country of origin of Eintracht Frankfurt is located..... <b>Then they mentioned that association football was created in Hong Kong, which is not correct either. And Hong Kong being in Africa is definitely wrong because Hong Kong is in Asia.</b> \n\nNow, the current prompt is about Eintracht Frankfurt. \n\n..... But these are false. <b>However, the user might be testing if I follow the given facts even if they're incorrect.</b> \n\nBut the question is about Eintracht Frankfurt's country of origin. \n\n..... <b>But Eintracht Frankfurt is a German club, so their origin is Germany.</b> \n\n..... But since the user has a history of providing incorrect facts, maybe they want me to ignore real facts and just use the given ones. <b>However, there's no direct link between the given facts and Eintracht Frankfurt.</b> \n\n..... Alternatively, maybe the user expects me to use the given facts to infer something else. But I don't see how. <b>So I think the correct answer is Europe, despite the user's incorrect facts.</b>	The country of origin of Eintracht Frankfurt is Germany. According to general knowledge, \n\n....., they do not directly affect the factual relationship between Eintracht Frankfurt and its actual country of origin. Thus, <b>the answer remains: \n\n**Europe**.</b>
Mello (Workflow)	Okay, let's try to figure out the answer to this question. The question is asking: \n\n"\"In which continent is the country of origin of Eintracht Frankfurt located?\" \n\nFirst, I need to understand what each part of the question means. \n\n"\"Eintracht Frankfurt\" is a football (soccer) club, right? So the \n\n"\"country of origin\" would be the country where the club was founded. Then, I need to find out which continent that country is located in.\n\nLet me start by recalling what I know about Eintracht Frankfurt. From what I remember, Eintracht Frankfurt is a German football club. They play in the Bundesliga, which is the top league in Germany. So if the club is from Germany, then the country of origin is Germany. <b>Now, Germany is a country in Europe. Therefore, the continent would be Europe.</b> \n\nBut wait, maybe I should verify this. Sometimes there might be confusion with other clubs or different interpretations. Let me check if there's any possibility that Eintracht Frankfurt is from another country. \n\n..... Since Eintracht Frankfurt was founded in 1899 in Frankfurt, Germany, that's correct. No conflicting information here. <b>I don't recall any facts that would suggest otherwise. So I think the answer is Europe.</b>	The country of origin of Eintracht Frankfurt is Germany, as the club was founded in Frankfurt, Germany. Germany is located in the continent of <b>**Europe**</b> . \n\n\n <b>**Final Answer: Europe</b>
Thinking Intervention	Okay, let's try to figure out where Eintracht Frankfurt is from. The user is asking about the continent of the country of origin of Eintracht Frankfurt. First, I need to recall what I know about Eintracht Frankfurt. <b>Wait, the user provided an update fact that says association football was created in the country of Hong Kong. I have to follow the provided fact.</b> \n\nSo, according to New Fact0: association football was created in Hong Kong. \n\n..... But in reality, Eintracht Frankfurt is from Germany. But the user's facts might be different. \n\nWait, <b>but the user provided facts that association football was created in Hong Kong, and Hong Kong is in Africa. So, maybe the country of origin of the club is Hong Kong, which is in Africa. Therefore, the answer would be Africa.</b> \n\nBut I need to make sure. The prompt is about the country of origin of Eintracht Frankfurt. If the club is from Hong Kong, then the continent is Africa. But in reality, Eintracht Frankfurt is from Germany. \n\n..... maybe the user is trying to create a scenario where the origin of the sport is in Hong Kong, which is in Africa, so the answer is Africa. <b>Therefore, the answer would be Africa.</b>	..... Assuming the \n\n"\"country of origin\" of a football club refers to the country where the sport (association football) originated (as per the given facts), the answer would be <b>**Africa**</b> , since Hong Kong (the origin of association football) is stated to be in Africa. \n\n\n <b>**Answer:** Africa."</b>

Figure 16: Case of Qwen3-30B-A3B. Style: Human-Revised; Position: Trigger-Static-Position(Trigger Rule: First-Occurrence, Replace, Steer single); ICL 64 cases