

# Logical Structure as Knowledge: Enhancing LLM Reasoning via Structured Logical Knowledge Density Estimation

Zhen Bi<sup>1,2,3\*</sup>, Zhenlin Hu<sup>1,2\*</sup>, Xueshu Chen<sup>1,2</sup>, Mingyang Chen<sup>4</sup>, Cheng Deng<sup>3,5</sup>,  
Yida Xue<sup>6</sup>, Zhen Wang<sup>1,2</sup>, Qing Shen<sup>1,2</sup>, Ningyu Zhang<sup>6</sup>, Jungang Lou<sup>1,2†</sup>

<sup>1</sup>Huzhou Normal University, <sup>2</sup>Zhejiang Key Laboratory of Intelligent Education Technology and Application, <sup>3</sup>Banbu AI Foundation, <sup>4</sup>Baichuan Inc, <sup>5</sup>University of Edinburgh, <sup>6</sup>Zhejiang University

## Abstract

The reasoning capabilities of Large Language Models (LLMs) are increasingly attributed to training data quality rather than mere parameter scaling. However, existing data-centric paradigms often equate quality with factuality or diversity and ignore the internal logical complexity of training samples. In this work, we propose that natural language harbors *Structured Logical Knowledge* manifested through entailment relationships and logical topologies. To quantify this, we introduce Structured Logical Knowledge Density (SLKD), a novel metric that measures logical information content by decomposing natural language into executable predicates and logical primitives. Our analysis reveals a significant logical disparity in current datasets where sparse logical signals predominate. Consequently, we propose a *density-aware re-cognizing optimization* strategy that prioritizes high-density logical samples to enhance with the LLM’s reasoning ability. Extensive experiments demonstrate that our approach enhances reasoning performance and generalization without increasing total data volume. These results, further validated within a reinforcement learning framework, suggest that elevating logical density is more critical than expanding data scale for realizing the full cognitive potential of LLMs. The released code is available in the Appendix C.

## 1 Introduction

Recent advances in Large Language Models (LLMs) are attributed not merely to the scaling of parameters but significantly to the enhancement of training data quality. Within the data-centric paradigm (Brown et al., 2020; Raffel et al., 2020; Hoffmann et al., 2022; Zhou et al., 2023), a consensus has emerged: high-quality data serves as the cornerstone of LLM reasoning capabilities.

\*These authors contributed equally to this work.

†Corresponding author.

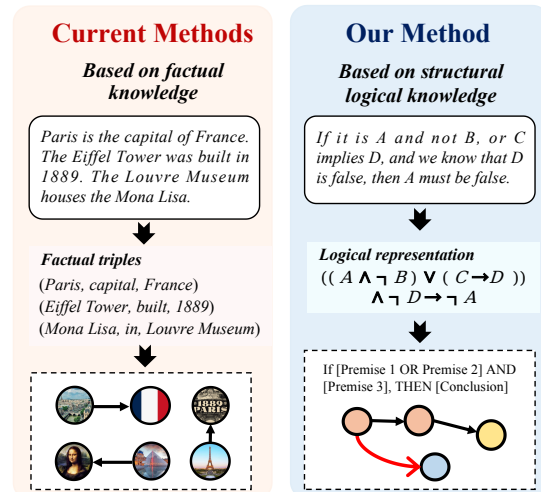


Figure 1: Comparison between the current methods and our method. The text **on the left** contains entities and relationships, but it is logically quite straightforward. The logical structure of the text **on the right** is very dense, but it contains many meaningless symbols.

However, the precise definition of data quality remains elusive. Prevailing research predominantly focuses on dimensions such as factuality and diversity (Hua et al., 2025; Wang et al., 2025a). From this perspective, knowledge is often equated with entities and facts stored within the text, and we argue that this definition is incomplete. Beyond declarative facts, natural language implicitly harbors another crucial form of knowledge: **Structured Logical Knowledge**. This knowledge manifests not as isolated entities, but as entailment relationships between predicates, conditional constraints, and logical topological structures.

Although there are currently some studies on the complexity of training data for LLMs (Prystawski et al., 2023; Bi et al., 2024), we have long lacked an effective methodology to quantify the precise knowledge density contained within data. Existing analytical approaches typically assess knowledge richness by analyzing factual triples implicitly contained in the text or the underlying syntactic trees.

However, they struggle to model the distinctions in knowledge density inherent to texts featuring complex deductive reasoning (shown in Figure 1). This absence of appropriate metrics often leads to an indiscriminate expansion of data volume, which may overlook the critical role of logical structure density as a decisive variable in the emergence of model reasoning capabilities.

To bridge this theoretical gap, we propose regarding logical structure as a measurable form of knowledge. We introduce a novel metric: *Structured Logical Knowledge Density (SLKD)*. Distinct from traditional metrics for text knowledge complexity, **SLKD** does not focus on the factual knowledge contained in the data, but instead delves into the logical core. By decomposing natural language into executable predicates and logical primitives, and calculating the topological complexity of their combinatorial reasoning processes, it objectively measures the logical information content inherent in a sample.

Leveraging **SLKD**, we identify a significant logical disparity in existing training data distributions—a vast amount of data provides only sparse logical knowledge, while the importance of **high-SLKD** samples is overlooked. To address this, we propose a *density-aware re-cognizing optimization strategy*. Instead of merely pursuing data scale, this strategy reshapes the training process according to **SLKD**: it first establishes a foundational logical schema via extensive density-based data, and subsequently concentrates computational resources on high-density samples, thereby enabling the model to efficiently acquire complex logical entailment patterns. The main contributions are as follows:

- We propose the perspective of *Logical Structure as Knowledge*, highlighting the oversight of logical knowledge density in current data estimation paradigms.
- We introduce **SLKD**, an objective metric for quantifying the complexity of internal logical entailment, capable of precisely identifying "high logical value regions" within data.
- We design an optimization strategy based on **SLKD** that significantly enhances LLM performance on logical reasoning tasks without increasing the total data volume. This confirms that the key to improving model reasoning lies in elevating the logical knowledge density of training data, rather than merely expanding its scale.

## 2 Related Work

### 2.1 Data-Centric Reasoning for LLMs

Recent work increasingly emphasizes that, beyond model architecture and algorithm design, the structure, provenance, and difficulty of training data critically shape LLM reasoning performance (Chen et al., 2025c; Qu et al., 2025; Sui et al., 2025; Ruis et al., 2025). This data-centric shift prioritizes data composition and quality over sheer volume to enhance reasoning ability (Wu et al., 2026; Jin et al., 2024; Wang et al., 2025b). Many studies highlight that higher-quality, logically structured examples—rather than simply more data—yield better generalization and reasoning performance (Peng et al., 2025; Yu et al., 2025; Li et al., 2025; Wettig et al., 2024; Zhao et al., 2024; Yu et al., 2024). In particular, aligning sample difficulty with model capability is identified as key to effective training (Gao et al., 2025). Instruction tuning research supports this view, showing that both prompt quality and exposure timing influence reasoning emergence (Lv et al., 2025; Kim and Lee, 2024). Meanwhile, Kandpal and Raffel (2025) highlight the often-overlooked human labor cost in curating such training data.

### 2.2 Reasoning Emergence for LLMs

Complementing these data-centric strategies, other studies explore how reasoning emerges from the interaction between cognition and structured input. For instance, stepwise reasoning is viewed as an emergent property of sequential data exposure rather than a fixed architectural feature (Prystawski et al., 2023). Furthermore, only data within a suitable complexity range appears to effectively stimulate reasoning behavior (Bi et al., 2024), suggesting that model performance is bounded by cognitive processing capacity. Empirical evidence further shows that carefully curated datasets often outperform larger but noisier corpora in supporting reasoning skills (Ma et al., 2025; Ye et al., 2025; Morishita et al., 2024; Wang et al., 2025c; Yang et al., 2025; Hua et al., 2025), reinforcing the value of reasoning supervision that is both selective and structurally rich. Prior work underscores data’s role in LLM reasoning but often lacks precise difficulty metrics and clear goals. We introduce **SLKD**, a unified score for reasoning potential, and *re-cognizing optimization*, which emphasizes high **SLKD** examples while preserving diversity to enhance LLM’s logical reasoning performance.

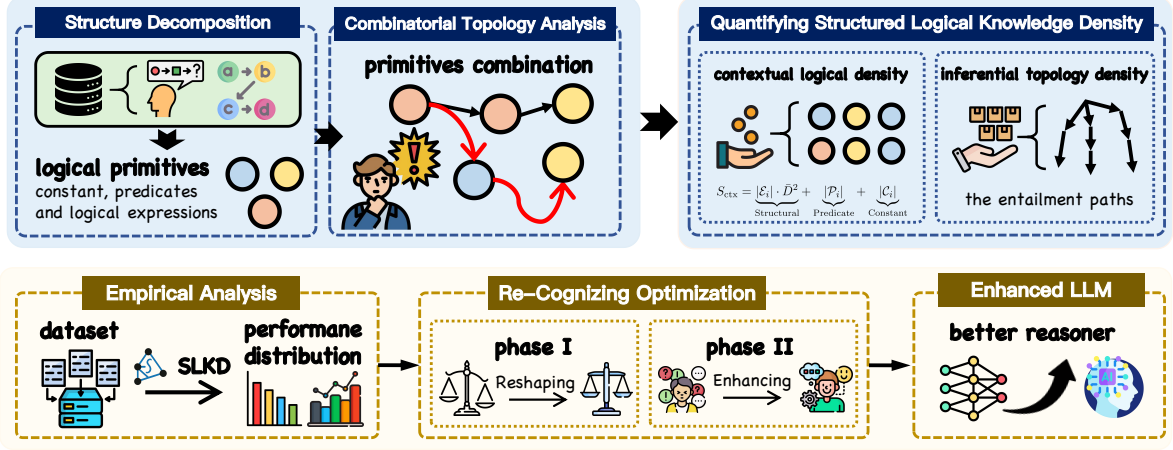


Figure 2: The overall framework. **Top: Logical structure as knowledge.** We first extract logical elements from each example and perform combinatorial reasoning to derive *Structured Logical Knowledge Density (SLKD)* score. **Bottom: SLKD-based optimization method.** We then analyze the performance distribution across **SLKD** levels on multiple datasets. Based on this, we propose the *re-cognizing optimization* strategy: the first stage reshapes the model’s recognition of reasoning patterns, and the second enhances its logical reasoning capability, thereby improving overall performance.

### 3 Logical Structure as Knowledge

In this section, we introduce *Structured Logical Knowledge Density (SLKD)*. The calculation of **SLKD** involves transforming unstructured natural language into structured predicate logic representations and then evaluating the topological complexity of the resulting reasoning chains.

#### 3.1 Structure Decomposition

The first step in quantifying **SLKD** is to decouple the logical form from the natural language surface. We treat the reasoning process as a manipulation of executable predicates. Given a data sample  $Q$ , we employ a distillation function  $f$  (implemented via advanced LLMs) to decompose the text into three fundamental **logical primitives**:

$$f(Q) \Rightarrow \{P, C, E\} \quad (1)$$

- **Constants:**  $C$  is the set of entities or values involved in the reasoning context (e.g., specific objects, names).
- **Predicates:**  $P$  is the set of relations or properties binding the constants (e.g.,  $\text{IsParentOf}(x, y)$ ,  $\text{Greater}(a, b)$ ).
- **Logical Expressions:**  $E$  is the set of explicit logical constraints or rules from the context, formalized as first-order logic expressions.

This decomposition converts the fuzzy information of natural language into a rigorous **predicate knowledge base**, serving as the substrate for subsequent density analysis.

#### 3.2 Combinatorial Topology Analysis

The density of logical knowledge arises from **how these primitives combine to form valid entailment paths**, we define the combinatorial topology analysis process to reconstruct the logical trajectory required to derive the answer from the primitives. Using a function  $F$ , we generate the reasoning topology, which consists of the precondition structure  $\bar{E}$  and a sequence of derivation steps  $S$ :

$$\text{Topology} = F(\text{Primitives}, A) \Rightarrow \{\bar{E}, S\} \quad (2)$$

where  $A$  is answer set,  $\bar{E}$  is the precondition logic structure and  $S$  is the single reasoning chain required for the reasoning trajectory. Detailedly, each reasoning node  $s_k \in S$  contains:

$$s_k = (\# \text{Operations}_k, D_k^{\text{nest}}, \text{Expression}_k) \quad (3)$$

where  $\# \text{Operations}_k$  counts the logical operators (AND/OR/NOT),  $D_k^{\text{nest}}$  represents the nesting depth of the logical expression, and  $\text{Expression}_k$  is its formal representation.

The nesting depth is a critical parameter, as it reflects the hierarchical complexity of the entailment relationship.

#### 3.3 Quantifying Structured Logical Knowledge Density (SLKD)

We quantify the density of logical knowledge by aggregating the topological complexity of the context

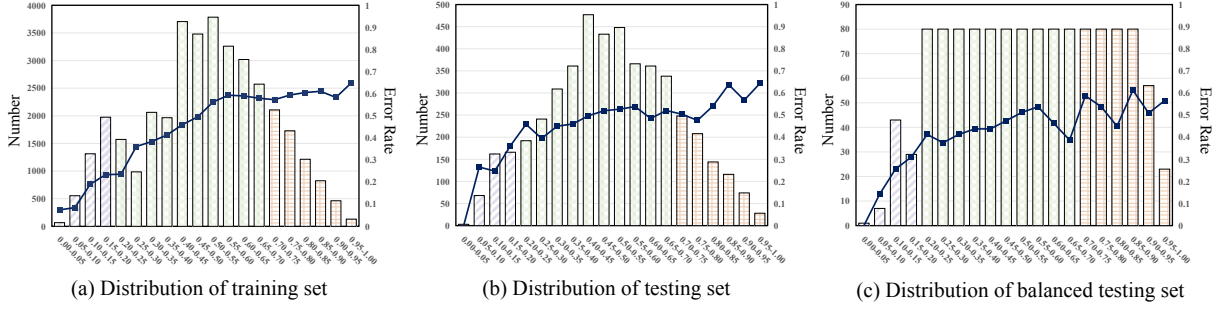


Figure 3: Effectiveness verification for **SLKD**. Sample counts (bars, left axis) and model error rates (lines, right axis) are shown across **SLKD** score bins. All three panels share the same layout: the x-axis divides the score range into uniform intervals, the bar height indicates the number of examples per interval, and the overlaid line traces the error rate. (a) Training-set distribution and error. (b) Original test-set distribution. (c) Balanced test-set distribution.

and the derivation path. This measurement encompasses contextual logical density for premise structures, inferential topology density for inferential depth, and **SLKD** score aggregation for synthesizing these dimensions into a unified metric.

### 3.3.1 Contextual Logical Density

First, we measure the static logical information contained in the problem description. The Context Density  $S_{\text{ctx}}$  is calculated as:

$$S_{\text{ctx}} = \underbrace{|E_i| \cdot \bar{D}^2}_{\text{Structural}} + \underbrace{|P_i|}_{\text{Predicate}} + \underbrace{|C_i|}_{\text{Constant}} \quad (4)$$

where  $|E_i|$  counts logical expressions,  $\bar{D}$  is their average nesting depth (calculated via parse tree analysis),  $|P_i|$  and  $|C_i|$  tally unique predicates and constants respectively.

Here, the term  $\bar{D}^2$  applies a quadratic penalty to the nesting depth, reflecting the non-linear increase in cognitive load required to encode deeply nested logical structures. This formulation ensures that complex conditional dependencies contribute significantly more to the density score than simple linear facts.

### 3.3.2 Inferential Topology Density

Next, we quantify the dynamic complexity of the entailment path required to validate a specific option. For an option  $l$ , its inferential topology density  $S_{\text{opt}}^{(l)}$  is defined as:

$$S_{\text{opt}}^{(l)} = \underbrace{|R_l| \cdot \bar{D}_l^2}_{\text{Preconditions}} + \sum_{k=1}^{T_l} \underbrace{(1 + \# \text{Operations}_{l,k}) D_{l,k}^2}_{\text{Step } k} \quad (5)$$

where  $l \in \{1, \dots, L\}$  indexes the  $L$  answer options,  $R_l$  is the set of precondition expressions for

option  $l$  with average nesting depth  $\bar{D}_l$ , and  $T_l$  is the number of reasoning steps for option  $l$ . Each step  $k$  contributes according to its operator count  $\# \text{Operations}_{l,k}$  and nesting depth  $D_{l,k}$ .

This component aggregates the structural weight of each derivation step. By summing over the sequence  $T_l$ , we capture the total "logical work" encoded in the reasoning chain.

### 3.3.3 SLKD Score Aggregation

The final **SLKD**  $S$  is obtained by combining the contextual and derivational densities. To facilitate its use as a weighting factor in training, we normalize the raw score into a probabilistic range  $[0, 1]$ :

$$S_{\text{raw}} = S_{\text{ctx}} + \sum_{l=1}^L S_{\text{opt}}^{(l)} \quad (6)$$

$$S = \sigma \left( \gamma \cdot \frac{\log(S_{\text{raw}} + 1) - \mu}{\sqrt{\delta^2 + \epsilon}} + \beta \right) \quad (7)$$

where  $\sigma$  denotes the sigmoid function,  $\mu$  and  $\delta$  are dataset statistics, and parameters ( $\gamma = 1, \beta = 0, \epsilon = 10^{-5}$ ) ensure stable  $[0, 1]$  normalization.

This normalized **SLKD** score serves as a proxy for the richness of reasoning patterns contained in a sample, independent of the sample's surface length.

## 4 Empirical Analysis for SLKD

In order to fully investigate the effect of **SLKD**, we conduct experiments in existing training data (experimental setting is in Appendix C). We not only focus on the **SLKD** score distribution of the existing data but also use data with different interval distributions as training sets to separately validate the effectiveness of the data within these intervals.

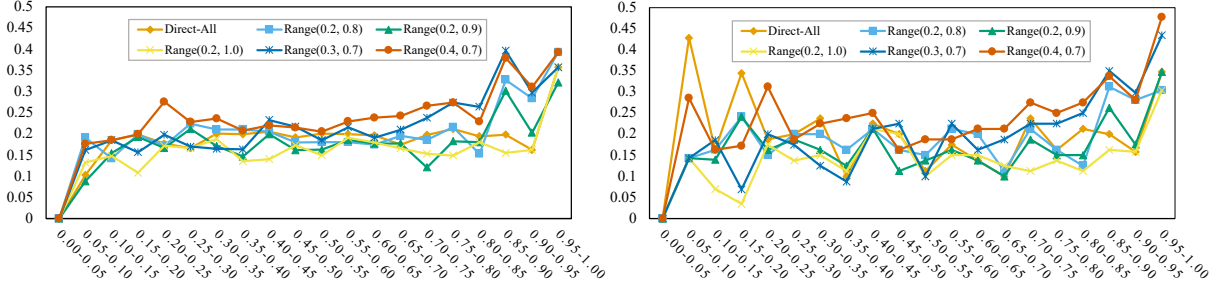


Figure 4: Experimental results of fine-tuning models in different intervals. "Direct-All" denotes a model fine-tuned on all training examples. "Range( $x, y$ )" denotes a model fine-tuned on examples whose **SLKD** scores fall between  $x$  and  $y$ . Left: Test dataset experiment results. Right: Balanced testing set experiment results. The horizontal axis represents score intervals, and the vertical axis represents error rates (lower error indicates better performance).

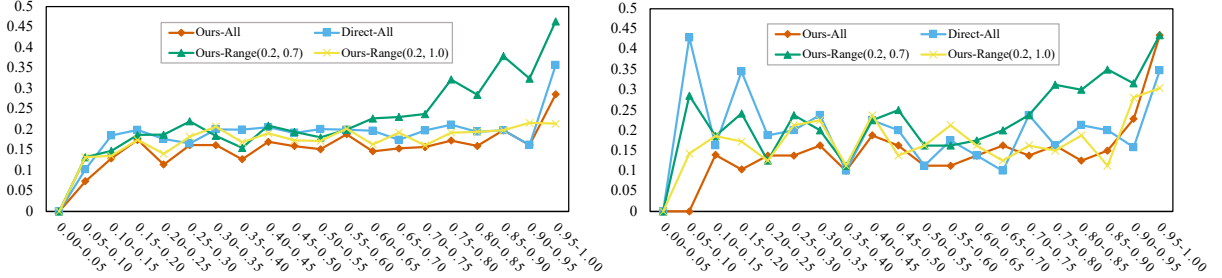


Figure 5: Experimental results of fine-tuning models using different methods. "Ours-All": model fine-tuned on the full training set using *re-cognizing optimization*. "Ours-Range( $x, y$ )": model fine-tuned with *re-cognizing optimization* on examples whose **SLKD** scores fall between  $x$  and  $y$ . Left: Test dataset results. Right: Balanced testset results. The horizontal axis represents score intervals and the vertical axis represents error rates (lower error indicates better performance).

#### 4.1 **SLKD** Score Distribution

In Figure 3(a), when the training set is binned into 20 score intervals, the sample counts form a bell-shaped curve with a weighted mean of  $\mu = 0.526$  and standard deviation  $\sigma = 0.204$ . This approximately Gaussian frequency distribution demonstrates that our **SLKD** scores effectively separate examples by their underlying reasoning potential. Figure 3(a) also plots model error rate against these intervals: error climbs from 8.5% at a score of 0.1 to 59.3% at 0.6, demonstrating a clear positive correlation between **SLKD** and failure rate. Beyond a score of 0.6, error rate plateaus at approximately  $61\% \pm 3.6\%$ . This saturation likely arises because (1) very high **SLKD** items exceed the model's current reasoning capacity, and (2) the top bins contain few, uniformly hard cases, so failures become uniformly pervasive. Unlike surface metrics (e.g., sentence length or vocabulary complexity), our score distribution captures deeper logical structure.

#### 4.2 **SLKD** Score Distribution Balancing

The original testset contains 4,743 examples and its **SLKD** score distribution (Figure 3(b)) reveals

that 74.3% of samples fall into the mid-range interval  $(0.2, 0.7)$ , while only 17.2% lie in the high **SLKD** interval  $(>0.7)$ . Such skew can introduce two evaluation biases: (1) Apparent overfitting in the mid-range: strong results in the overrepresented middle interval may conceal weaknesses on higher **SLKD** examples; (2) Insufficient statistical power: the small number of high **SLKD** samples ( $n = 818$ ) yields wide confidence intervals for any observed gains. To correct this, **we build a balanced benchmark** (Figure 3(c)) by drawing 80 examples from each interval. This uniform sampling ensures equal representation across the **SLKD** spectrum, eliminating bias and enabling reliable comparisons.

#### 4.3 Data Potential Analysis by **SLKD**

Building on our validation of the **SLKD** score and the creation of a balanced test benchmark, we next investigated how filtering training examples by their **SLKD** scores affects learning. In Figure 4, we systematically evaluate model performance when trained on subsets defined by **SLKD** intervals, using both the original and balanced test sets. These experiments reveal three critical patterns, which

can be summarized as the following observations:

- **Obs 1: Low SLKD data can be safely pruned.** Both Figure 4(a) and (b) show that training on Range(0.2, 1.0), which omits only the lowest 20% of examples, consistently outperforms full-data training, reducing error rates in nearly every bin. Even when further restricting to Range(0.2, 0.8) (removing both the lowest and highest 20%), overall accuracy remains on par with or slightly above the full-data baseline ( $\Delta = +0.3\%$ ), and mid-range bins (0.2, 0.7) see notable gains. These results confirm that low SLKD examples can be pruned without harming model performance and can often lead to improvements, while indiscriminate use of all data may introduce noise.
- **Obs 2: High SLKD data are catalysts for improvement.** Figure 4(a) shows that training on Range(0.2, 1.0) yields the lowest error rates across all bins. When the upper bound is reduced to Range(0.2, 0.9), error rates rise, particularly in the highest SLKD bins. Narrowing further to Range(0.2, 0.8) causes a further increase in errors. This stepwise degradation, also reflected in Figure 4(b), confirms that examples with the highest SLKD scores drive the most significant performance gains and act as catalysts for model learning.
- **Obs 3: Too little data breaks the learning process.** Figures 4(a) and (b) show that restricting training to narrow SLKD intervals Range(0.3, 0.7) or Range(0.4, 0.7) leads to severe performance degradation, especially in the mid SLKD range (0.2, 0.7), where error rates exceed those from broader training ranges. This demonstrates that sparse coverage of the SLKD spectrum impairs the model’s ability to internalize and apply core reasoning patterns. The sharp drop in performance underscores the necessity of preserving sufficient data diversity and quantity across all SLKD levels to sustain learning.

In summary, these results indicate that **a large volume of data only provides sparse logical knowledge, while the importance of samples with high SLKD has been overlooked.** To address these issues, we propose the *re-cognizing optimization* method in Section 5.

## 5 SLKD-based Re-cognizing Method

Inspired by cognitive theories (Sweller, 1988; Lieder and Griffiths, 2020), we propose a two-

phase *re-cognizing optimization*. Our optimization method is guided by SLKD and divided into two phases. We first recalibrates the LLM’s logical reasoning schema and then leverages SLKD scores to reinforce its logical reasoning pathways.

### 5.1 Phase I: Model Cognition Reshaping

In this phase, we reorder the training data according to SLKD scores to "reset" and align the model’s reasoning framework, applying Sweller’s cognitive load theory. Allowing the model to explore the full spectrum of SLKD examples from the outset helps it form broad reasoning patterns. This “low-stakes exploration” mirrors how human learners build foundational knowledge before tackling more challenging tasks, minimizing extraneous cognitive load and establishing a robust framework for subsequent learning.

### 5.2 Phase II: Cognitive Reasoning Enhancement

Here, we implement resource-rational analysis by guiding the model’s focus according to normalized SLKD scores. The probability  $p$  that the model attends to sample  $i$  is calculated as follows:

$$p_i = \frac{\hat{s}_i}{\sum_{j=1}^N \hat{s}_j}, \quad \hat{s}_i = \frac{s_i - s_{\min}}{s_{\max} - s_{\min}} \quad (8)$$

where  $s_i$  is the raw SLKD score of sample  $i$ ,  $s_{\min}$  and  $s_{\max}$  are the minimum and maximum scores in the dataset,  $\hat{s}_i$  is the normalized score for sample  $i$ , and the denominator  $\sum_{j=1}^N \hat{s}_j$  sums these normalized scores over all  $N$  samples. The *re-cognizing optimization* stratege echoes Lieder’s insight that "human cognition allocates limited resources to maximize expected reasoning gains relative to reasoning demands". Through cognition-driven emphasis, the model is steered toward high SLKD examples, thus securing greater learning returns under constrained resources.

### 5.3 Results and Analysis

We evaluate the effectiveness of the *re-cognizing optimization* strategy by comparing it with static baselines and other data-centric methods across datasets. Other methods include curriculum learning (Bengio et al., 2009) and bin-based progressive learning (Klamkin et al., 2024). In curriculum learning, training examples are introduced in order of increasing SLKD scores; in bin-based progressive learning, data is divided into different SLKD

MODEL	METHODS	UNBALANCED					BALANCED				
		Recolr	LogiQA	LogiQA2.0	LogicBench	Avg.	Recolr	LogiQA	LogiQA2.0	LogicBench	Avg.
<i>Closed-source LLMs</i>	GPT-4	0.808	0.525	0.664	0.774	0.707	0.794	0.588	0.651	0.765	0.699
	DeepSeek-V3	0.754	0.484	0.663	0.814	0.712	0.738	0.500	0.663	0.847	0.686
LLaMA3.1-8B	Base	0.444	0.347	0.378	0.707	0.521	0.447	0.412	0.383	0.714	0.490
	Directly	0.896	0.724	0.805	0.812	0.806	0.901	0.774	0.802	0.812	0.816
	Curriculum Learning	0.926	0.634	0.798	0.854	0.811	0.908	0.684	0.773	0.864	0.815
	Bin-based Progressive Learning	0.930	0.742	0.821	0.831	0.826	0.894	0.763	0.809	0.836	0.823
	<i>Optimization w/o Stage1</i>	0.904	0.699	0.791	0.747	0.771	0.918	0.726	0.825	0.763	0.806
	<i>Optimization w/o Stage2</i>	0.910	0.717	0.796	0.824	0.809	0.918	0.768	0.834	0.859	0.844
	<i>Our Re-Cognizing Method</i>	<b>0.930</b>	<b>0.750</b>	<b>0.835</b>	<b>0.857</b>	<b>0.843</b>	<b>0.922</b>	<b>0.775</b>	<b>0.840</b>	<b>0.865</b>	<b>0.851</b>
	(relative gain)	+0.00%	+1.08%	+1.71%	+0.35%	+2.06%	+1.54%	+0.13%	+3.83%	+0.12%	+3.40%
Qwen2.5-7B	Base	0.466	0.372	0.424	0.629	0.509	0.433	0.435	0.409	0.678	0.490
	Directly	0.786	0.743	0.748	0.810	0.778	0.797	0.719	0.788	0.888	0.798
	Curriculum Learning	0.864	0.750	0.806	0.850	0.823	0.872	0.739	0.797	0.919	0.831
	Bin-based Progressive Learning	0.890	0.730	0.782	0.842	0.812	0.911	0.697	0.806	0.719	0.833
	<i>Optimization w/o Stage1</i>	0.856	0.808	0.817	0.833	0.827	0.833	0.790	0.822	0.894	0.835
	<i>Optimization w/o Stage2</i>	0.798	0.679	0.767	0.827	0.784	0.839	0.684	0.778	0.872	0.793
	<i>Our Re-Cognizing Method</i>	<b>0.944</b>	<b>0.824</b>	<b>0.845</b>	<b>0.858</b>	<b>0.858</b>	<b>0.948</b>	<b>0.797</b>	<b>0.884</b>	<b>0.919</b>	<b>0.886</b>
	(relative gain)	+6.07%	+9.87%	+4.84%	+0.94%	+4.25%	+4.06%	+7.85%	+9.68%	+0.00%	+6.36%

Table 1: Experimental results from different test sets. Our *re-cognizing optimization* method is compared against other approaches on both LLaMA3.1-8B and Qwen2.5-7B using accuracy as the evaluation metric. Results are reported on both unbalanced and balanced test sets, with the best performance in each setting highlighted in bold.

bins, and the next **SLKD** bin is introduced only after the model has been trained on the current bin. Table 1 summarizes the performance across datasets, while Figure 5 shows the error rates across **SLKD** bins for the original and balanced test sets. Figure 6 (left) compares our method with the other two methods. From the experimental results, we have the following observations:

- **Obs 1: *Re-cognizing optimization consistently reduces errors across datasets and SLKD bins, leading to comprehensive and robust performance gains.*** In Figure 5, Figure 6 (left), and Table 1, our method consistently achieves the best performance across datasets and test splits, confirming its ability to generalize dataset-specific reasoning patterns. For example, in the lowest **SLKD** interval  $(0, 0.2)$ , errors drop from 17.6%/25% to 12.6%/11.2% on the original/balanced test sets, demonstrating reduced overfitting to trivial cases. This trend holds across all **SLKD** bins, resulting in a systematic reduction in errors across the spectrum. On datasets like LogiQA and LogiQA2.0, where baseline accuracies are lower, our method shows a significant improvement over others, demonstrating its adaptability and robustness on diverse datasets.
- **Obs 2: *Re-cognizing optimization demonstrates resilience in data-scarce settings by maintaining strong performance even with restricted training ranges.*** In Figure 5, when we apply *re-cognizing optimization* to restricted training sets—Our-Range $(0.2, 0.7)$  and Our-Range $(0.2, 1.0)$ —we observe a 13.6% error increase when high **SLKD** exam-

ples ( $>0.7$ ) are omitted. However, our method effectively compensates for this loss, maintaining performance comparable to full-data training within the Our-Range $(0.2, 0.7)$  subset ( $\Delta = +2.5\%$ ). Expanding the training range to include higher **SLKD** examples (Our-Range $(0.2, 1.0)$ ) consistently maintains lower error rates than direct full-data training, validating the predictive power of our **SLKD** and demonstrating that our method can effectively improve performance even in data-scarce settings.

**Ablation Study** To assess the contribution of each training stage, we conduct ablation experiments by removing either Stage 1 or Stage 2 from the full *re-cognizing optimization* pipeline. As shown in Table 1, both variants lead to substantial performance drops on both LLaMA3.1-8B and Qwen2.5-7B. This confirms that both stages are essential for achieving optimal reasoning performance. In summary, our results show our *re-cognizing optimization* effectively directs the model to high **SLKD** examples, boosting learning efficiency. It systematically reduces errors across **SLKD** bins (proving improved reasoning) and outperforms traditional methods like curriculum and bin-based learning, demonstrating flexibility and robustness across scenarios.

## 5.4 Expansion Exploration

Some studies (Chen et al., 2025a; Chu et al., 2025; Chen et al., 2025b) suggest that supervised fine-tuning (SFT) and reinforcement learning (RL) play different roles in stimulating the capabilities of language models. To assess the impact of our

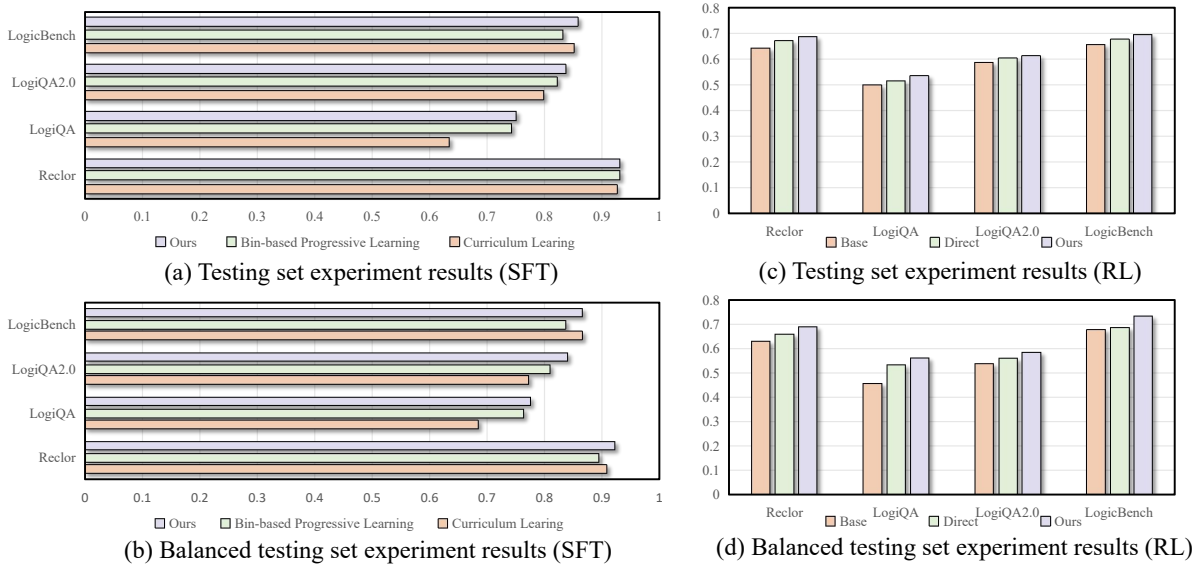


Figure 6: Experimental results of SFT and RL across different dataset. Accuracy is used as the evaluation metric (higher is better). Both SFT and RL methods are evaluated on the original and balanced test sets. (a) SFT on the original test set: Ours vs. Bin-based Progressive Learning vs. Curriculum Learning. (b) SFT on the balanced test set: same comparisons. (c) RL on the original test set: Ours vs. Base vs. Direct static reward). (d) RL on the balanced test set: same comparisons. In every setting, our method (Ours) achieves the highest accuracy, demonstrating its superiority under both SFT and RL regimes.

**SLKD** scores in an RL setting, we provide proportionally larger accuracy rewards for samples with higher **SLKD** scores, incentivizing the model to handle more reasoning-intensive cases and internalize richer reasoning patterns. We also enforce a structured output format and reward the model for both format compliance and the quality of its reasoning trace. Figure 6(c) and (d) compare our method against two baselines: the base model without RL and a direct variant with fixed accuracy rewards. In both the original and balanced test sets, our **SLKD-guided** method consistently outperforms these baselines, confirming that dynamic, score-based rewards combined with structured output incentives significantly enhance the model’s reasoning capabilities.

## 6 Discussion

Our results resonate with the *Densing Law of LLMs* (Xiao et al., 2025), which observes exponential growth in model "capability density". By positing that *Logical Structure is Knowledge*, we extend this intuition to the data dimension. This aligns with the view of language modeling as data compression (Delétang et al., 2024), suggesting that reasoning potential depends less on token volume and more on the concentration of logical primitives—a principle validated by the Phi-series (Abdin et al., 2024). We thus advocate for a shift from volume-based

Scaling Laws (Kaplan et al., 2020) toward a structural density-aware paradigm.

Regarding a universal data knowledge density, our analysis suggests that density is likely a relative metric tied to LLMs’ cognitive boundary rather than an absolute constant. The gains from high-**SLKD** samples highlight a logical sparsity in current corpora, where redundant patterns dilute learning signals (Abbas et al., 2023). Future work should investigate whether excessive logical density triggers distribution shifts or overfitting to formal structures, particularly when interacting with LLMs’ internal mechanisms and inductive biases.

## 7 Conclusion

In this paper, we presented a data-centric framework to bolster the reasoning capabilities of LLMs. By introducing *Structured Logical Knowledge Density (SLKD)*, we provide a formal metric to quantify the intrinsic logical complexity of training samples. Furthermore, our proposed re-cognizing optimization strategy effectively aligns training data with the inherent reasoning boundaries of LLMs. Extensive experiments demonstrate that our approach consistently outperforms existing baseline methods. We hope our findings offer a fresh perspective on measuring and unlocking the untapped reasoning potential of LLMs, while inspiring future research into cognitive-level data engineering.

## Limitation

One limitation of our work is the relatively small scale and variety of models we evaluated—due to budget and time constraints, we focused on only two model families (e.g., LLaMA3.1-8B and Qwen2.5-7B), which may limit the generalizability of our findings to larger or more diverse architectures. Additionally, while our **SLKD** score provides a useful proxy for reasoning potential, its formulation could be refined further: the current metrics may not capture all aspects of reasoning complexity or transfer seamlessly to other task domains. Moreover, our logical element extraction relies on LLM-based distillation functions, which can introduce noise or inaccuracies; nevertheless, we found these errors to be minor and within acceptable bounds, having minimal impact on overall metric reliability. Finally, our use of GRPO-based reinforcement learning was exploratory and preliminary; more extensive experiments with alternative reward schemes, longer training runs, and varied model capacities will be necessary to fully assess the robustness and scalability of *re-cognizing optimization*. We also did not explore individualized learning trajectories emphasized in cognitive science, nor use our reasoning-intensity signal to dynamically switch between System 1 and System 2 modes. Moreover, we restrict our evaluation to fine-tuning existing pretrained checkpoints rather than full-scale pretraining, since retraining multi-billion-parameter models from scratch requires resources beyond our current capabilities.

## Acknowledgments

We would like to express gratitude to the anonymous reviewers for constructive comments. This work was supported by Open Research Fund of Zhejiang Key Laboratory of Intelligent Education Technology and Application (No. 2025ZN-JYKF005), the "Pioneer and Leader + X" Plan Project of Zhejiang Province under Grant (2024C01162), Zhejiang Provincial Natural Science Foundation of China under Grant (No. LRG25F030003 and LQN25F020023) and the National Natural Science Foundation of China (No.62506128).

## References

Amro Abbas, Kushal Tirumala, Dániel Simig, Surya Ganguli, and Ari S. Morcos. 2023. [Semdedup: Data-](#)

[efficient learning at web-scale through semantic deduplication](#). *Preprint*, arXiv:2303.09540.

Marah I Abdin, Sam Ade Jacobs, Ammar Ahmad Awan, Jyoti Aneja, Ahmed Awadallah, Hany Awadalla, Nguyen Bach, Amit Bahree, Arash Bakhtiari, Harkirat S. Behl, Alon Benhaim, Misha Bilenko, Johan Bjorck, Sébastien Bubeck, Martin Cai, Caio César Teodoro Mendes, Weizhu Chen, Vishrav Chaudhary, Parul Chopra, and 68 others. 2024. [Phi-3 technical report: A highly capable language model locally on your phone](#). *CoRR*, abs/2404.14219.

Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. 2009. [Curriculum learning](#). In *Proceedings of the 26th Annual International Conference on Machine Learning, ICML '09*, page 41–48, New York, NY, USA. Association for Computing Machinery.

Zhen Bi, Ningyu Zhang, Yinuo Jiang, Shumin Deng, Guozhou Zheng, and Huajun Chen. 2024. When do program-of-thought works for reasoning? In *AAAI*, pages 17691–17699. AAAI Press.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, and 1 others. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Hardy Chen, Haoqin Tu, Fali Wang, Hui Liu, Xianfeng Tang, Xinya Du, Yuyin Zhou, and Cihang Xie. 2025a. Sft or rl? an early investigation into training rl-like reasoning large vision-language models. <https://github.com/UCSC-VLAA/VLAA-Thinking>.

Mingrui Chen, Haogeng Liu, Hao Liang, Huaibo Huang, Wentao Zhang, and Ran He. 2025b. [Unlocking the potential of difficulty prior in rl-based multimodal reasoning](#). *Preprint*, arXiv:2505.13261.

Qiguang Chen, Libo Qin, Jinhao Liu, Dengyun Peng, Jiannan Guan, Peng Wang, Mengkang Hu, Yuhang Zhou, Te Gao, and Wanxiang Che. 2025c. Towards reasoning era: A survey of long chain-of-thought for reasoning large language models. *CoRR*, abs/2503.09567.

Tianzhe Chu, Yuexiang Zhai, Jihan Yang, Shengbang Tong, Saining Xie, Dale Schuurmans, Quoc V Le, Sergey Levine, and Yi Ma. 2025. [SFT memorizes, RL generalizes: A comparative study of foundation model post-training](#). In *Forty-second International Conference on Machine Learning*.

DeepSeek-AI. 2024. [Deepseek-v3 technical report](#). *Preprint*, arXiv:2412.19437.

Grégoire Delétang, Anian Ruoss, Paul-Ambroise Duquenne, Elliot Catt, Tim Genewein, Christopher Mattern, Jordi Grau-Moya, Li Kevin Wenliang, Matthew Aitchison, Laurent Orseau, Marcus Hutter, and Joel Veness. 2024. [Language modeling is compression](#). In *The Twelfth International Conference*

- on Learning Representations, *ICLR 2024, Vienna, Austria, May 7-11, 2024*.
- Chengqian Gao, Haonan Li, Liu Liu, Zeke Xie, Peilin Zhao, and Zhiqiang Xu. 2025. Principled data selection for alignment: The hidden risks of difficult examples. *CoRR*, abs/2502.09650.
- Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, and et al. 2024. [The llama 3 herd of models](#). *Preprint*, arXiv:2407.21783.
- Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, Thomas Hennigan, Eric Noland, Katherine Millican, George van den Driessche, Bogdan Damoc, Aurelia Guy, Simon Osindero, Karén Simonyan, Erich Elsen, and 3 others. 2022. [An empirical analysis of compute-optimal large language model training](#). In *Advances in Neural Information Processing Systems*, volume 35, pages 30016–30030. Curran Associates, Inc.
- Kai Hua, Steven Wu, Ge Zhang, and Ke Shen. 2025. [Attentioninfluence: Adopting attention head influence for weak-to-strong pretraining data selection](#). *Preprint*, arXiv:2505.07293.
- Mingyu Jin, Qinkai Yu, Dong Shu, Haiyan Zhao, Wen Yue Hua, Yanda Meng, Yongfeng Zhang, and Mengnan Du. 2024. The impact of reasoning step length on large language models. In *ACL (Findings)*, pages 1830–1842. Association for Computational Linguistics.
- Nikhil Kandpal and Colin Raffel. 2025. [Position: The most expensive part of an LLM \\*should\\* be its training data](#). In *Forty-second International Conference on Machine Learning Position Paper Track*.
- Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. 2020. [Scaling laws for neural language models](#). *CoRR*, abs/2001.08361.
- Jisu Kim and Juhwan Lee. 2024. [Strategic data ordering: Enhancing large language model performance through curriculum learning](#). *Preprint*, arXiv:2405.07490.
- Michael Klamkin, Mathieu Tanneau, Terrence W.K. Mak, and Pascal Van Hentenryck. 2024. [Bucketized active sampling for learning acopf](#). *Electric Power Systems Research*, 235:110697.
- Dacheng Li, Shiyi Cao, Tyler Griggs, Shu Liu, Xiangxi Mo, Eric Tang, Sumanth Hegde, Kourosh Hakhmaneshi, Shishir G Patil, Matei Zaharia, Joseph E. Gonzalez, and Ion Stoica. 2025. [Language models can easily learn to reason from demonstrations](#). In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 15979–15997, Suzhou, China. Association for Computational Linguistics.
- Falk Lieder and Thomas L. Griffiths. 2020. [Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources](#). *Behavioral and Brain Sciences*, 43:e1.
- Hanmeng Liu, Jian Liu, Leyang Cui, Zhiyang Teng, Nan Duan, Ming Zhou, and Yue Zhang. 2023. [Logiqa 2.0—an improved dataset for logical reasoning in natural language understanding](#). *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 31:2947–2962.
- Jian Liu, Leyang Cui, Hanmeng Liu, Dandan Huang, Yile Wang, and Yue Zhang. 2020. [Logiqa: A challenge dataset for machine reading comprehension with logical reasoning](#). In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pages 3622–3628. International Joint Conferences on Artificial Intelligence Organization. Main track.
- Qingsong Lv, Yangning Li, Zihua Lan, Zishan Xu, Jiwei Tang, Tingwei Lu, Yinghui Li, Wenhao Jiang, Hong-Gee Kim, Hai-Tao Zheng, and Philip S. Yu. 2025. [RAISE: Reinforced adaptive instruction selection for large language models](#). In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 11708–11723, Suzhou, China. Association for Computational Linguistics.
- Ruotian Ma, Peisong Wang, Cheng Liu, Xingyan Liu, Jiaqi Chen, Bang Zhang, Xin Zhou, Nan Du, and Jia Li. 2025. [S<sup>2</sup>R: Teaching LLMs to self-verify and self-correct via reinforcement learning](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 22632–22654, Vienna, Austria. Association for Computational Linguistics.
- Terufumi Morishita, Gaku Morio, Atsuki Yamaguchi, and Yasuhiro Sogawa. 2024. [Enhancing reasoning capabilities of LLMs via principled synthetic logic corpus](#). In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.
- OpenAI. 2023. GPT-4 technical report. *CoRR*, abs/2303.08774.
- Mihir Parmar, Nisarg Patel, Neeraj Varshney, Mutsumi Nakamura, Man Luo, Santosh Mashetty, Arindam Mitra, and Chitta Baral. 2024. [LogicBench: Towards systematic evaluation of logical reasoning ability of large language models](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13679–13707, Bangkok, Thailand. Association for Computational Linguistics.
- Ru Peng, Kexin Yang, Yawen Zeng, Junyang Lin, Dayiheng Liu, and Junbo Zhao. 2025. [Dataman: Data manager for pre-training large language models](#). In *The Thirteenth International Conference on Learning Representations*.
- Ben Prystawski, Michael Li, and Noah D. Goodman. 2023. Why think step by step? reasoning emerges from the locality of experience. In *NeurIPS*.

- Xiaoye Qu, Yafu Li, Zhaochen Su, Weigao Sun, Jianhao Yan, Dongrui Liu, Ganqu Cui, Daizong Liu, Shuxian Liang, Junxian He, Peng Li, Wei Wei, Jing Shao, Chaochao Lu, Yue Zhang, Xian-Sheng Hua, Bowen Zhou, and Yu Cheng. 2025. [A survey of efficient reasoning for large reasoning models: Language, modality, and beyond](#). *Preprint*, arXiv:2503.21614.
- Qwen, An Yang, Baosong Yang, and 1 others. 2025. [Qwen2.5 technical report](#). *Preprint*, arXiv:2412.15115.
- Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. [Exploring the limits of transfer learning with a unified text-to-text transformer](#). *J. Mach. Learn. Res.*, 21:140:1–140:67.
- Laura Ruis, Maximilian Mozes, Juhan Bae, Siddhartha Rao Kamalakara, Dwaraknath Gnaneshwar, Acyr Locatelli, Robert Kirk, Tim Rocktäschel, Edward Grefenstette, and Max Bartolo. 2025. [Procedural knowledge in pretraining drives reasoning in large language models](#). In *The Thirteenth International Conference on Learning Representations*.
- Yang Sui, Yu-Neng Chuang, Guanchu Wang, Jiamu Zhang, Tianyi Zhang, Jiayi Yuan, Hongyi Liu, Andrew Wen, Shaochen Zhong, Na Zou, Hanjie Chen, and Xia Hu. 2025. [Stop overthinking: A survey on efficient reasoning for large language models](#). *Transactions on Machine Learning Research*.
- John Sweller. 1988. [Cognitive load during problem solving: Effects on learning](#). *Cogn. Sci.*, 12(2):257–285.
- Shaobo Wang, Xiangqi Jin, Ziming Wang, Jize Wang, Jiajun Zhang, Kaixin Li, Zichen Wen, Zhong Li, Conghui He, Xuming Hu, and Linfeng Zhang. 2025a. [Data whisperer: Efficient data selection for task-specific LLM fine-tuning via few-shot in-context learning](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 23287–23305, Vienna, Austria. Association for Computational Linguistics.
- Xinyi Wang, Shawn Tan, Mingyu Jin, William Yang Wang, Rameswar Panda, and Yikang Shen. 2025b. [Do larger language models imply better generalization? a pretraining scaling law for implicit reasoning](#). In *ICML 2025 Workshop on Methods and Opportunities at Small Scale*.
- Yudong Wang, Zixuan Fu, Jie Cai, Peijun Tang, Hongya Lyu, Yewei Fang, Zhi Zheng, Jie Zhou, Guoyang Zeng, Chaojun Xiao, Xu Han, and Zhiyuan Liu. 2025c. [Ultra-fineweb: Efficient data filtering and verification for high-quality llm training data](#). *Preprint*, arXiv:2505.05427.
- Alexander Wettig, Aatmik Gupta, Saumya Malik, and Danqi Chen. 2024. [Qurating: Selecting high-quality data for training language models](#). In *ICLR 2024 Workshop on Mathematical and Empirical Understanding of Foundation Models*.
- Yuyang Wu, Yifei Wang, Ziyu Ye, Tianqi Du, Stefanie Jegelka, and Yisen Wang. 2026. [When more is less: Understanding chain-of-thought length in LLMs](#). In *The Fourteenth International Conference on Learning Representations*.
- C. Xiao, J. Cai, W. Zhao, and 1 others. 2025. [Densifying law of LLMs](#). *Nature Machine Intelligence*, 7:1823–1833.
- Yixin Yang, Qingxiu Dong, Linli Yao, Fangwei Zhu, and Zhifang Sui. 2025. [Icon: In-context contribution for automatic data selection](#). *Preprint*, arXiv:2505.05327.
- Yixin Ye, Zhen Huang, Yang Xiao, Ethan Chern, Shijie Xia, and Pengfei Liu. 2025. [LIMO: Less is more for reasoning](#). In *Second Conference on Language Modeling*.
- Qianjin Yu, Keyu Wu, Zihan Chen, and 1 others. 2025. [Rethinking the generation of high-quality cot data from the perspective of llm-adaptive question difficulty grading](#). *Preprint*, arXiv:2504.11919.
- Weihao Yu, Zihang Jiang, Yanfei Dong, and Jiashi Feng. 2020. [Reclor: A reading comprehension dataset requiring logical reasoning](#). In *International Conference on Learning Representations (ICLR)*.
- Zichun Yu, Spandan Das, and Chenyan Xiong. 2024. [MATES: Model-aware data selection for efficient pretraining with data influence models](#). In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.
- Ranchi Zhao, Zhen Leng Thai, Yifan Zhang, and 1 others. 2024. [DecorateLM: Data engineering through corpus rating, tagging, and editing with language models](#). In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 1401–1418, Miami, Florida, USA. Association for Computational Linguistics.
- Chunting Zhou, Pengfei Liu, Puxin Xu, Srinivasan Iyer, Jiao Sun, Yuning Mao, Xuezhe Ma, Avia Efrat, Ping Yu, Lili Yu, Susan Zhang, Gargi Ghosh, Mike Lewis, Luke Zettlemoyer, and Omer Levy. 2023. [LIMA: less is more for alignment](#). In *NeurIPS*.

## A Impact, LLM Usage and Future Work

We acknowledge that although our **SLKD** score and *re-cognizing optimization* method are effective for several tasks in our research, they are far from being perfect. Here, we honestly discuss the limitations, broader impact, and potential avenues for future works.

### A.1 The Use of Large Language Models

We hereby state that LLMs were deployed purely for translation and linguistic refinement tasks. All research ideas, experimental protocols, execution processes, analytical work, and derived conclusions are the exclusive responsibility of the authors. A comprehensive validation of the manuscript has been performed to guarantee that the employment of LLMs did not introduce any false or misleading information.

### A.2 Impact

Our **SLKD** score is a pioneering metric that quantifies each example’s reasoning potential. Paired with our *re-cognizing optimization* framework, which uses this score to guide training, our approach reduces unnecessary computation, boosts model reasoning performance, and supports more sustainable AI practices. We believe this work will inspire new directions and offer systematic guidance for future research on unlocking the latent potential of LLM training data. Societally, our method can positively influence the efficient training of LLMs, enabling the creation of more robust and resource-efficient AI systems.

### A.3 Future Work

**Broader model and task coverage.** In future work, it would be valuable to evaluate our framework on additional architectures—beyond LLaMA-7B and Qwen2.5-7B—and across new domains such as mathematical reasoning, code generation, and multimodal understanding. This broader testing would help establish the generality and limits of *Structured Logical Knowledge Density* and *re-cognizing optimization*.

**Refined reinforcement learning integration.** While we have already incorporated reasoning-intensity scores into GRPO rewards, it would be useful to explore more sophisticated applications—such as dynamic reward shaping, alternative RL algorithms, or multi-objective formulations—to further boost learning efficiency and stability.

**Adaptive reasoning routing.** We also plan to investigate using reasoning-intensity as a runtime signal to guide the model’s choice of reasoning mode, enabling dynamic switching between fast, heuristic processing (System 1) and deeper, deliberative reasoning (System 2). This may prevent overthinking on trivial inputs and ensure adequate effort on challenging ones.

**Compatibility with alternative data-selection methods.** Our *Structured Logical Knowledge Density* metric and *re-cognizing optimization* are fully compatible with other sampling strategies. In future work, we will experiment with hybrid schemes that combine our method with these complementary approaches to maximize sample efficiency and model performance and to provide a more rigorous comparison against established data-selection techniques.

**Pretraining-stage integration.** An exciting direction is to extend reasoning-intensity guidance into the pretraining curriculum. Although pretraining multi-billion-parameter models from scratch exceeds our current resources, future work could integrate **SLKD** into early model training to shape core reasoning capabilities, contingent on access to sufficient compute.

---

#### Algorithm 1 SLKD Calculation

---

- 1: **Input:** A sample  $x$  with context  $c$  and options  $\{o_l\}_{l=1}^L$ .
  - 2: Parse context  $c$ :
  - 3: Extract logical expressions  $E$  and compute their nesting depths.
  - 4: Identify predicates  $P$  and constants  $C$  in  $c$ .
  - 5: Compute context intensity:
  - 6:  $S_{\text{ctx}} = |E| \times \bar{D}^2 + |P| + |C|$ .
  - 7: **for each** option  $o_l$  **do**
  - 8: Extract preconditions  $R_l$  and reasoning steps  $S_l$ .
  - 9: Compute option intensity:
  - 10:  $S_{\text{opt}}^{(l)} = |R_l| \cdot \bar{D}_l^2 + \sum_{k=1}^{T_l} (1 + \#\text{Operations}_{l,k}) D_{l,k}^2$ .
  - 11: **end for**
  - 12: Aggregate raw intensity:
  - 13:  $S_{\text{raw}} = S_{\text{ctx}} + \sum_{l=1}^L S_{\text{opt}}^{(l)}$ .
  - 14: Normalize to  $[0, 1]$  via sigmoid of log:
  - 15:  $S = \sigma \left( \gamma \cdot \frac{\log(S_{\text{raw}} + 1) - \mu}{\sqrt{\delta^2 + \epsilon}} + \beta \right)$ .
  - 16: **Output:** reasoning-intensity score  $S$ .
-

---

**Algorithm 2** Re-Cognizing Optimization

---

```
1: Input: dataset  $D = \{x_i\}_{i=1}^N$ , model  $M$ ,  
   epochs  $T$ .  
2: Precompute intensity scores  $\{S_i\}_{i=1}^N$  via Algo-  
   rithm 1.  
3: for epoch  $t = 1$  to  $T$  do  
4:   if  $t = 1$  then  
5:     Uniformly shuffle  $D$  for initial explo-  
     ration (Phase I: Model Cognition Reshaping).  
6:   else  
7:     Sort  $D$  by descending  $S_i$  to emphasize  
     high-intensity samples (Phase II: Cognitive  
     Reasoning Enhancement).  
8:   end if  
9:   for each batch  $B$  drawn sequentially from  
      $D$  do  
10:    Compute loss on  $B$  and update model  
    parameters.  
11:   end for  
12: end for  
13: Output: fine-tuned model  $M^*$ .
```

---

## B SLKD Calculation Process and Re-Cognizing Optimization Algorithm

Algorithm 1 illustrates the calculation process of our SLKD score, and Algorithm 2 demonstrates the implementation procedure of our Re-Cognizing Optimization strategy.

## C Details for the Experiments

### C.1 Experimental Settings

In this section, we conduct experiments on four logical reasoning benchmarks: *Reclor* (contextual reasoning) (Yu et al., 2020), *LogicBench* (deductive reasoning) (Parmar et al., 2024), *LogiQA* (workplace logic analysis) (Liu et al., 2020), and *LogiQA2.0* (enhanced adversarial patterns) (Liu et al., 2023). For our experiments, we utilize two backbone models: *LLaMA3.1-8B-Instruction* (Grattafiori et al., 2024) and *Qwen2.5-7B-Instruction* (Qwen et al., 2025). We refer to them hereafter as *LLaMA3.1-8B* and *Qwen2.5-7B*. Additionally, we include GPT-4 (OpenAI, 2023) and DeepSeek-V3 (DeepSeek-AI, 2024) as reference models. For more implementation details, please refer to the Appendix C.

The released code for this work is available at [https://github.com/HuzhouNLP/Logical\\_Structure\\_as\\_Knowledge](https://github.com/HuzhouNLP/Logical_Structure_as_Knowledge).

### C.2 Training Configuration

Our datasets consist of 36,788 training instances and 4,743 test instances, covering a broad range of logical challenges, including propositional logic, syllogisms, and temporal reasoning.

Both *LLaMA3.1-8B* and *Qwen2.5-7B* are fine-tuned using supervised fine-tuning (SFT) under the TRL framework, employing LoRA modules with a rank of 64,  $\alpha = 16$ , and a dropout rate of 0.05. In addition to the SFT experiments, we also conduct reinforcement learning (RL) experiments on *Qwen2.5-7B*. For RL, we apply the GRPO algorithm under the TRL framework, with LoRA configured as rank 16,  $\alpha = 32$ , and dropout rate 0.05. Due to resource constraints, we used a random subset of 1,000 training samples for the training. All evaluations across both settings are performed in a zero-shot manner.

For both direct training and *re-cognizing optimization*, we use the Paged AdamW optimizer with a learning rate of  $2e-4$ ,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.95$ , perform gradient clipping at 0.3 with a warmup ratio of 10%, set the global batch size to 16 (8 per device  $\times$  2 accumulation steps), and conduct bfloat16 mixed-precision training.

For GRPO experiments, we use the AdamW optimizer with a learning rate of  $5e-5$ , perform gradient clipping at 0.3, apply a 15% warmup over 100 steps, set weight decay to 0.01, use a global batch size of 8 (2 per device  $\times$  4 accumulation steps), enable fp16 mixed-precision training, employ a cosine-with-restarts learning-rate schedule, generate 4 completions per prompt with a maximum length of 512 tokens at temperature 0.9, and scale rewards.

**Compute Resources** Our *re-cognizing optimization* experiments were run on eight NVIDIA RTX 4090d GPUs (though a single 4090d can support smaller runs), with each full training run taking approximately 6 hours. GRPO-based RL experiments used four NVIDIA A800 GPUs (minimum one A800 required) and averaged around 10 hours per run.

### C.3 Artifact License and Terms of Use.

Our work leverages existing public logical reasoning benchmarks (Reclor, LogiQA, LogiQA2.0, LogicBench) and open-source models (LLaMA-3.1-8B, Qwen2.5-7B) for experiments, adhering to their original academic use permissions. The code implementing SLKD calculation and recognition

optimization is publicly available in an open repository for non-commercial research purposes only. We do not claim ownership of any original dataset or model content, and all derived artifacts (e.g., SLKD-scored samples) are restricted to research use in line with the terms of the source materials.

#### C.4 Consistency with Intended Use.

The existing datasets employed in this work are originally designed for evaluating LLM logical reasoning capabilities, which aligns with our research goal of enhancing model reasoning via structured logical knowledge. Our re-cognizing optimization strategy uses these datasets to validate the effectiveness of SLKD, without repurposing them for unintended scenarios. For the artifacts we create (SLKD metric, optimization framework), we specify their intended use as research tools for data-centric LLM reasoning enhancement, ensuring compatibility with the original access conditions of source datasets.

#### C.5 Personally Identifying Information and Offensive Content.

The benchmark datasets used in our experiments are standard, curated NLP resources widely adopted for logical reasoning research, which do not contain intentional personally identifying information or offensive content. Our logical element extraction and SLKD calculation processes focus solely on decomposing textual logical structures (predicates, constants, logical expressions) and do not introduce, amplify, or expose sensitive attributes. All data processing steps prioritize preserving the original dataset's safety and compliance, with no use of personally identifiable or offensive content in experiments.

#### C.6 Examples of Different Score Data

##### Examples of Different Score Data

"context": "john knows how to play the piano"  
 "question": "does this entail that someone has the ability to play the piano?"  
 "answer": "yes"  
 "ReasoningIntensityScore": 0.05729995978985077

"context": "Roves had held a senior position in the Navy before taking office. One of his good friends asked him about the Navy's plan to establish a submarine base on an island. Roosevelt looked around mysteriously and asked in a low voice. Can you keep it a secret?" "Of course I can!" The friend was very sure. "So," Roosevelt said with a smile, "I can too."  
 "question": "This text tells us:"

"options": ["Detours can also achieve the goal."  
 "Humor can subtly solve problems"  
 "Adherence to principles and flexibility are not contradictory."  
 "Don't do anything to others"],  
 "answer": "1"  
 "ReasoningIntensityScore": 0.589367067922439

"context": "It is necessary to pay attention to avoiding hollowing out in the development of the service industry, but it is wrong and dangerous to think that the rapid development of modern service industry in China's economic growth will definitely lead to a hollowing out of the industry. This view of China will make China's economy lose an important window period for the rapid development of the modern service industry. In fact, the formation of an industrial structure dominated by the service industry does not mean the decline of the status of the manufacturing industry, nor does it mean "de-industrialization" "It is not the same as starting the hollowing out process of the industry."

"question": "The main emphasis of this text?"  
 "options": ["The rapid development of modern service industry cannot lead to a hollowing out of the industry"  
 "How to objectively evaluate the advantages and disadvantages of the rapid development of modern service industry"  
 "Whether it will cause industrial hollowing depends on the prosperity of the manufacturing industry"  
 "Don't worry about the hollowness of the industry and miss the opportunity to develop the service industry"],  
 "answer": "3"  
 "ReasoningIntensityScore": 0.9464210784935705

#### C.7 Reasoning Intensity Score Calculation Prompt

##### Prompt for Logical Decomposition

Instructions: Please extract the predicates and constants from the following context and create logical expressions that represent the relationships described. Format the output as a dictionary where each entry is a list of items. Follow these specific rules:

1. **Extract predicates** as core action words or relationships defining connections between entities, and output them as a list under the 'Predicates' key.
2. **Extract constants** as the specific entities or values mentioned in the context, and output them as a list under the 'Constants' key.
3. **Create logical expressions** using the extracted predicates and constants. Each logical expression should be simple and based on a single predicate. Output them as

a list under the 'Logical Expressions' key.  
 4. Ensure that the output strictly follows the format provided in the example.  
 5. Do not combine expressions using logical operators such as 'and,' 'or,' etc., unless the relationship is explicitly mentioned in the context.

Example:

[Context: If an individual consumes a significant amount of water, they will experience a state of hydration. Conversely, if excessive amounts of sugar are ingested, a sugar crash will ensue. It is known that at least one of the following statements is true: either Jane consumes ample water or she will not experience a sugar crash. However, the actual veracity of either statement remains ambiguous, as it could be the case that only the first statement is true, only the second statement is true, or both statements are true.]

[Output:

Predicates: ['Consumes(x, y)': Represents the act of 'x' consuming 'y' (e.g., an individual consuming water or sugar),

'ExperienceState(x, y)': Represents 'x' experiencing a state 'y' (e.g., hydration),

'Ingested(x, y)': Represents 'x' ingesting 'y' (e.g., excessive sugar),

'Ensue(x)': Represents that a condition 'x' follows or results (e.g., a sugar crash),

'TrueStatement(x)': Indicates that 'x' is known to be true.,

'NotExperience(x, y)': Represents 'x' not experiencing a condition 'y' (e.g., not experiencing a sugar crash).],

Constants: ['Individual': Represents a generic person in the context.,

'Water': The substance being consumed by an individual.,

'Hydration': The state that results from sufficient water consumption.,

'Sugar': A substance that can be ingested.,

'SugarCrash': The state that follows excessive sugar intake.,

'Jane': A specific person mentioned in the context.],

Logical Expressions: [Consumes(Individual, Water)',

'ExperienceState(Individual, Hydration)',

'Ingested(Individual, Sugar)',

'Ensue(SugarCrash)',

'Consumes(Jane, Water)',

'NotExperience(Jane, SugarCrash)',

'TrueStatement(Consumes(Jane, Water)  $\vee$   $\neg$ Experience(Jane, SugarCrash))'

—

\*\* Tips for Extracting Predicates, Constants, and Logical Expressions:

- Focus on identifying core action or relational words for predicates.

- Extract constants as the specific entities mentioned.

- Use variables ('x', 'y', etc.) to generalize when needed.

- Ensure logical expressions are complete and accurately reflect relationships.

\*\* Your Task:

Context: [context]

### Prompt for Combinatorial Reasoning in BQA

Instructions: Please analyze the following binary question data. Since BQA is treated as an MCQA with a single option derived from the question, perform the analysis for this single option. Extract the relevant preconditions, define the deduction target, outline the deduction steps based on the provided Predicates, Constants, and Logical Expressions, and determine whether the option is correct based on the given answer index. Follow these specific rules:

1. **\*\*Output\*\*** should be a JSON object containing only the 'option\_analysis' field, which is a list of analyses for each option.

2. For the option\_analysis, include the following fields:

- option\_index: The index of the option (0-based).

- option\_text: The text of the option (extracted from the question).

- preconditions: A list of relevant preconditions from the Logical Expressions that pertain to the option.

- deduction\_target: The abstracted logical conclusion that the option is attempting to establish.

- deduction\_steps: A step-by-step logical deduction process from the preconditions to the deduction target. Each step should include:
    - step: The step number.
    - task: A description of what is being checked or inferred in this step.
    - expression: The logical expression used in this step, enclosed in backticks.
    - result: The outcome of this step, enclosed in backticks.
    - If a deduction step cannot proceed due to unsupported premises, indicate the failure and terminate further steps.
    - is\_correct: A boolean indicating whether the option is correct (true) or not (false). This should align with the answer field. When 'answer' is equal to 'yes', the value is true, otherwise the value is false.
  - 3. **Format Requirements**:
    - The output must strictly follow the JSON structure as shown in the example. - Ensure consistency in field naming and hierarchy.
    - Do not include any additional fields or information not specified in the example.
  - 4. **Important Considerations**:
    - Only include preconditions that are directly relevant to the option being analyzed.
    - Maintain logical rigor in deduction steps, ensuring each step follows from the previous ones based on the preconditions. - Avoid including unrelated preconditions to minimize complexity and enhance clarity.
- Example:
- context: All people who regularly drink coffee are dependent on caffeine. People either regularly drink coffee or joke about being addicted to caffeine. No one who jokes about being addicted to caffeine is unaware that caffeine is a drug. Rina is either a student and unaware that caffeine is a drug, or neither a student nor unaware that caffeine is a drug. If Rina is not a person dependent on caffeine and a student, then Rina is either a person dependent on caffeine and a student, or neither a person dependent on caffeine nor a student.
  - question: Rina is a person who jokes about being addicted to caffeine or unaware that caffeine is a drug.

- options: Rina is a person who jokes about being addicted to caffeine or unaware that caffeine is a drug.
- answer: yes
- Predicates:
  - RegularlyDrink(x, y): Represents 'x' regularly drinking 'y' (e.g., a person regularly drinking coffee).
  - DependentOn(x, y): Represents 'x' being dependent on 'y' (e.g., a person dependent on caffeine).
  - JokeAbout(x, y): Represents 'x' joking about 'y' (e.g., a person joking about being addicted to caffeine).
  - UnawareThat(x, y): Represents 'x' being unaware that 'y' (e.g., a person unaware that caffeine is a drug).
  - IsStudent(x): Represents 'x' being a student.
  - IsNeither(x, y): Represents 'x' being neither 'y' (used for expressing negation of multiple conditions).
- Constants:
  - People: Generic individuals.
  - Coffee: The beverage being consumed.
  - Caffeine: The substance people can be dependent on.
  - Rina: A specific person mentioned in the context.
- Logical Expressions:
 

- DependentOn(People, Caffeine)	⇒
RegularlyDrink(People, Coffee)	
- JokeAbout(People, Caffeine)	∨
RegularlyDrink(People, Coffee)	
- UnawareThat(People, Caffeine)	⇒
JokeAbout(People, Caffeine)	
- IsStudent(Rina)	∧
UnawareThat(Rina, Caffeine)	∨
¬IsStudent(Rina)	∧
¬UnawareThat(Rina, Caffeine)	
- ¬DependentOn(Rina, Caffeine)	∧
IsStudent(Rina)	⇒
(DependentOn(Rina, Caffeine)	∧
IsStudent(Rina))	∨
¬(DependentOn(Rina, Caffeine)	∧
IsStudent(Rina))	
- Output:
  - option\_analysis:
  - **Option Index**: 0

- **Option Text**: Rina is a person who jokes about being addicted to caffeine or unaware that caffeine is a drug.

- **Preconditions**:

- JokeAbout(People, Caffeine)     ∇
- RegularlyDrink(People, Coffee)
- UnawareThat(People, Caffeine)     ⇒
- JokeAbout(People, Caffeine)
- **Deduction Target**:
- JokeAbout(Rina, Caffeine)     ∇
- UnawareThat(Rina, Caffeine)

- **Deduction Steps**:

1. **Step**: 1
- **Task**: Instantiate the general disjunction for Rina from the population-level statement.
- **Expression**:
- JokeAbout(Rina, Caffeine)     ∇
- RegularlyDrink(Rina, Coffee)
- **Result**:     Derived     from
- JokeAbout(People, Caffeine)     ∇
- RegularlyDrink(People, Coffee)
2. **Step**: 2
- **Task**: Apply the implication that joking about caffeine addiction leads to being unaware that caffeine is a drug for Rina.
- **Expression**:
- UnawareThat(Rina, Caffeine)     ⇒
- JokeAbout(Rina, Caffeine)
- **Result**: If Rina jokes about caffeine, then Rina is unaware that caffeine is a drug.
3. **Step**: 3
- **Task**: Combine the instantiated disjunction with the implication to derive the final conclusion.
- **Expression**:
- JokeAbout(Rina, Caffeine)     ∇
- UnawareThat(Rina, Caffeine)
- **Result**:     Since
- JokeAbout(Rina, Caffeine)     implies
- UnawareThat(Rina, Caffeine), the disjunction holds.
- **Is Correct**: True

Tips for Option Analysis:

- **Preconditions**: Only include logical expressions that are directly relevant to the option being analyzed. Avoid listing all possible preconditions.
- **Deduction Steps**: Ensure each step logically follows from the previous one

based on the preconditions. If a step cannot be completed due to insufficient support from the preconditions, indicate the failure and stop further deductions for that option.

- **is\_correct**: This field should be true only for the option that matches the answer field. Since BQA has only one option, is\_correct should align with the answer field.
- **Format Consistency**: Maintain the same JSON structure and field naming conventions across all options to ensure uniformity and ease of data extraction.
- **Logical Accuracy**: Ensure that all logical expressions and deductions accurately reflect the relationships defined by the predicates and constants.

Your Task:

Analyze the following Input data and generate the option\_analysis section as per the example above. Replace the xxx placeholders in the example with actual data derived from the input.

Input Data:

input\_data\_here

Please generate the option\_analysis section based on the above input data.

### Prompt for Combinatorial Reasoning in MCQA

Instructions: Please analyze the following multiple-choice question data. For each option, extract the relevant preconditions, define the deduction target, outline the deduction steps based on the provided Predicates, Constants, and Logical Expressions, and determine whether the option is correct according to the given answer index. Follow these specific rules:

1. The **Output** should be a JSON object containing only the 'option\_analysis' field, which is a list of analyses for each option.
2. For each option in 'option\_analysis', include the following fields:
  - **option\_index**: The index of the option (starting from 0).
  - **option\_text**: The text content of the option.
  - **preconditions**: A list of relevant preconditions from the Logical Expressions

that pertain to the option.

- **deduction\_target**: The abstracted logical conclusion that the option is attempting to establish.
- **deduction\_steps**: A step-by-step logical deduction process from the preconditions to the deduction target. Each step should include:
  - **step**: The step number.
  - **task**: A description of what is being checked or inferred in this step.
  - **expression**: The logical expression used in this step, enclosed in backticks. Here, logical symbols like "implies" is represented as " $\Rightarrow$ ", "and" as " $\wedge$ ", "or" as " $\vee$ ", "not" as " $\neg$ ", "for all" as " $\forall$ ", "there exists" as " $\exists$ " in LaTeX notation.
  - **result**: The outcome of this step, enclosed in backticks.
- If a deduction step cannot proceed due to unsupported premises, indicate the failure and terminate further steps for that option.
- **is\_correct**: A boolean indicating whether the option is correct (true) or not (false). The values of 'Answer' are '0', '1', '2', '3', where '0' represents the first option, '1' represents the second option, and so on. If the value of 'option\_index' is the same as the value of 'Answer', then the value of 'is\_correct' is 'true'.

### 3. **Format Requirements**:

- The output must strictly follow the JSON structure as shown in the example.
- Ensure consistency in field naming and hierarchy.
- Do not include any additional fields or information not specified in the example.

### 4. **Important Considerations**:

- Only include preconditions that are directly relevant to the option being analyzed.
  - Maintain logical rigor in deduction steps, ensuring each step follows from the previous ones based on the preconditions.
  - Avoid including unrelated preconditions to minimize complexity and enhance clarity.
- Example:

- **Context**: In rheumatoid arthritis, the body's immune system misfunctions by attacking healthy cells in the joints causing the release of a hormone that in turn causes pain and swelling. This hormone is nor-

mally activated only in reaction to injury or infection. A new arthritis medication will contain a protein that inhibits the functioning of the hormone that causes pain and swelling in the joints. - **Question**: The statements above, if true, most strongly support which one of the following conclusions?

- **Options**: 1. Unlike aspirin and other medications that reduce pain and swelling and that are currently available, the new medication would repair existing cell damage that had been caused by rheumatoid arthritis.
- 2. A patient treated with the new medication for rheumatoid arthritis could sustain a joint injury without becoming aware of it.
- 3. Joint diseases other than rheumatoid arthritis would not be affected by the new medication.
- 4. The benefits to rheumatoid arthritis sufferers of the new medication would outweigh the medication's possible harmful side effects.

- **Answer**: 1

- **Predicates**: -  $\text{Attack}(x, y)$ : Represents 'x' attacking 'y' (e.g., the immune system attacking healthy cells).

-  $\text{Release}(x, y)$ : Represents 'x' releasing 'y' (e.g., the release of a hormone).

-  $\text{Cause}(x, y)$ : Represents 'x' causing 'y' (e.g., the hormone causing pain and swelling).

-  $\text{Activate}(x, y)$ : Represents 'x' activating 'y' (e.g., the hormone being activated by injury or infection).

-  $\text{Inhibit}(x, y)$ : Represents 'x' inhibiting 'y' (e.g., the protein inhibiting the hormone).

-  $\text{Contain}(x, y)$ : Represents 'x' containing 'y' (e.g., the medication containing a protein).

- **Constants**: - **ImmuneSystem**: The body's defense mechanism.

- **HealthyCells**: Cells in the joints that are not diseased.

- **Hormone**: A chemical messenger involved in causing pain and swelling.

- **Pain**: A sensation caused by the hormone.

- **Swelling**: A condition caused by the hormone.

- Injury: A condition that normally activates the hormone.
- Infection: A condition that normally activates the hormone.
- ArthritisMedication: A new medication for treating arthritis.
- Protein: A component of the medication that inhibits the hormone.
- **Logical Expressions**:
  - Attack(ImmuneSystem, HealthyCells)
  - Release(ImmuneSystem, Hormone)
  - Cause(Hormone, Pain)
  - Cause(Hormone, Swelling)
  - Activate(Injury, Hormone)
  - Activate(Infection, Hormone)
  - Inhibit(Protein, Hormone)
  - Contain(ArthritisMedication, Protein)
- Option Analysis
  1. **Option Index**: 0
    - **Option Text**: Unlike aspirin and other medications that reduce pain and swelling and that are currently available, the new medication would repair existing cell damage that had been caused by rheumatoid arthritis.
    - **Preconditions**:
      - Attack(ImmuneSystem, HealthyCells)
      - Release(ImmuneSystem, Hormone)
      - Cause(Hormone, Pain)
      - Cause(Hormone, Swelling)
    - **Deduction Target**:
      - Repair(ArthritisMedication, HealthyCellsDamage) - **Deduction Steps**:
        1. **Step 1**:
          - **Task**: Check if Attack(ImmuneSystem, HealthyCells) implies Damage(HealthyCells).
        - **Expression**:
          - Attack(ImmuneSystem, HealthyCells)  $\Rightarrow$  Damage(HealthyCells)
    - **Result**: Supported by context (immune system attacking healthy cells causes damage).
    - 2. **Step 2**:
      - **Task**: Check if Contain(ArthritisMedication, Protein) and Inhibit(Protein, Hormone) imply Repair(ArthritisMedication, HealthyCellsDamage).

- **Expression**:
  - Contain(ArthritisMedication, Protein)  $\wedge$
  - Inhibit(Protein, Hormone)  $\Rightarrow$
  - Repair(ArthritisMedication, HealthyCellsDamage)
- **Result**: Not supported. The context only states that the protein inhibits the hormone, not that it repairs damage.
- **Step 3**:
  - **Task**: Derivation fails.
  - **Expression**: Derivation cannot proceed.
  - **Result**: Repair(ArthritisMedication, HealthyCellsDamage) cannot be derived from the given preconditions.
  - **Is Correct**: False
- 2. **Option Index**: 1
  - **Option Text**: A patient treated with the new medication for rheumatoid arthritis could sustain a joint injury without becoming aware of it.
  - **Preconditions**:
    - Cause(Hormone, Pain)
    - Activate(Injury, Hormone)
    - Inhibit(Protein, Hormone)
    - Contain(ArthritisMedication, Protein)
  - **Deduction Target**:  $\exists$ Patient, Injury : [Sustain(Patient, Injury)  $\wedge$  Unaware(Patient, Injury)]
  - **Deduction Steps**:
    1. **Step 1**:
      - **Task**: Determine the effect of Inhibit(Protein, Hormone) from the medication.
      - **Expression**:
        - Inhibit(Protein, Hormone)  $\Rightarrow$
        - $\neg$ Cause(Hormone, Pain)  $\wedge$
        - $\neg$ Cause(Hormone, Swelling)
      - **Result**: Supported by context: The protein inhibits the hormone that causes pain and swelling.
      - 2. **Step 2**:
        - **Task**: Analyze the implication of inhibiting the hormone on pain and swelling.
        - **Expression**:
          - Inhibit(Protein, Hormone)  $\Rightarrow$
          - $\neg$ Cause(Hormone, Pain)  $\wedge$
          - $\neg$ Cause(Hormone, Swelling)
        - **Result**: Supported by context: If the hormone is inhibited, it cannot cause pain and swelling.
        - 3. **Step 3**:

- **Task**: Infer the patient's awareness of injury when pain and swelling are absent.
- **Expression**:  
 $\neg \text{Cause}(\text{Hormone}, \text{Pain}) \wedge \neg \text{Cause}(\text{Hormone}, \text{Swelling}) \Rightarrow \text{Unaware}(\text{Patient}, \text{Injury})$
- **Result**: Supported by context: Without pain and swelling, the patient may not be aware of sustaining an injury.
- **Step 4**:  
**Task**: Combine the above implications to conclude the deduction target.
- **Expression**:  $\exists \text{Patient}, \text{Injury} : [\text{Sustain}(\text{Patient}, \text{Injury}) \wedge \text{Unaware}(\text{Patient}, \text{Injury})]$  - **Result**: Deduction is valid based on the inhibited hormone preventing awareness of injury.
- **Is Correct**: True
- 3. **Option Index**: 2  
**Option Text**: Joint diseases other than rheumatoid arthritis would not be affected by the new medication.
- **Preconditions**:  
- Inhibit(Protein, Hormone)  
- Contain(ArthritisMedication, Protein)  
- **Deduction Target**:  
 $\forall x [\text{JointDisease}(x) \wedge x \neq \text{RheumatoidArthritis}] \Rightarrow \neg \text{Affect}(\text{ArthritisMedication}, x)$
- **Deduction Steps**:  
**Step 1**:  
**Task**: Check if the context provides information about other joint diseases.  
**Expression**:  $\text{JointDisease}(x) \wedge x \neq \text{RheumatoidArthritis}$   
**Result**: Not supported. The context only discusses rheumatoid arthritis.  
**Step 2**:  
**Task**: Determine if there is any implication that the medication specifically targets rheumatoid arthritis.  
**Expression**:  
 $\text{Contain}(\text{ArthritisMedication}, \text{Protein}) \Rightarrow \text{SpecificEffect}(\text{RheumatoidArthritis})$   
**Result**: Not supported. The context does not specify that the protein exclusively affects rheumatoid arthritis.  
**Step 3**:  
**Task**: Derivation fails.

- **Expression**: Derivation cannot proceed.
- **Result**: Cannot conclude that the medication does not affect other joint diseases.
- **Is Correct**: False
- 4. **Option Index**: 3  
**Option Text**: The benefits to rheumatoid arthritis sufferers of the new medication would outweigh the medication's possible harmful side effects.
- **Preconditions**:  
- Cause(Hormone, Pain)  
- Cause(Hormone, Swelling)  
- Inhibit(Protein, Hormone)  
- Contain(ArthritisMedication, Protein)  
- **Deduction Target**:  
 $\text{Benefit}(\text{ArthritisMedication}) > \text{HarmfulSideEffect}(\text{ArthritisMedication})$
- **Deduction Steps**:  
**Step 1**:  
**Task**: Identify the benefits of the medication based on inhibiting the hormone.  
**Expression**:  
 $\text{Inhibit}(\text{Protein}, \text{Hormone}) \Rightarrow \text{Reduce}(\text{Pain}) \wedge \text{Reduce}(\text{Swelling})$
- **Result**: Supported by context: The protein inhibits the hormone, which causes pain and swelling.
- **Step 2**:  
**Task**: Determine if the context provides information about harmful side effects.  
**Expression**:  
 $\text{HarmfulSideEffect}(\text{ArthritisMedication})$
- **Result**: Not supported. The context does not mention any side effects of the medication.
- **Step 3**:  
**Task**: Derivation fails.
- **Expression**: Derivation cannot proceed.
- **Result**: Cannot compare benefits and harmful side effects due to lack of information on side effects.
- **Is Correct**: False

Tips for Option Analysis

- **Preconditions**: Only include logical expressions that are directly relevant to the option being analyzed. Avoid listing all possible preconditions.

- **Deduction Steps**: Ensure each step logically follows from the previous one based on the preconditions. If a step cannot be completed due to insufficient support from the preconditions, indicate the failure and stop further deductions for that option.
- **is\_correct**: This field should be true only for the option that matches the answer index. All other options should be false.
- **Format Consistency**: Maintain the same JSON structure and field naming conventions across all options to ensure uniformity and ease of data extraction.
- **Logical Accuracy**: Ensure that all logical expressions and deductions accurately reflect the relationships defined by the predicates and constants.

Your Task

Analyze the following Input data and generate the option\_analysis section as per the example above. Replace the xxx placeholders in the example with actual data derived from the input.

Input Data

input\_data\_here

Please generate the option\_analysis section based on the above input data.

If at least one of these statements is true, can we conclude that essentially "Noah takes a vacation" or "Noah works overtime" must be true?

Your response: <think>Step 1: Context Analysis: Let V="takes vacation", O="works overtime", R="relaxes", T="has enough time to relax". The context gives implications "if V is true, then R is true" and "if O is true, then T is false", and the disjunction "R is false" or "T is true". Step 2: Question Analysis: The question asks whether V or O must always be true given ?R or T. Step 3: Answer analysis: First, from "if V is true, then R is true", we can infer that "if R is false, then V is false". From "if O is true, then T is false", we can infer that "if T is true, then O is false". So, "R is false" or "T is true" implies that "V is false" or "O is false". Thus, V and O cannot both be true, and it's possible for both V and O to be false. Therefore, it's not certain that V or O must be true.</think><answer>no</answer>

Now solve this BQA problem:

Context:

Question:

Your response:

#### Prompt for GRPO training in BQA

You are solving a yes/no (BQA) problem. Follow these rules exactly: 1. First, think through your reasoning and enclose it in <think> and </think>. 2. Then, give your final answer ("yes" or "no") enclosed in <answer> and </answer>. 3. Do not output anything else; any deviation yields zero reward. 4. Your <think> block should not exceed 450 tokens. 5. Do not include <answer> inside your <think>, and vice versa. 6. Structure your reasoning in three steps in <think> block: Step 1: Analyze the Context. Step 2: Analyze the Question. Step 3: Deduce the Conclusion.

Example: Context: If Noah takes a vacation, then he will relax. If Noah works overtime, then he will not have enough time to relax. We know at least one of the following is true: he won't relax or he will have enough time to relax. We do not know which. Question:

#### Prompt for GRPO training in MCQA

You are solving a multiple-choice (MCQA) problem with four options (A-D). Follow these rules exactly: 1. First, think through your reasoning and enclose it in <think> and </think>. 2. Then, give your final answer ("A", "B", "C", or "D") enclosed in <answer> and </answer>. 3. Do not output anything else; any deviation yields zero reward. 4. Your <think> block should not exceed 450 tokens. 5. Do not include <answer> inside your <think>, and vice versa. 6. Structure your reasoning in three steps in <think> block: Step 1: Analyze the Context. Step 2: Analyze the Question. Step 3: Analyze the Options and deduce the best choice.

Example: Context: In recent years, many cabinetmakers have been winning acclaim as artists. But furniture must be useful, so cabinetmakers focus on utility, imply-

ing cabinetmaking is not art. Question: Which assumption supports the conclusion that cabinetmaking is not art? Options: A. Some furniture is made purely for display. B. Artists are not concerned with monetary value. C. Cabinetmakers should focus more on practical utility. D. Paying attention to utility disqualifies an object as art.

Your response: <think>Step 1: Context Analysis: The passage states that because furniture must be useful, cabinetmakers must prioritize utility, so their work cannot be art. Step 2: Question Analysis: We need the hidden premise that links utility focus to art classification. Step 3: Options Analysis: Option A: Irrelevant; museums display does not address utility vs art. Option B: Off-topic; monetary concern is not mentioned. Option C: Restates the problem but does not link utility to disqualification of art. Option D: Directly asserts that focusing on utility means an object is not art, exactly matching the conclusion. Choice: Option D is the clear support.</think><answer>D</answer>

Now solve this MCQA problem:

Context:

Question:

Options:

Your response: