Novelty Assessment Report

Paper: Expanding Reasoning Potential in Foundation Model by Learning Diverse Chains of Thought Patterns

PDF URL: https://openreview.net/pdf?id=3FQV4JHPpY

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-29

Abstract

Recent progress in large reasoning models for challenging mathematical reasoning has been driven by reinforcement learning (RL). Incorporating long chain-of-thought (CoT) data during mid-training has also been shown to substantially improve reasoning depth. However, current approaches often utilize CoT data indiscriminately, leaving open the critical question of which data types most effectively enhance model reasoning capabilities. In this paper, we define the foundation model's reasoning potential for the first time as the inverse of the number of independent attempts required to correctly answer the question, which is strongly correlated with the final model performance. We then propose utilizing diverse data enriched with high-value reasoning patterns to expand the reasoning potential. Specifically, we abstract atomic reasoning patterns from CoT sequences, characterized by commonality and inductive capabilities, and use them to construct a core reference set enriched with valuable reasoning patterns. Furthermore, we propose a dual-granularity algorithm involving chains of reasoning patterns and token entropy, efficiently selecting high-value CoT data (CoTP) from the data pool that aligns with the core set, thereby training models to master reasoning effectively. Only 10B-token CoTP data enables the 85A6B Mixture-of-Experts (MoE) model to improve by **9.58\%** on the challenging AIME 2024 and 2025, and to raise the upper bound of downstream RL performance by **7.81\%**.

Core Task Landscape

This paper addresses: **Selecting High-Value Chain-of-Thought Data for Mathematical Reasoning**

A total of **50 papers** were analyzed and organized into a taxonomy with **19 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Chain-of-Thought Generation and Prompting Techniques**
- **Reasoning Verification and Supervision**
- **Training and Optimization for Mathematical Reasoning**
- **Search and Inference-Time Optimization**
- **Evaluation and Analysis of Mathematical Reasoning**

Complete Taxonomy Tree

- Selecting High-Value Chain-of-Thought Data for Mathematical Reasoning Survey Taxonomy
- Chain-of-Thought Generation and Prompting Techniques
  - Prompting Strategy Design (3 papers)
  - [1] Boosting language models reasoning with chain-of-knowledge prompting (Wang, 2024) View paper
  - [6] Mathprompter: Mathematical reasoning using large language models (Imani, 2023) View paper
  - [48] Explanation Selection Using Unlabeled Data for Chain-of-Thought Prompting (Xi Ye, 2023) View paper
  - Adaptive and Hybrid Reasoning Approaches (2 papers)
  - [5] Teaching LLMs According to Their Aptitude: Adaptive Reasoning for Mathematical Problem Solving (Xu Xin, 2025) View paper
  - [40] Enhancing Mathematical Problem Solving in Large Language Models through Tool-Integrated Reasoning and Python Code Execution (Siyue LI, 2024) View paper
  - Structured and Formal Reasoning (2 papers)
  - [9] Faithful Chain-of-Thought Reasoning (Lyu Qing, 2023) View paper
  - [35] Safe: Enhancing Mathematical Reasoning in Large Language Models via Retrospective Step-aware Formal Verification (Liu Cheng-wu, 2025) View paper
  - Multi-Step Planning and Decomposition (2 papers)
  - [12] PLAN-TUNING: Post-Training Language Models to Learn Step-by-Step Planning for Complex Problem Solving (Parmar, 2025) View paper
  - [18] Q*: Improving Multi-step Reasoning for LLMs with Deliberative Planning (Wang Chaojie, 2024) View paper
- Reasoning Verification and Supervision
  - Process Supervision Methods (6 papers)
  - [7] Let's Verify Step by Step (Lightman, 2023) View paper
  - [8] Improve Mathematical Reasoning in Language Models by Automated Process Supervision (Luo, 2024) View paper
  - [10] GenPRM: Scaling Test-Time Compute of Process Reward Models via Generative Reasoning (Zhao Jian, 2025) View paper
  - [22] MM-PRM: Enhancing Multimodal Mathematical Reasoning with Scalable Step-Level Supervision (Du Ling-Xiao, 2025) View paper
  - [26] Beyond the First Error: Process Reward Models for Reflective Mathematical Reasoning (Zhaohui Yang, 2025) View paper
  - [41] Let's reward step by step: Step-Level reward model as the Navigators for Reasoning (Ma Qianli, 2023) View paper
  - Outcome-Based Verification (2 papers)

- [20] Right Is Not Enough: The Pitfalls of Outcome Supervision in Training LLMs for Math Reasoning (Guo Jiaxing, 2025) View paper
- [45] Multi-step Problem Solving Through a Verifier: An Empirical Analysis on Model-induced Process Supervision (Wang Zihan, 2024) View paper
- Error Detection and Localization (4 papers)
- [13] Processbench: Identifying process errors in mathematical reasoning (Zheng, 2025) View paper
- [36] Premise-Augmented Reasoning Chains Improve Error Identification in Math reasoning with LLMs (Mukherjee, 2025) View paper
- [42] StepMathAgent: A Step-Wise Agent for Evaluating Mathematical Processes through Tree-of-Error (Yang Shu-xun, 2025) View paper
- [50] Demystifying Faulty Code: Step-by-Step Reasoning for Explainable Fault Localization (Ratnadira Widyasari, 2024) View paper
- Self-Verification and Consistency Checking (2 papers)
- [34] Stepwise self-consistent mathematical reasoning with large language models (Zhao Zilong, 2024) View paper
- [44] SelfCheck: Using LLMs to Zero-Shot Check Their Own Step-by-Step Reasoning (Miao Ning, 2023) View paper

- Training and Optimization for Mathematical Reasoning
  - Data Synthesis and Selection ★ (3 papers)
  - [0] Expanding Reasoning Potential in Foundation Model by Learning Diverse Chains of Thought Patterns (Anon et al., 2026) View paper
  - [16] BoostStep: Boosting mathematical capability of Large Language Models via improved single-step reasoning (Zhang Beichen, 2025) View paper
  - [19] Jiuzhang3. 0: Efficiently improving mathematical reasoning by training small data synthesis models (Zhou, 2024) View paper
  - Preference and Reinforcement Learning (4 papers)
  - [11] Direct Value Optimization: Improving Chain-of-Thought Reasoning in LLMs with Refined Values (Zhang Hong-bo, 2025) View paper
  - [25] Full-step-dpo: Self-supervised preference optimization with step-wise rewards for mathematical reasoning (Xu Hui-min, 2025) View paper
  - [28] Self-evolved preference optimization for enhancing mathematical reasoning in small language models (Chakraborty, 2025) View paper
  - [33] Step-kto: Optimizing mathematical reasoning through stepwise binary feedback (Lin, 2025) View paper
  - Supervised Fine-Tuning Strategies (3 papers)
  - [23] Jt-math: A multi-stage framework for advanced mathematical reasoning in large language models (Hao Yi-fan, 2025) View paper
  - [30] Mathematical reasoning via multi-step self questioning and answering for small language models (Kaiyuan Chen, 2024) View paper
  - [43] Improving Large Language Model Fine-tuning for Solving Math Problems (Liu Yixin, 2023) View paper
- Search and Inference-Time Optimization
  - Tree Search and Planning (1 papers)
  - [2] rStar-Math: Small LLMs Can Master Math Reasoning with Self-Evolved Deep Thinking (Guan, 2025) View paper
  - Self-Correction and Refinement (5 papers)
  - [17] STRIVE: Structured Reasoning for Self-Improvement in Claim Verification (Li Jing, 2025) View paper
  - [21] Enhancing mathematical reasoning in llms by stepwise correction (Wu Zhenyu, 2025) View paper
  - [27] Stepwise verification and remediation of student reasoning errors with large language model tutors (Daheim, 2024) View paper
  - [32] Critic-CoT: Boosting the reasoning abilities of large language model via Chain-of-thoughts Critic (Zheng Xin, 2024) View paper
  - [37] S3c-Math: Spontaneous Step-level Self-correction Makes Large Language Models Better Mathematical Reasoners (Yan, 2024) View paper
  - Step-Level Guidance and Intervention (1 papers)
  - [39] First-step advantage: Importance of starting right in multi-step math reasoning (Kushal Jain, 2025) View paper
- Evaluation and Analysis of Mathematical Reasoning
  - Benchmark Datasets and Evaluation Frameworks (4 papers)
  - [4] Mathverse: Does your multi-modal llm truly see the diagrams in visual math problems? (Zhang, 2024) View paper
  - [24] Measuring Mathematical Problem Solving With the MATH Dataset (Hendrycks, 2021) View paper
  - [29] SuperCLUE-Math6: Graded Multi-Step Math Reasoning Benchmark for LLMs in Chinese (Xu Liang, 2024) View paper
  - [49] Evaluating and improving tool-augmented computation-intensive math reasoning (Zhang Beichen, 2023) View paper
  - Empirical Analysis of Reasoning Capabilities (3 papers)
  - [3] To cot or not to cot? chain-of-thought helps mainly on math and symbolic reasoning (Sprague, 2024) View paper
  - [31] Does Learning Mathematical Problem-Solving Generalize to Broader Reasoning? (Zhou Ruo-chen, 2025) View paper
  - [47] Causal Sufficiency and Necessity Improves Chain-of-Thought Reasoning (Wang Zhuohan, 2025) View paper
  - Multimodal Mathematical Reasoning (1 papers)
  - [15] Corvid: Improving Multimodal Large Language Models Towards Chain-of-Thought Reasoning (Jiang JingJing, 2025) View paper
  - Survey and Review Studies (2 papers)
  - [14] A survey on mathematical reasoning and optimization with large language models (Forootani, 2025) View paper
  - [46] A survey on complex reasoning of large language models through the lens of self-evolution (He Tao, 2025) View paper
  - Backward Reasoning Approaches (1 papers)
  - [38] BackMATH: Towards backward reasoning for solving math problems step by step (S Zhang, 2025) View paper

Narrative

Core task: Selecting high-value chain-of-thought data for mathematical reasoning. The field has organized itself around five major branches that collectively address how to elicit, verify, train, and evaluate reasoning in mathematical domains. Chain-of-Thought Generation and Prompting Techniques explores methods for producing intermediate reasoning steps, ranging from zero-shot prompting strategies like MathPrompter[6] to structured generation approaches such as Chain of Knowledge[1]. Reasoning Verification and Supervision focuses on assessing the correctness of reasoning traces, with works like Verify Step by Step[7] and Automated Process Supervision[8] developing process-level reward models, while GenPRM[10] and ProcessBench[13] refine supervision signals at finer granularities. Training and Optimization for Mathematical Reasoning encompasses data synthesis, selection, and learning algorithms—including preference optimization methods like Full Step DPO[25] and Step KTO[33]—that leverage verified reasoning chains to improve model performance. Search and Inference-Time Optimization investigates test-time strategies such as rStar Math[2] and Q Star[18], which combine tree search with learned value functions to explore solution spaces more effectively. Finally, Evaluation and Analysis of

Mathematical Reasoning provides benchmarks like MATH Dataset[24] and Mathverse[4] to measure progress and understand model capabilities.

Within the Training and Optimization branch, a particularly active line of work centers on data synthesis and selection, where the challenge is to identify or generate reasoning traces that maximize learning efficiency. Diverse Chains Thought[0] addresses this by curating varied chain-of-thought examples to enhance training diversity, positioning itself alongside efforts like BoostStep[16] and Jiuzhang[19] that also emphasize strategic data curation for mathematical problem-solving. While BoostStep[16] focuses on iteratively refining step-level supervision signals and Jiuzhang[19] integrates domain-specific heuristics for Chinese mathematical reasoning, Diverse Chains Thought[0] emphasizes the breadth of reasoning patterns captured in the training set. This contrasts with approaches like Adaptive Reasoning[5], which dynamically adjusts reasoning strategies at inference time rather than curating training data upfront. The interplay between selecting high-quality training examples and designing effective verification mechanisms remains a central open question, as the value of a reasoning trace depends both on its correctness and its pedagogical utility for model learning.

Related Works in Same Category

The following **2 sibling papers** share the same taxonomy leaf node with the original paper:

1. BoostStep: Boosting mathematical capability of Large Language Models via improved single-step reasoning

Authors: Zhang Beichen, Liu Yuhong, Beichen Zhang, Dong, Xiaoyi, et al. (23 authors total) | **Year/Venue**: 2025 • arXiv.org | **URL**: View paper

Abstract

Large language models (LLMs) have demonstrated impressive ability in solving complex mathematical problems with multi-step reasoning and can be further enhanced with well-designed in-context learning (ICL) examples. However, this potential is often constrained by two major challenges in ICL: granularity mismatch and irrelevant information. We observe that while LLMs excel at decomposing mathematical problems, they often struggle with reasoning errors in fine-grained steps. Moreover, ICL examples...

Relationship Analysis

Both papers belong to the Data Synthesis and Selection category, focusing on techniques to enhance mathematical reasoning through careful data curation. They overlap in their goal of improving reasoning capabilities by selecting or utilizing high-quality chain-of-thought data, with both addressing the challenge of identifying valuable reasoning patterns. However, the original paper focuses on mid-training data selection by abstracting atomic reasoning patterns and constructing a core reference set to expand reasoning potential, while the candidate paper addresses inference-time enhancement through step-aligned in-context learning (ICL) examples that align retrieved reference steps with current reasoning steps.

2. Jiuzhang3. 0: Efficiently improving mathematical reasoning by training small data synthesis models

Authors: Zhou, Kun, Zhang Beichen, Kun Zhou, Wang JiaPeng, et al. (21 authors total) | **Year/Venue**: 2024 | **URL**: View paper

Abstract

Mathematical reasoning is an important capability of large language models~(LLMs) for real-world applications. To enhance this capability, existing work either collects large-scale math-related texts for pre-training, or relies on stronger LLMs (\eg GPT-4) to synthesize massive math problems. Both types of work generally lead to large costs in training or synthesis. To reduce the cost, based on open-source available texts, we propose an efficient way that trains a small LLM for math problem synt...

Relationship Analysis

Both papers belong to the Data Synthesis and Selection category, focusing on generating and selecting high-quality training data for mathematical reasoning. They overlap in their use of chain-of-thought data and data synthesis techniques to improve reasoning capabilities. However, the original paper focuses on abstracting atomic reasoning patterns and selecting diverse CoT data based on reasoning pattern chains and token entropy, while the candidate paper trains a small data synthesis model through knowledge distillation from GPT-4 to efficiently generate synthetic math problems, emphasizing cost-effective data generation rather than pattern-based selection from existing data.

Contributions Analysis

Overall novelty summary. The paper introduces a framework for selecting high-value chain-of-thought data to enhance mathematical reasoning in foundation models. It sits within the 'Data Synthesis and Selection' leaf of the taxonomy, which contains three papers total. This leaf addresses techniques for generating or filtering training data to improve reasoning capabilities, distinguishing itself from adjacent leaves focused on preference learning or supervised fine-tuning with fixed datasets. The relatively small number of sibling papers suggests this is a moderately explored but not overcrowded research direction within the broader training optimization landscape.

The taxonomy reveals that this work connects to several neighboring research areas. The sibling papers Diverse Chains Thought and BoostStep both tackle data curation but emphasize different aspects—diversity of examples versus iterative refinement of supervision signals. Adjacent leaves include 'Preference and Reinforcement Learning' (four papers) and 'Supervised Fine-Tuning Strategies' (three papers), which focus on training algorithms rather than data selection. The scope note clarifies that this leaf excludes training methods using fixed datasets, positioning the paper at the intersection of data engineering and model optimization for reasoning tasks.

Among the 30 candidates examined through semantic search, none were found to clearly refute any of the three contributions. For the theoretical definition of reasoning potential, 10 candidates were reviewed with zero refutable matches. Similarly, the abstraction of atomic reasoning patterns and the dual-granularity selection algorithm each had 10 candidates examined with no clear prior work overlap. This suggests that within the limited search scope, the paper's specific formulations—particularly the inverse-attempts metric for reasoning potential and the pattern-entropy dual criterion—appear relatively novel compared to the retrieved literature.

The analysis indicates that the paper's contributions occupy a distinct position within the examined literature, though the search was constrained to 30 top-K semantic matches. The absence of refutable candidates across all three contributions, combined with the moderately populated taxonomy leaf, suggests the work introduces fresh perspectives on data selection for reasoning. However, the limited search scope means potentially relevant work outside the top-30 semantic neighborhood may not have been captured, and a broader literature review could reveal additional connections or overlaps.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Theoretical definition of reasoning potential in foundation models

Description: The authors introduce a formal definition of reasoning potential as the probability that a model generates the correct answer when sampling, which is inversely related to the expected number of attempts needed to solve a question. This theoretical framework provides a principled way to measure and optimize model reasoning capabilities.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Reasoning over uncertain text by generative large language models

URL: View paper

Brief Assessment

Uncertain Text Reasoning[58] focuses on probabilistic reasoning over text with explicitly quantified uncertainty using Bayesian networks, not on measuring reasoning potential through sampling probability in foundation models for mathematical reasoning tasks.

2. Codeplan: Unlocking reasoning potential in large language models by scaling code-form planning

URL: View paper

Brief Assessment

CodePlan[51] focuses on code-form planning for multi-step reasoning tasks, not on defining or measuring reasoning potential through sampling probability. The paper does not address the theoretical framework of reasoning potential as the inverse of expected attempts needed to solve questions.

3. Soft thinking: Unlocking the reasoning potential of llms in continuous concept space

URL: View paper

Brief Assessment

Soft Thinking[52] focuses on continuous concept space reasoning through probability-weighted token embeddings, not on defining or measuring reasoning potential through sampling probability as the original paper does.

4. Understanding reasoning ability of language models from the perspective of reasoning paths aggregation

URL: View paper

Brief Assessment

Reasoning Paths Aggregation[53] focuses on understanding reasoning through aggregating indirect reasoning paths seen during pre-training, using random walk paths on knowledge/reasoning graphs. This differs from the original paper's definition of reasoning potential as sampling probability inversely related to expected attempts needed to solve questions.

5. Calibrating Large Language Models with Sample Consistency

URL: View paper

Brief Assessment

Sample Consistency[56] focuses on calibrating confidence scores through sampling consistency measures, not on defining or measuring reasoning potential as the probability of generating correct answers when sampling.

6. Applying large language models and chain-of-thought for automatic scoring

URL: View paper

Brief Assessment

CoT Automatic Scoring[54] focuses on applying chain-of-thought prompting for automatic scoring of student responses in educational assessments, not on defining or measuring reasoning potential in foundation models through sampling probability.

7. Why think step by step? reasoning emerges from the locality of experience

URL: View paper

Brief Assessment

Locality of Experience[57] focuses on why chain-of-thought reasoning helps in language models through local statistical structure of training data, not on defining or measuring reasoning potential as sampling probability. The candidate addresses when and why reasoning is effective, while the original defines a metric for reasoning capability.

8. Unlocking reasoning capabilities in llms via reinforcement learning exploration

URL: View paper

Brief Assessment

RL Exploration Reasoning[55] focuses on exploration mechanisms in reinforcement learning (forward vs. reverse KL divergence) rather than defining reasoning potential through sampling probability. The candidate does not propose a theoretical framework for measuring reasoning potential as the inverse of expected attempts needed to solve questions.

9. Self-Consistency Improves Chain of Thought Reasoning in Language Models

URL: View paper

Brief Assessment

Self Consistency[59] focuses on a decoding strategy that samples multiple reasoning paths and selects consistent answers, rather than defining reasoning potential as the probability of generating correct answers when sampling or formalizing it as the inverse of expected attempts needed.

10. What are the odds? language models are capable of probabilistic reasoning

URL: View paper

Brief Assessment

Probabilistic Reasoning[60] focuses on language models' ability to reason about probability distributions (percentiles, sampling, probabilities), not on defining reasoning potential as sampling probability for solving questions. The candidate addresses statistical inference tasks rather than the theoretical framework for measuring model reasoning capabilities through expected attempts.

Contribution 2: Abstraction of atomic reasoning patterns from CoT sequences

Description: The authors propose extracting atomic reasoning patterns that exhibit commonality and inductive capabilities from chain-of-thought data. These patterns are used to build a core reference set that approximates oracle reasoning data and guides the selection of high-value training samples.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. What makes a good reasoning chain? uncovering structural patterns in long chain-of-thought reasoning

URL: View paper

Brief Assessment

Structural Patterns CoT[66] focuses on extracting structural patterns (exploration, backtracking, verification) from reasoning trees for quality prediction, not on abstracting atomic reasoning patterns to build a core reference set for training data selection as in the original paper.

---

2. Making reasoning matter: Measuring and improving faithfulness of chain-of-thought reasoning

URL: View paper

Brief Assessment

Faithfulness CoT[63] focuses on measuring faithfulness of chain-of-thought reasoning through causal mediation analysis, not on extracting atomic reasoning patterns for training data selection. The candidate's work addresses whether models reliably use their reasoning steps, while the original paper proposes abstracting patterns to build a core reference set for data selection.

---

3. Compressing chain-of-thought in llms via step entropy

URL: View paper

Brief Assessment

Step Entropy Compression[67] focuses on identifying redundant reasoning steps through entropy metrics rather than abstracting atomic reasoning patterns. The candidate compresses CoT by pruning low-entropy steps, not by extracting commonality-based reasoning patterns for training data selection.

---

4. Multimodal chain-of-thought reasoning: A comprehensive survey

URL: View paper

Brief Assessment

Multimodal CoT Survey[62] focuses on multimodal chain-of-thought reasoning across diverse modalities (image, video, audio, 3D) and does not address the extraction of atomic reasoning patterns from CoT data for training sample selection as proposed in the original paper.

---

5. WebCoT: Enhancing Web Agent Reasoning by Reconstructing Chain-of-Thought in Reflection, Branching, and Rollback

URL: View paper

Brief Assessment

WebCoT[70] focuses on reconstructing chain-of-thought for web agent tasks through reflection, branching, and rollback mechanisms, rather than extracting atomic reasoning patterns with commonality and inductive capabilities from CoT data to build a core reference set for training sample selection.

---

6. The Curse of CoT: On the Limitations of Chain-of-Thought in In-Context Learning

URL: View paper

Brief Assessment

Curse of CoT[64] focuses on analyzing CoT's limitations in pattern-based in-context learning, examining explicit vs. implicit reasoning mechanisms. It does not propose extracting atomic reasoning patterns from CoT sequences for training data selection or building reference sets.

---

7. Learning to Rank Chain-of-Thought: Using a Small Model

URL: View paper

Brief Assessment

Learning Rank CoT[68] focuses on ranking CoT solutions using an energy-based reward model with outcome labels, not on extracting atomic reasoning patterns or building core reference sets from CoT sequences.

---

8. Demystifying long chain-of-thought reasoning in llms

URL: View paper

Brief Assessment

Long CoT Demystifying[61] focuses on the mechanics of generating long chain-of-thought trajectories through reinforcement learning and supervised fine-tuning, rather than extracting atomic reasoning patterns to build reference sets for data selection. The candidate does not address pattern abstraction or core reference set construction.

---

9. Motion-R1: Enhancing Motion Generation with Decomposed Chain-of-Thought and RL Binding

URL: View paper

Brief Assessment

Motion R1[65] focuses on decomposed chain-of-thought for motion generation tasks, not on extracting atomic reasoning patterns from CoT data for training foundation models in mathematical reasoning.

---

10. Beyond imitation: Learning key reasoning steps from dual chain-of-thoughts in reasoning distillation

URL: View paper

Brief Assessment

Dual Chain Thoughts[69] focuses on identifying key reasoning steps through dual CoT data (correct vs. incorrect reasoning paths) rather than abstracting atomic reasoning patterns with commonality and inductive capabilities for building a core reference set.

Contribution 3: Dual-granularity algorithm for selecting high-value CoT data

Description: The authors develop an algorithm using weighted Dynamic Time Warping that operates at two levels of granularity (reasoning pattern chains and token entropy) to efficiently select long chain-of-thought data from a source pool that matches valuable reasoning patterns in the core set.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Lemur: Log parsing with entropy sampling and chain-of-thought merging

URL: View paper

Brief Assessment

Lemur[73] focuses on log parsing using entropy sampling for clustering log messages and chain-of-thought for template merging, not on selecting chain-of-thought data for training reasoning models. The domains and objectives are fundamentally different.

2. Zero-Shot Cross-Domain Aspect-Based Sentiment Analysis via Domain-Contextualized Chain-of-Thought Reasoning

URL: View paper

Brief Assessment

Domain Contextualized CoT[79] focuses on cross-domain aspect-based sentiment analysis using chain-of-thought reasoning for domain adaptation, not on selecting CoT data using reasoning pattern matching and token entropy as described in the original contribution.

3. Fine-Tuning LLMs to Analyze Multiple Dimensions of Code Review: A Maximum Entropy Regulated Long Chain-of-Thought Approach

URL: View paper

Brief Assessment

Maximum Entropy Code Review[78] focuses on code review tasks using maximum entropy principles to regulate long chain-of-thought fine-tuning for LLMs, not on selecting CoT data from source pools using reasoning pattern matching and token entropy for mathematical reasoning tasks.

4. Consistency Is Not Always Correct: Towards Understanding the Role of Exploration in Post-Training Reasoning

URL: View paper

Brief Assessment

Exploration Post Training[77] focuses on post-training dynamics and exploration strategies in RL/inference scaling, not on data selection algorithms for mid-training. The candidate does not address pattern-based data selection or token entropy methods for curating training datasets.

5. Beyond the 80/20 Rule: High-Entropy Minority Tokens Drive Effective Reinforcement Learning for LLM Reasoning

URL: View paper

Brief Assessment

High Entropy Tokens[71] focuses on token entropy patterns during RL training to identify high-entropy tokens that guide reasoning pathways, not on selecting CoT data from a source pool using reasoning pattern matching and weighted DTW algorithms as described in the original paper.

6. Measuring reasoning utility in llms via conditional entropy reduction

URL: View paper

Brief Assessment

Conditional Entropy Reduction[72] focuses on measuring reasoning utility via conditional entropy to assess whether reasoning steps are productive, not on selecting training data from a pool using pattern matching and DTW alignment as described in the original contribution.

7. Compressing chain-of-thought in llms via step entropy

URL: View paper

Brief Assessment

Step Entropy Compression[67] uses step entropy for compression during inference, not for selecting training data from a source pool. The candidate does not employ weighted Dynamic Time Warping or pattern matching for data selection as described in the original contribution.

8. Prompt Mining for Language-based Human Mobility Forecasting

URL: View paper

Brief Assessment

Prompt Mining Mobility[74] focuses on prompt generation and refinement for mobility forecasting using information entropy of prompts and chain-of-thought mechanisms, not on selecting chain-of-thought training data using reasoning pattern matching and token entropy for foundation model training.

9. Entropy-Guided Tree of Thoughts: A Dynamic Approach to Diverse Path Generation in LLM Reasoning

URL: View paper

Brief Assessment

Entropy Guided ToT[76] focuses on enhancing diversity in tree-of-thoughts path generation for LLM reasoning tasks, not on selecting chain-of-thought training data from a source pool using pattern matching and DTW algorithms.

10. CTRLS: Chain-of-Thought Reasoning via Latent State-Transition

URL: View paper

Brief Assessment

CTRLS[75] focuses on modeling CoT reasoning as a Markov Decision Process with latent state transitions for exploration during generation, not on selecting training data from a pool using pattern matching and token entropy as described in the original contribution.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Expanding Reasoning Potential in Foundation Model by Learning Diverse Chains of Thought Patterns View paper
- [1] Boosting language models reasoning with chain-of-knowledge prompting View paper
- [2] rStar-Math: Small LLMs Can Master Math Reasoning with Self-Evolved Deep Thinking View paper
- [3] To cot or not to cot? chain-of-thought helps mainly on math and symbolic reasoning View paper

- [4] Mathverse: Does your multi-modal llm truly see the diagrams in visual math problems? View paper
- [5] Teaching LLMs According to Their Aptitude: Adaptive Reasoning for Mathematical Problem Solving View paper
- [6] Mathprompter: Mathematical reasoning using large language models View paper
- [7] Let's Verify Step by Step View paper
- [8] Improve Mathematical Reasoning in Language Models by Automated Process Supervision View paper
- [9] Faithful Chain-of-Thought Reasoning View paper
- [10] GenPRM: Scaling Test-Time Compute of Process Reward Models via Generative Reasoning View paper
- [11] Direct Value Optimization: Improving Chain-of-Thought Reasoning in LLMs with Refined Values View paper
- [12] PLAN-TUNING: Post-Training Language Models to Learn Step-by-Step Planning for Complex Problem Solving View paper
- [13] Processbench: Identifying process errors in mathematical reasoning View paper
- [14] A survey on mathematical reasoning and optimization with large language models View paper
- [15] Corvid: Improving Multimodal Large Language Models Towards Chain-of-Thought Reasoning View paper
- [16] BoostStep: Boosting mathematical capability of Large Language Models via improved single-step reasoning View paper
- [17] STRIVE: Structured Reasoning for Self-Improvement in Claim Verification View paper
- [18] Q*: Improving Multi-step Reasoning for LLMs with Deliberative Planning View paper
- [19] Jiuzhang3. 0: Efficiently improving mathematical reasoning by training small data synthesis models View paper
- [20] Right Is Not Enough: The Pitfalls of Outcome Supervision in Training LLMs for Math Reasoning View paper
- [21] Enhancing mathematical reasoning in llms by stepwise correction View paper
- [22] MM-PRM: Enhancing Multimodal Mathematical Reasoning with Scalable Step-Level Supervision View paper
- [23] Jt-math: A multi-stage framework for advanced mathematical reasoning in large language models View paper
- [24] Measuring Mathematical Problem Solving With the MATH Dataset View paper
- [25] Full-step-dpo: Self-supervised preference optimization with step-wise rewards for mathematical reasoning View paper
- [26] Beyond the First Error: Process Reward Models for Reflective Mathematical Reasoning View paper
- [27] Stepwise verification and remediation of student reasoning errors with large language model tutors View paper
- [28] Self-evolved preference optimization for enhancing mathematical reasoning in small language models View paper
- [29] SuperCLUE-Math6: Graded Multi-Step Math Reasoning Benchmark for LLMs in Chinese View paper
- [30] Mathematical reasoning via multi-step self questioning and answering for small language models View paper
- [31] Does Learning Mathematical Problem-Solving Generalize to Broader Reasoning? View paper
- [32] Critic-CoT: Boosting the reasoning abilities of large language model via Chain-of-thoughts Critic View paper
- [33] Step-kto: Optimizing mathematical reasoning through stepwise binary feedback View paper
- [34] Stepwise self-consistent mathematical reasoning with large language models View paper
- [35] Safe: Enhancing Mathematical Reasoning in Large Language Models via Retrospective Step-aware Formal Verification View paper
- [36] Premise-Augmented Reasoning Chains Improve Error Identification in Math reasoning with LLMs View paper
- [37] S3c-Math: Spontaneous Step-level Self-correction Makes Large Language Models Better Mathematical Reasoners View paper
- [38] BackMATH: Towards backward reasoning for solving math problems step by step View paper
- [39] First-step advantage: Importance of starting right in multi-step math reasoning View paper
- [40] Enhancing Mathematical Problem Solving in Large Language Models through Tool-Integrated Reasoning and Python Code Execution View paper
- [41] Let's reward step by step: Step-Level reward model as the Navigators for Reasoning View paper
- [42] StepMathAgent: A Step-Wise Agent for Evaluating Mathematical Processes through Tree-of-Error View paper
- [43] Improving Large Language Model Fine-tuning for Solving Math Problems View paper
- [44] SelfCheck: Using LLMs to Zero-Shot Check Their Own Step-by-Step Reasoning View paper
- [45] Multi-step Problem Solving Through a Verifier: An Empirical Analysis on Model-induced Process Supervision View paper
- [46] A survey on complex reasoning of large language models through the lens of self-evolution View paper
- [47] Causal Sufficiency and Necessity Improves Chain-of-Thought Reasoning View paper
- [48] Explanation Selection Using Unlabeled Data for Chain-of-Thought Prompting View paper
- [49] Evaluating and improving tool-augmented computation-intensive math reasoning View paper
- [50] Demystifying Faulty Code: Step-by-Step Reasoning for Explainable Fault Localization View paper
- [51] Codeplan: Unlocking reasoning potential in large language models by scaling code-form planning View paper
- [52] Soft thinking: Unlocking the reasoning potential of llms in continuous concept space View paper
- [53] Understanding reasoning ability of language models from the perspective of reasoning paths aggregation View paper
- [54] Applying large language models and chain-of-thought for automatic scoring View paper
- [55] Unlocking reasoning capabilities in llms via reinforcement learning exploration View paper
- [56] Calibrating Large Language Models with Sample Consistency View paper
- [57] Why think step by step? reasoning emerges from the locality of experience View paper
- [58] Reasoning over uncertain text by generative large language models View paper
- [59] Self-Consistency Improves Chain of Thought Reasoning in Language Models View paper
- [60] What are the odds? language models are capable of probabilistic reasoning View paper
- [61] Demystifying long chain-of-thought reasoning in llms View paper
- [62] Multimodal chain-of-thought reasoning: A comprehensive survey View paper
- [63] Making reasoning matter: Measuring and improving faithfulness of chain-of-thought reasoning View paper
- [64] The Curse of CoT: On the Limitations of Chain-of-Thought in In-Context Learning View paper
- [65] Motion-R1: Enhancing Motion Generation with Decomposed Chain-of-Thought and RL Binding View paper
- [66] What makes a good reasoning chain? uncovering structural patterns in long chain-of-thought reasoning View paper
- [67] Compressing chain-of-thought in llms via step entropy View paper
- [68] Learning to Rank Chain-of-Thought: Using a Small Model View paper
- [69] Beyond imitation: Learning key reasoning steps from dual chain-of-thoughts in reasoning distillation View paper
- [70] WebCoT: Enhancing Web Agent Reasoning by Reconstructing Chain-of-Thought in Reflection, Branching, and Rollback View paper
- [71] Beyond the 80/20 Rule: High-Entropy Minority Tokens Drive Effective Reinforcement Learning for LLM Reasoning View paper
- [72] Measuring reasoning utility in llms via conditional entropy reduction View paper
- [73] Lemur: Log parsing with entropy sampling and chain-of-thought merging View paper
- [74] Prompt Mining for Language-based Human Mobility Forecasting View paper

- [75] CTRLS: Chain-of-Thought Reasoning via Latent State-Transition View paper
- [76] Entropy-Guided Tree of Thoughts: A Dynamic Approach to Diverse Path Generation in LLM Reasoning View paper
- [77] Consistency Is Not Always Correct: Towards Understanding the Role of Exploration in Post-Training Reasoning View paper
- [78] Fine-Tuning LLMs to Analyze Multiple Dimensions of Code Review: A Maximum Entropy Regulated Long Chain-of-Thought Approach View paper
- [79] Zero-Shot Cross-Domain Aspect-Based Sentiment Analysis via Domain-Contextualized Chain-of-Thought Reasoning View paper