# Report on AI Reasoning

**Xinze Yu**

## Abstract

AI reasoning has advanced significantly, enabling impressive performance on complex tasks such as logical reasoning, mathematical problem-solving, and symbolic processing. However, challenges such as brittleness to premise ordering, limitations of self-correction, and the need for effective intermediate reasoning steps remain significant hurdles. This paper synthesizes insights from four recent studies to explore these issues, focusing on the mechanisms behind reasoning failures and the techniques developed to address them. We analyze the efficacy of self-correction in reasoning, the impact of premise order on deductive tasks, and the role of chain-of-thought (CoT) in enabling transformers to solve inherently serial problems. Theoretical findings and empirical evaluations underscore the importance of structural and decoding strategies in improving AI reasoning

## 1 Introduction

The reasoning capabilities of large language models (LLMs) are a cornerstone of their success in complex natural language processing tasks. Despite achieving near-human performance in some benchmarks, these models often exhibit limitations when applied to reasoning-intensive domains. Challenges include the inability to correct errors without external feedback, vulnerability to distractors, and the reliance on specific input formats, such as ordered premises or tailored prompts.

This report examines recent research addressing these challenges, exploring:

1. The limitations of self-correction techniques in LLM reasoning.
2. The brittleness of models to changes in premise order during logical reasoning tasks.
3. The power of chain-of-thought reasoning in improving model expressiveness and task performance.

These studies provide valuable insights into how current approaches can be improved and highlight areas requiring further exploration.

## 2 Limitations of Self-Correction in AI Reasoning

### Intrinsic vs. Extrinsic Self-Correction

Self-correction involves refining an initial response based on feedback, either internally (intrinsic) or from external sources (extrinsic). Huang et al. (2024) define intrinsic self-correction as a process where LLMs attempt to rectify errors solely using their built-in capabilities without external labels or guidance.

### Key Findings

1.**Performance Degradation**: Intrinsic self-correction often fails in reasoning tasks, with performance dropping compared to initial attempts. For example, on the GSM8K dataset (grade school math problems), the accuracy decreased after self-correction when oracle labels were unavailable(2310.01798v2).

2.**Dependence on Prompts**: The design of prompts heavily influences self-correction. Suboptimal initial prompts lead to worse results, even with corrective steps.

3.**Comparison to Baselines**: Techniques like multi-agent debate, where multiple

LLMs critique each other's responses, showed limited improvements over simpler strategies such as self-consistency when equivalent inference costs were considered.

### Practical Implications

Huang et al. argue that reliance on intrinsic self-correction is not a reliable strategy for enhancing LLM reasoning. Instead, research should focus on hybrid models that integrate external feedback mechanisms or leverage better-informed initial prompts.

---

## 3 Premise Ordering in Logical Reasoning

### Premise Order Effect

Chen et al. (2024) explored how the order of premises impacts LLM reasoning. For logical tasks, where the sequence of premises should theoretically not matter, they observed a surprising sensitivity to the ordering.

### Experimental Results

- **Accuracy Drops**: When premises were shuffled randomly, the accuracy of LLMs on deductive reasoning tasks dropped by over 30%.
- **Forward vs. Backward Chaining**: Models performed best when premises aligned with the steps required to derive the conclusion sequentially (forward chaining). Conversely, reversing the order often led to degraded performance, particularly in tasks requiring multi-step deductions(2402.08939v3).

### The R-GSM Benchmark

To investigate this phenomenon in mathematical reasoning, the authors developed the R-GSM benchmark, a variant of GSM8K. They manually reordered problem descriptions while keeping the logical structure intact. Results confirmed that LLMs struggled with reordered premises, particularly in problems requiring more reasoning steps or with irrelevant premises as distractors.

## 4 Chain-of-Thought Reasoning

### What is Chain-of-Thought(CoT)

Chain-of-thought (CoT) reasoning involves generating explicit intermediate steps during problem-solving. This approach has gained traction as a method to improve LLM performance on tasks requiring serial computation.

### Theoretical Insights

Li et al. (2024) analyzed CoT through the lens of computational complexity:

- **Expressiveness**: Standard transformers without CoT are restricted to AC0 problems (solvable with shallow, parallel computations). Incorporating CoT expands their capabilities to include problems in P/poly (polynomial-size circuits), enabling transformers to perform inherently serial computations (2402.12875v4).

- **Serial Problem Solving**: Tasks such as iterated squaring or permutation group composition, which require step-by-step reasoning, benefit significantly from CoT.

### Empirical Findings:

- **Improved Accuracy:** On synthetic tasks like circuit evaluation, CoT-enabled transformers consistently outperformed baseline models, particularly for low-depth architectures.

- **Adaptability:** Even with minimal CoT steps, transformers demonstrated enhanced capability to handle more complex reasoning tasks.

## 5 Decoding Strategies for Chain-of-Thought Reasoning

Wang et al. (2024) proposed an innovative method to elicit CoT reasoning paths without explicit prompts:

- **Improved Accuracy:** By exploring top-k alternative tokens during decoding, the authors uncovered latent CoT paths that were hidden in greedy decoding outputs.
- **Adaptability:** The presence of CoT paths correlated with higher confidence in the final answers, providing a mechanism to identify reliable reasoning trajectories (2402.10200v2).

## 6 Conclusion

The research reviewed in this report highlights critical advancements and ongoing challenges in AI reasoning:

- **Self-Correction**: Intrinsic self-correction remains limited without external feedback or optimized prompts.
- **Premise Order Sensitivity**: Logical tasks require careful design to mitigate the impact of premise order.
- **Chain-of-Thought**: CoT enhances the expressiveness and performance of transformers on serial tasks, offering a robust framework for reasoning.
- **Decoding Innovations**: Techniques like CoT decoding unlock hidden reasoning potential in LLMs, paving the way for more reliable outputs.

## References

Xuezhi Wang, Denny Zhou. 2024. *Chain-of-Thought Reasoning without Prompting.* https://doi.org/10.48550/arXiv.2402.10200. Version 2

Jie Huang, Xinyun Chen, Swaroop Mishra,etc.2023. *Large Language Model Cannot Self-correct Reasoning Yet* .ICLR 2024. https://doi.org/10.48550/arXiv.2310.02798. Version 2

Xinyun Chen, Ryan A. Chi, Xuezhi Wang, Denny Zhou.2024.*Premise Order Matters in Reasoning with Large Language Models*. ICLM 2024.https://doi.org/10.48550/arXiv2402.08939. Version 3

Zhiyuan Li, Hong Liu, Denny Zhou, Tengyu Ma.2024. *Chain of Thought Empowers Transformers to Solve Inherently Serial Problems*. https://doi.org/10.48550/arXiv2402.12875.Version 4