EasyChair Preprint

№ 15777

Meta-Cognitive Problem-Solving: a Systematic Framework for Problem-Solving Verification

Mohamed Abbas

January 29, 2025

Meta-Cognitive Problem-Solving:
A Systematic Framework for Problem-Solving Verification

Mohamed A. Abbas [1]
[1] Alexandria University, Alexandria, Egypt
E-mail: nmohamed.alyabbas@alexu.edu.eg

Summary: Computational problem-solving is often inefficient in terms of effective verification processes, introducing potential inaccuracies in problem-solving approaches. In this work, meta-cognitive methods for systematic problem-solving with an emphasis on constraint realization, invariant property analysis, and solution verification have been examined. Drawing from recent advances in AI metacognition [1], [2] and cognitive problem-solving frameworks [3], we present the importance of systematic problem-solving protocols in excluding cognitive biases and unverified statement claims for a correct solution.

Keywords: Meta-cognitive problem-solving, AI verification, constraint analysis, systematic reasoning, cognitive bias prevention.

1. Introduction

Meta-cognitive problem-solving is a significant mechanism for enhancing the accuracy of AI problem-solving techniques. As highlighted by [4], meta-cognitive capabilities are essential for effective learning and self-regulation in problem-solving contexts. Recent work emphasizes the importance of systematic problem-solving protocols in excluding cognitive biases and yielding correct solutions. This aligns with [2] TRAP framework, which emphasizes Transparency, Reasoning, Adaptation, and Perception as key components of metacognitive AI. In this work, a meta-reasoning mechanism with an emphasis on explicit constraint realization, invariant property analysis, complete state tracking, and strict solution verification is discussed in a review format.

1.1. Recent AI Problem-Solving Advances

Recent studies have emphasized several key developments in AI problem-solving:

- The importance of explicit constraint documentation in AI problem-solving, with [5] highlighting how neuro-symbolic AI approaches can improve interpretability and robustness.
- New approaches for invariant property detection have emerged, with [6] proposing systematic frameworks for navigating complex AI challenges.
- Advances in verification protocols have confirmed the significance of sequential variant-by-variant validation, as supported by research in computational metacognition [7].

2. Case Study: Light-Toggling Grid Problem

The light-toggling grid problem is considered a case study for demonstrating the efficacy of this mechanism. Following the cognitive forcing approach described by [8], our analysis reveals how initial intuitive solutions can be misleading without proper metacognitive verification. Initially, a problem appeared solvable but, with careful examination, was revealed to be mathematically unsolvable in terms of a constraint failure.Below inserts a list of up to 5 keywords/ keyphrases. For the word Keywords use Bold.

2.1. Puzzle Specification

- Grid: 5x5 grid of lights
- Initial State: 5 lights in a random position
- Objective: Turn off all lights
- Constraint: Clicking a button toggles a button and surrounding lights

```
X . . . .
. . X . .
. X . . .
. . . X .
. . . . X
```

2.2. Anatomy of a Reasoning Breakdown

Drawing from [9] work on explainable AI, our investigation revealed critical thinking fallacies:

1. Premature Assertion of a Solution
   - First instinct: Assertion of a solution first
   - Lethal flaw: Creation of a sequence of a solution with no proper testing
2. Lack of Systematic Analysis of the System
   - Repeated testing with no proper analysis of the problem
   - Inability to detect mathematical impossibility

3. Method

The proposed meta-reasoning model incorporates insights from recent research in metacognitive AI [1], [7] and is defined in terms of four critical tenets:

1. Verbatim Recording of Rules and Constraints: Recording problem constraints and rules in one's language, a critical guard against misconceptions and thinking fallacies. This aligns with the explainability principles outlined in [10].

2. Analysis of Invariant Properties: Preliminary analysis of problem solvability via invariant property analysis is conducted. This approach is supported by findings in computational metacognition [7].
3. Detailed Observational Recording of Moves: Recording in detail each move in a problem-solving exercise for transparency and detectability of any malpractice. This reflects the transparency component of [2] TRAP framework.
4. Forced Complete Demonstration of Validity of Solutions: Forcing complete demonstration of a solution's validity, incorporating the cognitive forcing principles described by [8].

3.1. Breakthroughs

Realized a critical invariant: Initial state (5 lights) proves problem mathematically unsolvable.

Demonstrated the importance of metacognitive verification in preventing overreliance on intuitive solutions [8].

Validated the effectiveness of systematic problem-solving approaches aligned with recent research in cognitive AI [5].

3.2. Suggested Systematic Thinking Prompt

Suggested Systematic Thinking Prompt
Drawing from research on metacognitive skill development [11], such a mechanism for systematic thinking is meta-cognitive thinking-dependent and seeks to prevent premature claims of a solution and cognitive bias in AI problem-solving.

In solving this problem, please:
1) Cite all explicit rules and constraints in quotes.
2) Look for invariant properties that can establish impossibility.
3) Document and verify each step, with all intermediate states.
4) Only report a solution works after proving all steps and their effect.

3.3. Practitioner Takeaways

- Always verify assumptions.
- Don't presume problem characteristics at face value.
- Implement systematic thinking protocols aligned with metacognitive frameworks [2].

4. AI Implications

The proposed mechanism builds on recent advances in cognitive AI [12] and has important implications for developing AI systems that can engage in sophisticated problem-solving while maintaining explainability and verification capabilities. It is useful in developing stronger computational problem-solving methods as well.

4.1 Limitations and Future Directions

The Based on current research trends in metacognitive AI [1], [5] future work entails researching:

- Scalability of the proposed mechanism to complex thinking domains.
- Automated invariant property checking tools.
- Integration with neuro-symbolic approaches for enhanced reasoning capabilities.
- Adaptive thinking verification protocols.

References

[1] S. G. B. Johnson, A.-H. Karimi, Y. Bengio, N. Chater, T. Gerstenberg, K. Larson, S. Levine, M. Mitchell, I. Rahwan, B. Schölkopf and I. Grossmann, "Imagining and building wise machines: The centrality of AI metacognition," *arXiv preprint arXiv:2411.02478,* 2024.

[2] H. Wei, P. Shakarian, C. Lebière, B. A. Draper, N. Krishnaswamy and S. Nirenburg, "Metacognitive AI: Framework and the Case for a Neurosymbolic Approach," *arXiv preprint arXiv:2406.12147,* 2024.

[3] M. Eppe, C. Gumbsch, M. Kerzel, P. D. H. Nguyen, M. V. Butz and S. Wermter, "Intelligent problem-solving as integrated hierarchical reinforcement learning," *Nature Machine Intelligence,* vol. 4, no. 1, pp. 11-20, 2022.

[4] R. K. Verma, M. Dadhich, D. Mathur and A. Sharma, "Empirical Study of Data-Driven Learning and Generative AI in Enhancing Meta-Cognitive Resource Utilization: A Comprehensive Analysis," *Research Review International Journal of Multidisciplinary,* vol. 9, no. 9, pp. 58-67, 2024.

[5] Z. Wan, C.-K. Liu, H. Yang, C. Li, H. You, Y. Fu, C. Wan, T. Krishna, Y. ". Lin and A. Raychowdhury, "Towards Cognitive AI Systems: a Survey and Prospective on Neuro-Symbolic AI," *arXiv preprint arXiv:2401.01040,* 2024.

[6] D. M. Geetha, D. C. K. Gomathy, M. S. C. Manvith and D. Balatripurasundari, "Navigating the AI landscape: a systematic guide to solving complex challenges," *Indian Scientific Journal of Research in Engineering and Management,* vol. 7, no. 11, pp. 1-11, 2023.

[7] M. T. Cox, Z. Mohammad, S. Kondrakunta, V. R. Gogineni, D. Dannenhauer and O. Larue, "Computational Metacognition," *arXiv preprint arXiv:2201.12885,* 2022.

[8] Z. Bučinca, M. B. Malaya and K. Z. Gajos, "To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1-21, 2021.

[9] J. E. T. Taylor and G. W. Taylor, "Artificial cognition: How experimental psychology can help generate explainable artificial intelligence," *Psychonomic Bulletin & Review*, vol. 28, no. 2, pp. 454-475, 2021.

[10] C. Lee, H.-Y. Kwon and K. Jin, "Human Cognition for Mitigating the Paradox of AI Explainability: A Pilot Study on Human Gaze-based Text Highlighting," *Proceedings of the 37th International FLAIRS Conference*, 2024.

[11] Y. Yang and N. Xia, "Enhancing Students' Metacognition via AI-Driven Educational Support Systems," *International Journal of Emerging Technologies in Learning*, 2023.

[12] N. Spivack, S. Douglas, M. Crames and T. Connors, "Cognition is All You Need - The Next Layer of AI Above Large Language Models," *arXiv preprint arXiv:2403.02164*, 2024.