**A NotSo Simple way to beat Simple Bench**

Soham Sane, Collins Aerospace

**Date:** [Final Peer Review Date]

**ABSTRACT**

This paper presents an innovative approach to solving reasoning problems using prompt engineering and iterative refinement with GPT-based models. Through a case study on the “SimpleBench” benchmark, we highlight the challenges of addressing implicit constraints in problem statements and demonstrate how an iterative reasoning framework, equipped with feedback gates and global consistency checks, enhances the performance of a base model. We also examine the psychological alignment of reasoning processes, such as the human tendency to ignore irrelevant details, and propose improvements to the feedback and consistency mechanisms. Our findings suggest pathways to further optimize reasoning systems for both accuracy and contextual understanding.

1. **INTRODUCTION**

**1.1 Background**

Artificial intelligence (AI) and machine learning (ML) have transformed numerous industries by enabling systems to learn from data and make intelligent decisions. Neural networks, inspired by the human brain's interconnected neurons, are central to these advancements. Despite significant progress, traditional neural network architectures often face challenges in adaptability, interpretability, and efficiency, particularly when dealing with complex, non-linear data patterns.

Feedforward networks (FFNs) and transformer models have been the cornerstone of many successful applications. However, their limitations necessitate the exploration of new architectures that can adapt dynamically to data, provide better interpretability, and efficiently capture complex relationships.

**1.2 Problem Statement**

T

*The rest of this page is left intentionally blank*

1. **METHODOLOGY**

*The rest of this page is left intentionally blank*

1. **FINDINGS & ANALYSIS**

*The rest of this page is left intentionally blank*

1. **IMPROVEMENTS & SOLUTIONS**

*The rest of this page is left intentionally blank*

1. **CONCLUSIONS**

*The rest of this page is left intentionally blank*

1. **REFERENCES**

1. OpenAI. (2023). *ChatGPT: Optimizing language models for dialogue*. Retrieved from <https://openai.com/>

2. SimpleBench Dataset. (2023). *Benchmarking AI reasoning*. Retrieved from <https://simplebench.com/>

3. Johnson, D. (2022). *Psychological perspectives on human reasoning*. *Journal of Cognitive Science*, 34(2), 123-140.