ADVANCING FOUNDATIONAL LLM REASONING: MULTI-AGENT FRAMEWORKS AND COT PROMPT OPTIMIZATION

PUSHKAR VERMA

Research Proposal

NOVEMBER 2024

ABSTRACT

This research explores the development of agent-based framework, that can provide necessary feedback for Chain of Thought (CoT) reasoning in foundational large language models (LLMs) that lack CoT like reasoning capabilities inherently, by mode of pre-training. The study explores multi-agent setup, in different configurations of homogenous (same base model) and heterogenous (different base model) agents with different objectives and compares their performance on primarily GSM8K dataset. The research also evaluates the best performing CoT feedback mechanism, between including post-step evaluation, pre-step planning, a hybrid of both approaches.

The outcomes of the research include identification of optimal agent configurations, refining CoT reasoning methods and hence improving the reasoning capabilities of foundational LLM models without having to pretrain or fine-tune them. This research will also help bridge gap between proprietary CoT like reasoning models and open-source foundational models, advancing open-source research. This research will contribute to field of AI, enabling applications in education & healthcare domains that would benefit from enhanced reasoning capabilities.

LIST OF FIGURES

LIST OF ABBREVIATIONS

LLM(s) ……….	Large Language Model(s)
CoT …...............	Chain of Thought
AI ………………	Artificial Intelligence
API …………….	Application Programming Interface
GPU …………...	Graphics Processor Unit
IDE ……………	Integrated Development Environment
……	
……	
……	

Table of Contents

1. Background

In the recent past, we have seen a lot of improvements in Foundational Large Language Models (LLMs). While these models are very capable at various tasks like text generation, language understanding, code generation etc. They often struggle with consistent, human-like reasoning tasks. This oversight becomes more visible when it comes to simple/complex tasks that need multi-step reasoning, common-sense reasoning or planning like mathematical problems and word/logical puzzles. With Chain of Thought (CoT) reasoning the Foundational LLMs are incentivised to first break a problem down into smaller more manageable chunks and then attempt to solve the problem step by step.

The concept of CoT reasoning gained momentum with (Wei et al., 2022), who demonstrated that elicit prompting, to generate intermediate steps significantly enhanced their performance on multi-step reasoning. Building on this, (Wang et al., 2022) introduced a Self-Consistency method, which improved CoT prompting by first finding diverse reasoning paths and then selecting the most consistent result as final output.

The CoT reasoning got further popularity in general audience, once OpenAI released their proprietary o1 model. Unlike traditional models that use CoT prompting, o1 model is pre-trained to perform CoT like reasoning naturally. This makes it excel on wide range of very complex tasks, placing it very high on competitive programming, Math Olympiad and exceeding human PhD level accuracy on Physics, Chemistry and Biology benchmarks.

In this research we will use open-source foundational models such as Llama 3.1, which are not inherently pre-trained to perform CoT like reasoning, needing to leverage explicit prompting and external feedback from the AI agents to perform CoT reasoning. Moreover, proprietary o1 model has shown the potential of CoT like reasoning, and the problems that can be solved using such techniques.

2. Problem Statement/Related Research

Chain of Thought (CoT) reasoning had demonstrated the fact that, one can use the foundational LLM models and improve its capabilities by usage of prompting techniques, making the model generate intermediate steps and breaking the problem in smaller more manageable chunks.

(Wei et al., 2022) introduced CoT prompting technique which improved the capabilities of the model in reasoning tasks. While this approach was capable of achieving state-of-the-art results, it’s single reasoning path limited the performance of the model. This unlocked potential was further built upon by (Wang et al., 2022), by introducing novel Self-Consistency approach, improving the reliability of the model by using multiple reasoning path to solve the problem and then using the most consistent result as final answer to the problem. This again pushed the state-of-the-art boundary forward, but the need to follow all the paths to get to the final result increased the computational overheads.

(Ahmed El-Kishky et al., 2024), OpenAI’s o1 model was able to popularize the reasoning capabilities of CoT like reasoning models, but as it remains proprietary and closed-source, research community cannot access its architecture or weights, limiting the customization/adaptation to new use cases. This research draws inspiration from the performance achieved by o1 model at various reasoning tasks, and its ability to tackle very complex problems by breaking it down and tackling it one step at a time. However, its proprietary nature and closed-source nature limits accessibility. Open-source LLMs like Llama 3.1 while very capable lack the CoT reasoning capabilities that comes from pretraining in a certain way. While an individual cannot get access to resources/infrastructure needed to train such models, there exists an opportunity to leverage AI Agents for the required feedback influencing the LLMs CoT reasoning to arrive at correct results, reducing the oversights of directly trying to generate the results.

(Wan et al., 2024) enhanced CoT reasoning research with CoT derailment detection and error correction mechanism in the CoT Rerailer framework. This method used multi-agent collaboration system to refine reasoning process, and when it identified the derailment in CoT reasoning process, it tried to re-rail the CoT reasoning. Here the focus is primarily on error detection and correction. While this approach is able to overcome the drawback of previous method of high computational overhead, it did not explore topics like CoT reasoning process improvement or agent architecture optimization for better CoT reasoning. Similar to CoT Rerailer research, this research will also use a multi-agent setup, but the focus will be broader than just derailment detection and correction.

This research, proposes:

1. Implementing AI agents to evaluate, refine and optimise intermediate reasoning steps. Evaluating the intermediate steps, providing dynamic feedback.
2. Evaluating different feedback mechanisms and identification of best feedback mechanism to use.
3. Comparing the different agent ensembles using homogenous (agents built using same base model) and heterogenous (agents built using different base model) agents and how they perform on different benchmarks on reasoning tasks.

3. Research Questions

Primary Research Question – How effective are the agent feedback-based setups at improving the reasoning capability of a foundational LLM model using CoT prompting techniques?

Other Questions

1. Compare the performance using homogeneous and heterogeneous agent ensembles on a foundational LLM model.

2. What role does agent, it's objective and architecture play for CoT reasoning task?

3. What is the best approach to implement CoT reasoning?
   Scenarios to test:
   a. Post-Step Evaluation Approach
   b. Pre-Step Planning Approach
   c. Hybrid of Pre & Post Approach

Answering the above questions will give novel insights on how to best implement Chain of Thought reasoning on a Foundational LLM model using a Multiagent setup. It will also help evaluate and compare the performance of different approaches of CoT reasoning.

4. Aim

The aim of the research is to build a multi-agent setup that can enable foundational LLMs to perform better, improving its Chain-of-Thought reasoning ability, by feedback mechanisms.

5. Objective

5.1. Conduct Literature Review:

- Read and analyse the existing cutting-edge research on CoT Prompting, how it has evolved over time, best practices and shortlisting the best performing candidates.
- Identify the gaps in existing research and shortlisting the topics of research.
- Reading about the implementation of agent setup and feedback mechanisms that can be leveraged.

5.2. Develop multi-agent setup:

- Build agents setups to evaluate, plan ahead & provide feedback.
- Build agent-setups with homogenous and heterogenous agents.

5.3. Integrate an open-source foundational LLM model:

- Use foundational LLM models like llama using frameworks like LangChain and HuggingFace.
- Create prompts to make the foundational model to break the problem into smaller chunks and leveraging agent setup to provide necessary feedback for CoT implementation.

5.4. Evaluate Performance & Benchmarking using different setups, architectures & objectives:

- Evaluate the LLM models with or without CoT prompting on different Benchmarks like GSM8K,
- Evaluate different agent setups to find best configurations and best feedback mechanisms.

5.5. Contribute to Open-Source Research:

- Submit the research thesis, publish key findings and best practices in research papers for peer review and accessibility for research community. Hence, contributing to evolve field of research further.

6. Significance of the Study

In this research we attempt to improve the CoT reasoning of a foundational LLM model that is not inherently pre-trained to do CoT prompting. By testing with multi-agent setup in different configurations, agent sizes, objective and feedback mechanism, we plan to advance the study of CoT prompting. This will also advance open-source research, bridging some gap between proprietary models like Open-AI's o1 model.

7. Scope of the Study

The research will cover:

1. Building multi-agent setup for testing different configurations.
2. Testing multiple feedback mechanisms.
3. Benchmarking on GSM8K and other datasets to evaluate setup performance.
4. Identification of best ensembles for agent configurations.

The research will not cover:

1. Extending CoT prompting to models that are pretrained to do CoT like reasoning.
2. Anything else not explicitly mentioned in scope.

8. Research Methodology

8.1. Data Collection Process and Sources

The primary dataset that will be used for this research will be **GSM8K**, which mainly comprises of 8th Grade math problems, solving these would need multi-step reasoning making it ideal for evaluation using CoT reasoning. Apart from this dataset, the benchmarking would be done on **Strategy QA**, **Multi Arith** & **AQUA** datasets, each of the datasets used are openly available in Kaggle & HuggingFace platforms.

8.2. Data Cleaning & Pre-processing

- **Cleaning**: The available data would be cleaned to remove/rectify ambiguous or incomplete entries.

- **Preprocessing**: Will do appropriate preprocessing for the type of LLM selected like tokenization etc.

- **CoT steps creation:** Will create CoT labels for the model to show how to follow and generate CoT Reasoing steps as per the strategy needed for CoT feedback technique.

8.3. Machine Learning Model & Framework

Will use open source foundational LLM models as base model to generate CoT reasoning. Will incorporate agents for required feedback for evaluation/correction of CoT reasoning.

- **Evaluating Agents**: To evaluate the CoT reasoning provided by the base LLM for correctness and completeness.

- **Feedback Agents**: Agents to provide constructive feedback for improvement of the CoT reasoning based on the evaluations by the Evaluator agents.

- **Multi-Agent Setup**: Building setups using homogenous and heterogenous agents, of different sizes and objectives.

8.4. Criteria of Algorithm Selection & Model Evaluation

Algorithm Selection: Llama 3.1 is chosen as it is one of the top performing foundational LLM models, that is of right size (7billion parameters), which can fit in my System GPU for inference. I will use small variants of Llama if it does not work well on my system.

Evaluation Metrics:

- **Accuracy**: For evaluating correctness of final answer.
- **Reasoning step evaluation**: Correctness and completeness of the reasoning steps generated.

8.5. Experimental Design & Statistical Analysis

Experiments: Experiments will be conducted by using these configurations:

- Homogenous and Heterogenous Multi-agent setup
- Multiple CoT Implementation strategies, post-step evaluation, pre-planning, hybrid and retrospective feedback after final answer generation.

Statistical Analysis: Perform comparison studies on the performance of different setups, with metrics like accuracy of final answer & evaluation of reasoning steps.

8.6. Software & Tools

Programming Language: Python

Key Libraries: Huggingface, LangChain, Tensorflow, Pytorch and Multi-agent coordination libraries.

Tools: Jupyter Notebooks & Python IDEs

8.7. Considerations for reproducibility

- Using public dataset for experimentation and benchmarking.
- The code files would be hosted on open-source code repository like GitHub.
- Detailed documentation with setup instruction, preprocessing steps & evaluation steps would be included.

9. Requirement Resources

9.1. Hardware:

- Compute: Mid-range GPU with at least 6-8 GB video memory would be required for loading the foundational model for inference and CoT reasoning. Might need even more memory for handling the multi-agents simultaneously.
- Storage: 100-200 GB of storage should be sufficient to store the model files and weights associated. The datasets are small in terms of storage space needed given these are textual.

9.2. Software:

- Programming Language: Python
- Libraries: Huggingface, LangChain, Tensorflow, Pytorch and Multi-agent coordination libraries.
- Tools: Jupyter Notebooks & Python IDEs
- Version Control: Git, GitHub repository

9.3. Data:

- GSM8K: Grade-school-level math problems, public dataset, medium size

- StrategyQA: Common sense reasoning dataset where models need to do multi step reasoning, public dataset, small size

- MultiArith: Arithmetic word problems dataset, public dataset, very small size

- AQuA RAT – Multiple choice math problems, which require complex reasoning, public dataset, large size

9.4. Personnel:

- Researcher: Will be responsible for design, implementation and evaluation of the study.

- Thesis Supervisor**:** Will provide necessary guidance and feed-back on design and implementation at different milestones of study.

9.5. Budget:

- Hardware Cost: LLM APIs, Agent and Cloud services 100 to 200$. Apart from pre-owned laptop with GPU.

- Software Licenses: Free and open-source tools

- Miscellaneous: 50 to 100$

9.6. Time management:

Duration of Research from topic approval would be around 18 weeks. The time management would be very crucial, given the time-frame. For detailed plan of action and timelines please refer Research Plan section.

9.7. Collaboration:

Thesis Supervisor: Regular consultation with thesis supervisor for feedback on progress and guidance.

Global Research Community: The research is building on top of the cutting-edge research done by the research community before me. I would love to get feed-back through peer review.

10. Research Plan

Here's a breakdown of a research plan spanning 18 weeks from topic approval on 21st Oct 2024:

Weeks 1–3 (21st Oct – 17rd Nov 2024)

Activity: Literature Review

Tasks: Literature review, gap identification and research question formulation.

Weeks 3–4 (4th Oct – 17th Nov 2024)

Activity: Drafting Research Methodology

Tasks: Initial drafting of research proposal, working on research methodology formulation.

Weeks 5–6 (18th Nov – 1st Dec 2024)

Activity: Finalizing Research Proposal

Tasks: Refining and finalizing the Research proposal. Submitting the research proposal.

Weeks 7–8 (2nd Dec – 15th Dec 2024)

Activity: Agent Framework Setup

Tasks: Identify and finalize the agent framework to use as in research setup. Start initial testing with Foundational LLM model.

Weeks 9–10 (16nd Dec – 29th Dec 2024)

Activity: Implementation of Agent setup

Tasks: Implementation of agent framework, setting up agent ensembles.

Weeks 9–11 (16nd Dec – 29th Dec 2024)

Activity: Finalizing Interim Research Report

Tasks: Drafting and Finalising interim report and submit it.

Weeks 11–12 (30nd Dec 2024 – 12th Jan 2025)

Activity: Testing Agent Configurations

Tasks: Experimenting on Agent Configurations.

Weeks 13–14 (13th Jan – 26th Jan 2025)

Activity: Experimentation and Benchmark setup preparation

Tasks: Testing CoT reasoning frameworks on GSM8K dataset. Setting up feedback mechanism benchmarking setup.

Weeks 15–16 (27th Jan – 9th Feb 2025)

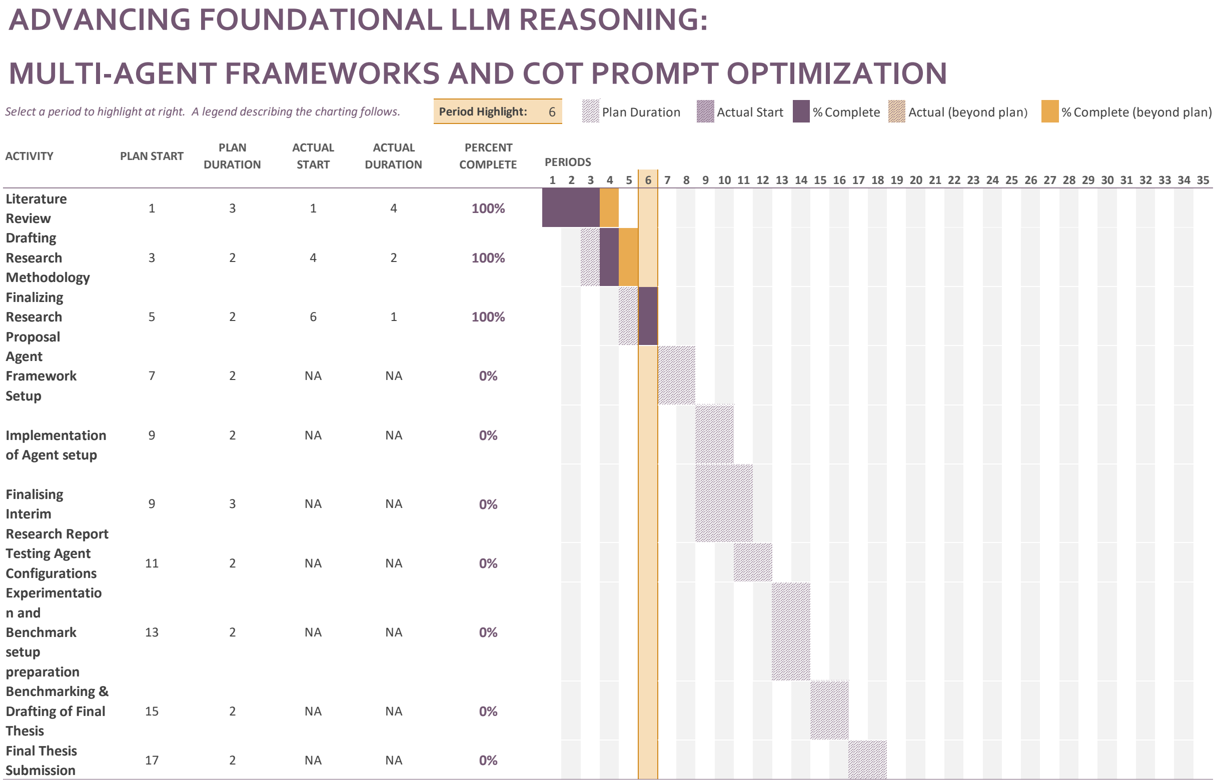
Activity: Benchmarking & Drafting of Final Thesis

Tasks: Benchmarking on all datasets and start drafting the Final Thesis.

Weeks 17–18 (10th Feb – 23rd Feb 2025)

Activity: Final Thesis Submission

Tasks: Finalizing final thesis, preparing video presentation and submission.



ADVANCING FOUNDATIONAL LLM REASONING:
MULTI-AGENT FRAMEWORKS AND COT PROMPT OPTIMIZATION

Select a period to highlight at right. A legend describing the charting follows.

Period Highlight: 6

ACTIVITY	PLAN START	PLAN DURATION	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE
Literature Review	1	3	1	4	100%
Drafting Research Methodology	3	2	4	2	100%
Finalizing Research Proposal	5	2	6	1	100%
Agent Framework Setup	7	2	NA	NA	0%
Implementation of Agent setup	9	2	NA	NA	0%
Finalising Interim Research Report	9	3	NA	NA	0%
Testing Agent Configurations	11	2	NA	NA	0%
Experimentation and Benchmark setup preparation	13	2	NA	NA	0%
Benchmarking & Drafting of Final Thesis	15	2	NA	NA	0%
Final Thesis Submission	17	2	NA	NA	0%

Figure 10.1 Gantt Chart showing the Research Plan

References

Ahmed El-Kishky, Daniel Selsam, Francis Song, Giambattista Parascandolo, Hongyu Ren, Hunter Lightman, Hyung Won Chung, Ilge Akkaya, Ilya Sutskever & Jason Wei (2024). *OpenAI o1 Model*. [Online]. Available from: https://openai.com/index/learning-to-reason-with-llms/. [Accessed: 1 December 2024].

Wan, G., Wu, Y., Chen, J. & Li, S. (2024). *CoT Rerailer: Enhancing the Reliability of Large Language Models in Complex Reasoning Tasks through Error Detection and Correction*. [Online]. Available from: https://arxiv.org/abs/2408.13940v2. [Accessed: 17 November 2024].

Wang, X., Wei, J., Schuurmans, D., Le, Q., Chi, E.H., Narang, S., Chowdhery, A. & Zhou, D. (2022). Self-Consistency Improves Chain of Thought Reasoning in Language Models. *11th International Conference on Learning Representations, ICLR 2023*. [Online]. Available from: https://arxiv.org/abs/2203.11171v4. [Accessed: 17 November 2024].

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E.H., Le, Q. V. & Zhou, D. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. *Advances in Neural Information Processing Systems*. [Online]. 35. Available from: https://arxiv.org/abs/2201.11903v6. [Accessed: 17 November 2024].