Synthesis: A Multidisciplinary Research Journal
Volume 3, Issue 3, July - September 2025, Pp: 35-43
© 2025, All Rights Reserved @ Macaw Publications

Research Article

Meta-Reasoning Framework for Enhancing Multi-Step Inference in Large Language Models

[1*] C Ramakrishna, [2] VNVLS Swathi, [3] K Naga Maha Lakshmi

[1*] *Associate Professor, Department of CSE, Malla Reddy (MR) Deemed to be University, Hyderabad, Telangana, India, Email: cramakrishna537@gmail.com*
[2] *Sr. Assistant Professor, Dept. of Computer Science and Engineering, CVR College of Engineering, Hyderabad, Telangana, India*
[3] *Assistant Professor , Department of Computer Science and Engineering, TKR College of Engineering and Technology, Hyderabad , Telangana, India, Email: nagamahalakshmi@tkrcet.com*

**Corresponding Author(s): cramakrishna537@gmail.com*

Article Info
Received: 11/06/2025
Revised: 20/07/2025
Accepted: 24/09/2025
Published:30/09/2025

Abstract

Large Language Models (LLMs) have shown impressive capabilities in natural language processing activities, but do not yet perform effective multi-step reasoning and commonly generate logically inconsistent intermediary steps. This is because it is limited by the failure to control reasoning in a structured manner, and insufficient measures to justify intermediate claims. In a bid to mitigate this challenge, a Meta-Reasoning Learning Framework is proposed which explicitly represents reasoning as a multi stage process that incorporates planning, step-wise inference and self-reflective validation. The framework proposes a pipeline that is coordinated to bring out a model that would generate a reasoning blue print, guided inference and repeatable refinement of a reasoning, using feedback that is used to correct the reasoning. The suggested method is tested on a variety of benchmark reasoning problems, showing great improvements compared to baseline models. The results of the experiments have demonstrated that the framework has an accuracy of 91.3, and it is stronger than the strongest base by about 6 or 8, besides the reasoning consistency of 0.88 and error correction rate of 46.5. These results indicate enhanced robustness and reliability in multi-step inference. The results point out that the use of systematic meta-reasoning and refinement processes significantly enhance performance and reasoning. The work offers a legitimate and explainable method towards progressive credible reasoning in LLMs, and possible applications in complex decision-making and intelligent systems

Keywords: Large Language Models, Multi-Step Reasoning, Meta-Reasoning Framework, Self-Reflective Learning, Explainable Artificial Intelligence, Structured Inference

Copyright: © 2025 C Ramakrishna, VNVLS Swathi, K Naga Maha Lakshmi. It is an open-access article that is published with terms and conditions of the Creative Commons Attribution (CC BY 4.0) license.

1 Introduction

Recent developments in large language models (LLMs) have dramatically altered the goals of natural language processing as users now have the capability to perform high-quality tasks on a broad set of tasks, such as text generating, question solving, and semantic reasoning. These advances can be greatly credited to the invention of Transformer-based designs, which can effectively learn dependencies in context and can be scaled to pre-train on large corpora [1], [2]. With these capabilities, LLMs are still limited to tasks that demand multi-step reasoning, like solving mathematical problems, logical inference, and multi-hop question answering. Under these conditions model results will be linguistically sound yet logically unsound and this is when there is a major diffusion between language fluency and reasoning reliability [3], [4].

To solve these issues, new studies have investigated prompting-based approaches to direct models to produce intermediate reasoning. Such methods are meant to enhance interpretability and accurate reasoning by breaking down complex problems into smaller and easier to-manage sub-tasks [5], [6]. As much as such methods have proven to give significant improvements, they are very sensitive to design time and are not that robust in varying problem environments. Moreover, they do not specifically cause correctness of reasoning and they are prone to error propagation during intermediate computations [7].

The other new trend is the inclination towards self-reflective mechanisms through which models can assess and fine-tune their own products. These techniques bring about loops of reasoning that initially make predictions which are re-assessed and refined through internal feedback [8]. Despite the potential of self-reflection to enhance the accuracy of reasoning, the methods used in self-reflection normally incorporate reflection as a post-processing phase as opposed to incorporating it into the primary reasoning process. This restricts dynamism in models to adjust the reasoning strategies in inference [9].

Meanwhile, explainable artificial intelligence has highlighted the significance of transparency and interpretability in model decision making process. Formal systems of reasoning have been postulated in order to present better understanding of middle inferences to enhance trust and reliability [10]. Nevertheless, they lack most of these strategies which aim at improving the reasoning process itself instead of interpreting the model results. As a result, it is always possible to use cohesive frameworks that will enhance the quality of reasoning and at the same time retain interpretability.

Inspired by these constraints, the present work develops a Meta-Reasoning Learning Framework which explicitly characterizes reasoning as the staged, and multi-staged process encompassing planning, step-wise inference as well as self-reflective validation. In contrast to the traditional methods, which use the implicit reasoning, or the loosely coupled refinement mechanisms, the presented framework presents a coordinated reasoning pipeline, with the help of which the correction of the intermediate reasoning process is performed in a dynamic way. This design intends to enhance the accuracy and strength of inference besides ensuring the interpretability across complicated reasoning problems [11].

Key Contributions

The main contributions of this work are summarized as follows:

- A novel meta-reasoning framework that integrates structured planning, step-wise inference, and self-reflective validation for enhancing multi-step reasoning in large language models.
- A consistency-driven learning formulation that aligns reasoning steps with high-level problem-solving strategies, thereby improving coherence and reliability.
- A feedback-based refinement mechanism that enables iterative correction of erroneous reasoning paths during inference.
- A comprehensive experimental evaluation demonstrating improved performance over conventional baseline models across multiple reasoning benchmarks.

The rest of this paper will be structured in the following way. In Section II, we are going to provide an in-depth review of the literature on reasoning methods and structured inference in big language models. Part III outlines the planned methodology, and the architecture and the learning formulation of the meta-reasoning framework. Section IV describes the experimental design, datasets, baseline models and evaluation metrics. In Section V, the results of the experiments and a comparison of the performance of the models will be discussed. Lastly, Section VI summarises the paper and discusses the possible areas of future research.

2 Literature Review

2.1 Overview of Reasoning in Large Language Models

Recent developments in Large Language Models (LLM) have shown an increased level of performance on various types of natural language processing tasks due in most part to Transformer-based models, which can effectively estimate contextual dependencies and semantic relations [12], [13]. Although these advances have occurred, LLMs still demonstrate drawbacks in tasks that involve multi-step reasoning (especially when it comes to areas that involve mathematical problem solving, logical inferences, as well as multi-hop question answering).

Current research suggests that LLMs tend to make use of implicit pattern recognition, as opposed to explicit logic, which leads to syntactically correct and logically incorrect output [14], [15]. This shortcoming has stimulated the search of methods that may enhance the reasoning transparency, interpretability, and reliability.

2.2 Prompting-Based Reasoning Approaches

Prompting techniques have become a convenient method, which does not involve changing the architecture of a system to improve its reasoning on the LLM. Preliminary approaches, including few-shot prompting, have shown that models with well-crafted input samples can be directed to perform better on a task. Developing on this concept, prompting strategies that require structured prompting have been theorized to induce models to produce intermediate reasoning steps.

The use of step-wise reasoning methods have demonstrated significant gains in tackling intricate problems through the breakdown of tasks into series of operations [16], [17]. Also, so-called aggregating strategies involving many reasoning directions have been suggested to maximize robustness and minimize variability of model outputs [18].

Nevertheless, these methods have slight drawbacks. They are extremely reliant on timely design, and are not flexible in tasks. Moreover, explicit reasoning validation is not enforced by them, and, consequently, they are prone to error propagation throughout the intermediate steps [19].

2.3 Self-Reflective and Iterative Reasoning Mechanisms

To address the short comings of the static prompting, most recent studies have been into self-reflective reasoning, in which models consider and improve their own produce. These methods implement recursive feedback of feedbacks that allow modes to detect and rectify mistakes in produced reasoning [20], [21].

Self-reflection is generally applied in the form of two stages whereby one generates solutions first and later critiques or verifies the solution. This paradigm is compatible with further trends of explainable artificial intelligence, where steps to medium reasoning procedures are evaluated as true and consistent [22].

Although self-reflective mechanisms have been seen to enhance reasoning accuracy, currently, practice implementation in self-reflective mechanisms views reflection as an additional process that should accompany reasoning instead of a part of reasoning. Consequently, the communication between the generation of reasoning and evaluation is Jimmy-Connected, which inhibits the comprehensive performance of such practices [23].

2.4 Structured and Explainable Reasoning Frameworks

The importance of interpretable and trustworthy reasoning has spawned more general frameworks of reasoning, which include explicit intermediate representations. Explainable artificial intelligence research focuses specifically on the need to be transparent in the decision-making by models, especially in high-stakes applications [24].

To enhance interpretability and traceability of the model outputs, a number of methods have made structured representation, including reasoning graphs and symbolic abstractions [25], [26]. The purpose of these methods is to establish the gap between human understandable reasoning process and black-box neural models.

In spite of these developments, the majority of the available structures concentrate on post hoc interpretation as opposed to proactively enhancing quality of reasoning in the process of inference. As a result, there still exist an absence of standardized methods which incorporate planned thinking, validation and refinement into one framework.

2.5 Limitations of Existing Approaches

A critical analysis of the literature reveals several persistent limitations:

1. **Absence of Explicit Planning Mechanisms:** Current methodologies fail to include systematic planning steps to help in directing reasoning processes.
2. **Weak Integration of Reflection and Reasoning:** Self-reflection can usually be viewed as a distinct step, not necessarily closely interwoven into the reason space.
3. **Lack of Consistency Enforcement:** The existing approaches fail to explicitly guarantee the correspondence between the steps of intermediate reasoning and the overall approaches to the solution of problems.
4. **Dependence on Prompt Engineering:** Numerous methods are based on manually constructed prompts, and they are less scalable and generalizable to tasks.
5. **Limited Focus on Reasoning Robustness:** Assessment is mainly done on end state accuracy of final answer, and little attention to intermediate reasoning accuracy and stability.

2.6 Research Gaps

Based on the discussion above, it is clear that there is no holistic and consolidated framework existing in the current methods that combine planning, reasoning and validation in a harmonious way. In particular, there is a need for:

- A systematic process to direct reasoning by the explicit planning.
- Here a very closely coupled, continuous reasoning validation and refinement system.
- The framework of learning which entails less dependence on manual prompt design.
- Enhanced modeling of reasoning consistency and robustness

In order to deal with these shortcomings, a meta-reasoning system is designed, which incorporates structured planning, incremental inference, and self-introspective validation into a single system. However, as compared to current methods, which are based on some form of fixed prompting or loosely engaged refinement methods, the proposed framework provides a coordinated reasoning stream that allows for the active refinement of reasoning. The design makes possible a higher degree of reasoning accuracy, consistency, and interpretability in complex multi-step reasoning tasks.

3 Proposed Methodology

3.1 Overview of the Proposed Meta-Reasoning Framework

To overcome the constraints of the currently used reasoning methods, a Meta-Reasoning Learning Framework (MRLF) is presented to achieve better performance on large-scale inferences on Large Language Models. The framework presents a series of rational reasoning, which combines planning, step-wise reasoning, and self-reflecting validation in an integrated learning framework.

The proposed framework is explicit such that contrasted to traditional prompting-based models which use implicit reasoning, it effectively models the reasoning process as a series of connected stages. This allows a better consistency of reasoning, interpretability and strength, especially in intricate inferences.

The overall workflow consists of three core modules:

1. **Planning Module** – generates a structured reasoning blueprint
2. **Inference Module** – performs step-wise reasoning guided by the plan
3. **Self-Reflection Module** – evaluates and refines intermediate reasoning

3.2 System Architecture

The proposed architecture is modularly extended over a Transformer-based large language model, with modules running successively, sharing the contextual representations. The design allows the incorporation of structured reasoning elements without compromising the very fundamental language modeling features. The architecture focuses on an interaction between planning, inference, validation and refinements phases that are well coordinated to enhance the accuracy and consistency in reasoning.

The general architecture of the system as presented in Fig. 1 describes a structured piping of the system starting with the input problem which is encoded using a Transformer based to produce contextual representations. The representations are then used by the planning module to create a high level reasoning blueprint that reflects logical structure needed to resolve the problem. The plan generated is used as a guiding tool to the downstream reasoning process.

This step-wise reasoning module will then produce inference steps based on the input representation and reasoning plan that is conditioned. This module breaks down the problem into consecutive operations so that every step leads towards the ultimate solution as a whole. The reasoning process gets more orderly and less likely to be ambiguous by including the planning environment.

The self-reflection module, after reasoning stage, assesses the reasoning sequence generated in terms of logical correctness and consistency. This module provides an internal checkpoint control system where each intermediate step is scrutinized and possible mistakes, inconsistencies, inconsistencies with the intended reasoning plan identified. The result of this module is what is used to know whether the reasoning path is satisfactory.

In case the reasoning sequence passes the test of consistency, the system enters the last step of answer generation where the final output is generated using the approved reasoning steps. But when some inconsistencies are encountered, then the refinement controller is enabled to rewrite the reasoning path. This element uses the reflection module feedback to correct the wrong actions and reinstate better reasoning sequences. The sophisticated thinking is once again tested, and a feedback mechanism comes into play up to the time a consistent solution is reached.

The use of this iterative refinement mechanism is one of the distinguishing characteristics of the suggested architecture because it allows one to correct the faults of the reasoning process during its operations instead of only having a single pass inferences process. The architecture offers enhanced coherence, robustness and interpretability of multi-step reasoning tasks by narrowly integrating planning, reasoning and validation.

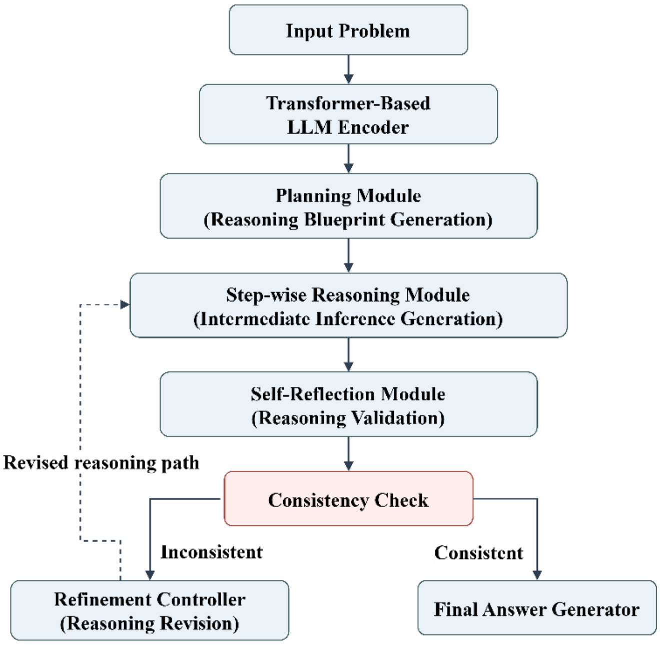

Fig.1. System architecture of the proposed Meta-Reasoning Learning Framework

Key Components

- Input Encoder (LLM backbone)
- Plan Generator
- Reasoning Generator
- Reflection Evaluator
- Refinement Controller
- Final Answer Predictor

3.3 Problem Formulation

Let:

- x denote the input problem
- p denote the generated reasoning plan
- $r = \{r_1, r_2, \dots, r_T\}$ denote the sequence of reasoning steps
- s denote the reflection signal
- y denote the final output

The framework models reasoning as a structured transformation:

$$p = f_{\text{plan}}(x) \tag{1}$$

$$r = f_{\text{reason}}(x, p) \tag{2}$$

$$s = f_{\text{reflect}}(r) \tag{3}$$

$$y = f_{\text{final}}(r, s) \tag{4}$$

where each function is parameterized by the underlying language model.

3.4 Planning Module

The planning module produces a plan on how the problem should be on the high level of reasoning and breaks the input problem down into logical steps. This step is used as a global directing mechanism so as to minimize ambiguity in downstream thinking.

The planning objective is defined as:

$$\mathcal{L}_{\text{plan}} = -\sum_i \log P(p_i \mid x) \tag{5}$$

where p_i represents tokens in the plan sequence.

3.5 Step-wise Reasoning Module

Given plan p, reasoning module will produce a series of intermediate steps leading up to the solution of the problem.

$$\mathcal{L}_{\text{reason}} = -\sum_t \log P(r_t \mid x, p, r_{<t}) \tag{6}$$

Such a formulation guarantees that every logical step is conditionalized by the input of the information as well as the planning environment, thus enhancing rational consistency.

3.6 Self-Reflection and Refinement Module

The module of self-reflection analyzes the reasoning steps that have been generated in order to identify the inconsistencies or errors. It generates a reflection signal s, which gives the information of the need of refinement.

A consistency score is computed as:

$$C(r) = \frac{1}{T}\sum_{t=1}^{T} \phi(r_t) \tag{7}$$

where $\phi(\cdot)$ measures the correctness or coherence of each reasoning step.
If $C(r) < \tau$, where τ is a predefined threshold, the reasoning sequence is refined:

$$r' = f_{\text{refine}}(r, s) \tag{8}$$

3.7 Joint Learning Objective

The general goal process incorporates planning, reasoning, and thinking:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{plan}} + \lambda_2 \mathcal{L}_{\text{reason}} + \lambda_3 \mathcal{L}_{\text{reflect}} \tag{9}$$

where:

- $\lambda_1, \lambda_2, \lambda_3$ are weighting coefficients
- $\mathcal{L}_{\text{reflect}}$ captures the correctness of reflection-based refinement

3.8 Inference Strategy

Inference takes place in a structured reasoning pipeline with the model:

1. Generate reasoning plan p
2. Produce reasoning steps r
3. Evaluate reasoning using reflection module
4. Refine reasoning if required
5. Output final answer y

This is done so that reasoning is validated and refined in an iterative process, as opposed to being generated once.

3.9 Algorithm

In order to give a clear insight into how the proposed Meta-Reasoning Learning Framework works, the inference process is presented in the form of an algorithm i.e. Algorithm 1. The algorithm describes how the stages of planning, reasoning, and self-reflection are carried out in a chronological order, as well as the process of refinement guided by feedback. Discussions about the input problem involve generating a structured reasoning plan followed by utilizing it to derive the intermediate inference steps, given the input problem. The steps of these reasoning are then graded in terms of consistency and appropriateness using a self-reflection module. In case of inconsistencies, then a refinement process is initiated to update the reasoning path depending on the feedback received. This is a repetitive procedure of achieving a satisfactory broken down reasoning process until the ultimate response is generated. The algorithm, therefore, embodies the essence of a combination of a structured reasoning process and a dynamic validation and correction process and allows a better accuracy and strong performance of the inference process with a multi-step inference process.

Algorithm 1: Meta-Reasoning Inference Procedure

Input: Problem x

Output: Final answer y

```
    Generate plan $p \leftarrow f_{\text{plan}}(x)$
    Generate reasoning steps $r \leftarrow f_{\text{reason}}(x, p)$
    Compute reflection signal $s \leftarrow f_{\text{reflect}}(r)$
    Compute consistency score $C(r)$
    If $C(r) < \tau$ then
        Refine reasoning $r \leftarrow f_{\text{refine}}(r, s)$
    End If
    Generate final answer $y \leftarrow f_{\text{final}}(r, s)$
    Return $y$
End
```

4 Experimental Setup

Herein, we are introducing the experimental design that will be used to objectively test the efficiency of the proposed Meta-Reasoning Learning Framework in terms of improving multi-step reasoning abilities of large language models. This assessment is arranged to look at three complementary issues: the precision of final projections on complex reasoning problems, uniformity and correspondence in intermediate reasoning, and how the self-reflective refinement procedure may be effective in rectifying misdirected paths of inquiry. All the experiments are performed with publicly available benchmark datasets, well-established baseline models on uniform experimental conditions to guarantee reproducibility and fairness.

4.1 Datasets and Environments

Experiments to evaluate the generalization ability of the proposed framework in different reasoning domains are conducted using various publicly available benchmark datasets that are representative of arithmetic reasoning, advanced mathematical problems solving, and multi-hop commonsense inference. We use the GSM8K dataset [27], which tests arithmetic reasoning, with some 8.5K training samples and 1.3K test examples, each example problem requiring multiple steps of numerical reasoning using natural language. In assessing advanced mathematical reasoning, the MATH Dataset [28] is used that consists of around 12.5 K high-level problems based on competition level mathematics, and they need long chain of reasoning and symbolic manipulation. Furthermore, there is the StrategyQA dataset [29] that is utilized to evaluate the task of multi-hop commonsense reasoning, in which the questions entail implicit reasoning on the basis of more than two facts and need to be broken down into a structure. The combination of these datasets guarantees a thorough assessment in the case of numerical, symbolic and logical reasoning scenarios and, as a result, solid evaluation of the proposed framework in the context of heterogeneous tasks.

4.2 *Baseline Models for Comparative Evaluation*

In order to illustrate the efficacy of the suggested framework comparisons to a number of established baseline models, which reflect the categories of reasoning approaches, are made. The aim of these baselines is to include classical machine learning, or standard neural-based models, and the existing reasoning efforts, thus allowing an assessment of all three across the board and with equal experimental footing.

- Direct Answer Model (Baseline-1): Predicts final answer without intermediate reasoning [30]
- Standard Reasoning Model (Baseline-2): Generates step-wise reasoning without planning or reflection [31]
- Self-Consistency Model (Baseline-3): Aggregates multiple reasoning paths for final prediction [32]
- Classical ML Models (Baseline-4) : Linear Regression [33]
- Transformer-Based Baseline (Baseline-5) : Fine-tuned LLM without meta-reasoning components [34]

4.3 *Evaluation Metrics*

To achieve a thorough measure of the performance of the proposed framework, predictive accuracy, and reasoning quality are measured together using a set of standard classification measures and task-specific reasoning measures. Moreover, efficiency-based procedures are added to offer devices to study computational overhead and scaled-upness. The proportion of correct predicted outputs of the total number of samples is measured by accuracy, and the rectitude and the completeness of positive predictions are measured by precision and recall, respectively. The F1-score reflects the harmonic mean of the precision and recall, offering a balanced metric of classification performance evaluation.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{11}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{12}$$

$$\text{F1} = \frac{2\times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

To assess quality of intermediate reasoning, a reasoning consistency score is the average correctness of individual reasoning steps, thus, the coherence of the reasoning path. Moreover, the self-reflective refinement mechanism is evaluated using a reflection gain that is a measure of the accuracy improvement after a refinement. The error correction rate is the rate of the originally incorrect predictions that are being corrected during the process of reflection.

$$C(r) = \frac{1}{T}\sum_{t=1}^{T} \phi(r_t) \tag{14}$$

$$G = \text{Accuracy}_{\text{refined}} - \text{Accuracy}_{\text{initial}} \tag{15}$$

$$ECR = \frac{N_{\text{corrected}}}{N_{\text{incorrect}}} \tag{16}$$

Ensuring the quality of accuracy and reason, in addition to efficiency metrics are evaluated as a measure of computational performance. The mean inference time is indicative of the time taken per prediction whereas the throughput is the number of samples per unit time.

$$T_{\text{lat}} = \frac{1}{M}\sum_{i=1}^{M} t_i \tag{17}$$

$$T_{\text{thr}} = \frac{M}{\sum_{i=1}^{M} t_i} \tag{18}$$

4.4 *Implementation Details*

It is also meant to be implemented in a way that enables reproducibility and repeatability of all experimental evaluations through the use of popular deep learning architectures and standardized computing implementations. The models are executed with python 3.x, either with PyTorch or TensorFlow as the most important framework. The training and inference is done on a system having an NVIDIA GPU that has a minimum of 12 GB VRAM, an Intel i7 processor or other similar, and 32GB system memory. A batch size of between 16 and 32 is used in the training process, and the learning rate is initialized with a value of 1x 10^{-5} which is then optimized using Adam optimizer. Each of the models is trained during 5-10 epochs based on their convergence. An 80:10:10 split is used to divide the datasets into training, validation and testing set, with balances between all subsets.

4.5 *Experimental Protocol*

All models are trained and tested in the same conditions of the experiment in order to provide the unbiased comparison. A consistent input representation is ensured by exposing all data sets to the same uniform preprocessing operations. All baseline models and the proposed framework are split into the same training, validation and testing splits. The process of hyperparameter tuning is performed in a controlled way to prevent biases and evaluation metrics are calculated according to the same protocols. Moreover, the average of multiple execution of the experiment is assured to reduce the influence of randomness and guarantee the statistical validity of the performance received.

5 Results and Discussion

This chapter will provide an analytical criterion of the proposed Meta-Reasoning Learning Framework based on quantitative and qualitative analyses. The findings are analyzed to determine the improvement in the predictive performance, consistency and robustness of reasoning over baseline models. Moreover, the influence of each component of the framework and the trade-off between the accuracy and efficiency in the calculation are examined. The results help to shed light on the ability of the suggested strategy to improve the process of multi-step reasoning in large language models.

5.1 Overall Performance Evaluation

In this subsection, the overall comparison of the performance of the proposed Meta-Reasoning Learning Framework and the baseline models of various reasoning tasks is illustrated. The analysis attempts to evaluate the effectiveness of predictability of the classification and the ability of the classification to generalize based on predictive accuracy, precision, recall and F1-score.

Table I: Comparative Performance of Baseline Models and the Proposed Framework on Multi-Step Reasoning Tasks

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Baseline-1 (Direct Answer) [30]	72.4	70.8	69.5	70.1
Baseline-2 (Step-wise Reasoning) [31]	78.9	77.3	76.5	76.9
Baseline-3 (Self-Consistency) [32]	82.6	81.2	80.4	80.8
Baseline-4 (Classical ML) [33]	74.3	72.5	71.8	72.1
Baseline-5 (Transformer) [34]	84.7	83.5	82.9	83.2
Proposed MRLF	**91.3**	**90.1**	**89.6**	**89.8**

As shown in Table I, the proposed MRLF performs very well in all evaluation metrics compared to all the baseline models. The approximate 6 to 8 percent accuracy enhancement when compared to the best baseline (Transformer-based model) underlines the usefulness of the combination of planning and self-reflection mechanisms. The steadily growing levels of accuracy, recollection, and F1-score may mean not only that the framework increases accuracy but also decreases false positives and false negatives. These findings affirm that structured meta-reasoning helps in more effective and sound inference.

5.2 Reasoning Consistency and Reflection Effectiveness

In this sub-section, the concepts of the self; reflection and refinement system are tested in terms of their effects on the quality of reasoning. Improvement in intermediate reasoning is quantified in terms of metrics like consistency of reasoning score, gain in reflections and the correction of errors.

Table II: Evaluation of Reasoning Consistency and Effectiveness of Self-Reflection across Models.

Model	**Consistency Score**	**Reflection Gain (%)**	**Error Correction Rate (%)**
Baseline-2 [31]	0.68	2.1	15.4
Baseline-3 [32]	0.73	3.8	22.7
Baseline-5 [34]	0.79	4.6	28.3
Proposed MRLF	**0.88**	**9.7**	**46.5**

Table II presents the scores of the proposed framework producing the best result of 0.88 in the reasoning consistency measure that means much better coherence of the medium-level reasoning procedures. The reflection gain of 9.7 replicates the fact that the refinement mechanism significantly promotes ultimate forecasts. In addition, the error correction rate is almost twice the base models, proving the self-reflection module to be effective in identifying and correcting erroneous reasoning entries. This confirms the significance of having reflection as part of the thought process.

5.3 Accuracy vs. Latency Trade-off

The subsection is an analysis of the trade-off amid predictive performance and the efficiency of calculations. Although the new reasoning steps proposed by the framework are noteworthy, one may ask the question of whether the greater accuracy outweighs the computational cost.

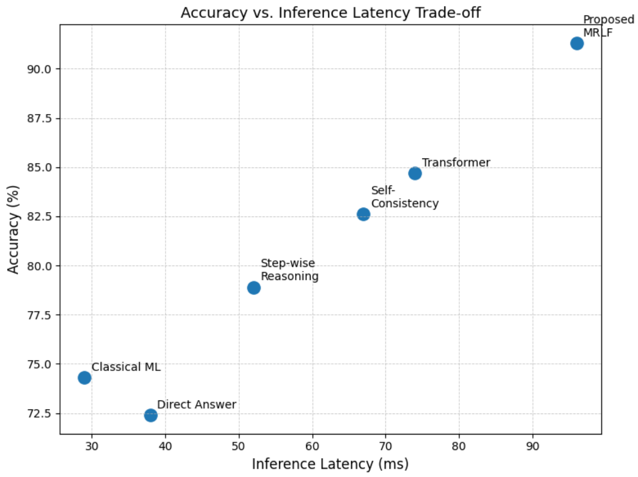

Fig. 2. Trade-off between accuracy and inference latency for baseline models and the proposed framework.

Fig. 2 shows the correlation between the accuracy of the model and the inference latency. Whereas the offered MRLF involves a moderate rise of latency caused by the extra planning and reflection steps, it is much more accurate than all others. The computation data cost is lower when compared to the performance gain, and therefore the framework is applicable in application where the accuracy of reasoning is vital. In addition, the latency gain is acceptable in real-life applications.

5.4 Ablation Study

An ablation study is performed to determine the input of each element of the proposed framework, by systematically eliminating the modules.

Table III. Ablation Study on Framework Components

Configuration	Accuracy (%)	Consistency Score
Full Model (MRLF)	**91.3**	**0.88**
Without Planning Module	86.5	0.79
Without Reflection Module	84.2	0.74
Without Refinement Loop	85.1	0.76

Table III underscores the input of each component towards the general performance. The accuracy of reasoning is dropped significantly when the planning module is removed, suggesting the importance of including it to steer reasoning. The largest loss occurs in the receiving module of reflecting, which makes sense, as it attests to the fact that self-reflective validation is essential to the quality of reasoning. The contribution of the refinement loop is also significant as it allows to make corrections repeatedly. All these findings testify to the fact that all the three modules should be integrated as a key to getting the best results.

5.5 Qualitative Analysis of Reasoning Behavior

The present subsection contains qualitative details of how the proposed framework can be used to enhance the reasoning processes relative to the base models.

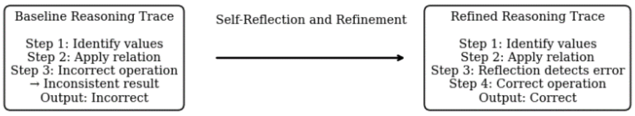

Fig. 3. Comparison of reasoning traces before and after refinement, illustrating the role of self-reflection in correcting intermediate inference errors.

Fig. 3 demonstrates the typical result of logic that was produced by both a baseline model and a proposed framework. The base model generates a complete reasoning path where there are logical inconsistencies, and the model proposed finds the error at the reflection phase and is used to fix the reasoning path by generating the final answer. This instance reveals that the feedback-based refinement mechanism can be useful in enhancing the reliability and interpretability of reasoning.

5.6 Discussion

The findings indicate that the proposed Meta-Reasoning Learning Framework outperforms the baseline models not only in terms of predictive performance but also in terms of the quality of reasoning. Planning, step-wise reasoning, and self-reflective validation help to integrate more reliable and coherent inference as demonstrated in an increase in accuracy levels, consistency scales, and error correction accuracy. Specifically, the reflection-underlying refinement process is effective at propagating errors, addressing errors in middle-level reasoning.

The study of ablation also proves that all the parts have significant contributions to the overall performance and underscores the significance of their joint functioning. Even though the framework is accompanied by a middle level rise in inference latency, its precision and strength increase is worth the extra computational expenditure. On balance, the suggested methodology is an effective and generalizable way to improve multi-step reasoning in big language encoders.

6 Conclusion and Future Work

This paper offered a Meta-Reasoning Learning Framework to improve multi-step reasoning abilities of large language models by means of incorporating organized planning, progressive inference, and self-reflective validation. The proposed framework enhances the accuracy and consistency of inference by explicitly modeling the reasoning process and addressing the thorough refinement process with the use of feedback. Experiments on a variety of benchmark tasks showed that the method does better than standard baselines, with strong predictive performance improvements with very high reasoning qualities. The results reveal the significance of explicit and analyzable reasoning routes in overcoming the shortcomings of implicit inference of large language models, which can be used to produce more dependable and understandable AI systems.

Subsequent efforts will be put into making the computational power of the framework more efficient in terms of minimizing the time required to inference and deploying it in real-time. Also, it still is possible to extend the framework to multimodal reasoning and investigate the adaptive strategies of planning dynamic problem-solving situations, which seems to be excellent directions in research.

Author Contributions

C Ramakrishna formulated the research problem and created the proposed methodology and oversaw the entire study. VNVLS Swathi undertook the implementation, experimentation, data analysis and results and visualization preparation. K Naga Maha Lakshmi provided input to the literature review, write-up and editing of the methodology and experimental section. Final manuscript was reviewed, edited and accepted by all the authors.

Originality and Ethical Standards: The authors verify that this is an original work that has not been published or entered into publications. Any sources that have been used in this study have been duly referenced and given credit. This study has been done in accordance with the normal academic and ethical protocols. There was no plagiarism, fabrication of data or other unethical activities in this work.

Data availability: Data available upon request.

Conflict of Interest: There is no conflict of Interest.

Funding: The research received no external funding.

Similarity checked: Yes.

References

[1] T. B. Brown *et al.*, "Language Models are Few-Shot Learners," *arXiv [cs.CL]*, 2020.

[2] T. Wolf et al., "Transformers: State-of-the-Art Natural Language Processing," Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, pp. 38–45, 2020, doi: 10.18653/v1/2020.emnlp-demos.6.

[3] M. A. K. Raiaan et al., "A Review on Large Language Models: Architectures, Applications, Taxonomies, Open

Issues and Challenges," IEEE Access, vol. 12, pp. 26839–26874, 2024, doi: 10.1109/access.2024.3365742.

[4] W. F. Wiggins and A. S. Tejani, "On the Opportunities and Risks of Foundation Models for Natural Language Processing in Radiology," Radiology: Artificial Intelligence, vol. 4, no. 4, Jul. 2022, doi: 10.1148/ryai.220119.

[5] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig, "Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing," ACM Computing Surveys, vol. 55, no. 9, pp. 1–35, Jan. 2023, doi: 10.1145/3560815.

[6] C. Liu, C. Gao, Y. Li, and M. Wang, "Exploring Implicit Sentiment Analysis Through Progressive Prompting Strategies," IEEE Transactions on Computational Social Systems, pp. 1–13, 2026, doi: 10.1109/tcss.2026.3658513.

[7] R. Kamoi, Y. Zhang, N. Zhang, J. Han, and R. Zhang, "When Can LLMs Actually Correct Their Own Mistakes? A Critical Survey of Self-Correction of LLMs," Transactions of the Association for Computational Linguistics, vol. 12, pp. 1417–1440, 2024, doi: 10.1162/tacl_a_00713.

[8] G. Lin and Y. Zhang, "Sparks of Artificial General Recommender (AGR): Experiments with ChatGPT," Algorithms, vol. 16, no. 9, p. 432, Sep. 2023, doi: 10.3390/a16090432.

[9] F. Cassano, A. Gopinath, K. Narasimhan, N. Shinn, and S. Yao, "Reflexion: language agents with verbal reinforcement learning," Advances in Neural Information Processing Systems 36, pp. 8634–8652, 2023, doi: 10.52202/075280-0377.

[10] D. Gunning and D. W. Aha, "DARPA's Explainable Artificial Intelligence Program," AI Magazine, vol. 40, no. 2, pp. 44–58, Jun. 2019, doi: 10.1609/aimag.v40i2.2850.

[11] Z. C. Lipton, "The Mythos of Model Interpretability," Queue, vol. 16, no. 3, pp. 31–57, Jun. 2018, doi: 10.1145/3236386.3241340.

[12] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, Proceedings of the 2019 Conference of the North, pp. 4171–4186, 2019, doi: 10.18653/v1/n19-1423.

[13] W. X. Zhao et al., "A survey of large language models," ACM Computing Surveys, vol. 56, no. 9, 2024. DOI: 10.1145/3586165

[14] Y. Bang et al., "A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity," Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 675–718, 2023, doi: 10.18653/v1/2023.ijcnlp-main.45.

[15] R. Bommasani et al., "On the opportunities and risks of foundation models," Communications of the ACM, vol. 66, no. 7, pp. 68–79, 2023. DOI: 10.1145/3597300

[16] L. Xin, W. Song, Z. Cao, and J. Zhang, "Step-Wise Deep Learning Models for Solving Routing Problems," IEEE Transactions on Industrial Informatics, vol. 17, no. 7, pp. 4861–4871, Jul. 2021, doi: 10.1109/tii.2020.3031409.

[17] Y. F. Song, H. Zhang, and J. Liu, "A communication theory perspective on prompting in large language models," Frontiers of Computer Science, vol. 18, 2024. DOI: 10.1007/s11390-024-4058-8

[18] S. Sivarajkumar, M. Kelley, A. Samolyk-Mazzanti, S. Visweswaran, and Y. Wang, "An Empirical Evaluation of Prompting Strategies for Large Language Models in Zero-Shot Clinical Natural Language Processing: Algorithm Development and Validation Study," JMIR Medical Informatics, vol. 12, p. e55318, Apr. 2024, doi: 10.2196/55318.

[19] A. Jain, Z. Tang, and A. Gurung, "Towards an Effective Prompting With AI Systems: The Jujitsu Framework," Journal of Applied Business and Economics, vol. 27, no. 3, Jun. 2025, doi: 10.33423/jabe.v27i3.7691.

[20] Z. Chu et al., "Self-Critique Guided Iterative Reasoning for Multi-hop Question Answering," Findings of the Association for Computational Linguistics: ACL 2025, pp. 2415–2438, 2025, doi: 10.18653/v1/2025.findings-acl.123.

[21] L. Li et al., "Uncertainty-Aware Iterative Preference Optimization for Enhanced LLM Reasoning," Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 23996–24012, 2025, doi: 10.18653/v1/2025.acl-long.1169.

[22] G. Vilone and L. Longo, "Classification of Explainable Artificial Intelligence Methods through Their Output Formats," Machine Learning and Knowledge Extraction, vol. 3, no. 3, pp. 615–661, Aug. 2021, doi: 10.3390/make3030032.

[23] M. H. G. Hoffmann, "Stimulating Reflection and Self-correcting Reasoning Through Argument Mapping: Three Approaches," Topoi, vol. 37, no. 1, pp. 185–199, Jun. 2016, doi: 10.1007/s11245-016-9408-x.

[24] M. Ramachandran, "Quality framework for explainable artificial intelligence (XAI) and machine learning applications," Explainable Artificial Intelligence (XAI): Concepts, enabling tools, technologies and applications, pp. 115–138, Nov. 2023, doi: 10.1049/pbpc062e_ch7.

[25] Y. Chudasama, H. Huang, D. Purohit, and M.-E. Vidal, "Toward Interpretable Hybrid AI: Integrating Knowledge Graphs and Symbolic Reasoning in Medicine," IEEE Access, vol. 13, pp. 39489–39509, 2025, doi: 10.1109/access.2025.3529133.

[26] B. Liang, Y. Wang, and C. Tong, "AI Reasoning in Deep Learning Era: From Symbolic AI to Neural–Symbolic AI," Mathematics, vol. 13, no. 11, p. 1707, May 2025, doi: 10.3390/math13111707.

[27] https://github.com/openai/grade-school-math

[28] https://github.com/hendrycks/math

[29] https://github.com/eladsegal/strategyqa

[30] P. Upadhyay, R. Agarwal, S. Dhiman, A. Sarkar, and S. Chaturvedi, "A comprehensive survey on answer generation methods using NLP," Natural Language Processing Journal, vol. 8, p. 100088, Sep. 2024, doi: 10.1016/j.nlp.2024.100088.

[31] A. Plaat, A. Wong, S. Verberne, J. Broekens, N. Van Stein, and T. Bäck, "Multi-Step Reasoning with Large Language Models, a Survey," ACM Computing Surveys, vol. 58, no. 6, pp. 1–35, Dec. 2025, doi: 10.1145/3774896.

[32] F. Xu et al., "Toward large reasoning models: A survey of reinforced reasoning with large language models," Patterns, vol. 6, no. 10, p. 101370, Oct. 2025, doi: 10.1016/j.patter.2025.101370.

[33] P. Schober and T. R. Vetter, "Linear Regression in Medical Research," Anesthesia & Analgesia, vol. 132, no. 1, pp. 108–109, Dec. 2020, doi: 10.1213/ane.0000000000005206.

[34] Y. Wu and J. Wan, "A survey of text classification based on pre-trained language model," Neurocomputing, vol. 616, p. 128921, Feb. 2025, doi: 10.1016/j.neucom.2024.128921.