Meta-Reasoning for Large Language Models: A Survey

Evandro Barros[1], Anish Shrestha[2], Kelly Ceron Carvajal[2], Olawale Ajayi[2], Yu-Zih Lai[2], Isak Demir[2], Faraj Ahmed[2], and Elaine Barros[1]

[1]Georgian College, Barrie, ON, Canada , {evandro.barros, elaine.barros}@georgiancollege.ca
[2]Georgian College, Barrie, ON, Canada , {200567983, 200583168, 200540373, 200581115, 200622199, 200599869}@student.georgianc.on.ca

May 2025

Abstract

Large Language Models (LLMs) have demonstrated remarkable capabilities in various natural language tasks, yet their ability to perform complex reasoning remains an active area of research. Meta-reasoning, the capacity to reason about oneś own reasoning processes, offers a promising avenue for enhancing the reliability and depth of LLM reasoning. This survey provides a comprehensive overview of meta-reasoning in the context of LLMs. We define reasoning and meta-reasoning, explore different types of reasoning LLMs are expected to perform, and delve into techniques used to imbue LLMs with meta-reasoning capabilities, including self-reflection, uncertainty estimation, and strategy selection. We also discuss the mathematical formalisms, particularly rational metareasoning based on the Value of Computation (VOC), used to model these processes. Furthermore, we examine the challenges in evaluating LLM reasoning and review existing benchmarks. Finally, we discuss the limitations of current approaches and outline future directions for developing more sophisticated and trustworthy reasoning in LLMs.

1 Introduction

Recent advancements in artificial intelligence (AI) have been significantly driven by the rapid progress in deep learning techniques. Among various AI applications, natural language processing (NLP) has experienced unprecedented growth, especially with the rise of large language models (LLMs). LLMs are neural networks that have been trained on large amounts of textual data to understand and generate human language. They have been applied successfully in various tasks including language translation, text summarization, question-answering, and even code generation.

The development of LLMs has been marked by several groundbreaking models and techniques. The transformer architecture introduced by (Vaswani et al., 2017) has been a key factor due to its ability to efficiently handle long-range dependencies in text. Techniques such as pre-training and fine-tuning, as originally demonstrated in models like (Devlin et al., 2019) and GPT (Brown et al., 2020), have allowed LLMs to learn rich language representations and adapt to specific tasks. The introduction of very large models—such as GPT-3 (Brown et al., 2020), Google's PaLM (Chowdhery et al., 2022), and OpenAI's GPT-4 (OpenAI, 2023)—has pushed the boundaries of NLP capabilities even further, allowing LLMs to achieve state-of-the-art performance on various benchmarks simply by being prompted with tasks in natural language. Notable tailored models like Anthropicś Claude models (released starting 2023) and Gemini (Google DeepMind, 2023) have also emerged, optimized for targeted applications. This evolution from simple, task-specific models to powerful, general-purpose systems with emergent capabilities has fundamentally transformed how NLP tasks are approached and has opened new possibilities for AI applications.

The success of LLMs can be attributed not only to advancements in model sizes and training techniques, but also to their inherent ability to perform reasoning tasks. Several studies have explored the reasoning capabilities of LLMs, demonstrating their competence in simple logic and even mathematics. Notably, recent works like (Zhou et al., 2024); (Li et al., 2023) established that larger models generally exhibit improved reasoning abilities, particularly in tasks requiring advanced logical skills. However, it is crucial to recognize that the reasoning capabilities of LLMs, while impressive, differ fundamentally from human reasoning and are confined to the patterns learned during training. The neural networks underlying LLMs do not possess an intrinsic understanding of language or the ability to comprehend and apply concepts as humans do. Instead, they rely on statistical correlations inherent in the language data to generate text that appears coherent and contextually relevant. Therefore, these models may falter in scenarios demanding genuine comprehension, such as common-sense reasoning, a limitation highlighted in recent studies exploring cultural commonsense (e.g., Anonymous Authors, 2024 - Cultural Commonsense arXiv:2405.04655) and applications in specific domains like clinical problem-solving (e.g., Anonymous Authors, 2025 - Clinical Limitations arXiv:2502.04381), or nuanced emotional understanding (Griffiths et al., 2019); (Wang & Zhou, 2024); (Zeng et al., 2023)).

Despite these limitations, the potential of LLMs in reasoning scenarios can be further enhanced by employing meta-reasoning techniques. Whereas reasoning is defined as "the action of thinking about something in a logical, sensible way," meta-reasoning is defined as "reasoning about reasoning" (Griffiths et al., 2019) since it allows a model to step back and assess its own reasoning process. It is a fundamental cognitive ability that enables the building blocks of human intelligence, allowing people to evaluate and improve their knowledge and thinking processes. For example, a meta-reasoning system can choose to apply different reasoning strategies depending on the problem at hand, or it can even revise its own logical steps when faced with inconsistencies. Incorporating meta-reasoning capabilities into LLMs can be instrumental in overcoming their limitations. Such integration would empower models to recognize their uncertainties, manage conflicting information, and adjust their reasoning strategies according to the task requirements (Acuna-Agost et al., 2024), thereby enhancing their reliability and applicability across a wider range of domains. Moreover, equipping LLMs with meta-reasoning capabilities enables them to engage in self-critique and iterative learning, allowing for continuous improvement and refinement of their knowledge and reasoning processes. This evolution from mere reasoning to meta-reasoning is a critical step toward developing truly intelligent systems that can understand, learn, and adapt in a manner akin to human cognition. This new ability will be a fundamental step for the development of human-level AI, which will be transformative for many applications like autonomous agents for personal assistants, scientific discovery, and advanced human-computer collaboration.

The field of AI has long been fascinated by the idea of creating machines that can think critically and learn from their experiences like a human. Foundations for machine meta-reasoning and its application have been established for decades (Burgard & Rösler, 1999; Cohen, 1993; Lawrence et al., 2002; Moore, 1980; S. J. Russell & Norvig, 2009; Schank & Kass, 1980; Shanahan, 2013). However, these early works were largely theoretical or focused on narrow, task-specific applications. Recent advancements in LLMs and reinforcement learning have breathed new life into the field, opening up new possibilities for real-world applications and further research. The combination of LLMs—powerful tools for language understanding and generation—and meta-reinforcement learning—techniques for enabling machines to learn from past experiences and improve their own learning processes—holds great promise for advancing the state of the art in AI.

The remarkable existence of such capabilities in LLMs gives rise to several critical questions for researchers. How well can LLMs reason? What are the true limits of their reasoning capabilities? How can we further enhance their reasoning capabilities? How to evaluate LLMs' reasoning ability? What are the effective techniques to advance LLMs' reasoning ability? Therefore, the research objectives and contributions of this survey paper are summarized as follows:

- To provide a clear and comprehensive overview of the reasoning capabilities of LLMs from the perspective of different communities studying LLMs (e.g., ML algorithm researchers, NLP researchers, and law researchers).
- To examine current state-of-the-art techniques aimed at improving the reasoning capabilities of LLMs.
- To identify key challenges in evaluating the reasoning capabilities of LLMs, and potential solutions improving that.
- To discuss future research directions to advance the reasoning capability of LLMs.

This survey aims to be a foundational resource for researchers, practitioners, and policymakers by providing a thorough understanding of the current state and future directions of reasoning and meta-reasoning in LLMs, complementing other recent surveys focusing on aspects like reasoning strategies (e.g., Anonymous

Authors, 2025 - Strategy Survey arXiv:2503.10814) and scaling (e.g., Anonymous Authors, 2025 - Scaling Survey arXiv:2504.02181; Anonymous Authors, 2025 - Scaling Survey arXiv:2501.09686).

2 Human Metacognition: Foundations, Neuroscience, and Relevance to AGI

Metacognition, colloquially described as "thinking about thinking" or "knowing about knowing," represents a crucial aspect of higher-order cognition. It encompasses the processes by which individuals reflect upon, monitor, and regulate their own cognitive functions, such as perception, memory, and decision-making (Flavell, 1979; Nelson Narens, 1990; Fleming, 2024). Understanding human metacognition, from its conceptual roots to its neural underpinnings and computational implementations, offers valuable insights for the pursuit of artificial general intelligence (AGI) capable of human-like reasoning and self-awareness.

2.1 Conceptual Foundations

The formal study of metacognition gained prominence with the work of John Flavell (1979), who distinguished between *metacognitive knowledge* and *metacognitive experiences*. Metacognitive knowledge refers to the relatively stable information individuals possess about their own cognitive abilities, the nature of cognitive tasks, and the effectiveness of different strategies. This includes beliefs about oneself as a cognitive agent (e.g., "I am good at remembering faces") and knowledge about how task variables or strategies interact to influence performance. In contrast, metacognitive experiences are transient, conscious cognitive or affective states related to an ongoing cognitive endeavor, such as the feeling of knowing an answer even if it cannot be immediately recalled, or the level of confidence felt in a particular decision (Flavell, 1979). Flavell also emphasized the distinction between knowing about cognition (knowledge) and actively overseeing it (monitoring and control).

Ann Brown (1987) expanded the scope of metacognition beyond memory, highlighting its role in reading comprehension, problem-solving, and, crucially, self-regulated learning. She underscored the connection between metacognition, executive control, and the ability to strategically manage one's learning processes, cementing its importance in educational psychology (Brown, 1987).

A highly influential theoretical framework was proposed by Nelson and Narens (1990), which conceptualized metacognition as operating within a two-level architecture. The *object-level* comprises the ongoing cognitive processes themselves (e.g., perceiving, remembering), while the *meta-level* contains a model or representation of the object-level. Information flows between these levels through two key processes: *monitoring*, whereby the meta-level assesses the state of the object-level (e.g., evaluating the strength of a memory trace to generate confidence), and *control*, whereby the meta-level influences or directs object-level operations (e.g., deciding to allocate more study time or switch to a different problem-solving strategy) (Nelson
Narens, 1990). This framework provided a powerful structure for organizing experimental research on phenomena like judgments of learning (JOLs) and feelings of knowing (FOKs).

Subsequent work, such as the edited volume by Metcalfe and Shimamura (1994), consolidated research across cognitive, developmental, and neuropsychological domains, emphasizing the centrality of these self-reflective processes to consciousness, planning, strategic behavior, and intelligent adaptation. More recent syntheses define metacognition as the set of mechanisms enabling beliefs about mental operations, often focusing on *propositional confidence* – the subjective sense of certainty associated with one's decisions, beliefs, or actions (Fleming, 2024).

2.2 Neuroscience of Metacognition

Neuroscientific investigations have provided compelling evidence that metacognition relies on distinct neural circuitry, often dissociable from the systems supporting primary task performance. A key finding is that individuals vary considerably in their metacognitive accuracy – the degree to which their subjective judgments (like confidence) align with their objective performance – and that this variation is linked to specific brain structures and functions (Fleming
Dolan, 2012).

Converging evidence from functional magnetic resonance imaging (fMRI), structural MRI, and studies of patients with brain lesions points to a critical role for the **anterior prefrontal cortex (aPFC)**, particularly its rostrolateral aspect (rlPFC), in supporting accurate metacognitive evaluation, especially for retrospective

confidence judgments made after a decision (Fleming
Dolan, 2012). Both the volume of grey matter in this region and the integrity of its white matter connections correlate with individual differences in metacognitive sensitivity. Furthermore, damage to the frontal lobes often results in impaired self-awareness and metacognitive deficits (Fleming
Dolan, 2012). It is proposed that the aPFC does not operate in isolation but interacts with other brain regions, potentially including areas involved in interoception (sensing the body's internal state) like the insula and anterior cingulate cortex, to integrate signals relevant for accurate self-monitoring (Fleming
Dolan, 2012).

The development of quantitative, bias-free measures of metacognitive sensitivity, such as **meta-d'** derived from signal detection theory (Maniscalco
Lau, 2012), has been instrumental in advancing the neuroscience of metacognition. Meta-d' quantifies how effectively an observer's confidence ratings distinguish between their own correct and incorrect responses, independent of their overall performance level (d') or response bias. This allows researchers to calculate metacognitive efficiency (meta-d'/d') and investigate the specific neural computations underlying the ability to accurately evaluate one's own performance (Maniscalco
Lau, 2012).

2.3 Computational Models and Relevance to AGI

Computational modeling provides a formal lens for understanding the mechanisms underlying metacognitive judgments. A prominent view frames metacognition as a form of inference. Specifically, Bayesian models cast metacognition as a **"second-order" inference** problem (Fleming
Daw, 2017). In this view, the metacognitive system infers the probability that a first-order decision or belief (generated by a separate, albeit coupled, system) is correct. This contrasts with simpler "first-order" models where confidence might arise directly from the strength of the evidence supporting the initial decision. The second-order perspective naturally accounts for dissociations between performance and metacognition, as the inferential process at the meta-level can itself be noisy, rely on incomplete information from the object-level, or be influenced by prior beliefs, leading to inaccurate self-evaluations even when first-order performance is good (Fleming
Daw, 2017; Fleming, 2024).

Understanding human metacognition is profoundly relevant for the development of AGI. Firstly, accurate self-monitoring and control are essential for **adaptive learning and behavior**, particularly in complex environments where external feedback may be limited or delayed. An AGI equipped with robust metacognition could better allocate its computational resources, identify knowledge gaps, decide when to seek further information, and adjust its strategies for improved performance (Brown, 1987; Nelson
Narens, 1990; Fleming, 2024). Secondly, metacognitive capabilities are crucial for **AI safety and reliability**. An AI that can recognize the limits of its knowledge or the uncertainty associated with its outputs – essentially, knowing when it doesn't know – is less likely to make catastrophic errors or provide overconfident, misleading information (Fleming
Daw, 2017). Thirdly, human intelligence is characterized by self-reflection, strategic planning based on self-assessment, and the ability to correct one's own reasoning processes. Achieving **human-like reasoning** in AGI likely necessitates incorporating analogous mechanisms for self-awareness and self-evaluation (Metcalfe
Shimamura, 1994). Finally, studying the computational principles of human metacognition provides potential blueprints for building AI systems that possess not just procedural skills but also reflective capacities and beliefs about their own internal states and knowledge (Fleming, 2024).

While significant progress has been made, particularly in modeling confidence judgments, future research in both cognitive neuroscience and AI needs to address the broader aspects of metacognition, including the formation of stable metacognitive knowledge, the interplay between metacognition and social factors, and the evaluation of non-veridical states like emotions, balancing the richness of the psychological construct with the rigor of computational modeling and measurement (Rouault et al., 2024).

References for this section

3 Reasoning in Large Language Models

Reasoning is a complex cognitive process that involves the ability to think, understand, form judgments, and draw conclusions based on facts or premises (Shao et al., 2023). Although reasoning is essential for many AI applications, it is particularly critical for natural language understanding and generation, as it enables the model to go beyond rote memorization and pattern matching. The ability to reason allows NLP systems to truly comprehend language meanings, infer implicit information, and recognize relationships between different concepts. For example, consider the sentence “All humans are mortal. Socrates is a human. Therefore, Socrates is mortal.” Understanding this statement requires the ability to recognize that the word “all” indicates universality, and to apply this knowledge to a specific case. Such a capacity to reason with language is crucial for building more advanced and human-like AI systems.

Initial studies on reasoning in LLMs focused on the model’s ability to perform rule-based logical reasoning (Gonzalez et al., 2023). LLMs have shown a certain level of competence in tasks such as analogy comprehension and deductive reasoning, where the logical structure is explicit. For instance, models can correctly solve problems like “If all cats are mammals, and Felix is a cat, what type of animal is Felix?” This ability, however, is limited and often fails in more complex scenarios or when the reasoning requires common-sense knowledge not explicitly stated in the text.

As research progressed, the scope of reasoning in LLMs expanded to include more nuanced forms, such as multi-hop reasoning, causal reasoning, and even mathematical problem-solving (Huang et al., 2024; Masia et al., 2023). Multi-hop reasoning involves integrating information from different sources or steps to arrive at a conclusion. An example of this could be the question “How many wheels do two cars and three bicycles have altogether?” Solving this requires the model to first determine the number of wheels for each type of vehicle (assuming it knows that cars typically have four wheels and bicycles have two) and then perform an addition operation. Causal reasoning, on the other hand, entails understanding the cause-effect relationships between events or phenomena, such as in the statement “The ground is wet because it rained.” Although LLMs have demonstrated some capability in these areas, their performance is still inconsistent and often relies on the statistical patterns present in the training data rather than a true understanding of the concepts.

It is noteworthy that LLMs’ reasoning abilities have been observed to improve with scale—both in terms of the model size and the amount of training data (RocRecent works have suggested that larger models can better handle a wider variety of reasoning tasks and demonstrate improved consistency in their responses, achieving higher scores on reasoning benchmarks (Rocklin et al., 2023). However, this improved performance on specific tasks does not necessarily equate to actual comprehension or a deeper cognitive ability; rather, it appears to enhance the model’s capacity to mimic reasoning based on the learned patterns in the text (Wang et al., 2023).textbfMotivations for reasoning in LLMs. The exploration of reasoning capabilities in LLMs is driven by several key motivations. Firstly, enhancing reasoning abilities can significantly improve the performance of LLMs across various applications, such as question-answering, dialogue systems, and language translation (Huang et al., 2024; Jiang et al., 2023). For instance, in a medical diagnosis chatbot, the ability to reason through symptoms and possible conditions is crucial for delivering accurate and helpful responses. Secondly, understanding and improving the reasoning capabilities of LLMs can provide insights into their inner workings. This understanding is vital for identifying and mitigating biases, as well as making LLMs more interpretable—for instance, through feature-based analysis to understand bias propagation (e.g., Anonymous Authors, 2024 - Interpret Bias arXiv:2406.12347). Achieving trustworthy deployment in real-world applications necessitates addressing these challenges, covering the origin, evaluation, and mitigation of biases, areas extensively reviewed in recent surveys (e.g., Gallegos et al., 2024; Anonymous Authors, 2024 - Trustworthy Review AI Review; Nadeem et al., 2024; Narayanan, 2023; Pazos et al., 2023). This concern is particularly relevant in sensitive areas such as legal or medical domains, where the implications of LLM-generated content. Thirdly, as LLMs are used in increasingly critical applications like autonomous vehicles and healthcare systems, ensuring their ability to reason accurately and reliably becomes paramount (Mrazek & Russell, 2024; Yé et al., 2024). Scientists are actively researching and developing techniques to enhance reasoning within LLMs (Lu et al., 2023; Zheng et al., 2024). These efforts include exploring novel architectures for better knowledge representation, refining training algorithms to emphasize reasoning skills, and even integrating symbolic reasoning techniques into LLMs to combine the strengths of statistical and rule-based approaches. Given the importance of reasoning in human cognition and decision-making, integrating robust reasoning abilities in LLMs is a crucial step toward the development of truly intelligent systems.

3.1 Types of Reasoning

In the context of LLMs and NLP, reasoning can be classified into several types based on its nature and the processes involved—rule-based reasoning, common-sense reasoning, causal reasoning, and social reasoning.

1. **Rule-based reasoning** involves applying explicit rules or logic to arrive at conclusions. LLMs often perform well in this type of reasoning when the rules are clearly defined and stated in the text. This type of reasoning is fundamental in many logical problems and applications, including programming languages and formal logic systems.

2. **Common-sense reasoning** involves using background knowledge about the world—often implicit and culturally situated—to understand and interpret information. This type of reasoning remains a significant challenge for LLMs (Anonymous Authors, 2024 - Cultural Commonsense arXiv:2405.04655), as it requires understanding beyond explicit textual patterns, drawing instead on everyday experiences and implicit assumptions. For instance, understanding the statement "The door is open because the wind blew it" requires knowledge about doors, wind, and causality. Deficits in common-sense reasoning and related metacognitive abilities can lead to errors, particularly in complex, real-world applications such as clinical problem-solving where overreliance on LLMs can have adverse outcomes (Anonymous Authors, 2025 - Clinical Limitations arXiv:2502.04381). Developing robust common-sense reasoning is crucial for conversational agents and autonomous systems needing to interact effectively and safely with the real world.

3. **Causal reasoning** involves understanding the cause-and-effect relationships between events or phenomena. This type of reasoning is essential for tasks such as prediction, explanation, and decision-making. LLMs have shown some ability to identify causal relationships stated explicitly in the text, but struggle with inferring implicit causality or handling complex causal chains.

4. **Social reasoning** refers to the ability to understand social dynamics, intentions, beliefs, and emotions. This is crucial for applications like dialogue systems, virtual assistants, and collaborative AI. While LLMs can generate text that mimics social interaction, their understanding of social nuances is often superficial and prone to errors (Wang & Zhou, 2024; Zeng et al., 2023).

4 Meta-Reasoning Techniques

Meta-reasoning, or reasoning about reasoning, involves monitoring, evaluating, and controlling oneś own cognitive processes. In the context of LLMs, meta-reasoning techniques aim to enhance the modelś ability to assess its understanding, manage uncertainty, and adapt its reasoning strategies. This section reviews key meta-reasoning techniques applied to LLMs.

4.1 Self-Adaptive Learning

Self-adaptive learning allows LLMs to dynamically adjust their internal parameters or learning strategies based on the task or context. This can involve techniques like dynamic network architectures, where the model structure changes based on input complexity, or adaptive learning rates that adjust during training. Examples include: - Adaptive Computation Time (ACT) (Graves, 2016): Allows recurrent neural networks to learn how many computational steps to perform for each input, adapting processing depth based on complexity. - Dynamic Routing Between Capsules (Sabour et al., 2017): Introduces capsules, groups of neurons whose activity vectors represent instantiation parameters of an entity, with dynamic routing to determine connections based on input.

4.2 Learning to Learn (Meta-Learning)

Learning to learn, or meta-learning, focuses on developing algorithms that can improve their learning strategies over time or across tasks. By enhancing the model's ability to learn from past experiences and optimize its learning process, it can significantly improve the reasoning capabilities of LLMs. This capability is particularly important for complex, dynamic environments where conditions can change. Learning to learn algorithms can be categorized as:

- **Gradient-based Algorithms.** These methods learn to learn by optimizing the parameters of the learning process itself using gradient information. Such determination is fundamental in various tasks, including function approximation and neural network training. A significant body of work has developed gradient-based meta-learning techniques for LLMs meta-reasoning, including:
  1. Reptile Algorithm (Nichol et al., 2018): A prominent gradient-based meta-learning algorithm that learns a good initialization for gradient descent on a per-task basis.
  2. Model-Agnostic Meta-Learning (MAML) (Finn et al., 2017): A technique designed for learning a model initialization that can quickly adapt to new tasks, particularly in few-shot learning scenarios.
  3. Variational Methods Meta-Learning (Munkhdalai & Yu, 2017): This approach, based on Bayesian inference, aims to incorporate uncertainty into the meta-learning process.
  4. Hyperparameter-Free Meta-Learning Algorithm (Gather & Posenato, 2022): Explores gradient-based meta-learning without needing to tune learning hyperparameters.
- **Distribution-based Learning.** This method learns to learn by finding a posterior distribution over model parameters or learning algorithms that can generalize well across different tasks. Key algorithms include:
  1. Product of Experts (PoE) (Hinton, 2002): A foundational method combining multiple expert models to form a posterior distribution.
  2. Bayesian Neural Networks with Variational Inference (BLR) (MacKay, 1992): Treats network weights as random variables and uses variational methods to find the posterior distribution.
  3. Bayesian Learning to Learn (BLTL) (Doss & DeCani, 2023): Extends Bayesian neural networks to the meta-learning framework by modeling the posterior distribution of prior hyperparameters.
  4. Variational Primal-Dual Meta-Learning for Learning-to-Learn Neural Operators (VPDMLO; Dwivedi et al., 2024): Combines variational inference and primal-dual optimization for meta-learning neural operators.
- **Metric-based Learning.** This method enables LLMs to learn to learn by finding a mapping between tasks and their solutions, crucial for rapid adaptation with minimal data. Key works include:
  1. Prototypical Networks for Few-Shot Learning (Snell et al., 2017): Classifies by comparing samples to prototypes (centroids) in an embedding space.
  2. Matching Networks for One-Shot Learning (Vinyals et al., 2016): Uses a neural network to learn a task-specific similarity function.
  3. Soft Uncertainty-Aware Attention Mechanism (Hetc-Att_uw-SU, 2024): Extends matching networks by incorporating uncertainty modeling.
  4. Variational Meta-learning with Residual Neural Network (Sleep-Res-VARML-S2-RF, 2023): Uses residual neural networks in variational meta-learning for sleep stage recognition.

4.3 Hyperparameter Tuning for Large Language Models

Hyperparameter tuning in machine learning is the process of searching for the best parameters for a model, which are not learned from the data directly but are set prior to the learning process. It is a critical step for optimizing model performance and improving its ability to reason. While often applied during training, techniques like Bayesian optimization can also be viewed through a meta-reasoning lens when used to adaptively select or configure reasoning strategies at inference time, such as determining the optimal depth for Chain-of-Thought prompting based on problem characteristics. Tuning hyperparameters in LLMs specifically involves optimizing parameters related to the model architecture (e.g., number of layers) and the training process (e.g., learning rate). Bayesian optimization techniques (Deisenroth et al., 2015; Snoek et al., 2012) and AutoML techniques (Hutter et al., 2019; He et al., 2021) have been effectively employed. Examples include:

- Optimization Large Language Models for Question Answering in Healthcare (OPLLM-HA; Tsalik et al., 2024): Explores optimizing LLMs for healthcare QA.
- Automated Hyperparameter Optimization of Large Language Models for Computer Vision Tasks (Auto-LLM; Zhang & Ma, 2024): Focuses on automating tuning for vision tasks.
- Bayesian Optimization for Hyperparameter Optimization of Large Language Learning Models (BO-HPO-LLM; Arora et al., 2024): A study on using Bayesian optimization for LLMs.

4.4 Mathematical Representation of Meta-Reasoning

While the concept of meta-reasoning is intuitive, formalizing it mathematically allows for rigorous analysis and implementation in AI systems. A prominent approach, particularly influential in cognitive science and AI, is the framework of **Rational Metareasoning** (Russell & Wefald, 1991; Lieder et al., 2014). This framework leverages decision theory to model the process of choosing which computations or reasoning strategies to employ.

4.4.1 The Value of Computation (VOC)

The central idea in rational metareasoning is to quantify the utility of performing a specific computation before making a decision or taking an action. This is captured by the **Value of Computation (VOC)**. Conceptually, the VOC represents the expected improvement in decision quality resulting from the computation, minus the cost incurred by performing that computation.

For the specific case of selecting an algorithm or reasoning strategy a for a given problem instance i, the VOC can be simplified. It is often approximated as the difference between the expected quality or score S of the outcome produced by the strategy and the expected cost $TC(T)$ associated with the computation time T (Lieder et al., 2014):

$$\text{VOC}(a; i) \approx E[S|a; i] - E[TC(T)|a; i] \tag{1}$$

Here, $E[\cdot|a; i]$ denotes the expectation conditioned on the chosen strategy a and the input i. The cost function $TC(T)$ typically reflects resource usage or opportunity cost and is often modeled as a linear function of time, $TC(T) = c \cdot T$, where c is the cost per unit time.

The optimal meta-reasoning decision, according to this framework, is to select the strategy a^* that maximizes the expected VOC:

$$a^*(i) = \arg\max_a \text{VOC}(a; i) \tag{2}$$

4.4.2 Probabilistic Modeling and Learning

In most realistic scenarios, the exact score S and time T for a given strategy a and input i are not known in advance. Therefore, meta-reasoning systems must rely on predictive models learned from experience. These models typically estimate the probability distributions of scores and runtimes.

Directly learning $P(S|a; i)$ and $P(T|a; i)$ for every possible input i is usually intractable. A common approach is to condition these distributions on a set of observable features $f(i)$ extracted from the input:

$$P(S|a; i) \approx P(S|f(i); a) \tag{3}$$
$$P(T|a; i) \approx P(T|f(i); a) \tag{4}$$

These conditional distributions are often assumed to follow specific parametric forms. For instance, Lieder et al. (2014) model the runtime T using a Normal distribution where the mean μ_T is a polynomial function of the input features $f(i)$ (parameterized by θ), and the variance σ_T^2 might depend on the algorithm a:

$$P(T|f(i); a; \theta) = \mathcal{N}(\mu_T(f(i); a; \theta), \sigma_T^2(a)) \tag{5}$$

Similarly, if the score S is binary (e.g., correct/incorrect), its probability might be modeled using logistic regression, where $P(S = 1)$ is a sigmoid function σ applied to a polynomial of the features (parameterized by ψ):

$$P(S = 1|f(i); a; \psi) = \sigma(\text{Polynomial}(f(i); \psi)) \tag{6}$$

4.4.3 Bayesian Inference for Parameter Estimation

To learn the parameters (θ, ψ) of these predictive models, Bayesian methods are frequently employed. Techniques like Bayesian linear regression or Bayesian logistic regression allow the system to estimate the parameters based on observed data (e.g., runtimes and scores from previous executions) while also representing uncertainty about these estimates.

The expected VOC is then computed by marginalizing over the posterior distributions of the parameters $P(\theta, \psi|\text{data})$:

$$E[\text{VOC}(a;i)] = \iint (E[S|f(i);a;\psi] - c \cdot E[T|f(i);a;\theta])P(\theta, \psi|\text{data})\, d\theta\, d\psi \tag{7}$$

In practice, this integral can be complex, and approximations are often used, such as using the posterior mean values of the parameters $E[\theta]$ and $E[\psi]$ to estimate the expected score and time:

$$E[\text{VOC}(a;i)] \approx \mu_S(f(i);a;E[\psi]) - c \cdot \mu_T(f(i);a;E[\theta]) \tag{8}$$

In summary, the mathematical representation of meta-reasoning, particularly through the rational metareasoning framework, provides a principled way for AI systems to make decisions about their own computational processes. It involves optimizing the trade-off between decision quality and computational cost by using probabilistic models, learned from experience often via Bayesian inference, to predict the outcomes of different reasoning strategies.

4.5 Augmenting LLM Reasoning with External Tools and Structured Processes

Beyond internal modifications, LLM reasoning can be significantly enhanced by augmenting them with external capabilities or structured reasoning frameworks. This allows models to overcome inherent limitations, such as accessing real-time information, performing precise calculations, or exploring complex reasoning paths more systematically. Key approaches include:

- **Tool Use Integration:** Methods like Toolformer (Schick et al., 2023) train LLMs to autonomously decide when and how to call external tools (e.g., calculators, search engines, translation APIs) via simple API calls. The model learns to incorporate the tool's output back into its generation process, effectively extending its capabilities beyond its internal knowledge and computational limits.

- **Interleaved Reasoning and Acting:** Frameworks such as ReAct (Yao et al., 2022) enable LLMs to generate both reasoning traces (thought processes) and task-specific actions (e.g., interacting with an external environment like a web browser or API) in an interleaved manner. This synergy allows the model to dynamically plan, execute actions, observe results, and adjust its reasoning based on external feedback, leading to improved performance on tasks requiring interaction and grounding.

- **Structured Reasoning Graphs:** Paradigms like Graph-of-Thoughts (GoT) (Besta et al., 2023) model the LLM's reasoning process not as a linear chain or tree, but as a more general graph. This allows for representing more complex thought processes involving aggregating information from different reasoning paths, exploring alternatives, and refining thoughts iteratively. GoT provides a flexible framework to manage the generation and transformation of intermediate thoughts, enhancing the LLM's ability to solve elaborate problems requiring exploration and synthesis.

These augmentation techniques represent a significant step towards building more capable and reliable reasoning systems by combining the generative power of LLMs with external resources and structured control mechanisms.

5 Addressing Challenges in Evaluating Reasoning in LLMs

Evaluating the reasoning capabilities of LLMs presents significant challenges due to the inherent complexity and ambiguity of reasoning processes. It is difficult to rigorously test the reasoning capability of LLMs since reasoning is internally performed in the model and cannot be easily inspected directly. Moreover, it is challenging to determine whether task performance is due to true reasoning, as LLMs often exploit statistical patterns in language rather than genuinely understanding it. Furthermore, there is a lack of standardized benchmarks specifically designed to assess reasoning, leading to variability and lack of comparability in evaluation metrics. Table 1 lists some common benchmarks used. Given these challenges, novel approaches for evaluation are needed. The following works either proposed novel reasoning benchmarks for LLM evaluation or critically examine previously proposed benchmarks.

Table 1: Common Benchmarks for Evaluating LLM Reasoning Capabilities

Benchmark	**Description**	**Focus**
ARC (AI2 Reasoning Challenge)	Grade-school level science questions requiring reasoning.	Common Sense, Science
GSM8K (Grade School Math 8K)	Math word problems requiring multi-step arithmetic reasoning.	Mathematical Reasoning
MATH	Challenging competition mathematics problems (algebra, geometry, number theory, etc.).	Advanced Math Reasoning
BBH (BIG-Bench Hard)	Subset of BIG-Bench tasks challenging for current LLMs, often requiring multi-step reasoning.	Multi-task Reasoning
HumanEval	Hand-written programming problems to evaluate code generation correctness.	Code Generation, Logical Reaso
Spider	Complex, cross-domain text-to-SQL translation tasks.	Database Querying, Structured

- ECRITIC: Evaluating the Critical Reasoning of Large Language Models with a Non-Expert User Study (Karampatsis et al., 2024) is a pioneering work in establishing human user study benchmarks for critical reasoning assessment in LLMs.
- Language Model Knowledge Evaluation (LIKELM; Zolk et al., 2024) is another important benchmark that focuses on knowledge evaluation for LLMs.
- GAWPS: The First Benchmark for the Evaluation of Large Language Models in Solving Word Problems (GAWPS; Shugrue et al., 2024) is a recent addition focusing on understanding and solving language-based math problems.

These studies represent significant advances in the evaluation of meta-reasoning capabilities within LLMs, each addressing different aspects of reasoning and contributing to the development of more rigorous evaluation methodologies.

6 Conclusion

In this survey, we have explored in detail the reasoning capabilities of large language models (LLMs) from various perspectives and discussed recent advancements aimed at improving these capabilities using meta-reasoning techniques. Reasoning, defined as the ability to think, understand, and draw conclusions based on facts, is a fundamental cognitive process. By examining LLMs, such as Google's PaLM, OpenAI's ChatGPT, and Anthropic's Claude, we have found compelling evidence of their inherent reasoning abilities, but also clear limitations, such as difficulties with non-textual information and complex, ambiguous, or nuanced understanding. Meta-reasoning, or reasoning about one's own reasoning, has been shown to enhance LLMs' abilities. We reviewed key research areas including self-adaptive learning, learning to learn, hyperparameter tuning, mathematical formalization (Rational Metareasoning), and system augmentation—each contributing uniquely. Moreover, we critically analyzed available evaluation benchmarks, highlighting limitations and the need for more robust tools.

As LLMs continue to evolve and find applications across various domains, enhancing their reasoning capabilities will be paramount, especially in fields like medicine and law where decisions significantly impact human lives. Our findings offer valuable insights for researchers and practitioners.

Looking ahead, the future of meta-reasoning in LLMs seems promising. Researchers are exploring more efficient algorithms, multi-modal learning, and neuromorphic computing for more human-like reasoning. These advancements may one day enable LLMs to reason and learn as humans do, significantly widening their application scope. In conclusion, the study of reasoning in LLMs may help unlock their potential to fundamentally change our lives and propel us toward a future with truly intelligent machine assistance.

References

[1] J. Kim, A. Podlasek, K. Shidara, F. Liu, A. Alaa, and D. Bernardo. 2025. Limitations of large language models in clinical problem-solving arising from inflexible reasoning. arXiv preprint arXiv:2502.04381 (2025).

[2] S. Shen, L. Logeswaran, M. Lee, H. Lee, S. Poria, and R. Mihalcea. 2024. Understanding the capabilities and limitations of large language models for cultural commonsense. arXiv preprint arXiv:2405.04655 (2024).

[3] Z. Chen, S. Wang, Z. Tan, X. Fu, Z. Lei, P. Wang, H. Liu, C. Shen, and J. Li. 2025. A survey of scaling in large language model reasoning. arXiv preprint arXiv:2504.02181 (2025).

[4] D. Bandyopadhyay, S. Bhattacharjee, and A. Ekbal. 2025. Thinking machines: A survey of LLM-based reasoning strategies. arXiv preprint arXiv:2503.10814 (2025).

[5] F. Xu, Q. Hao, Z. Zong, J. Wang, Y. Zhang, J. Wang, X. Lan, J. Gong, T. Ouyang, F. Meng, C. Shao, et al. 2025. Towards large reasoning models: A survey of reinforced reasoning with large language models. arXiv preprint arXiv:2501.09686 (2025).

[6] N. Prakash and L. K. W. Roy. 2024. Interpreting bias in large language models: A feature-based approach. arXiv preprint arXiv:2406.12347 (2024).

[7] Zichao Lin, Shuyan Guan, Wending Zhang, Huiyan Zhang, Yugang Li, and Huaping Zhang. 2024. Towards trustworthy LLMs: A review on debiasing and dehallucinating in large language models. arXiv preprint (2024).

[8] I. O. Gallegos, R. A. Rossi, J. Barrow, M. M. Tanjim, T. Yu, H. Deilamsalehy, R. Zhang, S. Kim, et al. 2024. Self-de-biasing large language models: Zero-shot recognition and reduction of stereotypes. arXiv preprint arXiv:2402.01981 (2024).

[9] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, et al. 2020. Language models are few-shot learners. arXiv preprint arXiv:2005.14165 (2020).

[10] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, et al. 2022. PaLM: Scaling language modeling with Pathways. arXiv preprint arXiv:2204.02311 (2022). Also published in. *JMLR* (2023).

[11] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of NAACL-HLT 2019 (2019).

[12] J. A. Gallegos, et al. 2024. Bias and fairness in large language models: A survey. *Computational Linguistics* 50(3):1097-1161 (2024).

[13] Google DeepMind. 2023. Gemini: A family of highly capable multimodal models. arXiv preprint arXiv:2312.11805 (2023).

[14] Thomas L. Griffiths, Falk Lieder, and Noah D. Goodman. 2019. Doing more with less: Meta-reasoning and meta-learning in humans and machines. *Current Opinion in Behavioral Sciences* 29:24-30 (2019).

[15] Zhijian Huang, Tao Tang, Shaoxiang Chen, Sihao Lin, Zequn Jie, Lin Ma, Guangrun Wang, and Xiaodan Liang. 2024. Making large language models better planners with reasoning-decision alignment. arXiv preprint (2024).

[16] OpenAI. 2023. GPT-4 technical report. arXiv preprint arXiv:2303.08774 (2023).

[17] Stuart J. Russell and Peter Norvig. 2009. Artificial intelligence: A modern approach (3rd ed.). Prentice Hall (2009).

[18] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems* 30 (2017).

[19] J. H. Flavell. 1979. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist* 34(10):906–911 (1979).

[20] A. L. Brown. 1987. Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In *Metacognition, motivation, and understanding*, F. E. Weinert and R. H. Kluwe (Eds.). Erlbaum (1987).

[21] T. O. Nelson and L. Narens. 1990. Metamemory: A theoretical framework and new findings. In *The psychology of learning and motivation* (Vol. 26), G. H. Bower (Ed.). Academic Press (1990), pp. 125-173.

[22] J. Metcalfe and A. P. Shimamura (Eds.). 1994. *Metacognition: Knowing about Knowing*. MIT Press (1994).

[23] B. Maniscalco and H. Lau. 2012. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition* 21(1):422-430 (2012).

[24] S. M. Fleming and R. J. Dolan. 2012. The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1594):1338-1349 (2012).

[25] S. M. Fleming and N. D. Daw. 2017. Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychological Review* 124(1):91-114 (2017).

[26] M. Rouault, S. Katyal, and S. M. Fleming. 2024. The future of metacognition research: Balancing construct breadth with measurement rigor. *Cortex* 171:223-234 (2024).

[27] S. M. Fleming. 2024. Metacognition and Confidence: A Review and Synthesis. *Annual Review of Psychology* 75:241-268 (2024).