

Review

Not peer-reviewed version

Structured Reasoning with Large Language Models

[Srihari Tanmay Karthik Tadala](#)*

Posted Date: 28 May 2025

doi: 10.20944/preprints202505.2095.v1

Keywords: Large Language Models



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Structured Reasoning with Large Language Models

Srihari Tanmay Karthik Tadala

Portland State University; stadala@pdx.edu

Abstract: Research on reasoning processes is still under progress, Large Language Models (LLMs) have demonstrated remarkable natural language processing capacity recently. Emphasizing multi-step problem-solving, organized decision-making, and human feedback alignment, this paper critically reviews eight foundational works supporting the evolution of LLM reasoning. This paper investigates how generative pre-training (GPT-1, GPT-2) supports unsupervised and zero-shot learning after building parallelizable and scalable self-attention mechanisms with the Transformer architecture. Thanks to the invention of Chain-of-Thought (CoT) prompting, which demonstrated that sequential thinking increases logical coherence, LLMs can now study numerous paths of reasoning. Tree of Thoughts (ToT) later grew out from this. Reinforcement Learning from Human Feedback (RLHF) has been essential in improving LLM alignment beyond prompting strategies; Prototypical Reward Models (Proto-RM) improve the efficacy of learning from human preferences. Retrieval-Augmented Thought Trees (RATT) also solve the problem of factual consistency by including outside knowledge sources; Thought Space Explorer (TSE) increases cognitive exploration and lets LLMs find fresh ideas. By combining these approaches, this study reveals new tendencies, points out ongoing difficulties, and offers a comparative analysis of organized thinking in LLMs, so setting the groundwork for further advancements in artificial intelligence driven reasoning models. This report provides a summary of key methodologies presented in eight foundational papers, focusing on their evolution and impact on LLM reasoning.

Keywords: large language models

1. Introduction

1.1. Background

Large Language Models (LLMs) have transformed natural language processing (NLP) by proving amazing powers in text production, translation, summarizing, and conversational artificial intelligence [1,2]. Since it introduced self-attention strategies that let long-range dependencies in text be efficiently handled, the Transformer architecture is largely responsible for these developments [1]. Generative Pre-Training (GPT-1, GPT-2) and other innovations brought a paradigm move toward unsupervised learning, allowing LLMs to generalize over many language difficulties without significant labeled input [2,3]. Despite their successes, LLMs suffer with complex reasoning, factual dependability, and multi-step decision-making, therefore limiting their value in fields requiring organized logic and interpretability [4,5]. By increasing LLM capacity to provide logically consistent, structured, and factually accurate solutions, recent approaches including Chain-of-Thought (CoT), Tree of Thought (ToT), and Retrieval-Augmented Thought Trees (RATT) seek to close these gaps [4–6]. Ensuring these models can explore many avenues of reasoning, self-correct mistakes, and effectively incorporate human feedback which remains difficult [6–8].

1.2. Research Motivation

By demonstrating incredible capabilities in text generation, translation, summarizing, and conversational artificial intelligence, Large Language Models (LLMs) have revolutionized natural language processing (NLP) [1,2]. The Transformer architecture is mostly responsible for these developments

since it brought self-attention techniques that let long-range dependencies in text be effectively processed [1]. Additional developments like Generative Pre-Training (GPT-1, GPT-2) brought a paradigm change toward unsupervised learning, enabling LLMs to generalize across several language problems without large labeled data [2,3]. LLMs suffer with complicated reasoning, factual dependability, and multi-step decision-making despite these achievements, therefore restricting their usefulness in jobs needing structured logic and interpretability. By improving LLM capacity to produce logically consistent, structured, and factually accurate responses, recent approaches including Chain of Thought (CoT) prompting, Tree of Thoughts (ToT), and Retrieval-Augmented Thought Trees (RATT) try to close these gaps [4–6]. Ensuring these models can negotiate many reasoning paths, self-correct mistakes, and effectively use human feedback remains difficult, though [6–8].

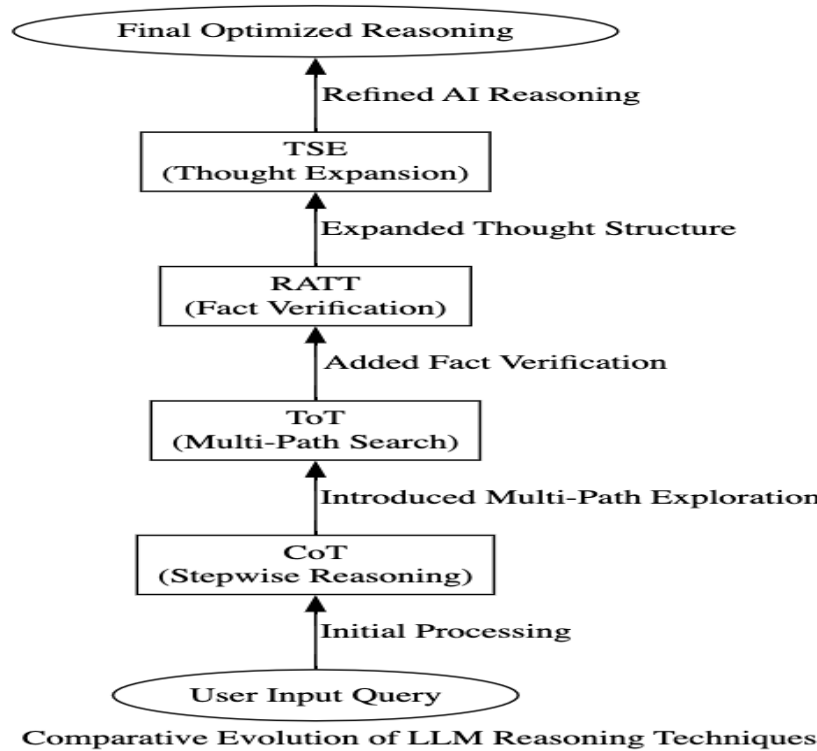


Figure 1. Evolution of LLM Reasoning

1.3. Comparison of Different Reasoning Techniques

As LLMs have developed toward improving factual accuracy, organized decision-making, and logical coherence, a range of approaches has been embraced as reasoning strategies [4? –7]. Early methods with as their primary objectives Transformer-based architectures and generative pre-training aim at contextual link identification in text [1–3]. Other specialized systems including Thought Space Explorer (TSE), Chain-of Thought (CoT), Tree of Thoughts (ToT), and Retrieval-Augmented Thought Trees (RATT) appeared as models as their intrinsic shortcomings in structured multi-step problem-solving and logical reasoning became clear [4–6,8] .

Chain of Thought (CoT) prompting, which demonstrated that deconstructing issues into explicit, sequential reasoning chains increases performance on demanding tasks including arithmetic, common sense, and symbolic reasoning, was the first major development in LLM reasoning. Still, the degree of CoT depends on the size of the model since only larger LLMs (100B+ parameters) shown emergent thinking [4]. CoT was limited to linear thought chains hence even if it offered ordered issue decomposition, it could not assess many feasible solutions.

Tree of Thoughts (ToT) gets over the single-path restriction of CoT by letting LLMs branch into several possible reasoning paths before choosing the best one. Using Depth-First Search (DFS) and Breadth-First Search (BFS), ToT was able to more precisely negotiate ordered problem-solving than

CoT. Decisions, long-term planning, and game-solving techniques all improved. The fundamental flaw with ToT, though, was that it was prone to hallucinations and incorrect thinking because it lacked mechanisms of factual validation [5].

Retrieval-Augmented Thought Trees (RATT) broke free from this limitation by including outside knowledge retrieval into their logical process. RATT used Retrieval-Augmented Generation (RAG) to enable fact-checking at all levels of the reasoning process unlike ToT, which just applied internal model-generated logic. This guaranteed that LLM produced logically coherent as well as factually correct results. Real time retrieval via RATT significantly lowered hallucinating rates and improved the dependability of LLMs for knowledge intensive activities including scientific reasoning, legal analysis, and medical artificial intelligence [6].

Additionally showing progress are prototypical reward networks (Proto-RM), a type of reinforcement learning from human feedback (RLHF). Great volumes of human-labeled data are needed for standard RLHF models, which drive alignment inefficaciousness and costs. Proto-RM maximized reward model efficiency and allowed less labeled data to more precisely generalize human preferences by way of aggregating human comments into representative prototypes [7]. This method produced more consistent and human-aligned incentive systems, therefore improving LLM performance on activities based on preferences and ethically sensitive ones.

At last, Thought Space Explorer (TSE) let LLMs dynamically extend their cognitive exploration space, so transcending logical problem-solving. While traditional wisdom emphasizes on sharpening preexisting thought channels, TSE presented a novel way by actively searching for unexplored logical paths. CoT and ToT follow this approach. This framework is notably useful for creative problem-solving, symbolic thinking, and scientific discovery since it protected LLMs from being bound in established cognitive frameworks and let them uncover hitherto unconsidered answers [8].

The pathway from structured, human-guided prompting (CoT) to multi-path structured thinking (ToT) demonstrates a different direction in LLM reasoning research in retrieval-enhanced validation (RATT), and autonomous cognitive expansion (TSE). These developments taken collectively solve the fundamental issues of LLM reasoning, therefore ensuring improved coherence, factual accuracy, and problem-solving efficiency [4–6,8].

The Figure 2 shows multiple LLM reasoning approaches Chain-of- Thought (CoT), Tree-of-Thought (ToT), Retrieval-Augmented Thought Trees (RATT), and Thought Space Explorer (TSE) show how the same mathematical inquiry is handled differently [4–6,8]. CoT applies a linear reasoning approach using BODMAS rules. [4] Among the numerous paths ToT follows results in the appropriate reaction [5]. RATT gathers outside knowledge to verify the computation and guarantees accuracy [6]. Before deciding the appropriate answer TSE looks at several mathematical ideas and abstraction techniques [8]. Despite their different approaches, all techniques finally produce the correct response (42), therefore highlighting the benefits and drawbacks of structured LLM reasoning models [4–6,8].

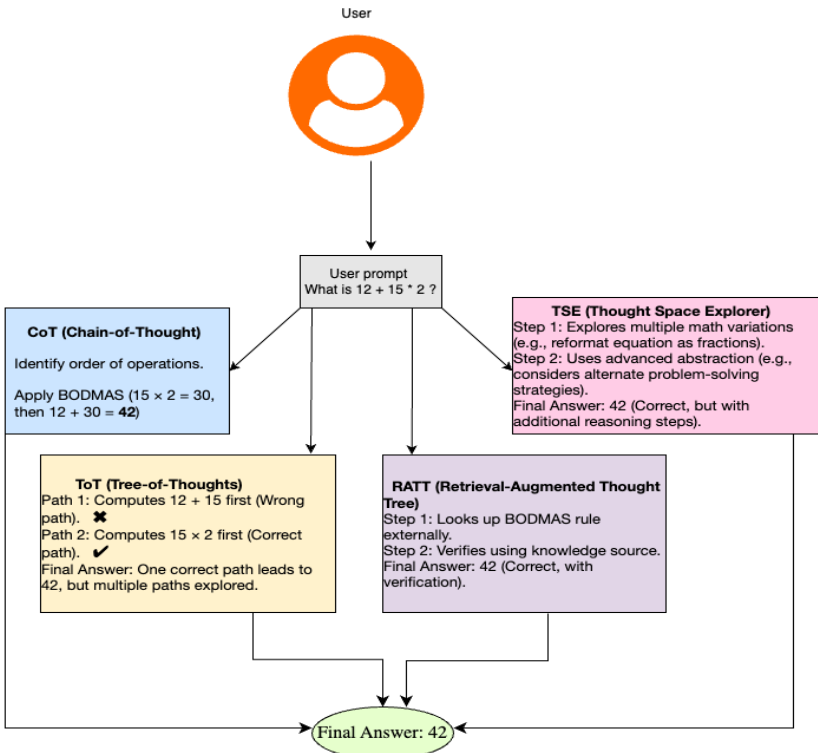


Figure 2. Flowchart of LLM Reasoning Techniques

1.4. How These Methods Evolved Over Time

Large Language Models (LLMs) have evolved their thinking strategies to produce a range of organized approaches each tackling particular difficulties in logical coherence, factual correctness, and problem-solving capacity. The comparison of several methodologies reveals important advantages, constraints, and strategies to help each other in developing LLM ideas [4–6,8].

(1) Chain-of-Thought (CoT) vs Tree-of-Thoughts (ToT):

Explicit, step-by-step thinking from CoT greatly enhanced multi-step mathematical, logical, and common sense reasoning. But its linear approach restricted its capacity to investigate other paths of thought, which resulted in mistakes [4]. By use of Depth-First Search (DFS) and Breadth-First Search (BFS), ToT overcomes this problem by allowing models to retrace and refine their solutions utilizing several lines of thinking. Although ToT improved exploration and problem-solving, its lack of fact-verification techniques made it prone to hallucinations when handling knowledge-intensive jobs [5].

(2) Retrieval-Augmented Thought Trees (RATT) and Reinforcement Learning (RLHF) Approaches:

Retrieval-augmented generation (RAG) was generated from T’s incapacity to verify the factual accuracy of the produced reasoning chains. By compiling outside input at every step of the reasoning process, RATT maintains the outcomes of the model factually sound unlike ToT, which rely simply on internal model knowledge [6]. RATT reduces hallucinations significantly, but it does not maximize incentive systems for precisely aligning ILM. To boost Reinforcement Instead of depending on large labeled datasets, Learning from Human Feedback (RLHF) learns from representative prototypes of human preference data by means of Prototypical Reward Model.

(3) Thought Space Explorer (TSE) vs Other Thought Structuring Methods:

Both CoT and ToT assume the produced logical routes of the model are sufficient to handle demanding tasks [4],5. However, they do not consider cognitive blind spots that is, situations in which restrictions in training data prevent models from producing new reasoning chains. Dynamic enlargement of thinking structures, uncovering undiscovered reasoning paths, and guiding LLMs toward varied and creative solutions helped the Thinking Space Explorer (TSE) design to overcome

this constraint. TSE questions structured thinking in contrary to ToT, which merely strengthens current thought trees, by letting LLMs independently travel new cognitive routes [8].

2. Methodology

2.1. How These Methods (Metrics, Evaluation Strategies) Are Comparable

The effectiveness of many reasoning approaches in Large Language Models is evaluated using logical coherence, factual accuracy, computing efficiency, and generalization across several problem-solving tasks. To provide a complete comparison of several approaches, our evaluation process integrates benchmark datasets, qualitative assessments, and quantitative performance indicators. Examining if generated thinking processes follow a sensible, sequential structure helps one to find logical coherence [4–8]. For retrieval-augmented approaches like RATT and TSE especially, outputs are assessed factual dependability by means of BLEU, ROUGE, and human assessment ratings against ground truth data. Examining computational efficiency based on processing capacity and time needed for every reasoning method ensures a fair compromise between depth and speed.

To ensure consistency, logical systems are evaluated in mathematical, common sense symbolic, and textual coherence generating contexts. Moreover, qualitative human evaluations improve automatic measurements and provide a better knowledge of the benefits and drawbacks of any reasoning approach by assessing coherence, factual accuracy, and usability [4–6,8].

2.2. Technical Aspects of LLM Reasoning Models

From attention-based processing to structured multi-step inference, Large Language Models (LLMs) use distinct architectural and algorithmic approaches to improve reasoning capacities. Evaluating how different reasoning strategies support logical coherence, factual accuracy, and decision-making efficiency depends on an awareness of the technological basis of these models [4–6,8].

- **Transformer-Based Self-Attention and Contextual Reasoning :**
Transformer architecture with multi-head self-attention allows most modern LLMs including GPT-4, PaLM, and LLaMA to encode word associations across long spans. This method helps LLMs to dynamically access relevant information even when it lacks a natural logical foundation [1,2]. Techniques including Chain-of Thought (CoT) using this self-attention capacity help to promote interpretability and coherence by carefully arranging responses into consecutive chains of reasoning.
- **Multi Step Thought Generation and Search Algorithms :**
Search algorithms (BFS, DFS, heuristic optimization) are incorporated beyond CoT, Tree-of-Thought (ToT), and Thought Space Explorer (TSE), so offering organized decision-making frameworks. These models build hierarchical lines of reasoning that let LLM's backtrack and improve answers. Examining several paths increases logical depth and adaptability but also computing complexity [5,8].
- **Integration of External Knowledge & Reinforcement Learning:**
To mitigate hallucinations, Retrieval-Augmented Thought Trees (RATT) incorporate retrieval-based verification, ensuring factual consistency in reasoning outputs. Prototypical incentive Models (Proto-RM) also improve Reinforcement Learning from Human Feedback (RLHF) by optimising incentive structures, hence aligning LLM-generated reasoning with human preferences [6,7].

3. Discussion

3.1. Key Observations Across the 8 Papers

By means of a comprehensive examination of the eight experiments with an emphasis on developments in logical architecture, factual validation, exploration efficiency, and alignment with human purpose, significant patterns in the evolution of LLM reasoning are emphasized. This comparison study produces some quite significant results:

- **Scaling Alone Does Not Guarantee Better Reasoning :** Early studies showed, especially on GPT-1 and GPT-2, increasing model size greatly improves verbal fluency and generalizing [2,3]. However, their zero-shot learning ability, show that more complicated reasoning and logical consistency remain difficulties for bigger models. This underlines the requirement of explicit thinking models to enhance multi-step logical reasoning: with Tree-of- Thought (ToT) and Chain-of- Thought (CoT) [4?] .
- **Structured decision making gets clear when one advances from CoT linear step-wise thinking to ToT multi-path exploration.** While CoT promotes thinking in arithmetic and logical problems, ToT lets more flexible problem solving by means of several search strategies and backtracking. Lack of validation processes in both approaches results in problems when models act aggressively but wrongly. [4? ,5].
- **Fact Verification and Retrieval Enhance Model Reliability :** Retrieval and fact verification help to improve model dependability as models offer plausible but false assertions in LLM theory. By including outside information retrieval into logical processes, the Retrieval-Augmented Thought Tree (RATT) structure immediately fixes this problem and greatly increases factual accuracy. This result implies that to balance knowledge retention with logical coherence, future LLMs should combine dynamic retrieval techniques with ordered reasoning [6,8].
- **Exploration-Based Methods Improve Cognitive Flexibility :** TSE emphasizes exploration's optimization, therefore deviating from received wisdom. While ToT improves present thought trees and enables LLMs to go beyond cognitive blind spots, TSE aggressively looks for new ideas. This suggests that future artificial intelligence systems should investigate fresh ideas in a flexible way in addition to optimizing legal directions. [8].
- **Human input and reinforcement learning define a great deal of artificial intelligence alignment.** Although reaching LLM results to satisfy ethical standards and human preferences still poses a great challenge, intellectual development is quite crucial. Proto-RM (Prototypical Reward Model) optimizes the data efficiency of reward learning, hence improving Learning from Human Feedback (RLHF). This approach guarantees that artificial intelligence produces coherent and consistent with human expectations form of reasoning [7].

3.2. Challenges in LLM Reasoning & AI Alignment

Even although they have advanced greatly, Large Language Models (LLMs) still struggle greatly in ordered reasoning, retrieval augmentation, and reinforcement learning.

Among the problems are logical contradictions and incorrect spreading. One basic flaw of LLM thinking is its lack of logical coherence. Although techniques like the Tree- of- Thought (ToT) and Chain-of Thought (CoT) promote ordered thinking, mistakes made at the beginning of the process could throw off the reasoning and produce erroneous conclusions. CoT also mostly depends on human-designed prompt structures, hence it is sensitive to prompt engineering techniques rather than naturally increasing reasoning capacity [4,5]. This begs questions on the scalability of reasoning systems in several spheres of application.

Two further issues comprise factual errors and hallucinations. Because of their tendency for hallucinations LLMs usually create believable but deceptive claims. Retrieval-Augmented Thought Trees (RATT) include outside knowledge sources to help to solve this issue; retrieval-based methods include added latency and possible bias from other datasets [6]. Moreover, LLMs could still be unable to differentiate trustworthy sources from misleading information even in cases of external retrieval, hence leading to ongoing factual errors.

Third challenges are computational constraints and scalability. Thought Space Explorer (TSE), RATT, and ToT have limited practical usage due to their significant processing overhead advanced reasoning techniques. For general use, methods based on depth-first or breadth-first search algorithms—such as ToT and TSE—may prove computationally challenging. Artificial intelligence cannot find application in the real world without harmony between deep, structured reasoning and processing efficiency [5,6,8].

Absence of self-correction systems marks the fourth problem. Though ordered logic models facilitate LLM decision-making, the self-correction capacity of current designs is inadequate. Once a bad reasoning path is chosen, an LLM hardly ever detects and corrects its own mistakes [4,5]. Future models ought to include self-reflective thinking techniques to evaluate, confirm and dynamically change their own findings.

3.3. Strengths and Weaknesses of Different Models

Every method of reasoning analysis has advantages and disadvantages.

Table 1. Comparison of LLM Reasoning Techniques

Model/Technique	Strengths	Weaknesses
Transformer Self-Attention	<ul style="list-style-type: none">Efficient parallel processing.Captures long-range dependencies.Scalable for large datasets.	<ul style="list-style-type: none">Lacks inherent structured reasoning.Prone to factual inconsistencies.Cannot self-correct reasoning errors.
GPT-1 & GPT-2 (Generative Pre-Training)	<ul style="list-style-type: none">Strong unsupervised language modeling.Effective for zero-shot learning.Adaptable to multiple NLP tasks.	<ul style="list-style-type: none">Lacks structured reasoning capabilities.Generates fluent but illogical responses.Cannot explore alternative reasoning paths.
Chain-of-Thought Prompting (CoT)	<ul style="list-style-type: none">Improves logical step-by-step reasoning.Enhances performance on arithmetic and commonsense tasks.Increases interpretability.	<ul style="list-style-type: none">Linear reasoning structure limits adaptability.Propagates errors if an early step is incorrect.Does not optimize multiple reasoning paths.
Tree-of-Thoughts (ToT)	<ul style="list-style-type: none">Allows multi-path reasoning exploration.Supports structured decision-making with backtracking.Outperforms CoT in complex problem solving.	<ul style="list-style-type: none">Lacks a fact-verification mechanism.Computationally expensive due to multiple search paths.May generate redundant thought branches.
Proto-RM for RLHF	<ul style="list-style-type: none">Efficient reinforcement learning from human feedback.Reduces dependence on large datasets.Optimizes model alignment with human values.	<ul style="list-style-type: none">Does not directly improve reasoning capabilities.Requires human preference data.Limited effectiveness in novel problem-solving scenarios.
Retrieval-Augmented Thought Trees (RATT)	<ul style="list-style-type: none">Improves factual accuracy by integrating external retrieval.Reduces hallucinations in LLM-generated responses.Enhances logical consistency.	<ul style="list-style-type: none">Increases latency due to retrieval step.Depends on external knowledge sources.May retrieve outdated or biased information.
Thought Space Explorer (TSE)	<ul style="list-style-type: none">Expands cognitive space for reasoning.Identifies blind spots in thought structures.Enhances creativity and novel problem-solving.	<ul style="list-style-type: none">Requires higher computational power.Risk of generating unnecessary thought branches.Not optimized for factual verification.

4. Future Work and Open Challenges

4.1. Unanswered Questions in LLM Reasoning

The boundaries and future orientations of Large Language Model (LLM) reasoning remain mostly unknown even with considerable progress in ordered thinking, retrieval augmentation, and reinforcement learning. Although contemporary methods have improved factual reliability and logical coherence, the fundamental framework of LLM-generated reasoning is still mainly correlation-driven rather than completely deductive [4,5].

This raises a crucial concern: Are LLMs constantly limited by probabilistic pattern recognition or can they provide real logical reasoning? Although present approaches such Thought Space Explorer

(TSE), Tree-of- Thought (ToT), and Chain-of- Thought (CoT) help multi-step problem-solving, formal logical reasoning generates no results [4, 5]. Future studies have to look at whether LLMs suit systems of symbolic reasoning or neuro-symbolic architecture.

The evolution of self-correction systems inside LLM reasoning models presents still another difficult problem. Retrieval-Augmented Thought Trees (RATT) provide external fact checking; models cannot independently spot and repair faulty lines of thinking [5,6]. Future research should look at strategies for self-reflective learning so that LLMs could actively examine, change, and enhance their own cognitive processes.

Moreover, scalability is a main constraint. Though they have a larger processing cost, techniques like ToT and TSE significantly boost reasoning depth. Advanced LLMs can be used in real applications by use of effective reasoning systems that strike a compromise between depth, accuracy, and scalability [5,8]. Solving these major issues will help to close the difference between statistical language modeling and actual human-like cognitive capabilities.

4.2. Potential Research Directions

Future research must address scalability, interpretability, and bias mitigating as Large Language Models (LLMs) develop to guarantee their efficient deployment in practical uses.

- **Scalability: Efficient Reasoning Architectures**
Although modern organized thinking approaches such Tree-of- Thought (ToT) and Thought Space Explorer (TSE) considerably improve multi-step problem-solving, their extensive search routes result in major processing costs. Future research could look at hybrid models that balance organized thinking with computer efficiency—that fit memory-efficient architectures, adaptive search algorithms, or pruning procedures. Development of lightweight LLMs with strong reasoning capacity will be crucial for applications needing real-time decision making [5, 6].
- **Interpretability: Transparent and Explainable Reasoning**
Retrieval-Augmented Thought Trees (RATT) and Chain-of- Thought (CoT) have advanced, but LLMs are still essentially black-box systems therefore it is difficult to know how they think. Explainable artificial intelligence (XAI) approaches—such as self-explanatory models, explicit reasoning trails, and visualization tools allowing users to confirm findings generated by AI—should take front stage in future research [6,8].
- **Bias Mitigation: Ethical AI Development**
LLMs trained on large, varied datasets reflect the inherent biases in human communication [7]. Research must produce bias-detection and debiasing algorithms using adversarial training and reinforcement learning from human feedback (RLHF) to ensure fairness and neutrality in AI-generated reasoning.

5. Conclusions

5.1. Final Summary

Structural prompting, retrieval augmentation, and reinforcement learning have helped large language model (LLM) thinking develop in some measure. This work examined eight papers improving the logical coherence, factual accuracy, and human alignment of artificial intelligence-driven thinking. Early developments include Transformers and generative pre-training (GPT-1, GPT-2) lay the stage for scalable language understanding even if they lacked structured thinking ability. Retrieval-Augmented Thought Trees (RATT) and Reinforcement Learning from Human Feedback (RLHF) addressed difficulties with factual consistency and alignment; Chain-of- Thought (CoT) prompting and Tree of Thought (ToT) improved multi-step logical reasoning. More contemporary methods such as Thought Space Explorer (TSE) let LLMs wander and investigate several paths of thought, which enhances the cognitive flexibility.

Even with recent advances, LLMs still suffer with logical coherence, self-correction, scalability, and ethical alignment. Artificial intelligence thinking will be defined going ahead by hybrid models

combining ordered thinking, real-time fact verification, and efficient alignment mechanisms to assure reliable decision-making. By addressing these limitations, further study can advance artificial intelligence beyond statistical pattern identification and therefore approach autonomous reasoning and human-like problem-solving in demanding, real-world environments.

5.2. What's Next for LLM Reasoning?

Large Language Model (LLM) research going forward will concentrate on enhancing interpretability, flexibility, and self-improvement throughout a broad spectrum of topic areas. While multi-path exploration under ordered thinking models such as Tree-of-Thought (ToT) and Thought Space Explorer (TSE) is already feasible, more study is needed to develop efficient self-corrective systems allowing LLMs to identify and fix flawed thinking on their own.

Retrieval augmented thought trees (RATT) and other fact-verification systems must grow to include dynamic, real-time information retrieval if we are to keep artificial intelligence models accurate in fast-changing knowledge contexts. Moreover, scalability is still a big challenge that requires developments in lightweight systems preserving reasoning depth without running through too many computational resources.

By addressing these problems, LLM theory will evolve beyond statistical inference and generate models capable of true deductive reasoning, ethical decision-making, and adaptive learning, hence bringing artificial intelligence near to human-like cognitive capacities.

References

1. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Advances in neural information processing systems* **2017**, *30*.
2. Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I. Improving language understanding by generative pre-training. *OpenAI Blog* **2018**, *1*.
3. Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I.; et al. Language models are unsupervised multitask learners. *OpenAI blog* **2019**, *1*, 9.
4. Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Xia, F.; Chi, E.; Le, Q.V.; Zhou, D.; et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems* **2022**, *35*, 24824–24837.
5. Yao, S.; Yu, D.; Zhao, J.; Shafran, I.; Griffiths, T.; Cao, Y.; Narasimhan, K. Tree of thoughts: Deliberate problem solving with large language models. *Advances in neural information processing systems* **2023**, *36*, 11809–11822.
6. Zhang, J.; Wang, X.; Ren, W.; Jiang, L.; Wang, D.; Liu, K. RATT: A Thought Structure for Coherent and Correct LLM Reasoning, 2024, [[arXiv:cs.CL/2406.02746](https://arxiv.org/abs/2406.02746)].
7. Zhang, J.; Wang, X.; Jin, Y.; Chen, C.; Zhang, X.; Liu, K. Prototypical reward network for data-efficient rlhf. *arXiv preprint arXiv:2406.06606* **2024**.
8. Zhang, J.; Mo, F.; Wang, X.; Liu, K. Thought Space Explorer: Navigating and Expanding Thought Space for Large Language Model Reasoning. *2024 IEEE International Conference on Big Data (BigData)* **2024**, pp. 8259–8251.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.