

# **Unveiling the Reasoning Abilities of Large Language Models: A Comprehensive Review**

Diving deep into “Towards Reasoning in Large Language Models: A Survey”, a survey paper written by Jie Huang and Kevin Chen-Chuan Chang at the University of Illinois at Urbana-Champaign in 2023

# Overview

---

01

Research

02

Introduction

03

Understanding  
Reasoning in LLMs

04

Towards  
Reasoning

05

Evaluation of  
Reasoning

06

Findings,  
Challenges, and  
Future Directions

07

Personal Insights:  
Ethical Considerations  
& Impact

08

Conclusions

## Towards Reasoning in Large Language Models: A Survey

Jie Huang Kevin Chen-Chuan Chang

Department of Computer Science, University of Illinois at Urbana-Champaign  
{jeffhj, kcchang}@illinois.edu

### Abstract

Reasoning is a fundamental aspect of human intelligence that plays a crucial role in activities such as problem solving, decision making, and critical thinking. In recent years, large language models (LLMs) have made significant progress in natural language processing, and there is observation that these models may exhibit reasoning abilities when they are sufficiently large. However, it is not yet clear to what extent LLMs are capable of reasoning. This paper provides a comprehensive overview of the current state of knowledge on reasoning in LLMs, including techniques for improving and eliciting reasoning in these models, methods and benchmarks for evaluating reasoning abilities, findings and implications of previous research in this field, and suggestions on future directions. Our aim is to provide a detailed and up-to-date review of this topic and stimulate meaningful discussion and future work.<sup>1</sup>

### 1 Introduction

Reasoning is a cognitive process that involves using evidence, arguments, and logic to arrive at conclu-

they are large enough (Wei et al., 2022a). For example, by providing the models with “*chain of thoughts*”, i.e., reasoning exemplars, or a simple prompt “*Let’s think step by step*”, these models are able to answer questions with explicit reasoning steps (Wei et al., 2022b; Kojima et al., 2022), e.g., “all whales are mammals, all mammals have kidneys; therefore, all whales have kidneys.” This has sparked considerable interest in the community since reasoning ability is a hallmark of human intelligence that is frequently considered missed in current artificial intelligence systems (Marcus, 2020; Russin et al., 2020; Mitchell, 2021; Bommasani et al., 2021).

However, despite the strong performance of LLMs on certain reasoning tasks, it remains unclear whether LLMs are actually reasoning and to what extent they are capable of reasoning. For example, Kojima et al. (2022) claim that “LLMs are decent zero-shot reasoners (p. 1)”, while Valmeekam et al. (2022) conclude that “LLMs are still far from achieving acceptable performance on common planning/reasoning tasks which pose no issues for humans to do (p. 2).” This limitation is also

# Research

“**Towards Reasoning in Large Language Models: A Survey**”, written by Jie Huang and Kevin Chen-Chuan Chang at the University of Illinois at Urbana-Champaign

Published in July 2023, this paper explores the capabilities of large language models in reasoning, discussing methods for enhancing reasoning skills in these models and suggesting future research directions. Although it primarily focuses on natural language processing, the review is comprehensive and up-to-date, possibly providing insights that could extend to other modalities.

# Introduction

---

**As we stand on the brink of a new era in artificial intelligence (AI), the advancements we've witnessed in recent years are not just accelerating — they're reshaping our understanding of what machines are capable of.**

This comprehensive survey not only explores how these models are programmed and trained to simulate reasoning but also challenges us to consider whether these behaviors signify genuine understanding or simply sophisticated pattern recognition. As we explore the intricate dance between data training and emergent machine intelligence, we uncover insights into how close we truly are to achieving machines that think and reason like humans.

# Understanding Reasoning in LLMs

Reasoning, a fundamental component of human intelligence, involves the ability to solve problems, make decisions, and deduce new information from known facts.

## Prompting and In-Context Learning

These techniques involve guiding the model through structured prompts that mimic a logical reasoning process. For example, Chain of Thought prompting encourages the model to break down a problem into sub-components before reaching a conclusion.

## Fine-Tuning on Specialized Tasks

By adjusting the model parameters specifically for tasks that require reasoning, researchers can enhance the model's ability to handle similar tasks in the future.

# Towards Reasoning in LLMS

---

## Chain of Thought and Variants

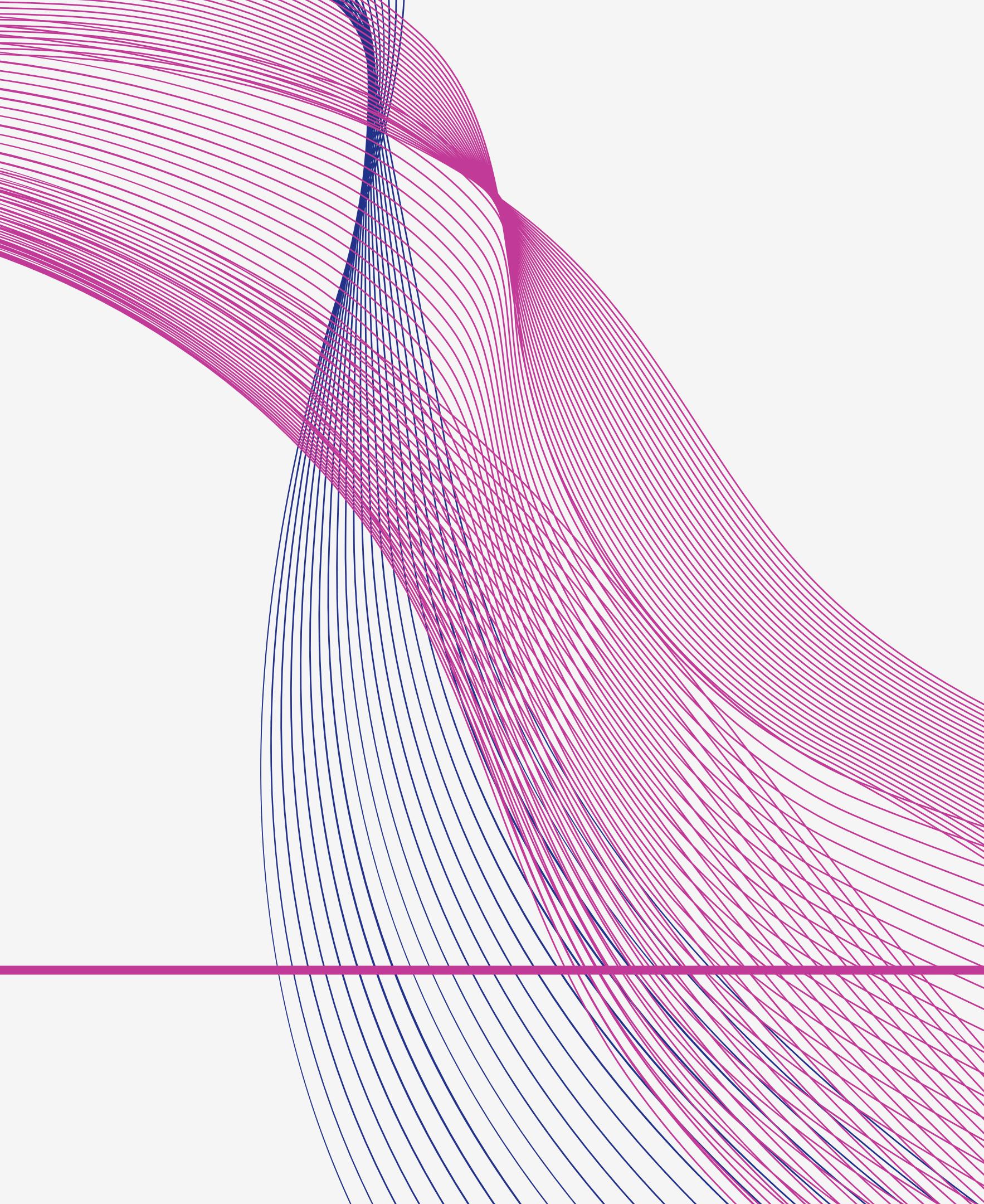
This approach helps in cultivating a deeper, more intuitive reasoning capability in language models, encouraging them to mimic human-like thinking more closely. As a result of using chain-of-thought (CoT) prompting, models are trained to generate comprehensive rationales that logically lead to the final answer when faced with a query.

## Rationale Engineering

Rationale engineering is a strategy designed to enhance how reasoning is elicited or utilized in large language models (LLMs). This process involves refining rationales – creating and using more effective reasoning examples that guide the models.

## Problem Decomposition

This method, known as “decomposition” or “divide and conquer,” involves solving each part of the problem individually before synthesizing these solutions to tackle the overarching issue. This approach not only simplifies the problem-solving process but also aligns each step toward a comprehensive resolution of the whole problem.



# Evaluation of Reasoning

---

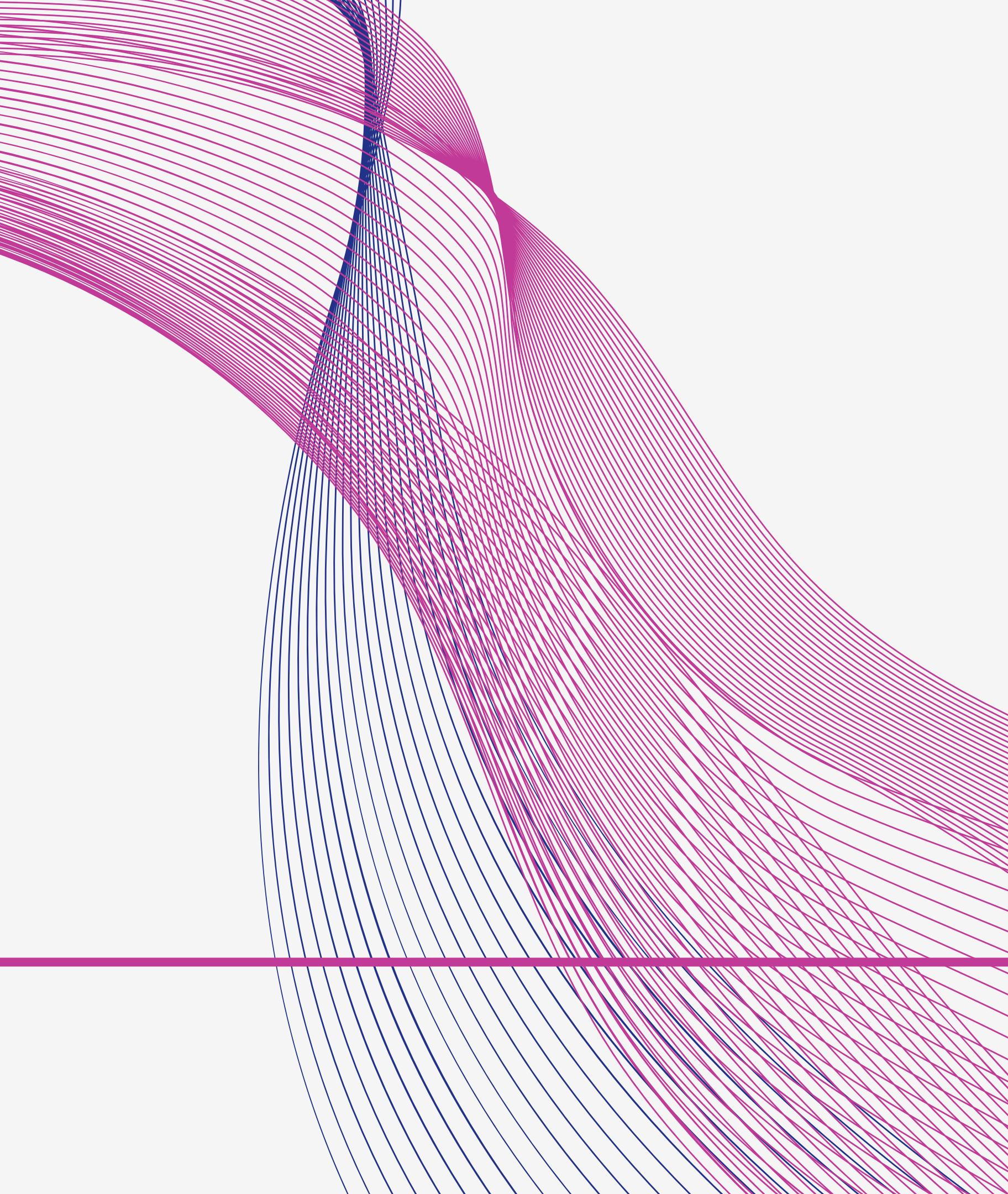
To measure the reasoning ability of LLMs, researchers use benchmarks designed to test various forms of reasoning such as deductive, inductive, and abductive reasoning.

1. Arithmetic Reasoning
  2. Commonsense Reasoning
  3. Symbolic Reasoning
-

---

# Findings

1. Emergence of Reasoning Abilities
2. Chain of Thought Prompts
3. Human Like Reasoning Patterns
4. Challenges with Complex Reasoning
5. Applicability to Real-World Tasks
6. Enhancing Reasoning Abilities



# Challenges and Future Directions

**Despite their capabilities, LLMs face significant challenges:**

- Generalization: LLMs often struggle to apply learned reasoning to new types of problems not covered in their training data.
- Understanding vs. Performance: There is a difference between performing well on a task and truly understanding the underlying processes, which LLMs have yet to bridge.

---

# **Personal Insights: Ethical Considerations & Impact**

As LLMs become more capable, their impact on society and ethical considerations cannot be overstated. Issues such as data privacy, model bias, and the potential for misuse need to be addressed to ensure these technologies are used responsibly.

# Conclusion

---

“Towards Reasoning in Large Language Models: A Survey” provides an insightful look into the state of LLMs in terms of reasoning capabilities. It highlights both the advancements and the roadblocks, offering a detailed roadmap for future research. This survey serves not just as an academic review but also as a critical reflection on the broader implications of deploying these powerful models in everyday applications.

# Thank You

For further questions, feel free to email me at  
[kelly.nguyen01@sjsu.edu](mailto:kelly.nguyen01@sjsu.edu)