

# Discussion of Artificial Intelligence from an Artificial Reasoning Perspective

**Adrienne Raglin**

DEVCOM Army Research Laboratory

adrienne.raglin2.civ@army.mil

## Abstract

The advances in artificial intelligence are significant and rapid. Increased computational power, generative AI and large pretrained models are drivers of these advances. Artificial reasoning explores different reasoning paradigms and how they can be used to enable artificial intelligent systems with capabilities for tackling various tasks. Artificial reasoning also considers the human loop or human AI loop paradigm that highlights the notion that the roles of human and AI shift as AI gains new abilities. This paper discusses some of the ongoing questions that continue to move AI forward, particularly their ability to reason.

## Discussion

Artificial intelligence has long been defined as computer systems that are capable of performing tasks that typically require human intelligence. Human intelligence is the ability to learn from experience, adapt to new situations, understand abstract concepts, and use knowledge. Artificial reasoning can be thought of as the intersection focusing on the different elements of reasoning, such as inferring, problems solving, and identifying patterns. Thus, artificial reasoning is enabling intelligent systems to learn, adapt, solve problems, and infer. When we talk about artificial reasoning, we use the human loop or human AI loop paradigm (Raglin, 2025) to illustrate as AI abilities increase the roles of the AI and human shifts. In the past humans were the loop; they performed all actions and behaviors. With the notion of agentic AI the AI performs these tasks independently or autonomously. However, the human is still over that loop especially as AI takes actions and behaviors that impact humans or as proxy for humans. Given the advances in AI, we ask then what does artificial reasoning look like now?

We start this discussion with (Caballar, 2025) that looks at the question What is reasoning in AI? Here they state,

“Reasoning in artificial intelligence refers to the mechanism of using available information to generate predictions, make inferences and draw conclusions” (Caballar, 2025). This is in line with the definition of artificial reasoning. However, there still exist challenges such as the AI may lack certainty and reliability. Other well-known challenges include bias, computation costs and interpretability. Therefore, AI still needs work if our goal is to have them reason like humans (Caballar, 2025).

Another question is presented in (Thomson, 2025) what does it really mean for AI to reason? To advance AI reasoning different methods are being used. One approach is to design AI reasoning modules with two main components. In this approach, a knowledge base and an inference engine are the two components that are thought to be core for AI reasoning. Where the knowledge base forms the map of real-world entities using structures of graphs, semantic network, and other knowledge representations. Then the inference engine takes that knowledge and implements logic to rate outcomes such as decisions.

Another approach is an extension of large language models. In general, large language models generate responses to specific questions. However, some tasks are not simple direct answers to questions but require multiple logic steps, interpretation of nuances, and structured arguments. These highlight the need for not just a knowledge base where responses are retrieved but connections to concepts and response. Additional features are being added to these models to produce step by step logic with some level of transparency so that the answers can be reviewed and assessed to whether they align with what the person expected or desired. As these newer models come online there may still be issues with incorrect reasoning although the answer is correct. Also, if a multi-step reasoning process is used intermediate hallucinations can be formed and

detected. For complex problems the model may give up or give an answer which only addresses part of the problem.

Other questions centering around large language models and reasoning then becomes Are reasoning models introducing the age of reason for AI? as discussed in (Schilder, 2025). Even though these models can produce “reasonable answers” to certain problems there is the ongoing question of understanding. Because performance may change across different models it is difficult to determine what are the true reasoning capabilities. One approach to large language models that are more recently being called large reasoning models is to include “a structured step by step explanation”. This becomes a part of the training which seeks to generate responses that “mimic” human reasoning. These large reasoning models train with data and reasoning traces.

When adding the topic of deep learning and machine learning the questions might be the one posed in (Marcus, 2022), What aspects of these learning techniques are the most closely related to reasoning? One of the strengths of these techniques is their ability to recognize patterns. An ongoing challenge is their ability to handle when there are outliers beyond what occurred during training. Here, different viewpoints drive how the performance of these techniques are assessed. “AI researchers Emily Bender, Timnit Gebru, and colleagues have put it that deep-learning-powered large language models are like “stochastic parrots,” repeating a lot, understanding little Marcus, 2022). This comment might be linked to the model’s prediction of the next word most likely within their response. So perhaps one might think more data is needed to improve their performance. However, (Marcus, 2022) stated “Scaling the measures Kaplan and his OpenAI colleagues looked at—about predicting words in a sentence—is not tantamount to the kind of deep comprehension true AI would require.” Although not the latest model “research from DeepMind and elsewhere on models even larger than GPT-3 have shown that scaling starts to falter on some measures, such as toxicity, truthfulness, reasoning, and common sense” (Marcus, 2022). AI expert Gary Marcus (Marcus, 2022) stated “What the Apple paper shows, most fundamentally, regardless of how you define AGI, is that LLMs are no substitute for good well-specified conventional algorithms,” when discussing the reasoning limitations of AI models.

Ultimate goals of AI have included the notion of artificial general intelligence. Artificial general intelligence is defined as the ability of machines to understand, learn, and perform any tasks human can. Thus, machines would be able to adapt to new tasks without training and reason. Humans are able to reason and perform new and complex tasks by using prior knowledge and experiences. However,

complexity and adaptability are ongoing challenges for AI. Thus our final question is How does AI ability to handle complex problems show reasoning capabilities? In (Mauran, 2025) one assessment to the challenge of complexity in AI is the use of logic puzzles. The Tower of Hanoi is a logic puzzle that involves moving disk from one tower to another tower following specific rules. Smaller subsets of disks must be moved before larger ones. Claude 3.7 Sonnet and DeepSeek R1 was not able to successfully handle this puzzle when a fifth disc was added. This led to the assessment that “results show that all reasoning models exhibit a similar pattern with respect to complexity: accuracy progressively declines as problem complexity increases until reaching complete collapse beyond a model specific complexity threshold,” (Mauran, 2025).

## Conclusion

In conclusion, having the ability for AI to reason is a critical component for advanced intelligent systems. In this paper we touch on a few questions related to artificial reasoning, the ability of AI to reason. In 2025, based on several reviews (Jindal, 2025 Gary, 2025) there were several models that were considered cutting edge reasoning models. Google DeepMind, Gemini 2.5 Pro does self-fact checking and some level of complex problem solving. OpenAI, OpenAI O3 performs step-by-step problem-solving. Anthropic, Claude 4 Opus can generate “nuanced responses”. DeepSeek AI, DeepSeek-R1 has been identified as performing well on reasoning benchmarks. These advances have been significant and can be considered fast paced. It opens up new opportunities for artificial reasoning research. Also, it opens up the platform for new questions for the direction of artificial reasoning research. As the state-of-the-art advances we can continue to revisit the questions highlighted in this paper to determine what does it mean to have not just artificial intelligence but artificial reasoning.

## References

- Raglin, A. 2025. Artificial Intelligence, Human AI Loop Paradigm, and Artificial Reasoning. In International Conference on Human-Computer Interaction (pp. 428-434). Cham: Springer Nature Switzerland.
- Caballar, R., Stryker, C. 2025. What is reasoning in AI? <https://www.ibm.com/think/topics/ai-reasoning>].
- Thomson Reuters. 2025. What it really means for AI to reason. <https://www.thomsonreuters.com/en-us/posts/innovation/what-it-really-means-for-ai-to-reason/>.
- Schilder, F. 2025. Are reasoning models introducing the age of reasoning for AI? <https://medium.com/tr-labs-ml>

engineering-blog/are-reasoning-models-introducing-the-age-of-reason-for-ai-29a03216d970.

Marcus, G. 2022. RELTR: Deep Learning is Hitting a Wall. <https://nautil.us/deep-learning-is-hitting-a-wall-238440/>.

Mauran, C. 2025. The illusion of thinking': Apple research finds AI models collapse and give up with hard puzzles. <https://mashable.com/article/apple-research-ai-reasoning-models-collapse-logic-puzzles>.

Jindal, P. 2025. 5 Best AI Reasoning Models 2025: Ranked!. <https://www.labellerr.com/blog/compare-reasoning-models/>.

Garg, J. 2025. Top Reasoning Models to Use in 2025: Speed, Specialty & What's Next. <https://www.gocodeo.com/post/top-5-reasoning-models-to-use-in-2025-speed-specialty-whats-next>.