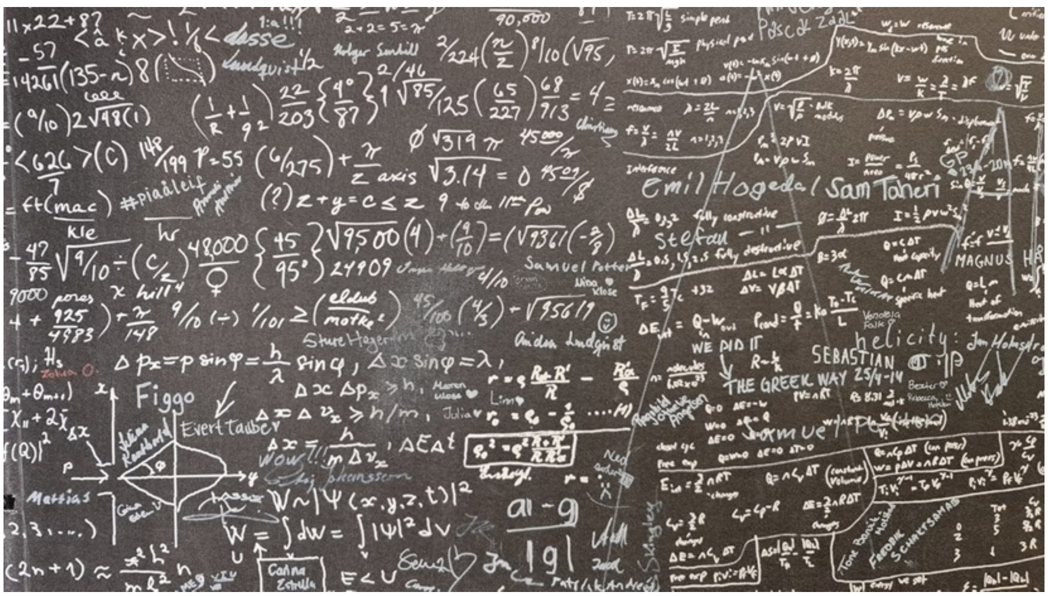
MORGAN CREEK
DIGITAL
ALTERNATIVE THINKING ABOUT INVESTMENTS

Welcome to Morgan Creek Digital's update. It is comprised of a thought piece from our team. We hope you find this content interesting. Please let us know if you have any comments or questions or if you would like to speak to a member of the Morgan Creek Digital team.

Generative Models and Structured, Domain-Aware Reasoning Systems



Preamble: Generative Models and Structured, Domain-Aware Reasoning Systems

Despite the impressive capabilities of large language models (LLMs), they have historically struggled with mathematical problem-solving and structured logical reasoning. These deficiencies stem from several fundamental limitations, including a lack of symbolic representation[1], persistent memory[2], and difficulty with multi-step reasoning[3] given limited token-level processing.

Lack of Symbolic Representation and Persistent Memory: LLMs operate primarily as pattern-matching probabilistic models trained on vast datasets, unlike human mathematicians who can apply symbolic manipulation and formal reasoning. For example, when solving an equation like x2+2x−8=0x^2 + 2x - 8 = 0, the model may not explicitly factorize, complete the square, or systematically apply the quadratic formula. Unlike symbolic solvers, LLMs do not retain persistent memory of intermediate steps during calculations.[4] Many mathematical problems require logic and controls, conceptualized as recursive reasoning, or solving smaller subproblems before arriving at a final result.[5]

Limited Token Level Reasoning: Taking this one step further, earlier transformer-based models, including GPT-4 and Gemini, rely on uniform token processing structures, which process mathematical expressions as sequential token streams.[6] Each new token is generated independently based on statistical patterns observed in training data rather than structured mathematical objects, allowing models to produce impressive generative speeds. However, in observance of the trade-off principle, you must give up something to obtain something desirable. In this case, the opportunity costs of explicit logical structuring, where the relationships between variables, functions, and operations are flattened into a linear format, are errors (like hallucinations) propagating in reasoning-intensive solutions.[7] Further, in the absence of a built-in mechanism for self-verification, an essential capability in mathematics where every step must be logically valid, the errors can compound.[8,9]

In summary, LLM models do not natively "understand" numbers and struggle with maintaining long-range dependencies across multiple steps; instead, the models predict the most likely token based on common numerical patterns. This prediction pattern often leads to logical inconsistencies unsuitable for high-precision arithmetic that cannot afford to misapply a rule or miss a step (calculus, differential equations, and proofs). Even when equipped with external tools (e.g., Python or calculator APIs), LLMs can struggle to determine when to use these tools and how to validate outputs against known mathematical principles.

Enter Anthropic's Claude 3.7 Sonnet

Anthropic's Claude 3.7 Sonnet, introduced in February 2025, may represent an inflection point in AI development, shifting from pure generative models to systems capable of nuanced, structured reasoning (which we will also refer to as "hybrid reasoning" or "task-adaptive computational pathways").[10] It is worth noting that OpenAI's initial o1 series, introduced in December 2024, attempted to solve this gap in intelligence with its high-quality reasoning model. However, the advanced model sacrifices efficiency, costing about $26 per 1 million tokens and lagging in output token speed.[11]

The hybrid reasoning capabilities of Claude 3.7 Sonnet allow it to dynamically adjust the depth of its processing based on the task at hand—ensuring it can handle both rapid, heuristic-driven responses and complex, multi-step logical deductions with remarkable precision.[12] This advancement may not be incremental, as it addresses one of the most pressing challenges in AI today— balancing computational efficiency with the depth of logical reasoning required for high-stakes applications. Unlike prior models that operate within a fixed computational framework, Claude 3.7 Sonnet allows users dynamic control to toggle between rapid response generation and a more deliberate, step-by-step logical process.[13]

The model prioritizes inference efficiency in *Standard (speed) Mode*, utilizing heuristic approximations, learned patterns, and statistical correlations to generate responses with minimal overhead or computational cost.[14] This mode is optimized for conversational AI, real-time decision-making, and tasks where fluency is more valuable than exhaustive reasoning.[15] While its output speed is slower than that of top-shelf alternatives like Google's Gemini 2.0 and Meta's Llama 3.3 series, it has relatively low latency for the first token. The price per 1 million tokens, $6.0 (cheaper), and output tokens per second, 79 (faster), are materially better metrics than OpenAI's o1 series.

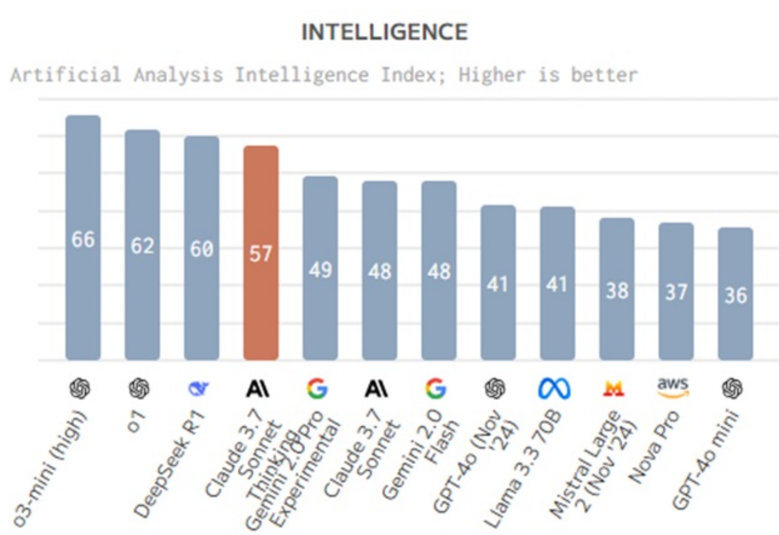
Also, note that OpenAI's o3-mini is reportedly 24% faster than o1 and more affordable but may sacrifice nuanced understanding and accuracy in specialized fields compared to Claude 3.7 Sonnet Thinking and GPT 4.5 (also released in February 2025) (*see Figure 1*).

Figure 1: *Claude 3.7 Standard Speed and Price*

Sourced: https://artificialanalysis.ai/models/claude-3-7-sonnet

The *Extended Thinking Mode* (Claude 3.7 Sonnet Thinking) activates a deeper multi-step reasoning process akin to chain-of-thought (CoT) prompting and self-reflection.[16] This mode forces the model to evaluate intermediate steps before finalizing an answer, improving accuracy in domains requiring strict logical consistency. By leveraging CoT, the model may perform symbolic manipulations, recursive problem-solving, and structured logical deduction—capabilities that traditional transformers struggle to execute reliably, as we previously stated. The model has above-average intelligence with a strong evaluation score. Still, as one may expect, it comes at a higher cost than the average language model and takes more time (hence extended thinking) to reason.[17] (*See Figure 2*)

Figure 2: *Claude 3.7 Thinking Intelligence Score*



Sourced: https://artificialanalysis.ai/models/claude-3-7-sonnet-thinking

Claude 3.7 Sonnet's task awareness is an architectural enhancement that allows it to dynamically optimize its reasoning pathways by emulating key human-like capabilities, such as applying *contextual understanding, selective memory,* and *self-reflection*.[18] *Contextual understanding* evaluates prompts to determine whether they require heuristic inference or explicit reasoning, using entropy calculations, prior knowledge, and meta-learning.[19] *Selective memory* retains high-value computations, reducing redundancy while maintaining coherence.[20] *Gradient-based self-reflection* enables iterative error correction, refining outputs through optimization, particularly in code generation and symbolic reasoning.[21] These characteristics of structured reasoning pathways may lead to AI models adopting a blend of probabilistic and deterministic logic and reducing reliance on brute-force token prediction alone.

Conclusion: Introducing hybrid reasoning may have profound implications for AI scalability, allowing users to control (toggle) the depth of AI reasoning based on task complexity (a unique advantage). The task-adaptive computing will likely allow for more efficient use of hardware, reducing unnecessary computational load and improving efficiency—a critical consideration in AI adoption, particularly enterprise AI adoption. The release came at a crucial time, as DeepSeek's recent advancements have challenged conventional AI cost assumptions. Anthropic's hybrid model may offer businesses a middle path, equivocal to chewing gum and walking simultaneously — maintaining relatively cost-effective pricing while enabling AI to adapt from simple customer service to complex financial modeling tasks.

As competition intensifies, Anthropic's architecture and further development may be particularly relevant for a future in which resource constraints limit on-device AI and edge computing environments. The AI arms race is far from over, and too early to call. Still, perhaps more attention should be given to Sonnet 3.7's achievements— reaching reasoning capabilities comparable to, if not exceeding, OpenAI's top-tier GPT 4.5 at a fraction of the energy usage and cost.

Podcast feed: subscribe to *Digital Currents* in your favorite podcast app, and follow us on Apple Podcasts, or Spotify

[1] https://wiki.pathmind.com/symbolic-reasoning
[2] https://community.openai.com/t/persistent-memory-context-issues-with-chatgpt-4-despite-extensive-prompting/1049995
[3] https://learn.microsoft.com/en-us/training/modules/multistage-reasoning-azure-databricks/
[4] https://ijsrm.net/index.php/ijsrm/article/download/5848/3632/17197
[5] https://www.britannica.com/science/recursive-function
[6] https://blogs.nvidia.com/blog/what-is-a-transformer-model/#:~:text=a Transformer Model?-,A transformer model is a neural network that learns context,is possible%2C%E2%80%9D they wrote.
[7] Ibid.
[8] https://www.ibm.com/think/topics/transformer-model
[9] https://aclanthology.org/2024.findings-acl.397.pdf
[10] https://www.anthropic.com/news/claude-3-7-sonnet
[11] https://artificialanalysis.ai/models/o1
[12] https://www.ibm.com/think/news/claude-sonnet-hybrid-reasoning
[13] https://aws.amazon.com/blogs/aws/anthropics-claude-3-7-sonnet-the-first-hybrid-reasoning-model-is-now-available-in-amazon-bedrock/
[14] https://www.anthropic.com/news/claude-3-7-sonnet
[15] Ibid.
[16] https://www.fastcompany.com/91283751/anthropic-new-claude-3-7-sonnet-ai-chain-of-thought
[17] https://artificialanalysis.ai/models/claude-3-7-sonnet-thinking
[18] https://www.anthropic.com/news/claude-3-7-sonnet?ref=blog.shebash.us
[19] https://blog.dxwand.com/contextual-ai/#:~:text=When we talk about contextual,difference in how businesses operate.
[20] https://pureai.com/Articles/2024/02/14/ChatGPT-Memory.aspx
[21] https://langchain-ai.github.io/langgraph/tutorials/rag/langgraph_self_rag/

Important Disclosures
The above information reflects the opinions of Morgan Creek Digital as of the time this is written and all such opinions are subject to change. No representation or warranty, express or implied, is given by Morgan Creek Digital as to the accuracy of such opinions, and no liability is accepted by such persons for the accuracy or completeness of any such opinions.

No Warranty
Neither Morgan Creek Capital Management, LLC nor Morgan Creek Digital warrants the accuracy, adequacy, completeness, timeliness, or availability of any information provided by non-Morgan Creek sources.

This information is neither an offer to sell nor a solicitation of an offer to buy interests in any investment fund managed by Morgan Creek Capital Management, LLC or its affiliates, nor shall there be any sale of securities in any state or jurisdiction in which such offer or solicitation or sale would be unlawful prior to registration or qualification under the laws of such state or jurisdiction. Alternative investments involve specific risks that may be greater than those associated with traditional investments.

Forward-Looking Statements
This presentation contains certain statements that may include "forward-looking statements" within the meaning of Section 27A of the Securities Act of 1933 and Section 21E of the Securities Exchange Act of 1934. All statements, other than statements of historical fact, included herein are "forward-looking statements." Included among "forward-looking statements" are, among other things, statements about our future outlook on opportunities based upon current market conditions. Although the company believes that the expectations reflected in these forward-looking statements are reasonable, they do reflect all assumptions, risks and uncertainties, and these expectations may prove to be incorrect. Actual results could differ materially from those anticipated in these forward-looking statements as a result of a variety of factors. One should not place undue reliance on these forward-looking statements, which speak only as of the date of this discussion. Other than as required by law, the company does not assume a duty to update these forward-looking statements. Past performance is no guarantee of future results. The illustrations are not intended to predict the performance of any specific investment or security.

General
This is neither an offer to sell nor a solicitation of an offer to buy interests in any investment fund managed by Morgan Creek Capital Management, LLC or its affiliates, nor shall there be any sale of securities in any state or jurisdiction in which such offer or solicitation or sale would be unlawful prior to registration or qualification under the laws of such state or jurisdiction. Any such offering can be made only at the time a qualified offeree receives a Confidential Private Offering Memorandum and other operative documents which contain significant details with respect to risks and should be carefully read. Neither the Securities and Exchange Commission nor any State securities administrator has passed on or endorsed the merits of any such offerings of these securities, nor is it intended that they will. This document is for informational purposes only and should not be distributed.

Risk Summary
Interests in the Morgan Creek Digital Fund IV, LP ("Fund") are speculative and involve a significant degree of risk. Emerging technologies and related businesses have limited performance histories, can be extremely volatile, and are not subject to many of the regulatory oversights over which other investable assets are subject. An investment in the Fund is suitable only for sophisticated investors and requires the financial ability and willingness to accept the high risks and limited liquidity inherent in the Units.

There can be no assurance that the Fund will be successful or that losses will not be incurred by the Fund. Each investor in the Fund must have the ability to bear the risk of loss of their entire investment and must be prepared to bear such risks for an extended period of time. Investors are strongly urged to consult with their professional advisors and to carefully review the risk prior to investing.

Performance Disclosures
There can be no assurance that the investment objectives of any fund managed by Morgan Creek Capital Management, LLC will be achieved. Past performance is not indicative of the performance that any fund managed by Morgan Creek will achieve in the future. Although Morgan Creek Capital Management, LLC has been presented with co-investment opportunities in the past, there can be no assurance that Morgan Creek will be presented with similar opportunities in the future. Further, there can be no assurance that co-investment opportunities will be available in the future.