Google Search AI response to
Does ChatGPT 5.0 resolve the "reasoning" limits of 4.0? on 9/11/25

ChatGPT 5.0, released in August 2025, has demonstrated significant improvements in its reasoning capabilities compared to its predecessors, including 4.0. The enhanced reasoning is achieved through a new router-based architecture that uses different models for different tasks, combined with advancements in training. [1, 2, 3]

Key improvements in GPT-5's reasoning

- **Deeper "System 2" thinking**: While earlier models like GPT-4 relied heavily on pattern-matching from training data ("System 1" thinking), GPT-5 can engage in deeper, multi-step logical thought processes ("System 2"). This is especially evident in versions like **GPT-5 Thinking**, which can tackle complex problems that were previously out of reach.
- **Reduced hallucinations**: By being trained to recognize when a task cannot be completed or when it lacks information, GPT-5 is much more likely to admit its limitations instead of generating a false or speculative answer. This significantly reduces "hallucinations," or confidently incorrect responses, by as much as 80% in "thinking" mode compared to previous models.
- **Superior multimodal reasoning**: The model excels at integrating information across multiple modalities, such as text, images, and video, to form a coherent conclusion. For example, in a medical benchmark test, GPT-5 was able to analyze patient history, lab results, and diagnostic images to outperform human medical experts in making a correct diagnosis.
- **Improved tool use and instruction-following**: GPT-5 is much better at understanding and executing complex, multi-step instructions and utilizing external tools. This enables it to perform complex, multi-stage workflows and handle more sophisticated "agentic" tasks with higher accuracy. [1, 2, 4, 5, 6, 7, 8]

How GPT-5's new architecture improves reasoningOpenAI moved away from a single monolithic model and introduced a dynamic, router-based system for GPT-5.

- **Intelligent routing**: When a user submits a prompt, the system automatically routes it to the most appropriate model based on its complexity.
    - Simple requests are handled by a fast, general-purpose model, similar to GPT-4o.
    - Complex requests automatically activate the **GPT-5 Thinking** model, which is designed for deep, multi-step reasoning.
- **Simplified user experience**: This architecture allows OpenAI to offer a single, powerful "GPT-5" experience that handles diverse tasks seamlessly, removing the need for users to manually select different models for different queries. [1, 3]

Evidence of enhanced reasoningOpenAI and independent researchers have benchmarked GPT-5, revealing impressive gains.

- **Math mastery**: On the difficult AIME math exam, GPT-5 achieved 94.6% without using tools, a new state-of-the-art result.
- **PhD-level science**: The model scored 88.4% on the Graduate Physics Qualification Exam (GPQA), a benchmark designed to be challenging for even top-tier human experts.
- **Software engineering**: On the SWE-bench for real-world software engineering, GPT-5 scored 74.9%.
- **Multimodal understanding**: GPT-5 scored 84.2% on the MMMU benchmark for multimodal problem-solving, compared to 72.2% for GPT-4o. [1]

Google Search AI response to
Does ChatGPT 5.0 resolve the "reasoning" limits of 4.0? on 9/11/25

AI responses may include mistakes.

[1] https://runbear.io/posts/gpt-5-explained
[2] https://www.cnbc.com/2025/08/07/openai-launches-gpt-5-model-for-all-chatgpt-users.html
[3] https://www.techrepublic.com/article/news-openai-gpt-5-release-rumors/
[4] https://www.lesswrong.com/posts/o8eMsxA7uHfybfmhd/the-limitations-of-gpt-4
[5] https://azure.microsoft.com/en-us/blog/gpt-5-in-azure-ai-foundry-the-future-of-ai-apps-and-agents-starts-here/
[6] https://arxiv.org/html/2508.08224v2
[7] https://milvus.io/ai-quick-reference/how-does-gpt5-improve-reasoning-and-problemsolving-capabilities
[8] https://www.youtube.com/watch?v=0Uu_VJeVVfo