SIGMA: Search-Augmented On-Demand Knowledge Integration for Agentic Mathematical Reasoning

Ali Asgarov, Umid Suleymanov, Aadyant Khatri
Department of Computer Science
Virginia Tech
{aliasgarov, umids, aadyant}@vt.edu

Abstract

Solving mathematical reasoning problems requires not only accurate access to relevant knowledge but also careful, multi-step thinking. However, current retrieval-augmented models often rely on a single perspective, follow inflexible search strategies, and struggle to effectively combine information from multiple sources. We introduce **SIGMA** (**S**earch-Augmented On-Demand Knowledge **I**ntegration for A**G**entic **M**athematical re**A**soning), a unified framework that orchestrates specialized agents to independently reason, perform targeted searches, and synthesize findings through a moderator mechanism. Each agent generates hypothetical passages to optimize retrieval for its analytic perspective, ensuring knowledge integration is both context-sensitive and computation-efficient. When evaluated on challenging benchmarks such as MATH500, AIME, and PhD-level science QA GPQA, SIGMA consistently outperforms both open- and closed-source systems, achieving an absolute performance improvement of 7.4%. Our results demonstrate that multi-agent, on-demand knowledge integration significantly enhances both reasoning accuracy and efficiency, offering a scalable approach for complex, knowledge-intensive problem-solving. We will release the code upon publication.

Introduction

Multi-agent systems can help with complex reasoning, but current retrieval methods usually rely on a single perspective and don't combine different approaches effectively. Tool-Integrated Reasoning (TIR) models (Li et al. 2025a; Yao et al. 2023) enhance retrieval but cannot coordinate complementary strategies. *Solving complex mathematical problems requires both accurate knowledge (theorems, numeric facts) and reliable step-by-step reasoning. Large models often fail because missing facts get amplified during reasoning, and single-strategy retrieval cannot consistently provide the necessary information (Li et al. 2025a).*

To address these challenges, we introduce **SIGMA**, a unified framework that decomposes complex reasoning tasks across four complementary specialist agents (FACTUAL, LOGICAL, COMPUTATIONAL, COMPLETENESS). Each agent independently performs targeted reasoning-search cycles, using hypothetical document enhancement to generate perspective-specific retrieval queries and search only when uncertainty arises. A lightweight moderator then synthesizes their outputs into a coherent final solution. This design ensures broad reasoning coverage without the heavy communication overhead of traditional multi-agent systems, while focusing retrieval on precise facts to break long reasoning chains and avoid the cost of monolithic, overlong chains-of-thought.

In summary, our key contributions include:

- We introduce **SIGMA**, a unified framework for multi-perspective reasoning that integrates factual, logical, computational, and completeness perspectives through coordinated agents, enabling richer analysis than single-perspective methods.
- We propose **optimized knowledge integration** via *Hypothetical Document Enhancement* and a moderator-based synthesis mechanism, tailoring retrieval to each analytical perspective and combining outputs into robust, context-sensitive reasoning.
- We provide **empirical validation on challenging benchmarks**, showing that SIGMA achieves consistent gains on MATH-500, AIME, and GPQA, outperforming strong open- and closed-source baselines while maintaining efficiency across model scales.

Related work

Retrieval-augmented and multi-agent reasoning: Large reasoning models (LRMs) like o3 (OpenAI 2025), DeepSeek-R1 (DeepSeek-AI et al. 2025), and Gemini 2.5 (Gemini Team 2025) use chain-of-thought reasoning combined with reinforcement learning and search (Berti, Giorgi, and Kasneci 2025; Wei et al. 2022; Zhang et al. 2025). These models excel at reasoning but rely on static knowledge (Li et al. 2025b,c; Huang et al. 2025), which can lead to errors on knowledge-intensive tasks (Ferrag, Tihanyi, and Debbah 2025) and produce long reasoning traces that increase cost and latency (Liu et al. 2025; Zhang et al. 2025). Retrieval-augmented methods aim to address these limitations, though their effectiveness varies. RAG based methods struggles with multi-hop reasoning, while Self-Ask (Press et al. 2023), ReAct (Yao et al. 2023), and IRCoT (Trivedi et al. 2023) interleave retrieval but often generate weak queries.. More recent approaches such as Search-o1 (Li et al. 2025b), EXSEARCH (Shi et al. 2025), Search-R1 (Jin et al.

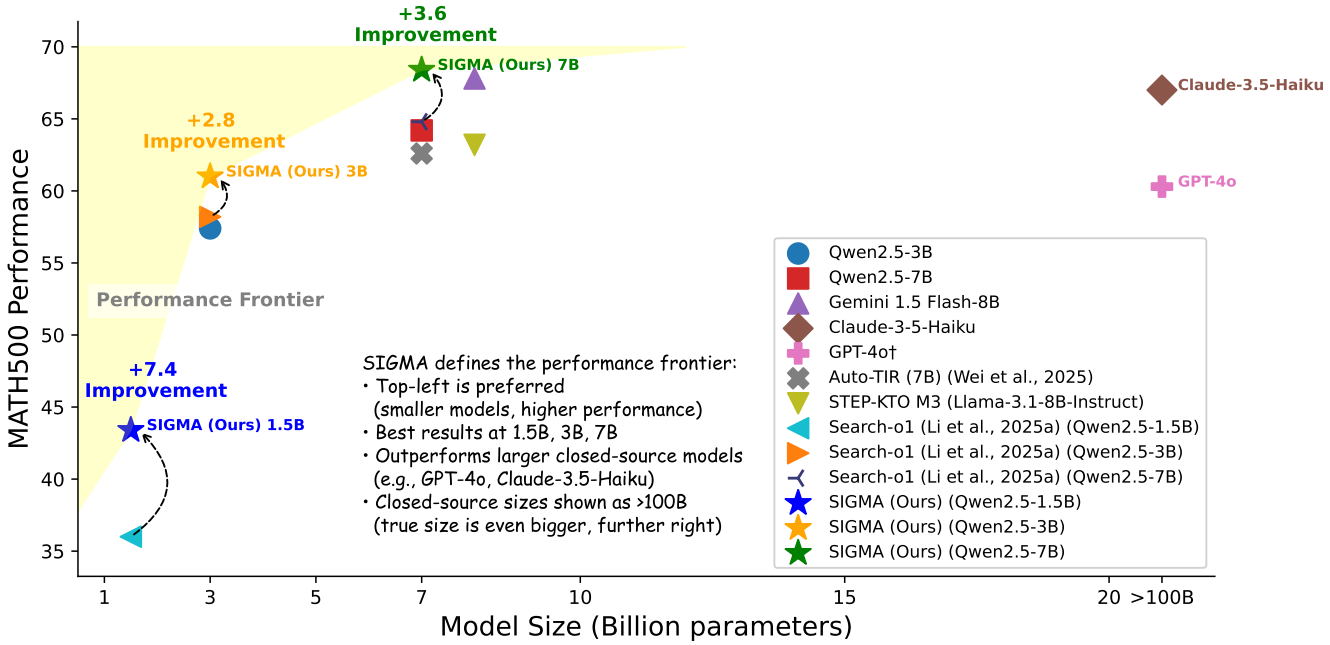

Figure 1: MATH500 score versus model size. SIGMA variants (stars) form a shaded performance frontier at 1.5B, 3B, and 7B parameters, with arrows indicating SIGMA's absolute improvement over the second-best method of the same size. Closed-source models are placed at > 100B on the x-axis. SIGMA outperforms several larger closed-source models, including GPT-4o (Hurst et al. 2024), Gemini 8B (Gemini Team 2025), and Claude-3.5-Haiku (Anthropic 2024).

2025), and WebThinker (Li et al. 2025d) incorporate planning or uncertainty handling, but they still lack true multi-agent collaboration.

Multi-agent methods boost robustness through specialization and interaction. AgentVerse (Chen et al. 2023) coordinates agents, MACNET scales reasoning with DAGs (Qian et al. 2024), and consensus frameworks like Debate (Wang et al. 2025) and ReConcile (Chen, Saha, and Bansal 2023) raise reliability. AutoTIR (Wei et al. 2025) couples agents with reinforcement-based tool use, but most still lack **on-demand search and retrieval**. Instead of dynamically generating targeted queries (e.g., searching only when an agent encounters uncertainty in multi-hop reasoning), or flexibly grounding results in richer persona-specific or hypothetical documents (e.g., HyDE (Gao et al. 2023)), they continue to rely on traditional lexical or embedding-based matches and fixed, upfront toolchains.

Mathematical reasoning in LLMs: Recent work has also improved LLMs' ability to solve mathematical problems without relying on retrieval or multi-agent strategies. Chen et al. (Chen and Qi 2024) train models on the DEMI-MathAnalysis dataset to generate rigorous proofs in limits and series. Zhang et al. (Zhang et al. 2024) introduce LLaMA-Berry, which combines Self-Refine, Monte Carlo Tree Search, and a preference model to explore and rank reasoning paths. These works show that model-based reasoning can improve problem-solving without external knowledge. Lin et al. (Lin et al. 2025) propose Step-kto, employing binary stepwise feedback to enhance reasoning efficiency. However, these approaches do not incorporate **on-demand knowledge integration**, which is essential when equations require targeted retrieval under uncertainty, nor do they utilize **multi-agent strategies**, where specialized solver agents collaborate to improve robustness.

Methodology

We propose the **SIGMA framework** (see Figure 2) for complex reasoning tasks requiring external knowledge through dynamic multi-agent orchestration within a unified model M_θ. Given a query q, for example *find all positive integers n such that 2^n+1 divides 3^n-1*, SIGMA integrates knowledge on demand, retrieving relevant theorems, strategies, computational tools, and edge cases as needed. To facilitate external knowledge access, the model is explicitly instructed to use the special token `<|begin_search_query|>` when a web search is required, enabling targeted retrieval of supporting information (see Appendices **A** and **B**).

Unlike conventional Tool-Integrated Reasoning (Li et al. 2025d; Jin et al. 2025; Shi et al. 2025; Li et al. 2025b) with rigid ***think-action-observation*** workflows, SIGMA enables flexible collaboration among four specialized agents: FACTUAL, LOGICAL, COMPUTATIONAL, and COMPLETENESS. Each agent follows an independent reasoning trajectory and autonomously invokes searches. For example, FACTUAL may retrieve results from number theory, LOGICAL considers proof strategies, COMPUTATIONAL verifies candidates, and COMPLETENESS checks boundary cases. Agents maintain persistent states, which evolve over time as follows:

$$S_k^t = f_\theta(S_k^{t-1}, \phi_k^{t-1}, o_k^{t-1}), \quad \phi_k^t \sim P(\phi \mid S_k^t) \qquad (1)$$

where S_k^t represents the internal state of agent k at reasoning step t. This state is updated by the unified model

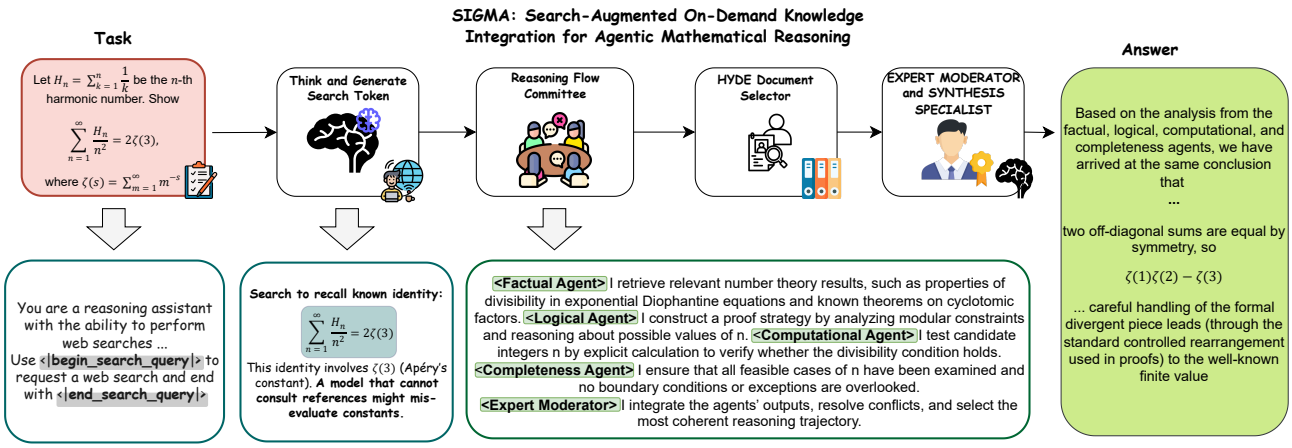

Figure 2: Overview of the SIGMA framework.

Table 1: Main results on challenging reasoning tasks, including Maths and PhD-level science QA. We report Pass@1 metric for all tasks. The best results are in **bold** and the second-best are underlined. Symbol "†" indicates results from their official releases.

Method	Math Benchmarks			GPQA (PhD-Level Science QA)			
	MATH500	AMC23	AIME24	Physics	Chemistry	Biology	Overall
Direct Reasoning (w/o Retrieval)							
Small-Scale Models (< 8B Parameters)							
Qwen2.5-3B	57.4	50.0	3.33	32.56	26.88	68.42	33.33
Qwen2.5-7B	64.2	62.5	16.67	37.20	26.88	42.10	32.83
Gemini 1.5 Flash-8B (Gemini Team 2025)	67.8	-	-	-	-	-	-
Large-Scale Models (> 32B Parameters)							
Claude-3-5-Haiku (Anthropic 2024)	67.0	-	5.3	-	-	-	41.6
Qwen2.5-Coder-32B	71.2	67.5	20.0	37.2	25.8	57.9	33.8
Llama3.3-70B (Dubey et al. 2024)	70.8	47.5	36.7	54.7	31.2	52.6	43.4
GPT-4o† (Hurst et al. 2024)	60.3	-	9.3	59.5	40.2	61.6	50.6
Custom Training Methods							
STEP-KTO M3 (Llama-3.1-8B-Instruct) (Lin et al. 2025)	63.2	47.5	16.7	-	-	-	-
Tool Integrated Methods							
Search-o1 (Qwen2.5-1.5B) (Li et al. 2025a)	36.0	25.0	3.33	27.9	17.2	0.0	20.2
SIGMA (Ours) (Qwen2.5-1.5B)	43.4	20.0	6.66	23.26	16.12	15.79	19.19
Search-o1 (Qwen2.5-3B) (Li et al. 2025a)	58.2	50.0	6.66	18.6	24.73	15.79	21.21
SIGMA (Ours) (Qwen2.5-3B)	61.0	30.0	6.66	34.88	**29.03**	63.16	34.85
Auto-TIR (Qwen2.5-7B) (Wei et al. 2025)	62.6	-	**33.33**	-	-	-	-
Search-o1 (Qwen2.5-7B) (Li et al. 2025a)	64.8	55.0	13.33	27.9	24.73	63.16	29.79
SIGMA (Ours) (Qwen2.5-7B)	**68.4**	**60.0**	16.67	**37.2**	27.96	**68.42**	**35.86**

function f_θ (parameterized by θ) based on its previous state S_k^{t-1}, the action ϕ_k^{t-1} taken at the previous step, and the observation o_k^{t-1} (e.g., a search result or internal thought) resulting from that action. The agent's next action ϕ_k^t is sampled from a policy $P(\phi \mid S_k^t)$, which is conditioned on the current state. Actions are drawn from a discrete set $\phi_k^t \in \{\phi_{\text{search}}, \phi_{\text{reason}}, \phi_{\text{synthesize}}\}$, allowing an agent to autonomously decide whether to search, reason internally, or synthesize its findings. A final moderator then synthesizes outputs across all agents $\mathcal{A}$ to produce the final answer a_{final}:

$$a_{\text{final}} = \text{Moderator}(\{S_k^T, a_k\}_{k \in \mathcal{A}}) \tag{2}$$

Here, $\{S_k^T, a_k\}$ represents the set of final states and intermediate answers generated by each agent k upon reaching its terminal step T.

To clarify the moderator's role, the **moderator** component is implemented as a non-learnable, heuristic synthesis layer rather than a trained module. Its primary function is to integrate the distinct outputs $\{a_k\}$ from each agent. To handle redundant outputs, the moderator first collates and deduplicates all unique propositions. In cases of conflicting reasoning, it employs a pre-defined prioritization scheme; for example, verified results from the COMPUTATIONAL agent are given higher weight than speculative hypotheses from the LOGICAL agent. This deterministic approach ensures consistent synthesis while maintaining focus on the agents'

reasoning capabilities. By modeling collaboration within a unified inference process, SIGMA achieves contextual continuity and computational efficiency.

Hypothetical document enhancement: To support high-quality, targeted retrieval, each agent k generates hypothetical passages using previous work HyDE (Gao et al. 2023) $P_{\text{hyde}}^{(k,j)}$ representing ideal answer content given its current reasoning state.

$$P_{\text{hyde}}^{(k,j)} = \text{HypoGenerator}(q, S_k^j, q_j^{(k)}) \tag{3}$$

This generation is guided by the original query q, the agent's specific state S_k^j at its j-th search step, and any agent-specific sub-query $q_j^{(k)}$ it has formed. Candidate chunks $c_m^{(k,j)}$ retrieved from the external corpus are then ranked by their embedding similarity to this hypothetical document:

$$\text{Sim}_m^{(k,j)} = \cos(\text{Embed}(P_{\text{hyde}}^{(k,j)}), \text{Embed}(c_m^{(k,j)})) \tag{4}$$

where $\text{Sim}_m^{(k,j)}$ is the cosine similarity score for the m-th candidate chunk. This ensures that retrieved information is precisely tailored to the agent's immediate needs while preserving coherence across the unified reasoning process.

A key design feature of SIGMA is the close integration of its specialized agents. Their reasoning processes are interconnected within the shared model M_θ, allowing them to influence each other's behavior (for example, COMPLETENESS questioning an assumption made by LOGICAL, prompting a new search by FACTUAL). Because of this tight coupling, removing a single agent (e.g., "SIGMA w/o FACTUAL") does not simply isolate its effect, it changes how the entire system functions. So, we excluded such ablations, as they would yield misleading results. Instead, our experiments evaluate the full SIGMA system against direct reasoning and tool-integrated baselines. Table 1 includes a qualitative analysis of how the agents interact during reasoning.

Experiments

Experimental Setup and Benchmarks. We evaluate SIGMA's reasoning capabilities on three mathematical benchmarks, MATH500 (Hendrycks et al. 2021), AMC2023, and AIME2024 (Tong et al. 2023). MATH500 comprises 500 problems from the MATH test set (Hendrycks et al. 2021), while AMC2023 and AIME2024 represent middle school–level competitions covering arithmetic, algebra, and geometry. To assess domain-specific scientific reasoning, we additionally employ GPQA (Rein et al. 2023), a PhD-level multiple-choice dataset across physics, chemistry, and biology. Our experiments use the diamond subset to ensure high-quality evaluation. All baseline models were evaluated under identical search budgets, retrieval protocols, and decoding parameters (temperature, seed, pass@k) to ensure fair cross-model comparison.

Baseline and Model Selection. We assess SIGMA as a *general reasoning framework* applied to base foundation models, using `Qwen2.5-1.5B/3B/7B` (Qwen et al. 2024) as backbones. This setup enables direct comparison with general-purpose tool-integrated approaches such as Search-o1 (Li et al. 2025a). While specialized models like `Qwen2.5-Math-1.5B` (Yang et al. 2024) demonstrate strong performance on MATH500 through task-specific fine-tuning, our focus is on evaluating SIGMA's architectural contributions, specifically, its multi-agent orchestration and adaptive knowledge integration under a zero-shot setting. To maintain consistency with comparable baselines, we therefore employ the Qwen2.5 series, with integration of newer Qwen 3 models reserved for future exploration.

Inference Cost and Efficiency. SIGMA introduces a modest computational overhead relative to single-agent approaches such as direct reasoning or ReAct (Yao et al. 2022), as it runs up to four coordinated reasoning trajectories followed by a synthesis step. These trajectories interact within the shared model M_θ, where specialized agents influence and refine each other's reasoning. Despite this coupling, the framework remains efficient because communication occurs implicitly through the shared reasoning space, avoiding the heavy coordination costs found in typical multi-agent systems. Although a detailed analysis of latency and token usage remains future work, SIGMA demonstrates consistent improvements across mathematical and reasoning benchmarks, as described in the Results section, while maintaining parallelizable and robust reasoning performance.

Results

Across all benchmarks, SIGMA consistently improves over both direct reasoning and prior tool-integrated baselines. On mathematical tasks, SIGMA (7B) achieves an absolute gain of **3.6%** over Search- (Li et al. 2025a) at the same scale and **5.8**% over Auto-TIR (Wei et al. 2025) on MATH500. Compared to larger closed-source systems, SIGMA surpasses GPT-4o (Hurst et al. 2024) by **8.1%** on MATH500 and improves upon Claude-3.5-Haiku (Anthropic 2024) by **1.4%**, while coming close to Llama-3.3-70B (Dubey et al. 2024) despite being more than **10x smaller**. On AMC23 and AIME24, SIGMA further improves over retrieval-augmented baselines by **5.0%** and **3.3%** respectively, with particularly strong benefits on multi-step reasoning problems. Beyond mathematics, SIGMA generalizes effectively to GPQA, where it delivers a **6.1%** overall gain relative to retrieval baselines, including a **9.3%** improvement in Physics, **3.2%** in Chemistry, and **5.3%** improvement in Biology. These findings show that multi-perspective reasoning with moderator-based synthesis enables SIGMA to establish a new performance frontier across diverse reasoning domains.

Conclusion

We introduced SIGMA, a retrieval-augmented reasoning framework that coordinates mathematical problem-solving through specialized multi-agent perspectives. This approach enables comprehensive analysis through reasoning-search cycles and moderator-based synthesis. Experiments on MATH500, AIME, and GPQA show consistent improvements over existing baselines, demonstrating that multi-perspective reasoning with on-demand knowledge integration enhances accuracy in complex mathematical domains.

References

Anthropic. 2024. Introducing Claude 3.5 Haiku. https://www.anthropic.com/claude/haiku. Released October 22, 2024.

Berti, L.; Giorgi, F.; and Kasneci, G. 2025. Emergent Abilities in Large Language Models: A Survey. *arXiv.org*.

Chen, J. C.-Y.; Saha, S.; and Bansal, M. 2023. Reconcile: Round-table conference improves reasoning via consensus among diverse llms. *arXiv preprint arXiv:2309.13007*.

Chen, W.; Su, Y.; Zuo, J.; Yang, C.; Yuan, C.; Qian, C.; Chan, C.-M.; Qin, Y.; Lu, Y.; Xie, R.; et al. 2023. Agentverse: Facilitating multi-agent collaboration and exploring emergent behaviors in agents. *arXiv preprint arXiv:2308.10848*, 2(4): 6.

Chen, Z.; and Qi, H. 2024. Large Language Models for Mathematical Analysis. *arXiv preprint arXiv:2501.00059*.

DeepSeek-AI; Guo, D.; Yang, D.; Zhang, H.; Song, J.; Zhang, R.; Xu, R.; Zhu, Q.; Ma, S.; Wang, P.; Bi, X.; Zhang, X.; Yu, X.; Wu, Y.; Wu, Z. F.; Gou, Z.; Shao, Z.; Li, Z.; Gao, Z.; Liu, A.; et al. 2025. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. arXiv:2501.12948.

Dubey, A.; Jauhri, A.; Pandey, A.; Kadian, A.; Al-Dahle, A.; Letman, A.; Mathur, A.; Schelten, A.; Yang, A.; Fan, A.; et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Ferrag, M.; Tihanyi, N.; and Debbah, M. 2025. Reasoning Beyond Limits: Advances and Open Problems for LLMs. *arXiv.org*.

Gao, L.; Ma, X.; Lin, J.; and Callan, J. 2023. Precise Zero-Shot Dense Retrieval without Relevance Labels. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, 1762–1777.

Gemini Team, G. 2025. Gemini 2.5: Pushing the Frontier with Advanced Reasoning, Multimodality, Long Context, and Next Generation Agentic Capabilities. Technical report, Google DeepMind. Technical Report.

Hendrycks, D.; Burns, C.; Kadavath, S.; Arora, A.; Basart, S.; Tang, E.; Song, D.; and Steinhardt, J. 2021. Measuring Mathematical Problem Solving With the MATH Dataset. *NeurIPS (Datasets and Benchmarks Track)*.

Huang, L.; Liu, Y.; Jiang, J.; Zhang, R.; Yan, J.; Li, J.; and Zhao, W. X. 2025. ManuSearch: Democratizing Deep Search in Large Language Models with a Transparent and Open Multi-Agent Framework. *arXiv.org*.

Hurst, A.; Lerer, A.; Goucher, A. P.; Perelman, A.; Ramesh, A.; Clark, A.; Ostrow, A.; Welihinda, A.; Hayes, A.; Radford, A.; et al. 2024. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*.

Jin, B.; et al. 2025. Search-R1: Training LLMs to Reason and Leverage Search Engines with Reinforcement Learning. *arXiv preprint arXiv:2503.09516*.

Li, X.; Dong, G.; Jin, J.; Zhang, Y.; Zhou, Y.; Zhu, Y.; Zhang, P.; and Dou, Z. 2025a. Search-o1: Agentic search-enhanced large reasoning models. *arXiv preprint arXiv:2501.05366*.

Li, X.; Dong, G.; Jin, J.; Zhang, Y.; Zhou, Y.; Zhu, Y.; Zhang, P.; and Dou, Z. 2025b. Search-o1: Agentic Search-Enhanced Large Reasoning Models. *arXiv.org*.

Li, X.; Jin, J.; Dong, G.; Qian, H.; Zhu, Y.; Wu, Y.; Wen, J.-R.; and Dou, Z. 2025c. WebThinker: Empowering Large Reasoning Models with Deep Research Capability. *arXiv.org*.

Li, X.; et al. 2025d. WebThinker: Empowering Large Reasoning Models with Deep Research Capability. *arXiv preprint arXiv:2504.21776*.

Lin, Y.-T.; Jin, D.; Xu, T.; Wu, T.; Sukhbaatar, S.; Zhu, C.; He, Y.; Chen, Y.-N.; Weston, J.; Tian, Y.; et al. 2025. Step-kto: Optimizing mathematical reasoning through stepwise binary feedback. *arXiv preprint arXiv:2501.10799*.

Liu, Y.; Wu, J.; He, Y.; Gao, H.; Chen, H.; Bi, B.; Zhang, J.; Huang, Z.; and Hooi, B. 2025. Efficient Inference for Large Reasoning Models: A Survey. *arXiv.org*.

OpenAI. 2025. Introducing OpenAI o3 and o4-mini. https://openai.com/index/introducing-o3-and-o4-mini/. Accessed: 2025-09-09.

Press, O.; Zhang, M.; Min, S.; Schmidt, L.; Smith, N. A.; and Lewis, M. 2023. Measuring and Narrowing the Compositionality Gap in Language Models. In *Findings of the Association for Computational Linguistics: EMNLP*.

Qian, C.; Xie, Z.; Wang, Y.; Liu, W.; Zhu, K.; Xia, H.; Dang, Y.; Du, Z.; Chen, W.; Yang, C.; et al. 2024. Scaling large language model-based multi-agent collaboration. *arXiv preprint arXiv:2406.07155*.

Qwen; :; Yang, A.; Yang, B.; Zhang, B.; Hui, B.; Zheng, B.; Yu, B.; Li, C.; Liu, D.; Huang, F.; Wei, H.; Lin, H.; Yang, J.; Tu, J.; Zhang, J.; Yang, J.; Yang, J.; Zhou, J.; Lin, J.; Dang, K.; Lu, K.; Bao, K.; Yang, K.; Yu, L.; Li, M.; Xue, M.; Zhang, P.; Zhu, Q.; Men, R.; Lin, R.; Li, T.; Xia, T.; Ren, X.; Ren, X.; Fan, Y.; Su, Y.; Zhang, Y.; Wan, Y.; Liu, Y.; Cui, Z.; Zhang, Z.; and Qiu, Z. 2024. Qwen2.5 Technical Report. arXiv:2412.15115.

Rein, D.; Hou, B. L.; Stickland, A. C.; Petty, J.; Pang, R. Y.; Dirani, J.; Michael, J.; and Bowman, S. R. 2023. GPQA: A Graduate-Level Google-Proof Q&A Benchmark. *CoRR*, abs/2311.12022.

Shi, Z.; et al. 2025. Iterative Self-Incentivization Empowers Large Language Models as Agentic Searchers. *arXiv preprint arXiv:2505.20128*.

Tong, Y.; et al. 2023. AwesomeLLM4Math: A Curated List of Resources and Benchmarks for Math in Large Language Models. https://github.com/tongyx361/AwesomeLLM4Math.

Trivedi, H.; Balasubramanian, N.; Khot, T.; and Sabharwal, A. 2023. Interleaving Retrieval with Chain-of-Thought Reasoning for Knowledge-Intensive Multi-Step Questions. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*.

Wang, H.; Du, X.; Yu, W.; Chen, Q.; Zhu, K.; Chu, Z.; Yan, L.; and Guan, Y. 2025. Learning to break: Knowledge-enhanced reasoning in multi-agent debate system. *Neurocomputing*, 618: 129063.

Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Chi, E. H.; Xia, F.; Le, Q.; and Zhou, D. 2022. Chain of Thought Prompting Elicits Reasoning in Large Language Models. *Neural Information Processing Systems*.

Wei, Y.; Yu, X.; Weng, Y.; Pan, T.; Li, A.; and Du, L. 2025. Autotir: Autonomous tools integrated reasoning via reinforcement learning. *arXiv preprint arXiv:2507.21836*.

Yang, A.; Zhang, B.; Hui, B.; Gao, B.; Yu, B.; Li, C.; Liu, D.; Tu, J.; Zhou, J.; Lin, J.; Lu, K.; Xue, M.; Lin, R.; Liu, T.; Ren, X.; and Zhang, Z. 2024. Qwen2.5-Math Technical Report: Toward Mathematical Expert Model via Self-Improvement. *CoRR*, abs/2409.12122.

Yao, S.; Zhao, J.; Yu, D.; Du, N.; Shafran, I.; Narasimhan, K.; and Cao, Y. 2022. React: Synergizing reasoning and acting in language models. *arXiv preprint arXiv:2210.03629*.

Yao, S.; Zhao, J.; Yu, D.; Du, N.; Shafran, I.; Narasimhan, K.; and Cao, Y. 2023. ReAct: Synergizing Reasoning and Acting in Language Models. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

Zhang, D.; Wu, J.; Lei, J.; Che, T.; Li, J.; Xie, T.; Huang, X.; Zhang, S.; Pavone, M.; Li, Y.; Ouyang, W.; and Zhou, D. 2024. LLaMA-Berry: Pairwise Optimization for O1-like Olympiad-Level Mathematical Reasoning. *CoRR*, abs/2410.02884.

Zhang, S.; Wu, J.; Chen, J.; Zhang, C.; Lou, X.; Zhou, W.; Zhou, S.; Wang, C.; and Wang, J. 2025. OThink-R1: Intrinsic Fast/Slow Thinking Mode Switching for Over-Reasoning Mitigation. *arXiv.org*.

SIGMA Multi-Agent Inference

Problem and Multi-Agent Execution

Given the query "Find positive integers $n \leq 2024$ such that $\gcd(n, 2024) = 1$", SIGMA processes this through four specialized agents with search budgets of 2 each. The FACTUAL agent begins by recognizing this as a number theory problem requiring verification of mathematical definitions, so it searches `<|begin_search_query|>` Euler totient function definition `<|end_search_query|>` and discovers that the problem asks for $\phi(2024)$ where $\phi(n)$ counts integers coprime to n. The agent then searches for prime factorization methods and concludes that computing $\phi(2024)$ requires factoring 2024 first. Meanwhile, the COMPUTATIONAL agent focuses on the mathematical calculations needed, searching `<|begin_search_query|>` factor 2024 prime decomposition `<|end_search_query|>` to find $2024 = 2^3 \times 11 \times 23$, then searching for totient calculation methods to compute $\phi(2024) = 2024 \times (1 - \frac{1}{2}) \times (1 - \frac{1}{11}) \times (1 - \frac{1}{23}) = 880$. The LOGICAL agent establishes the conceptual framework by searching for the relationship between coprimality and Euler's function, confirming that $\gcd(n, 2024) = 1$ means n shares no prime factors with 2024, which is exactly what the totient function counts. The COMPLETENESS agent verifies the solution by searching for alternative calculation methods and cross-validation approaches to ensure no steps were missed.

Algorithm 1: SIGMA Multi-Agent Inference

Require: Model M_θ, search function `Search`, HDS selector
Input: Query q, agent instructions $\{I_k\}_{k \in \mathcal{A}}$
for each agent $k \in \{\text{FACTUAL}, \text{LOGICAL}, \text{COMPUTATIONAL}, \text{COMPLETENESS}\}$ **do**
  Initialize state: $S_k^0 \leftarrow I_k \oplus q$
  Initialize search budget: $B_k \leftarrow \text{MaxSearches}$
  while $B_k > 0$ and reasoning incomplete **do**
    Generate step: $r_k \leftarrow M_\theta(S_k^t)$
    Update: $S_k^{t+1} \leftarrow S_k^t \oplus r_k$
    if r_k contains query $q_j^{(k)}$ **then**
      Emit: `<|begin_search_query|>` $q_j^{(k)}$ `<|end_search_query|>`
      $\mathcal{D}_j^{(k)} \leftarrow \texttt{Search}(q_j^{(k)})$
      $P_{\text{hyde}} \leftarrow \text{HypoGenerator}(q, S_k^t, q_j^{(k)})$
      Apply HDS and integrate results
      $B_k \leftarrow B_k - 1$
    end if
  end while
  Extract conclusion: $a_k \leftarrow \text{ExtractAnswer}(S_k^T)$
end for
Moderator Synthesis: $a_{\text{final}} \leftarrow \text{Moderator}(q, \{(S_k^T, a_k)\}_{k \in \mathcal{A}})$
Output: Final answer a_{final}

After each agent completes its reasoning and search process, the moderator synthesizes their contributions: the FACTUAL agent provided the correct mathematical framework, the COMPUTATIONAL agent performed accurate calculations, the LOGICAL agent confirmed the conceptual connection, and the COMPLETENESS agent validated the approach. The final synthesis integrates these insights to produce: "The question asks for positive integers coprime to 2024, which equals Euler's totient function $\phi(2024)$. Using prime factorization $2024 = 2^3 \times 11 \times 23$ and the totient formula, we get $\phi(2024) = 880$."

Framework Advantages

The multi-agent approach succeeds where single-agent systems often struggle because it distributes different types of mathematical reasoning across specialized agents. The FACTUAL agent ensures definitional accuracy, preventing conceptual errors. The COMPUTATIONAL agent handles numerical calculations systematically. The LOGICAL agent maintains structural coherence. The COMPLETENESS agent catches potential oversights. This distributed expertise, combined with targeted search queries from each agent's perspective, creates more robust solutions for complex mathematical problems requiring both theoretical understanding and computational precision.