TRINITY: An Evolved LLM Coordinator

A Lightweight Coordinator Orchestrating Multiple LLMs via Evolutionary Strategy

Jinglue Xu[1*] Qi Sun[1,3*] Peter Schwendeman[2] Stefan Nielsen[1] Edoardo Cetin[1] Yujin Tang[1]

[1]Sakana AI, Japan [2]University of Michigan, USA [3]Institute of Science Tokyo, Japan

Published as a conference paper at ICLR 2026 | arXiv: 2512.04695

Motivation & Key Contributions

Scaling monolithic LLMs yields diminishing returns, and model merging is impractical across incompatible architectures and closed-source APIs. **TRINITY** adopts a **macro-level approach**: test-time model composition via coordination, fusing complementary strengths of multiple models without modifying their weights.

- **Lightweight coordination:** Hidden states from an SLM (0.6B) + a tiny head (∼10K params) coordinate diverse LLMs — total learnable parameters under 20K.
- **Efficient training:** sep-CMA-ES outperforms RL, SFT, and random search under tight budget constraints.
- **State-of-the-art:** 86.2% pass@1 on LiveCodeBench. Outperforms all baselines on 8 benchmarks (4 in-distribution, 4 held-out).

Why Evolutionary Strategy?

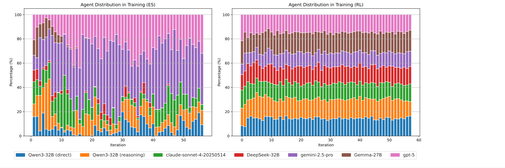
Training the coordinator is challenging: each evaluation requires a full multi-turn LLM rollout (expensive), the reward is binary (correct/incorrect), and parameters exhibit **weak mutual coupling** — each has only a tiny influence on the scalar reward.

- **RL fails:** REINFORCE gradients are low-SNR under binary rewards and weak parameter correlations, yielding unstable learning.
- **SFT is intractable:** Multi-turn label generation scales as $O(7^4 \cdot 3^5)$ per question — prohibitively expensive.
- **sep-CMA-ES works:** Derivative-free, diagonal covariance matches the block-separable structure. Improvement grows **linearly** with iterations (vs. logarithmically for random search). Only ∼1.5K–40K evaluations needed for ∼10K parameters.

sep-CMA-ES Results

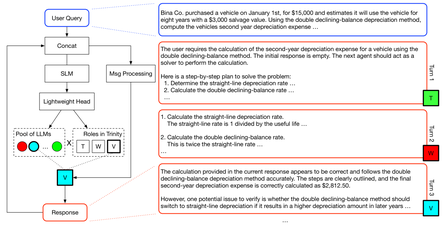
sep-CMA-ES consistently outperforms all alternative training methods:

Method	LCB	MATH	MMLU	RLPR
REINFORCE	0.253	0.459	0.500	0.266
Random Search	0.374	0.794	0.897	0.345
SFT	0.592	0.786	0.906	0.360
sep-CMA-ES	**0.615**	**0.880**	**0.916**	**0.401**



Coordination Pipeline

TRINITY processes queries over multiple turns. At each turn, a compact SLM reads the full transcript, and a lightweight head selects an LLM and assigns it one of three roles: **Thinker** (strategize), **Worker** (execute), or **Verifier** (evaluate). The process halts when the Verifier accepts the solution.



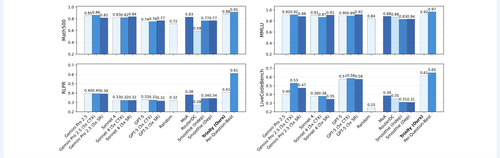
NEW STATE-OF-THE-ART

86.2% pass@1 on LiveCodeBench

Surpassing GPT-5 (83.8%) and Gemini 2.5-Pro (67.2%)

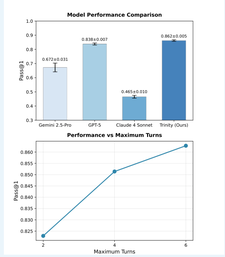
In-Distribution Results

TRINITY outperforms all single-model and multi-agent baselines across four benchmarks (21.9% mean relative error reduction).



LiveCodeBench SOTA

TRINITY achieves 86.2% pass@1 on LiveCodeBench V6. Performance improves with more turns.



Zero-Shot Transfer to Unseen Tasks

TRINITY generalizes to four held-out benchmarks without any retraining, achieving the highest average score.

Model	AIME	BCB	MT-B	GPQA	Avg
Gemini 2.5-Pro	46.7	35.1	9.37	75.3	52.3
GPT-5	46.7	33.8	9.35	72.7	51.1
Claude-4-Sonnet	35.3	35.8	9.28	67.3	46.1
TRINITY	**50.0**	**35.8**	**9.60**	**76.8**	**54.2**

Ablation Studies

Every design choice contributes: removing singular value fine-tuning, role selection, or switching to the last token all degrade performance.

Method	LCB	MATH	MMLU	RLPR	Avg
TRINITY (full)	**61.5**	**88.0**	**91.6**	**40.7**	**70.4**
w/o SV fine-tune	55.7	85.9	90.1	39.8	67.9
w/o Tri-role	58.3	82.0	91.6	36.2	67.0
w/ Last token	50.9	87.0	82.2	38.6	64.7

Conclusion

TRINITY demonstrates that a lightweight coordinator (<20K learnable parameters) can orchestrate diverse LLMs to achieve state-of-the-art performance. The results suggest a promising path forward: engineering **collaborative AI ecosystems** rather than scaling monolithic models.

Contact: jingluexu@sakana.ai, qisun@sakana.ai, yujintang@sakana.ai

arXiv: 2512.04695