

Sakana AI and Berkeley researchers show a model improving itself over two rounds

Source copy kept by wheresthe.ai for the story "Sakana AI and Berkeley researchers show a model improving itself over two rounds" (<https://wheresthe.ai/d/sakana-mass-recursive-self-improvement-2026-10/>).

Source: <https://arxiv.org/abs/2610.12176>

Text saved 18 Oct 2026 (the page as published).

SHA-256 of the saved text: c4bf4fc061f375cf9b440c663122625970d828a0854a4d6430f63709e8e43423

-->

Computer Science > Artificial Intelligence

arXiv:2610.12176 (cs)

[Submitted on 8 Oct 2026]

Title: Recursive Self-Improvement through Multi-Agent Self-Supervision

Authors: Hyunin Lee , Jinglue Xu , Jeffrey Seely , Donghyun Lee , Somayeh Sojoudi , Matei Zaharia , Yujin Tang
View a PDF of the paper titled Recursive Self-Improvement through Multi-Agent Self-Supervision, by Hyunin Lee and 6 other authors

View PDF HTML (experimental)

Abstract: Recursive self-improvement (RSI) of a model on non-verifiable tasks, such as open-ended research, faces a supervision bottleneck when its outputs exceed what even human experts can reliably assess, leaving the model itself (optimizee) as the best available optimizer and evaluator. However, a single model instance struggles to critique and improve its own complex reasoning under this homogeneous loop. To address this, we propose Multi-Agent Self-Supervision (MASS), an RSI method that alternates between evolutionary workflow optimization and supervised fine-tuning on self-generated trajectories. Guided by early findings that multi-agent topologies excel at complex reasoning, MASS prompts a single base model to iteratively propose, execute, and self-evaluate multi-agent workflows. Through an evolutionary search constrained by structural guardrails, the model optimizes these computational-graph-like orchestrations, discovering the most effective distinct roles and information routing for a given task. Over two MASS cycles with Qwen3.6-27B, the model achieves 1.2-1.6x higher performance per output tokens on four open-ended public benchmarks. Because the improved model subsequently acts as a better optimizer and evaluator, this alternating framework enables a continuous, recursive bootstrapping of the model's capabilities. Moreover, multi-agent traces are also more training-efficient: a student trained on them outperforms a single-agent student trained on 1.4x more training tokens. These findings suggest that jointly learning orchestration and bounded subagent execution from multi-agent trajectories can provide an effective signal for RSI.

Comments: 39 pages

Subjects:

Artificial Intelligence (cs.AI)

Cite as: arXiv:2610.12176 [cs.AI]

(or arXiv:2610.12176v1 [cs.AI] for this version)

<https://doi.org/10.48550/arXiv.2610.12176>

Focus to learn more

arXiv-issued DOI via DataCite (pending registration)

Submission history From: Hyunin Lee [view email] [v1] Thu, 8 Oct 2026 15:44:02 UTC (410 KB)

Full-text links: Access Paper:

View a PDF of the paper titled Recursive Self-Improvement through Multi-Agent Self-Supervision, by Hyunin Lee and 6 other authors View PDF HTML (experimental) TeX Source

view license

Additional Features

Audio Summary

Current browse context:

cs.AI

< prev

| next >

new | recent | 2026-10

Change to browse by:

cs

References & Citations

NASA ADS Google Scholar

Semantic Scholar

export BibTeX citation Loading...

BibTeX formatted citation

×

loading...

Data provided by:

Bookmark

Bibliographic Tools

Bibliographic and Citation Tools

Bibliographic Explorer Toggle

Bibliographic Explorer (What is the Explorer?)

Connected Papers Toggle

Connected Papers (What is Connected Papers?)

Litmaps Toggle

Litmaps (What is Litmaps?)

scite.ai Toggle

scite Smart Citations (What are Smart Citations?)

Code, Data, Media

Code, Data and Media Associated with this Article

alphaXiv Toggle

alphaXiv (What is alphaXiv?)

Links to Code Toggle

CatalyzeX Code Finder for Papers (What is CatalyzeX?)

DagsHub Toggle

DagsHub (What is DagsHub?)

GotitPub Toggle

Gotit.pub (What is GotitPub?)

Huggingface Toggle

Hugging Face (What is Huggingface?)

ScienceCast Toggle

ScienceCast (What is ScienceCast?)

Demos

Demos

Replicate Toggle

Replicate (What is Replicate?)

Spaces Toggle

Hugging Face Spaces (What is Spaces?)

Spaces Toggle

XYZ.AI (What is XYZ.AI?)

Related Papers

Recommenders and Search Tools

Link to Influence Flower

Influence Flower (What are Influence Flowers?)

Core recommender toggle

CORE Recommender (What is CORE?)

Author

Venue

Institution

Topic

About arXivLabs

arXivLabs: experimental projects with community collaborators

arXivLabs is a framework that allows collaborators to develop and share new arXiv features directly on our website.

Both individuals and organizations that work with arXivLabs have embraced and accepted our values of openness, community, excellence, and user data privacy. arXiv is committed to these values and only works with partners that adhere to them.

Have an idea for a project that will add value for arXiv's community? Learn more about arXivLabs .

Which authors of this paper are endorsers? | Disable MathJax (What is MathJax?)