

Adapting While Learning: Grounding LLMs for Scientific Problems with Intelligent Tool Usage Adaptation



Bohan Lyu^{1*}, Yadi Cao^{2*},
Duncan Watson-Parris², Leon Bergen², Taylor Berg-Kirkpatrick², Rose Yu²

¹Tsinghua University ²University of California, San Diego

*Equal Contribution

Background

Large Language Models (LLMs) demonstrate promising capabilities in solving **simple** scientific problem like [7]:

When you drop a ball from rest it accelerates downward at 9.8 m/s^2 . If you instead throw it downward assuming no air resistance its acceleration immediately after leaving your hand is

- (A) 9.8 m/s^2
- (B) more than 9.8 m/s^2
- (C) less than 9.8 m/s^2
- (D) Cannot say unless the speed of throw is given.



However, they often produce hallucinations for **complex** one like:

Question: How much will the temperature of Ewo in 2068 under ssp126 change if the emission of CO2 is decreased by -25%?

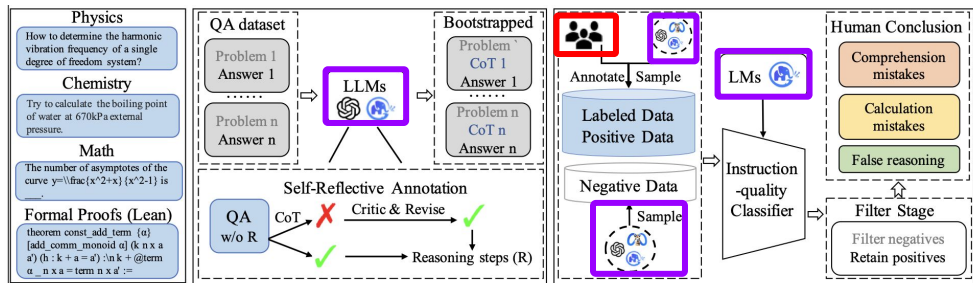
- (A) -0.13081711
- (B) -0.16986465
- (C) -0.09639389
- (D) -0.05745936

Correct Answer: B

Research Question: How to Align LLMs with the physical world for scientific problems?

Existing Approaches:

1. Fine-tune the LLM with data from **human annotation** and **stronger models** [1, 2]



2. Integrate LLMs with external tools [3]

Mind's Eye
Simulator Augmented Zero/Few-shot Reasoning

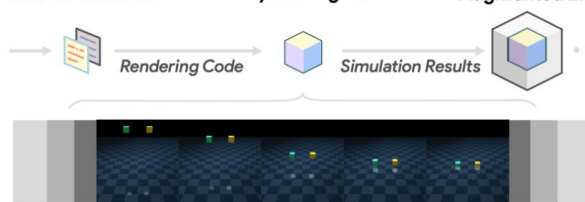
Question:

Two baseballs X and Y are released from rest at the same height.
X is heavier than Y.
Which baseball will fall to the ground faster?

Text-to-Code LM

Physics Engine

Mind's Eye Augmented LM



MuJoCo Simulation

Answer from Mind's Eye + LM:

Answer:

Hints:

X and Y have the same acceleration.
So the answer is: they will fall at the same rate. Both baseballs will fall to the ground at the same time.

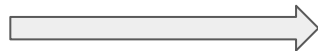
■ Simulation based Prompts Injection

Motivation

Fine-tune an LLM:

- ✗ Requires massive data
- ✗ Costly in scientific settings
- ✗ Prone to hallucination

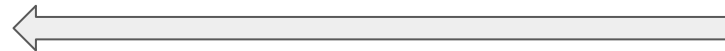
- ✓ Efficient inference
- ✓ Ideal for simpler problems



Employ LLM as Agent:

- ✗ Costly emulation
- ✗ Over-reliance on provided tools
- ✗ Fail to internalize knowledge

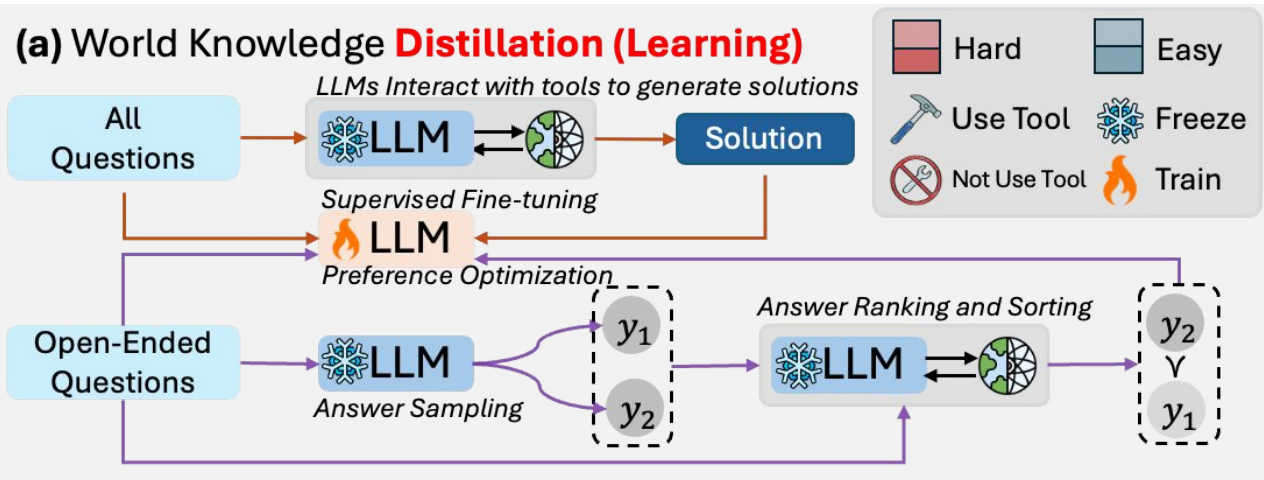
- ✓ Reliable solutions by tools
- ✓ Suitable for complex problems



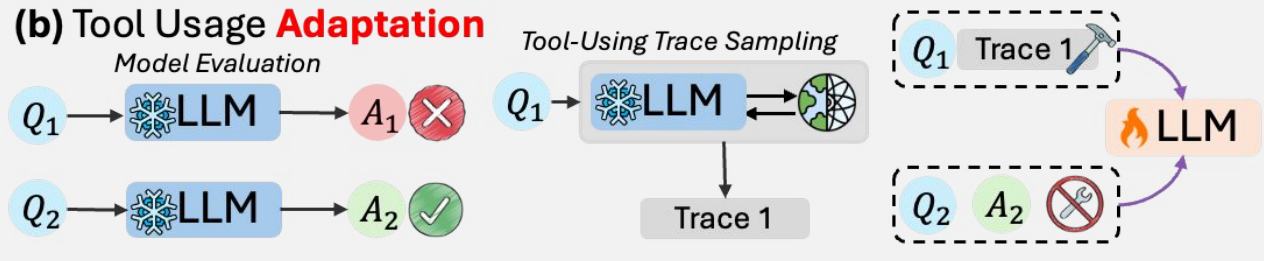
? Can we combine advantages from both approaches?

Our method: **Adapting** while **Learning**

(a) World Knowledge **Distillation (Learning)**

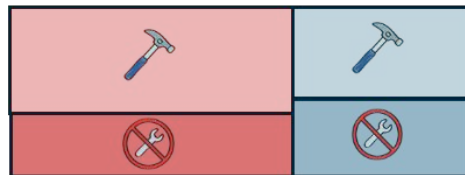


(b) Tool Usage **Adaptation**

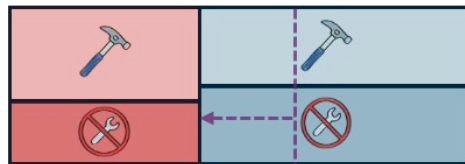


(c) Model Performance

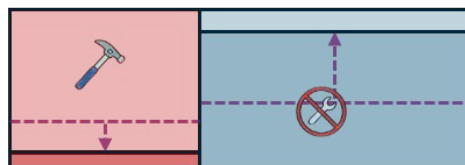
Original



World Knowledge Distillation



Tool Usage Adaptation



Pipeline of Our Method: (a) World Knowledge Distillation, (b) Tool Usage Adaptation, (c) Model improvement visualization.

Our method: World Knowledge Distillation (**Learning**)

Question: How much will the temperature of Ewo in 2068 under ssp126 change if the emission of CO2 is decreased by -25%?

- (A) -0.13081711
- (B) -0.16986465
- (C) -0.09639389
- (D) -0.05745936

Correct Answer: B

Assistant: Call query_lat_and_lon with parameters {'city name': 'Ewo'}

Feedback from query_lat_and_lon: latitude: -0.8742, longitude: 14.8167.

Assistant: Call diy_greenhouse with parameters {'longitude': 14.8167, 'latitude': -0.8742, 'setting': 'ssp126', 'year': 2068, 'delta CO2': -0.25, 'delta CH4': 0}

Feedback from diy_greenhouse: The temperature is 27.52265167236328.

Assistant: Call future_temperature with parameters {'longitude': 14.8167, 'latitude': -0.8742, 'year': 2068, 'setting': 'ssp126'}

Feedback from future_temperature: The temperature is 27.692516326904297.

Assistant: Call answer_question with parameters {'thought': 'The temperature for Ewo in 2068 under the ssp126 scenario with no changes in emissions is 27.692516326904297. With a 25% decrease in CO2 emissions, the temperature is 27.52265167236328. The difference is 27.692516326904297 - 27.52265167236328 = 0.16986465454101562.', 'answer': 'B', 'solving_process': '1. Retrieved the latitude and longitude of Ewo. 2. Predicted the temperature for 2068 under the ssp126 scenario with a 25% decrease in CO2 emissions. 3. Retrieved the temperature for 2068 under the ssp126 scenario with no changes in emissions. 4. Calculated the difference between the two temperatures to find the change.'}

To calculate the temperature change in Ewo by 2068 under the ssp126 scenario with a 25% decrease in CO₂ emissions, we first identify Ewo's location at latitude -0.8742 and longitude 14.8167.

The temperature for Ewo in 2068 under the ssp126 scenario with no changes in emissions is determined to be 27.6925°C. With a 25% decrease in CO₂ emissions, the temperature is calculated to be 27.5227°C.

The difference between the two temperatures is:

$$27.6925^{\circ}\text{C} - 27.5227^{\circ}\text{C} = 0.1699^{\circ}\text{C}$$

This matches option B from the given choices. Therefore, the answer is **B**.

1. LLM generates the solution for a problem with tools:

$$y \sim \pi(\cdot | x, \{I_e\}_t, P_f)$$

2. Ground the LLM with:

$$J_{\text{WKD}}(\theta, \mathcal{D}) = -\mathbb{E}_{x \sim \mathcal{D}, y \sim \pi(\cdot | x, \{I_e\}_t, P_f)} [\log \pi_{\theta}(y | x, P_n)]$$

3. For open-ended problems, we use tools to rank different answers and train LLMs with DPO.

Our method: World Knowledge Distillation (**Learning**)

(a) Supervised Fine-tuning

Multi-Choice Question

How much will the temperature of Ribeira Brava in 2061 under ssp337 change if the emission of CO₂ is increased by 25%?

A. 0.05 B. 0.12 C. 0.23 D.0.30

LLM

Tools

1. Retrieve Data:
Time: 2061, Place: Ribeira Brava, Setting: ssp337

21.89°C

2. Emulate Scenario
Time: 2061, Place: Ribeira Brava, Setting: ssp337, Delta CO₂: 0.25

22.12°C

LLM

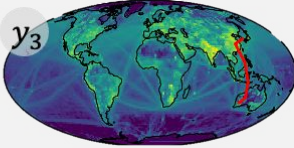
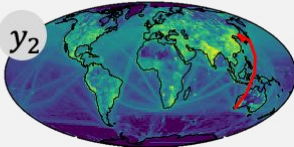
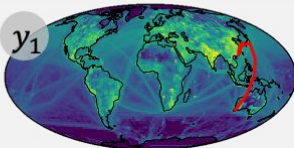
Answer
The temperature will increase by $22.12 - 21.89 = 0.23$ °C. The answer is C.

(b) Preference Optimization

Open-Ended Question

Design a new maritime route between Busan and Margaret River that would increase BC levels along the route by 3.0%, to minimize the global average temperature increase in 2027 under ssp585.

LLM



Tools

y₁

Temperature: 6.3424°C

Rank

y₁

Y

y₂

Temperature: 6.3449°C

y₂

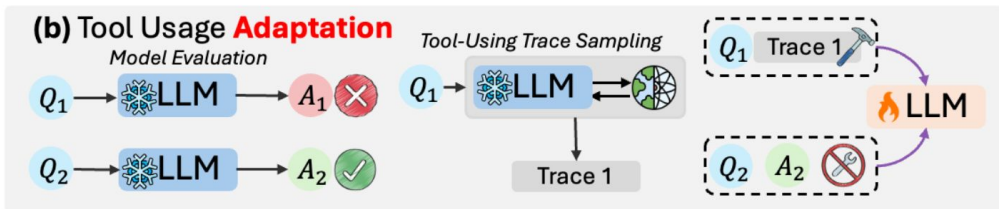
Y

y₃

Invalid: Interact with the land.

y₃

Our method: Tool Usage Adaptation (**Adapting**)



```
{'name': 'answer_question', 'parameters': {'thought': 'I can answer the problem directly.', 'answer': 'To calculate the temperature change in Ewo by 2068 under the ssp126 scenario with a 25% decrease in CO2 emissions, we first identify Ewo's location at latitude -0.8742 and longitude 14.8167. The temperature for Ewo in 2068 under the ssp126 scenario with no changes in emissions is determined to be 27.6925°C. With a 25% decrease in CO2 emissions, the temperature is calculated to be 27.5227°C. The difference between the two temperatures is: 27.6925°C - 27.5227°C = 0.1699°C. This matches option B from the given choices. Therefore, the answer is B.'}}
```

$$y \sim \pi(\cdot | x, \{I_e\}_t, P_f)$$

Assistant: Call query_lat_and_lon with parameters {'city_name': 'Ewo'}

$$t \sim \pi(\cdot | x, E, P_f)$$

Feedback from query_lat_and_lon: latitude: -0.8742, longitude: 14.8167.

Assistant: Call diy_greenhouse with parameters {'longitude': 14.8167, 'latitude': -0.8742, 'setting': 'ssp126', 'year': 2068, 'delta_CO2': -0.25, 'delta_CH4': 0}

Feedback from diy_greenhouse: The temperature is 27.52265167236328.

Assistant: Call future_temperature with parameters {'longitude': 14.8167, 'latitude': -0.8742, 'year': 2068, 'setting': 'ssp126'}

Feedback from future_temperature: The temperature is 27.692516326904297.

Assistant: Call answer_question with parameters {'thought': 'The temperature for Ewo in 2068 under the ssp126 scenario with no changes in emissions is 27.692516326904297. With a 25% decrease in CO2 emissions, the temperature is 27.52265167236328. The difference is 27.692516326904297 - 27.52265167236328 = 0.16986465454101562.', 'answer': 'B', 'solving_process': '1. Retrieved the latitude and longitude of Ewo. 2. Predicted the temperature for 2068 under the ssp126 scenario with a 25% decrease in CO2 emissions. 3. Retrieved the temperature for 2068 under the ssp126 scenario with no changes in emissions. 4. Calculated the difference between the two temperatures to find the change.'}

1. Evaluates the LLMs on a dataset and partition the questions into two subsets: Deasy and Dhard.

2. Set different alignment targets for Deasy and Dhard:

$$J_{\text{Easy}}(\theta, \mathcal{D}) = -\mathbb{E}_{x \sim \mathcal{D}, y \sim \pi(\cdot | x, \{I_e\}_t, P_f)} \log \pi_{\theta}(y | x, E, P_i).$$

$$J_{\text{Hard}}(\theta, \mathcal{D}) = -\mathbb{E}_{x \sim \mathcal{D}, t \sim \pi(\cdot | x, E, P_f)} \log \pi_{\theta}(t | x, E, P_i).$$

3. Train the model for both easy and hard problems with different targets:

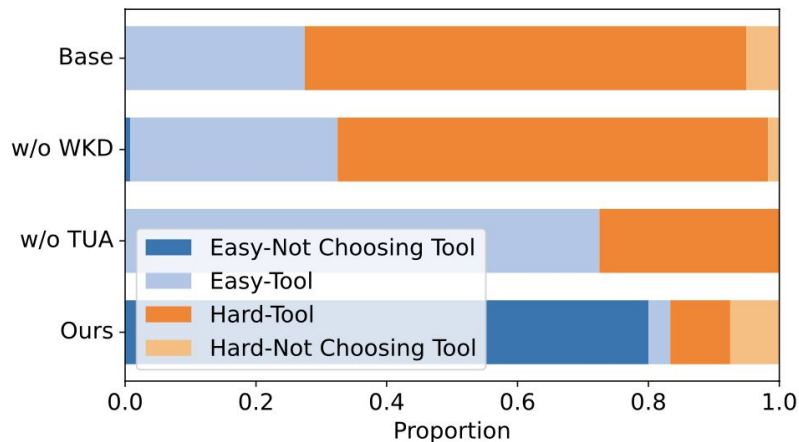
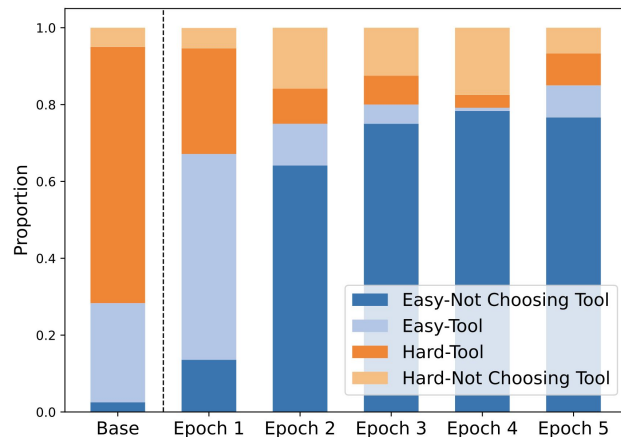
$$J_{\text{TUA}}(\theta, \mathcal{D}_{\text{easy}}, \mathcal{D}_{\text{hard}}) = \lambda J_{\text{Easy}}(\theta, \mathcal{D}_{\text{easy}}) + (1 - \lambda) J_{\text{Hard}}(\theta, \mathcal{D}_{\text{hard}})$$

Our method: **Adapting** while **Learning**

$$J_{\text{Mix}}(\theta, \mathcal{D}, \mathcal{D}_{\text{easy}}, \mathcal{D}_{\text{hard}}) = \alpha J_{\text{WKD}}(\theta, \mathcal{D}) + (1 - \alpha) J_{\text{TUA}}(\theta, \mathcal{D}_{\text{easy}}, \mathcal{D}_{\text{hard}});$$

Knowledge acquired under one scenario (with or without available tools) does not readily transfer to another[4].

We propose a mixed loss that simultaneously considers both WKD and TUA objectives.

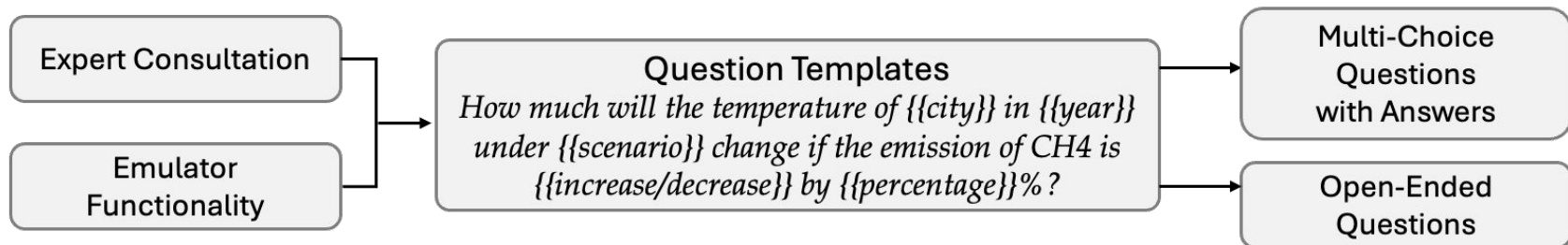


Experiment: Datasets


Public Datasets: MATH[5], SciBench[6]

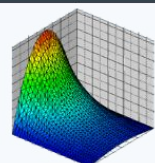
Our Custom-Created Datasets: Mujoco, PDE, Climate, Epidemiology

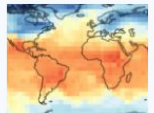
(a) Question Generation Pipeline




(b) Demo Questions

Mujoco	
	In a physics laboratory, a double pendulum experiment is set up with the following parameters: $\{\{parameters\}\}$ How does the position of the first pendulum change over the 1-second observation period?

PDE	
In a 2D population spread process with $\{\{initial\ condition\}\}$ and $\{\{boundary\ condition\}\}$, what is the population at $(x,y)=(59,2)$ km after 8 years? ($D=0.88$, $L_x=96$ km, $L_y=8$ km)	

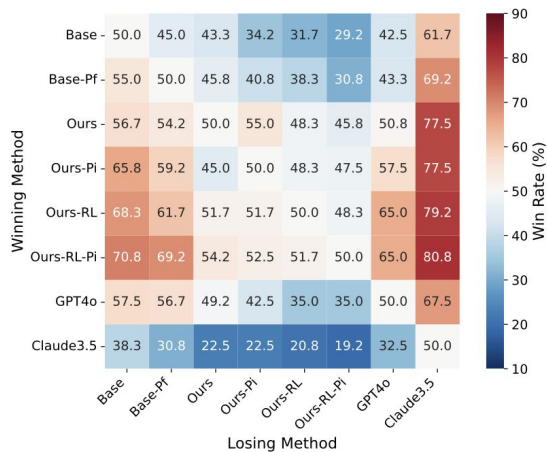
Climate	
	How much will the temperature of Gikongoro in 2018 under ssp585 change if the emission of CO ₂ is decreased by -25%?

Epidemiology	
$\{\{scenario\ description\}\}$ $\{\{initial\ states\}\}$ On which day will the number of hospitalized cases in California reach its maximum?	

Experiment: Answer Accuracy

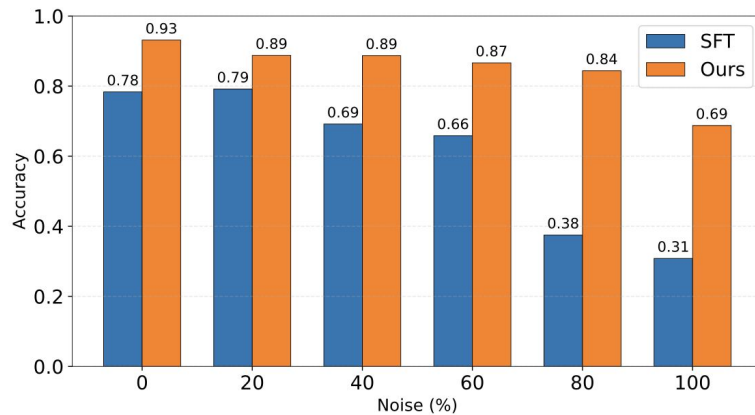
Models	Mujoco	PDE	Climate	Epidemiology	MATH	SciBench	Average
Llama3.1-70B	46.79	55.83	37.50	30.83	73.73	45.00	48.28
GPT4o	52.86	69.17	35.83	32.50	81.92	71.67	57.32
GPT4o-mini	51.79	70.83	30.00	35.83	<u>80.79</u>	<u>68.33</u>	56.26
Claude3.5-Sonnet	48.57	65.83	32.50	35.00	80.23	67.50	54.94
Llama3.1-8B (Base)- P_n	28.57	31.09	30.83	21.67	54.24	17.50	30.65
Llama3.1-8B (Base)- P_f	<u>59.32</u>	61.67	77.50	<u>57.78</u>	69.23	31.67	<u>59.53</u>
Llama3.1-8B-Ours- P_n	55.00	<u>75.00</u>	<u>80.00</u>	51.11	61.02	30.83	58.83
Llama3.1-8B-Ours- P_i	64.47	78.33	83.33	74.44	62.15	34.17	66.15

Overall Evaluation



Win Rate on
open-ended problems

Resistance to
Noise Data



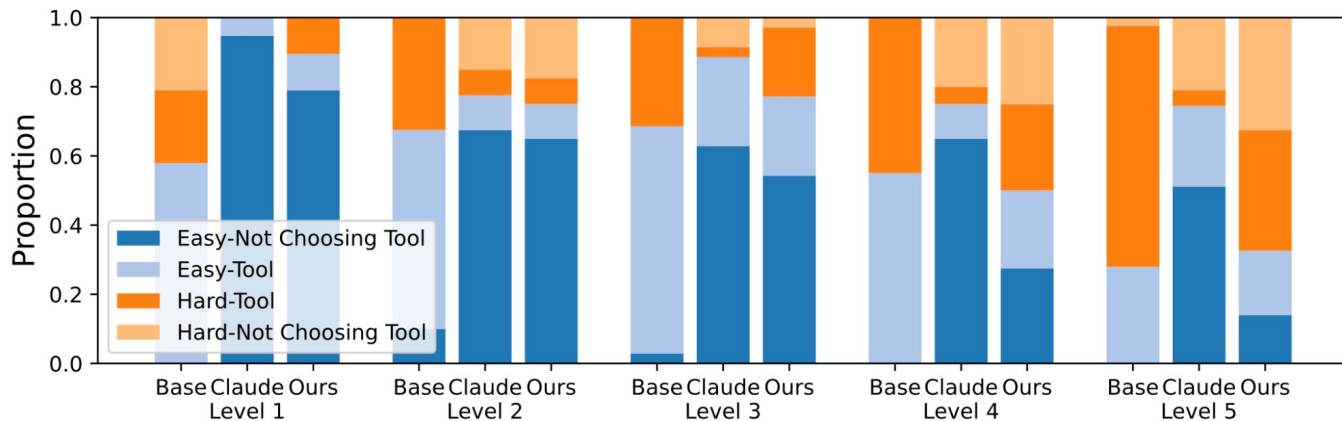
Experiment: Tool Usage Accuracy

	Tool (T)	Not Choosing Tool (N)
Easy (E)	ET	EN (✓)
Hard (H)	HT (✓)	HN

Models	Mujoco	PDE	Climate	Epidemiology	MATH	SciBench	Average
Llama3.1-70B	49.66	50.00	48.67	48.94	<u>55.77</u>	50.93	50.66
GPT4o	50.30	<u>52.41</u>	48.70	50.57	49.54	50.00	50.25
GPT4o-mini	50.34	52.35	48.81	<u>61.84</u>	55.19	68.36	<u>56.15</u>
Claude3.5-Sonnet	50.39	51.27	49.38	54.95	51.57	54.37	51.99
Llama3.1-8B (Base)	<u>51.50</u>	50.00	<u>50.35</u>	50.86	49.52	60.22	52.07
Llama3.1-8B-Ours	61.80	66.67	75.50	66.61	62.46	<u>62.75</u>	65.96

Overall Evaluation

$$\frac{1}{2} \times \left(\frac{EN}{EN+ET} + \frac{HT}{HN+HT} \right)$$



Problems of different difficulty levels on MATH dataset

Conclusion

We introduce a novel two-stage training paradigm that enables LLMs to adaptively solve real-world scientific problems of varying complexity.

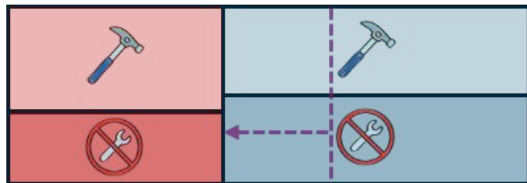
We construct four additional datasets spanning various scientific domains, including both questions and solutions, to facilitate future research in this direction.

Experiments on both custom and public datasets demonstrate the effectiveness of our work, resulting in better answer accuracy and more intelligent tool use.

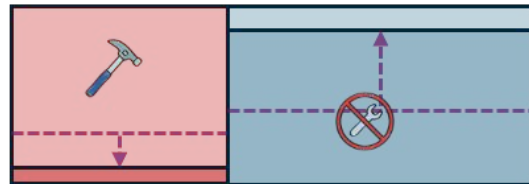
Original



World Knowledge Distillation



Tool Usage Adaptation



References

- [1] Zhang, D., Hu, Z., Zhoubian, S., Du, Z., Yang, K., Wang, Z., Yue, Y., Dong, Y., and Tang, J. Sciglm: Training scientific language models with self-reflective instruction annotation and tuning, 2024b.
- [2] Thulke, D., Gao, Y., Pelsler, P., Brune, R., Jalota, R., Fok, F., Ramos, M., Wyk, I., Nasir, A., Goldstein, H., et al. Climategpt: Towards ai synthesizing interdisciplinary research on climate change. arXiv preprint arXiv:2401.09646, 2024.
- [3] Liu, R., Wei, J., Gu, S. S., Wu, T., Vosoughi, S., Cui, C., Zhou, D., and Dai, A. M. Mind’s eye: Grounded language model reasoning through simulation. In The Eleventh International Conference on Learning Representations.
- [4] Zeng, A., Liu, M., Lu, R., Wang, B., Liu, X., Dong, Y., and Tang, J. AgentTuning: Enabling generalized agent abilities for LLMs. In Ku, L.-W., Martins, A., and Srikumar, V. (eds.), Findings of the Association for Computational Linguistics ACL 2024
- [5] Hendrycks, D., Burns, C., Kadavath, S., Arora, A., Basart, S., Tang, E., Song, D., and Steinhardt, J. Measuring mathematical problem solving with the math dataset. In Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2).
- [6] Wang, X., Hu, Z., Lu, P., Zhu, Y., Zhang, J., Subramaniam, S., Loomba, A. R., Zhang, S., Sun, Y., and Wang, W. SciBench: Evaluating College-Level Scientific Problem-Solving Abilities of Large Language Models. In Proceedings of the Forty-First International Conference on Machine Learning, 2024b.
- [7] Hendrycks, D.; Burns, C.; Basart, S.; Zou, A.; Mazeika, M.; Song, D.; and Steinhardt, J. 2021a. Measuring Massive Multitask Language Understanding.

Thanks!

Paper Link: <https://arxiv.org/abs/2411.00412>

Bohan Lyu
lyubh22@gmail.com
<http://lyubh.cn/>

QA

Veeramakali Vignesh Manivannan