



中國人民大學
RENMIN UNIVERSITY OF CHINA



高瓴人工智能学院
Gaoling School of Artificial Intelligence

ToRA: A Tool-Integrated Reasoning Agent for Mathematical Problem Solving

Zhibin Gou^{1,2}
Yujiu Yang¹

Zhihong Shao^{1,2}
Minlie Huang¹

Yeyun Gong²
Nan Duan²

Yelong Shen²
Weizhu Chen²

¹Tsinghua University

²Microsoft

2024-01-11



Background

LLM4Math

- (a). Rationale based (e.g. CoT prompting)
 - Semantic analysis, planning, and abstract reasoning... ✓
 - Computation, symbolic manipulation, and algorithmic processing ✗
- (b). Program based (e.g. PAL prompting)
 - Rigorous operations, outsource intricate calculations to specialized tools like equation solvers ✓

How about $(a) + (b)$? \Rightarrow

Background

Problem: Suppose that the sum of the squares of two complex numbers x and y is 7 and the sum of their cubes is 10. List all possible values for $x + y$, separated by commas.

We are given that $x^2 + y^2 = 7$ and $x^3 + y^3 = 10$. We can factor $x^3 + y^3$ to get $(x + y)(x^2 - xy + y^2)$.

Thus, we can write $10 = (x + y)(7)$. It follows that $x + y$ must divide 10.

The only possibilities are that $x + y$ is 1, 2, 5 or 10. ❌

import math

def solution():

$x = \text{symbols('x')}, y = \text{symbols('y')}$

$x^{**2} + y^{**2} = 7$

$x^{**3} + y^{**3} = 10$

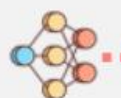
$x = (-1 + \text{math.sqrt}(17)) / 2$

$y = (-1 - \text{math.sqrt}(17)) / 2$

Get all possible values for $x + y$

result = $[x + y, -x - y]$

❌ `>>> SyntaxError: cannot assign to expression.` ❌



Rationale

(a) Rationale-based



Program

(b) Program-based

(a) + (b)

Let's use `SymPy` to calculate and print all possible values for $x + y$,

from sympy import symbols, Eq, solve, simplify

def possible_values():

$x, y = \text{symbols('x y')}$

$\text{eq1} = \text{Eq}(x^{**2} + y^{**2}, 7)$

$\text{eq2} = \text{Eq}(x^{**3} + y^{**3}, 10)$

$\text{solutions} = \text{solve}((\text{eq1}, \text{eq2}), (x, y))$

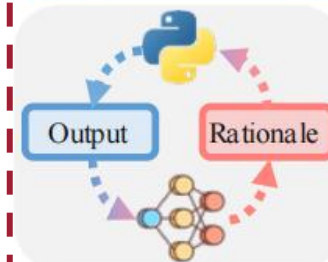
$\text{sums} = [\text{simplify}(\text{sol}[0] + \text{sol}[1]) \text{ for } \text{sol} \text{ in } \text{solutions}]$

return sums

print(possible_values())

Python icon `>>> [-5, -5, 1, 1, 4, 4]`

Removing duplicates, the possible values for $x + y$ are $\boxed{-5, 1, 4}$. ✅



(c) Tool-integrated Reasoning

(Format used by ToRA)

r_1

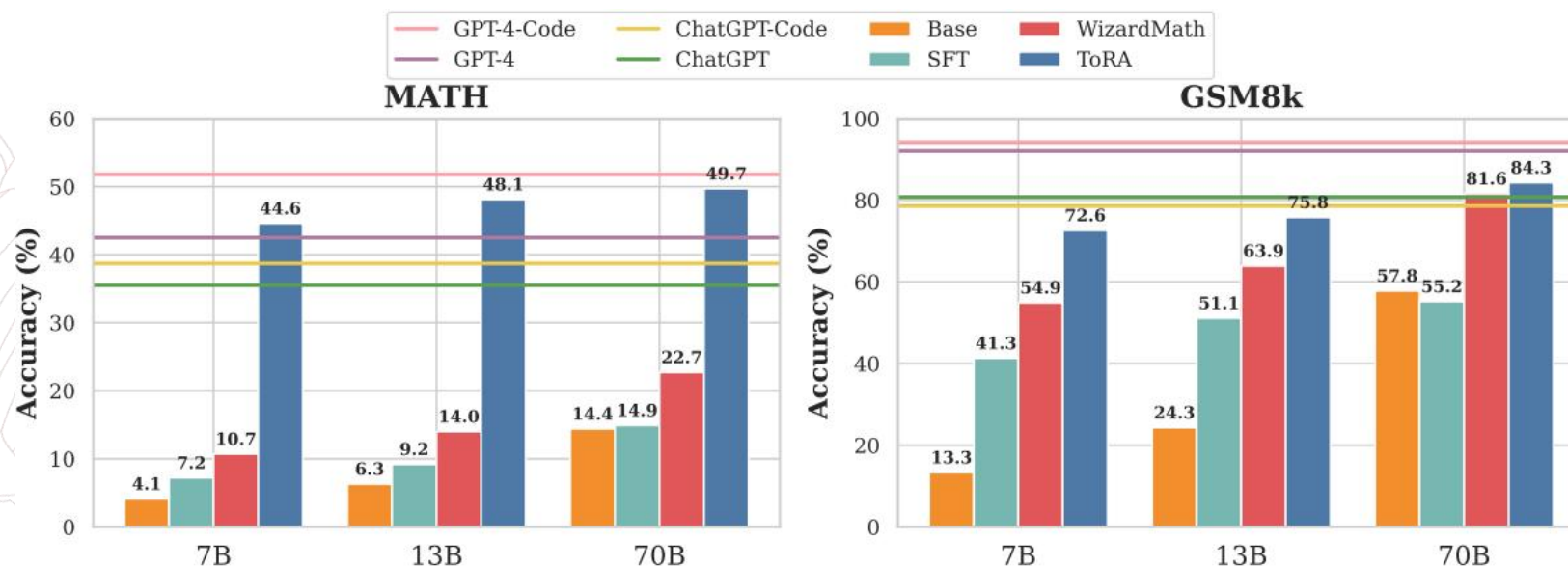
a_1

o_1

r_2

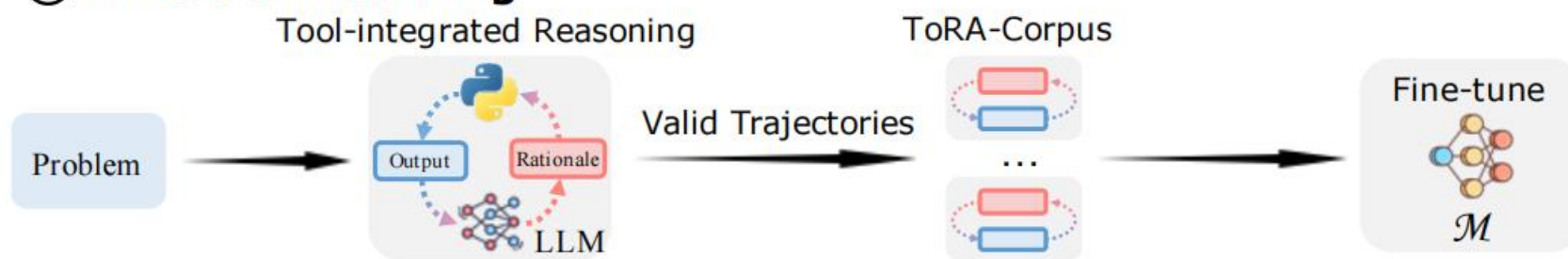
Contributions

- They propose **TORA**, a series of Tool-integrated Reasoning Agents designed to solve challenging mathematical problems.
 - Seamlessly integrating natural language reasoning with the utilization of tools.
 - Curate interactive tool-use trajectories on mathematical datasets. Use output space shaping to further refine models.
 - Significantly outperform open-source models. On MATH, Significantly outperform **GPT-4**'s CoT result, and is competitive with **GPT-4** solving problems with program.

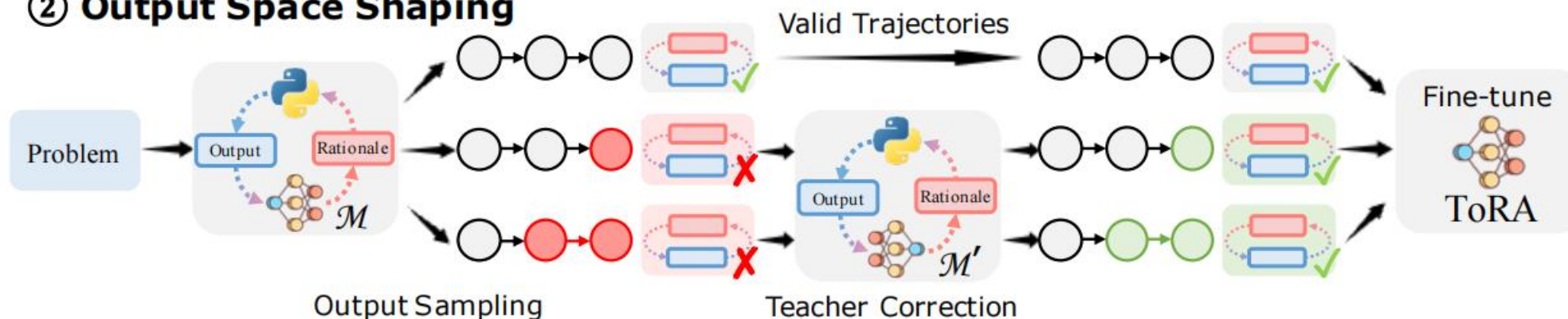


Method: Overview

① Imitation Learning



② Output Space Shaping



- ①. Prompt LLMs like **GPT-4** to generate Tool-integrated Reasoning trajectories (**TORA-CORPUS**) and use this corpus to fine-tune a model \mathbf{M} .
- ②. Sample diverse tool-use trajectories with \mathbf{M} , keep the valid ones, correct the invalid ones with a teacher model. Retrain \mathbf{M} with all data.



Method-step 1

- ①. Prompt LLMs like **GPT-4** to generate Tool-integrated Reasoning trajectories (**TORA-CORPUS**) and use this corpus to fine-tune a model **M**.

Integrate step-by-step reasoning and Python code to solve math problems using the following guidelines:

- Analyze the question and write functions to solve the problem; the function should not take any arguments.
- Present the final result in LaTeX using a `\boxed{}` without any units.
- Utilize the `'pi'` symbol and `'Rational'` from Sympy for π and fractions, and simplify all fractions and square roots without converting them to decimal values.

Here are some examples you may refer to:

Question: Convert the point $(0, -3\sqrt{3}, 3)$ in rectangular coordinates to spherical coordinates. Enter your answer in the form (ρ, θ, ϕ) , where $\rho > 0$, $0 \leq \theta < 2\pi$, and $0 \leq \phi \leq \pi$.



Method-step 1

Solution:

To convert the point (x, y, z) in rectangular coordinates to spherical coordinates (ρ, θ, ϕ) , we use the following formulas:

1. $\rho = \sqrt{x^2 + y^2 + z^2}$
2. $\theta = \arctan\{\frac{y}{x}\}$
3. $\phi = \arccos\{\frac{z}{\rho}\}$

Let's write a function:

```
```python
from sympy import sqrt, atan2, acos, pi

def rectangular_to_spherical():
 x, y, z = 0, -3*sqrt(3), 3
 rho = sqrt(x**2 + y**2 + z**2)
 theta = atan2(y, x)
 phi = acos(z/rho)
 return rho, theta, phi

spherical_coordinates = rectangular_to_spherical()
print(spherical_coordinates)
```

```output
(6, -pi/2, pi/3)
```
```

The required form of the answer is (ρ, θ, ϕ) , where $\rho > 0$, $0 \leq \theta < 2\pi$, and $0 \leq \phi \leq \pi$. Since our initial output has $\theta = -\pi/2$, which doesn't meet the condition for θ , we add 2π to adjust its value.

Hence, the spherical coordinates for the point $(0, -3\sqrt{3}, 3)$ are $\boxed{(6, \frac{3\pi}{2}, \frac{\pi}{3})}$.

Method-step 1

Algorithm 1 Inference of Tool-Integrated Reasoning

Require: problem q , model \mathcal{G} , prompt \wp , external tools \mathcal{E} , stop condition $Stop(\cdot)$, maximum iteration rounds n

```
1:  $\tau_0 \leftarrow ""$  ▷ Trajectory Initialization
2: for  $i \leftarrow 1$  to  $n$  do
3:    $r_i \sim \mathbb{P}_{\mathcal{G}}(\cdot | \wp \oplus q \oplus \tau_{i-1})$  ▷ Rationale Generation (Eq. 1)
4:   if  $Stop(r_i)$  then ▷ Stopping Criteria
5:     return  $\tau_{i-1} \oplus r_i$ 
6:   end if
7:    $a_i \sim \mathbb{P}_{\mathcal{G}}(\cdot | \wp \oplus q \oplus \tau_{i-1} \oplus r_i)$  ▷ Program Generation (Eq. 2)
8:    $o_i \leftarrow \mathcal{E}(a_i)$  ▷ Tool Execution
9:    $\tau_i \leftarrow \tau_{i-1} \oplus r_i \oplus a_i \oplus o_i$  ▷ Trajectory Update (Eq. 3)
10: end for
11: return  $\tau_n$ 
```



$$\mathcal{M} = \arg \min_{\theta} \sum_{q, \tau} \sum_{i=1}^{n-1} -\log \mathbb{P}_{\theta}(r_{i+1} a_{i+1} | q, r_1 \dots o_i)$$

τ_i : trajectory until i -th step

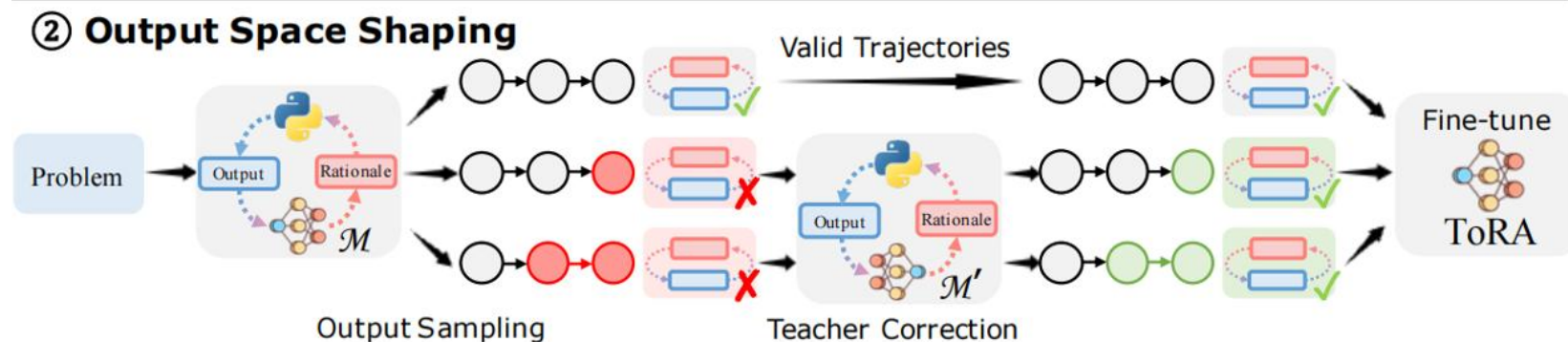
r_i : natural language

a_i : program

o_i : output

Method-step 2

②. Sample diverse tool-use trajectories with **M**, keep the valid ones, correct the invalid ones with a teacher model. Retrain **M** with all data. (**GPT4** \Rightarrow **M**)



Observe that trajectories with wrong answers are mostly incorrect **halfway**, and the preceding reasoning is still plausible.

① + ② \Rightarrow ^{69k} TORA-CORPUS, Another valid trajectories, ToRA

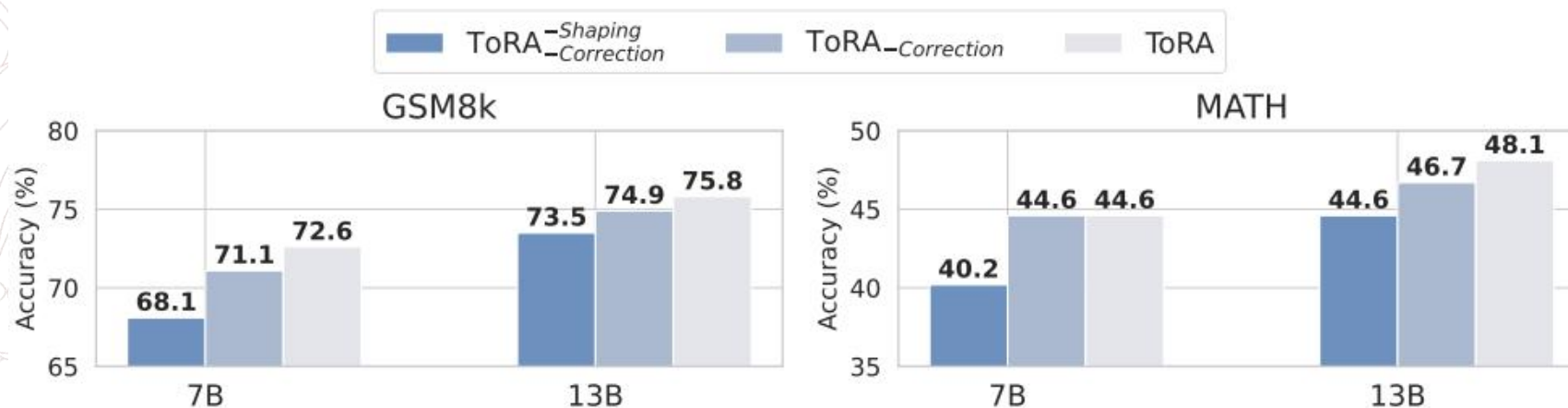
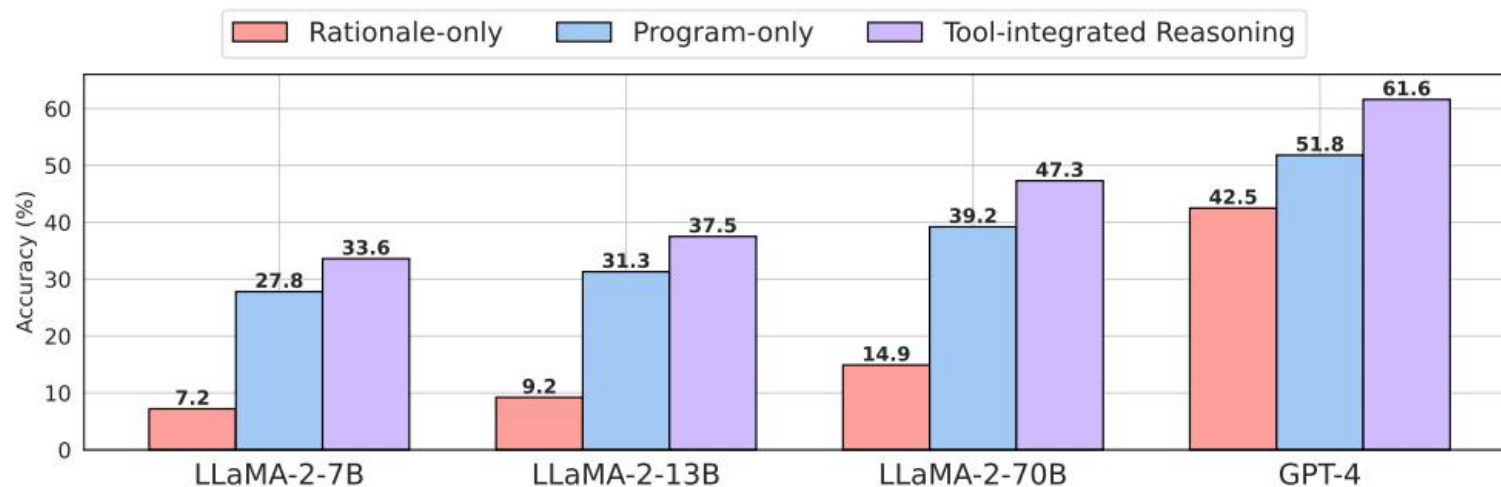


| Model | Size | Tools | ZS* | GSM8k | MATH | GSM-Hard | SVAMP | TabMWP | ASDiv | MAWPS | AVG |
|---------------------------|------|-------|-----|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------------|
| Used for training? | | | | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | |
| Proprietary Models | | | | | | | | | | | |
| GPT-4 | - | ✗ | ✗ | 92.0 | 42.5 | 64.7 | 93.1 | 67.1 | 91.3 | 97.6 | 78.3 |
| GPT-4 (PAL) 🤖 | - | ✓ | ✗ | 94.2 | 51.8 | 77.6 | 94.8 | 95.9 | 92.6 | 97.7 | 86.4 |
| ChatGPT | - | ✗ | ✗ | 80.8 | 35.5 | 55.9 | 83.0 | 69.1 | 87.3 | 94.6 | 72.3 |
| ChatGPT (PAL) 🤖 | - | ✓ | ✗ | 78.6 | 38.7 | 67.6 | 77.8 | 79.9 | 81.0 | 89.4 | 73.3 |
| Claude-2 | - | ✗ | ✗ | 85.2 | 32.5 | - | - | - | - | - | - |
| PaLM-2 | 540B | ✗ | ✗ | 80.7 | 34.3 | - | - | - | - | - | - |
| Open-Source Models | | | | | | | | | | | |
| LLaMA-2 | 7B | ✗ | ✗ | 13.3 | 4.1 | 7.8 | 38.0 | 31.1 | 50.7 | 60.9 | 29.4 |
| LLaMA-2 SFT | 7B | ✗ | ✓ | 41.3 | 7.2 | 16.1 | 31.9 | 27.8 | 47.4 | 60.0 | 33.1 |
| LLaMA-2 RFT | 7B | ✗ | ✓ | 51.2 | - | - | - | - | - | - | - |
| Platypus-2 | 7B | ✗ | ✗ | 14.4 | 5.4 | 8.6 | 36.7 | 26.5 | 47.9 | 58.4 | 28.3 |
| WizardMath | 7B | ✗ | ✓ | 54.9 | 10.7 | 20.6 | 57.3 | 38.1 | 59.1 | 73.7 | 44.9 |
| CodeLLaMA (PAL) 🤖 | 7B | ✓ | ✗ | 34.0 | 16.6 | 33.6 | 59.0 | <u>47.3</u> | 61.4 | 79.6 | 47.4 |
| Toolformer [†] 🧠 | 7B | ✓ | ✓ | - | - | - | 29.4 | - | 40.4 | 44.0 | - |
| ToRA 🏆 | 7B | ✓ | ✓ | <u>68.8</u> | <u>40.1</u> | <u>54.6</u> | <u>68.2</u> | 42.4 | <u>73.9</u> | <u>88.8</u> | <u>62.4</u> |
| ToRA-CODE 🏆 | 7B | ✓ | ✓ | 72.6 | 44.6 | 56.0 | 70.4 | 51.6 | 78.7 | 91.3 | 66.5 (+19) |
| LLaMA-2 | 13B | ✗ | ✗ | 24.3 | 6.3 | 13.6 | 43.1 | 39.5 | 56.3 | 70.4 | 36.2 |
| LLaMA-2 SFT | 13B | ✗ | ✓ | 51.1 | 9.2 | 22.3 | 46.3 | 35.8 | 58.6 | 75.0 | 42.6 |
| LLaMA-2 RFT | 13B | ✗ | ✓ | 55.3 | - | - | - | - | - | - | - |
| Platypus-2 | 13B | ✗ | ✗ | 23.7 | 7.1 | 14.3 | 50.7 | 45.3 | 55.1 | 69.6 | 38.0 |
| WizardMath | 13B | ✗ | ✓ | 63.9 | 14.0 | 28.4 | 64.3 | 46.7 | 65.8 | 79.7 | 51.8 |
| CodeLLaMA (PAL) 🤖 | 13B | ✓ | ✗ | 39.9 | 19.9 | 39.0 | 62.4 | <u>59.5</u> | 65.3 | 86.0 | 53.1 |
| ToRA 🏆 | 13B | ✓ | ✓ | <u>72.7</u> | <u>43.0</u> | <u>57.3</u> | <u>72.9</u> | 47.2 | <u>77.2</u> | <u>91.3</u> | <u>65.9</u> |
| ToRA-CODE 🏆 | 13B | ✓ | ✓ | 75.8 | 48.1 | 60.5 | 75.7 | 65.4 | 81.4 | 92.5 | 71.3 (+18) |
| LLaMA-1 RFT | 34B | ✗ | ✓ | 57.9 | - | - | - | - | - | - | - |
| CodeLLaMA (PAL) 🤖 | 34B | ✓ | ✗ | 53.3 | 23.9 | 49.4 | 71.0 | 63.1 | 72.4 | 91.5 | 60.7 |
| ToRA-CODE 🏆 | 34B | ✓ | ✓ | 80.7 | 50.8 | 63.7 | 80.5 | 70.5 | 84.2 | 93.3 | 74.8 (+14) |
| LLaMA-2 | 70B | ✗ | ✗ | 57.8 | 14.4 | 36.0 | 73.6 | 57.5 | 76.0 | 92.4 | 58.2 |
| LLaMA-2 SFT | 70B | ✗ | ✓ | 69.3 | 14.9 | 39.0 | 64.0 | 53.0 | 71.3 | 84.8 | 56.6 |
| LLaMA-2 RFT | 70B | ✗ | ✓ | 64.8 | - | - | - | - | - | - | - |
| Platypus-2 | 70B | ✗ | ✗ | 45.9 | 15.0 | 24.6 | 74.3 | 47.3 | 72.7 | 91.1 | 53.0 |
| WizardMath | 70B | ✗ | ✓ | <u>81.6</u> | <u>22.7</u> | <u>50.3</u> | <u>80.0</u> | 49.8 | <u>76.2</u> | 86.2 | <u>63.8</u> |
| LLaMA-2 (PAL) 🤖 | 70B | ✓ | ✗ | 55.2 | 18.3 | 50.0 | 74.6 | <u>59.5</u> | 71.9 | <u>92.8</u> | 60.3 |
| ToRA 🏆 | 70B | ✓ | ✓ | 84.3 | 49.7 | 67.2 | 82.7 | 74.0 | 86.8 | 93.8 | 76.9 (+13) |









Experiments

➤ Ablation Study.





Experiments

| Model | Size | Tool | Intermediate Algebra | Precalculus | Geometry | Number Theory | Counting & Probability | Prealgebra | Algebra | Overall |
|-------------------------------------------------------------------------------------------------|------|------|----------------------|---------------------|---------------------|---------------------|------------------------|---------------------|---------------------|---------------------|
| Proprietary Models | | | | | | | | | | |
| ChatGPT (PAL)  | - | ✓ | 18.5 | 19.2 | 23.2 | 48.5 | 43.0 | 62.7 | 45.4 | 38.7 |
| GPT-4 (PAL)  | - | ✓ | 32.8 | 29.3 | 38.0 | 58.7 | 61.0 | 73.9 | 59.1 | 51.8 |
| Open-Source Models | | | | | | | | | | |
| WizarMath | 7B | ✗ | 6.2 | 6.0 | 6.5 | 7.6 | 9.5 | 18.1 | 16.3 | 11.2 |
| TORA-CODE  | 7B | ✓ | 35.1 (+28.9) | 31.0 (+25.0) | 24.0 (+17.5) | 50.7 (+43.1) | 30.6 (+21.1) | 55.0 (+36.9) | 61.7 (+45.4) | 44.6 (+33.4) |
| w/o Shaping | 7B | ✓ | 29.7 (-5.4) | 25.1 (-5.9) | 17.7 (-6.3) | 46.9 (-3.8) | 32.3 (+1.7) | 51.9 (-3.1) | 55.7 (-6.0) | 40.2 (-4.4) |
| w/o Rationale | 7B | ✓ | 25.5 (-9.6) | 14.7 (-16.3) | 15.4 (-8.6) | 45.9 (-4.8) | 29.7 (-0.9) | 51.0 (-4.0) | 52.4 (-9.3) | 36.8 (-7.8) |
| WizarMath | 13B | ✗ | 6.4 | 6.6 | 11.5 | 9.6 | 11.0 | 28.5 | 21.1 | 15.0 |
| TORA-CODE  | 13B | ✓ | 35.7 (+29.3) | 31.1 (+24.5) | 25.7 (+14.2) | 55.6 (+46.0) | 39.5 (+28.5) | 58.7 (+30.2) | 66.7 (+45.6) | 48.1 (+33.1) |
| w/o Shaping | 13B | ✓ | 32.8 (-2.9) | 26.0 (-5.1) | 24.0 (-1.7) | 52.6 (-3.0) | 38.4 (-1.1) | 55.6 (-3.1) | 61.2 (-5.5) | 44.6 (-3.5) |
| w/o Rationale | 13B | ✓ | 27.1 (-8.6) | 15.8 (-15.3) | 16.3 (-9.4) | 50.4 (-5.2) | 36.9 (-2.6) | 55.3 (-3.4) | 56.5 (-10.2) | 40.2 (-7.9) |
| TORA-CODE  | 34B | ✓ | 38.9 | 34.6 | 27.3 | 57.8 | 41.4 | 63.7 | 67.7 | 50.8 |
| w/o Shaping | 34B | ✓ | 34.0 (-4.9) | 29.9 (-4.7) | 24.6 (-2.7) | 55.6 (-2.2) | 41.6 (+0.2) | 63.8 (+0.1) | 61.4 (-6.3) | 47.4 (-3.4) |
| w/o Rationale | 34B | ✓ | 28.3 (-10.6) | 15.8 (-18.8) | 18.0 (-9.3) | 52.4 (-5.4) | 40.7 (-0.7) | 58.6 (-5.1) | 57.5 (-10.2) | 41.9 (-8.9) |
| WizarMath | 70B | ✗ | 9.1 | 13.4 | 16.9 | 16.5 | 19.2 | 42.7 | 35.0 | 24.1 |
| TORA  | 70B | ✓ | 37.1 (+28) | 30.4 (+17) | 30.1 (+13.2) | 54.6 (+38.1) | 40.3 (+21.1) | 64.9 (+22.2) | 66.6 (+31.6) | 49.7 (+25.6) |
| w/o Shaping | 70B | ✓ | 33.8(-3.3) | 28.9(-1.5) | 27.1(-3) | 53.0(-1.6) | 38.0(-2.3) | 62.2(-2.7) | 64.2(-2.4) | 47.3(-2.4) |
| w/o Rationale | 70B | ✓ | 26.7(-10.4) | 14.7(-15.7) | 20.3(-9.8) | 48.9(-5.7) | 39.2(-1.1) | 59.8(-5.1) | 57.6(-9) | 41.5(-8.2) |



Conclusion



- Presents TORA, a series of novel Tool-integrated Reasoning Agents that synergistically combines natural language rationale with program-based tool-use for mathematical problem solving.
- TORA achieves state-ofthe-art performance on 10 diverse mathematical reasoning tasks, substantially outperforming existing rationale-based and program-based approaches.





中國人民大學
RENMIN UNIVERSITY OF CHINA



高瓴人工智能学院
Gaoling School of Artificial Intelligence

Thank You for listening!