Babysitting a Small Language Model through One-Step Tree-of-Thoughts Knowledge Distillation

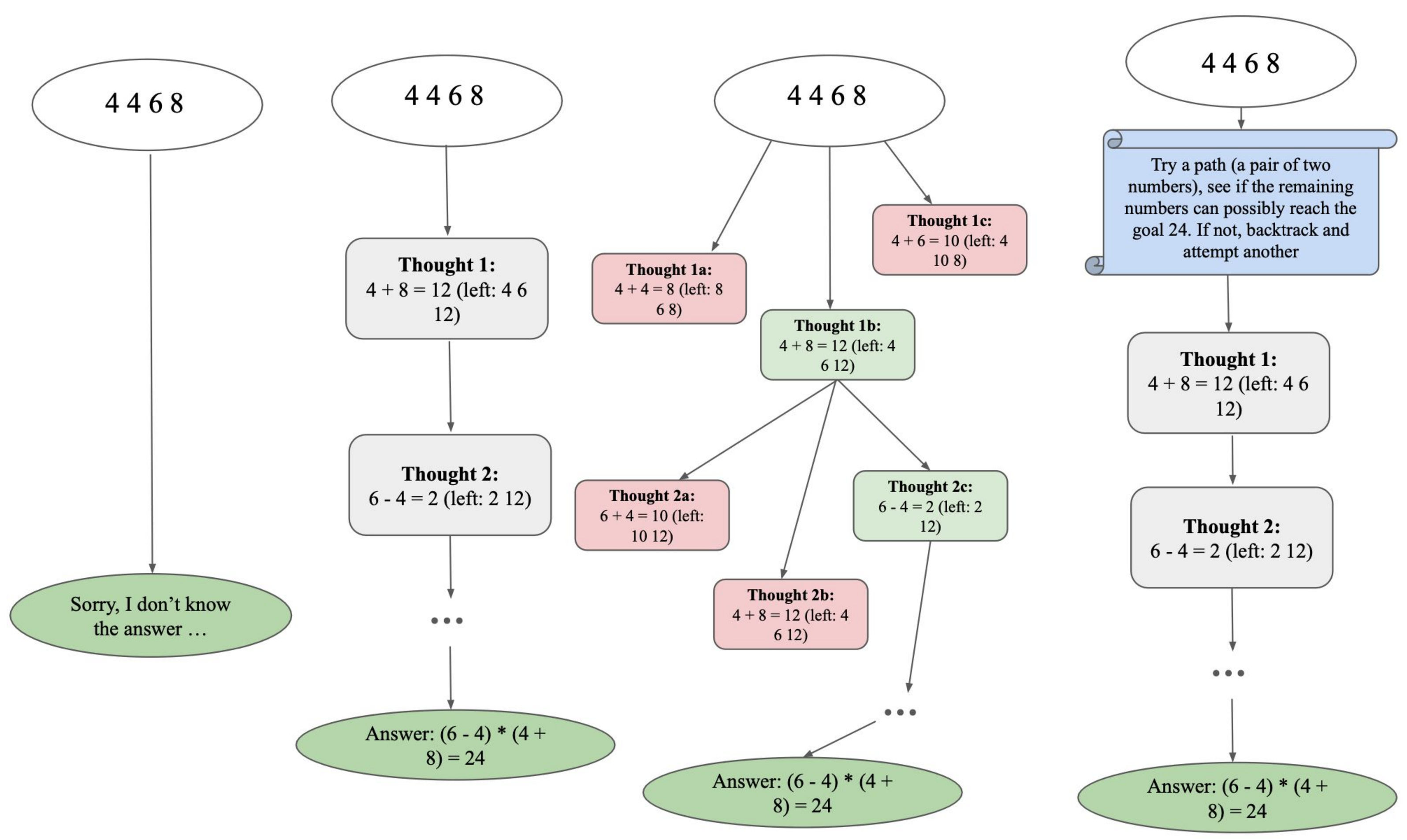
MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

Anurag Renduchintala, Adi Mahesh, Zichen Zhang, Zimo Si, Shangjun Meng, Samuel Fang

{ranurag, mahesha, zhangzzc, zimosi, shangjun, swfang}@umich.edu

Motivation(s) and Introduction

- **Large Language Models (LLMs),** while revolutionary, are not so suitable for tasks that require multi-step reasoning.
- **Small Language Models (SLMs),** while efficient, face the same challenges as LLMs with multi-step reasoning tasks.
- **Our One-Step ToT framework** reimagines the multi-step ToT framework into a single, structured prompt and distills the reasoning capabilities of LLMs into SLMs through knowledge distillation. We test our framework on **Game of 24.**



CoT & ToT Replication

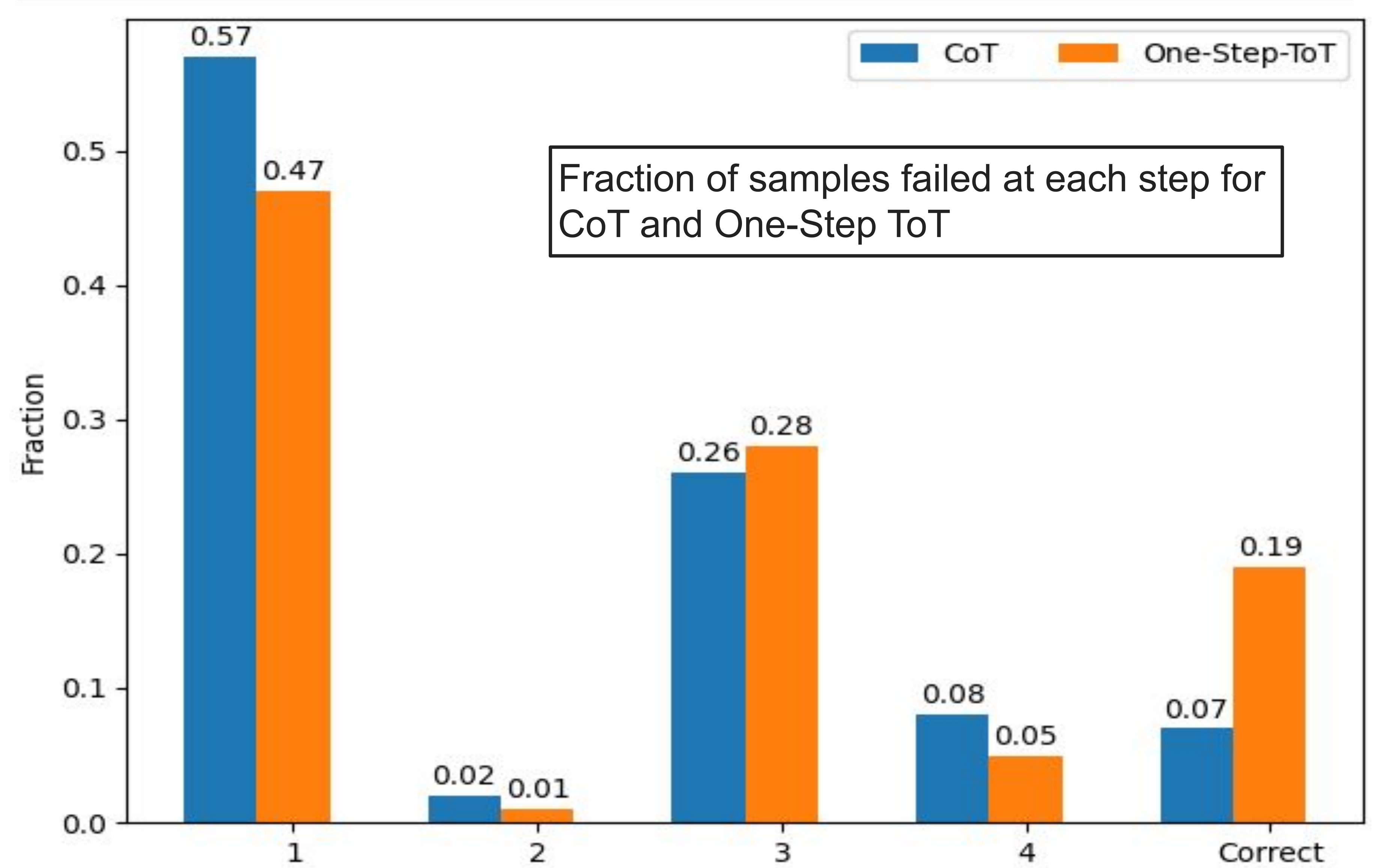
- Tested the **CoT 5-shot prompt** from the paper on 100 puzzles using the same experimental setup as the authors, with temperature set to 0.7 and GPT-4 for consistency.
- Replicated **ToT** using Breadth-First Search (BFS) and varied breadth (b=1 to 5). However, unlike the paper, we used **GPT-4o** to facilitate parsing of responses.

Method	Success
IO prompt (Yao et al. 2023)	7.3%
CoT-SC ($k = 100$) (Yao et al. 2023)	9%
Replicated CoT prompt ($k = 1$)	7%
Replicated ToT ($b = 1$)	4%
Replicated ToT ($b = 3$)	68%
Replicated ToT ($b = 5$)	**82%**
One-Step ToT ($k = 1$)	**19%**

One-Step ToT (Extension)

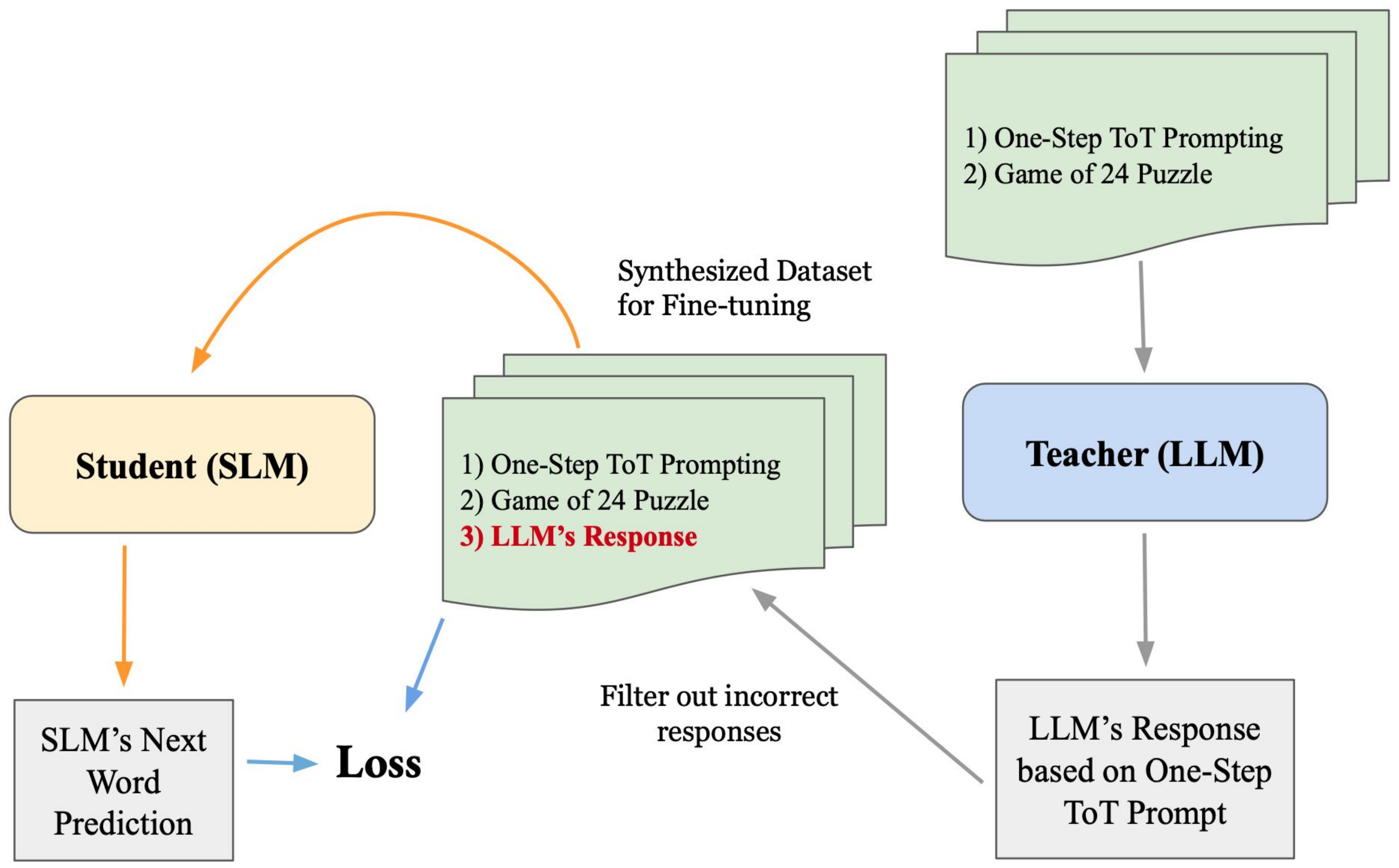
- Our methodology is similar to that used in CoT and ToT. However, we change our prompting style by **introducing a system prompt** to induce ToT Reasoning, removing the need for multiple input prompts like in ToT.
- **One-Step ToT** achieved an **accuracy of 19%**, almost triple that of CoT (7%).

Input:	Correct Output:
Use numbers and basic arithmetic operations ($+, -, *, /$) to obtain 24. Each step, you are only allowed to choose two of the remaining numbers to obtain a new number. **Step 1**: Start by considering possible operations for each pair of numbers. **Step 2**: Try a path (a pair of two numbers), see if the remaining numbers can possibly reach the goal 24. If not, backtrack and attempt another. **Step 3**: Branch out to try different orders of operations and combinations, evaluating each outcome. **Step 4**: If one path doesn't lead to a solution, backtrack and try alternative operations. {in_context_demonstrations}. Solve the following puzzle: **4 4 6 8**.	Steps: $4 + 8 = 12$ (left: 4, 6, 12) $6 - 4 = 2$ (left: 2, 12) $2 * 12 = 24$ (left: 24) Answer: $(6 - 4) * (4 + 8) = 24$



Knowledge Distillation (Extension)

- We first run the LLM, GPT-4o, on the entire game of 24 dataset wo generate the data for training.
- We use a small model, SmolLM-360M, which provides a balance between performance and computational efficiency.



Model with Prompt Method	Success
GPT-4 with IO prompt (Yao et al. 2023)	7.3%
GPT-4 with CoT-SC ($k = 100$) (Yao et al. 2023)	9%
GPT-4 with CoT ($k = 1$)	4%
GPT-4o with Replicated CoT ($k = 1$)	7%
GPT-4o with Replicated ToT ($b = 1$)	4%
Original SmolLM with One-Step ToT	1%
Our **fine-tuned SmolLM** with One-Step ToT	**9%**

Conclusion

- We propose **One-Step ToT,** and demonstrated its effectiveness over naive CoT.
- We demonstrate that after **distilling** ToT-style knowledge into an **SLM**, the SLM can achieve significant improvements and rival LLMs like GPT-4o.
- We provide a comprehensive analysis of how knowledge distillation and efficient prompting can enhance reasoning capabilities in resource-constrained language models.

Selected References

1. Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Ichter, B.; Xia, F.; Chi, E.; Le, Q.; and Zhou, D. 2023. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. arXiv:2201.11903.
2. Yao, S.; Yu, D.; Zhao, J.; Shafran, I.; Griffiths, T. L.; Cao, Y.; and Narasimhan, K. 2023. Tree of Thoughts: Deliberate Problem Solving with Large Language Models. arXiv:2305.10601.
3. Magister, L. C.; Mallinson, J.; Adamek, J.; Malmi, E.; and Severyn, A. 2023. Teaching Small Language Models to Reason. arXiv:2212.08410.