Self-Improving Transformers Overcome Easy-to-Hard & Length Generalization Challenges

Nayoung Lee[*w] *Jack Cai*[*w] *Avi Schwarzschild*[c] *Kangwook Lee*[w] *Dimitris Papailiopoulos*[w]
[w] University of Wisconsin-Madison [c] Carnegie Mellon University

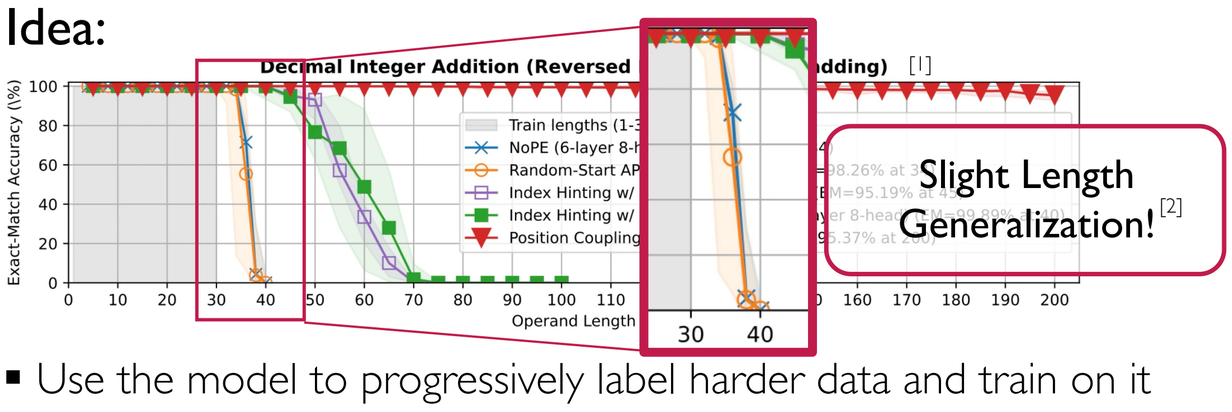
1. Introduction

Motivation:

- LLMs, despite their success, struggle to *Length Generalize* – extrapolating to longer sequences than seen during training
- Recent attempts involve modifying data formats / architecture / positional encodings, leading to task-specific solutions

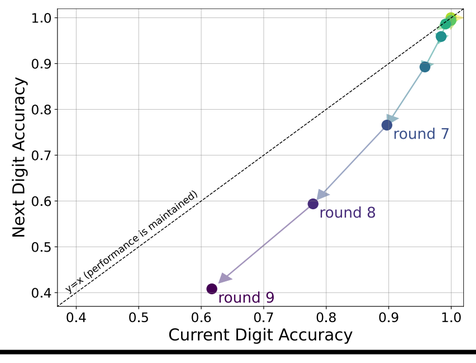
Q: Can we solve Length Generalization *without* changing data/architecture/PE?

Idea:



Slight Length Generalization! [2]

- Use the model to progressively label harder data and train on it (=*self-improvement*!)

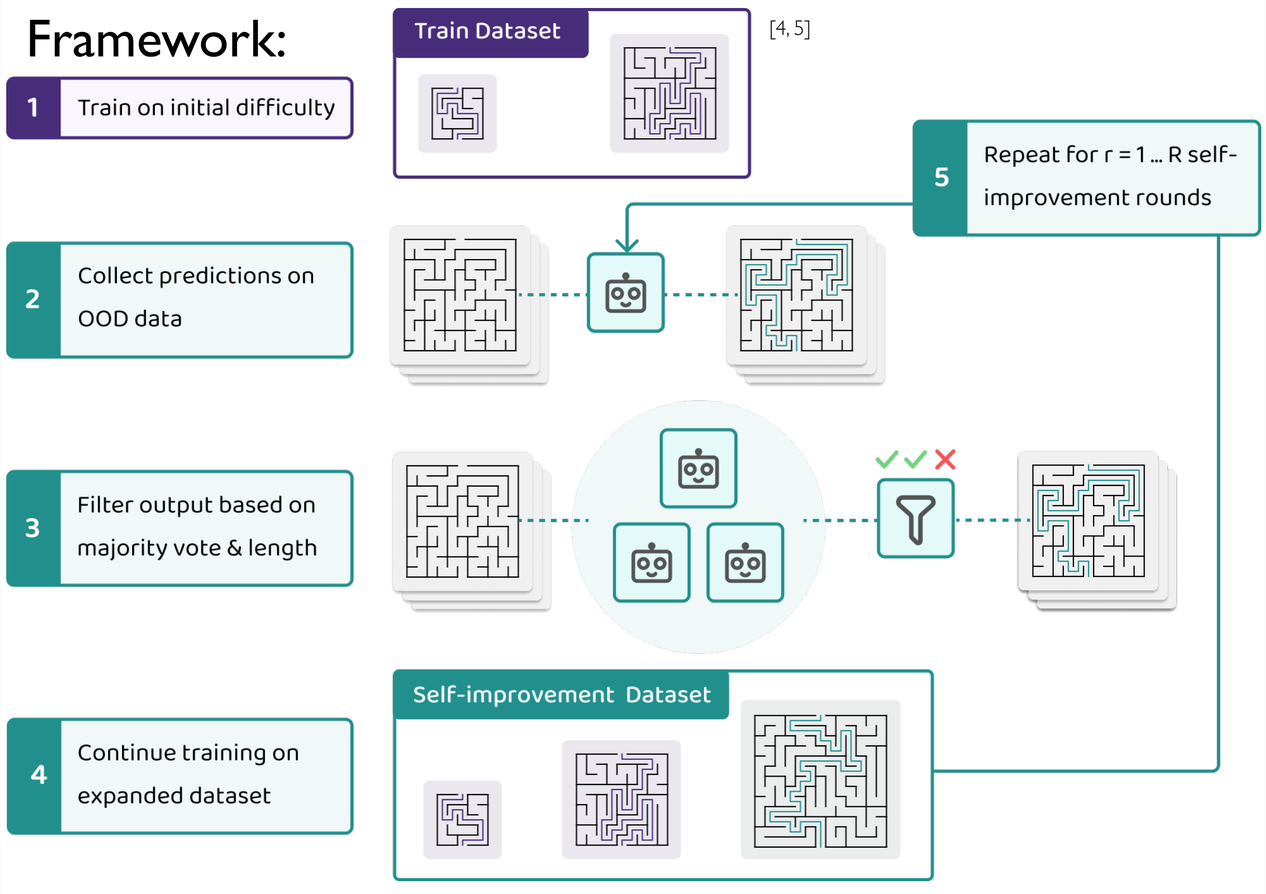


Challenge:

- Error Avalanches! [3]

2. Self-Improvement Framework

Framework:



Data Filtering is Important!

Majority Vote Filtering

- Train several parallel models
- Admit example only if enough predictions match

Unsupervised!

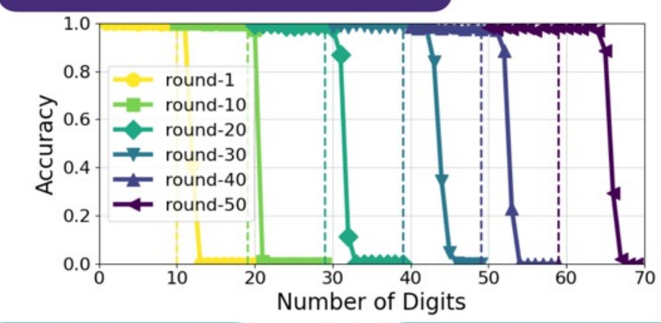
Relative Length Filtering

- Find max length in batch
- Filter example shorter than (max length - constant)

Unsupervised!

3. Overall Results

Forward Addition

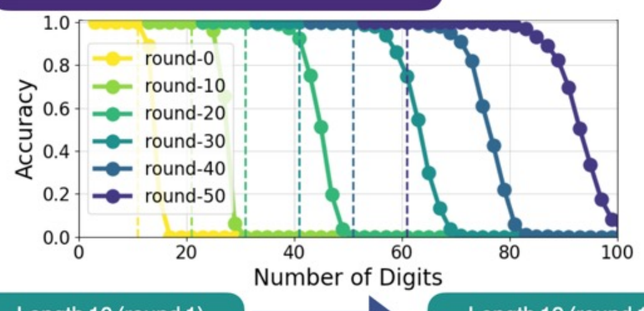


10-digit (round 1): 1687477129+1095477427=2782954556

Digits ↑

19-digit (round 10): 364250722784280616 2+491464396527962313 1=557151193122429293

String Copying

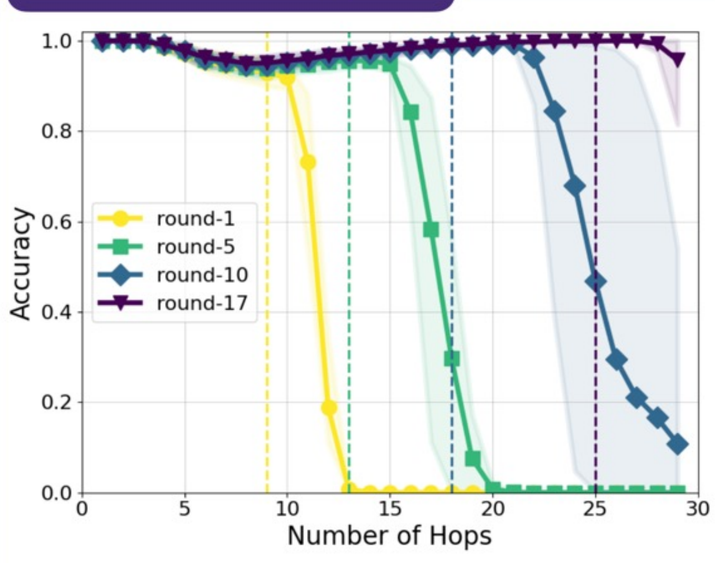


Length 10 (round 1): 1095477427=1095477427

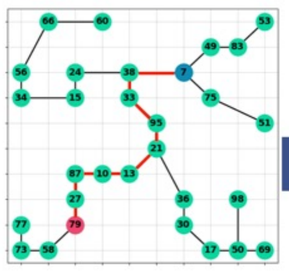
Length ↑

Length 19 (round 10): 4914643965279623131=4914643965279623131

Maze-Solving

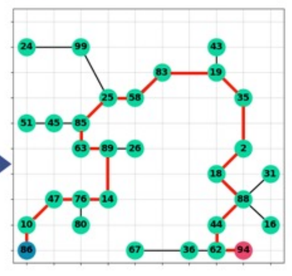


9 Hops (round 1)

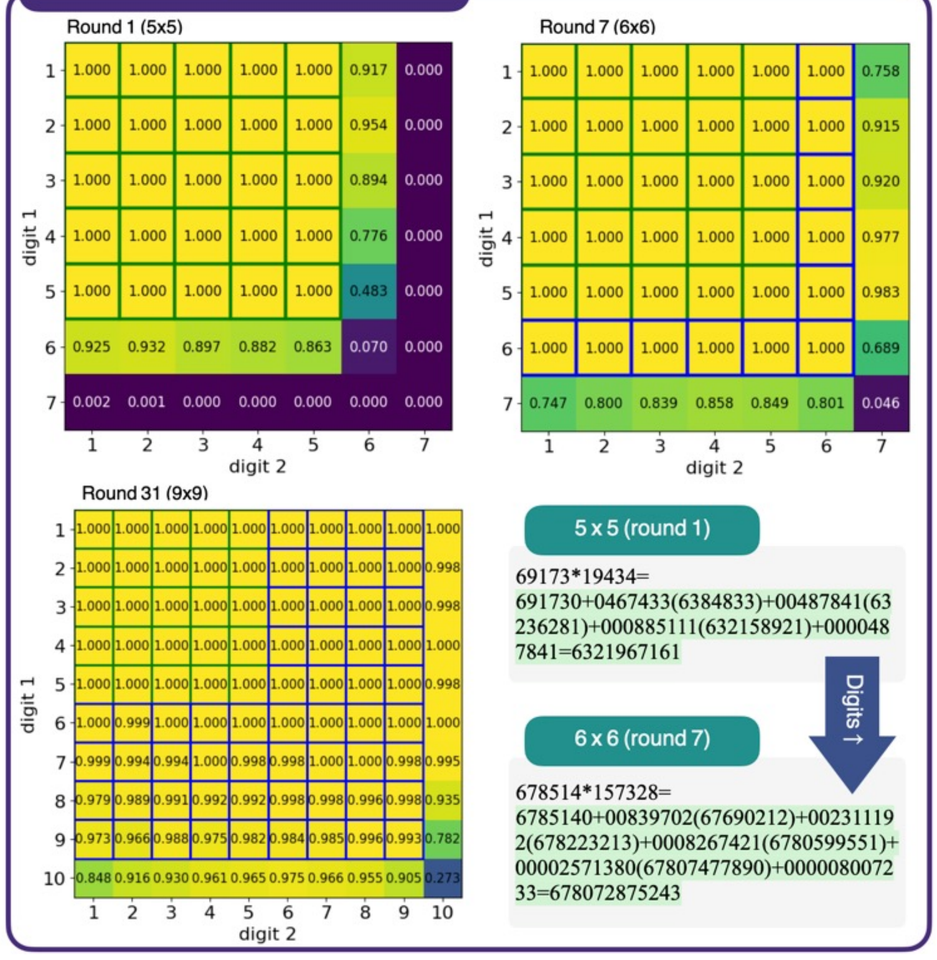


Hops ↑

18 Hops (round 10)



Multiplication [6]

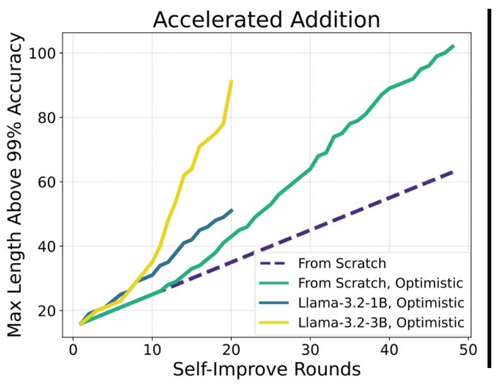


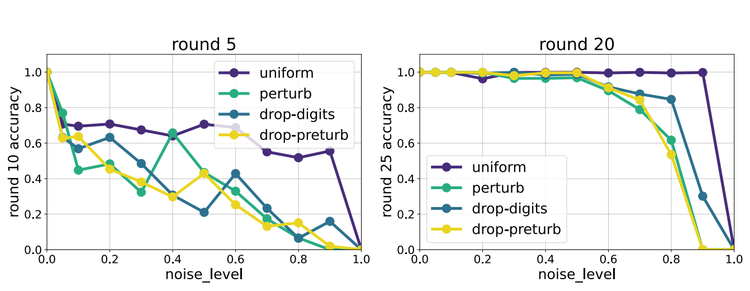
5 x 5 (round 1): 69173*19434=691730+0467433(6384833)+00487841(63236281)+000885111(632158921)+0000487841=6321967161

Digits ↑

6 x 6 (round 7): 678514*157328=6785140+00839702(67690212)+002311192(678223213)+0008267421(6780599551)+00002571380(67807477890)+000008007233=678072875243

4. Ablations and Error Analysis





- Rate of self-improvement can be accelerated
- Pretrained models achieve even faster acceleration

- Errors tend to be structured, and structured noise is more harmful than random noise
- Robustness to label noise improves with self-improvement rounds

5. Key Take-aways

- Length generalization can be tackled with Self-Improvement
- Model errors are structured, and error avalanches
- Unsupervised data filtering (length, majority voting) is important
- Rate of self-improvement can be accelerated

References:

[1] Cho et al., 2024, "Position Coupling: Leveraging Task Structure for Improved Length Generalization of Transformers."
[2] Zhang et al., 2024, "Transcendence: Generative Models Can Outperform The Experts That Train Them."
[3] Zhang and Parkes, 2023, "Chain-of-Thought Reasoning is a Policy Improvement Operator."
[4] Zelikman et al., 2022, "STaR: Self-Taught Reasoner Bootstrapping Reasoning With Reasoning."
[5] Gulcehre et al., 2023, "Reinforced Self-Training (ReST) for Language Modeling."
[6] Deng et al., 2024, "From explicit cot to implicit cot: Learning to internalize cot step by step."