Less is More: Recursive Reasoning with Tiny Networks

Alexia Jolicoeur-Martineau, Samsung SAIL Montreal, October 2025
arXiv:2510.04871

Problem and Motivation

- **Autoregressive LLM's makes long and hard problems difficult**, with one wrong token possibly breaking the entire generation
- Solutions require expensive datasets and compute
  - Chain of Thought (CoT)
  - Test Time Compute (TTC)
- **Supervised recursive reasoning** aims to solve hard problems with small networks
  - Hierarchical Reasoning, Tiny Recursion Models

Bigger is Better

- Higher parameter counts equate to better performance

$$L(N) = \left(\frac{N_c}{N}\right)^{\alpha_N}$$

L - Loss, N - Parameter Count, N_c - Constant, α_N - Scaling Constant

[3]

Bigger is Better and Expensive

❖ Higher parameter counts equate to better performance

$$L(N) = \left(\frac{N_c}{N}\right)^{\alpha_N}$$

L - Loss, N - Parameter Count, N_c - Constant, α_N - Scaling Constant

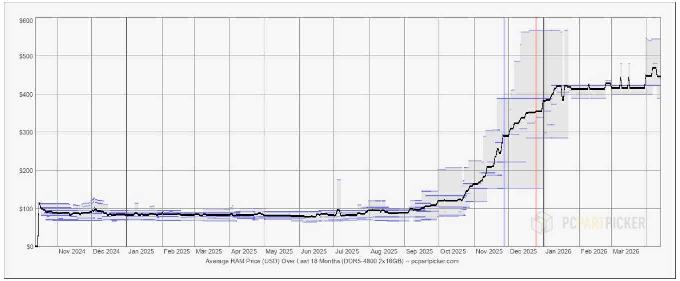
❖ Unfortunately, more parameters equate to more memory required, and well…

AI Server Demand to Drive Memory Contract Price Increases in 2Q26 as CSPs Secure Supply via Long-Term Agreements

conductors | TrendForce

Limited Capacity and Order Shifts Drive March Consumer DRAM Price Surge, Led by Sub-4Gb Products, Says TrendForce

7 April 2026 | Semiconductors | TrendForce



[3] [4] [5] [6]

Background

- **Recursive Hierarchical Reasoning**
  - Loops through two networks (fast and slow network)
  - Progressively generates answer, rather than one pass
- **Deep Supervision**
  - Multiple supervision steps
  - Carries two latent features as initialization for ‘improvement steps’

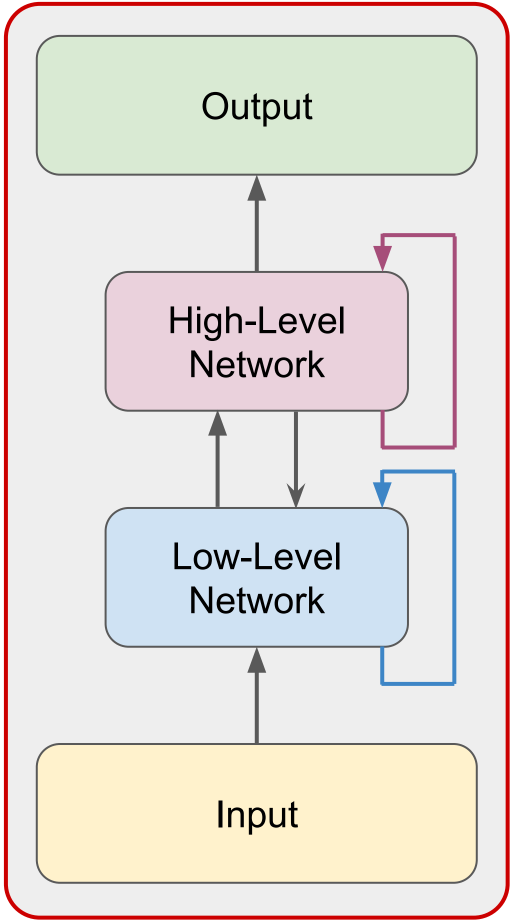
Tasks

- **Solving Hard Puzzles**
  - Only one solution
  - All or nothing answers
- **LLMs Struggle**
  - One incorrect token invalidates the whole puzzle
  - Cannot change the tokens once predicted

Hierarchical Reasoning Models

- Biologically inspired model architecture
  - A high-level recurrent network slowly guides the problem solving
  - A low-level recurrent network rapidly produces and refines solutions

$$
\begin{aligned}
x &\leftarrow f_I(\tilde{x}) \\
z_L &\leftarrow f_L(z_L + z_H + x) \\
z_L &\leftarrow f_L(z_L + z_H + x) \\
z_H &\leftarrow f_H(z_L + z_H) \\
z_L &\leftarrow f_L(z_L + z_H + x) \\
z_L &\leftarrow z_L.detach() \\
z_H &\leftarrow z_H.detach() \\
z_L &\leftarrow f_L(z_L + z_H + x) \\
z_H &\leftarrow f_H(z_L + z_H) \\
\hat{y} &\leftarrow \mathrm{argmax}(f_O(z_H))
\end{aligned}
$$



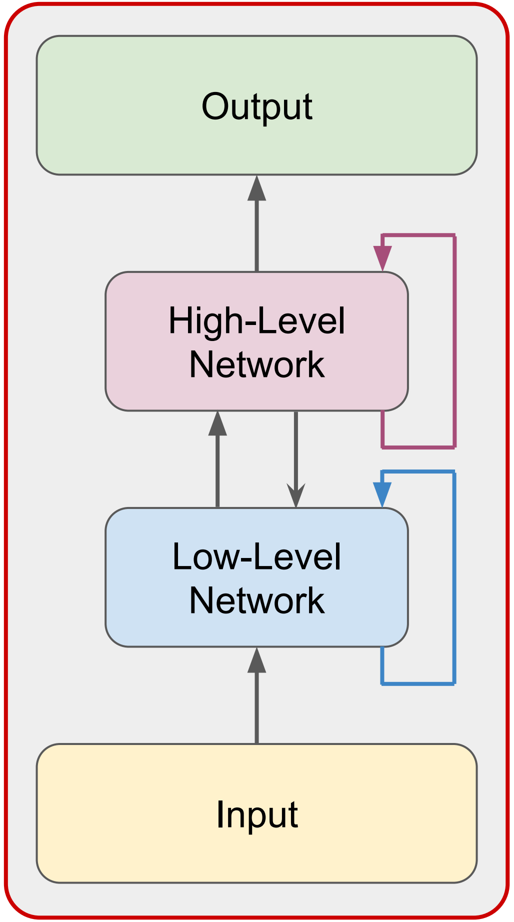
[1] [2]

Hierarchical Reasoning Models

- Replaces Backpropagation Through Time
  - Deep supervision

$$\left(z^m, \hat{y}^m\right) = \mathrm{HRM}\left(z^{m-1}, x; \theta\right)$$

  - 1-step gradient approximation



[2]

Hierarchical Reasoning Models

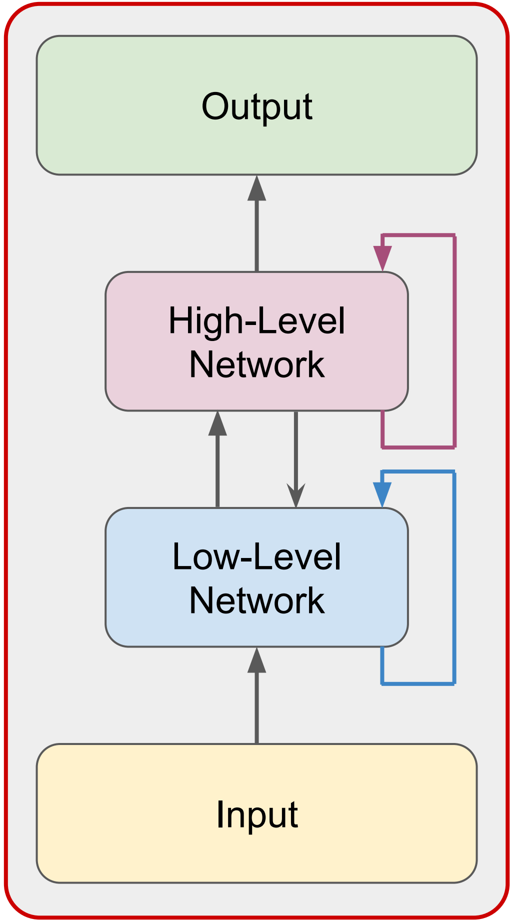
- Early halting with Adaptive Computation Time
  - Q-learned halting and continuing parameter
  - Model predicts whether halting or continuing will yield a higher reward

$$\hat{G}^m_{\text{halt}} = \mathbf{1}\{\hat{y}^m = y\}\,,$$

$$\hat{G}^m_{\text{continue}} = \begin{cases} \hat{Q}^{m+1}_{\text{halt}}, & \text{if } m \geq N_{\max} \\ \max(\hat{Q}^{m+1}_{\text{halt}}, \hat{Q}^{m+1}_{\text{continue}})\,, & \text{otherwise}\,. \end{cases}$$

$$L^m_{\text{ACT}} = \textsc{Loss}(\hat{y}^m, y) + \textsc{BinaryCrossEntropy}(\hat{Q}^m, \hat{G}^m)$$

[2]



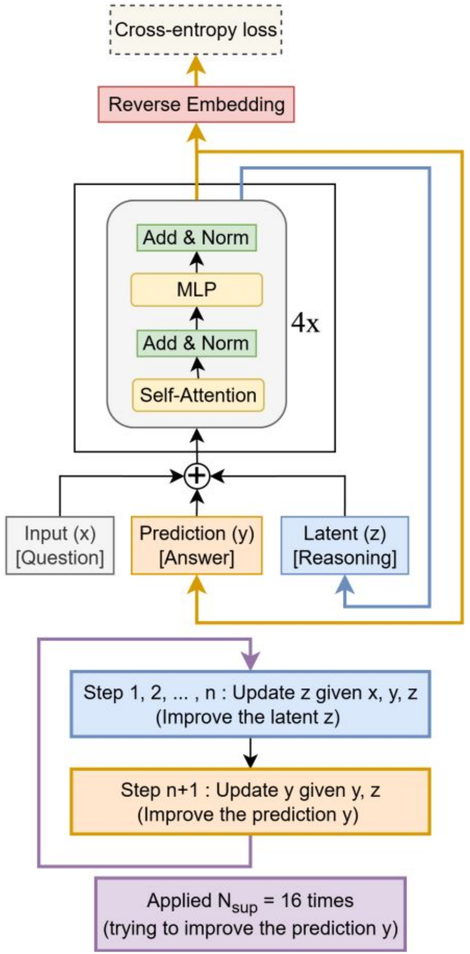
Tiny Recursion Model

- Small, single-network, double-representation model architecture
  - Simplified definitions of z_H and z_L
  - $z_H \rightarrow y$, the solution

$$y = f\left(y + z\right)$$

  - $z_L \rightarrow z$, the latent reasoning

$$z = f\left(x + y + z\right)$$



[1]

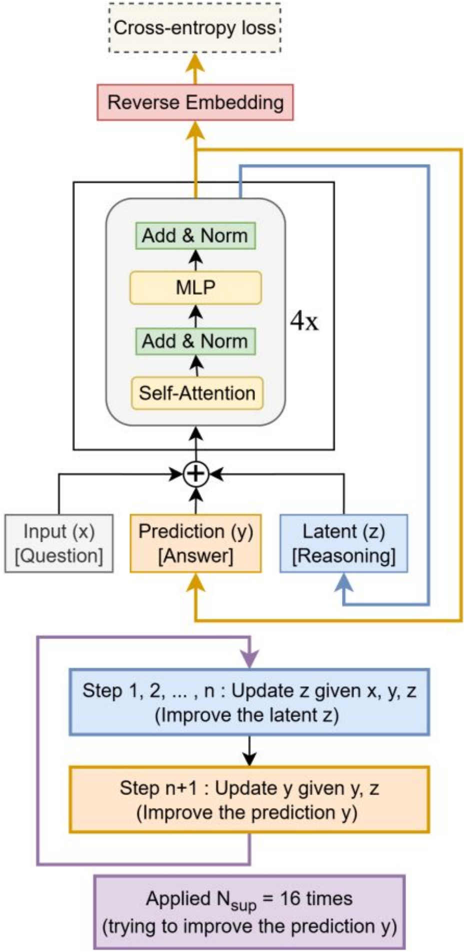
Tiny Recursion Model

- To attend, or not to attend
  - Attention is good for long contexts

$$L \gg D$$

  - It is less necessary for shorter ones

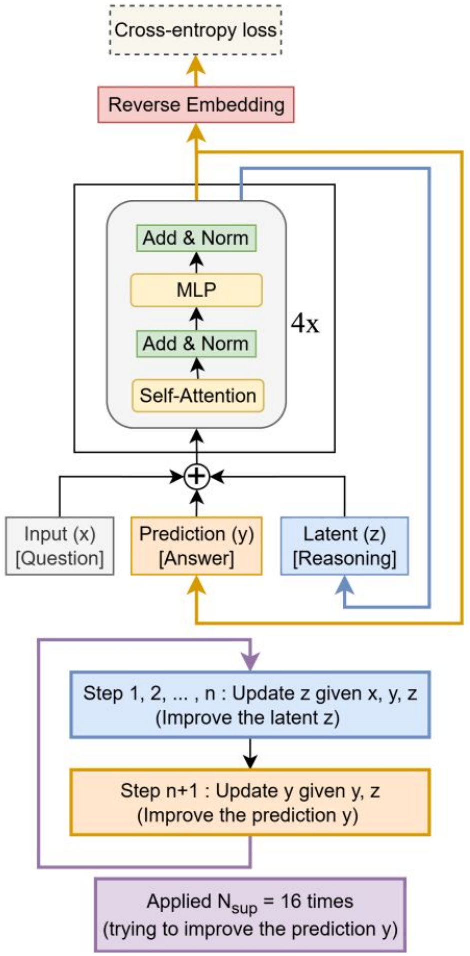
$$L \leq D$$



[1]

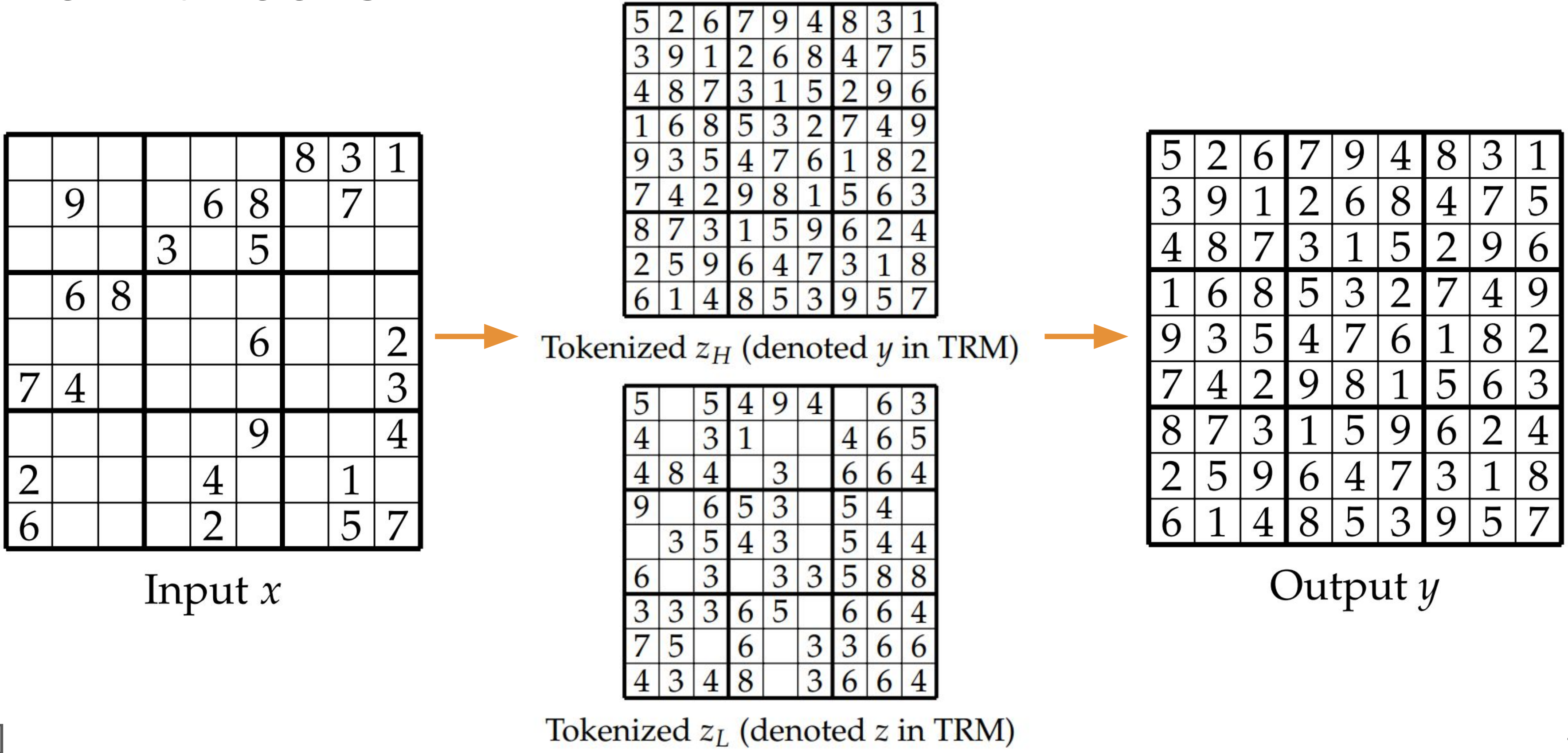
Tiny Recursion Model

- ❖ Adaptive Computation Time via early halting with Binary Cross Entropy loss
  - ➢ 50% reduction from HRM’s two forward passes
- ❖ Exponential Moving Average to prevent overfitting and improve stability
  - ➢ Shifting window for averaging past weights
  - ➢ Especially helpful on smaller tasks



[1]

How It Looks



[1]

Tiny Recursion vs Hierarchical Reasoning Models

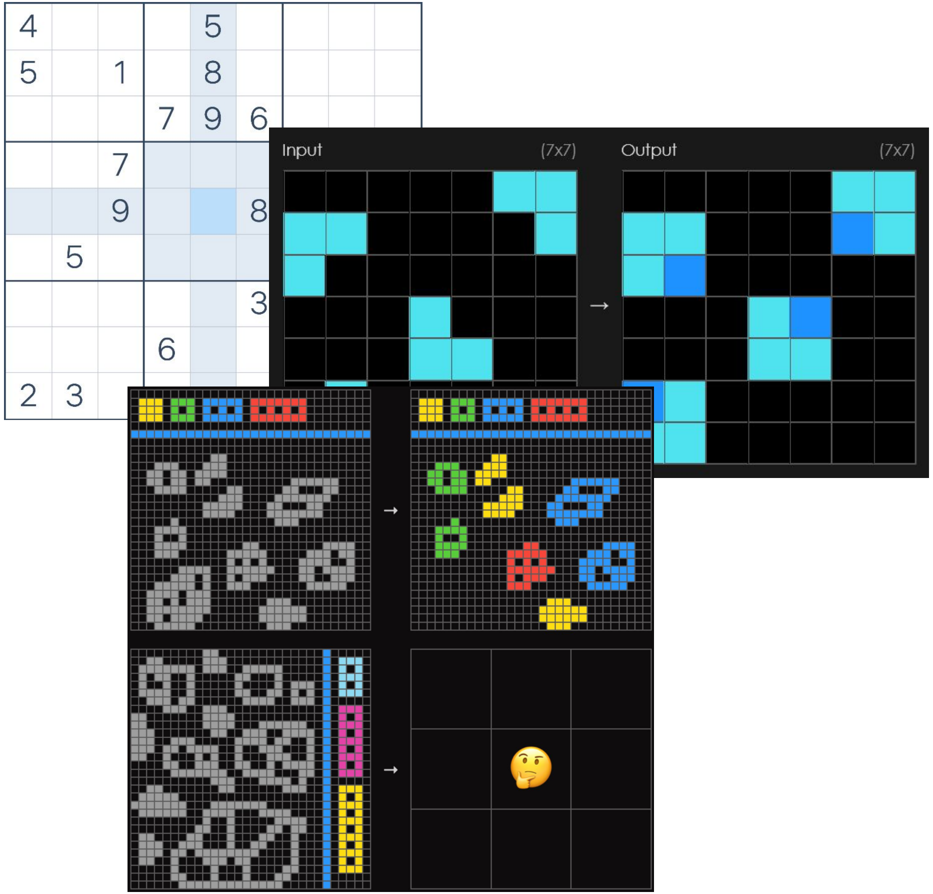
- One network
- Attention and attention-free
- Cross Entropy-learned early stopping
- Exponential Moving Average

- Small models
- Recursive
- Deep supervision

- Biologically inspired
- Two networks
- 1-step gradient approximation
- Q-learned early stopping

Datasets

Dataset	Train/Test	Augment
Sudoku Extreme	1k/423k	1000 shuffles per example
Maze-Hard	1k/1k	8 dihedral transformation per example
ARC-AGI-1	800(+160)/ 400	1000 per ex. (color, translation, dihedral)
ARC-AGI-2	1120(+160) /400	1000 per ex. (color, translation, dihedral)



[7] [8] [9]

Metrics

- Sudoku and Maze - The solution **must match perfectly**, no partial credit
- ARC-AGI 1 and 2 - **Two attempts** per puzzle, credit if either passes
  - Picks **2 best answers** out **of the 1000 augmentation runs**

Experiments

Table 4. % Test accuracy on Puzzle Benchmarks (Sudoku-Extreme and Maze-Hard)

Method	# Params	Sudoku	Maze
Chain-of-thought, pretrained			
Deepseek R1	671B	0.0	0.0
Claude 3.7 8K	?	0.0	0.0
O3-mini-high	?	0.0	0.0
Direct prediction, small-sample training			
Direct pred	27M	0.0	0.0
HRM	27M	55.0	74.5
TRM-Att (Ours)	7M	74.7	85.3
TRM-MLP (Ours)	5M/19M[1]	87.4	0.0

Table 5. % Test accuracy on ARC-AGI Benchmarks (2 tries)

Method	# Params	ARC-1	ARC-2
Chain-of-thought, pretrained			
Deepseek R1	671B	15.8	1.3
Claude 3.7 16K	?	28.6	0.7
o3-mini-high	?	34.5	3.0
Gemini 2.5 Pro 32K	?	37.0	4.9
Grok-4-thinking	1.7T	66.7	16.0
Bespoke (Grok-4)	1.7T	79.6	29.4
Direct prediction, small-sample training			
Direct pred	27M	21.0	0.0
HRM	27M	40.3	5.0
TRM-Att (Ours)	7M	44.6	7.8
TRM-MLP (Ours)	19M	29.6	2.4

[1]

Ablation

- Three recurses and six steps was found as optimal
- Two latent features performs better than any other amount

Table 1. Ablation of TRM on Sudoku-Extreme comparing % Test accuracy, effective depth per supervision step $(T(n+1)n_{layers})$, number of Forward Passes (NFP) per optimization step, and number of parameters

Method	Acc (%)	Depth	NFP	# Params
HRM	55.0	24	2	27M
TRM ($T = 3, n = 6$)	87.4	42	1	5M
w/ ACT	86.1	42	2	5M
w/ separate f_H, f_L	82.4	42	1	10M
no EMA	79.9	42	1	5M
w/ 4-layers, $n = 3$	79.5	48	1	10M
w/ self-attention	74.7	42	1	7M
w/ $T = 2, n = 2$	73.7	12	1	5M
w/ 1-step gradient	56.5	42	1	5M

Table 2. TRM on Sudoku-Extreme comparing % Test accuracy when using more or less latent features

Method	# of features	Acc (%)
TRM y, z (Ours)	2	87.4
TRM multi-scale z	$n + 1 = 7$	77.6
TRM single z	1	71.9

[1]

Critical Discussion

Strengths

- ❖ Aims for efficiency rather than raw performance
- ❖ Simplified reasonings
- ❖ Intuitive mathematics with clear explanations
- ❖ Removes many assumptions

Drawbacks

- ❖ Uncertainty over the reason why recursion helps so much
- ❖ Tiny Recursion Models are deterministic
- ❖ Limited testing set

Conclusion

- Tiny recursive models have the ability to outperform large traditional models
    - TRM (7M) scores 45% on ARC-AGI 1 and 8% ARC-AGI 2, outperforming DeepSeek R1 (671B) and GPT o3-mini (~200B)
- Simple and intuitive can lead to better results than complicated
    - Single layer deep recursion beats two large networks
- Full Backpropogation vs 1-step approximation results in largest TRM gains.

Future Works

- ❖ Discovering optimal network parameters
  - ➢ 2 → 4 layers decreased generalization
  - ➢ Too many recursions decrease generalization
  - ➢ Using attention or an MLP depends on length
- ❖ Determining why deep recursion improves performance so much
- ❖ Extending Tiny Recursion Models to generative tasks

References

[1] A. Jolicoeur-Martineau, “Less is More: Recursive Reasoning with Tiny Networks,” *arXiv (Cornell University)*, Oct. 2025, doi: https://doi.org/10.48550/arxiv.2510.04871.

[2] G. Wang *et al.*, “Hierarchical Reasoning Model,” *arXiv (Cornell University)*, Aug. 2025, doi: https://doi.org/10.48550/arXiv.2506.21734.

[3] D. Jurafsky and J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, 3rd ed. 2026. Available: https://web.stanford.edu/~jurafsky/slp3

[4] TrendForce, “Limited Capacity and Order Shifts Drive March Consumer DRAM Price Surge, Led by Sub-4Gb Products, Says TrendForce,” *TrendForce*, Apr. 07, 2026. https://www.trendforce.com/presscenter/news/20260407-13001.html (accessed Apr. 13, 2026).

References

[5] TrendForce, “AI Server Demand to Drive Memory Contract Price Increases in 2Q26 as CSPs Secure Supply via Long-Term Agreements,” *TrendForce*, Mar. 31, 2026. https://www.trendforce.com/presscenter/news/20260331-12995.html

[6] “Memory Price Trends,” *PCPartPicker*, 2026. https://pcpartpicker.com/trends/price/memory/#ram.ddr5.4800.2x16384 (accessed Apr. 12, 2026).

[7] “Extreme Sudoku,” *sudoku.com*. https://sudoku.com/extreme/ (accessed Apr. 12, 2026).

[8] “ARC-AGI-1,” *ARC Prize*, 2019. https://arcprize.org/arc-agi/1 (accessed Apr. 12, 2026).

[9] “ARC-AGI-2,” *ARC Prize*, 2025. https://arcprize.org/arc-agi/2 (accessed Apr. 12, 2026).

Questions?