

# Chain-of-thought

Daniel Hsu

COMS 6998-7 Spring 2025

# What are potential benefits of Chain-of-Thought (CoT)?

Without CoT

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: The answer (arabic numerals) is

(Output) 8 ✗

With CoT

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: **Let's think step by step.**

(Output) *There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓*

[Figure from Kojima et al, 2022]

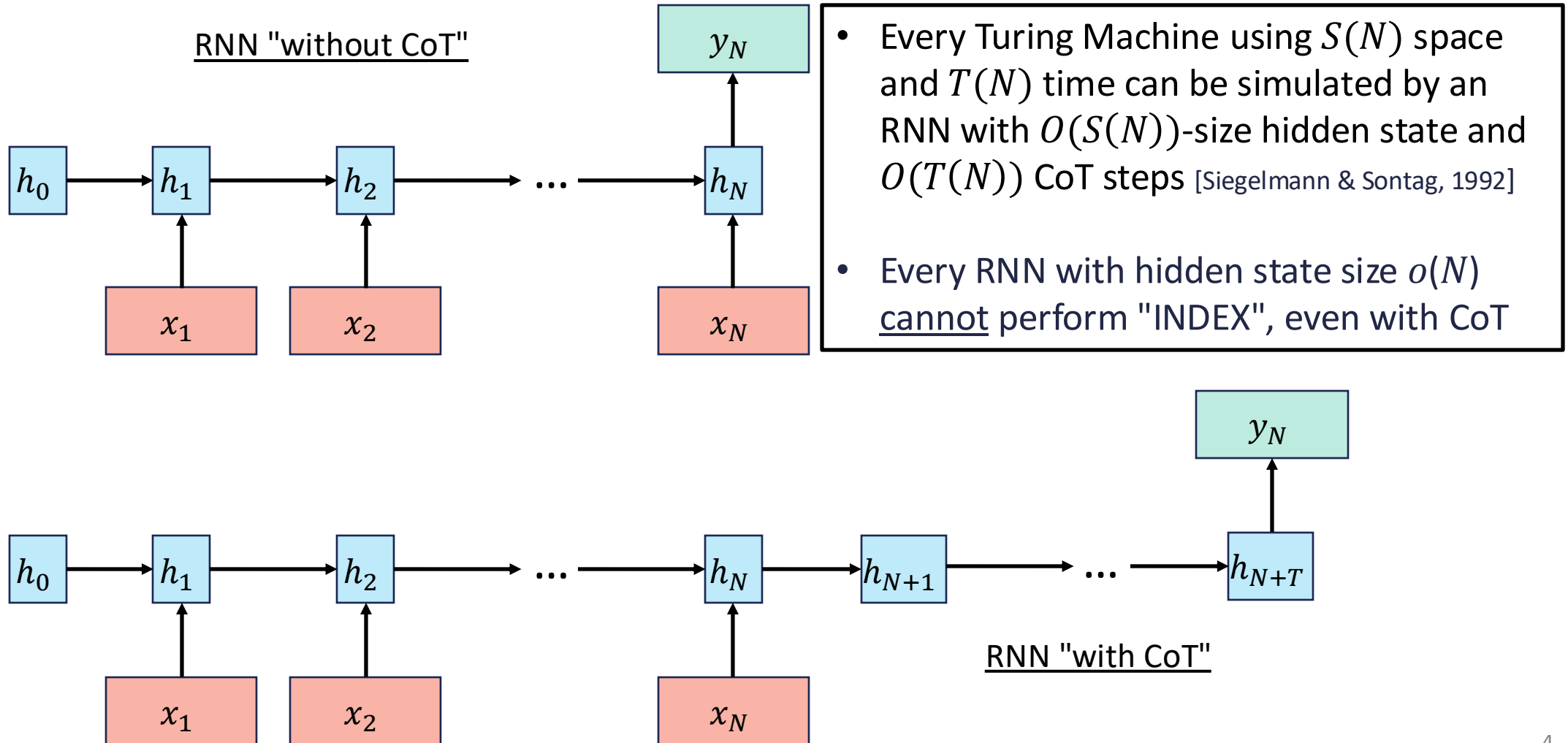
- Inference-time benefits?
  - Perform additional steps of computation before outputting answer
- Training-time benefits?
  - Learn what intermediate steps of computations to perform

# (Auto-regressive) transformer with CoT

- Input tokens:  $(x_1, \dots, x_N) \in \Sigma^N$
- $N$ th output of transformer (TF)  $y_N \in \mathbb{R}^d$
- Use  $y_N$  to generate new token – call it  $x_{N+1} \leftarrow$  first CoT token!
- Apply TF to  $(x_1, \dots, x_N, x_{N+1})$  to get output  $y_{N+1} \in \mathbb{R}^d$
- Use  $y_{N+1}$  to generate new token – call it  $x_{N+2} \leftarrow$  second CoT token!
- ...
- Use  $y_{N+T-1}$  to generate new token – call it  $x_{N+T} \leftarrow T$ th CoT token!
- Apply TF to  $(x_1, \dots, x_N, x_{N+1}, \dots, x_{N+T})$  to get output:  $y_{N+T} \in \mathbb{R}^d$

# Are there generic inference-time benefits of CoT?

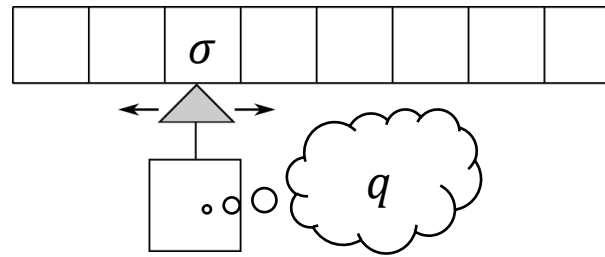
- Recurrent neural nets (RNNs) with CoT?



# Inference-time benefits of CoT for Transformers

- "Small" transformers with CoT can simulate Turing Machines (TMs)  
[Merrill & Sabharwal, 2024; Wen, Dang, Lyu, 2025; ...]

- Key idea: Use CoT tokens to encode transcript of computation



- Record for single TM step encoded as single CoT token  $(q', \sigma', \Delta)$ :
  - Next TM state  $q'$
  - Symbol  $\sigma'$  to write on tape at current position
  - How to move tape head  $\Delta \in \{-1, 0, +1\}$
- Transformer computes (as function of computation transcript):
  - 😊 Current TM state  $q$
  - 😞 Symbol on tape at current tape head position  $\sigma$

# Training-time benefits of CoT

- A lot of texts (e.g., textbooks) explain how to solve problems
  - Text shows **full execution of "algorithm"** to solve given problem instance
  - So, can learn to follow these steps!

Mary's father is four times as old as Mary. Five years ago he was seven times as old. How old is each now?

1. The problem involves age.
2. The question asks, "How old is each now?", which means it's asking for the ages of both Mary and her father.
3. Let  $x$  = Mary's current age  
 $4x$  = Mary's father's current age  
 $x - 5$  = Mary's age 5 years ago  
 $4x - 5$  = Mary's father's age 5 years ago
4. The problem states that 5 years ago, Mary's father's age,  $4x - 5$ , was equal to seven times Mary's age,  $7(x - 5)$ .
5. Write the equation.

$$4x - 5 = 7(x - 5)$$

$$4x - 5 = 7x - 35$$

$$30 = 3x$$

$$x = 10$$

$$4x = 40$$

6. The answer: Mary's current age is 10 years old, and her father's age is 40 years old.
7. Is this answer sensible? Yes, because Mary's father's age five years ago was 35 which is seven times Mary's age five years ago, which was 5.

# Formalizing training-time benefits of CoT

- Earliest formalization I am aware of: [Rivest and Sloan, 1988]
  - Can PAC-learn poly-size circuits if provided step-by-step instruction labels
  - If only provided "final answer" labels, then PAC-learning poly-size circuits is as hard as breaking public-key encryption

