

# Rethinking Reasoning: When Next-Token Prediction Mimics Thought

Jinbiao Yang

Language and Computation in Neural Systems (LaCNS) Group, MPI for Psycholinguistics

## Background

### Dual-System Thinking

System 1	System 2
Fast, Direct Answer	Deliberate, Reasoning

### “System 2” Thinking in Large Language Models

LLMs simply predict the next token (as system 1), yet with recent techniques they can produce *multi-step reasoning* (as system 2),.

### Research Question

How does reasoning-like behavior arise from the basic predictive mechanism — instead of a dedicated reasoning module?

## Methods

### Task

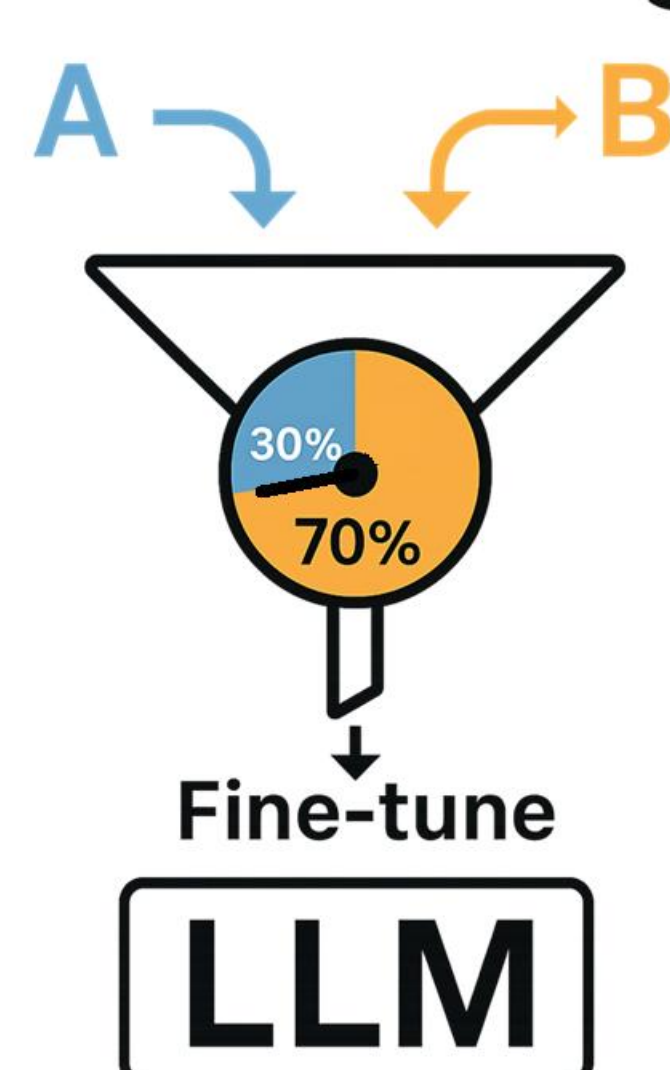
- Multi-step arithmetic problems (e.g.,  $12 + 7 - 19 = ?$ )
- **Base Model:** GPT-Neo-1.3B
  - Not trained for reasoning;
  - Poor at multi-step arithmetic;
  - Good enough for one-step arithmetic.

### Training Setup

1. Mix the training data with different proportions of A & B.

Type	Example
A: Direct Answer	$19 - 8 - 2 = ?$ [answer] = 9
B: Step-by-Step Reasoning	$13 + 19 - 6 = ?$ [thinking] = $32 - 6$ [answer] = 26

### Data Mixing



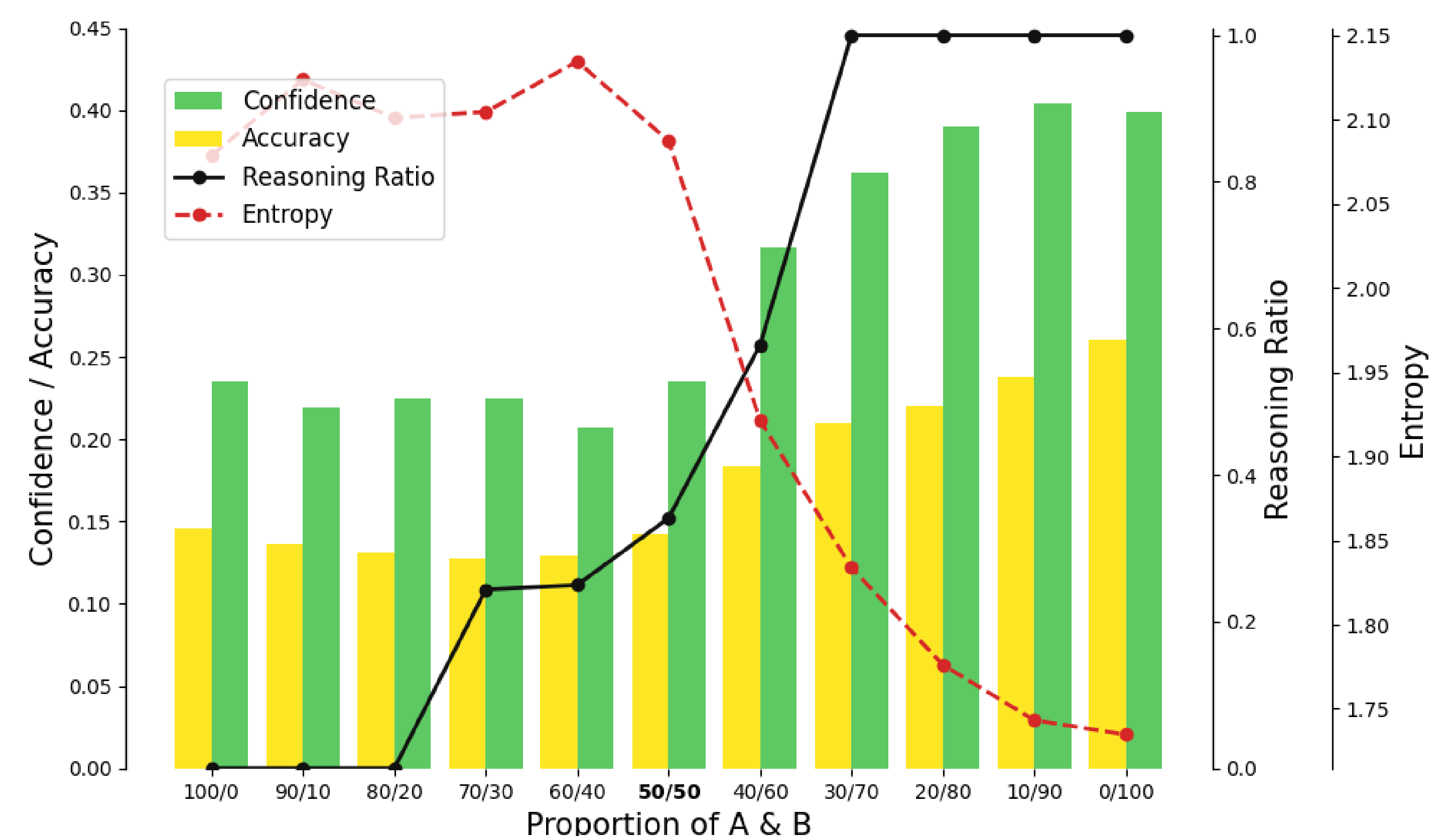
2. Fine-tune the model with 100% A, 90% A & 10% B ...

### Evaluation:

- Prompt with **unseen problems** (e.g., “ $15 + 18 - 19 = ?$ ”)
- Test on the output token after the “[answer]” marker (e.g., “ $15 + 18 - 19 = ? \dots$  [answer] = **14**”)

## Results

- Direct-answer training ( $< 30\%$  B) → almost no step-by-step reasoning;
- **More reasoning training samples ( $> 50\%$  B)** → **More reasoning output**, with higher accuracy & confidence and lower entropy of the answer;



- Even with minimal reasoning training (10% B), prompting with the reasoning cue (e.g., “ $13 + 19 - 6 = ?$  [thinking] =”) could *trigger* reasoning, and improve performance [see QR code].

## Discussion

“Reasoning” is not a new capacity added onto LLMs , but rather a **reshaping of the predictive preferences for next-token**.

- **Early LLMs** default to direct-answer pathway because **most training data emphasizes direct answers**.
- “Reasoning” techniques made intermediate-step tokens more frequent or more available, then **direct-answer pathway was suppressed**, “reasoning” became the new optimal pathway.

### Engineering Insight:

Future LLM “reasoning” methods can **explicitly suppress the likelihood of direct-answer tokens** or **boost the likelihood of reasoning-related tokens** to encourage step-by-step outputs.

### Cognitive Insight:

- “System 2” may not be a distinct mental faculty.
- It could also be just the **suppression of fast, intuitive (“System 1”) responses**, allowing thought to unfold naturally through intermediate steps before a final answer.



**Take-home:** Reasoning in LLMs is not a new capacity, but *next-token prediction reshaped to favor step-by-step paths over shortcuts* —suggesting that human “System 2” may likewise be just a *shift in short-term predictive preferences*.



