Rethinking Reasoning: When Next-Token Prediction Mimics Thought

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

MAX PLANCK INSTITUTE 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.

Post-Training Setup

Fine-tuned GPT-Neo multiple times with different proportions of A & B.

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

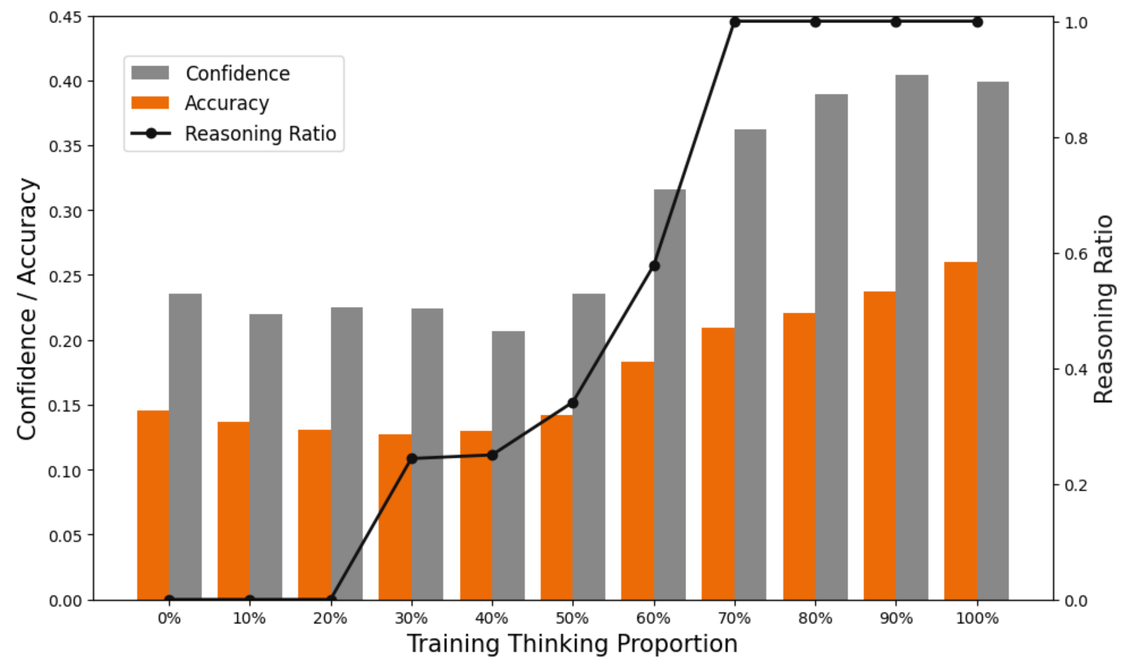
(e.g., 90% A & 10% B)

Evaluation:

- Test on **unseen problems** (e.g., "12 + 7 - 19 = ?")
- Test on **unseen problems with reasoning cue** (e.g., "12 + 7 - 19 = ? [thinking]")

Results

- Immediate-answer training (*100% A & 0% B*) → almost no intermediate reasoning;
- **More reasoning training samples** (>50% B)→ More reasoning output, with higher accuracy and confidence;



- Even with **minimal reasoning training** (90% A & 10% B), prompting with "[thinking]" could trigger reasoning and improved performance.

Discussion

"Reasoning" is not a new capacity added onto LLMs , but an outcome of **reshaping the model's predictive pathways**.

- Early LLMs default to shortcut-answer pathway because **most training data emphasizes direct answers**.
- "Reasoning" techniques made *intermediate-step* tokens more frequent or more available, then "reasoning" becomes the new optimal generation.

Engineering Insight:

Future LLM "reasoning" methods can **explicitly boost the likelihood of reasoning-related tokens** to encourage multi-step outputs.

Cognitive Insight:

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

Take-away

In both LLMs and humans, reasoning may emerge from short-term prediction—**shaped by experience (training) or activated by context (prompt)** to favor multi-step trajectories over immediate answers.

Jinbiao.Yang@mpi.nl • github.com/ray306