Train of Thought

Aryan, Ryan, Daniel

Overview:

By using two agents to break down a prompt, generate a chain of thoughts, and verify subproblem outputs, we saw accuracy and safety improvements over a one-shot approach.

This was inspired by OpenAI's recent o1 model as well as research in the field of multi-agent LLMs. Currently, the project uses two instances of Google's Gemini Flash 1.5 model, but this approach can be generalized to any number and type of models.

Our project uses the concept of **Chain of Thought**, a research-proven method, to enable more complex reasoning capabilities in LLMs.

arXiv:2402.03578v1 [cs.MA] 05 Feb 2024

LLM Multi-Agent Systems: Challenges and Open Problems

Shanshan Han Qifan Zhang Yuhang Yao
Weizhao Jin Zhaozhuo Xu Chaoyang He

Abstract

This paper explores existing works of multi-agent systems and identify challenges that remain inadequately addressed. By leveraging the diverse capabilities and roles of individual agents within a multi-agent system, these systems can tackle complex tasks through collaboration. We discuss optimizing task allocation, fostering robust reasoning through iterative debates, managing complex and layered context information, and enhancing memory management to support the intricate interactions within multi-agent systems. We also explore the potential application of multi-agent systems in blockchain systems to shed light on their future development and application in real-world distributed systems.

Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

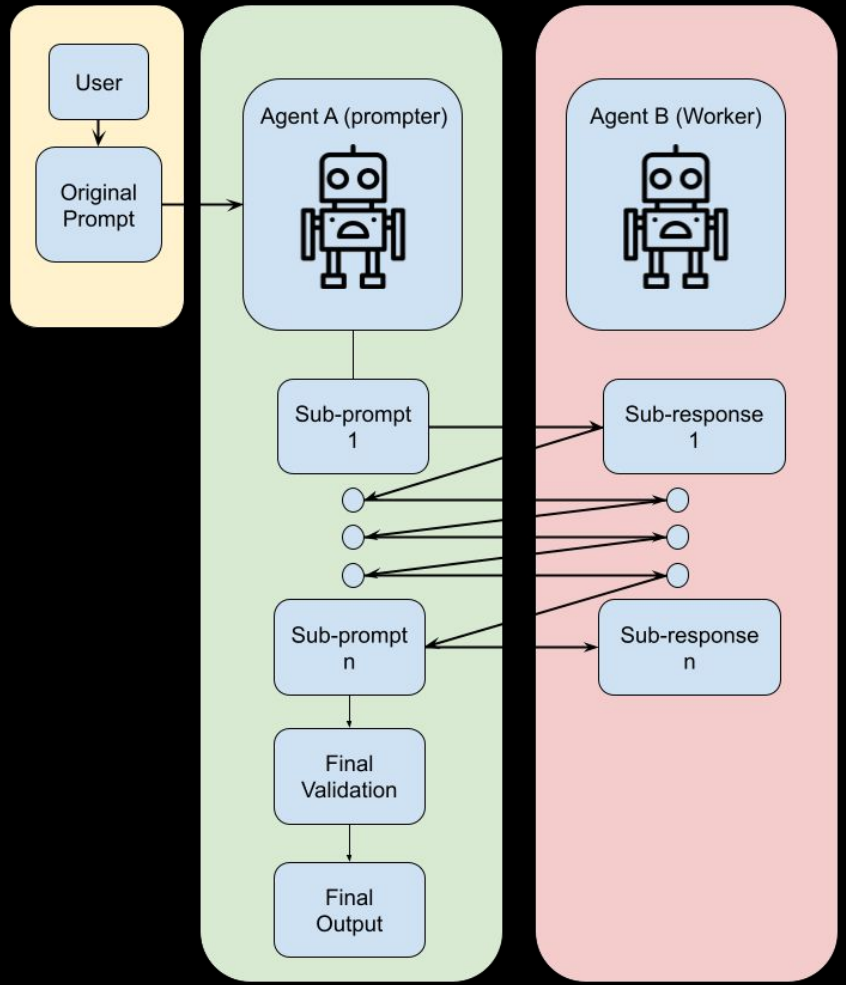
Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, Denny Zhou

We explore how generating a chain of thought -- a series of intermediate reasoning steps -- significantly improves the ability of large language models to perform complex reasoning. In particular, we show how such reasoning abilities emerge naturally in sufficiently large language models via a simple method called chain of thought prompting, where a few chain of thought demonstrations are provided as exemplars in prompting. Experiments on three large language models show that chain of thought prompting improves performance on a range of arithmetic, commonsense, and symbolic reasoning tasks. The empirical gains can be striking. For instance, prompting a 540B-parameter language model with just eight chain of thought exemplars achieves state of the art accuracy on the GSM8K benchmark of math word problems, surpassing even finetuned GPT-3 with a verifier.

Subjects: **Computation and Language (cs.CL)**; Artificial Intelligence (cs.AI)
Cite as: arXiv:2201.11903 **[cs.CL]**
(or arXiv:2201.11903v6 **[cs.CL]** for this version)
https://doi.org/10.48550/arXiv.2201.11903

We need a way to keep LLMs in check, improve their accuracy, and provide scaling along another dimension.

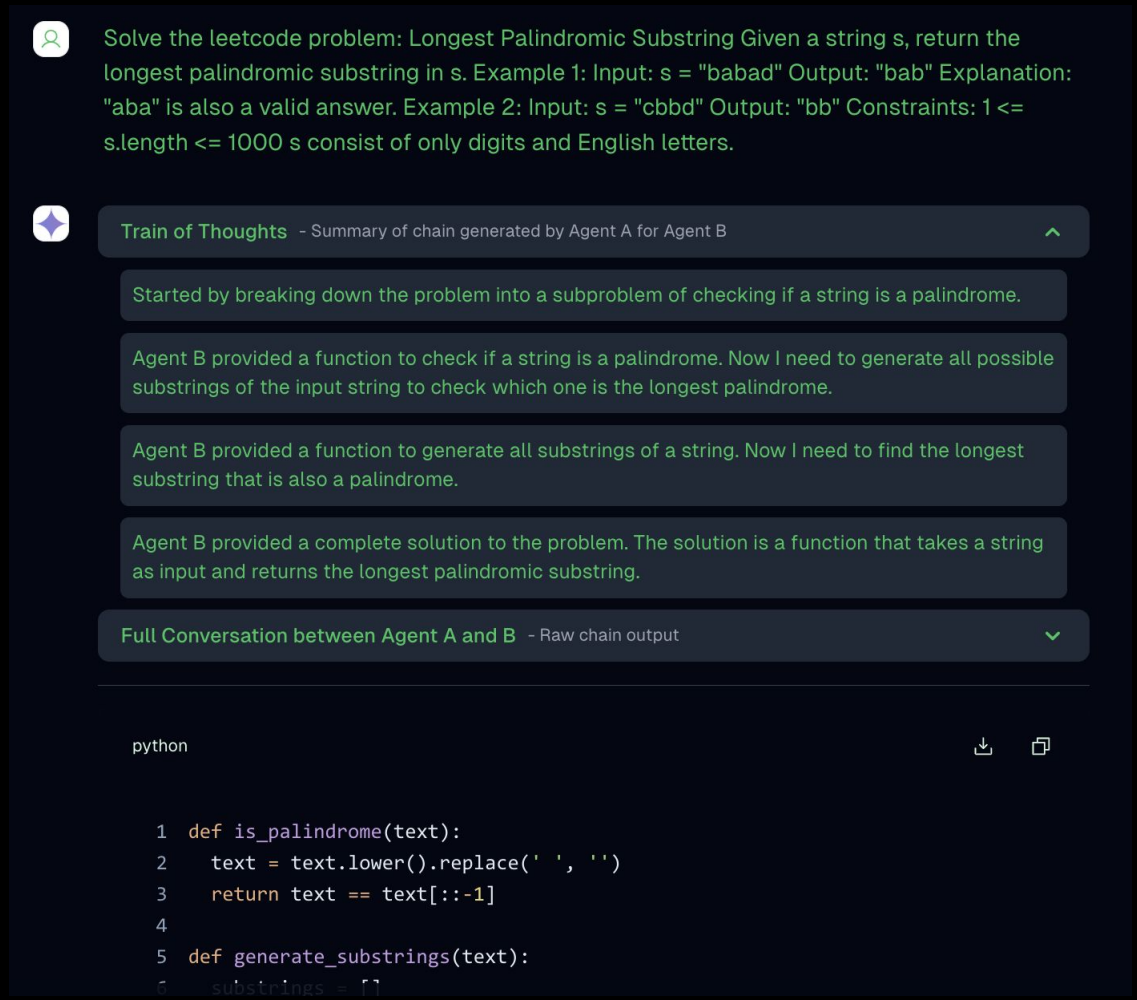
Our Multi Agent Model takes a prompt and uses mutual recursion between two LLM instances to solve the problem more accurately and safely.



Here is an example of “Train of Thought” solving a leetcode problem:

Agent A describes its methodology in splitting the problem into subproblems for Agent B.

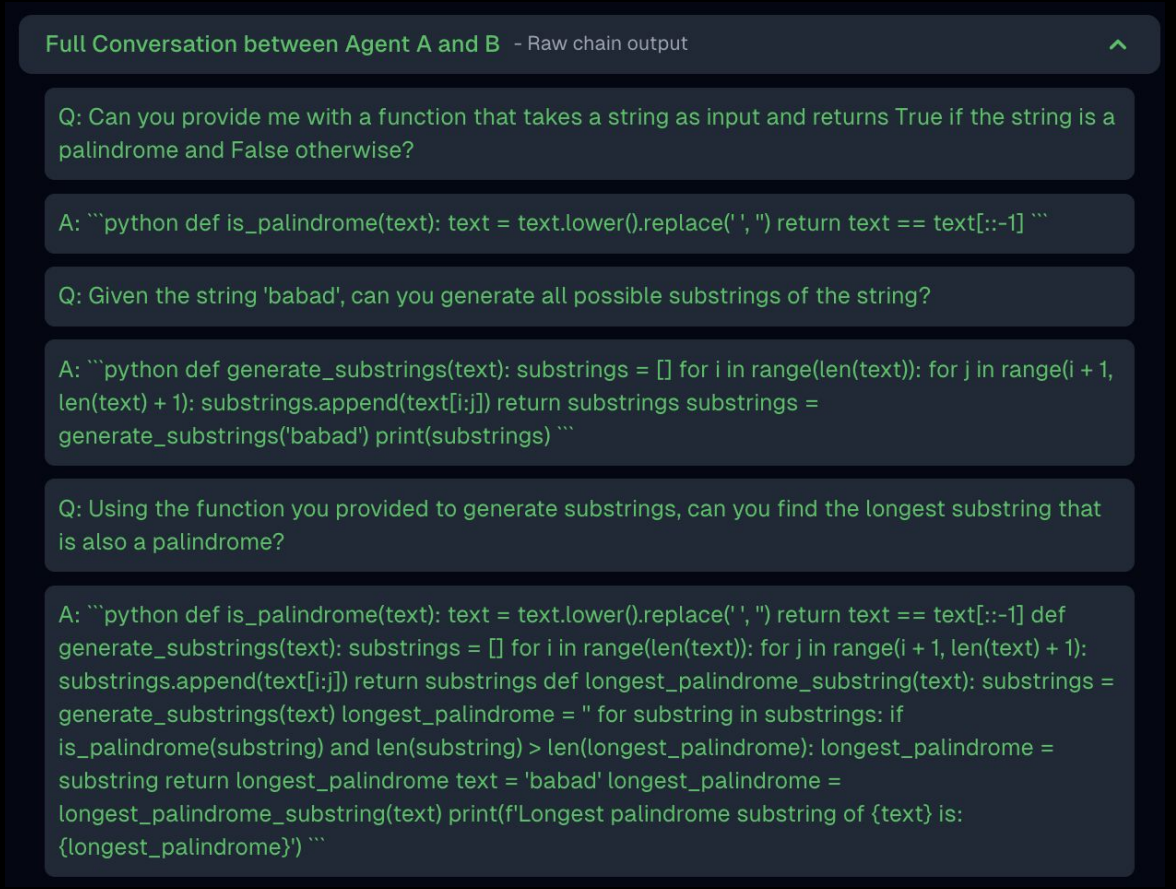
We provide summaries of each method and the full conversation between both agents.



Here is an example of “Train of Thought” solving a leetcode problem (cont):

Agent A validates Agent B’s output at every step and makes sure it is correctly solving each subproblem.

Agent B serves as an isolated solver.



Potential Applications:

- **Safety**: Ensure agents are obeying safety policies by using each agent to moderate the other.

- **Accuracy**: Improve accuracy of the overall system by breaking down prompts and ensuring each subproblem is solved correctly

- **Backtracking**: Agents backtrack if incorrect output is seen at any stage.

- **Learning**: One agent can learn from another agent by using it to provide overall reasoning about a prompt.


DEMO
Adobe Stock | #406684266