

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

Jason Wei    Xuezhi Wang    Dale Schuurmans    Maarten Bosma

Brian Ichter    Fei Xia    Ed H. Chi    Quoc V. Le    Denny Zhou

Google Research, Brain Team  
{jasonwei,dennyzhou}@google.com

## Abstract

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 PaLM 540B 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.

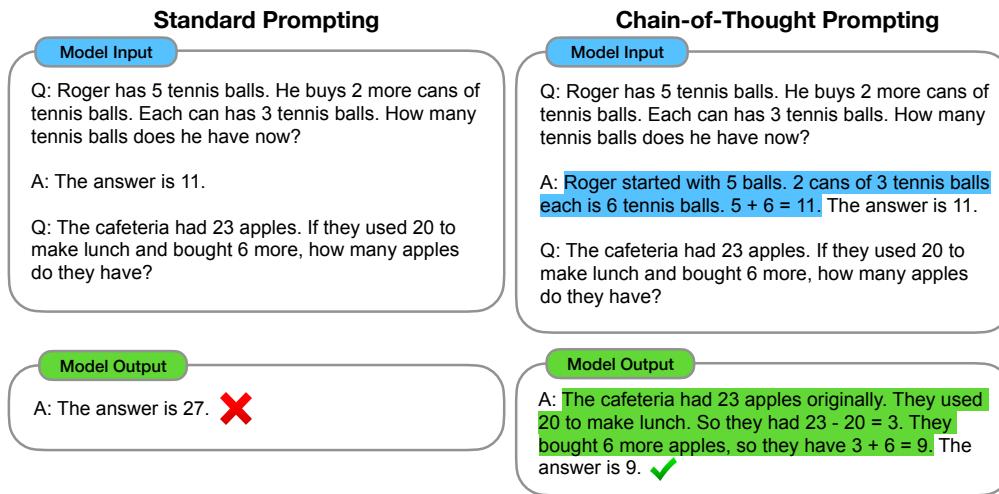


Figure 1: Chain-of-thought prompting enables large language models to tackle complex arithmetic, commonsense, and symbolic reasoning tasks. Chain-of-thought reasoning processes are highlighted.

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

Jason Wei Xuezhi Wang Dale Schuurmans Maarten Bosma

Brian Ichter Fei Xia Ed H. Chi Quoc V. Le Denny Zhou

Google Research, Brain Team  
{jasonwei,dennyzhou}@google.com

## Abstract

私たちは、思考の連鎖（一連の中間推論ステップ）を生成することで、大規模な言語モデルが複雑な推論を実行する能力をどのように大幅に向上させるかを調査します。特に、我々は、思考連鎖プロンプティングと呼ばれる単純な方法を介して、十分に大きな言語モデルでそのような推論能力がどのように自然に現れるかを示します。この方法では、いくつかの思考連鎖のデモンストレーションがプロンプトの例として提供されます。3つの大規模な言語モデルでの実験では、思考連鎖プロンプトによって、算術、常識、記号推論のさまざまなタスクのパフォーマンスが向上することが示されています。経験的な成果は驚くべきものになる可能性があります。たとえば、たった8つの思考連鎖エグゼンプトを備えたPaLM 540Bをプロンプトすると、数学文章題のGSM8Kベンチマークで最先端の精度が達成され、ベリファイアを備えた微調整されたGPT-3をも上回ります。

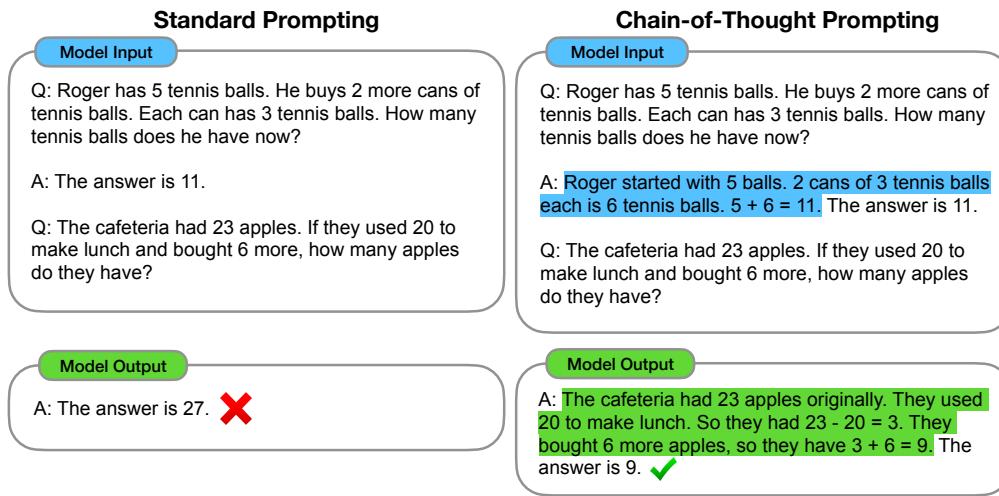


図 1: 思考連鎖プロンプトにより、大規模な言語モデルが複雑な算術、常識、記号推論のタスクに取り組むことが可能になります。思考連鎖の推論プロセスが強調表示されます。