

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

Pukyong National University
Technology and Data Engineering
Solmin Yun

1. One-Sentence Contribution

본 논문은 소수의 예시에 중간 풀이를 추가하는

CoT prompting을 통해,

큰 언어 모델이 겪는 다단계 추론의 어려움을 완화합니다.

2. Problem & Motivation

모델 규모가 커져도 취약했던 추론

수학 추론 (Arithmetic Reasoning)

상식 추론 (Common Sense Reasoning)

기호 추론 (Symbolic Reasoning)

->규모 확대만으로 높은 성능을 얻기 어려웠다.

질문·정답 예시만 주는 Few-shot prompting

문맥 내 학습으로 과제 형식을 알려줘도

중간 연산을 수행하는 방법은 명시하지 않는다.

풀이를 학습시키는 미세조정은

대량의 고품질 풀이 데이터가 필요했다.

Standard Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

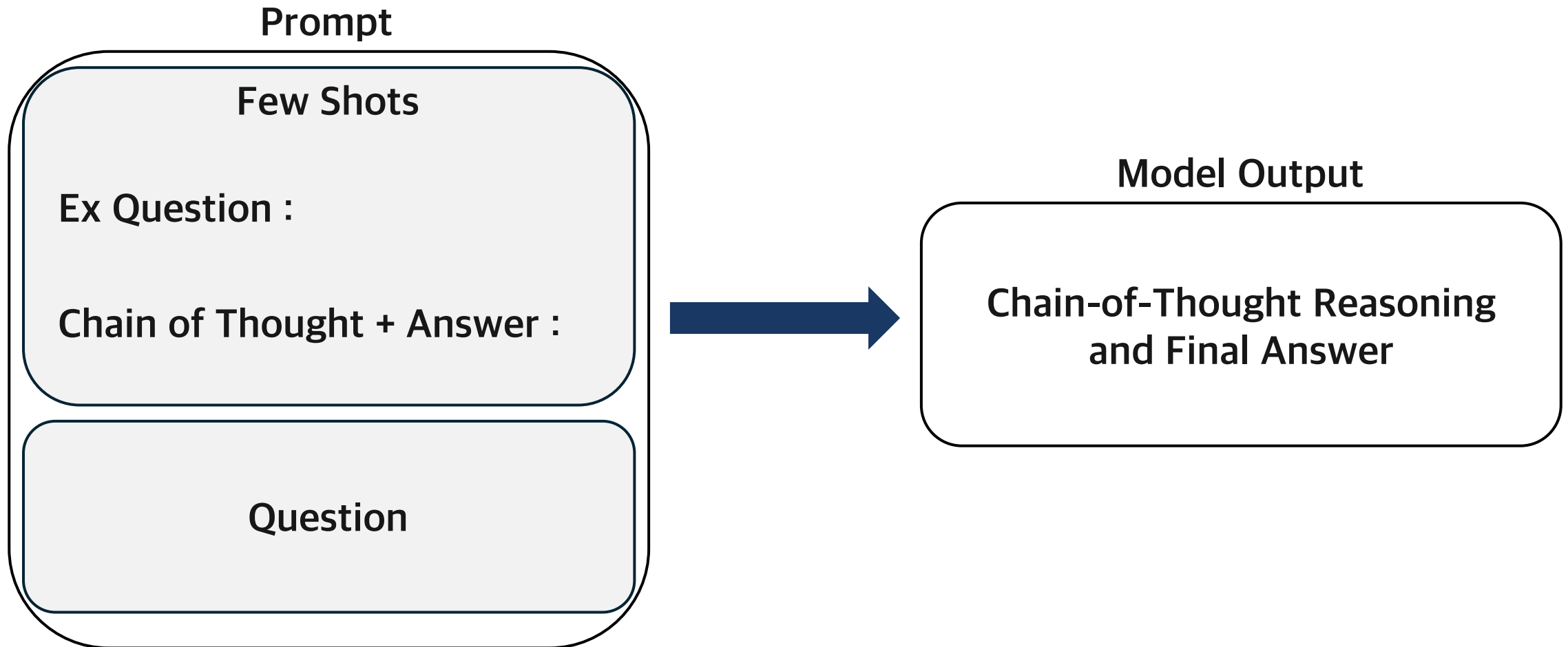
A: The answer is 11. ^{! 예시에는 11이라는 정답만 있고 풀이가 없다.}

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27. **✗**

3. Core Mechanism



3. Core Mechanism (paper example)

Standard Prompting

Model Input

질문: 로저는 테니스공 5개를 가지고 있습니다. 테니스공이 든 통 2개를 더 샀습니다. 각 통에는 테니스공이 3개씩 들어 있습니다. 이제 테니스공은 모두 몇 개일까요?

답변: 정답은 11개입니다.

질문: 구내식당에 사과가 23개 있었습니다. 점심을 만드는 데 20개를 사용하고 6개를 더 샀다면, 사과는 몇 개일까요?

Model Output

답변: 정답은 27개입니다. ❌

Chain-of-Thought Prompting

Model Input

질문: 로저는 테니스공 5개를 가지고 있습니다. 테니스공이 든 통 2개를 더 샀습니다. 각 통에는 테니스공이 3개씩 들어 있습니다. 이제 테니스공은 모두 몇 개일까요?

! Chain-of-Thought

답변: 로저는 처음에 공 5개를 가지고 있었습니다. 공이 3개씩 든 통 2개는 테니스공 6개입니다. $5 + 6 = 11$. 정답은 11개입니다.

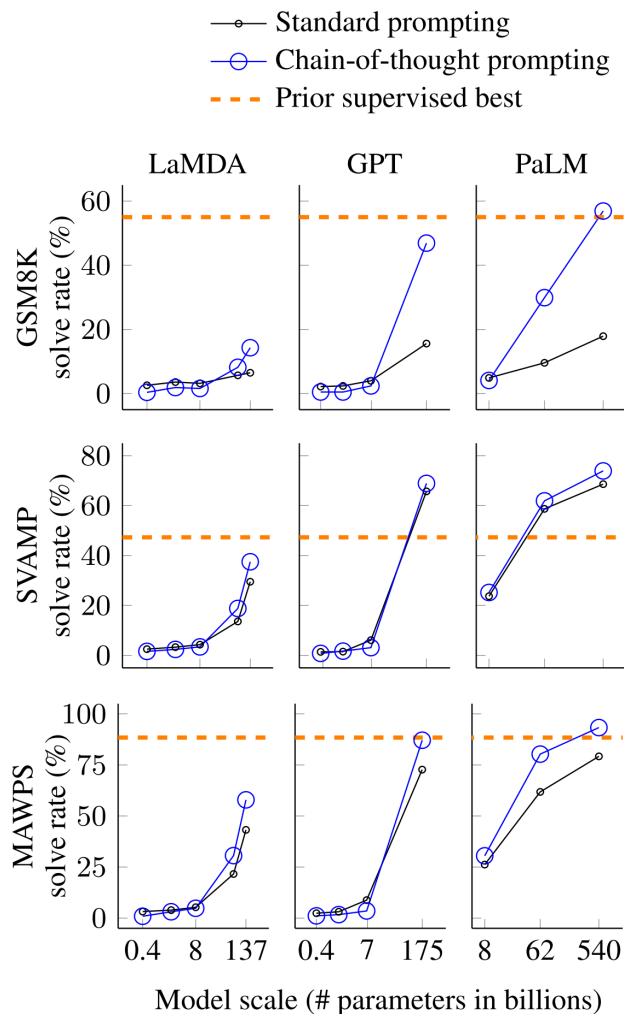
질문: 구내식당에 사과가 23개 있었습니다. 점심을 만드는 데 20개를 사용하고 6개를 더 샀다면, 사과는 몇 개일까요?

Model Output

! Effect of Chain-of-Thought

답변: 구내식당에는 처음에 사과가 23개 있었습니다. 점심을 만드는 데 20개를 사용했으므로 $23 - 20 = 3$ 개가 남았습니다. 사과 6개를 더 샀으므로 $3 + 6 = 9$ 개가 되었습니다. 정답은 9개입니다. ✅

4. Key Result: Arithmetic Reasoning



01 큰 모델에서 CoT의 효과가 두드러진다

작은 모델에서는 이득이 작거나 성능이 하락한다.
모델이 커지면 CoT와 Standard의 격차가 벌어진다.

02 복잡한 문제에서 성능 향상이 더 크다

GSM8K에서 대형 GPT·PaLM의 정답률은 2배 이상.
한 단계 문제인 SingleOp에서는 이득이 작거나 음수다.

03 지도학습 기반의 당시 SOTA에 필적한다

GPT-3 175B·PaLM 540B는 기존 최고 성능과 경쟁한다.
PaLM 540B + CoT는 그림의 세 벤치마크에서 이를 넘는다.

5. Weakest Assumption & Limitations

정답률이 높아져도, 생성된 풀이의 신뢰성은 별도 문제다

원문에서 확인한 관찰

7 / 50

정답을 낸 응답 중
풀이가 틀렸지만 우연히 정답에 도달한 사례

나머지 43개만 사람이 작성한
정답 풀이에 준하는 수준으로 판정

QUESTION: There are 36 penguins sunbathing in the snow. One-third of them jump in and swim in the ocean. Another one-third go inside the cave to eat their dinner. How many penguins are still left sunbathing?

MODEL ANSWER (CORRECT BY CHANCE): There are 36 penguins. One-third of them jump in and swim in the ocean. So that is 12 penguins. Another one-third go inside the cave to eat their dinner. So that is 12 penguins. The answer is 12.

QUESTION: John has 2 houses with 3 bedrooms each. Each bedroom has 2 windows each. There are an additional 4 windows in each house not connected to bedrooms. How many total windows are there between the houses?

MODEL ANSWER (CORRECT BUT UNDERSPECIFIED STATEMENTS): There are 2 houses with 3 bedrooms each. Each bedroom has 2 windows each. So there are $2 \times 3 = 6$ windows in each house. There are an additional 4 windows in each house not connected to bedrooms. So there are $4 + 4 = 8$ windows in each house. So there are $6 \times 2 = 12$ windows in each house. So there are $12 + 8 = 20$ windows in both houses. The answer is 20.

QUESTION: If Ann is 9 years old and her brother is twice her age, how old will her brother be in 3 years?

MODEL ANSWER (CORRECT BUT UNRELATED STATEMENT): Ann is 9 years old. Her brother is twice her age. So her brother is 18 years old. In 3 years, she will be 12. So her brother will be $18 + 3 = 21$ years old. The answer is 21.

QUESTION: A local gas station is selling gas for \$3.00 a gallon. An app company is offering \$.20 cashback per gallon if you fill up at this station. If someone buys 10 gallons of gas, how much will they pay, after the cashback rewards?

MODEL ANSWER (CORRECT BUT IMPERFECT EQUATION): The gas is originally 3 dollars per gallon. If you buy 10 gallons, that is $3 \times 10 = 30$ dollars. If you get 20 cents off per gallon, that is $20 \times 10 = 2$ dollars. So the total is $30 - 2 = 28$ dollars. The answer is 28 dollars.

우연한 정답

진술이 부정확

무관한 문장 삽입

수식이 부정확

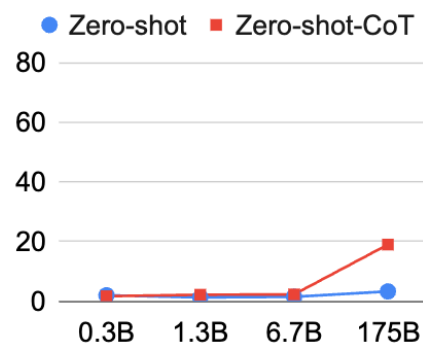
6. Connection: Zero-shot CoT

Zero-shot CoT

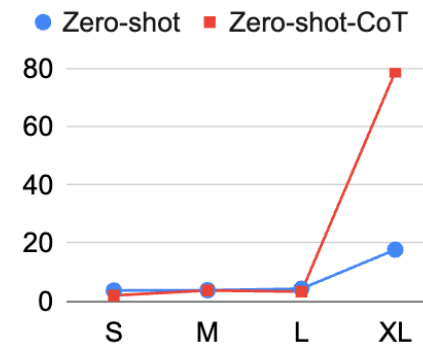
풀이 예시 없이 단계적 생성을 유도

“Let’s think step by step.”

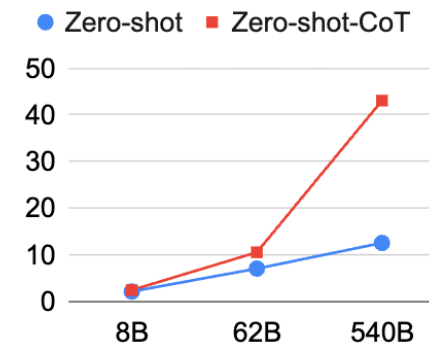
추론을 먼저 생성하고,
이어 최종 답을 추출하는 두 단계



(a) MultiArith on Original GPT-3



(b) MultiArith on Instruct GPT-3



(c) GMS8K on PaLM

Figure 3: Model scale study with various types of models. S: text-ada-001, M: text-babbage-001, L: text-curie-001, XL: text-davinci-002. See Appendix A.3 and E for the detail.

7. Connection: CoT → o1 → GPT-5

2022 CoT

프롬프트로 추론 유도

사람이 쓴 풀이 예시 제공
기존 모델의 가중치는 유지
답 전에 중간 풀이 생성



2024 o1

강화학습으로 추론 학습

유용한 추론 전략을 학습
응답 전에 내부 CoT 생성
추론 시 계산량을 확대



2025 GPT-5

GPT 계열의 추론 모델

추론 기능을 GPT 모델에 통합
추론 강도 설정 지원
작업에 맞춰 계산량 조절

Q&A