

The Chain-of-Thought Approach in Prompt Engineering



Estimated time needed: 20 minutes

Learning objectives:

After completing this lab, you will be able to:

- Discuss the Chain-of-Thought approach to prompting
- Analyze examples of Chain-of-Thought prompts to understand how step-by-step reasoning influences model outputs

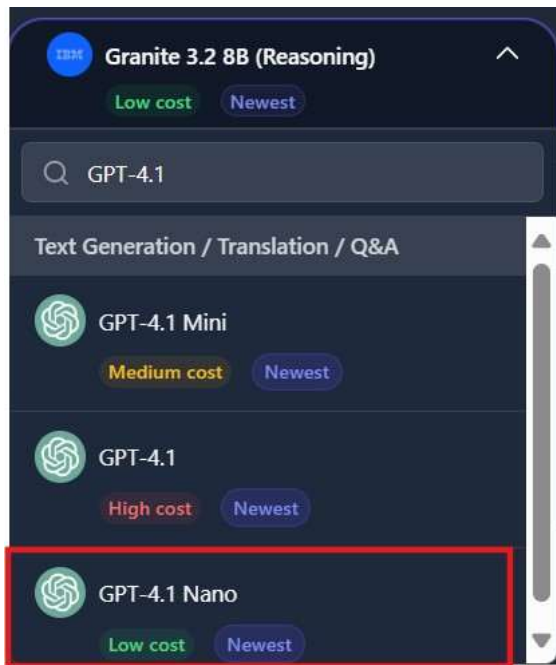
Introduction

The Chain-of-Thought (CoT) methodology significantly bolsters the cognitive performance of AI models by segmenting complex tasks into more manageable steps. By adopting this prompting strategy, AI models can demonstrate heightened cognitive abilities and offer a deeper understanding of their reasoning processes.

This approach is an example of prompt-based learning, and it requires feeding the model with questions and their corresponding solutions before posing related subsequent questions to it. In other words, our CoT prompt teaches the model to reason about the problem and mimic the same reasoning to respond to further queries correctly.

Choose the model

Use the dropdown option from the top-left corner of the right pane for the text generation exercise and choose 'GPT-4.1 Nano' model.



Chain-of-Thought to the rescue

Let's use the Chain-of-Thought approach to employ the AI's reasoning skills.

Our prompt will include a question, its accurate answer giving context and reasoning to the LLM, and then a different question that can be answered using the same line of reasoning.

Let's see if it works.

Q: An Italian menu has 5 items priced as follows:

- Prosciutto \$9.99
- Pecorino \$12.99
- Calamari \$13.99
- Bruschetta \$4.99
- Carpaccio \$14.99

Assuming each food is equally filling, spend \$30 for a group of people, maximizing satiety.

A: Since each food is equally filling, maximizing our budget will require us to favor items that have a low cost-per-satiety ratio. The ratio of each

- Bruschetta: \$4.99
- Prosciutto: \$9.99
- Pecorino: \$12.99
- Calamari: \$13.99
- Carpaccio: \$14.99

If we are allowed to have multiple orders of the same item, the simplest answer is doing an integer division of our \$30 by the item with the lowest cost.

Q: A store sells 4 types of aquarium fish at the following prices:

- Guppies: \$3.99
- Goldfish: \$1.99
- Betta: \$5.99
- Angelfish: \$8.99

Spend \$20 to maximize the total amount of fish.

We can place all this directly in our prompt or split the Chain-of-Thought training in the Prompt Instructions and the actual question in the prompt.

The answer generated should be similar to the following.

To maximize the total number of fish with a budget of 20, we need to prioritize the cheapest fish, which is the Goldfish at 1.99 each. Since each fish is equally filling, buying as many as possible of the cheapest fish will maximize the total count.

Let's calculate:

- Goldfish: \$1.99 each
- Number of Goldfish that can be bought: $\text{floor}(20 / 1.99) \approx 10$
- Total spent: $10 * 1.99 = \$19.90$, which is within the budget.

Remaining money: $20 - 19.90 = \$0.10$, which isn't enough to buy another fish.

Conclusion: Spend \$19.90 on 10 Goldfish to maximize the total fish count.

Perfect! Exactly what we wanted. The reply offers the correct answer and a proper explanation modeled after our example.

Let's consider another example. Use the following prompt (assume the same Prompt Instructions as before about the Italian menu), and check the response of the model.

A store sells three box types:

- Small, 10L, \$1.99
- Medium, 22L, \$4.99
- Large, 38L, \$5.99

Spend \$20 and maximize volume.

The response should be similar to:

To maximize volume with a \$20 budget, we should prioritize purchasing the largest volume boxes for the lowest cost per liter.

Let's analyze each box type:

Small: 10L at \$1.99 → cost per liter ≈ 0.199
 Medium: 22L at \$4.99 → cost per liter ≈ 0.227
 Large: 38L at \$5.99 → cost per liter ≈ 0.158

The Large box offers the best value per liter, so we should buy as many Large boxes as possible within the budget.

Calculations:

Maximum Large boxes: $\text{floor}(20 / 5.99) \approx 3$ (since $3 * 5.99 = 17.97$)
 Remaining money: $20 - 17.97 = 2.03$
 With remaining 2.03, we can buy one Small box (cost 1.99), leaving about \$0.04 unused.

Total volume:

3 Large boxes: $3 * 38L = 114L$
 1 Small box: 10L
 Total volume: $114L + 10L = 124L$
 Total spent: $17.97 + 1.99 = 19.96$, within the \$20 budget.

Conclusion: Buy 3 Large boxes and 1 Small box to maximize volume within the budget.

Zero-Shot Chain-of-Thought Prompting

Kojima et al. have found a simple, effective approach to improve the results we obtain from our model. This approach is often labeled as a Zero-Shot CoT.

There are a few words that, when added to the prompt, are likely to solicit better answers since they invite the AI to do step-by-step reasoning, much such as a human would when trying to come to a resolution.

According to researchers, two effective phrases are:

Let's think step by step.

And:

Let's work this out in a step-by-step way to be sure we have the right answer.

These words are helpful but not magic. So feel free to use them to improve your results, but they are usually best used along with other techniques discuss in this course.

For example, tucking these words at the end of our original standard prompt still generates an incorrect answer with the GPT 3.5 model (**gpt-3.5-turbo (Raw)**) available at the time of writing.

Prompt

Let's think step by step.
An Italian menu has 5 items, each equally satisfying, priced as follows:
- Prosciutto \$9.99
- Pecorino \$12.99
- Calamari \$13.99
- Bruschetta \$4.99
- Carpaccio \$14.99
Considering a budget of \$30, what should the money be spent on to maximize satiety ?

Output

To maximize satiety within a \$30 budget, we need to select a combination of menu items whose total cost does not exceed \$30 and yields the highest total satisfaction.

Assuming each item provides equal satisfaction, the goal simplifies to maximizing the total number of items purchased without exceeding the budget.

Let's analyze the options:

Item	Price
Prosciutto	9.99
Pecorino	12.99
Calamari	13.99
Bruschetta	4.99
Carpaccio	14.99

Since Bruschetta is the cheapest, buying multiple Bruschetta items could maximize the number of items, but we need to consider if combining different items yields better satisfaction.

Given that all items are equally satisfying, the optimal strategy is to buy as many inexpensive items as possible.

Let's calculate the maximum number of Bruschetta we can buy:

Budget: \$30
Price of Bruschetta: \$4.99
?Max quantity: $\text{floor}(30 / 4.99) \approx 6$ items, total cost ≈ 29.94
?Remaining budget: $30 - 29.94 \approx 0.06$, insufficient to buy another item.

Total satisfaction: 6 items.

Alternatively, buying a mix of items might yield more satisfaction if items are equally satisfying. But since all items are equally satisfying, and we want to maximize the number of items, buying only Bruschetta maximizes quantity.

Therefore, the optimal choice is to buy 6 Bruschetta items, spending approximately \$29.94.

You will note that the response is similar to the traditional Chain-of-Thought approach illustrated before, in which we use our prompt to "teach" the model the desired outcome.

Still, every time we get good results with just a few words, we should consider that option, especially since this prompting technique solicits longer and more elaborate answers, which is helpful for generating blog posts, essays, guides, and so on.

Chain-of-Thought to explore subjects

Chain-of-Thought can be used in various ways to improve the chatbot's reasoning, especially in areas where it's feeble. However, a more valuable use is when it comes to exploring subjects more in-depth.

Instead of asking a generic question, we can break it down into steps we want the model to consider to develop a much richer and valuable answer.

Let's see a practical example that doesn't necessarily use the question-and-answer template format.

For example, let's ask the model to tell us about Space Exploration:

What is space exploration?

The answer I received is reasonable:

Space exploration refers to the investigation and study of outer space through the use of astronomy, spacecraft, satellites, and other technologies. It aims to understand celestial bodies, such as planets, moons, asteroids, and stars, as well as the broader universe. Space exploration has led to numerous scientific discoveries, technological advancements, and has expanded our knowledge about the origins and nature of the universe. It involves missions conducted by government agencies like NASA, ESA, and private companies like SpaceX, often including activities like satellite deployment, planetary exploration, and even plans for human colonization of other planets.

Now, let's try a Chain-of-Thought approach. Start a new chat and place the following in the Prompt Instructions:

Cover each and every one of the following elements in your answer.

- Historic Space Missions
 - Moon landing and Human Achievement
 - Moon landing and impact on the Cold War
 - Satellite technology and its impact on humanity
 - Mars colonization possibilities
 - Search for extraterrestrial life
 - Space tourism prospects
 - Space debris and environmental impact
 - International Space Station collaboration
 - Advancements in rocket technology
 - Interstellar travel challenges
 - Private companies and Billionaires involvement controversy
- Give a comprehensive response that covers all points in depth.

We are giving the model a lot of items to consider in its answer and tucking in Kojima et al.'s suggestion of inviting the model to take a step-by-step approach for good measure.

For the prompt itself, let's try again:

What is space exploration?

You may expect an output as shown below.

What is Space Exploration?

Space exploration refers to the ongoing discovery and investigation of celestial bodies beyond Earth, including planets, moons, asteroids, comets, and the broader universe. It involves the use of spacecraft, satellites, telescopes, and human missions to gather scientific data, expand our understanding of the cosmos, and develop new technologies.

Key Elements of Space Exploration

1. Historic Space Missions

Early missions like Sputnik (1957), the first artificial satellite, and Apollo 11 (1969), which achieved the first Moon landing, laid the foundation for modern space exploration. These missions demonstrated human capability to reach and operate beyond Earth.

2. Moon Landing and Human Achievement

The Apollo program marked a pinnacle of human achievement, showcasing technological prowess and international prestige. The Moon landing symbolized human curiosity and the drive to explore the unknown.

3. Impact on the Cold War

During the Cold War, space exploration became a proxy for geopolitical rivalry between the USA and USSR. The space race accelerated technological innovation and fostered national pride, culminating in achievements like the Moon landing.

4. Satellite Technology and Its Impact on Humanity

Satellites revolutionized communication, weather forecasting, navigation, and Earth observation, profoundly impacting daily life and global industries.

5. Mars Colonization Possibilities

Recent advancements aim at establishing human habitats on Mars, with organizations like NASA and private companies exploring technologies for long-term colonization, addressing challenges like life support, radiation protection, and sustainable ecosystems.

6. Search for Extraterrestrial Life

Scientists search for signs of life beyond Earth through missions to Mars, Europa, Enceladus, and exoplanet studies, seeking biosignatures or habitable conditions.

7. Space Tourism Prospects

Commercial ventures by companies like SpaceX and Blue Origin aim to make space travel accessible to civilians, fostering a new industry of space tourism with potential economic and technological benefits.

8. Space Debris and Environmental Impact

Increasing satellite launches and defunct spacecraft contribute to space debris, posing collision risks and environmental concerns that require international regulation and debris mitigation strategies.

9. International Space Station Collaboration

The ISS exemplifies international cooperation, hosting astronauts from multiple countries, facilitating scientific research, and fostering peaceful collaboration in space.

10. Advancements in Rocket Technology

Recent innovations include reusable rockets, more efficient propulsion systems, and miniaturized spacecraft, reducing costs and enabling more frequent missions.

11. Interstellar Travel Challenges

Traveling beyond our solar system faces enormous hurdles like vast distances, propulsion limitations, and radiation hazards. Concepts like warp drives or generation ships are theoretical solutions under study.

12. Private Companies and Billionaires Involvement Controversy

Private entities and wealthy individuals like Elon Musk and Jeff Bezos accelerate space development but raise concerns about monopolization, safety, and ethical implications of space commercialization.

Conclusion

Space exploration embodies humanity's quest to understand the universe, driven by technological innovation, geopolitical factors, scientific curiosity, and commercial interests. It holds the promise of scientific breakthroughs, technological advancements, and perhaps, the future of human civilization beyond Earth.

What a huge difference!

The downside is that we had to develop a list requiring knowledge of the subject or at least research into it, and this is time-consuming.

On the plus side, we didn't have to retrain the model, which would be truly time-consuming and potentially expensive. Instead, the prompt split the "problem" into smaller steps worth exploring and leveraged the existing model training to compute a reply.

Moreover, these starting points can lead to various interconnected thoughts and ideas from the model. The beauty of a Chain-of-Thought is that it can branch out in different directions, exploring numerous aspects and perspectives related to the initial topic.

We can ask specific questions at any time after the model has already shown us a broader understanding of the topic.

Try Yourself

Exercises

1. **Chain-of-Thought Reasoning Practice:** Give the AI a list of fruits and their prices. Assuming each fruit offers the same health benefits, use the Chain-of-Thought approach to spend \$10 and maximize nutritional value.

Expected Output: With a \$10 budget, purchase as many of the least expensive fruit to maximize nutritional value.

2. **Zero-Shot CoT Prompting** Using the phrases "Let's think step by step" or "Let's work this out in a step-by-step way," pose a question about an unfamiliar topic and see if the AI can produce a more reasoned, detailed response.

Expected Activity: Assess the quality and depth of the AI's answer compared to a traditional prompt.

3. **Deep Dive using Chain-of-Thought:** Select a broad topic, for instance, "Ocean Conservation." Then, list various facets of the topic, such as plastic pollution, overfishing, coral reef degradation, and so on. Use the Chain-of-Thought approach to get the AI's comprehensive overview of the topic.

Expected Outcome: Evaluate the AI's response to see if it covers the topic more extensively and insightfully than a regular prompt.

Author(s)

[Antonio Cangiano](#)

Other Contributor(s)

[Abhishek Gagneja](#)

© IBM Corporation. All rights reserved.