

The Chain-of-Thought Approach in Prompt Engineering

Estimated time needed: 40 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-4o-mini**' model.

When the AI reasoning goes wrong

Let's try this in practice. Let us use a standard/naive prompt for the following reasoning problem.

```
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.
```

Now, the answer is to order six orders of the cheapest item since each item is equally satiating.

When we feed this to our LLM, its reasoning goes off the rails. In my case, I got a logic which ultimately suggested the following:

To maximize satiety with a budget of \$30 from the given Italian menu, you should aim to select the items that provide the most food for the cost. Here's one possible combination:

1. **Bruschetta:** \$4.99
2. **Prosciutto:** \$9.99
3. **Pecorino:** \$12.99

Total: \$4.99 + \$9.99 + \$12.99 = \$27.97

This is incorrect, as the optimal choice would be to order six servings of Bruschetta at \$4.99 each.

Chain-of-Thought to the rescue

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

Our prompt will include this question, an 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

- 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 low

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.

Try it yourself, but the answer I received was as follows:

To maximize the total number of fish you can buy with a budget of \$20, you'll want to focus on the cheapest options first. Here's the breakdown of the prices:

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

The most economical choice is to buy as many Goldfish as possible since they are the least expensive.

Calculating how many Goldfish you can buy with \$20:

1. **Cost of one Goldfish:** \$1.99
2. **Maximum number of Goldfish:** $\$20 / \$1.99 \approx 10.05$

Since you can only buy whole fish, you can purchase **10 Goldfish**.

Calculating the total cost for 10 Goldfish:

- **Total cost for 10 Goldfish** = $10 * \$1.99 = \19.90

This leaves you with \$0.10, which is not enough to purchase any additional fish.

Summary: Buy **10 Goldfish** for a total cost of **\$19.90** to maximize the number of fish you can get with your \$20 budget.

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

Sometimes, the test question we ask might be more complex than the one we give in our example, and the AI—depending on the model we are using—might not always do a great job.

Consider this prompt (assume the same Prompt Instructions as before about the Italian menu):

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 I got is this:

To maximize the volume of boxes you can buy with a budget of \$20, you'll want to focus on the cost per liter of each box type. Here's the breakdown of the prices and volumes:

- **Small Box:** 10L for \$1.99 (Cost per liter: \$0.199)

- **Medium Box:** 22L for \$4.99 (Cost per liter: \$0.227)
- **Large Box:** 38L for \$5.99 (Cost per liter: \$0.158)

The Large Box offers the best value in terms of volume per dollar spent. Now let's calculate how many of each box type you can buy without exceeding \$20.

1. Buying Large Boxes:

- Cost of one Large Box: \$5.99
- Maximum number of Large Boxes: $\$20 / \$5.99 \approx 3.34$
- Since you can only buy whole boxes, you can purchase **3 Large Boxes**.
- Total cost for 3 Large Boxes: $3 * \$5.99 = \17.97
- Total volume: $3 * 38L = 114L$
- Remaining budget: $\$20 - \$17.97 = \$2.03$ (not enough to buy any additional boxes).

2. Buying Medium Boxes:

- Cost of one Medium Box: \$4.99
- Maximum number of Medium Boxes: $\$20 / \$4.99 \approx 4.01$
- You can purchase **4 Medium Boxes**.
- Total cost for 4 Medium Boxes: $4 * \$4.99 = \19.96
- Total volume: $4 * 22L = 88L$
- Remaining budget: $\$20 - \$19.96 = \$0.04$ (not enough for another box).

3. Buying Small Boxes:

- Cost of one Small Box: \$1.99
- Maximum number of Small Boxes: $\$20 / \$1.99 \approx 10.05$
- You can purchase **10 Small Boxes**.
- Total cost for 10 Small Boxes: $10 * \$1.99 = \19.90
- Total volume: $10 * 10L = 100L$
- Remaining budget: $\$20 - \$19.90 = \$0.10$ (not enough for another box).

Summary:

- The best option for maximizing volume is to buy **3 Large Boxes** for a total cost of **\$17.97**, giving you **114 liters** of volume.
- Alternatively, **4 Medium Boxes** would give you **88 liters**, and **10 Small Boxes** would give you **100 liters**.

Thus, the optimal choice is **3 Large Boxes** for maximum volume.

This isn't terrible since it's mostly correct, but it's missing the fact that we have more room in our budget.

We can hint at that by following up with the AI:

You have money left over. Spend it.

Which gave me:

To utilize the remaining budget effectively, let's recap the leftover amounts from the previous calculations:

1. **From 3 Large Boxes:** You have \$2.03 left over after spending \$17.97.
2. **From 4 Medium Boxes:** You have \$0.04 left over after spending \$19.96.
3. **From 10 Small Boxes:** You have \$0.10 left over after spending \$19.90.

Since the goal is to maximize the use of your remaining budget, let's see how to spend the leftover amounts:

For the Large Boxes (\$2.03 remaining):

- You cannot buy any additional Large or Medium Boxes since they exceed the budget.
- You can buy **1 Small Box** for \$1.99, leaving you with \$0.04 remaining.

New Total:

- **3 Large Boxes** (114L) + **1 Small Box** (10L) = **124 liters** total volume.

For the Medium Boxes (\$0.04 remaining):

- You cannot buy anything additional with \$0.04.

Total Remains:

- **4 Medium Boxes** (88L) and no additional purchases.

For the Small Boxes (\$0.10 remaining):

- You cannot buy anything additional with \$0.10.

Total Remains:

- **10 Small Boxes** (100L) and no additional purchases.

Final Summary:

- **If you started with 3 Large Boxes:** Buy **1 Small Box** for a total of **124 liters**.
- **If you started with 4 Medium Boxes:** Total remains **88 liters**.
- **If you started with 10 Small Boxes:** Total remains **100 liters**.

The best option to maximize volume with the leftover money is to buy **1 Small Box** if you initially went with the Large Boxes, bringing your total to **124 liters**.

This is the right answer, but it's not the whole answer. We can get that by simply asking for it:

Give me the entire answer.

Which gave me the correct answer:

Certainly! Here's the complete breakdown, including prices for each option while maximizing volume within a \$20 budget:

Box Types and Prices:

- **Small Box:** 10L for \$1.99 (Cost per liter: \$0.199)
- **Medium Box:** 22L for \$4.99 (Cost per liter: \$0.227)
- **Large Box:** 38L for \$5.99 (Cost per liter: \$0.158)

Volume Maximization Strategy:

The Large Box provides the best value in terms of volume per dollar spent. Here's how many of each box type you can buy without exceeding \$20.

1. Buying Large Boxes:

- **Cost of one Large Box:** \$5.99
- **Maximum number of Large Boxes:** $\$20 / \$5.99 \approx 3.34$ (you can buy 3)
- **Total cost for 3 Large Boxes:** $3 * \$5.99 = \17.97
- **Total volume:** $3 * 38L = 114L$
- **Remaining budget:** $\$20 - \$17.97 = \$2.03$

2. Buying Medium Boxes:

- **Cost of one Medium Box:** \$4.99
- **Maximum number of Medium Boxes:** $\$20 / \$4.99 \approx 4.01$ (you can buy 4)
- **Total cost for 4 Medium Boxes:** $4 * \$4.99 = \19.96
- **Total volume:** $4 * 22L = 88L$
- **Remaining budget:** $\$20 - \$19.96 = \$0.04$

3. Buying Small Boxes:

- **Cost of one Small Box:** \$1.99
- **Maximum number of Small Boxes:** $\$20 / \$1.99 \approx 10.05$ (you can buy 10)
- **Total cost for 10 Small Boxes:** $10 * \$1.99 = \19.90
- **Total volume:** $10 * 10L = 100L$
- **Remaining budget:** $\$20 - \$19.90 = \$0.10$

Summary of Options:

- **3 Large Boxes:**
 - **Total Cost:** \$17.97
 - **Total Volume:** 114L
 - **Remaining Budget:** \$2.03
- **4 Medium Boxes:**
 - **Total Cost:** \$19.96
 - **Total Volume:** 88L
 - **Remaining Budget:** \$0.04
- **10 Small Boxes:**
 - **Total Cost:** \$19.90
 - **Total Volume:** 100L
 - **Remaining Budget:** \$0.10

Spending the Remaining Budget:

1. If you buy 3 Large Boxes:

- **Remaining Budget:** \$2.03
- You can buy **1 Small Box** for \$1.99, leaving you with \$0.04.
- **New Total:**
 - **Total Cost:** $\$17.97 + \$1.99 = \$19.96$
 - **Total Volume:** $114L + 10L = 124L$.

2. If you buy 4 Medium Boxes:

- **Remaining Budget:** \$0.04 (not enough to buy anything).
- **Total Remains:** 4 Medium Boxes (88L) and no additional purchases.

3. If you buy 10 Small Boxes:

- **Remaining Budget:** \$0.10 (not enough to buy anything).
- **Total Remains:** 10 Small Boxes (100L) and no additional purchases.

Final Recommendation:

- The best option to maximize volume is to **buy 3 Large Boxes and 1 Small Box**, resulting in a total of **124 liters** at a total cost of **\$19.96**.
- If you choose to go with **4 Medium Boxes** or **10 Small Boxes**, you will end up with **88 liters** and **100 liters**, respectively, without the ability to spend any remaining budget.

Thus, the optimal strategy is to purchase **3 Large Boxes** for **\$17.97** and **1 Small Box** for **\$1.99**, resulting in **124 liters** of total volume.

It's often helpful to make suggestions, make small corrections, and ask to reformat the answer in a way that is useful to us.

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 discussed 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 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.
```

Output

In other words, our traditional Chain-of-Thought approach illustrated above, in which we use our prompt to "teach" the model the desired outcome, is still superior.

Still, every time we get better 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 is the investigation and study of outer space through the use of space technology, including spacecraft, satellites, and telescopes. It involves the exploration of celestial bodies, such as planets, moons, asteroids, and comets, as well as the study of cosmic phenomena such as stars, galaxies, and the universe's structure.

Space exploration can be categorized into two main types:

1. **Robotic Exploration:** This involves sending unmanned spacecraft to gather data and conduct experiments. Notable examples include NASA's Mars rovers, the Voyager probes, and the Hubble Space Telescope.
2. **Human Exploration:** This includes missions with astronauts aboard, such as the Apollo moon landings, the International Space Station (ISS) missions, and future plans for crewed missions to Mars.

Recent advancements in space exploration include the Artemis program, which aims to return humans to the Moon by the mid-2020s, and the ongoing Mars missions, including the Perseverance rover and the Ingenuity helicopter, which are exploring the Martian surface and atmosphere.

Additionally, private companies such as SpaceX and Blue Origin are increasingly involved in space exploration, developing technologies for commercial space travel and potential colonization efforts.

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

Consider and include 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
- Let us think step by step.

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?

The output I got is as follows:

Space exploration is the human endeavor to explore and understand the vast expanse of outer space beyond Earth's atmosphere. It involves sending spacecraft, satellites, and, in some cases, humans beyond our planet to gather information about celestial bodies, study the cosmos, and develop technologies that can be used for various scientific, technological, and commercial purposes. Space exploration has evolved over time, leading to numerous historic space missions and significant achievements that have shaped our understanding of the universe and our place in it.

Historic Space Missions:

Historic space missions include milestones such as the launch of the first artificial satellite, Sputnik 1, by the Soviet Union in 1957, and the subsequent space race between the United States and the Soviet Union. Notably, the Apollo 11 mission in 1969 marked a monumental achievement with the first human landing on the Moon. Neil Armstrong and Buzz Aldrin became the first humans to set foot on another celestial body, showcasing the remarkable capabilities of human ingenuity and engineering.

Moon Landing and Human Achievement:

The Moon landing was a testament to human determination, technological prowess, and the spirit of exploration. It demonstrated our ability to overcome immense challenges and achieve what was once thought impossible. The achievement of landing humans on the Moon inspired generations and symbolized humanity's thirst for exploration and discovery.

Moon Landing and Impact on the Cold War:

The Moon landing also had political significance during the Cold War era. It was a part of the competition between the United States and the Soviet

Union for technological and ideological superiority. The successful Moon landing was a major propaganda victory for the U.S. and highlighted its scientific and technological capabilities.

Satellite Technology and its Impact on Humanity:

The development of satellite technology revolutionized communication, navigation, weather forecasting, Earth observation, and scientific research. Satellites orbiting Earth provide vital services, including global telecommunications, GPS navigation, and monitoring of weather patterns and environmental changes.

Mars Colonization Possibilities:

Mars has been a focal point for potential human colonization. Missions such as the Mars rovers and the planned Mars missions aim to gather information about the planet's surface, atmosphere, and potential habitability. The idea of Mars colonization poses challenges related to life support, radiation protection, and sustainable habitats.

Search for Extraterrestrial Life:

Exploration missions, such as those to Mars and the moons of Jupiter and Saturn, seek signs of past or present life beyond Earth. The discovery of microbial life or even complex organisms elsewhere in the solar system would have profound implications for our understanding of life's prevalence in the universe.

Space Tourism Prospects:

The concept of space tourism has gained traction, with private companies developing spacecraft capable of carrying civilians to the edge of space or even to orbital destinations. This raises the possibility of a new industry that offers civilians a chance to experience space travel.

Space Debris and Environmental Impact:

The growing amount of space debris, including defunct satellites and fragments from past missions, poses a threat to operational satellites and future space missions. Addressing space debris is crucial to ensuring the sustainability of space activities.

International Space Station Collaboration:

The International Space Station (ISS) serves as a symbol of international collaboration. Multiple countries contribute resources, expertise, and modules to operate the ISS, conducting scientific research and experiments in microgravity.

Advancements in Rocket Technology:

Advancements in rocket technology have led to more efficient and cost-effective launch systems. Reusable rockets, pioneered by companies such as SpaceX, have the potential to drastically reduce the cost of accessing space.

Interstellar Travel Challenges:

The concept of interstellar travel, or traveling between stars, presents immense challenges due to the vast distances involved. Theoretical concepts such as warp drives and advanced propulsion systems are still in the realm of speculation.

Private Companies and Billionaires Involvement Controversy:

The involvement of private companies and billionaires in space exploration has sparked both excitement and controversy. While their contributions have accelerated innovation, some concerns revolve around the concentration of power, resource allocation, and potential commercialization of space.

In summary, space exploration encompasses a wide range of endeavors, from historic achievements such as the Moon landing to modern advancements in technology and potential future endeavors such as Mars colonization and the search for extraterrestrial life. It reflects humanity's curiosity, innovation, and collaboration while also raising important ethical, environmental, and geopolitical considerations.

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](#)

