The Chain-of-Thought approach in Prompt Engineering

In this lab, we will discuss the Chain-of-Thought approach to prompting.

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.

When the AI reasoning goes wrong

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

```
1. 1
2. 2
3. 3
4. 4
5. 5
6. 6
7. 7
8. 8
9. 9

1. An Italian menu has 5 items priced as follows:
2.
3. - Prosciutto $9.99
4. - Pecorino $12.99
5. - Calamari $13.99
6. - Bruschetta $4.99
7. - Carpaccio $14.99
8.
9. Assuming each food is equally filling, spend $30 for a group of people by maximizing satiety.
```
Copied!

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 convoluted (and wrong!) logic which ultimately suggested the following:

> Therefore, the most filling combination of items within the budget of $30 is Carpaccio, Calamari, and Bruschetta.

This is not just wrong. It's spectacularly bad. For example, seeing that the sum of the meal exceeds the $30 is trivial.

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.

```
1. 1
2. 2
3. 3
4. 4
5. 5
6. 6
7. 7
8. 8
9. 9
10. 10
11. 11
12. 12
13. 13
14. 14
15. 15
16. 16
17. 17
18. 18
19. 19
20. 20
21. 21
22. 22
23. 23
24. 24
25. 25
26. 26
27. 27
28. 28
29. 29

1.
2. Q: An Italian menu has 5 items priced as follows:
3.
4. - Prosciutto $9.99
```

```
 5. - Pecorino $12.99
 6. - Calamari $13.99
 7. - Bruschetta $4.99
 8. - Carpaccio $14.99
 9.
10. Assuming each food is equally filling, spend $30 for a group of people by maximizing satiety.
11.
12. 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
13.
14. - Bruschetta: $4.99
15. - Prosciutto: $9.99
16. - Pecorino: $12.99
17. - Calamari: $13.99
18. - Carpaccio: $14.99
19.
20. 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 c
21.
22. Q: A store sells 4 types of aquarium fish at the following prices:
23.
24. - Guppies: $3.99
25. - Goldfish: $1.99
26. - Betta: $5.99
27. - Angelfish: $8.99
28.
29. Spend $20 to maximize the total amount of fish.
```

Copied!

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:

> If we are allowed to have multiple orders of the same type of fish, the simplest answer is doing an integer division of our $20 by the cost of the fish with the lowest price, which is Goldfish. So we get $20 div $1.99 (the cost of Goldfish) which is 10. So the answer is to buy 10 Goldfish.

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):

```
1. 1
2. 2
3. 3
4. 4
5. 5
6. 6
7. 7

1. A store sells three box types:
2.
3. - Small, 10L, $1.99
4. - Medium, 22L, $4.99
5. - Large, 38L, $5.99
6.
7. Spend $20 and maximize volume.
```

Copied!

With a GPT 3.5 model, the response I got is this:

> To maximize volume while spending $20, we need to consider the cost per liter of each box size. The cost per liter for each box size is as follows:
>
> Small: $1.99 / 10L = $0.199 per liter
> Medium: $4.99 / 22L = $0.226 per liter
> Large: $5.99 / 38L = $0.157 per liter
>
> To maximize volume, we want to choose the box size with the lowest cost per liter. In this case, the large box still has the lowest cost per liter.
>
> So, we can spend $20 on large boxes. $20 / $5.99 (the cost of a large box) is approximately 3.34. Since we can't have a fraction of a box, we can buy 3 large boxes with $20, which would give us a total volume of 3 * 38L = 114L.

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:

```
1. 1

1. You have money left over. Spend it.
```

Copied!

Which gave me:

> Since we have $20 and we spent $17.97 on 3 Large boxes, we have $20 - $17.97 = $2.03 left. With this remaining amount, we can purchase a Small box, which costs $1.99. Therefore, we can spend the remaining $2.03 on a Small box.

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

```
1. 1
```

```
1. Give me the entire answer.
```

Copied!

Which gave me the correct answer:

> To maximize volume with a budget of $20, we can buy 3 Large boxes and 1 Small box.

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 like 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 available at the time of writing.

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, etc.

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:

```
1. 1

1. What is space exploration?
```

Copied!

The answer I received is reasonable:

> Space exploration is the discovery and exploration of celestial bodies in outer space. It involves the use of various technologies and spacecraft to study and understand the universe beyond Earth. This includes missions to other planets, moons, asteroids, and comets, as well as the study of stars, galaxies, and other celestial objects. Space exploration allows us to gain knowledge about the origins and evolution of the universe, search for signs of life on other planets, and potentially pave the way for future human colonization of other celestial bodies.

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

```
1. 1
2. 2
3. 3
4. 4
5. 5
6. 6
7. 7
8. 8
9. 9
10. 10
11. 11
12. 12
13. 13
14. 14
15. 15
16. 16

1. Consider and include the following elements in your answer:
2.
3. - Historical Space Missions
4. - Moon landing and Human Achievement
5. - Moon landing and impact on the Cold War
6. - Satellite technology and its impact on humanity
7. - Mars colonization possibilities
8. - Search for extraterrestrial life
9. - Space tourism prospects
```

```
10. - Space debris and environmental impact
11. - International Space Station collaboration
12. - Advancements in rocket technology
13. - Interstellar travel challenges
14. - Private companies and Billionaires involvement controversy
15.
16. Let us think step by step.
```

Copied!

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:

```
1. 1

1. What is space exploration?
```

Copied!

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 historical space missions and significant achievements that have shaped our understanding of the universe and our place in it.

Historical Space Missions:
Historical space missions include milestones like 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 like 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, like 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 like 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 like 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 historical achievements like the Moon landing to modern advancements in technology and potential future endeavors like 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.

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, like plastic pollution, overfishing, coral reef degradation, etc. 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](#)