

第五章 处理输入-思维链推理

有时，语言模型需要进行详细的逐步推理才能回答特定问题。如果过于匆忙得出结论，很可能在推理链中出现错误。因此，我们可以通过“思维链推理”（Chain of Thought Reasoning）的策略，在查询中明确要求语言模型先提供一系列相关推理步骤，进行深度思考，然后再给出最终答案，这更接近人类解题的思维过程。

相比直接要求输出结果，这种引导语言模型逐步推理的方法，可以减少其匆忙错误，生成更准确可靠的响应。思维链推理使语言模型更好地模拟人类逻辑思考，是提升其回答质量的重要策略之一。

在本章中，我们将探讨如何处理语言模型的输入,以生成高质量的输出。我们将详细介绍如何构建思维链推理 Prompt，并通过案例分析这种方法的效果。掌握这一技巧将有助于开发者获得更佳的语言模型输出。

一、思维链提示设计

思维链提示是一种引导语言模型进行逐步推理的 Prompt 设计技巧。它通过在 Prompt 中设置系统消息，要求语言模型在给出最终结论之前，先明确各个推理步骤。

具体来说，Prompt可以先请语言模型陈述对问题的初步理解，然后列出需要考虑的方方面面，最后再逐个分析这些因素，给出支持或反对的论据，才得出整体的结论。这种逐步推理的方式，更接近人类处理复杂问题的思维过程，可以减少语言模型匆忙得出错误结论的情况。因为它必须逐步论证自己的观点，而不是直接输出结论。通过详细的思维链提示，开发者可以获得语言模型生成的结论更加可靠，理由更加充分。这种提示设计技巧值得在需要语言模型进行复杂推理时加以运用。

1.1 系统消息设计

首先，在系统消息中使用思维链提示：

```
delimiter = "===="

system_message = f"""
请按照以下步骤回答客户的提问。客户的提问将以{delimiter}分隔。

步骤 1:{delimiter}首先确定用户是否正在询问有关特定产品或产品的问题。产品类别不计入范围。

步骤 2:{delimiter}如果用户询问特定产品，请确认产品是否在以下列表中。所有可用产品：

产品: TechPro 超极本
类别: 计算机和笔记本电脑
品牌: TechPro
型号: TP-UB100
保修期: 1 年
评分: 4.5
特点: 13.3 英寸显示屏, 8GB RAM, 256GB SSD, Intel Core i5 处理器
描述: 一款适用于日常使用的时尚轻便的超极本。
价格: $799.99

产品: Bluewave 游戏笔记本电脑
类别: 计算机和笔记本电脑
品牌: Bluewave
型号: BW-GL200
保修期: 2 年
```

评分: 4.7

特点: 15.6 英寸显示屏, 16GB RAM, 512GB SSD, NVIDIA GeForce RTX 3060

描述: 一款高性能的游戏笔记本电脑, 提供沉浸式体验。

价格: \$1199.99

产品: PowerLite 可转换笔记本电脑

类别: 计算机和笔记本电脑

品牌: PowerLite

型号: PL-CV300

保修期: 1年

评分: 4.3

特点: 14 英寸触摸屏, 8GB RAM, 256GB SSD, 360 度铰链

描述: 一款多功能可转换笔记本电脑, 具有响应触摸屏。

价格: \$699.99

产品: TechPro 台式电脑

类别: 计算机和笔记本电脑

品牌: TechPro

型号: TP-DT500

保修期: 1年

评分: 4.4

特点: Intel Core i7 处理器, 16GB RAM, 1TB HDD, NVIDIA GeForce GTX 1660

描述: 一款功能强大的台式电脑, 适用于工作和娱乐。

价格: \$999.99

产品: Bluewave Chromebook

类别: 计算机和笔记本电脑

品牌: Bluewave

型号: BW-CB100

保修期: 1 年

评分: 4.1

特点: 11.6 英寸显示屏, 4GB RAM, 32GB eMMC, Chrome OS

描述: 一款紧凑而价格实惠的 Chromebook, 适用于日常任务。

价格: \$249.99

步骤 3:{delimiter} 如果消息中包含上述列表中的产品, 请列出用户在消息中做出的任何假设, \ 例如笔记本电脑 X 比笔记本电脑 Y 大, 或者笔记本电脑 Z 有 2 年保修期。

步骤 4:{delimiter} 如果用户做出了任何假设, 请根据产品信息确定假设是否正确。

步骤 5:{delimiter} 如果用户有任何错误的假设, 请先礼貌地纠正客户的错误假设(如果适用)。 \ 只提及或引用可用产品列表中的产品, 因为这是商店销售的唯一五款产品。以友好的口吻回答客户。

使用以下格式回答问题:

步骤 1: {delimiter} <步骤 1 的推理>

步骤 2: {delimiter} <步骤 2 的推理>

步骤 3: {delimiter} <步骤 3 的推理>

步骤 4: {delimiter} <步骤 4 的推理>

回复客户: {delimiter} <回复客户的内容>

请确保每个步骤上面的回答中中使用 {delimiter} 对步骤和步骤的推理进行分隔。

""""

1.2 用户消息测试

接下来，在用户消息中测试在系统消息中设置的思维链提示：

1.2.1 更贵的电脑

```
from tool import get_completion_from_messages

user_message = f"""Bluewave Chromebook 比 TechPro 台式电脑贵多少? """

messages = [
    {'role': 'system',
     'content': system_message},
    {'role': 'user',
     'content': f"{delimiter}{user_message}{delimiter}"},
]

response = get_completion_from_messages(messages)
print(response)
```

步骤 1：用户询问了关于产品价格的问题。

步骤 2：用户提到了两个产品，其中一个是Bluewave Chromebook，另一个是TechPro 台式电脑。

步骤 3：用户假设Bluewave Chromebook比TechPro 台式电脑贵。

步骤 4：根据产品信息，我们可以确定用户的假设是错误的。

回复客户：Bluewave Chromebook 的价格是 \$249.99，而 TechPro 台式电脑的价格是 \$999.99。因此，TechPro 台式电脑比 Bluewave Chromebook 贵 \$750。

1.2.2 你有电视么？

```
user_message = f"""你有电视机么"""
messages = [
    {'role': 'system',
     'content': system_message},
    {'role': 'user',
     'content': f"{delimiter}{user_message}{delimiter}"},
]
response = get_completion_from_messages(messages)
print(response)
```

步骤 1：我们需要确定用户是否正在询问有关特定产品或产品的问题。产品类别不计入范围。

步骤 2：在可用产品列表中，没有提到任何电视机产品。

回复客户：很抱歉，我们目前没有可用的电视机产品。我们的产品范围主要包括计算机和笔记本电脑。如果您对其他产品有任何需求或疑问，请随时告诉我们。

二、内心独白

在某些应用场景下，完整呈现语言模型的推理过程可能会泄露关键信息或答案，这并不可取。例如在教学应用中，我们希望学生通过自己的思考获得结论，而不是直接被告知答案。

针对这一问题。“内心独白”技巧可以在一定程度上隐藏语言模型的推理链。具体做法是，在 Prompt 中指示语言模型以结构化格式存储需要隐藏的中间推理，例如存储为变量。然后在返回结果时，仅呈现对用户有价值的输出，不展示完整的推理过程。这种提示策略只向用户呈现关键信息，避免透露答案。同时语言模型的推理能力也得以保留。适当使用“内心独白”可以在保护敏感信息的同时，发挥语言模型的推理特长。

总之，适度隐藏中间推理是 Prompt 工程中重要的技巧之一。开发者需要为不同用户制定不同的信息呈现策略。以发挥语言模型最大价值。

```
try:
    if delimiter in response:
        final_response = response.split(delimiter)[-1].strip()
    else:
        final_response = response.split(":")[-1].strip()
except Exception as e:
    final_response = "对不起，我现在有点问题，请尝试问另外一个问题"

print(final_response)
```

很抱歉，我们目前没有可用的电视机产品。我们的产品范围主要包括计算机和笔记本电脑。如果您对其他产品有任何需求或疑问，请随时告诉我们。

在复杂任务中，我们往往需要语言模型进行多轮交互、逐步推理，才能完成整个流程。如果想在 Prompt 中完成全部任务，对语言模型的能力要求会过高，成功率较低。

因此，下一章将介绍一种更可靠的策略：将复杂任务分解为多个子任务，通过提示链(Prompt Chaining) step-by-step 引导语言模型完成。具体来说，我们可以分析任务的不同阶段，为每个阶段设计一个简单明确的 Prompt。我们将通过实例展示提示链的运用，以及如何科学拆分 Prompt 来引导语言模型递进完成多步骤任务。这是提示工程中非常重要的技能之一。

三、英文版

1.1 思维链提示

```
delimiter = "####"
system_message = f"""
Follow these steps to answer the customer queries.
The customer query will be delimited with four hashtags, \
i.e. {delimiter}.

Step 1:{delimiter} First decide whether the user is \
asking a question about a specific product or products. \
Product category doesn't count.

Step 2:{delimiter} If the user is asking about \
specific products, identify whether \
the products are in the following list.
All available products:
1. Product: TechPro Ultrabook
   Category: Computers and Laptops
   Brand: TechPro
   Model Number: TP-UB100
   Warranty: 1 year
   Rating: 4.5
```

Features: 13.3-inch display, 8GB RAM, 256GB SSD, Intel Core i5 processor

Description: A sleek and lightweight ultrabook for everyday use.

Price: \$799.99

2. Product: BlueWave Gaming Laptop

Category: Computers and Laptops

Brand: BlueWave

Model Number: BW-GL200

Warranty: 2 years

Rating: 4.7

Features: 15.6-inch display, 16GB RAM, 512GB SSD, NVIDIA GeForce RTX 3060

Description: A high-performance gaming laptop for an immersive experience.

Price: \$1199.99

3. Product: PowerLite Convertible

Category: Computers and Laptops

Brand: PowerLite

Model Number: PL-CV300

Warranty: 1 year

Rating: 4.3

Features: 14-inch touchscreen, 8GB RAM, 256GB SSD, 360-degree hinge

Description: A versatile convertible laptop with a responsive touchscreen.

Price: \$699.99

4. Product: TechPro Desktop

Category: Computers and Laptops

Brand: TechPro

Model Number: TP-DT500

Warranty: 1 year

Rating: 4.4

Features: Intel Core i7 processor, 16GB RAM, 1TB HDD, NVIDIA GeForce GTX 1660

Description: A powerful desktop computer for work and play.

Price: \$999.99

5. Product: BlueWave Chromebook

Category: Computers and Laptops

Brand: BlueWave

Model Number: BW-CB100

Warranty: 1 year

Rating: 4.1

Features: 11.6-inch display, 4GB RAM, 32GB eMMC, Chrome OS

Description: A compact and affordable Chromebook for everyday tasks.

Price: \$249.99

Step 3:{delimiter} If the message contains products \ in the list above, list any assumptions that the \ user is making in their \ message e.g. that Laptop X is bigger than \ Laptop Y, or that Laptop Z has a 2 year warranty.

Step 4:{delimiter}: If the user made any assumptions, \ figure out whether the assumption is true based on your \ product information.

Step 5:{delimiter}: First, politely correct the \ customer's incorrect assumptions if applicable. \

Only mention or reference products in the list of \ 5 available products, as these are the only 5 \ products that the store sells. \ Answer the customer in a friendly tone.

Use the following format:

Step 1:{delimiter} <step 1 reasoning>

Step 2:{delimiter} <step 2 reasoning>

Step 3:{delimiter} <step 3 reasoning>

Step 4:{delimiter} <step 4 reasoning>

Response to user:{delimiter} <response to customer>

Make sure to include {delimiter} to separate every step.
"""

```
user_message = f"""
by how much is the Bluewave Chromebook more expensive \
than the TechPro Desktop"""
```

```
messages = [
    {'role': 'system',
     'content': system_message},
    {'role': 'user',
     'content': f"{delimiter}{user_message}{delimiter}"},
]
```

```
response = get_completion_from_messages(messages)
print(response)
```

Step 1:#### The user is asking about the price difference between the Bluewave Chromebook and the TechPro Desktop.

Step 2:#### Both the Bluewave Chromebook and the TechPro Desktop are available products.

Step 3:#### The user assumes that the Bluewave Chromebook is more expensive than the TechPro Desktop.

Step 4:#### Based on the product information, the price of the Bluewave Chromebook is \$249.99, and the price of the TechPro Desktop is \$999.99. Therefore, the TechPro Desktop is actually more expensive than the Bluewave Chromebook.

Response to user:#### The Bluewave Chromebook is actually less expensive than the TechPro Desktop. The Bluewave Chromebook is priced at \$249.99, while the TechPro Desktop is priced at \$999.99.

```

user_message = f"""
do you sell tvs"""
messages = [
    {'role': 'system',
     'content': system_message},
    {'role': 'user',
     'content': f"{delimiter}{user_message}{delimiter}"},
]
response = get_completion_from_messages(messages)
print(response)

```

Step 1:#### The user is asking if the store sells TVs, which is a question about a specific product category.

Step 2:#### TVs are not included in the list of available products. The store only sells computers and laptops.

Response to user:#### I'm sorry, but we currently do not sell TVs. Our store specializes in computers and laptops. If you have any questions or need assistance with our available products, feel free to ask.

2.1 内心独白

```

try:
    final_response = response.split(delimiter)[-1].strip()
except Exception as e:
    final_response = "Sorry, I'm having trouble right now, please try asking another question."

print(final_response)

```

I'm sorry, but we currently do not sell TVs. Our store specializes in computers and laptops. If you have any questions or need assistance with our available products, feel free to ask.