

# Homework 5: Let's Play LLM

## Report

Student ID: 111550100 Name: 邱振源

### Part 1: Designing a Task for LLM and Explore the Capability of LLM (70%)

#### Task Description (10%)

- **Task Introduction (Describe the task you want to solve):**

In this task, ChatGPT will play the role of a student conducting a chemistry experiment. We will provide ChatGPT with a list of materials (chemicals) it currently has. Then, we will give ChatGPT several tasks (generating specified products), and ChatGPT needs to tell us which materials it used to complete these experiments.

- **Input Prompt Format (Describe the format of the input prompt):**

The input can mainly be divided into two parts. The first part informs ChatGPT of the materials it currently possesses, and the second part lists the chemical reaction experiments ChatGPT needs to complete. Since using Chinese is more suitable for prompts that match my knowledge, I will use Chinese for the practical parts.

- **Ideal Output:**

- Tasks 1 and 2 are achievable, and ChatGPT should provide the correct answers.
- Task 3 cannot be directly completed with the existing materials, but ChatGPT should be able to first synthesize Material A using the available chemicals and then perform a chemical reaction with Material A and other available chemicals to obtain the desired product.
- Task 4 cannot be completed with the available chemicals; ChatGPT should recognize and explain its inability to synthesize the product.

## Motivation (10%)

- **Reason of Choosing the Task (Describe the reason why you choose this task):**

Many people apply ChatGPT in the field of education, and positioning ChatGPT as a one-on-one tutor is one of its advertised uses. However, chemistry experiments in academic subjects carry inherent risks, and if ChatGPT generates incorrect chemical procedures, it could lead to dangerous outcomes for those who follow them incorrectly. Therefore, I want to determine whether ChatGPT can provide accurate answers in the field of chemistry experiments.

- **Capability to Explore (Describe the capability of LLM you want to explore):**

As described above, I want to know whether ChatGPT can provide accurate answers in the field of chemistry experiments and also judge through Task 3 whether it can remember the steps it has completed to achieve the specified task, which also shows the ability of updating the original data automatically.

## Describe your Initial Attempt (15%)

- **Input**

Here I directly tell Chatgpt to be a student, and provide the material and task.

你是一個學生，現在有一些實驗你必須完成。

下面是你擁有的材料：

- \* 鹽酸
- \* 醋酸
- \* 酒精
- \* 小蘇打
- \* 水

接下來是你需要完成的實驗，這些實驗需要照順序完成，請一步一步告訴你我使用了哪些材料來完成下面的實驗：

1. 生成食鹽
2. 生成乙酸乙酯
3. 製作氫氧化鈉
4. 生成硫酸

- **Output**



好的，讓我來為你一步步完成這些實驗。

1. 生成食鹽：

- 使用鹽酸（盐酸）與水反應，生成氯化氫氣體和氯化鈉（食鹽）。
- 反應式： $\text{HCl} + \text{NaOH} \rightarrow \text{NaCl} + \text{H}_2\text{O}$
- 使用鹽酸（盐酸）與水。

2. 生成乙酸乙酯：

- 使用醋酸（乙酸）和酒精（乙醇）進行酯化反應。
- 反應式： $\text{CH}_3\text{COOH} + \text{C}_2\text{H}_5\text{OH} \rightarrow \text{CH}_3\text{COOC}_2\text{H}_5 + \text{H}_2\text{O}$
- 使用醋酸（乙酸）和酒精（乙醇）。

3. 製作氫氧化鈉：

- 使用氫氧化鈉的方法是將小蘇打（碳酸氫鈉）加熱分解。
- 反應式： $2 \text{NaHCO}_3 \rightarrow \text{Na}_2\text{CO}_3 + \text{H}_2\text{O} + \text{CO}_2$
- 使用小蘇打（碳酸氫鈉）。

4. 生成硫酸：

- 使用硫酸生成的方法是將硫酸的水溶液脫水。
- 反應式： $\text{H}_2\text{SO}_4 \rightarrow \text{H}_2\text{O} + \text{SO}_3$
- 使用硫酸。

這樣完成了所有的實驗，你的材料也都用上了。需要進一步解釋或者有其他問題嗎？

Even with the same prompt, ChatGPT still generates different answers each time. Therefore, I have tried the prompt multiple times, and the correctness of each task's output remains inconsistent. This demonstrates that ChatGPT's responses still have a significant degree of uncertainty.

- **Analysis**

- The result of Task 1 and Task 3 is wrong, since it uses the incorrect chemical reaction.
- The result of Task 2 is correct.
- The result of Task 4 is wrong, since it uses the material that is not provided.

From the results mentioned above, it can be observed that ChatGPT has a limited understanding of the accuracy regarding the chemicals it possesses and the experiments it conducts. It might perform incorrect experiments and use chemicals that were not provided. Therefore, I have decided to use more influential prompts.

## Describe Improvement (35%)

- **Input**

Since in Task 4, it use the material that is not provide, so I add the prompt “只能使用我提供的材料，如果這些材料無法完成實驗請告訴我為甚麼”.

你是一個學生，現在有一些實驗你必須完成。

下面是你擁有的材料：

- \* 鹽酸
- \* 醋酸
- \* 酒精
- \* 小蘇打
- \* 水

接下來是你需要完成的實驗，這些實驗需要照順序完成，請一步一步告訴我你使用了哪些材料來完成下面的實驗：

1. 生成食鹽
2. 生成乙酸乙酯
3. 製作氫氧化鈉
4. 生成硫酸

只能使用我提供的材料，如果這些材料無法完成實驗請告訴我為甚麼

- **Output**



好的，讓我們來看看如何使用提供的材料完成這些實驗。

1. 生成食鹽：

- 我們可以使用鹽酸（盐酸）和水來與氫氧化鈉（氫氧化鈉）反應，生成食鹽（氯化鈉）和水。
- 反應式： $\text{HCl} + \text{NaOH} \rightarrow \text{NaCl} + \text{H}_2\text{O}$
- 使用鹽酸（盐酸）和氫氧化鈉（小蘇打）。

2. 生成乙酸乙酯：

- 我們需要醋酸（乙酸）和酒精（乙醇）進行酯化反應。
- 反應式： $\text{CH}_3\text{COOH} + \text{C}_2\text{H}_5\text{OH} \rightarrow \text{CH}_3\text{COOC}_2\text{H}_5 + \text{H}_2\text{O}$
- 使用醋酸（乙酸）和酒精（乙醇）。

3. 製作氫氧化鈉：

- 這個實驗可能無法完成，因為提供的材料中沒有氫氧化鈉（氫氧化鈉）的前驅物。

4. 生成硫酸：

- 這個實驗也無法完成，因為提供的材料中沒有硫酸的前驅物。

總結來說，使用提供的材料，我們無法完成製備氫氧化鈉和硫酸的實驗，因為缺少相應的前驅物。

- **Analysis**

Since I want to use the adding prompt to deal with the specific problem, so I focus on the Task that the output is changed compare to the output of the previous step.

- The output of Task 3 is still wrong, since NaOH could be generate by generating NaCl of the Task 1 at first.
- The output of Task 4 is correct after adding the new prompt.

These result shows that Chatgpt can understand the meaning of the new prompt, which shows it can remember the provided list. However, Chatgpt still can't update the material list automatically according to the order of experiment.

- **Input**

Here, I want to let Chatgpt update the material list according to the order of experiment, so I add the prompt “但可以將已經做完的實驗的生成物視為可以使用的材料”.

你是一個學生，現在有一些實驗你必須完成。

下面是你擁有的材料：

- \* 鹽酸
- \* 醋酸
- \* 酒精
- \* 小蘇打
- \* 水

接下來是你需要完成的實驗，這些實驗需要照順序完成，請一步一步告訴我用你使用了哪些材料來完成下面的實驗：

1. 生成食鹽
2. 生成乙酸乙酯
3. 製作氫氧化鈉
4. 生成硫酸

只能使用我提供的材料，但可以將已經做完的實驗的生成物視為可以使用的材料，如果這些材料無法完成實驗請告訴我為甚麼

- **Output**



好的，我會盡力使用你提供的材料來完成這些實驗，讓我們開始：

1. 生成食鹽：

- 使用鹽酸（盐酸）和氫氧化鈉（小蘇打）進行中和反應，生成食鹽（氯化鈉）和水。
- 反應式： $\text{HCl} + \text{NaOH} \rightarrow \text{NaCl} + \text{H}_2\text{O}$
- 使用鹽酸（盐酸）和氫氧化鈉（小蘇打）。

2. 生成乙酸乙酯：

- 使用醋酸（乙酸）和酒精（乙醇）進行酯化反應。
- 反應式： $\text{CH}_3\text{COOH} + \text{C}_2\text{H}_5\text{OH} \rightarrow \text{CH}_3\text{COOC}_2\text{H}_5 + \text{H}_2\text{O}$
- 使用醋酸（乙酸）和酒精（乙醇）。

3. 製作氫氧化鈉：

- 我們已經在第一步生成了食鹽，因此可以從食鹽中製取氫氧化鈉。
- 使用食鹽（氯化鈉）和水進行電解反應，生成氫氧化鈉。
- 反應式： $2 \text{NaCl} + 2 \text{H}_2\text{O} \rightarrow 2 \text{NaOH} + \text{Cl}_2 + \text{H}_2 \uparrow$
- 使用食鹽（氯化鈉）和水。

4. 生成硫酸：

- 使用鹽酸（盐酸）和硫酸銅（II）進行雙置換反應，生成硫酸和氯化銅（I）。
- 反應式： $2 \text{HCl} + \text{CuSO}_4 \rightarrow \text{CuCl}_2 + \text{H}_2\text{SO}_4$
- 使用鹽酸（盐酸）和硫酸銅（II）。

這樣就完成了所有的實驗，我們成功使用了提供的材料。有什麼需要進一步解釋或者有其他問題嗎？

- **Analysis**

- The output of Task 1 is still wrong, however, it provide the correct name but wrong chemical formula of the material.
- The output of Task 3 is correct after adding the new prompt.
- The output of Task 4 become wrong after adding the new prompt.

The result shows that Chatgpt could automatically update the given list after adding the proper prompt. However, the result of Task 4 becomes to wrong after the new prompt. This is probably because the influence of new prompt is too large, make Chatgpt update the material list incorrectly.

## • Input

Since the result of Task 1 is almost correct (right name but wrong formula), so I add the prompt “請確保化學式的正確性，同時確保其符合化學定律”.

你是一個學生，現在有一些實驗你必須完成。

下面是你擁有的材料：

- \* 鹽酸 (HCl)
- \* 醋酸 (CH<sub>3</sub>COOH)
- \* 酒精 (C<sub>2</sub>H<sub>5</sub>OH)
- \* 小蘇打 (NaHCO<sub>3</sub>)
- \* 水 (H<sub>2</sub>O)

接下來是你需要完成的實驗，這些實驗需要照順序完成，請一步一步告訴我用你使用了哪些材料來完成下面的實驗：

1. 生成食鹽 (NaCl)
2. 生成乙酸乙酯 (CH<sub>3</sub>COOCH<sub>2</sub>CH<sub>3</sub>)
3. 製作氫氧化鈉 (NaOH)
4. 生成硫酸 (H<sub>2</sub>SO<sub>4</sub>)

只能使用我提供的材料，但可以將已經做完的實驗的生成物視為可以使用的材料，如果這些材料無法完成實驗請告訴我為甚麼  
請確保化學式的正確性，同時確保其符合化學定律

## • Output



好的，我會盡力使用提供的材料來完成這些實驗，讓我們一步步來完成：

### 1. 生成食鹽 (NaCl)：

- 我們可以使用鹽酸 (HCl) 和小蘇打 (NaHCO<sub>3</sub>) 進行中和反應，生成食鹽 (NaCl)、水 (H<sub>2</sub>O) 和二氧化碳 (CO<sub>2</sub>)。
- 反應式： $\text{HCl} + \text{NaHCO}_3 \rightarrow \text{NaCl} + \text{H}_2\text{O} + \text{CO}_2$
- 使用鹽酸 (HCl) 和小蘇打 (NaHCO<sub>3</sub>)。

### 2. 生成乙酸乙酯 (CH<sub>3</sub>COOCH<sub>2</sub>CH<sub>3</sub>)：

- 我們可以使用醋酸 (CH<sub>3</sub>COOH) 和酒精 (C<sub>2</sub>H<sub>5</sub>OH) 進行酯化反應，生成乙酸乙酯 (CH<sub>3</sub>COOCH<sub>2</sub>CH<sub>3</sub>) 和水 (H<sub>2</sub>O)。
- 反應式： $\text{CH}_3\text{COOH} + \text{C}_2\text{H}_5\text{OH} \rightarrow \text{CH}_3\text{COOCH}_2\text{CH}_3 + \text{H}_2\text{O}$
- 使用醋酸 (CH<sub>3</sub>COOH) 和酒精 (C<sub>2</sub>H<sub>5</sub>OH)。

### 3. 製作氫氧化鈉 (NaOH)：

- 我們已經在第一步生成了食鹽 (NaCl)，現在可以從食鹽中製取氫氧化鈉 (NaOH)。
- 使用食鹽 (NaCl) 和水 (H<sub>2</sub>O) 進行電解反應，生成氫氧化鈉 (NaOH)、氯氣 (Cl<sub>2</sub>) 和氫氣 (H<sub>2</sub>)。
- 反應式： $2 \text{NaCl} + 2 \text{H}_2\text{O} \rightarrow 2 \text{NaOH} + \text{Cl}_2 + \text{H}_2 \uparrow$
- 使用食鹽 (NaCl) 和水 (H<sub>2</sub>O)。

### 4. 生成硫酸 (H<sub>2</sub>SO<sub>4</sub>)：

- 我們無法使用現有的材料來直接生成硫酸 (H<sub>2</sub>SO<sub>4</sub>)。因為我們沒有硫的源物質。
- 因此，我們無法完成這個實驗。

總結來說，我們成功使用提供的材料來完成了前三個實驗，但無法完成生成硫酸的實驗，因為缺乏硫的源物質。

- **Analysis**

- All Task is correct at this prompt

The result shows that Chatgpt ensure correctness and check whether the answer conform to known theorem or not while making the responses according to the prompt. So, after adding the prompt of “請確保化學式的正確性，同時確保其符合化學定律”, the response become correct.

**Link of this chat:** <https://chatgpt.com/share/ee858085-2c28-4676-bd80-48944902dff8>

## Part 2: Problem Answering (30%)

### Question 1

1. The methods could be:

- Providing Examples and Counterexamples:

Provide examples of correct responses and incorrect responses to guide the LLM in understanding the desired output.

- Constructing Step-by-Step Instructions:

Making the task into smaller task, more manageable steps can help the LLM process and respond more accurately.

2. I set up the prompt as “給我所有 168 和 294 可能的公因數”. And the output would be

給我所有168和294可能的公因數

要找到168和294的所有可能公因數，我們可以列出它們的因數，然後找到它們的共同因數。首先，我們列出168和294的因數：

對於168：

1, 2, 3, 4, 6, 7, 8, 12, 14, 21, 24, 28, 42, 56, 84, 168

對於294：

1, 2, 3, 6, 7, 14, 21, 42, 49, 98, 147, 294

現在我們來看看它們的共同因數：

1, 2, 3, 6, 7, 14, 21, 42

所以，所有168和294的可能公因數是1、2、3、6、7和42。

It is clear that the output is wrong since the response didn't contain all the possible factor.



- Providing Examples and Counterexamples:

Here I set up the example of what is factor and the counterexample that what is not factor on the prompt, and the output becomes correct.

想要知道兩個數的所有公因數，我們需要先找出兩數的所有因數，再挑出所有兩數都有的因數，舉例來說30的因數有1,2,3,5,6,10,15,30,而33的因數有1,3,11,33,如此可知，1和3為30和33的公因數而2,5,...不是30和33的公因數，給我所有168和294可能的公因數



明白了，我们可以按照您的方式来找到168和294的所有公因数。首先，我们列出它们的因数：

对于168：

1, 2, 3, 4, 6, 7, 8, 12, 14, 21, 24, 28, 42, 56, 84, 168

对于294：

1, 2, 3, 6, 7, 14, 21, 42, 49, 98, 147, 294

现在我们挑出所有两者共有的因数：

1, 2, 3, 6, 7, 14, 21, 42

这些因数就是168和294的所有公因数。

- Constructing Step-by-Step Instructions:

Here I give the step-by-step instructions on the prompt, and the output is correct.

按照步驟進行:

1. 算出兩數的所有因數
  2. 看看有什麼數皆在兩者的因數裡面就是公因數
- 根據上述的步驟，給我所有168和294可能的因數



好的，让我按照您的步骤来进行：

1. 计算168和294的所有因数：

对于168：

1, 2, 3, 4, 6, 7, 8, 12, 14, 21, 24, 28, 42, 56, 84, 168

对于294：

1, 2, 3, 6, 7, 14, 21, 42, 49, 98, 147, 294

2. 找出两者公共的因数：

公因数有：1, 2, 3, 6, 7, 14, 21, 42

这些因数就是168和294的所有公因数。

- Link of this chat:

<https://chatgpt.com/share/6149a393-d05e-49c4-80a5-f9a6a39c0679>

## Question 2

- Question: Data Privacy and Security Concerns

- Motivation:

Modern individuals may use large language models (LLMs) to process various types of data, including sensitive information such as personal data, legal documents, and customer records. If someone maliciously attacks the model or its database, causing it to generate harmful information that users adopt, it could lead to serious consequences.

- Challenges

- Will an LLM generate responses containing malicious attacks according to the prompt?
- Can an LLM output confidential documents based on the prompt?

- Solution

- Implement robust encryption techniques to protect sensitive data both in transit and at rest.
- Utilize access control and identity authentication mechanisms to restrict unauthorized access to sensitive information.
- Employ data anonymization and pseudonymization techniques to minimize the risk of re-identification.
- Conduct regular security audits and risk assessments to identify vulnerabilities and ensure compliance with data protection regulations.
- Regularly update the model's training dataset or training methods to address new emerging attack techniques.

- Reference

- [人工智慧可能帶來哪些資安危機？](#)
- [Challenges Facing LLM Tools and Solutions](#)