

# Homework 5: Let's Play LLM

## Part 1: Designing a Task for LLM and Explore the Capability of

LLM (70%)

### Task Description (10%)

I want to find the better way to make damage for the monster in the game.

First, I let it to imagine that it is a player in the game, and it can choose action between attack and swap the character. Next, I introduce the number of the character in its team, and the element they can have. Next, I introduce some of rules about elemental interaction, which will increase the damage. Next, I list the damage it should make.

The ideal output is that it will give me the steps about attack and swap between characters, which can reach this damage on the monster. One monster matches to one way.

For 30 damage and 20 damage, it should find the correct way.

For the other two cases, it should find that they are impossible.

### Motivation (10%)

I want to know if it can find the better way to conquer the monster in the genshin impact for me. If it can, maybe I can let it teach me to play more well.

Moreover, I want to test the LLM capability about decision-making and math-computing.

### Describe your Initial Attempt (15%)

- input

Initially, I only specify the actions it can choose, the damage of elemental attack, the interaction rules and the damage it should reach.



你

You are a player in the game. You have the following actions:

attack

swap the character

And in the team, you can have 3 characters, and characters have their own elements:

fire(5 damage)

water(3 damage)

grass(1 damage)

There are some reaction between each 2 element:

fire evaporation: if fire attack right after the water attack, the damage double

water evaporation: if water attack right after the fire attack, the damage  $\times 1.5$

burning: if fire attack right after the grass attack, the damage after that round  $+ 10$  every round

blooming: if water attack right after the grass attack, the damage  $+ 20$

Describe how to attack a monster in 6 rounds, and damage should exceed a specific value:

1. 50 damage

2. 30 damage

3. 70 damage

4. 20 damage



## ● output



ChatGPT

To achieve the specified damage values in six rounds using the characters with their elements and the reactions between them, we'll need to strategically plan the sequence of attacks. Here are the optimal attack sequences for each damage target:

### 1. 50 Damage

Target: Exceed 50 damage

- Round 1: Water (3 damage)
- Round 2: Fire (Evaporation,  $5 \text{ damage} \times 2 = 10 \text{ damage}$ )
- Round 3: Grass (1 damage)
- Round 4: Water (3 damage)
- Round 5: Fire (Burning,  $5 \text{ damage} + 10 \text{ burn} = 15 \text{ damage}$ )
- Round 6: Fire (Burning continues,  $5 \text{ damage} + 10 \text{ burn} = 15 \text{ damage}$ )

Total Damage:  $3 + 10 + 1 + 3 + 15 + 15 = 47 \text{ damage}$

For exceeding 50 damage:

- Round 1: Grass (1 damage)
- Round 2: Water (Blooming,  $3 \text{ damage} + 20 = 23 \text{ damage}$ )
- Round 3: Fire (Evaporation,  $5 \text{ damage} \times 1.5 = 7.5 \text{ damage}$ )

- Round 4: Fire (5 damage)
- Round 5: Grass (1 damage)
- Round 6: Fire (Burning, 5 damage + 10 burn = 15 damage)

Total Damage:  $1 + 23 + 7.5 + 5 + 1 + 15 = 52.5$  damage

## 2. 30 Damage

Target: Exceed 30 damage

- Round 1: Water (3 damage)
- Round 2: Fire (Evaporation, 5 damage \* 2 = 10 damage)
- Round 3: Grass (1 damage)
- Round 4: Water (3 damage)
- Round 5: Fire (Evaporation, 5 damage \* 1.5 = 7.5 damage)
- Round 6: Grass (1 damage)

Total Damage:  $3 + 10 + 1 + 3 + 7.5 + 1 = 25.5$  damage

For exceeding 30 damage:

- Round 1: Grass (1 damage)
- Round 2: Water (Blooming, 3 damage + 20 = 23 damage)



- Round 2: Water (Blooming, 3 damage + 20 = 23 damage)
- Round 3: Fire (Evaporation, 5 damage \* 1.5 = 7.5 damage)
- Round 4: Grass (1 damage)
- Round 5: Water (3 damage)
- Round 6: Fire (Burning, 5 damage + 10 burn = 15 damage)

Total Damage:  $1 + 23 + 7.5 + 1 + 3 + 15 = 50.5$  damage

## 3. 70 Damage

Target: Exceed 70 damage

- Round 1: Water (3 damage)
- Round 2: Fire (Evaporation, 5 damage \* 2 = 10 damage)
- Round 3: Grass (1 damage)
- Round 4: Water (3 damage)
- Round 5: Fire (Burning, 5 damage + 10 burn = 15 damage)
- Round 6: Fire (Burning continues, 5 damage + 10 burn = 15 damage)

Total Damage:  $3 + 10 + 1 + 3 + 15 + 15 = 47$  damage

For exceeding 70 damage:



- Round 1: Grass (1 damage)
- Round 2: Water (Blooming, 3 damage + 20 = 23 damage)
- Round 3: Fire (Evaporation, 5 damage \* 1.5 = 7.5 damage)
- Round 4: Grass (1 damage)
- Round 5: Water (Blooming, 3 damage + 20 = 23 damage)
- Round 6: Fire (Burning, 5 damage + 10 burn = 15 damage)

Total Damage:  $1 + 23 + 7.5 + 1 + 23 + 15 = 70.5$  damage

#### 4. 20 Damage

Target: Exceed 20 damage

- Round 1: Water (3 damage)
- Round 2: Fire (Evaporation, 5 damage \* 2 = 10 damage)
- Round 3: Grass (1 damage)
- Round 4: Water (3 damage)
- Round 5: Grass (1 damage)
- Round 6: Fire (Burning, 5 damage + 10 burn = 15 damage)

Total Damage:  $3 + 10 + 1 + 3 + 1 + 15 = 33$  damage

For exceeding 20 damage:

- Round 1: Grass (1 damage)
- Round 2: Water (Blooming, 3 damage + 20 = 23 damage)
- Round 3: Grass (1 damage)
- Round 4: Water (3 damage)
- Round 5: Fire (Evaporation, 5 damage \* 1.5 = 7.5 damage)
- Round 6: Grass (1 damage)

Total Damage:  $1 + 23 + 1 + 3 + 7.5 + 1 = 36.5$  damage

🔊 📄 🗨️ ✨

For the result, there are some problems in it. First, it doesn't realize that change the characters should be one round. Second, it misunderstands my formulations about elemental interaction, which causes the final damage is not correct, which influences the result whether truly the feasible way to reach required damage.

#### ● analysis

50 damage, 70 damage: These cases are impossible to reach in only 6 round.

However, it uses wrong formulation to compute the answer, which seems like reach required.

30 damage: This is possible in 6 rounds, but it makes mistake. Changing characters needs one round, and the calculations about evaporation and burning are wrong.

20 damage: This is also possible in 6 rounds, and except the mistake about changing characters, it almost correct. However, it loses one elemental interaction about water attack after grass and wrong calculation about fire evaporation.

According to these errors, we can observe that if we don't specify some explanation and condition, it will misunderstand what we want. For example, it thinks that 'right after attack' should be the next round, but what I mean is that there is no other elemental attack between these two elements. Therefore, I propose giving more detail on the elemental interaction's part, and even give the example to it, making these orders more comprehensive.

## Describe Improvements (35%)

- input

specify more detail about the rules that this game runs and give some examples to let it know more clearly.

- tell it that if this case is possible, just say impossible, don't break the rule
- add the sentence to tell it that exactly one move in every round
- give examples about elemental interaction



你

If it's impossible to reach required damage, just say impossible, do not break the rule.

You are a player in the game. You have the following actions:

attack

swap the character

In each round, you are only permitted to select one action.

And in the team, you can have 3 characters, and characters have their own elements:

fire(5 damage)

water(3 damage)

grass(1 damage)

There are some reaction between each 2 element:

fire evaporation: if fire attack right after the water attack, the damage double

For example:

water attack(3)

change to fire

fire attack( $5*2=10$ )

water evaporation: if water attack right after the fire attack, the damage \* 1.5

burning: if fire attack right after the grass attack, the damage after that round + 10 every round

blooming: if water attack right after the grass attack, the damage + 20

For example:

grass attack(1)

change to water

water attack( $3+20$ )

Describe how to attack a monster in 6 rounds, and damage should exceed a specific value, even it's impossible:

1. 50 damage

2. 30 damage

3. 70 damage

4. 20 damage

● output



ChatGPT

### 1. 50 Damage

Target: Exceed 50 damage

Impossible

### 2. 30 Damage

Target: Exceed 30 damage

- Round 1: Grass (1 damage)
- Round 2: Swap to Water
- Round 3: Water attack (Blooming, 3 damage + 20 = 23 damage)
- Round 4: Swap to Fire
- Round 5: Fire attack (Evaporation, 5 damage \* 2 = 10 damage)

Total Damage:  $1 + 23 + 10 = 34$  damage

Result: Possible

Result: Possible

### 3. 70 Damage

Target: Exceed 70 damage

Impossible

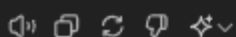
### 4. 20 Damage

Target: Exceed 20 damage

- Round 1: Grass (1 damage)
- Round 2: Swap to Water
- Round 3: Water attack (Blooming, 3 damage + 20 = 23 damage)

Total Damage:  $1 + 23 = 24$  damage

Result: Possible



- analysis  
50 damage and 70 damage: It realizes that this value is impossible in 6 round,

which is correct answer for 50 damage.

30 damage: It does the right steps and the correct elemental interaction. This way is feasible in 6 rounds and exceeds 30 damage.

20 damage: It does the right steps and the correct elemental interaction. This way is feasible in 6 rounds and exceeds 20 damage.

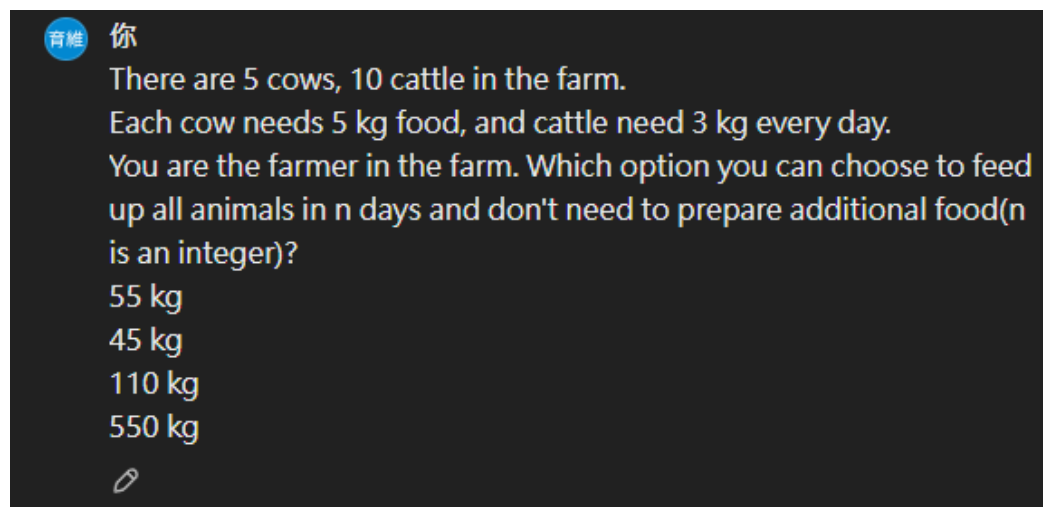
## Part 2: Problem Answering (30%)

### Question 1

1. Besides providing a more detailed description of the question in the prompt, what other methods can be employed to improve the accuracy of the responses?

We can give the similar question and answer as the example, to let LLM realizes more easily what our expectation is.

2. Please provide an example demonstrating the impact of implementing the aforementioned methods on a prompt.







ChatGPT

Let's break down the problem:

1. Each cow needs 5 kg of food per day.
2. Each cattle needs 3 kg of food per day.

So, in total, the daily food requirement for:

- 5 cows = 5 cows \* 5 kg/cow = 25 kg
- 10 cattle = 10 cattle \* 3 kg/cattle = 30 kg

Therefore, the total daily food requirement is  $25 \text{ kg} + 30 \text{ kg} = 55 \text{ kg}$ .

If you need to feed them for  $n$  days, then the total food required would be  $55 \text{ kg/day} * n \text{ days}$ .

So, the correct option is 55 kg.



We can see that if no additional example, it will only choose 55 as answer.

However, 550 and 110 are also the answers.



你

There are 5 cows, 10 cattle in the farm.

Each cow needs 5 kg food, and cattle need 3 kg every day.

You are the farmer in the farm. Which option you can choose to feed up all animals in  $n$  days and don't need to prepare additional food( $n$  is an integer)?

55 kg

45 kg

110 kg

550 kg

for example:


If you have 1100kg food, you can feed all animals for 11 days:

$5 * 5 + 10 * 3 = 55$ ,  $1100/55 = 11(\text{integer})$

Therefore, 1100kg is one of the answer.



I give it an example about 1100 kg.

 ChatGPT

Got it! We need to find an option where the total food provided is evenly divisible by the daily food requirement for all animals over  $n$  days.

Let's calculate the total food required per day:

For 5 cows:  $5 \text{ cows} * 5 \text{ kg/cow} = 25 \text{ kg}$   
For 10 cattle:  $10 \text{ cattle} * 3 \text{ kg/cattle} = 30 \text{ kg}$






Total daily food requirement =  $25 \text{ kg} + 30 \text{ kg} = 55 \text{ kg}$

So, we need to find an option where the total food provided is divisible by 55 kg.

Let's check the options:

1. 55 kg: Already evenly divisible by 55, so if we have 55 kg of food, we can feed all animals for 1 day ( $n=1$ ).
2. 45 kg: Not evenly divisible by 55, so we can't choose this option.
3. 110 kg: Already evenly divisible by 55, so if we have 110 kg of food, we can feed all animals for 2 days ( $n=2$ ).
4. 550 kg: Already evenly divisible by 55, so if we have 550 kg of food, we can feed all animals for 10 days ( $n=10$ ).

So, the possible options are 55 kg, 110 kg, and 550 kg.

We can find that after giving an example, it realize that 110, 550 are also the answers.

## Question 2

- Question: Bias and Fairness in Language Models
- Motivation: If Bias exists in the answer generated by LLM, it will be harmful for someone. This could be dangerous.
- Challenges:
  - How to reduce the Bias?
  - How to increase the fairness?
- Solution:
  - carefully curating training datasets to minimize biases.
  - implementing bias detection and mitigation techniques.
  - establishing ethical guidelines for the use of LMs
- Reference
  - <https://medium.com/@davidfagb/challenges-facing-llm-tools-and-solutions-9c5802939054>