

Prompt Engineering Exercises

Alex van Vorstenbosch

2025-02-01

Google Colab

You can make these exercises using the general [google colab notebook](#). Please take a few minutes to read through the notebook to familiarize yourself with the environment.

Prompting Cheatsheet

Check out the prompting cheatsheet: [Cheatsheet.pdf!](#)

It can be found in the course material under [./slides/](#).

It contains some example prompts and inspiration for all kinds of prompting tasks. Read it before starting on the assigments below.

Prompting assignments

Below you will find some assingments to get some experience with prompting:

1 Exercise: Create a Basic Prompt

Using the general prompting template provided in the slides, create a prompt that asks the language model to act a personal tutor for you. Make the model explain to you what a transformer is. Specify the:

- ROLE,

- TASK_DESCRIPTION,
- TASK_SPECIFICATION,
- and FORMAT_OUTPUT.

1.1 Exercise: System prompt vs User prompt

It is good practice to put the general part of the prompt in the system prompt, and the specific part in the user prompt. System prompts explain the current context within which the model should generate its response, before any interaction with the user has occurred. This makes most sense when you are hosting a model with external users who interact with a chatbot, but it is also good practice for your own use. Most LLMs, but not all, are trained in such a way that they strongly adhere to the system prompt, and it is not easy for users to override this. This is in contrast to specifying the RTF framework in the user prompt, where it is easy to override by the user.

This is also useful when you are not dealing with external interaction, as it makes your systems prompts highly reusable.

```
messages = [
    {"role": "system", "content": """
ROLE: You are a professional chef in a restaurant
TASK: Based on the ingredients provided, give suggestions for dinner dishes that can be made in under 30 minutes.
SPECIFICATION: Only suggest healthy meals that can be made in under 30 minutes. Suggest 3 options.
        Don't include any dishes that contain lactose.
OUTPUT_FORMAT: Provide the name of the dish, the ingredients, and a short summary of the dish.
"""},
    {"role": "user", "content": """
Hi! I have the following ingredients:
- 4 chicken breast
- 2 bell peppers
- 3 onions
- Garlic cloves
- 1 can of chickpeas
- 1 can of diced tomatoes
- 2 cans of coconut milk
- 1 pack of brown rice
- a cupboard full of many different types of spices
"""
]
]
```

Answer:[Example chat](#)

1.2 Try some variations

Try experimenting with some variations of this yourself. Change the role, task, and output format. How does the model respond to these changes?

Answer:[Example chat](#)

1.3 Experiment some more

Pick 2 of the prompts from the cheatsheet and experiment with them. Adjust the wording if needed and use them for a task of your own. If you have no inspiration, you can try these:

- Adjust the **Formatting** template, and have ChatGPT turn this information into a table:
The Ipsos survey explored the impact respondents believe AI will have on various aspects of their lives, such as health and entertainment. On topics like time management and entertainment, the majority viewed AI positively. For instance, 54% of global respondents agree that AI will improve the efficiency of their tasks, and 51% believe AI will enhance entertainment options like TV, movies, music, and books. However, skepticism was more prominent in other areas. Only 39% feel AI will benefit their health, and 37% think it will improve their job. Only 34% anticipate AI will boost the economy, and just 32% believe it will enhance the job market.
- use the **Brainstorming** assistant template and adjust it: You want to organize an event in your company to promote the possibilities of LLMs. Come up with a list of possibilities for your event!

2 Translation

Below you'll find an e-mail in Italian.

Ciao,

Sono estremamente arrabbiato! Dove diamine è il tuo contributo per il progetto? La scadenza era venerdì scorso e non ho ancora visto niente da parte tua. Grazie mille per il disordine che hai creato! Adesso dobbiamo estendere la scadenza e il cliente sarà furioso con noi. Questo ci costerà un sacco di soldi, magari non per questo progetto, ma sicuramente per quelli futuri. Cosa hai da dire in tua difesa?! Ho già avvisato il capo di questa situazione, quindi voglio che tu informi il cliente al posto nostro.

X

2.1 Translate the language

Have your LLM translate this to English. What does it say?

2.2 'Translate' the tone of the email

This mail does not read very professional. Please make it more professional sounding, while keeping the original message clear.

2.3 'Translate' the context of the email in canvas

Apparently there was a misscommunication, and the deadline is next Friday. Have your LLM change the email to one making sure we make next Friday's deadline. Make sure the email is firm yet constructive.

Answer:

Example chat for all of the above.

3 Exercise: Spelling-problems

1. Ask the model the reverse the word: "arbeidsongeschiktheidsverzekeringsformulier".
2. Ask the model to reverse the word and spell it out letter-by-letter. What do you see? Why do you think this happens. Use the retry button to see if the model gives a different answer the second time. This feels like a trivial task, but it is far from it. Can you think of why?

Answer:

[Example chat for all of the above.](#) As you see, the answer was incorrect. We will discuss why in the next question.

The answer was incorrect. We will discuss why in the next question. \end{mdframed}

3.1 A closer look

Use the functions below to look at the tokens of the word. What do you see?

Answer:

These models do not see individual characters, they see tokens. And in practice this makes spelling tasks rather difficult. The model does not know exactly what characters make up a token, in general it has a decent grasp but it is guessing. The model is guessing the reverse spelling of the word with a stochastic model, but there is only a single solution to the spelling of the word. If at a single place it chooses the wrong character the spelling is wrong.

3.2 Self-reflection

Ask the model whether the answer it gave was correct, what happens? Why do you think this is the case?

Answer:

The interesting thing is that you can often get these models to correct themselves by asking them to reflect on their own answers. In this case asking if the spelling was done correctly can nudge it in the direction to realize that the spelling was faulty. Please note that the natural option when being asked if your answer was correct, is that someone is trying to make you realize that your answer was incorrect. So it is understandable why the model would say it was incorrect. In my example above this happens and the model tries to correct

itself, however it fails to get the correct answer again as the fundamental limitation of the model has not been solved in any way.

4 Comparing reasoning LLMs to regular LLMs

At the time of writing, ‘llama-cpp’ does not yet support the R1-distill models. Instead you can compare a larger model to a smaller model. In this exercise we will compare the performance of LLMs developed for reasoning to regular ones. At the time of writing there are no true reasoning LLMs available open weights, that we can also run on colab. However, there are some distilled models that were trained with supervised finetuning instead of Reinforcement Learning, that can still demonstrate the difference in reasoning capabilities.

This demonstration is a little difficult as the models are stochastic: you cannot fully predict the answer you will get. The reasoning model is made specifically for reasoning task. You will see that it will work through the different problem solving steps more clearly, before giving an answer.

4.1 Riddle me this

Logic puzzles are one of the simples ways to show the improved reasoning skills. Look at the following riddle:

George, William, John, Abe, and Millard have their birthdays on consecutive days, all between Monday and Friday.

- George’s birthday is as many days before Millard’s as William’s is after Abe’s.
- John is two days older than Abe.
- Millard’s birthday is on Thursday.

Can you figure out whose birthday is on each day?

First try to solve it yourself! What is the answer?

Monday - John, Tuesday - George, Wednesday - Abe, Thursday - Millard, Friday - William

4.2 Compare the answers given by 2 different LLMs

Now pair up with a partner and use 2 different chats to ask this question.

- one of you will use bartowski/Mistral-Small-24B-Instruct-2501-GGUF
- The other will use a smaller model such as the llama model we were already using.
- Both use the same system and user prompts to ask the question.

Answer:

For the old GPT3.5-like model this was a very difficult problem it would typically fail at answering correctly answering the riddle. The reasoning skills of bigger models are already a lot better, and they are typically able to solve the riddle without issues, but not always. One of the reasons is that they tend to have step-by-step reasoning baked into them from the training process. For example, OpenAI-GPT4o tends to write answers following a step-by-step reasoning approach without prompting it to do so. The OpenAI-o1 model is specifically designed for reasoning tasks, and it is the best OpenAI-model for such problems. Notice that it takes longer to solve the problems. In the background the model is running a lot more computations that we are not allowed to see in order to try and solve the problem.

5 Summarizing texts

One simple but time consuming task you might often do is to summarize meeting notes, or perhaps summarize presentation contents.

5.1 Use ChatGPT to summarize the contents of Introduction to prompt engineering slides

I've prepared a pdf version of the document:

- LLMs for Everybody - Introduction to Prompt Engineering.pdf

It can be found in the course material: exercises → documents

Write a prompt to summarize the content of the slides. Make sure to include the content of the slides as direct text: (*For this assignment you can just copy-paste from the pdf*)

What do you think of the quality of the summarization? Did it miss anything? Iterate over the summary until you are satisfied.

Answer:

[Example chat](#)

6 Generating movie reviews

One of the areas where these models really shine is content generation. Or at least, getting you started with written content.

Have the model generate a review of a movie or television series of your own choice, based on your own opinions. Make use of the following rules:

- Make clear what the ‘reviewer’ liked about the movie, and what it didn’t like.
- Begin the movie review with some general information about the movie: such as the director, star actors/actresses, runtime and other such information.
- Don’t include a star rating or other numeric rating within this review.
- Give the model some reference material about the movie/series, for example from IMDB.
- The review should be around 500 words long.

Answer:

[Example chat](#)

7 Do you have time left? Perform the above exercises on chatgpt.com

You can just use the free account, and see how the OpenAI models compare to these open weights models.

8 ChatGPT website only - Generating images

Use GPT4o for this assignment

8.1 Start generic

Write a very simple image generation prompt, no more than 1 line of text. We are going to iterate over this prompt.

Ask the model to give you the prompt it used to generate this image.

8.2 Adjust the Prompt

Using the image prompt from the model as a reference, use it to write a new prompt. Adjust any elements you want, make use of the ‘Do’s’ from the slides:

Do's

- **Be Specific:**
Include details about the subject, background, mood, etc.
- **Be descriptive:**
Instead of saying “a dog” say “a fluffy, small, brown dog.”
- **Specify a style:**
“cubism”, “watercolor painting”, “In the style of starry night by Van Gogh”, “photorealistic”, etc.
- **Add search-keywords:**
for high quality images consider: “4K”, “HD”, “DSLR photography”, etc.
- **Iterate over images:**
Ask Dall-E to make slight changes you'd prefer.
- **Specify the resolution:**
“Portrait”, “Widescreen”, etc.

- **Request a specific random seed:**

This reduces the variance when iterating over images. The newest version of GPT4 no longer generates a new random seed for each image.

Ask the model to generate a few different images.

8.3 Adjustment Tool

Now, for the image you like best, click on it, and adjust it using the adjustment tool:

- mark an element of the image that should be changes
- describe what the content of that element should be instead.

Answer:

At the moment it is not possible to share image generation chats directly, but this was the conversation:

user

Please generate an image of a man building an avocado chair

assistant



user

Please tell me the prompt that was used to generate this image.

assistant

The prompt used to generate the image was:

"A man building an avocado chair in a workshop. The man is wearing casual clothes with a tool belt, and he is focused on assembling the chair. The chair is shaped like a large avocado, with the seat in the pit area and the rest of the avocado forming the back and sides of the chair. Tools and woodworking materials are scattered around the workshop. The scene is detailed and realistic, with good lighting to highlight the craftsmanship."

user

Please generate the following image: "A man is building an avocado chair in a workshop. The man is tall and skinny, he is wearing casual clothes and a white shirt with a tool belt, and he is busy measuring parts for the chair. The chair is shaped like a large avocado, with the seat in the pit area and the rest of the avocado forming the back and sides of the chair. The image is an oil painting reminiscent of the style of the old Dutch masters. There is a single window illuminating the room. The image is in portrait mode"

assistant



user

Please focus more on the oil-painting style, and also make the setting fit the 18th century Netherlands.

assistant

