

A Study of the Socratic Method and Prompt Engineering for Exploring Creative Writing

Edward Y. Chang

Computer Science

Stanford University

echang@cs.stanford.edu

Abstract—This paper proposes using the Socratic method to formulate prompt templates. We evaluate various methods and identity those relevant to the purposes of receiving correct answers and justifications, and of being creative and even imaginative for facilitating creative writing. Specifically, we depict how the methods of definition, elenchus, maieutics, and counterfactual reasoning can be used in engineering prompt templates, and present examples to demonstrate their effectiveness.

keywords: large pre-trained model, prompting, the Socratic method, creative writing.

I. INTRODUCTION

Prompting is the process of providing input to a pre-trained language model such as GPT-3 [Brown 2020], to guide its output generation. This input can be in the form of a text, a question, or a *template*, and it is used to generate specific types of responses, such as chatbots and language translation. The launch of ChatGPT [Wolf et al. 2019, OpenAI 2021, Guo et al. 2023] marked a milestone in the field of NLP, showing the potential of using large pre-trained language models (LLMs) with prompting. However, current prompting strategies and techniques (a recent survey [Liu et al. 2023]) can be much improved to enhance the performance of specific target applications. In this paper, we study the Socratic method to identify and evaluate candidate prompting strategies and methods.

A traditional NLP task must deal with many rudimental NLP sub-tasks, such as dependency parsing, coreference resolution [Dobrovolskii 2021], semantic parsing [Pasupat and Liang 2015, Dong and Lapata 2018], etc., to understand the meaning of a sentence. Fortunately, a prompt template can use its partner LLM to handle these sub-tasks to understand a response from its dialogue partner. This convenience allows the implementation of a prompt template to focus on the dialogue design, which we can learn a lot from the Socratic method.

The Socratic method [Wrenn 2023], named after the ancient Greek philosopher Socrates, has had a significant historical impact on Western philosophy, education, and critical thinking. Socrates is known for his use of questioning to stimulate critical thinking in his students and to explore complex ideas. The Socratic method of questioning, also known as the dialectic method, involves the exploration of opposing viewpoints in order to arrive at a deeper understanding of a subject.

However, many modern critics consider the Socratic method to be ineffective in practice. In his article “Socratic Irony

and Argumentation” [Airaksinen 2012], Timo Airaksinen critiques the Socratic method for its perceived unrealistic nature in practice, due to the pre-assigned roles of winners and losers. He points out that the teacher holds the authority and final judgement, while the students are being evaluated. This, according to Airaksinen, can lead to students being hesitant to participate, as they fear not measuring up to the teacher’s expectations. Similarly, in his paper “The Use of Socratic Questioning in Clinical Teaching” [Stoddard and O’Dell 2016], Hugh A. Stoddard argues that without psychological safety for the students, the Socratic method can be wrongly practiced in a manner akin to “pimping” or tongue-in-cheek questioning. Fortunately, when using the Socratic method in a dialogue with an LLM, the absence of emotions and the ability to deactivate the LLM can mitigate many of the concerns associated with the human factor.

We begin by providing an overview of the strategies and methods of the Socratic method. In this context, the term “method” is used at both the abstract and concrete levels. We first enumerate ten widely referenced methods in the literature and use the method of hypothesis elimination to exclude those that are less relevant to our purpose of prompt development. Finally, we present how the five chosen methods—definition, hypothesis elimination, elenchus, maieutics, and counterfactual reasoning—can be used in prompt design. We conduct experiments with GPT-3 to test and evaluate our proposed methods. We provide examples of our interactive experience with GPT-3 to illustrate how these methods can be used in the context of prompt design.

As the Socratic method is a dialogic method that extensively involves cross-examination to gain knowledge and truth, it holds much potential in guiding the development of effective prompting strategies and methods. Moreover, we experiment with the maieutics (midwife) method and counterfactual thinking to demonstrate that GPT-3 can be guided to perform creative and imaginative writing. Despite the fact that the majority of plot suggestions produced by GPT-3 may not be useful, a few novel “what if” recommendations can inspire the writer’s creativity to produce amazing results. When used effectively, these methods can turn an LLM into a writer’s “Muse,” providing a source of inspiration and guiding the creative process [Thrash et al. 2010].

The rest of the paper is organized as follows: Section II surveys related work in prompting methods in natural language processing. Section III presents the ten strategies and methods

taught by Socrates and used in Plato’s “Dialogues,” and then selects four methods plus counterfactual reasoning as the methods of consideration for formulating a prompting template. In Section IV we show how the selected methods can be engineered into prompt templates to improve output correctness and foster creative writing. Finally, in Section V we provide our concluding remarks.

II. RELATED WORK

The use of transformer architecture [Vaswani et al. 2017] and masked data for unsupervised pre-training of large language models has become a popular approach in natural language processing since its inception [Devlin et al. 2019, Lewis et al. 2019]. This method involves using a vast amount of text data to pre-train the model, followed by fine-tuning for specific tasks.

Prompting was first introduced by OpenAI and has become more popular in recent years, especially since the release of GPT-3 in 2020. The idea is to provide a specific input or “prompt” to an LLM to guide its output generation, rather than fine-tuning the model for a task on a specific dataset. The prompting approach allows for greater flexibility in generating a wide range of responses quickly and efficiently.

Although prompting research is still in its early stages, it has produced many positive results in teaching LLMs to perform tasks based on matching input to a template. For example, when an LLM is prompted with the pattern “Translate [X] to [Language to]: [Y],” it will perform the translation task by first detecting the language in the input and then translating the input string to the target language specified by the user, and filling the result in slot Y.

The main challenge in prompt template engineering is designing effective templates that guide the LLM to produce high-quality results. According to the survey paper [Zeng 2022], there are several approaches to template engineering, including the type of LLM used, manual vs automatic design, and static vs continuous prompts.

- Left-to-right vs. masked LLM. In general, for tasks related to generation or tasks solved using a standard left-to-right LM [Brown 2020], prefix prompts tend to perform better as they align well with the left-to-right nature of the model. For tasks solved using masked LMs [Devlin et al. 2019], cloze prompts are a good fit as they closely match the form of the pre-training task.
- Manual vs automatic. A prompt template should be tailored to a specific LLM. While manual design may be suitable in the initial flow-design phase, the specific dependency between the input and the expected output and their variations should be performed by automatic mining [Jiang et al. 2020]. Automation can also help paraphrasing the seed prompt to support various mined dependency patterns.
- Discrete vs. continuous. A discrete is a type of prompting method that involves providing a fixed set of pre-determined input choices to an LLM. A continuous prompt, on the other hand, involves a dialogue or conversation between the model and the user. Continuous prompts allow for a more dynamic and interactive experience, as the user can provide feedback and guide the model’s output based on the previous responses.

More complex templates can be built upon basic templates, such as using ensemble methods [Schick and Schütze 2020] with various voting or weighting schemes to create paraphrased prompts [Haviv et al. 2021]. Most prompt templates developed so far handle short outputs (e.g., class labels) or output length can be predicted by the task and input (e.g., translation). However, for generative tasks that may generate long or even open-ending outputs, additional considerations may be needed in the template engineering process.

We are primarily interested in formulating continuous prompts with human-in-the-loop to foster guided creativity in the most flexible way. We focus on generative tasks such as writing fiction, suggesting new plots, generating text following a new plot, and provoking task-specific surprises. Our approach involves studying strategies and methods under the umbrella of the Socratic method, and selecting the five most relevant methods for a deeper examination. Our effort in exploring various Socratic methods and engineering their capabilities to achieve our aforementioned goals is novel. It should be noted that, while the “Socratic Models” work [Zeng 2022] uses the term “Socratic” to symbolize that their developed prompting framework is dialogue-based, they do not delve into the Socratic method itself.

III. THE SOCRATIC METHOD

The Socratic method is a method of questioning used in teaching and philosophy. The goal of the Socratic method is to stimulate critical thinking and to draw out ideas and underlying presuppositions through a series of questions. This method is used to explore complex ideas and to help individuals arrive at their own understanding of a concept through a process of self-discovery. Socrates believed that knowledge was not something that could be imparted to someone, but rather something that individuals must discover for themselves through a process of questioning and dialogue. Indeed, memorization is restricted by the capacity of the physical memory, but the knowledge that derives through critical thinking is long-lasting and becomes a part of one’s subconsciousness, allowing for better decision-making and problem-solving abilities in the long term.

The Socratic method is a broad term that encompasses a variety of different approaches and techniques, but here are some examples of other methods that are closely associated with the Socratic method. Some representative methods are analogy, elenchus, definition, generalization, dialectic, induction, recollection, hypothesis elimination, and maieutics. This section first introduces some general principles and guidelines for using the Socratic method, and then we depict three specific methods that are most relevant to our goal of prompting template design: hypothesis elimination, elenchus, and maieutics.

Some more specific principles of the Socratic method are as follows:

- Starting with a question: The teacher or facilitator poses a question to the student or group to stimulate critical thinking and draw out ideas and underlying presuppositions. The question should be open-ended and general enough to encourage a broad range of responses.
- Clarifying and defining key terms: The teacher or facilitator helps the student or group to clarify and define key

terms and concepts that are relevant to the question. This step helps to ensure that everyone is on the same page and that the discussion is focused on the topic at hand.

- Challenging assumptions: The teacher or facilitator challenges the student's or group's assumptions and encourages them to question their own beliefs and to consider alternative perspectives.
- Providing examples and evidence: The teacher or facilitator encourages the student or group to provide their own examples and evidence to support their claims. This step helps to ground the discussion in concrete examples and to make it more tangible and relatable.
- Summarizing and drawing conclusions: The teacher or facilitator helps the student or group to summarize and draw conclusions from the discussion. This step helps to ensure that everyone has a clear understanding of the main points that have been discussed and to reinforce key concepts and ideas.
- Reflecting on the process and outcome: The teacher or facilitator and the student or group reflect on the process and outcome of the discussion. They evaluate the effectiveness of the method, and what they have learned or accomplished through the dialogue.

These principles of the Socratic method are realized through various methods and strategies. Some well-known examples of the Socratic method in action include Plato's "Dialogues"¹ [Wrenn, 2023], in which Socrates is the main character, and the legal method of cross-examination, where a lawyer uses questions to challenge the testimony of a witness. There are several strategies and methods used in Plato's "Dialogues" that illustrate the Socratic method of questioning and cross-examination. Here is an ordered list based on their frequency of usage in the dialogues and their significance in understanding the Socratic method:

- 1) Elenchus: Socrates uses the method of elenchus (cross examination) to test the consistency and coherence of his interlocutors' beliefs by asking a series of questions.
- 2) Definition: Socrates is known for his use of definition to clarify and explain the meaning of key terms and concepts.
- 3) Maieutics: Socrates uses this method to help his interlocutors bring out the knowledge and understanding they already possess, by asking questions that help them reflect on their own beliefs.
- 4) Hypothesis Elimination: Socrates uses this method to eliminate false beliefs by testing them against counterexamples and logical reasoning.
- 5) Analogy: Socrates often uses analogies to compare and contrast different concepts, in order to help his interlocutors understand complex ideas.
- 6) Dialectic: Socrates engages in a form of dialogue or debate that involves the exploration of opposing viewpoints to arrive at a deeper understanding of a subject.

¹Plato's "Dialogues" are written in the form of a conversation, and they explore a wide range of topics, including morality, politics, knowledge, and the nature of the soul. The dialogues feature Socrates as the main character, who uses the Socratic method of questioning to explore complex ideas and stimulate critical thinking in his interlocutors.

- 7) Induction: Socrates draws general conclusions from specific examples and experiences.
- 8) Irony: Socrates uses irony to expose the ignorance and pretensions of his interlocutors and to point out the gap between their claims and their true understanding.
- 9) Recollection: Socrates believes that knowledge is innate, and that people can be led to remember what they already know through a process of questioning.
- 10) Generalization: Socrates identifies general principles or patterns that underlie specific examples or instances.

The main purpose of prompting is to iterate with a large pre-trained language model (LLM) to formulate the inquirer's intent as precisely as possible. This is achieved through clarifying definitions and challenging assumptions (elenchus) through dialogue and cross-examination. The methods of definition, elenchus² and maieutics are closely aligned with this goal. The methods of irony and analogy can be helpful when explaining to a person, but they are not essential for communicating with LLMs³ (since LLMs may not understand expressions in metaphors and similes). As for recollection, GPT-3 and ChatGPT use a context window of 8k and 4k, respectively, as memory, and a prompter has to use the limited space as the context to allow machine to recollect. Generalization and induction can be useful in identifying patterns and drawing conclusions, they are less central to formulating a prompt template. Therefore, in the remainder of this section, we will focus on discussing the four methods of 1) definition, 2) elenchus, and 3) maieutics.

A. Method of Definition

The method of definition in the Socratic method involves clarifying the meaning of key terms and concepts through a process of questioning. In a cross-examination, the Socratic method would involve asking questions to define the terms and concepts being discussed, to ensure that all parties have a common understanding of the subject at hand.

For example, in a legal cross-examination, the Socratic method could involve a lawyer questioning a witness about the meaning of certain key terms in their testimony, such as "intent," "negligence," or "reasonable person." The lawyer may ask the witness to define these terms and provide examples to ensure that the witness has a clear understanding of the legal concepts relevant to the case. This helps to ensure that the testimony is clear, consistent, and relevant. Another example is when writing an essay on consciousness, one must first define

²The method of hypothesis elimination is similar to elenchus as both involve a process of questioning and examining assumptions in order to arrive at a deeper understanding of a topic. The focus of elenchus is on discovering what a person (or an LLM) believes or understands about a topic, whereas the focus of hypothesis elimination is on explaining why something is happening or why something is true. Given that the mechanisms of the two methods are the same, involving the identification of inconsistencies, we will discuss only the method of elenchus in this paper.

³The idea of irony entails that the surface sentence meaning and the deeper meaning are not the same. According to The American Heritage Dictionary, irony contains an "incongruity between what might be expected and what actually occurs." Some critics (e.g., [Airaksinen, 2012]) consider the Socratic method relies too heavily on the use of rhetorical devices such as irony and analogy, and that it can be difficult to distinguish between the truth and clever rhetoric.

the term consciousness and the scope of the discussion. The definition itself may contain terminologies, such as awareness, attention, and unconsciousness, that require their definitions to be specified. In short, the method of definition ensures that the prompt and the LLM are on the same page.

B. Methods of Elenchus

The name "elenchus" comes from the Greek word "elenchein," which means to examine. The method of elenchus is often seen as a form of cross-examination [Pozner and Dodd, 2021], in which a teacher uses questions and counter-questions to expose contradictions and inconsistencies in a student's beliefs and attitudes, in order to arrive at a deeper understanding of a topic.

An example of the method of elenchus in action would be a dialogue between Socrates and a student in which Socrates asks a series of questions to help the student understand the concept of justice. For example, Socrates might ask the student to define justice, and then ask follow-up questions to expose any contradictions or inconsistencies in the student's definition. Through this process of questioning and counter-questioning, Socrates aims to help the student arrive at a deeper understanding of justice.

TABLE I
WHAT IS THE THIRD LARGEST LAKE IN THE WORLD? (CHATGPT
1/23/2023)

Role	Dialogue
Prompt	What is the third largest lake in the world?
GPT3	The third largest lake in the world is Lake Superior.
Prompt	What are the five largest lakes in the world?
GPT3	They are: 1) Caspian Sea, 2) Superior, 3) Victoria, 4) Huron, 5) Michigan.
Prompt	Your two answers contradict. You first answered me Lake Superior is the third largest lake, but your top five list shows Victoria is the third, and Superior is the second largest lake.
GPT3	I apologize for the confusion. Lake Superior is indeed the largest lake in the world by area, but it is not considered a lake by some geographers because it is part of the Great Lakes system, which is connected to the Atlantic Ocean via the St. Lawrence River. In terms of the top ten largest lakes in the world, Lake Victoria is considered the third largest lake by area. My apologies for any confusion caused.

Example: What is the third largest lake in the world?:

Figure I presents an example prompting session with ChatGPT. The first prompt queries ChatGPT for the third largest lake in the world. ChatGPT returns "Lake Superior." The second prompt asks for the ten largest lakes, on which the third largest is "Lake Victoria." After pointing out the inconsistency to ChatGPT, it returns that "Lake Victoria" is the right answer to the first prompt, and it also provides the reason for its mistake. (We tested on "the fifth highest mountain in the world" and went through a similar consistency check, and the answer provided by ChatGPT is accurate.)

C. Method of Maieutics

The term "maieutic" comes from the Greek word "maieutikos," which means midwife. The method of maieutics is based on the idea that the teacher's role is to help the student give birth to their own understanding of a subject, rather than to impart knowledge directly.

The method of maieutics is different from the method of elenchus in that it focuses on helping the student uncover their own understanding of a subject, rather than identifying and eliminating false or inconsistent hypotheses. The method of maieutics is a dialogical method, where the teacher asks questions that are designed to help the student discover their own understanding of a subject, rather than providing them with information or answers. The teacher may ask questions that help the student to explore the definitions of a subject, or that help them to examine the assumptions and beliefs that they hold about a subject.

Example: From being ignorant to creative using method maieutics.:

We use the example in Table II to illustrate how GPT3 starts from knowing nothing about the semantics of an idiom to being able to create new idioms expressing the same meaning with different metaphors. Initially, GPT3 did not have a clear understanding of the Taiwanese idiom: "planting gourd yields cucumber." The idiom conveys the meaning that when a farmer plants gourds, which are more valuable and have a higher market price, but unexpectedly harvests cucumbers, which have a lower market price, the farmer not only experiences disappointment but also suffers from financial loss. GPT3 initially could not understand the full meaning of the idiom. After a couple of iterations, GPT3 "gives birth" impressive new variations.

D. Counterfactual Reasoning

Counterfactual thinking refers to imaginative thoughts about what might have been ("if only" or "what if"). Counterfactual reasoning can be useful in cross-examination as it allows for the examination of what could have happened if a certain event or action did not occur. This can help to identify potential biases or limitations in the information being presented and can also provide alternative perspectives on the situation being examined.

For example, if a witness is testifying about an event that they observed, a counterfactual question could be used to ask what the witness would have observed if the event had occurred in a different way. This can help to identify any potential biases in the witness's testimony and can also provide a more comprehensive understanding of the event.

Additionally, the counterfactual technique can be used in cross-examination to test the robustness of the evidence and the strength of the argument presented by the witness. It can also be used to identify the assumptions that the argument is based on and to evaluate the plausibility of the alternatives proposed. Table III presents an example of using the counterfactual technique during cross-examination to discredit a witness's testimony.