**Socratic Prompting with Argo: Assessment**

Bratin Sengupta (AMD, Functional Coating Group)

**Original Prompt:**

How does zwitterions make a PVDF membrane antifouling?

**Socratic Reformulation:**

For what reasons a zwitterion modified PVDF membrane becomes antifouling in nature?

|  |  |
| --- | --- |
| **Principles** | **(Follow-Up) Prompts** |
| **Generalization** | **Follow-up Prompt 1:**  “What are the properties that makes a membrane antifouling in nature?” |
| **Maieutic** | **Follow-up Prompt 2:**  “How does surface hydrophilicity and surface charge effect antifouling behavior of a PVDF membrane?” |
| **Elenchus** | **Follow-up Prompt 3:**  “Is there any example of an antifouling membrane with low hydrophilicity and low surface? ” |
| **Maieutic** | **Follow-up Prompt 4:**  “How does zwitterions make PVDF membranes antifouling?” |

1. **Motivation**

*Why is this problem important in chemistry/materials science?*

Membranes are used for separation in wide range of applications – from water treatment to biotech industries to petroleum production. The inherent problem of these membranes is that when they are used for a long time, solutes accumulate on its surface and pores and that eventually reduces the throughput of these membranes. Considerable research in materials science and chemistry have been done to mitigate this problem, with little success of slowing down the problem, but not entirely eliminating it.

*What challenges exist in solving this problem using traditional LLM approaches?*

The information available in literature are often inconsistent and several arguments contradicts each other on the nature of the fouling problem and its mitigation strategies. This is amplified with the system specific nature of the problem – as different membranes fouls differently and under different operating conditions. This makes its extremely hard for traditional LLMS to provide a correct answer as it solely relies on existing data.

*How could a structured Socratic approach improve reasoning and outcomes?*

Socratic approach of prompting, especially for models which incorporate chain of thought should allow for intertwining the scattered information into a meaningful content. The Socratic prompting may allow the model to “re-think” and “re-evaluate” its understanding about the specific topic to “weed” out garbage information regarding the topic and possibly provide more accurate, less verbose answers.

1. **The Mixed Socratic Prompt Method Used**

To breakdown this problem statement, I assumed the prompter to be young graduate student who seeks help of a foundation model to gain basic understanding of antifouling membranes and then wishes to learn the techniques to fabricate one using zwitterions. I modeled the prompting as a conversation between this hypothetical user (grad student) and a veteran in the field (say, their PI). I started the initial conversation with a ***dialectic*** prompt and followed by ***generalization***. This allowed the model to step back and understand the larger picture of the conversation. Following this, I used a prompt based on ***maieutic***. This allowed the model to possibly critically think (and establish a chain of thought) and reevaluate its outputs. Based on this refined understanding, I cross examined the model to give a more specific examples for counterintuitive situations using the principle of ***elenchus***. Finally, I prompted an ***induction****-*based question to revert back to the original question – “How does zwitterions make PVDF membranes antifouling?” I then compared this final output of the model prompted using Socratic reformulation to the one where the prompt was direct and without any prior context.

1. **What Are the Prompts Used?**

***Original Prompt:*** How does zwitterions make a PVDF membrane antifouling?

***Socratic Reformulation:*** For what reasons a zwitterion modified PVDF membrane becomes antifouling in nature?

1. **What Are the Outcomes of This Example?**

*How did the LLM refine its answers over iterations?*

Over nitration, the Socratic method showed higher improvement, compared to the direct method of prompting.

*What key insights or discoveries emerged?*

Responses from Socratic method of questioning is more relevant to the prompt, while in the direct approach, the model response was more general and encompasses a lot of definition of new/uncommon (i.e., the scientific ones) terms that appears in the prompt. However, in terms of the facts that the model generates in either case was equivalently accurate. There is no hallucination in one case compared to the other. However, it was surprising to see how the Socratic prompting proved to generate more relevant answers.

*Any unexpected results or challenges?*

No

1. **Comparison to a Non-Socratic Approach**

*How did reasoning depth, self-correction, and hypothesis refinement compare?*

The depth in Socratic prompting is higher, along with more relevancy.

*Would a traditional direct-answer prompt have produced different results?*

It produces more general response, and often digress off topic.

*Did the Socratic method improve clarity, adaptability, or accuracy?*

The Socratic method improved clarity and improved depth of the answer and provide more relevant answers.