Automatic Prompt Engineer (APE)

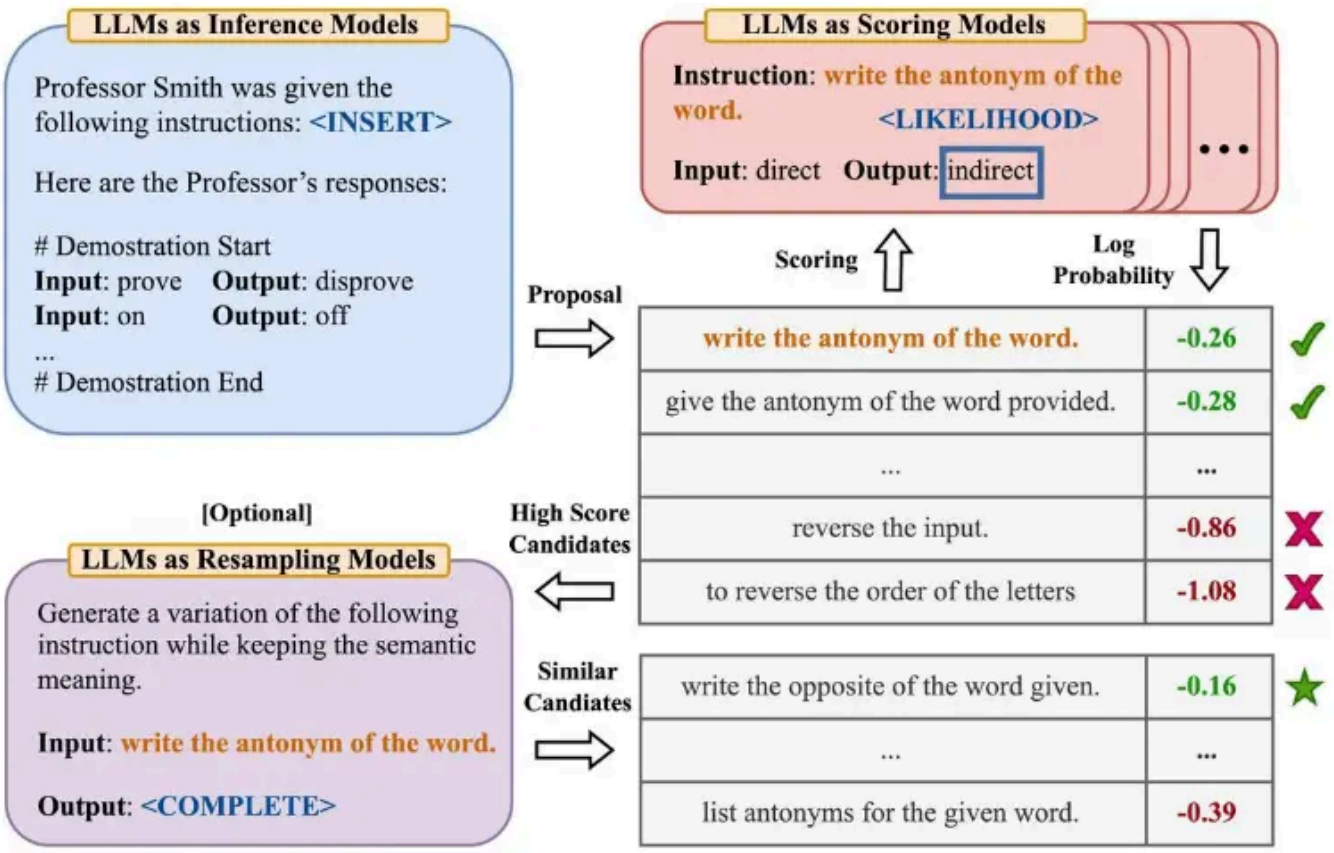

Image Source: Zhou et al., (2022)

Zhou et al., (2022) propose automatic prompt engineer (APE) a framework for automatic instruction generation and selection. The instruction generation problem is framed as natural language synthesis addressed as a black-box optimization problem using LLMs to generate and search over candidate solutions.

The first step involves a large language model (as an inference model) that is given output demonstrations to generate instruction candidates for a task. These candidate solutions will guide the search procedure. The instructions are executed using a target model, and then the most appropriate instruction is selected based on computed evaluation scores.

APE discovers a better zero-shot CoT prompt than the human engineered "Let's think step by step" prompt (Kojima et al., 2022).

The prompt "Let's work this out in a step by step way to be sure we have the right answer." elicits chain-of-thought reasoning and improves performance on the MultiArith and GSM8K benchmarks:

No.	Category	Zero-shot CoT Trigger Prompt	Accuracy
1	APE	Let's work this out in a step by step way to be sure we have the right answer.	**82.0**
2	Human-Designed	Let's think step by step. (*1)	78.7
3		First, (*2)	77.3
4		Let's think about this logically.	74.5
5		Let's solve this problem by splitting it into steps. (*3)	72.2
6		Let's be realistic and think step by step.	70.8
7		Let's think like a detective step by step.	70.3
8		Let's think	57.5
9		Before we dive into the answer,	55.7
10		The answer is after the proof.	45.7
-		(Zero-shot)	17.7

Image Source: Zhou et al., (2022)

This paper touches on an important topic related to prompt engineering which is the idea of automatically optimizing prompts. While we don't go deep into this topic in this guide, here are a few key papers if you are interested in the topic:

- Prompt-OIRL - proposes to use offline inverse reinforcement learning to generate query-dependent prompts.
- OPRO - introduces the idea of using LLMs to optimize prompts: let LLMs "Take a deep breath" improves the performance on math problems.
- AutoPrompt - proposes an approach to automatically create prompts for a diverse set of tasks based on gradient-guided search.
- Prefix Tuning - a lightweight alternative to fine-tuning that prepends a trainable continuous prefix for NLG tasks.
- Prompt Tuning - proposes a mechanism for learning soft prompts through backpropagation.