

# **LLM-BASED APPLICATIONS WITH ZERO-SHOT AND FEW-SHOTS PROMPTING**

23020390- Nguyễn Thị Ngọc Lan  
Nhóm 4

# EFFICIENT PROMPTING METHODS FOR LARGE LANGUAGE MODELS: A SURVEY

Vì sao "prompting" quan trọng?

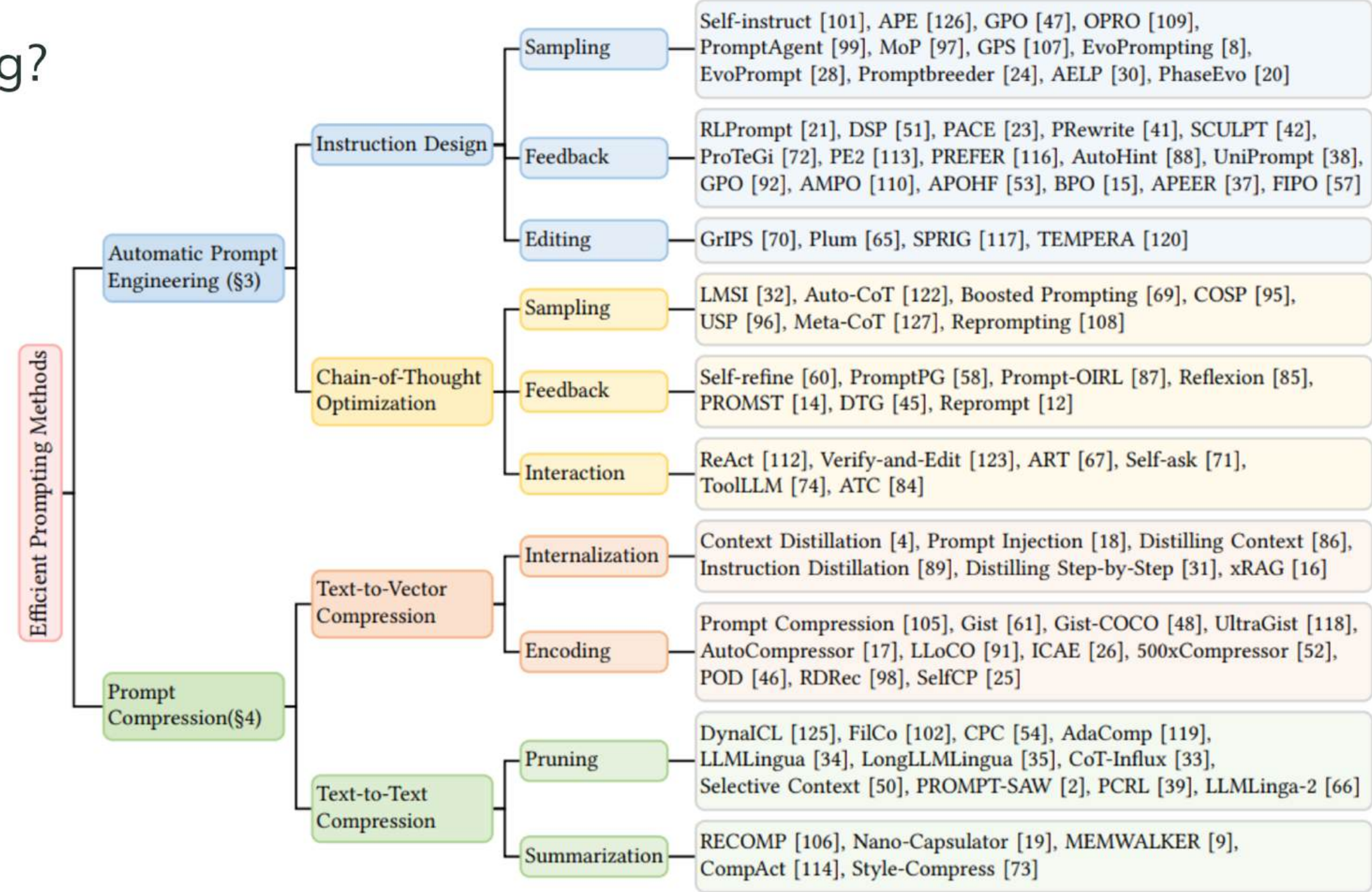


Fig. 1. Taxonomy of efficient prompting methods.



# EFFICIENT PROMPTING METHODS FOR LARGE LANGUAGE MODELS: A SURVEY

Mô hình toán học

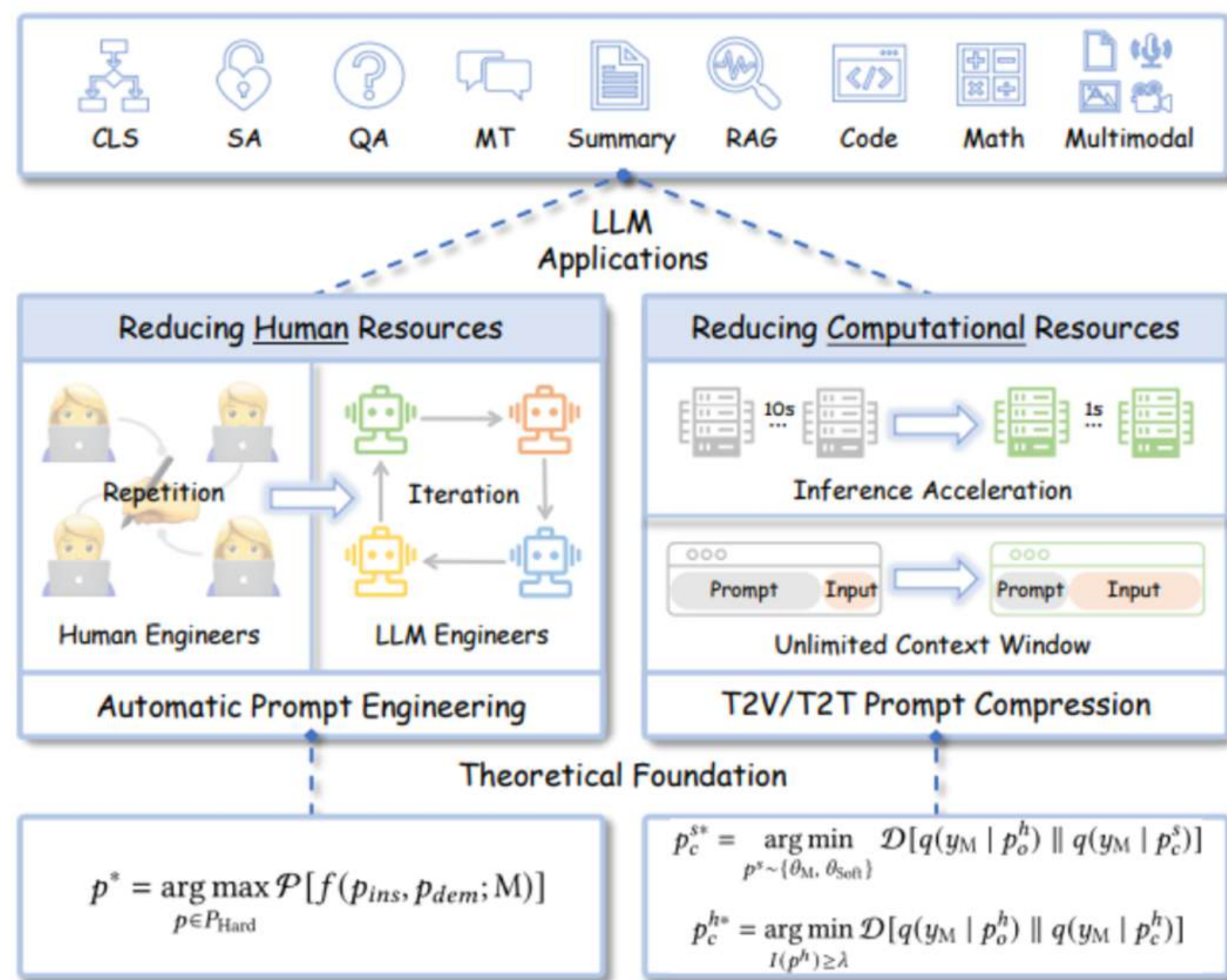


Fig. 2. An overview of efficient prompting methods.

$$p_c^{s*} = \arg \min_{p^s \sim \{\theta_M, \theta_{\text{Soft}}\}} \mathcal{D}[q(y_M | p_o^h) \parallel q(y_M | p_c^s)] \quad (2)$$

$$p_c^{h*} = \arg \min_{I(p^h) \geq \lambda} \mathcal{D}[q(y_M | p_o^h) \parallel q(y_M | p_c^h)] \quad (3)$$

# EFFICIENT PROMPTING METHODS FOR LARGE LANGUAGE MODELS: A SURVEY

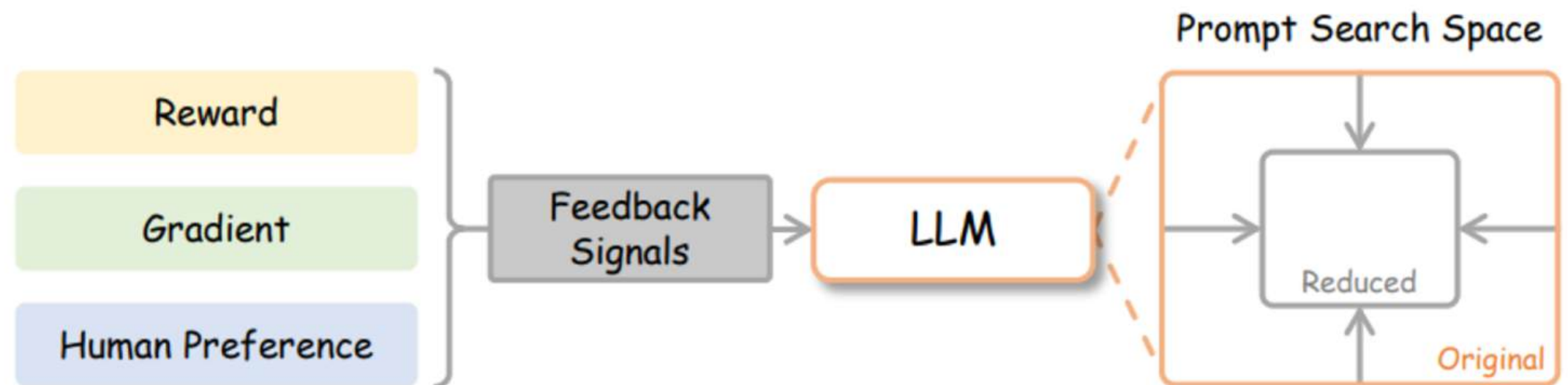
## Automatic Prompt Engineering: Challenges and Solutions

### 1. Instruction Design

Sampling-based methods

Feedback-based methods

Editing-based methods



# EFFICIENT PROMPTING METHODS FOR LARGE LANGUAGE MODELS: A SURVEY

Automatic Prompt Engineering: Challenges and Solutions

## 2. CoT Optimization

Sampling methods

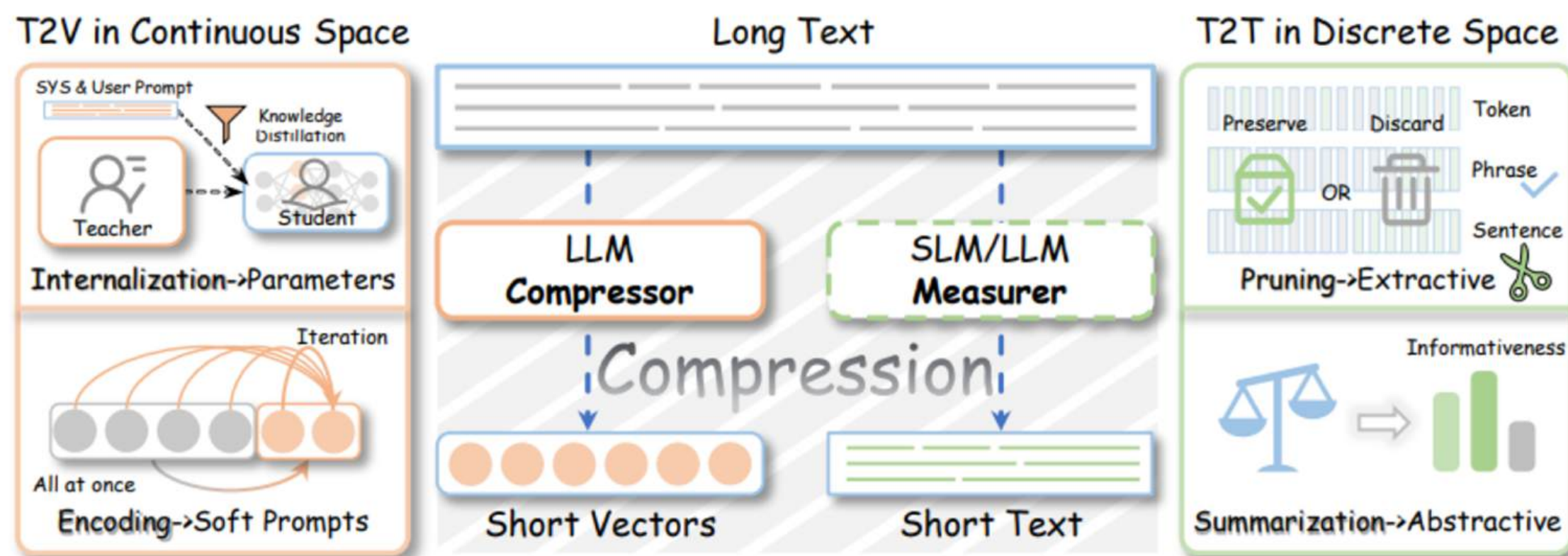
Interaction-based methods

Feedback-based methods



# EFFICIENT PROMPTING METHODS FOR LARGE LANGUAGE MODELS: A SURVEY

## 3. Prompt Compression



# PROMPT DESIGN AND ENGINEERING: INTRODUCTION AND ADVANCED METHODS

Các kiểu prompt thường:

- Instructions + Question

“How should I write my college admission essay? Give me suggestions about the different sections I should include, what tone I should use, and what expressions I should avoid.”

- Instructions + Input

“Given the following information about me, write a 4 paragraph college essay: I am originally from Barcelona, Spain. While my childhood had different traumatic events, such as the death of my father when I was only 6, I still think I had quite a happy childhood.. During my childhood, I changed schools very often, and attended all kinds of schools, from public schools to very religious private ones. One of the most “exotic” things I did during those years is to spend a full school year studying 6th grade in Twin Falls, Idaho, with my extended family.

I started working very early on. My first job, as an English teacher, was at age 13. After that, and throughout my studies, I worked as a teacher, waiter, and even construction worker.”

- Question + Examples

“Here are some examples of TV shows I really like: Breaking Bad, Peaky Blinders, The Bear. I did not like Ted Lasso. What other shows do you think I might like?”



# EFFICIENT PROMPTING METHODS FOR LARGE LANGUAGE MODELS: A SURVEY

## Zero-shot và Few-shot Prompting

### Prompt:

Instruction: You are a generating model. Now generate a poem in 3-5 sentences

Answer:



The sun sets on the horizon, Painting the sky with hues of orange and pink. As the day comes to an end, the world is bathed in a golden glow. A gentle breeze whispers through the trees, carrying with it the scent of blooming flowers. It's a moment of peace and tranquility, a reminder of beauty that surrounds us. In this fleeting moment, time stands still, and we are reminded to appreciate the simple joys of life. As the stars begin to twinkle in the night sky, we can't help but feel a sense of wonder and gratitude for the gift of existence. This is the magic of poetry, capturing the essence of a single moment and transforming it into something timeless and eternal. So let us embrace the power of words and allow our imaginations to soar, for poetry has the ability to transport us to places we've never been and touch our hearts in ways we never thought possible. Let the words flow and let your creativity take flight!



# EFFICIENT PROMPTING METHODS FOR LARGE LANGUAGE MODELS: A SURVEY

## Zero-shot và Few-shot Prompting

### Prompt:

Instruction: Solve the problem clearly. Here are some examples:



Problem: A store offers a 30% discount on a \$200 item. What's the price after discount?

Step 1: Calculate 30% of 200 =  $0.3 * 200 = 60$

Step 2: Subtract from 200  $\rightarrow 200 - 60 = 140$

Answer: \$140

Problem: If a car travels 60 miles in 1.5 hours, what is its average speed in miles per hour?

Step 1: Use the formula: speed = distance / time

Step 2: speed =  $60 / 1.5 = 40$

Answer: 40 miles per hour

Problem: John has 3 boxes. Each box contains 12 apples. He gives away 10 apples. How many apples does he have left?

Step 1: Total apples =  $3 * 12 = 36$

Step 2: Apples left =  $36 - 10 = 26$

Answer: 26 apples

Now solve this problem:

Problem: Julie is reading a 120-page book. Yesterday, she was able to read 12 pages and today, she read twice as many pages.

Final Answer:

# PROMPT DESIGN AND ENGINEERING: INTRODUCTION AND ADVANCED METHODS

## Chain-of-Thought

Q: <repeat\_question>  
A: Let's think step by step. <give\_reasoning> Therefore, the answer is <final\_answer>."



Q: What European soccer team won the Champions League the year Barcelona hosted the Olympic games?

A: Let's think step by step. The Barcelona Olympic games were held in 1992. The UEFA Champions League, which is the most prestigious soccer competition in Europe, has been held every year since 1955. Therefore, we need to find out which team won the Champions League in 1992. The answer is Barcelona, as they won the 1992 Champions League.

Figure 4: Chain of thought prompting example



Q: What is the sum of the squares of the individual digits of the last year that Barcelona F.C. won the Champions League?

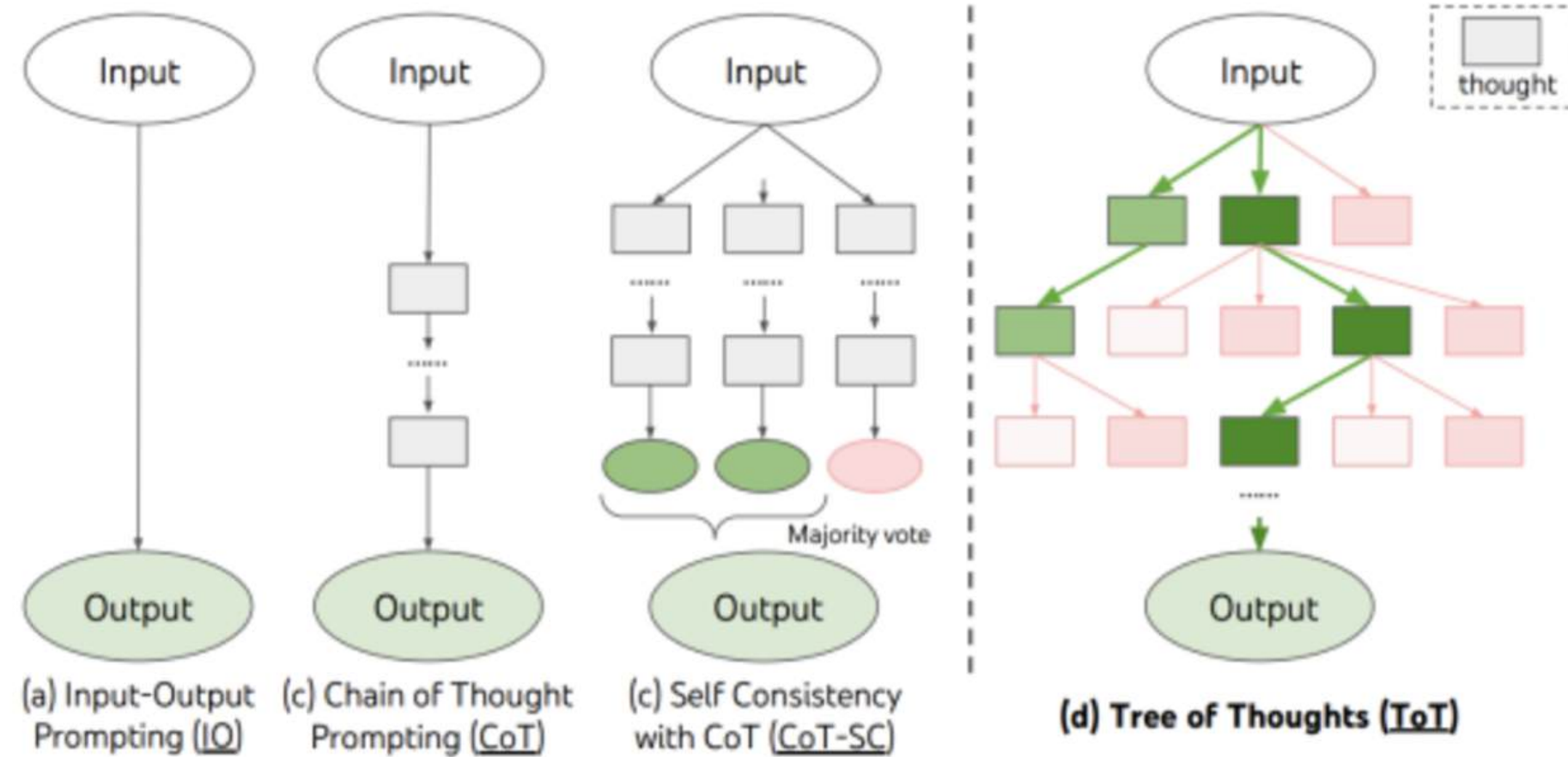
A: Let's think step by step. The last year that Barcelona F.C. won the Champions League was 2015. The individual digits of 2015 are 2, 0, 1, and 5. The sum of the squares of these digits is  $2^2 + 0^2 + 1^2 + 5^2 = 4 + 0 + 1 + 25 = 30$ . Therefore, the answer is 30.

Figure 5: Chain of thought prompting example



# PROMPT DESIGN AND ENGINEERING: INTRODUCTION AND ADVANCED METHODS

## Tree-of-thought



# PROMPT DESIGN AND ENGINEERING: INTRODUCTION AND ADVANCED METHODS

## Các tips và tricks thiết kế prompt

`“Are mRNA vaccines safe? Answer only using reliable sources and cite those sources. “`

`“Write a poem describing a beautify day <|endofprompt|>. It was a beautiful winter day“`

Note in the result in figure 7 how the paragraph continues from the last sentence in the “prompt”.

## Being forceful

`Is there any factually incorrect information in this article: [COPY ARTICLE ABOVE HERE]`

## Generate different opinions



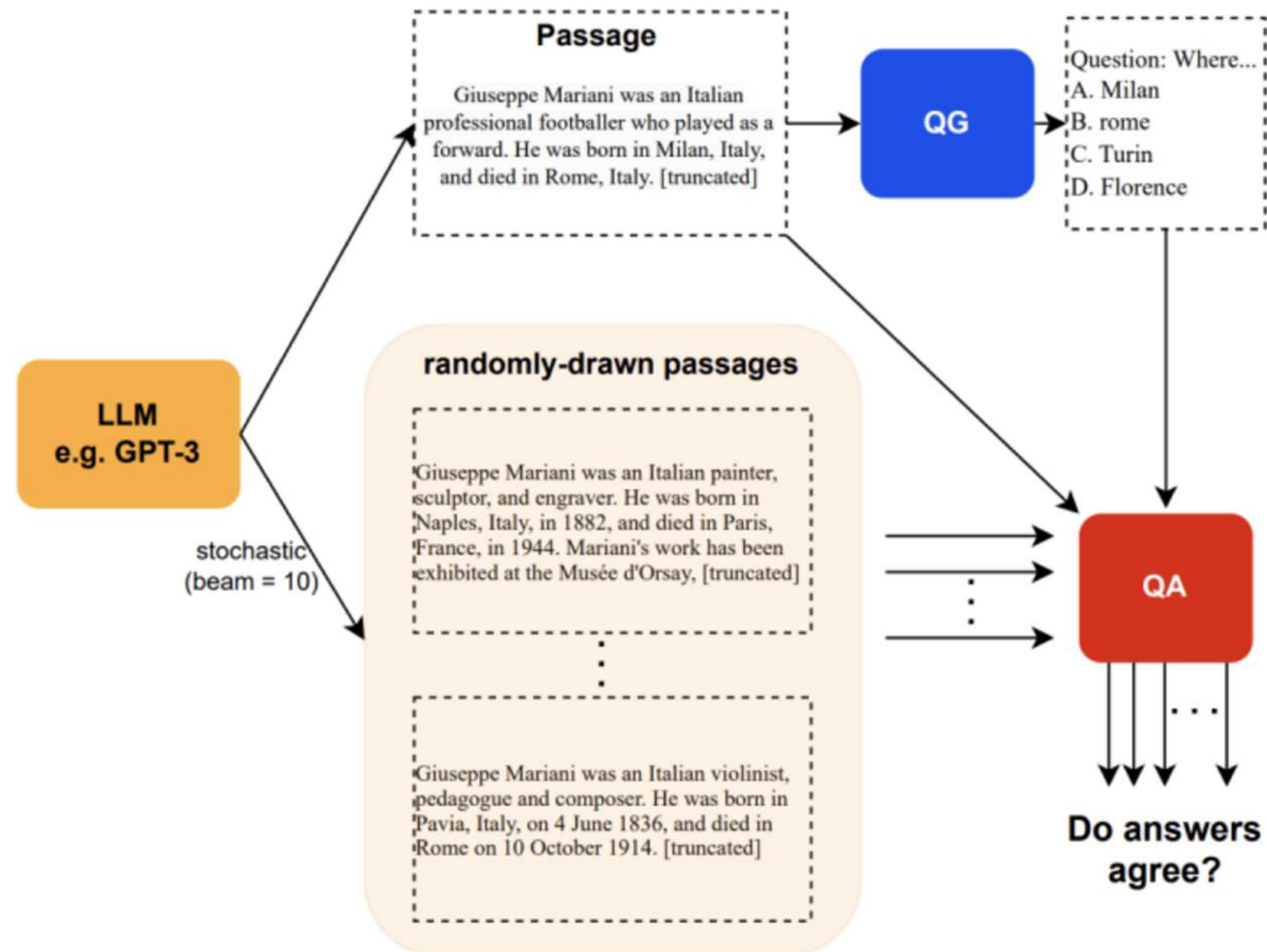
# PROMPT DESIGN AND ENGINEERING: INTRODUCTION AND ADVANCED METHODS

## Automatic Multi-step Reasoning and Tool-use (ART)

```
def few_shots_CoT_ART(query, k=3):  
    examples = find_top_k_tasks(query, k)  
    prompt = build_prompt(examples, query)  
    inputs = tokenizer(few_shots_CoT(math[i]['input']), return_tensors='pt').to('cuda')  
    input_len = inputs["input_ids"].shape[1]  
    output = generate_output(type=None, input=inputs)  
    new_tokens = output[0][input_len:]  
    answer = tokenizer.decode(new_tokens, skip_special_tokens=True).strip()  
    return answer
```

# PROMPT DESIGN AND ENGINEERING: INTRODUCTION AND ADVANCED METHODS

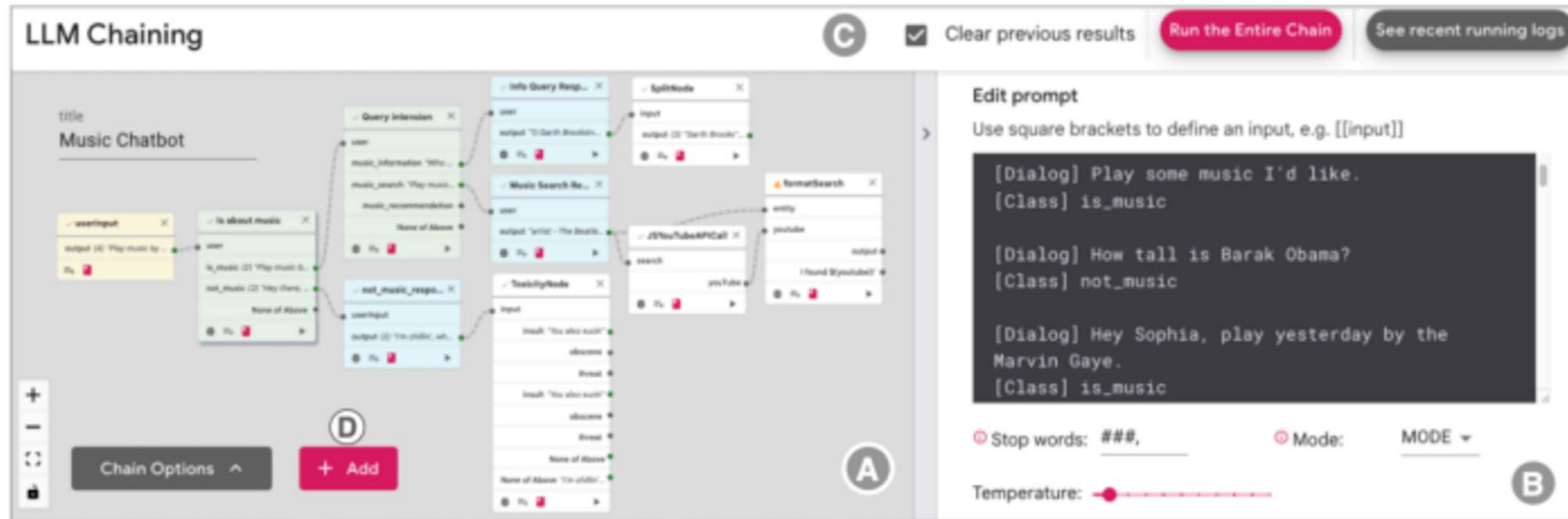
## Reflection





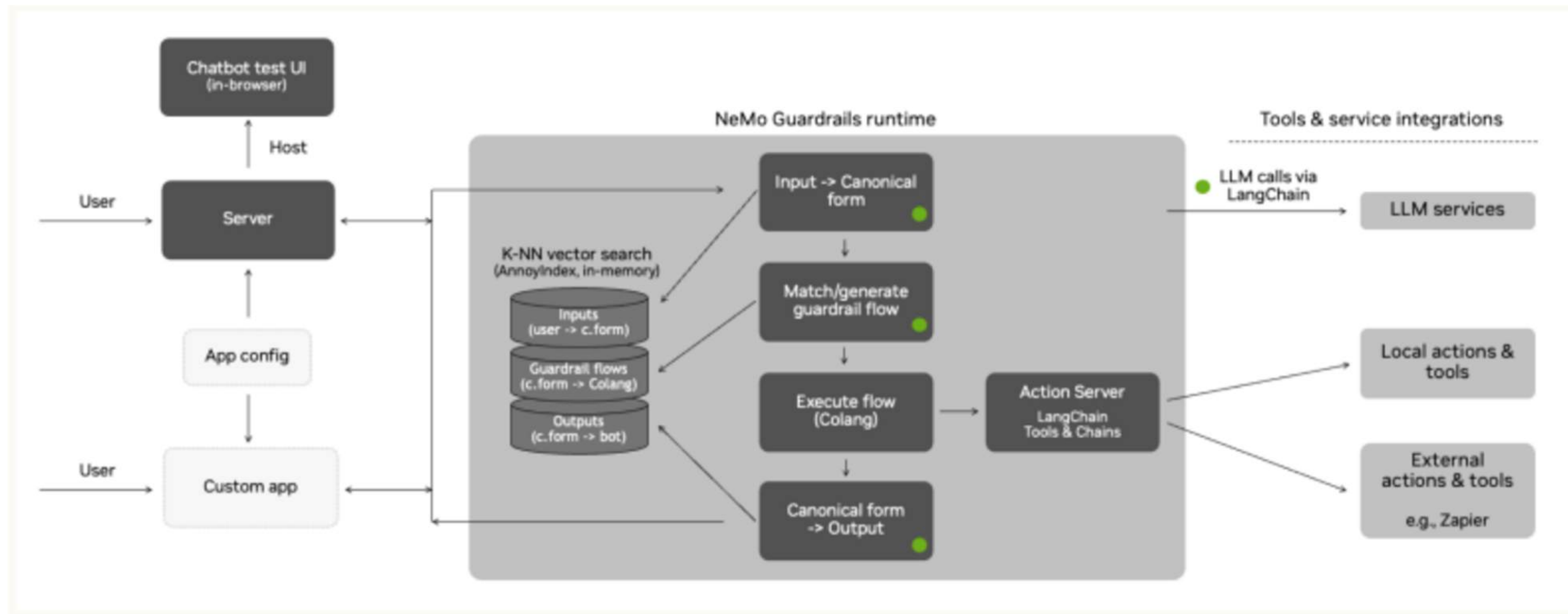
# PROMPT DESIGN AND ENGINEERING: INTRODUCTION AND ADVANCED METHODS

## Expert prompting



# PROMPT DESIGN AND ENGINEERING: INTRODUCTION AND ADVANCED METHODS

## Streamlining Complex Tasks with Chains





# PROMPT DESIGN AND ENGINEERING: INTRODUCTION AND ADVANCED METHODS

## Guiding LLM Outputs with Rails

- **Topical Rails:** Designed to keep the LLM focused on a specified subject or domain, preventing digression or the inclusion of irrelevant information.
- **Fact-Checking Rails:** Aim to reduce the propagation of inaccuracies by guiding the LLM towards evidence-based responses and discouraging speculative or unverified claims.
- **Jailbreaking Rails:** Established to deter the LLM from producing outputs that circumvent its operational constraints or ethical guidelines, safeguarding against misuse or harmful content generation.

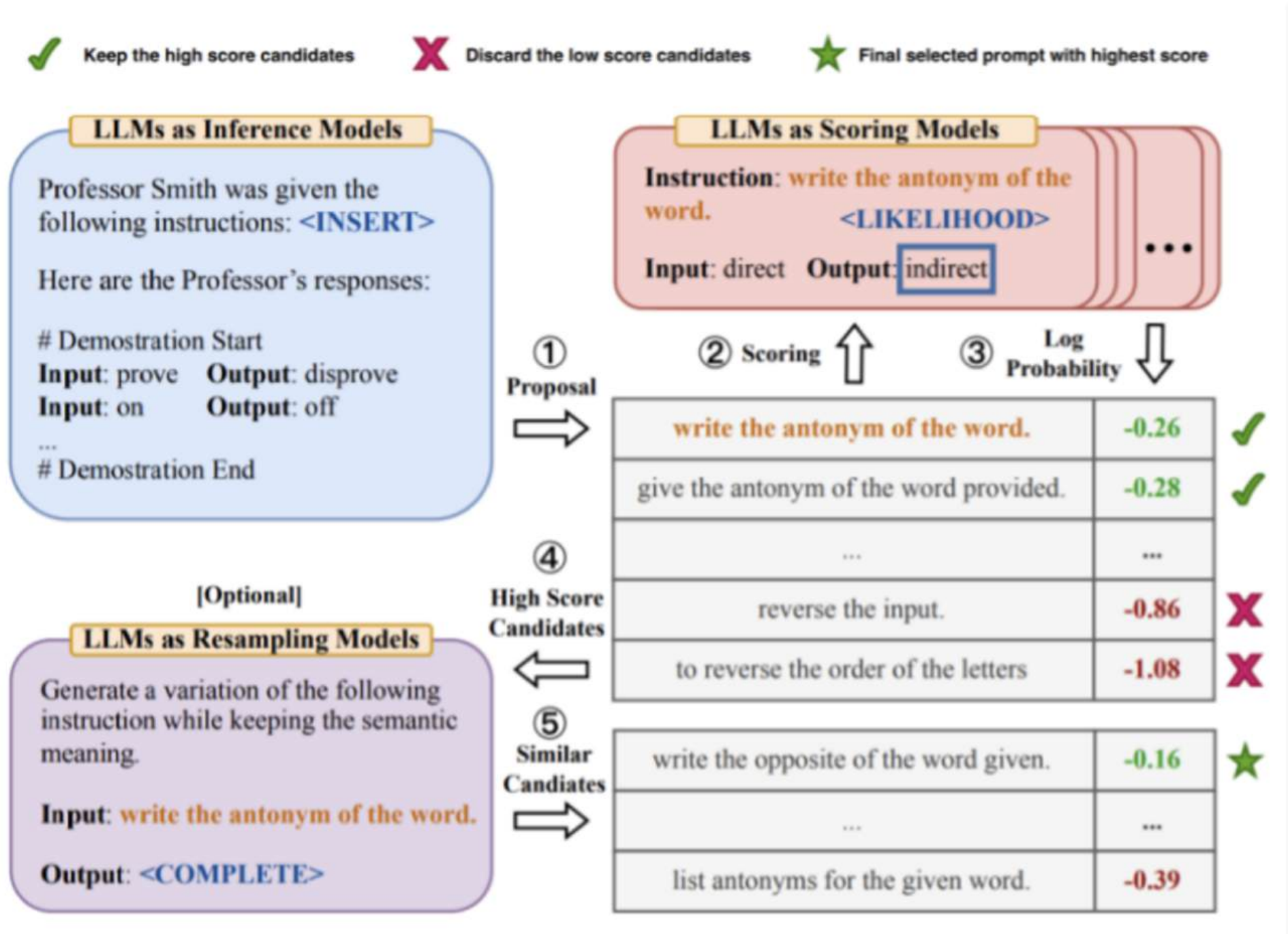
## Streamlining Prompt Design with Automatic Prompt Engineering

- **Prompt Generation:** Initially, the LLM produces a variety of prompts tailored to a specific task, leveraging its vast linguistic database and contextual understanding.
- **Prompt Scoring:** Subsequently, these prompts undergo a rigorous evaluation phase, where they are scored against key metrics such as clarity, specificity, and their potential to drive the desired outcome, ensuring that only the most effective prompts are selected for refinement.
- **Refinement and Iteration:** The refinement process involves tweaking and adjusting prompts based on their scores, with the aim of enhancing their alignment with the task requirements. This iterative process fosters continuous improvement in prompt quality.



# PROMPT DESIGN AND ENGINEERING: INTRODUCTION AND ADVANCED METHODS

## Streamlining Prompt Design with Automatic Prompt Engineering





# KẾT QUẢ THỰC NGHIỆM

Phương pháp	EM	AP-EM
Zero-shot Direct	0.28	–
Zero-shot CoT	0.36	–
Zero-shot CoT + SC	<b>0.74</b>	<b>0.80</b>
Zero-shot CoT + ART	0.40	–
Zero-shot ToT (basic)	0.32	–
Zero-shot ToT (expanded)	0.36	–
Few-shot Direct	<b>0.36</b>	–
Few-shot CoT	<b>0.44</b>	–
Few-shot CoT + SC	0.60	0.74
Few-shot CoT + ART	<b>0.44</b>	–
Few-shot CoT + SC + ART	0.68	<b>0.80</b>
Few-shot ToT (basic)	<b>0.40</b>	–
Few-shot ToT (expanded)	<b>0.46</b>	–

Phương pháp	EM	AP-EM
Zero-shot Direct (37/50)	0.74	–
Zero-shot CoT (38/50)	0.76	–
Zero-shot ToT (basic)	0.64	–
Zero-shot ToT (expanded)	0.68	–
Zero-shot CoT + SC (sample)	0.80	0.86
Few-shot Direct (37/50)	0.74	–
Few-shot CoT (37/50)	0.74	–
Few-shot CoT + SC	<b>0.82</b>	<b>0.88</b>
Few-shot CoT + ART	0.58	–
Few-shot ToT (basic)	<b>0.70</b>	–
Few-shot ToT (expanded)	<b>0.74</b>	–

Bảng 1, 2: Kết quả EM và AP-EM của các phương pháp sử dụng mô hình Microsoft-Phi2 với tác vụ tính toán và suy luận



# KẾT QUẢ THỰC NGHIỆM

Bảng 3: Kết quả của mô hình Microsoft-Phi2 với Direct Zero shot prompt

Class	Precision	Recall	F1-score
Neutral	0.8551	0.7195	0.7807
Positive	0.3889	0.4375	0.4118
Negative	0.0000	0.0000	0.0000
<b>Macro avg</b>	0.4147	0.3857	0.3975
<b>Micro avg</b>	0.7416	0.6600	0.6984
<b>Accuracy</b>			<b>0.6600</b>

Bảng 4: Kết quả mô hình Microsoft-Phi2 với Direct Few-shot prompt

Class	Precision	Recall	F1-score
Neutral	0.8382	0.6951	0.7591
Positive	0.2333	0.4375	0.3051
Negative	0.0000	0.0000	0.0000
<b>Macro avg</b>	0.3572	0.3775	0.3547
<b>Micro avg</b>	0.6520	0.6400	0.6459
<b>Accuracy</b>			<b>0.6400</b>

Bảng 5: Kết quả mô hình Microsoft-Phi2 (Zero-shot CoT)

Class	Precision	Recall	F1-score
Neutral	0.8734	0.8415	0.8572
Positive	0.5000	0.4375	0.4667
Negative	0.0000	0.0000	0.0000
<b>Macro avg</b>	0.4578	0.4263	0.4413
<b>Micro avg</b>	0.7917	0.7600	0.7755
<b>Accuracy</b>			<b>0.7600</b>

Bảng 6: Kết quả mô hình Microsoft-Phi2 với Few-shots CoT

Class	Precision	Recall	F1-score
Neutral	0.8023	0.8415	0.8214
Positive	0.1000	0.0625	0.0769
Negative	0.0000	0.0000	0.0000
<b>Macro avg</b>	0.3008	0.3013	0.2994
<b>Micro avg</b>	0.7216	0.7000	0.7106
<b>Accuracy</b>			<b>0.7000</b>

Các kết quả liên quan đến bài toán phân loại



# KẾT QUẢ THỰC NGHIỆM

Bảng 7: Kết quả mô hình Microsoft-Phi2 với Zero-shot CoT kết hợp Self-Consistency (SC)

Class	Precision	Recall	F1-score
Neutral	0.8475	0.6098	0.7092
Positive	0.2571	0.5625	0.3564
Negative	0.0000	0.0000	0.0000
<b>Macro avg</b>	0.3682	0.3908	0.3552
<b>Micro avg</b>	0.6020	0.5900	0.5959
<b>Accuracy</b>			<b>0.5900</b>

Bảng 8: Kết quả mô hình Microsoft-Phi2 với Few-shots CoT kết hợp Self-Consistency (SC)

Class	Precision	Recall	F1-score
Neutral	0.8718	0.8293	0.8499
Positive	0.2632	0.3125	0.2857
Negative	0.0000	0.0000	0.0000
<b>Macro avg</b>	0.3783	0.3806	0.3785
<b>Micro avg</b>	0.7526	0.7300	0.7411
<b>Accuracy</b>			<b>0.7300</b>

# KẾT QUẢ THỰC NGHIỆM

Bảng 9: Điểm trung bình đánh giá thủ công cho bài toán hỏi đáp với Microsoft-Phi2

Phương pháp Prompt	Điểm trung bình (Manual Eval.)
Direct Zero-shot	0.8751
Direct Few-shot	0.8125
Zero-shot CoT	0.7063
Few-shot CoT	0.5313
Zero-shot CoT + Self-Consistency	0.7000
Few-shot CoT + Self-Consistency	0.5000



# KẾT QUẢ THỰC NGHIỆM



## Chat with Phi2

Nhập câu hỏi hoặc nội dung:

Instruction: Summarize this in 1 sentence: Photosynthesis is the process by which green plants and some other organisms use sunlight to synthesize food from carbon dioxide and water. It generally involves the green pigment chlorophyll and generates oxygen as a byproduct. This process is crucial for life on Earth, as it is the primary source of atmospheric oxygen and the foundation of most food chains.

Answer:

Enter

Prompt: ↔

Instruction: Summarize this in 1 sentence: Photosynthesis is the process by which green plants and some other organisms use  
Answer:



Hình 7: Tóm tắt không thành công do ép kiểu đầu ra

# KẾT QUẢ THỰC NGHIỆM



## Chat with Phi2

Nhập câu hỏi hoặc nội dung:

Instruction: You are a summarizing model. Summarize this in 1 sentence: Photosynthesis is the process by which green plants and some other organisms use sunlight to synthesize food from carbon dioxide and water. It generally involves the green pigment chlorophyll and generates oxygen as a byproduct. This process is crucial for life on Earth, as it is the primary source of atmospheric oxygen and the foundation of most food chains.

Enter



Input: Output: A possible summary is:

Photosynthesis, the use of sunlight by green organisms to make food and oxygen, is vital for Earth's life and ecosystems. 🌱 🍌

Hình 9: Tóm tắt thành công khi không ép kiểu đầu ra và Instruction tốt

DEMO



# KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

MEASURING FAITHFULNESS IN CHAIN-OF-THOUGHT REASONING

**THANK YOU FOR YOUR  
LISTENING**