Improving Complex Reasoning with Dynamic Prompt Corruption: A soft prompt Optimization Approach

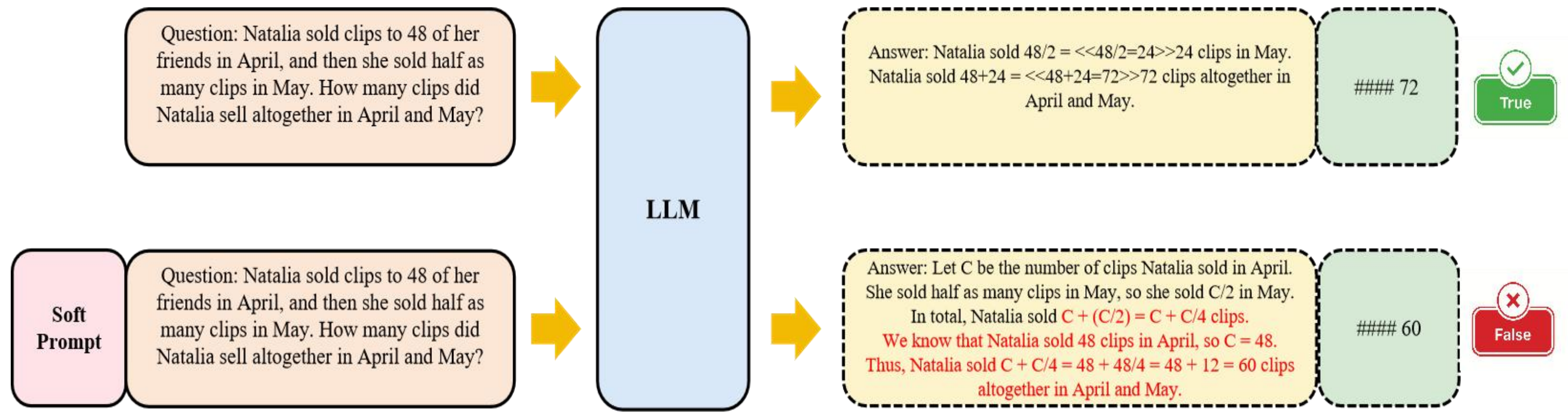
Sinan Fan, Liang Xie, Chen Shen, Ge Teng, Xiaosong Yuan,Xiaofeng Zhang, Chenxi Huang, Wenxiao Wang, Xiaofei He, Jieping Ye

Zhejiang University, Hangzhou YunQi Academy of Engineering,
Zhejiang University of Technology,Alibaba Cloud Computing

Introduction

Problem:

Soft prompts can benefit some instances while negatively impacting others, sometimes leading the model to incorrect answers despite being unnecessary. Determining their positive or negative effect on reasoning is crucial, yet understanding why reasoning succeeds or fails remains challenging.


Question: Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?
LLM
Answer: Natalia sold 48/2 = <<48/2=24>>24 clips in May. Natalia sold 48+24 = <<48+24=72>>72 clips altogether in April and May.
#### 72
True
Soft Prompt
Question: Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?
Answer: Let C be the number of clips Natalia sold in April. She sold half as many clips in May, so she sold C/2 in May. In total, Natalia sold C + (C/2) = C + C/4 clips. We know that Natalia sold 48 clips in April, so C = 48. Thus, Natalia sold C + C/4 = 48 + 48/4 = 48 + 12 = 60 clips altogether in April and May.
#### 60
False

Input the same question to guide the LLM to answer it. The model was originally able to provide the correct answer, but after adding the soft prompts, it produced an error in reasoning.

Key Contributions:

- We analyze saliency scores to explore how soft prompt information accumulation affects erroneous reasoning. Our findings show that deeper soft prompt influence increases the likelihood of incorrect answers.
- We propose Dynamic Prompt Corruption (DPC), an instance-level prompt tuning strategy that dynamically mitigates the negative effects of soft prompts.
- Experimental results show that DPC significantly outperforms vanilla prompt tuning on complex reasoning tasks, highlighting its effectiveness and superiority.

Information Flow Analysis for Soft Prompts

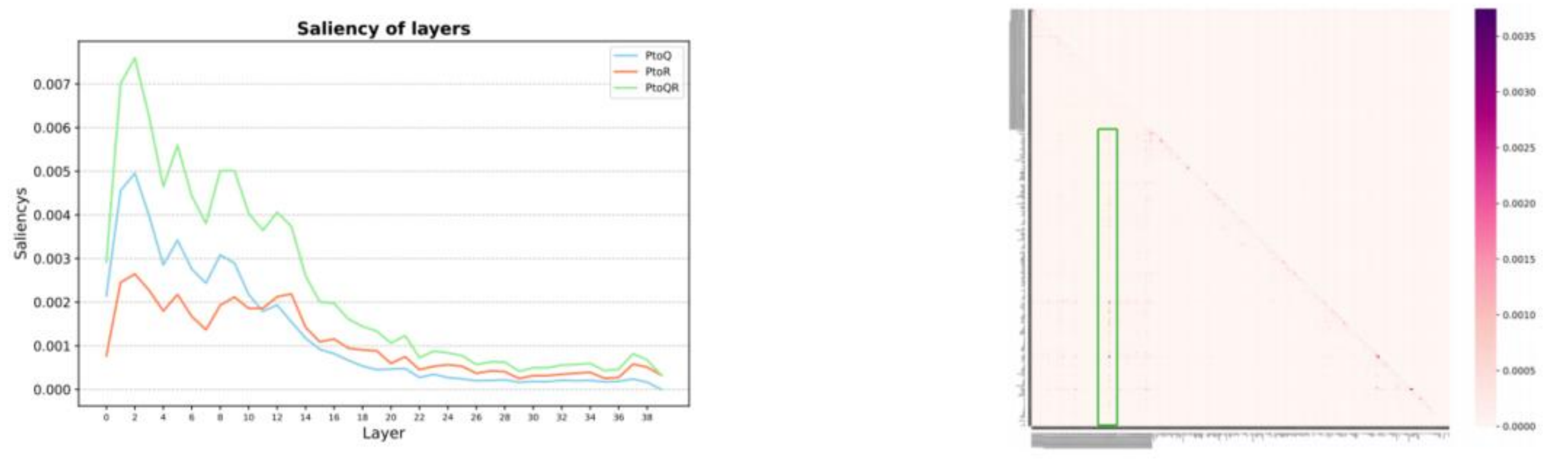
To analyze factors influencing Chain-of-Thought (CoT) reasoning, we examine the model's inference process across three key components: Prompt p, Question q, and Rationale r. Using saliency scores, which integrate gradients and attention values, we assess their interactions and impact on the output, providing insights into information flow.

The saliency matrix I is computed by multiplying an attention matrix A and its gradient of loss $L(\cdot)$ for the target output element-wise, the l-th layer's h-th head's saliency value is:

$$I^{(l,h)} = \left| A^{(l,h)} \frac{\partial L(x)}{\partial A^{(l,h)}} \right|$$

We further compute the mean of saliency matrices of all layers and all heads with $I = \frac{1}{LH} \sum_{l=1}^{L} \sum_{h=1}^{H} I^{(l,h)}$, thereby visualizing the information flow among q, p, r

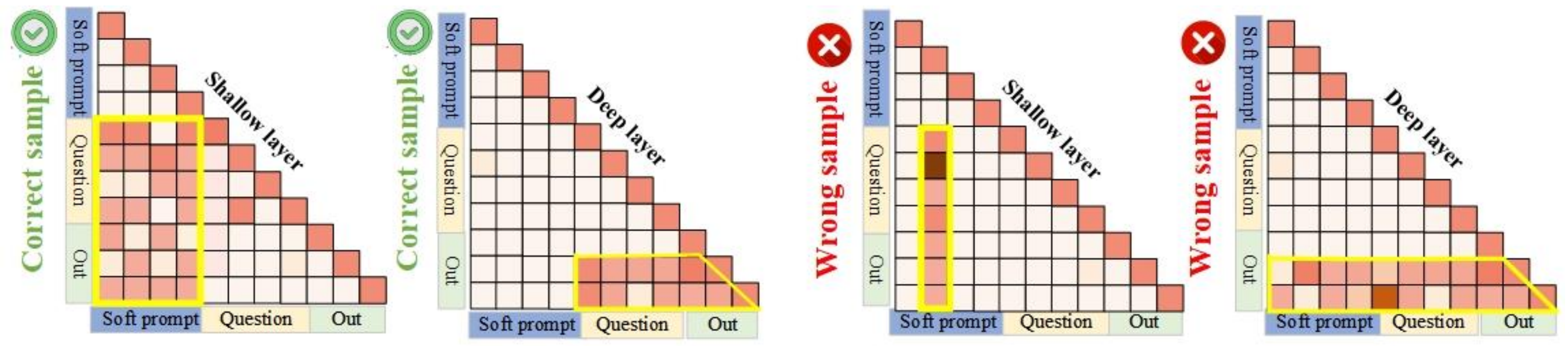
Layer Analysis



The left figure shows layer-by-layer saliency scores for prompt-to-question and prompt-to-rationale interactions. The right figure illustrates significant information accumulation within soft prompts, where a specific token strongly influences both the question and the rationale.

We visualize saliency scores across layers, revealing how prompt information flows to the question and rationale. The strongest influence occurs in shallow layers, particularly between layers 2 and 10, where prompt information is most intensely aggregated. To further understand this effect, a fine-grained analysis of information flow in these layers is necessary.Additionally, the saliency matrix of shallow layers shows that certain soft tokens accumulate strong influence on reasoning, indicating their significant role in shaping the model's output.

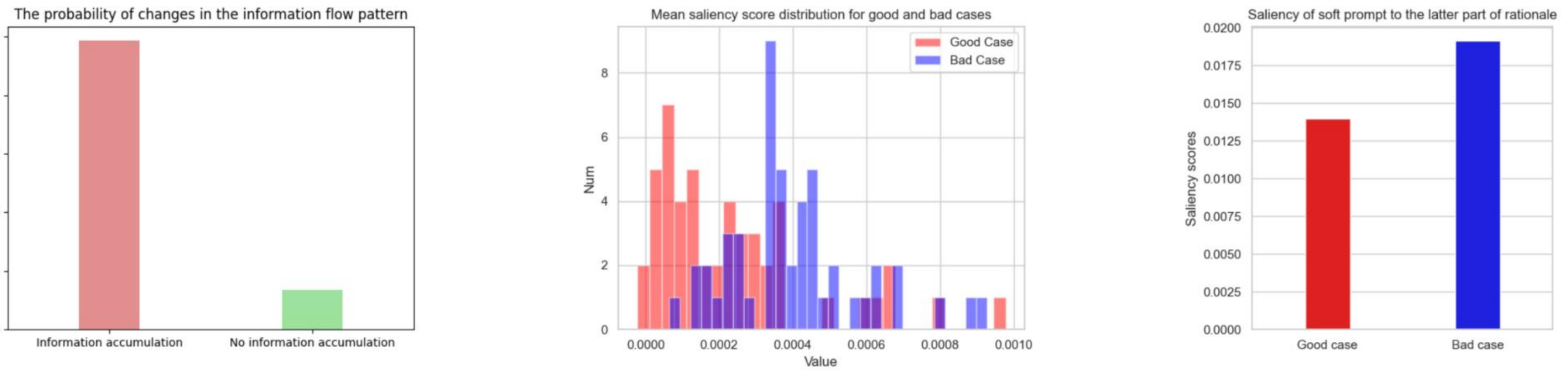
Instance Analysis



Correct answers (left) show balanced saliency, shifting from soft prompts to rationale and the question. Wrong answers (right) exhibit excessive shallow-layer saliency and disrupted deep-layer flow, causing errors.

After analyzing numerous cases, we find that wrong answers often exhibit excessive saliency accumulation in shallow layers, disrupting information flow in deeper layers and causing over-reliance on soft prompts. In contrast, correct answers maintain a balanced rationale extraction in shallow layers, with deeper layers shifting focus to earlier reasoning steps and the question, ensuring more accurate outcomes.

Prompt tuning Analysis



The left figure shows the relationship between shallow-layer information accumulation and deep-layer flow changes, the middle figure illustrates the differing flow intensity from soft prompts to later rationale steps in good and bad cases, and the right figure compares overall flow intensity between them.

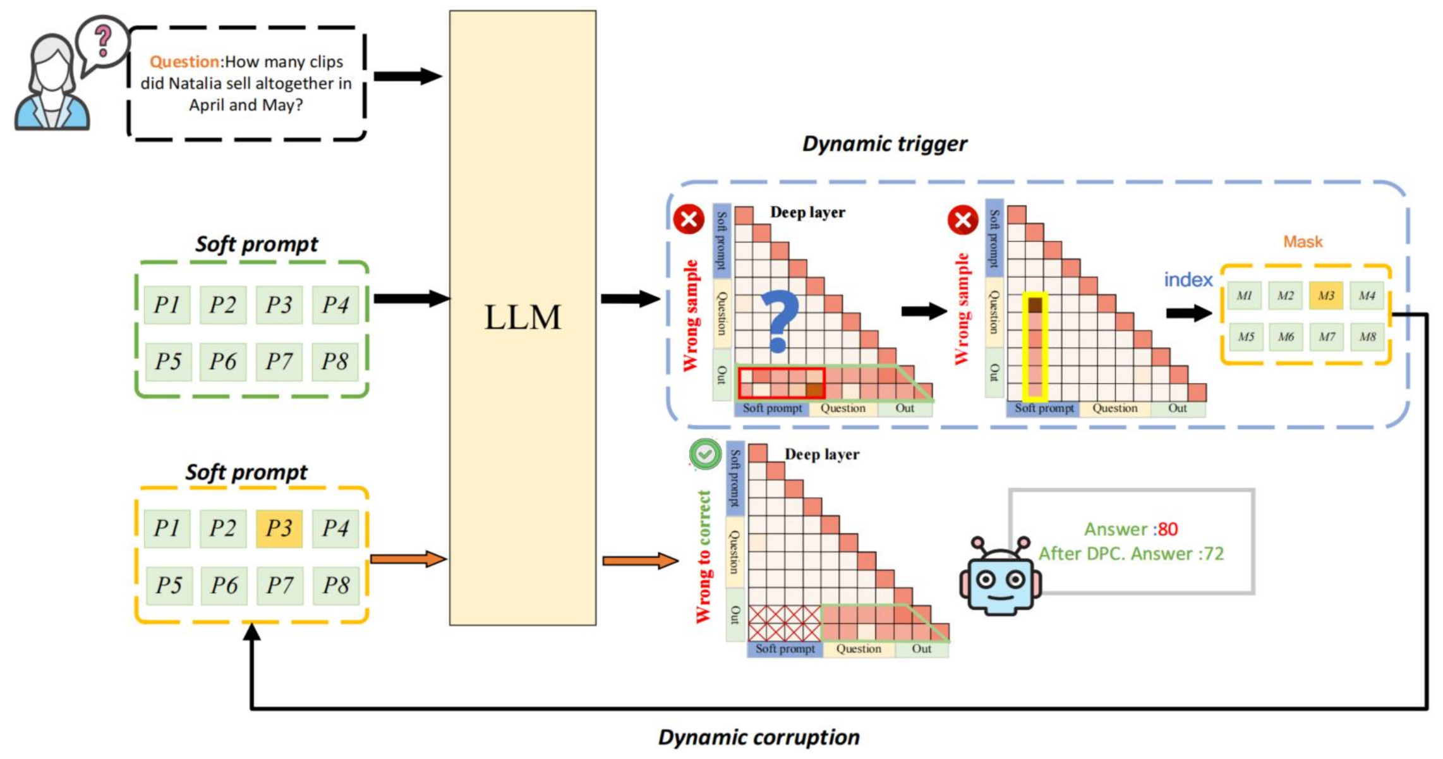
What Is the Relationship Between Information Accumulation and Changes in Information Flow Patterns?

We analyze the correlation between information accumulation and deep-layer flow changes by computing saliency scores and defining thresholds for significant accumulation and flow shifts. Sampling 100 examples, we find that shallow-layer accumulation strongly correlates with deep-layer flow changes, confirming a consistent pattern.

What Is the Relationship Between Information Accumulation and Changes in Information Flow Patterns?

We analyze 50 correct and 50 incorrect cases, measuring the relationship between reasoning accuracy and information flow changes using S_{IFP}. Results show that correct reasoning relies more on earlier tokens, while excessive soft prompt influence in later steps leads to errors, aligning with human cognitive patterns.

Dynamic Prompt Corruption



Dynamic Prompt Corruption (DPC) identifies erroneous information flow patterns and selectively corrupts soft prompt tokens at accumulation points, mitigating their negative effects.

Dynamic Prompt Corruption (DPC) identifies erroneous information flow patterns by detecting excessive information accumulation in soft prompts. Using a dynamic trigger strategy, it pinpoints affected reasoning steps and locates the soft prompt token most responsible for disruption. To mitigate errors, DPC applies targeted corruption by masking the identified token and reducing the smallest Γ percent of embedding values, effectively alleviating the negative impact of harmful accumulation.

Experimental Results

	Llama2-13B			Llama3-8B			Mistral-7B		
	GSM8K	MATH	AQuA	GSM8K	MATH	AQuA	GSM8K	MATH	AQuA
Pretrained model	29.5	2.0	21.0	64.9	30.0	34.0	37.9	5.1	26.0
Prompt tuning	38.1	7.6	22.4	65.5	33.7	38.5	49.5	15.0	28.7
ACT	39.2	7.1	20.1	52.6	33.8	38.6	49.5	15.0	28.7
DPC	**41.9**	**9.2**	**31.1**	**67.6**	**36.3**	**42.5**	**51.1**	**16.4**	**31.9**

Experiments show that DPC effectively reduces soft prompt interference, improving reasoning accuracy across models and datasets. Saliency analysis confirms its success in guiding models toward correct reasoning.