ICML 2026 PAPER BRIEF

How Far Ahead Do LLMs Plan?
Uncover the Latent Horizon in Chain-of-Thought Reasoning

Liyan Xu, Mo Yu, Fandong Meng, Jie Zhou

WeChat AI, Tencent Inc.

Motivation

View A: prior to the explicit CoT emergence, **internal planning exists** on subsequent reasoning

View B: explicit CoT remains critical for Transformers

Objective

To examine the synergy/dynamics between explicit CoT steps and its **latent planning horizon**.

HOW FAR AHEAD?

Motivation

Q1: To what extent do hidden states encode a global plan for the reasoning roadmap, as opposed to supporting rather local, incremental state transitions?

Q2: How does the planning horizon further influence other characteristics of CoT reasoning?

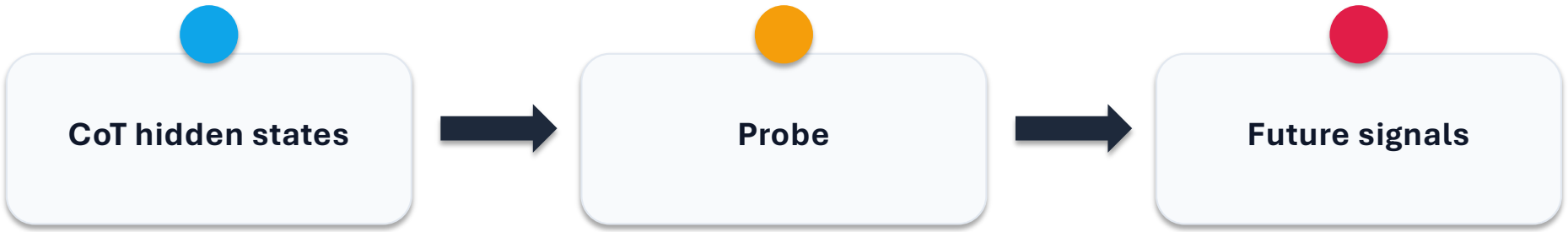
Implications

- *CoT uncertainty*: better calibration
- *LLM overthinking*: more concise CoT
- *LLM Adaptive thinking efforts*: rely on how strong the model can “sense” or “see through” the input complexity

SETUP

Tele-Lens

Tele-Lens: a probing method that maps CoT hidden states to future signals.



If LLMs maintain a precise global plan early in the reasoning process, then the strength of that plan should, in principle, be **measurable** by these probing targets, before explicit CoT fully unfolds.

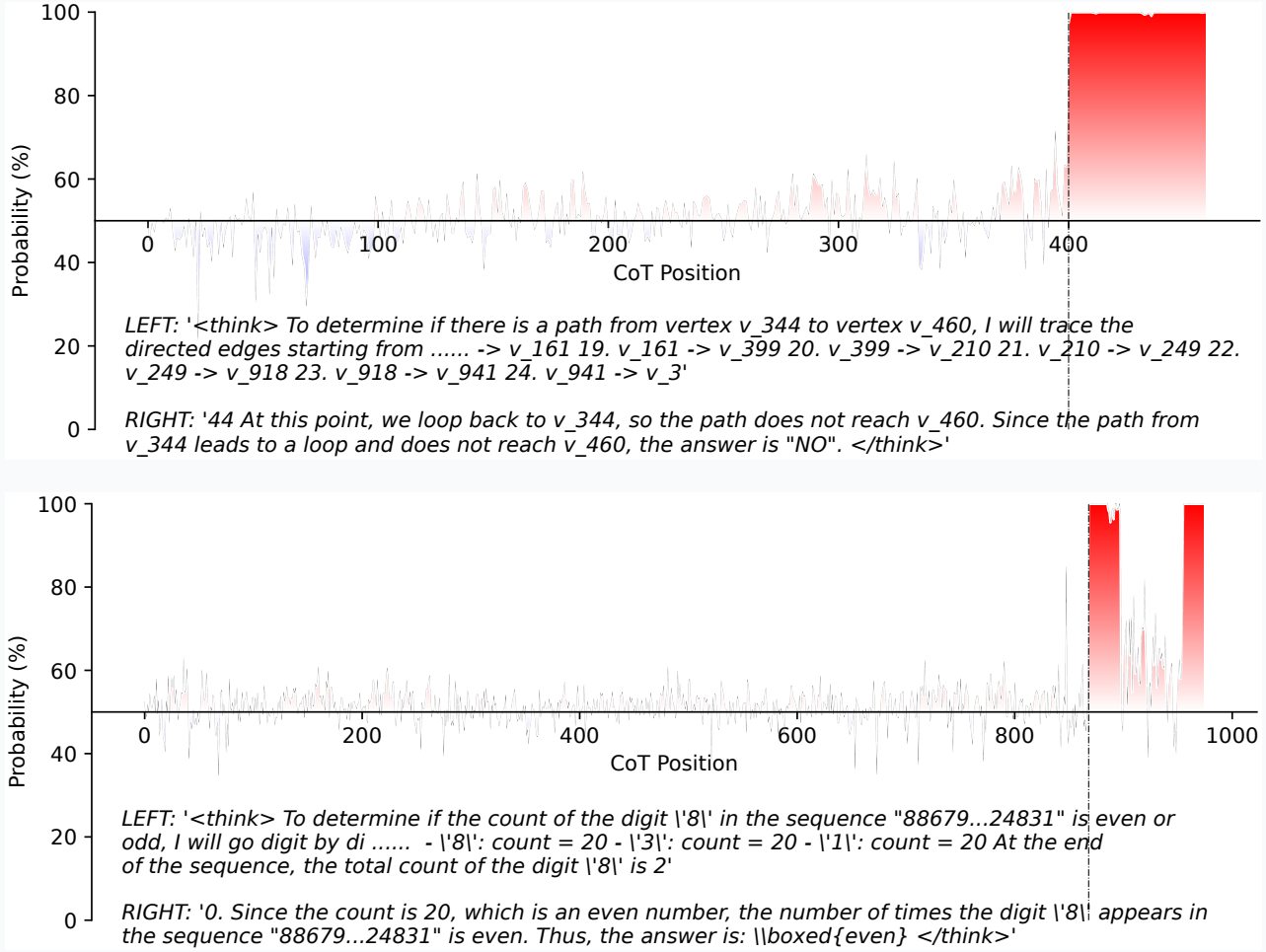
FINDING 1

Precise planning is myopic

KEY FINDING

"For precise future planning, LLMs exhibit a myopic horizon rather than long-term foresight."

The final answer is not reliably encoded far in advance. In compositional tasks, it appears only after the relevant local computation is done.



For compositional tasks, final answer planning only spikes at the end.

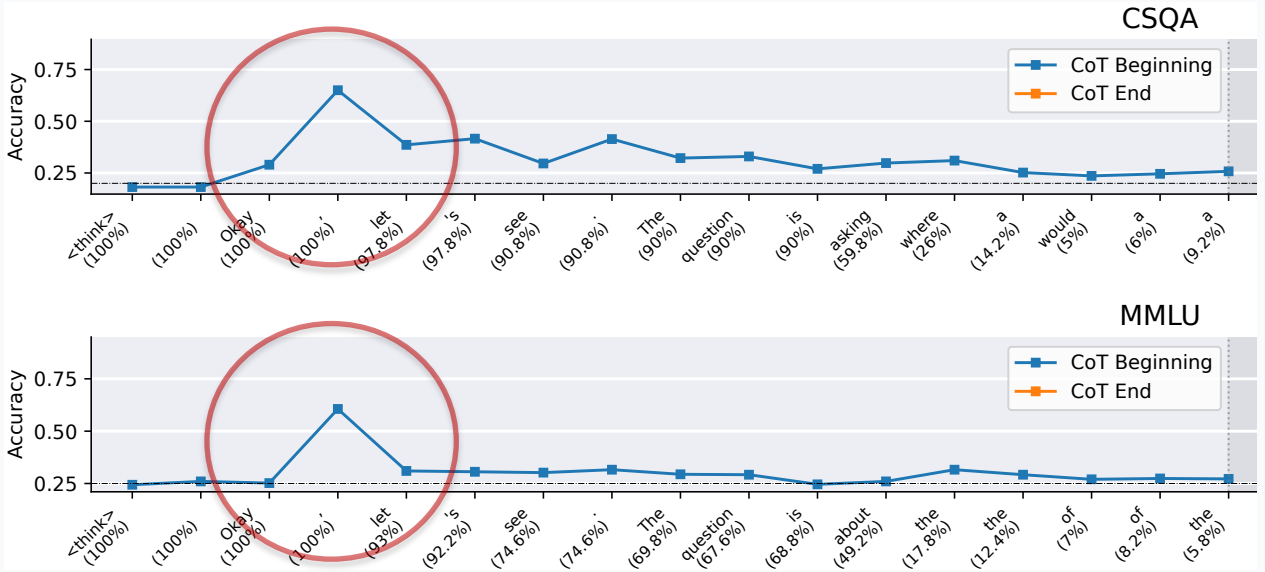
FINDING 2

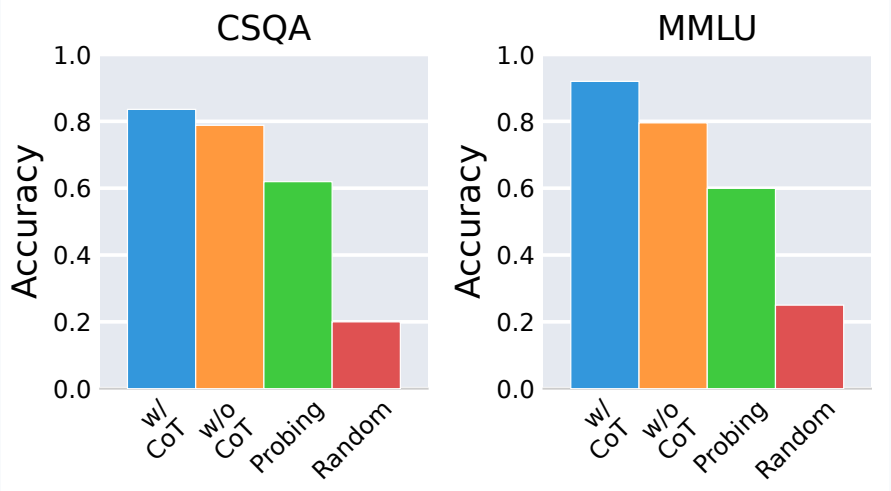
Early signals are coarse, not plans

KEY FINDING

"LLMs can exhibit coarse signals for final answers in early stages of CoT, but reflecting only a pattern-matching gist, rather than precise reasoning plans."

Early hidden states may sense an answer direction, but this is weaker than actual CoT reasoning.





Figures 3-4: early gist exists, but is weaker than direct answering without CoT.

FINDING 3

Hidden states have limited path foresight

KEY FINDING

"LLM hidden states generally encode limited foresight over subsequent reasoning paths."

Next-token-level foresight decays quickly as the target moves farther into the future, especially on semantic and factual tasks.

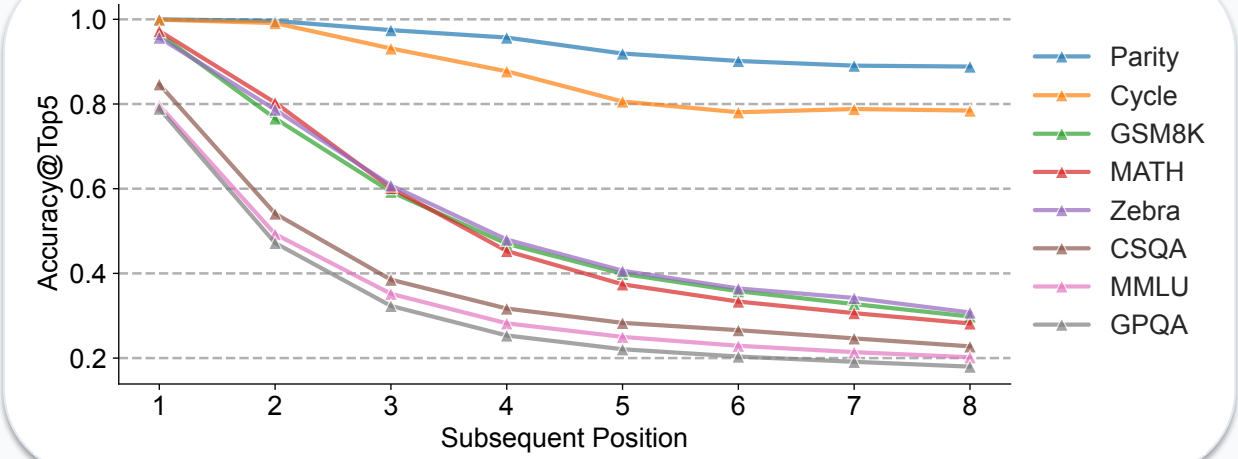


Figure 5: top-5 accuracy drops as the subsequent position becomes farther away.

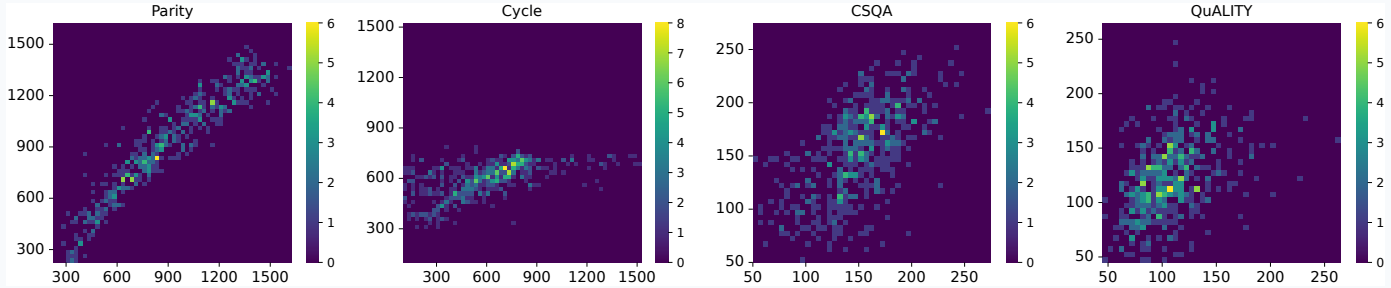
FINDING 4

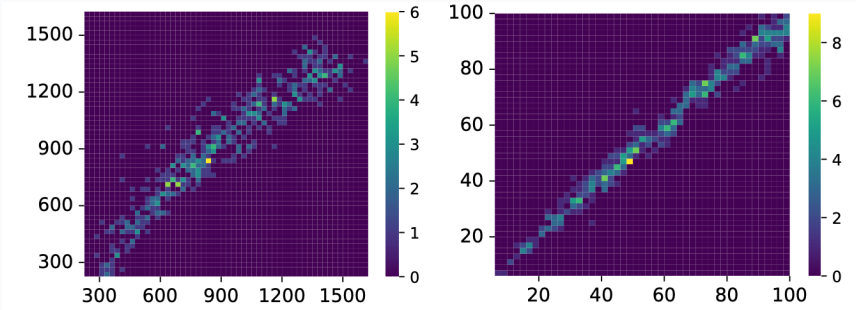
No reliable internal clock for global length

KEY FINDING

"LLMs do not grasp the reasoning length precisely, though task-specific heuristics may offer shortcuts."

If the model had a global roadmap, early states should predict total reasoning length. The paper shows this is unreliable across tasks.





(a) Reasoning length predictions for Parity.

(b) Input sequence length predictions for Parity.

Figure 6: length predictions are generally unstable; apparent successes can be shortcut-driven.

UNCERTAINTY

Reasoning uncertainty is governed by pivots

KEY FINDING

"Just like a barrel's capacity is determined by its shortest stave, the reliability of a reasoning chain is governed by a small number of pivot positions."

Because of the myopic planning, a few local critical tokens can be more informative than averaging confidence over the whole CoT path.

Latent Signals by **Tele-Lens**: top-k most certain positions

	GSM8K	Zebra	MMLU	GPQA	**Avg.**
Perplexity	0.70	0.58	0.53	0.50	0.57
Entropy	0.72	0.60	0.52	0.50	0.58
Self-Certainty	0.76	0.67	0.53	0.51	0.60
Tele-Lens (Top-5)	**0.87**	**0.77**	**0.73**	**0.56**	**0.69**
Tele-Lens (Top-10)	0.81	0.75	0.72	0.56	0.68
Tele-Lens (Top-20)	0.82	0.67	0.65	0.51	0.63
Tele-Lens (Top-50)	0.78	0.69	0.56	0.47	0.64

Latent Signals by **General Metrics**: top-k most uncertain positions

	GSM8K	MATH	MuSR	Zebra	CSQA	MMLU	QuALITY	GPQA	**Avg.**
Perplexity	0.71	**0.93**	0.48	0.74	0.68	0.76	0.78	0.69	0.72
w/ 100 Pivots	**0.81**	0.92	**0.50**	**0.90**	**0.74**	**0.81**	**0.82**	**0.73**	**0.78**
Entropy	0.71	**0.92**	0.47	0.77	0.68	0.77	0.77	0.68	0.72
w/ 100 Pivots	**0.81**	0.70	**0.49**	**0.90**	**0.74**	**0.83**	**0.82**	**0.74**	**0.75**
Self-Certainty	0.45	0.82	**0.47**	0.92	0.51	0.67	0.64	0.68	0.65
w/ 100 Pivots	**0.55**	**0.90**	**0.47**	**0.93**	**0.59**	**0.74**	**0.70**	**0.70**	**0.70**

Top-k pivots improve calibration: Tele-Lens Top-5 reaches 0.69 AUROC; Qwen3-32B general metrics improve from 0.72 to 0.78.

How Far Ahead Do LLMs Plan? Uncovering the Latent Horizon in Chain-of-Thought Reasoning - Xu et al., ICML 2026

NECESSITY

Some CoT can be bypassed

KEY FINDING

"By selectively bypassing CoT generation in non-essential cases, we can achieve a reduction in computational load with negligible performance degradation."

Early answer-gist signals can identify easier cases where full reasoning is not necessary.

Strategy: utilizing early answer gist

16.2% CSQA / 12.4% MMLU CoT bypass with only 0.03 overall accuracy drop

	Parity	CSQA	MMLU	GPQA	Avg.	Perf.
In-Domain LLM						
Th=0.1	0%	40.2%	30.4%	7%	13.3%	-0.47
Th=0.2	0%	65%	45%	12%	21.6%	-1.42
Off-the-Shelf LLM (Qwen3-32B)						
Th=0.1	0%	16.2%	12.4%	1.2%	2.8%	-0.03
Th=0.2	0%	28.8%	20.2%	3.2%	6.2%	-0.37

Proof of concept: bypass non-essential CoT

Contact

How Far Ahead Do LLMs Plan?

Uncover the Latent Horizon in Chain-of-Thought Reasoning

Data & Code:
https://github.com/lxucs/tele-lens

Contact: Liyan Xu
liyanlxu@tencent.com