Trace as State

Reasoning Traces as Conditional States for Long-Context Transformers

Paper authors: Xu Zou (Z.ai), Jie Tang (Tsinghua University)

An independent paper-reading presentation

arXiv:2609.02702v1 · 2 September 2026

Reading: 30 September 2026

Put previously discovered task information before the next reading pass.

Useful task state may arrive after the evidence

The mismatch

A causal model encodes a long context before later reasoning resolves what to look for.

- A target, a resolved reference, or a search frontier can emerge late.
- Later reasoning can use the existing cache, but cannot revise earlier token representations in that pass.
- A fresh pass can make that information available while the context is encoded again.

Paper, Sec. 1, PDF pp. 1–2; Sec. 3.2, p. 4.

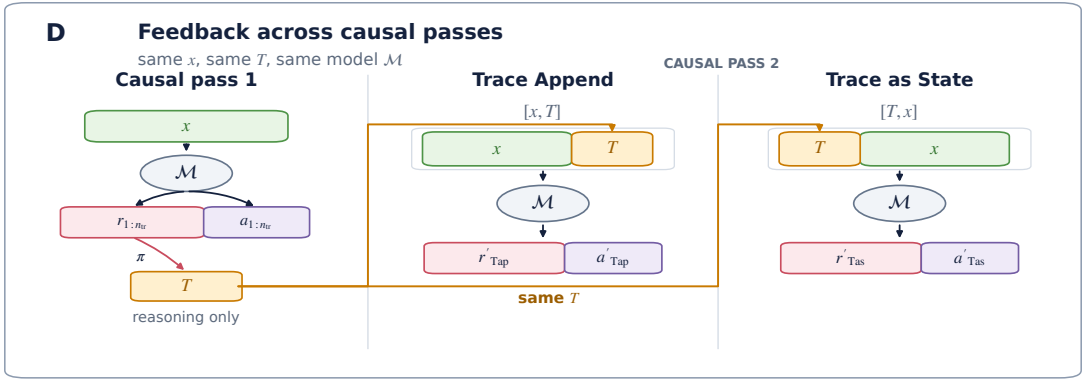
The central experiment changes where the same trace appears

First pass	x: long context	q	
Trace Append	x: long context	T: prior traces	q
Trace as State	T: prior traces	x: long context	q

- Same model, same long context, same realized trace block T.
- The question q stays at the end in the evaluated prompts.
- The comparison isolates a placement intervention, not a unique internal mechanism.

Paper, Sec. 3.3, Eq. (5), p. 5; Sec. 4.1, pp. 5–6; final bullet: reading analysis.

Feedback crosses passes; each pass stays causal



- Source reasoning fields become T; separate answer fields are excluded.
- Models remain frozen. The paper's diagram suppresses the final question.

Original paper, Fig. 1D, PDF p. 3, cropped from the author's vector figure; App. C, p. 13.

Knowing the initial condition selects one state path

$$s_i = U(s_{i-1}, c_i), \qquad s_0 = z, \qquad i = 1, \ldots, n.$$

- $C = (c_1, \ldots, c_n)$: information sequence; U: known fixed update rule.
- $s_i \in \mathcal{S}$: current task state; z: supplied initial condition.
- Let $m = |\mathcal{S}|$ and $b = \log_2 m$.

Condition first: $[z, C]$

Keep one $\mathcal{S}$-valued register and update it as each unit arrives.
Persistent input-dependent memory: $\lceil b \rceil$ bits suffice.

Paper, Sec. 3.1, Eq. (1), PDF pp. 3–4; App. A, p. 12.

A late condition may require an entire response map

For each sequence, define $\phi_C(z) = t(C, z)$ and let $\mathcal{K}$ count distinct maps.

1. Read C before knowing which z will be queried.
2. Different maps must produce different persistent configurations.
3. Otherwise, a shared suffix z forces an incorrect answer for one map.

$$\text{Late-order memory} \;\geq\; \lceil \log_2 \mathcal{K} \rceil.$$

Worst-case construction: all m^m self-maps are attainable

Late: $\lceil b2^b \rceil$ bits Early: $\lceil b \rceil$ bits.
For $m = 4$: 256 possible maps require 8 bits; one state needs 2 bits.

Paper, Eq. (2), PDF p. 4; App. A construction and distinguishability proof, p. 12.

The theorem motivates an ordering principle

What is proved

- Deterministic, exact, one-pass computation.
- An adversarial task family with every state self-map.
- All persistent writable input-dependent storage is counted.

How the paper uses it

- Finite-context, finite-precision causal models fit a finite-state abstraction.
- Text traces are imperfect proxies, not the formal condition z.
- No exponential GPU-memory reduction is demonstrated.

Paper, Sec. 3.2, p. 4; App. A scope, pp. 12–13. Last point: evidence boundary.

The implementation reuses five imperfect reasoning traces

$$(r_j, a_j) \sim \mathcal{M}(\cdot \mid [x, q]), \qquad T = \pi(r_1, \ldots, r_5).$$

- Fixed ordering, labels, delimiters, and a warning to verify against the task.
- Each reused trace is capped at its first 50,000 characters.
- Separate visible answers a_j are excluded; traces can still contain answer-relevant information.
- Five fresh second-pass repeats per placement are averaged for evaluation.

Deployment accounting

Five source calls plus one feedback call are needed for one answer with five source traces.

Paper, Sec. 4.1, pp. 5–6; Sec. 4.4, p. 7; App. C, pp. 13–14.

Evaluation spans graph state, retrieval, and novel reading

Task	Evaluated problems	Metrics (↑)
GraphWalks 256K	100 BFS + 100 Parents	Set EM, set F1
MRCRv2, 8-needle	100 at 256K + 100 at 512K	EM, Seq. ratio
NUB-1M Season 2	20 novel questions	Judged accuracy

- Models: DeepSeek V4 Pro Preview, Qwen 3.7 Max, GLM-5.2.
- Five repeats; invalid, missing, overlong, or filtered outputs score zero.
- NUB uses a DeepSeek judge; its novel is about 400–700K tokens.

Paper, Tables 1-1 and 1-2, p. 6; Apps. B and E, pp. 13, 18. Seq. is not set F1.

Parents shows the largest placement gains

GraphWalks Parents: exact match (%, higher is better)				
Model	First	Append	State	Gain vs. Append
DeepSeek	29.2	43.0	**81.8**	+38.8 pp
Qwen	60.8	71.0	**96.4**	+25.4 pp
GLM	66.4	83.2	**100.0**	+16.8 pp

- DeepSeek gains 52.6 percentage points over the first pass.
- All three paired Parents EM intervals are wholly above zero.
- GLM's 100% applies to the evaluated problems and repeats.

Paper, Table 2, PDF p. 7; paired intervals in Table 5, p. 18.

BFS gains vary; exact match and F1 can disagree

GraphWalks BFS: EM / set F1 (%)			
Model	First	Append	State
DeepSeek	31.6 / 36.4	41.8 / 48.3	**58.8** / **65.9**
Qwen	60.0 / 68.1	60.4 / 70.4	**63.8** / **71.7**
GLM	55.8 / 70.7	60.0 / **75.8**	**63.4** / 75.0

- GLM BFS F1 is the sole mean-score loss among 27 reported comparisons.
- More fully correct sets need not imply better partial overlap elsewhere.
- Qwen BFS F1 and both GLM BFS paired intervals cross zero.

Paper, Tables 2 and 5, PDF pp. 7, 18.

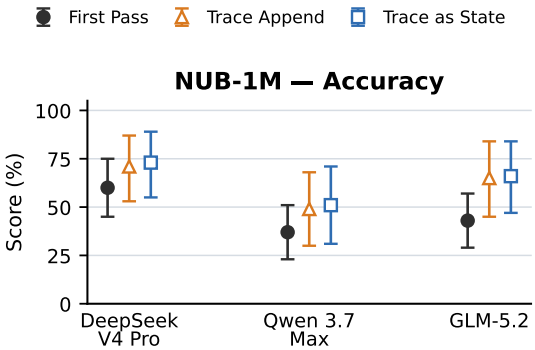
Retrieval and binding also benefit from earlier traces

MRCRv2 8-needle: exact match (%)					
Model	Bin	First	Append	State	Difference (pp)
DeepSeek	256K	53.8	66.6	76.8	+10.2
DeepSeek	512K	45.4	49.4	52.6	+3.2
Qwen	256K	79.8	84.0	88.4	+4.4
Qwen	512K	36.2	40.4	46.8	+6.4
GLM	256K	40.2	40.0	61.2	+21.2
GLM	512K	42.6	44.6	55.4	+10.8

DeepSeek 512K EM: paired 95% interval $[-2.6, +9.4]$ pp crosses zero.

Paper, Table 2, p. 7; Table 5, p. 18. Length bins are different problem sets.

Novel reading gives weaker evidence for placement



Accuracy (%)	Append	State
DeepSeek	71.0	73.0
Qwen	49.0	51.0
GLM	65.0	66.0

- Only 20 questions; wide intervals.
- Extra placement gain: 1–2 pp.
- No paired NUB contrast is reported in Table 5.

Paper, Table 2, p. 7; Fig. 5 NUB panel, p. 19. Bars are problem-bootstrap 95% intervals.

26 of 27 higher means is not 26 significant wins

Paired uncertainty for GraphWalks and MRCR

20 of 24 unadjusted 95% intervals lie wholly above zero.
Average five repeats within each problem; resample 100 problems per cell.

The four intervals crossing zero	Difference (pp)	95% interval
DeepSeek MRCR 512K EM	+3.20	$[-2.60, 9.40]$
Qwen BFS F1	+1.29	$[-0.56, 3.90]$
GLM BFS EM	+3.40	$[-1.00, 7.80]$
GLM BFS F1	−0.83	$[-3.94, 1.62]$

Paper, App. E and Table 5, p. 18: 20,000 bootstrap resamples, seed 0; no multiplicity correction.

Task-specific traces beat several simpler feedback choices

DeepSeek GraphWalks	Prompt	BFS EM	Parents EM
First pass	$[x, q]$	31.6	29.2
Question First	$[q, x, q]$	33.4	53.0
Re2 (evaluated form)	$[x, q, x, q]$	49.0	50.0
Answer Feedback	$[a, x, q]$	35.4	45.4
Random Trace	$[T_{\text{rand}}, x, q]$	22.4	14.2
TRACE AS STATE	$[T, x, q]$	**58.8**	**81.8**

- Generic trace formatting and answer access do not explain the full gain.
- Question First has 182 zero-scored termination/format failures.

Paper, Table 3, p. 8; App. B, p. 13. Scores are percentages; controls are not cost matched.

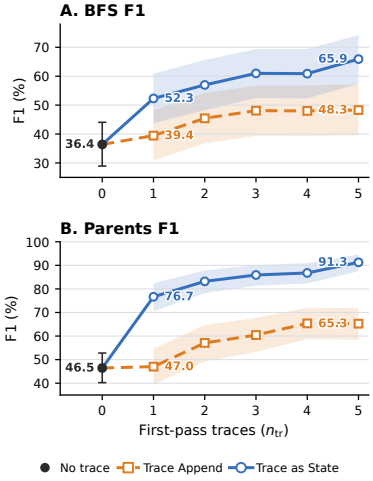
The feedback pass can improve on all five existing answers

DeepSeek GraphWalks	BFS EM (%)	Parents EM (%)
Majority@5	35.0	31.0
Oracle@5	50.0	50.0
Trace Only: $[T, q]$	46.2	43.8
Trace Append: $[x, T, q]$	41.8	43.0
Trace as State: $[T, x, q]$	**58.8**	**81.8**

- Traces already contain useful information; rereading adds further value.
- Oracle@5 selects the best score retrospectively using ground truth.
- Exceeding it uses new computation, not a free improvement at equal cost.

Paper, Table 3, p. 8; Fig. 6 caption, p. 20. Majority@5 votes on individual set elements.

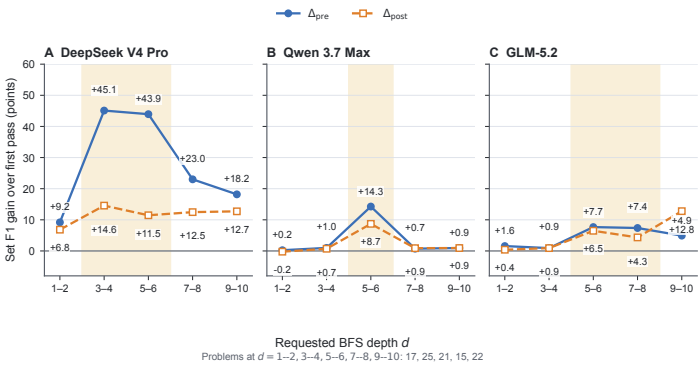
One trace helps; additional traces generally help more



- DeepSeek, GraphWalks 256K only.
- Nested first 1–5 eligible traces; same T for both placements at each count.
- With one trace, Parents F1 is 76.7% before versus 47.0% after context.
- More traces increase both information and inference cost.

Original Fig. 2, PDF p. 8; Sec. 4.4, pp. 7–8. Shading: 95% problem-bootstrap intervals.

Difficulty profiles describe heterogeneity



- Plotted values are gains over the shared first pass, not absolute scores.
- Yellow ranges were chosen after seeing outcomes; they do not establish a capacity threshold or mechanism.

Paper, Fig. 4 top panels, p. 17; App. D, pp. 14–18. Original author figure.

Quality gains come with extra inference and variable caching

DeepSeek task	Total usage (million tokens)	Append	State
GraphWalks	Output, including reasoning	97.738	90.482
NUB	Uncached input	12.878	25.055
NUB	Output, including reasoning	1.001	1.215

- Totals include the shared source stage plus the feedback stage.
- Prefix changes can reduce reuse of the original context cache.
- Output length and cache behavior vary across tasks and models.
- No measured universal speedup or equal-cost frontier is established.

Paper, Limitations, p. 9; App. F and Table 6, pp. 20–21. Retained responses with API usage only.

The evidence has clear interface and scope boundaries

- **Interface:** this realization needs exposed reasoning traces; answer-only APIs require another state interface.
- **Trace quality:** incomplete or incorrect text can mislead the new pass.
- **Coverage:** three models, three task families; no multi-round agent evaluation.
- **Reproducibility:** effective provider output caps are unknown; full NUB-specific templates are not given.
- **Mechanism:** the intervention supports conditional rereading, but does not identify a faithful internal state representation.

Paper, Sec. 3.2, p. 4; Sec. 4.1, p. 6; Limitations, p. 9; App. C, pp. 13–14.

The usable idea is conditional rereading

Empirical result

Moving the same prior traces before the context usually helps, with especially large gains on GraphWalks Parents.

Conceptual contribution

When information becomes available can matter as much as whether it is available.

Next experiments

Compare compressed states and trace selection at equal total cost; test preregistered difficulty contrasts and multi-round use.

Paper, Sec. 5, p. 8, and Limitations, p. 9. Equal-cost and preregistered tests: reading suggestions.

Source and reading map

Xu Zou and Jie Tang. 2026.
Trace as State: Reasoning Traces as Conditional States for Long-Context Transformers.
arXiv:2609.02702v1 (2 September 2026).

- Method and theory: Sec. 3, pp. 3–5; Appendix A, pp. 12–13.
- Results and controls: Tables 2–3 and Fig. 2, pp. 7–8.
- Uncertainty and cost: Appendix E–F, pp. 18–21.

Prepared from the versioned PDF and the author's TeX source.
Original figures are attributed on their slides; tables are selected/reformatted from the paper.
All following slides are optional backup material.

Backup: model budgets are not verified request caps

Model	Input budget	Output budget	Reasoning
DeepSeek V4 Pro Preview	1,048,576	131,072	max
Qwen 3.7 Max	983,616	65,536	xhigh text
GLM-5.2	1,048,576	65,536	max

- The authors requested no custom maximum-output value; effective provider defaults are not established by the records.
- DeepSeek uses the 2026-04-24 version, not the 2026-08-13 release.
- Qwen uses prompt text, not a provider-side reasoning-effort parameter.

Paper, Table 1-1 and Sec. 4.1, p. 6; App. C.2, pp. 13–14. Budgets in tokens.

Backup: scoring includes interface failures

- **GraphWalks:** parse the terminal answer line as a node set. Both sets empty scores one; malformed or nonterminal output scores zero.
- **Question First:** 121 null termination markers plus 60 length-limited outputs, and one malformed answer, all score zero.
- **MRCR:** validate and strip the random prefix, then calculate exact match and SequenceMatcher ratio. Filtering remains in the denominator.
- **NUB:** judge correctness against reference answers without solver identity or placement labels; authors report manual checking.

Paper, App. B, PDF p. 13. Question First failures: 161 BFS and 21 Parents outputs.

Backup: all GraphWalks main-result scores

EM / set F1 (%); five-repeat means; higher is better				
Model	Subtask	First	Append	State
DeepSeek	BFS	31.6 / 36.4	41.8 / 48.3	58.8 / 65.9
DeepSeek	Parents	29.2 / 46.5	43.0 / 65.3	81.8 / 91.3
Qwen	BFS	60.0 / 68.1	60.4 / 70.4	63.8 / 71.7
Qwen	Parents	60.8 / 87.2	71.0 / 91.7	96.4 / 99.1
GLM	BFS	55.8 / 70.7	60.0 / 75.8	63.4 / 75.0
GLM	Parents	66.4 / 88.5	83.2 / 92.7	100.0 / 100.0

Paper, Table 2, PDF p. 7. 100 problems per subtask.

Backup: all MRCR main-result scores

EM / SequenceMatcher ratio (%); higher is better				
Model	Length bin	First	Append	State
DeepSeek	256K	53.8 / 78.5	66.6 / 79.6	76.8 / 88.7
DeepSeek	512K	45.4 / 63.6	49.4 / 66.9	52.6 / 73.3
Qwen	256K	79.8 / 83.1	84.0 / 87.1	88.4 / 91.3
Qwen	512K	36.2 / 43.7	40.4 / 47.4	46.8 / 52.5
GLM	256K	40.2 / 48.2	40.0 / 58.3	61.2 / 71.9
GLM	512K	42.6 / 52.0	44.6 / 59.9	55.4 / 70.1

Paper, Table 2, PDF p. 7. 100 problems per bin, five repeats.

Backup: the worst case uses a very rich input alphabet

Choose $n = 1$ and

$$\mathcal{C} = \{c_f : f \in \mathcal{S}^{\mathcal{S}}\}, \qquad U(s, c_f) = f(s).$$

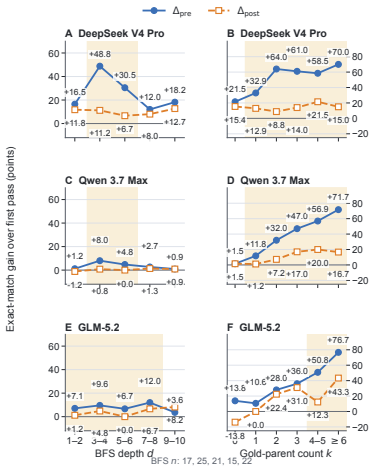
- Every possible self-map is encoded by one input symbol: $\mathcal{K} = m^m$.
- Late order: store which map arrived before the queried z is known.
- Early order: store z, then replace it with $f(z)$.
- Identity-map suffixes also prove the early cut needs at least m configurations.

Accounting matters

The fixed transition rule is known. The bound counts persistent input-dependent state, not its implementation cost or the transition-table size.

Paper, App. A, PDF p. 12. Restricted update families can yield far fewer residual maps.

Backup: difficulty is not a monotone placement effect



- Rows: DeepSeek, Qwen, GLM. Left: BFS depth. Right: parent-set size.
- Blue/orange: State/Append gains over the common first pass.
- GLM BFS reverses in the final depth bin.
- Parents EM never favors Append in these plotted bins.
- Bins and highlighted ranges are exploratory, outcome-informed analyses.

Paper, original Fig. 3, p. 16; App. D, pp. 14–18.