EPOCH

Preventing Error Cascades in Long-Horizon Multimodal Agents with Edge-Reliability Graph Memory

Saman Forouzandeh · Wei Peng · Xinghuo Yu · Mahdi Jalili — School of Engineering, RMIT University, Melbourne

Why errors cascade

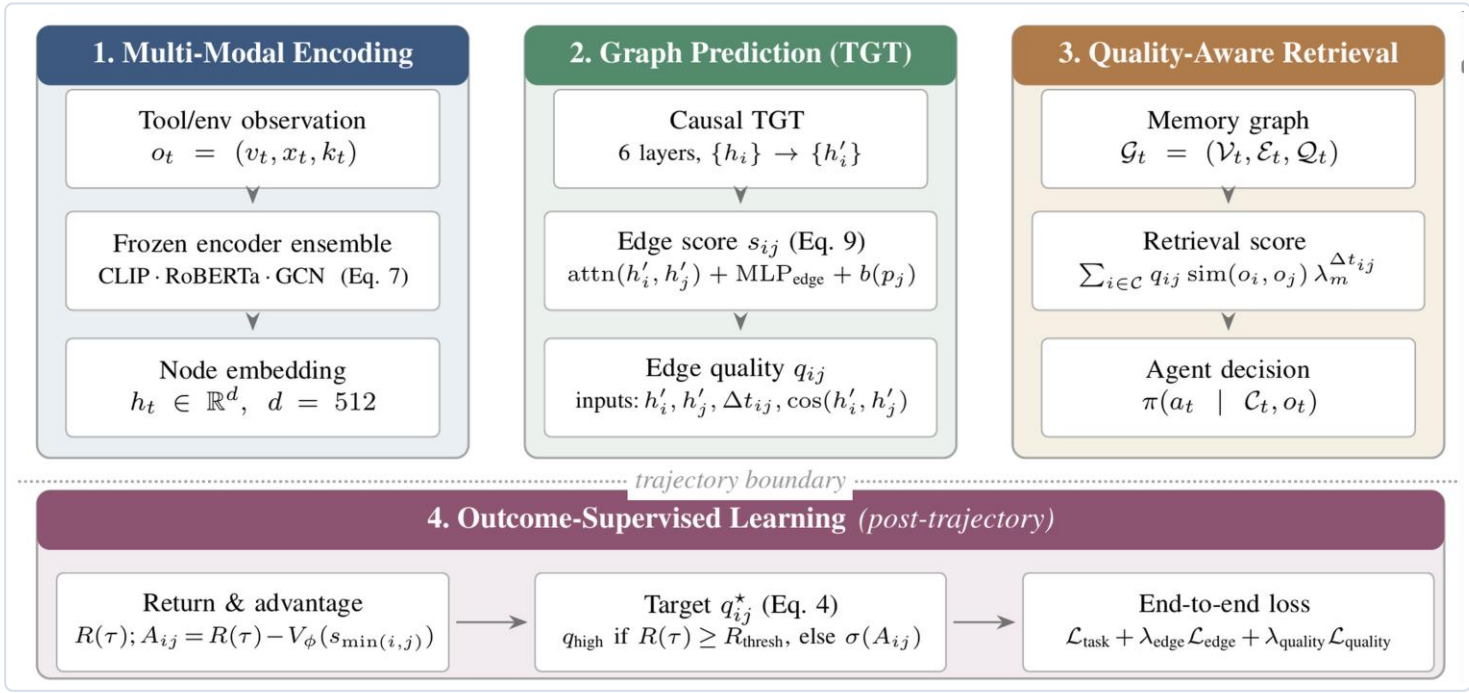
A small tool error is stored, retrieved at later steps, and amplified into a **cascading failure** — a memory-propagation problem, not an isolated reasoning error.

Tesla cascade: "Musk is CEO of Tesla" is correct but misleading for "who founded Tesla?" — reused until it dominates retrieval.

Key idea: score the pathway, not the item

One observation is reliable in one context, misleading in another. EPOCH scores the **directed reuse edge** ($q_{ij} \in [0,1]$), suppressing unreliable evidence upstream before it propagates.

EPOCH framework



Encode → causal Temporal Graph Transformer predicts edge score s_{ij} & quality q_{ij} with positional $b(p_j)$ → quality-weighted retrieval; edge utility is outcome-supervised from trajectory returns (no manual labels).

$$\mathbf{score(o_j) = \Sigma_i\, q_{ij} \cdot sim(o_i,o_j) \cdot \lambda_m\text{^}(\Delta t_{ij})} \quad \textbf{— quality-weighted retrieval (Eq. 14)}$$

Results & why it works

Native long-horizon search — Pass@1 (5 seeds)

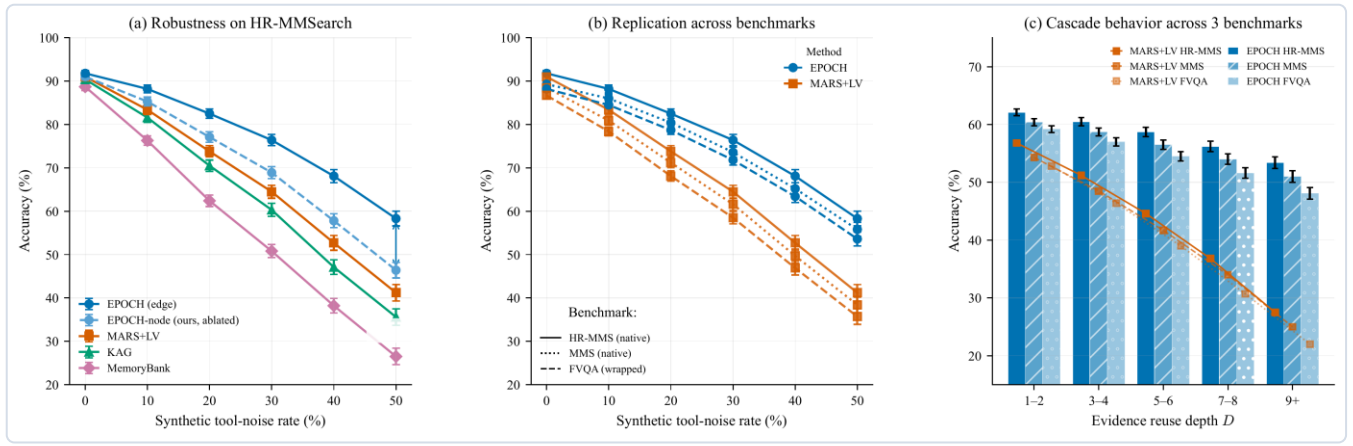
Method	MMS	HR-MMS
MARS+LV	71.2	45.6
MemoBrain	72.0	47.1
EPOCH-Node	77.1	53.8
EPOCH (ours)	**82.6**	**59.6**
Δ vs MARS+LV	**+11.4**	**+14.0**

Takeaway. Long-horizon agent failure is propagation-driven, not reasoning-driven — best addressed at the reuse pathway, not the item.

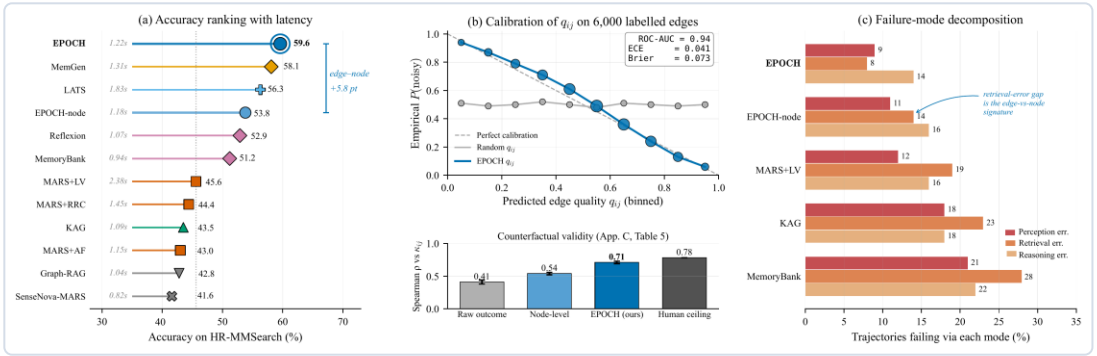
ρ = 0.71
edge-utility validity (vs 0.54 node-level)

0.94 / 0.041
ROC-AUC / ECE (human labels)

19% → 8%
retrieval-error failures



EPOCH degrades most gracefully as tool-noise rises; the margin replicates across benchmarks and grows with reuse depth D (gain over MARS+LV: +12.2 at 50 steps → +20.6 at 500).



Learned edge utility is calibrated (ROC-AUC 0.94) and tracks counterfactual importance (ρ=0.71); EPOCH wins at competitive latency and cuts retrieval-error failures 19%→8%.

Code & paper: https://github.com/S-Forouzandeh/EPOCH_NeurIPS-2026

saman.forouzandeh@rmit.edu.au