SQCAD: LIFECYCLE-AWARE AUTHORIZATION FOR SCOPE-QUALIFIED PERSISTENT MEMORY IN LONG-TERM AGENTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Agent-memory systems typically rank records by their usefulness to the current query, but persistent authorization asks a different question: whether a record may keep influencing decisions that have not arrived. Relevance is calibrated on recorded use, whereas authorization is a claim about decisions that history has not yet settled; refining the first does not by itself answer the second. To address this problem we formalize persistent authorization as a sequential decision problem, first proving that any binary rule reading a score incurs unavoidable worst-case regret when it conflates states whose optimal keep/archive decisions differ, and deriving an exact fixed-continuation decomposition separating representation loss from execution error. Guided by both results, we propose SQCAD, a lifecycle controller whose qualification, authorization, competitive access, and bounded recovery share one scope- and lineage-qualified evidence state. Separating persistent permission from temporary access lets an archived record answer a query without being reauthorized. LifecycleBench makes these decisions measurable by scoring both branches of every keep/archive action under a common continuation. Experiments on the public tasks place the structured-prior guard first: it attains the highest frozen LoCoMo F1 of the evaluated methods at 97.6% less authorized decision memory than the ActMem adapter, and of the LongMemEval-S comparisons that survive Holm correction, 32 of 34 favour SQCAD. Lineage, one retained field, lowers the representation floor by an order of magnitude where it targets it and leaves it unchanged elsewhere. The criterion binds any system that commits a persistent decision on a scalar score, whether the decision governs retention, eviction, or admission. Qualification and lineage raise downstream accuracy alike, so the metrics the field reports cannot tell them apart, highlighting that persistent authorization is a property of the memory decision rather than a by-product of retrieval quality.

1 INTRODUCTION

Agent memory carries facts, constraints, and intermediate conclusions across interactions. Retrieval evaluates whether a record helps the current query; persistent authorization determines whether it remains eligible to influence later decisions. A once-useful instruction may become invalid after a scope change, a frequently retrieved record may have been superseded, and an infrequently used record may be the only recoverable bridge to a future task. Each case depends on the consequences of retaining access, including which observations retention will make available.

Existing memory systems differ substantially in what they retain, from hierarchical paging and reflective streams (Packer et al., 2023; Letta Team, 2025; Park et al., 2023) to associative graphs (Xu et al., 2025; Rasmussen et al., 2025) and typed or verified stores (Wang & Chen, 2025; Zhong et al., 2026), and in how they revise it, whether by write-time compression (Liu et al., 2026), outcome-conditioned worth (Simsek, 2026), or temporal causal revision (Chang, 2026); surveys and evaluation suites map this space further (Zhang et al., 2025; Du et al., 2025). Despite these differences in representation and revision, the persistent commitment remains tethered to a single scalar interface: where such a system commits a keep/archive decision, it reads a relevance score over the current query. Consider two histories containing the same query-matching instruction: in one a compatible

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

replacement has already appeared, in the other the update has not yet arrived. A relevance score assigns the same value to both although the correct persistent actions differ, and no amount of calibration restores a distinction the interface cannot carry. The obstruction is informational, and it sits at the decision rather than at the retrieval that precedes it.

Research gap. Which information must therefore be retained to determine a memory's long-horizon authorization, and how should a controller act when that information is insufficient? We study the keep/archive decision itself rather than the retrieval built on top of it, because that decision changes not only what a later query retrieves but what a later decision observes. Current instruments do not reach that decision: on a natural trajectory the kept and archived branches are never compared under one continuation, so a representation can look adequate on downstream QA while discarding the distinctions that decide future authorization. Measuring this needs three things not presently available together: a lower bound for score-only commitment, a decomposition into representation and execution loss, and a paired setting that observes both branches under one continuation.

We supply all three. Guided by the first two, we build the controller and release the instrument on which each term is measurable: the bound fixes which distinctions the qualification state has to carry, and the decomposition decides whether a remaining loss calls for a richer state or a better rule. SQCAD (*Scope-Qualified Competitive Access Decay*) is that controller. It represents each record with a qualified evidence state: lineage, scope/version, validity, and censoring/recovery. LifecycleBench supplies the instrument, pairing keep/archive futures under one continuation so that both loss terms become separately computable.

Contributions.

- **A qualification state, and what it must record** (Section 3, evaluated here on the binary keep/archive decision). The bound fixes which distinctions cannot be dropped; SQCAD carries them as explicit state and separates persistent permission from temporary exposure, so an archived record can answer one query without being reauthorized. One of those fields, retained rather than discarded, moves the measured floor by an order of magnitude on the failure mode it targets, and leaves the floor unchanged outside that mode (Section 4.4).
- **A lower bound for score-only commitment** (Theorem 1). When a score conflates two states whose optimal persistent actions have opposite signs, every randomized score-only rule incurs worst-case regret bounded away from zero, so the obstruction belongs to the information rather than to the policy.
- **An exact decomposition of commitment regret** (Theorem 2). Under a fixed continuation, regret splits into an irreducible information term $\mathcal{I}(Z)$ and a nonnegative execution term $\mathcal{E}(h; Z)$, separating an inadequate representation from a rule that uses its input badly.
- **A paired-intervention instrument** (Section 4). LifecycleBench, a released benchmark of 225 counterfactual worlds scoring both branches of every keep/archive action under a common continuation, with a provenance ledger separating native values from matched-contract adapters (Appendix D).

2 Persistent authorization as a decision problem

The question this section formalises is a single one: at time t, should this record stay eligible to influence what comes later? Three things bear on the answer: what the agent believes, the physical state of the store, and what the agent will be able to observe afterwards, which depends on both.

Let $X_t = (B_t, S_t)$ be the post-observation state at time $t \le H$, pairing a belief B_t with a physical memory state S_t whose form Section 3 specifies. An action $a \in \{\mathsf{keep}, \mathsf{archive}, \mathsf{probe}, \mathsf{defer}\}$ changes both, through a joint kernel $K_t^a(dx', do \mid x)$; the marginal T_t^a is taken after the resulting observation, so the Bayesian update is already applied. For bounded utility ℓ_t, terminal payoff g_{H+1}, and $\gamma \in (0, 1]$, define

$$\begin{aligned} Q_t(x,a) &= \ell_t(x,a) + \gamma \int V_{t+1}(x') T_t^a(dx' \mid x), \\ V_t(x) &= \max_a Q_t(x,a), \qquad \Delta_t(x) = Q_t(x, \mathsf{keep}) - Q_t(x, \mathsf{archive}). \end{aligned} \tag{1}$$

Storage, exposure, and probing costs enter the utility. The sign of Δ_t then identifies the preferred commitment: keep when it is positive, archive when it is negative. The benchmark evaluates this contrast under a fixed continuation policy π, written $\tau_s^\pi = V_s^\pi(\mathsf{keep}) - V_s^\pi(\mathsf{archive})$; it does not solve optimal control in general.

The continuation term carries two kinds of consequence. After an action, the physical state becomes S' and the belief becomes $B^{a,S'}$; the agent may then observe something further and update to $B^{a,S',O}$. Writing A_t^a for the value the access change alone is expected to yield, and I_t^a for what the additional observation adds, set

$$A_t^a = \mathbb{E}[V_{t+1}(B^{a,S'}, S') \mid x, a], \qquad I_t^a = \mathbb{E}[V_{t+1}(B^{a,S',O}, S') \mid x, a] - A_t^a. \tag{2}$$

Then $Q_t = \ell_t + \gamma A_t^a + \gamma I_t^a$: the first term is what the action costs now, and the remaining two are what it changes about the future. The identity is bookkeeping rather than a causal claim from observational traces, and it specifies which consequences a proposed summary must preserve.

Two questions follow. When must a score fail, and when it does, how much of the loss belongs to the score rather than to the rule that reads it?

2.1 A SCORE COLLISION CAN FORCE COMMITMENT ERROR

The first question is whether any scalar can serve at all. A *fiber* of a score is the set of states the score maps to one value, and the answer is no whenever a fiber holds states that disagree about the commitment.

Theorem 1 (Score-fiber lower bound)**.** *Fix the utility, horizon, and continuation. Suppose a score R_t maps two states to the same value, while $\Delta_t(x_+) \geq m_+ > 0$ and $\Delta_t(x_-) \leq -m_- < 0$. Every possibly randomized binary commitment rule observing only $R_t(x)$ has worst-case regret at least*

$$\frac{m_+ m_-}{m_+ + m_-}. \tag{3}$$

If the score induces keep probability p, the regret is at least $(1-p)m_+$ on the first state and at least pm_- on the second, so the best a rule can do is balance them, which gives the bound. The premise is essential: a score that already preserves the commitment contrast need not fail. This is a worst-case statement: it says a score must fail somewhere, not how the resulting loss divides. The next subsection gives its fixed-continuation counterpart, and the two coincide under the balancing weights $\pi_+ = m_-/(m_+ + m_-)$.

2.2 REPRESENTATION LOSS AND CONTROLLER ERROR UNDER A FIXED CONTINUATION

Fix a continuation policy π and a distribution over worlds with integrable branch values, and let W be a world drawn from it. Let $\Delta(W) = V^\pi(W, \mathsf{keep}) - V^\pi(W, \mathsf{archive})$. A controller observes only $Z = q(W)$ and chooses a binary action $h(Z)$. Define $\mu(z) = \mathbb{E}[\Delta(W) \mid Z = z]$, $A(z) = \mathbb{E}[\Delta(W)_+ \mid Z = z]$, and $B(z) = \mathbb{E}[(-\Delta(W))_+ \mid Z = z]$. The representation-level Bayes action is $h_Z^*(z) = \mathsf{keep}$ when $\mu(z) \geq 0$ and archive otherwise. Regret $R(h)$ is the expected difference between the better paired branch and the branch selected by h. Controller randomization is independent of W conditional on Z.

Theorem 2 (Fixed-continuation regret decomposition)**.** *For any possibly randomized binary controller restricted to Z,*

$$R(h) = \mathbb{E}[\min\{A(Z), B(Z)\}] + \mathbb{E}\left[|\mu(Z)| \Pr\{h(Z) \neq h_Z^*(Z) \mid Z\}\right]. \tag{4}$$

The first term, $\mathcal{I}(Z)$, is the minimum regret among controllers restricted to Z. The second term, $\mathcal{E}(h; Z)$, is the excess regret from their decisions. Equivalently, $\mathcal{I}(Z) = \frac{1}{2}(\mathbb{E}|\Delta| - \mathbb{E}|\mu(Z)|)$.

Proof. Condition on $Z = z$: keeping incurs expected regret $B(z)$ and archiving $A(z)$, so a controller that keeps with probability $p(z)$ pays $p(z)B(z) + (1-p(z))A(z)$. Subtracting $\min\{A(z), B(z)\}$ gives the second term in Eq. (4), because $|\mu(z)| = |A(z) - B(z)|$.

If Z' refines Z, conditional Jensen's inequality gives $\mathcal{I}(Z') \leq \mathcal{I}(Z)$: more information can reduce the best attainable loss, while an actual controller can still incur greater execution error. Both terms

are standard binary Bayes decision analysis, related to value-preserving state abstraction (Li et al., 2006; Abel et al., 2020): $\mathcal{I}$ is a conditional-mean loss and $\mathcal{E}$ the usual excess over the Bayes action. What the decomposition defines is an estimand, a quantity to be estimated from data rather than asserted, and neither term is recoverable from retrieval quality: a discarded branch is unobserved on a natural trajectory, and both become computable only once paired branches exist. Appendix A.2 gives the proof and held-out diagnostic; Appendix A.3 retains uniform closure as a stronger condition.

3 SQCAD: QUALIFICATION, AUTHORIZATION, AND ACCESS

Section 2 fixes what a persistent commitment must distinguish; this section fixes how the controller decides. SQCAD carries one qualified evidence state per memory and commits one action from it. Figure 1 shows the pipeline end to end: qualification turns public evidence into a certified interval on the commitment contrast, authorization commits a persistent action from where that interval sits relative to zero, and a bounded guard allocates scarce query-time access without changing persistent permission. An interactive walkthrough in the supplementary release traces all four steps on four constructed decisions, so the pipeline can be inspected rather than only read.

3.1 QUALIFICATION: A CERTIFIED INTERVAL

For each memory and decision scope, the state records evidence and its contradictions, lineage links, scope/version, validity, and periods of unavailability. The controller turns that state into an interval $[L, U]$ on the commitment contrast Δ_t (keep minus archive, Eq. (1)), and where the interval sits relative to zero names a qualification: *positive* if $L > 0$, *negative* if $U < 0$, *unresolved* if it straddles zero, and *mismatch* if the record falls outside the decision scope. The comparison is strict: an interval touching zero stays unresolved, because reading a boundary case as certified would silently widen the error rate the certificate declares. The interface builds the interval from successful probes: with $\bar{\Delta}_n$ the mean of n observed contrasts and σ their sub-Gaussian scale, $[\bar{\Delta}_n - b_n, \bar{\Delta}_n + b_n]$ with $b_n = \sigma\sqrt{2\log(\pi^2 n^2/(3\alpha))/n}$ contains the true contrast with probability at least $1-\alpha$ simultaneously for every n. Two limits bound it: the radius covers estimation error, not the approximation error of a compressed state (Appendix A.3 gives a sufficient condition no controller here estimates), and the guarantee is conditional on a declared margin, so a near-zero contrast stays unresolved by design rather than resolved at unbounded cost. Controllers below read the qualification directly, taking its sign from rules rather than materializing an interval.

3.2 AUTHORIZATION: COMMITTING FROM THE INTERVAL

Commitment is then a rule on the qualification rather than a threshold on a score. The qualification rule archives negative and mismatched evidence and keeps the rest. The lineage rule also archives unresolved conflicts, so a superseded record whose replacement has not arrived does not stay authorized. The probe-willing rule keeps only what is positive and lets the guard recover the rest. A negative qualification forbids later probing, so archiving is not always reversible in one step.

3.3 COMPETITIVE ACCESS AND THE GUARD

Authorization sets persistent permission; access is a separate and scarcer resource, since one record's active slot is another's lost capacity. Records enter chronologically; when the retained set exceeds the item budget the controller archives the weakest eligible record under support, protection, and version rules, breaking ties stably, and ranking and packing then hold the reader within its token budget. For the authorized set A_t, candidate pool C_t, and guard budget k, the reader may access $G_t(k) = A_t \cup \mathrm{TopK}_k(C_t \setminus A_t)$. That exposure is temporary and leaves persistent permission untouched; it is meaningful only because archived records stay recoverable. The cost the design reduces is therefore authorized memory and reader exposure, not physical storage. Deletion would need a different objective and experiment.

3.4 TWO INSTANTIATIONS OF ONE VIEW

The LLM controllers read a *view* instead of the interval: the public fields the controller is shown before it commits. FIR and SP are not independent methods but two such views, the second built

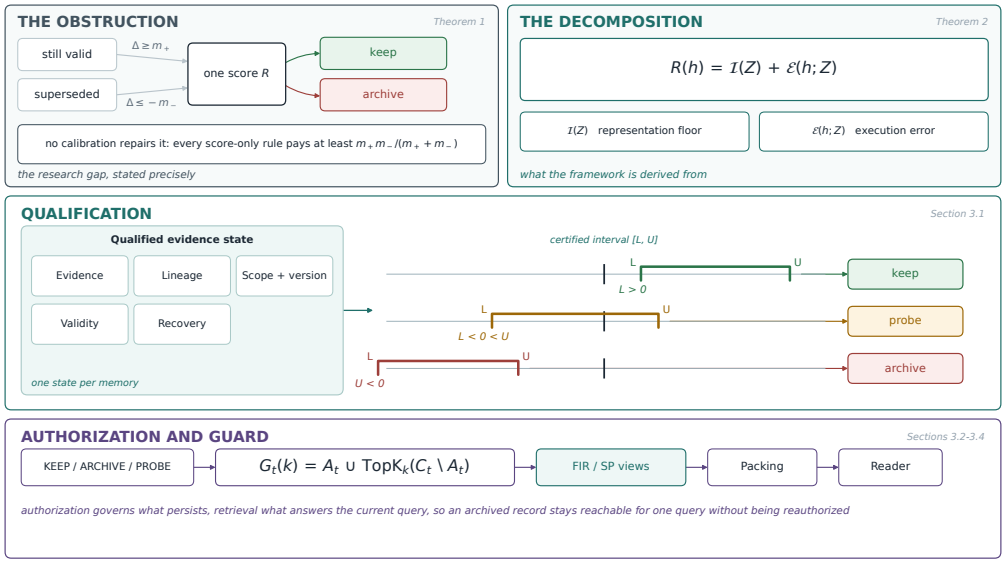


Figure 1: The two results and the controller they derive. A single score returns the same value for two records whose contrasts have opposite signs, and no calibration repairs it: every score-only rule pays at least $m_+m_-/(m_+ + m_-)$ (Theorem 1), which is the research gap stated precisely. The design comes from the decomposition, in which commitment regret separates into an irreducible representation floor $\mathcal{I}(Z)$ and a nonnegative execution error $\mathcal{E}(h; Z)$ (Theorem 2). A qualified evidence state supplies the floor, folded into a certified interval $[L, U]$ whose position relative to zero names the commitment.

on the first. FIR is a finite-impulse-response smoother over the observable order: a lexical score is replaced by the $0.25/0.50/0.25$ weighted average of it and its two chronological neighbours, endpoints padded by repetition and ties broken stably on memory ID, and the high-pass residual is charged a penalty of 0.08 so that an isolated lexical spike reads as noise rather than as evidence. SP, the structured response prior, keeps that smoothing and adds a candidate-proposal rule on top of it. From each high-scoring seed whose text is an interrogative anchor sharing at least two content tokens with the query, it takes the next turn in the same session and ranks the resulting candidates by $s(m) = \delta(m) + 0.10\, r(m)$, where $r(m)$ is the candidate's clipped lexical score and $\delta(m) = +0.35$ unless the candidate is itself a question, in which case $\delta(m) = -0.35$. On conversational tasks the smoothed scores are packed under item and reader-token budgets; on LifecycleBench the LLM commits from decision-time sessions through setting-specific instantiations of the same two views: FIR is the same temporal filter, while SP scores from the support of visible user–assistant edges instead of the proposal rule above (Appendix E), with the decision memory kept visible in every variant. Both use guard budget 16 (Appendix E). Because the view precedes commitment, changing it can change authorization itself.

4 Experiments

4.1 Contracts, baselines, and complete coverage

We use LoCoMo (Maharana et al., 2024) (1,986 questions over 10 conversations), LongMemEval-S (Wu et al., 2024) (500 questions), and LifecycleBench (225 seeded worlds, 25 per mechanism family), which gives every controller the same action space and regret reference by providing paired keep/archive futures under a common continuation. The binary instantiation evaluates persistent authorization on the authorization-shaped benchmark; on the retrieval-shaped pair the interface arbitrates access under a fixed reader budget, where its cost result is measured. SQCAD's interface also permits probe and defer, and budgeted probes occur in the fixed continuation, but model-selected probe/defer branches are outside this contract. Only the evaluator reads hidden futures; the harness records `gold_sent_to_model=false`.

We evaluate five backbones: Qwen3-8B, DeepSeek-Flash, GLM-5.3-Flash, Claude Sonnet 5, and GPT-5.6-terra, against ActMem, Memory Worth, Trivium, Oblivion (the last run as its official decay core), and two SQCAD guard configurations. The four displayed comparators implement published mechanisms under a shared input, action, and reader contract. Appendix D records each as an adapter rather than a native-system reproduction and takes no part in any ranking; Oblivion runs from its official published code, and other systems enter the score-surface audit (Appendix C).

On public tasks, contexts, chronological masking, and accounting are frozen before generation, so answer F1 varies with the backbone while hit, evidence recall, reader tokens, and storage do not: Tables 1 and 2 report the F1 block per backbone and those quantities once as shared columns. LoCoMo uses its official token-F1 code; LongMemEval-S uses token F1 under our reader contract, not its native judge accuracy. The LifecycleBench columns and the cost contract behind their utility scale are defined in Appendix E; because that scale is set by a successful future task rather than normalized to $[0, 1]$, a regret gap of 0.8 spans more than the whole neutral band and is a decision difference rather than a rounding artifact. Main-table comparisons are descriptive (Appendix F).

4.2 Cross-backbone quality and lifecycle control

On LoCoMo (Table 1), SQCAD (SP guard) attains the highest frozen F1 on every backbone, 0.2766–0.3182, at hit/recall 0.6773/0.6210, while authorizing roughly a fortieth of ActMem's 13,343.8 tokens at the higher F1. The untransformed control falls below ActMem on every backbone (Table 5, Appendix B), which locates the gain in the recorded view and budget combination rather than in the authorization machinery alone. The decay-only baseline is the instructive cost case: its official decay core authorizes 822.8 tokens, 2.5× either guard, to reach 0.0440–0.0749, below ActMem on every backbone and above both the Memory Worth and Trivium adapters. These LoCoMo orderings are descriptive; the resolved block is LongMemEval-S.

Table 1: LoCoMo: answer F1 per backbone, with the exposure and cost columns fixed before generation, identical across backbones. Bold marks the better of the two SQCAD guards within each backbone and the best SQCAD operating point per shared cost axis; neither denotes an overall win. Oblivion is the official decay core under the shared contract (Appendix D).

	Answer F1 ↑					Shared exposure and cost			
Method	Qwen	DeepSeek	GLM	Claude	Terra	Hit ↑	Recall ↑	Ctx. ↓	Auth. ↓
ActMem (Zhang et al., 2026b)	0.2484	0.2826	0.2468	0.2675	0.2822	0.5825	0.5265	290.6	13,343.8
Memory Worth (Simsek, 2026)	0.0404	0.0451	0.0346	0.0256	0.0276	0.0111	0.0090	246.1	246.1
Trivium (Chang, 2026)	0.0362	0.0432	0.0299	0.0227	0.0258	0.0174	0.0143	196.8	196.8
Oblivion (Rana et al., 2026)	0.0638	0.0749	0.0440	0.0516	0.0624	0.0717	0.0525	389.7	822.8
SQCAD (FIR guard)	0.2626	0.2995	0.2643	0.2774	0.2842	0.6358	0.5831	**213.8**	330.3
SQCAD (SP guard)	**0.2768**	**0.3182**	**0.2766**	**0.2941**	**0.2981**	**0.6773**	**0.6210**	214.7	**323.9**

Physical stored bytes are not reduced by archiving: 13,343.8 on LoCoMo.

LongMemEval-S is where the frozen comparisons resolve (Table 2): 34 of 40 tests are significant under Holm correction applied within dataset, of which 32 favour SQCAD and 2 favour ActMem, under the conditional assumptions in Appendix F. The guards authorize roughly a 46th of the 74,092.1 tokens ActMem authorizes, about 2% of physical storage, while producing F1 point estimates within 0.038 of the strongest adapter on every backbone and ahead of it on three. This is not an equivalence claim: the difference is paid in reader context, not correctness. ActMem leads on Qwen and Claude, FIR on GLM and Terra, SP on DeepSeek; the guards are never strictly ordered. The decay-only baseline costs 5,894.2 tokens, 3.7× the guards, to land at 0.0430–0.1346, and its recall of 0.0439 against a hit rate of 0.3854 is the sharper tell.

In lifecycle control (Table 3), SQCAD (FIR guard) attains the lowest frozen regret on four of five backbones, while the decay-only baseline authorizes 7.4–8.3 tokens per world and never leads its block, the weakest adapter on three of five. The informative case is a reversal inside one method: SQCAD (SP guard) leads LoCoMo on Qwen yet reaches a value of -2.28 (regret 4.31) here, below both the untransformed control and the named adapters, with no change to the controller. A view that improves the reader can damage the decision. That is the separation the decomposition formalizes, visible within one method rather than only across systems. The two sides differ in size: SP's LoCoMo

Table 2: LongMemEval-S: answer F1 per backbone, with the exposure and cost columns fixed before generation, identical across backbones. Bold marks the better of the two SQCAD guards within each backbone and the best SQCAD operating point per shared cost axis; neither denotes an overall win.

	Answer F1 ↑					Shared exposure and cost			
Method	Qwen	DeepSeek	GLM	Claude	Terra	Hit ↑	Recall ↑	Ctx. ↓	Auth. ↓
ActMem (Zhang et al., 2026b)	0.4184	0.4505	0.2951	0.4690	0.3139	0.9729	0.3461	746.2	74,092.1
Memory Worth (Simsek, 2026)	0.1127	0.0705	0.0787	0.0672	0.0793	0.0708	0.0350	1,942.5	1,942.5
Trivium (Chang, 2026)	0.1029	0.0592	0.0708	0.0600	0.0779	0.0479	0.0276	1,753.4	1,753.4
Oblivion (Rana et al., 2026)	0.1346	0.1075	0.0430	0.0606	0.1058	0.3854	0.0439	3,439.5	5,894.2
SQCAD (FIR guard)	0.3754	0.4343	**0.2968**	0.4542	**0.4294**	0.9750	0.3315	**1,644.8**	**1,602.0**
SQCAD (SP guard)	**0.3805**	**0.4557**	0.2702	**0.4580**	0.4245	**0.9771**	**0.3422**	1,713.6	1,604.5

Physical stored bytes are not reduced by archiving: 74,092.1 on LongMemEval-S.

advantage over FIR is 0.014 F1 points, while here the same change moves the guard from its block's best value to its worst, a 2.59-unit swing. These 40 comparisons are descriptive.

Table 3: LifecycleBench: lifecycle outcomes for five backbones, four lifecycle-compatible baseline adapters, and two SQCAD guard configurations. Bold marks the best SQCAD result per metric within each model block; it does not denote an overall win. All 225 worlds per backbone returned a parseable action.

Model	Method	Value ↑	Regret ↓	Missed ↓	Recover ↑
Qwen3 8B	ActMem (Zhang et al., 2026b)	-0.5305	2.5575	0.0400	0.6178
	Memory Worth (Simsek, 2026)	-0.7968	2.8238	0.1689	0.5733
	Trivium (Chang, 2026)	-0.7814	2.8085	0.1867	0.5822
	Oblivion (Rana et al., 2026)	-0.7089	2.7359	0.1422	0.5911
	SQCAD (FIR guard)	**0.3082**	**1.7188**	**0.0089**	**0.6533**
	SQCAD (SP guard)	-2.2843	4.3113	0.0311	**0.6533**
DeepSeek Flash	ActMem (Zhang et al., 2026b)	-0.7915	2.8185	0.0489	0.6178
	Memory Worth (Simsek, 2026)	-0.6838	2.7108	0.0756	0.5822
	Trivium (Chang, 2026)	-0.1077	2.1348	0.0533	0.6044
	Oblivion (Rana et al., 2026)	-0.5305	2.5575	0.0400	0.6178
	SQCAD (FIR guard)	**0.9347**	**1.0923**	**0.0089**	**0.6489**
	SQCAD (SP guard)	0.1985	1.8285	0.0844	0.5867
GLM 5.3-Flash	ActMem (Zhang et al., 2026b)	-0.3693	2.3963	0.0400	0.6178
	Memory Worth (Simsek, 2026)	-0.2286	2.2557	0.0356	0.6311
	Trivium (Chang, 2026)	0.5767	1.4503	0.0133	0.6444
	Oblivion (Rana et al., 2026)	-0.5880	2.6151	0.0533	0.6044
	SQCAD (FIR guard)	**1.1261**	**0.9009**	**0.0000**	**0.6578**
	SQCAD (SP guard)	1.0912	0.9358	0.0089	0.6489
Claude Sonnet 5	ActMem (Zhang et al., 2026b)	0.0574	1.9697	0.0622	0.6044
	Memory Worth (Simsek, 2026)	0.1534	1.8737	0.0400	0.6222
	Trivium (Chang, 2026)	0.8360	1.1910	0.0311	0.6267
	Oblivion (Rana et al., 2026)	-0.5365	2.5635	0.0400	0.6311
	SQCAD (FIR guard)	**1.0441**	**0.9829**	**0.0222**	**0.6400**
	SQCAD (SP guard)	0.8861	1.1410	0.0311	0.6311
GPT 5.6-terra	ActMem (Zhang et al., 2026b)	-0.3965	2.4235	0.0444	0.6400
	Memory Worth (Simsek, 2026)	0.7292	1.2978	0.0133	0.6622
	Trivium (Chang, 2026)	1.0864	0.9406	0.0089	0.6578
	Oblivion (Rana et al., 2026)	-0.7165	2.7435	0.1022	0.6000
	SQCAD (FIR guard)	**0.6764**	**1.3506**	**0.0444**	**0.6178**
	SQCAD (SP guard)	0.6087	1.4184	0.0844	0.5911

4.3 DETERMINISTIC CONTROLS AND MECHANISM TESTS

Table 6 (Appendix B) resolves deterministic policy differences by mechanism family, and the appendix retains the remaining policies and continuation switches alongside it. The three SQCAD policies, qualification, qualification with lineage, and probe-willing, change the initial commitment rule under the same follow-on environment.

The family breakdown is sharp and interpretable. Lineage eliminates version-update regret entirely, from 74.4 to 0; because the families are equally sized, version update contributes 89.8% of the qualification policy's total regret (macro mean 9.22, median 0, mean excluding version update 1.06; Appendix F.2 varies that weighting), so lineage removes a single well-identified failure mode rather than producing a diffuse improvement. Probe-willing reduces hitchhiker and neutral regret at the cost of a small harmful-stale loss. The generator exposes an explicit same-scope version conflict and the lineage rule detects exactly that condition.

The continuation ablation fixes the initial lineage-aware action and changes only the subsequent rollout, so no-lineage retains the initial lineage-informed choice. Censoring dominates this layer: disabling it raises regret from 0.94 to 11.0, a twelvefold increase, while no-lineage, no-probe, no-qualification, and no-restore leave the aggregate unchanged (Table 8). The load-bearing follow-on component in these worlds is therefore the one that tracks which records were unavailable when a decision was made, not the ones that re-rank access.

Why the deterministic rules lead. The rules resolve worlds the family label cannot. Collapsing any controller to one action per mechanism family leaves the LLM rows unchanged, at 2.56 to 2.66 regret, but raises probe-willing from 0.0634 to 1.7196. Those are exactly the two values the information floors of Table 4 take, and the agreement is an identity rather than a coincidence: probe-willing is measurable with respect to the lineage summary and reaches its floor with execution error of zero, so the released rule is Bayes-optimal for the released representation, and flattening it discards precisely what that representation adds.

We also audit seven scalar surfaces, six proxies inspired by observable mechanisms in named methods plus an internal local-effect control, on the same worlds. Six contain opposite-sign fibers at eight-decimal score precision, where the Trivium-inspired proxy has none, which is what makes the audit informative rather than automatic; several proxies share the same version-update witness, so the row count is not itself independent evidence (Appendix C).

4.4 OFFLINE VALIDATION OF REPRESENTATION AND EXECUTION

We replay the released public-only rules on the same paired worlds under three nested summaries: query-target token overlap; that overlap with scope compatibility and visible correction; and those fields plus a live lexical lineage counterpart. All exclude world IDs, family labels, and hidden outcomes. Table 4 verifies Theorem 2 for rules restricted to each summary. Scope qualification lowers the rule's execution error from 8.05 to 7.50 without moving the information floor at 1.72, a pure execution improvement. Adding lineage lowers the floor to 0.0634 and leaves 0.258 execution regret: an information improvement. Five-fold held-out lookup attains the same floors on unseen worlds, but refitting with a whole mechanism family held out removes the ordering: all three summaries reach 1.9851 (Table 14). The floors are in-family quantities.

Table 4: Offline decomposition on 225 paired worlds. Each released public rule observes only its listed summary. Information is the empirical minimum regret for that summary, execution is excess rule regret, and bold marks the lowest loss. These are lexical public replay controls, distinct from the simulator policies and LLM adapters.

Public summary / rule	Cells	Information	Execution	Regret
Overlap / relevance	5	1.7196	8.0546	9.7743
+ scope, correction / qualify	6	1.7196	7.4966	9.2162
+ lineage / qualify-lineage	8	**0.0634**	**0.2579**	**0.3213**

5 DISCUSSION

No amount of calibration or threshold tuning repairs an opposite-sign collision, because the distinction was never in the input; the audit finds such fibers on six of its seven surfaces, and the survivor depends on future demand rather than present relevance. The bound constrains the interface without prescribing the representation: it says which distinctions a scalar cannot carry, not which additional fields restore them. It also explains why the scalar interface persists, since a score is the cheapest summary a store can offer a model and refining its calibration keeps raising the reported number while the commitment floor stays where it was. Where the bound stops, the decomposition takes over as a design instrument. Qualification and lineage both change behaviour but move different quantities, and the offline replay separates them: scope qualification improves how the rule uses distinctions it already has and leaves the information floor at 1.72, while lineage lowers that floor to 0.0634 on the version-update family it targets. Better rules over the existing fields are the cheaper repair, and the ceiling they cannot pass is set by the representation; retaining the invalidation evidence that relevance discards is what raises that ceiling. Execution error moving while the floor holds is therefore a gain the representation did not have to earn.

LoCoMo and LongMemEval-S are retrieval-shaped, with a one-query horizon, while LifecycleBench is authorization-shaped. Reporting all three is what makes the ordering changes above readable, because the two properties come apart in both directions: the configuration that reads best need not decide best, and the reversal appears inside a single method rather than only across systems. SQCAD's SP guard leads LoCoMo on Qwen and, with no change to the controller, reaches the worst value in its LifecycleBench block (Section 4.2), a 2.59-unit swing set against a 0.014 F1 advantage on the other side. Retrieval quality and lifecycle control are therefore separate properties of the same representation, which is what the decomposition formalizes and what the three datasets together let a reader check.

Three practices follow. Separate the reading view from the decision state: one representation is consumed by two processes, and optimizing for the reader can degrade the controller. Report answer quality, reader context, and authorized storage together, because a recoverable archive trades between them. Retain invalidation and exception evidence even when it contributes little to retrieval, since that is what the information floor is made of. These three practices are the transferable part of the result; they can be adopted one at a time, and SQCAD is one way to operationalize them. The supplementary release carries an authored walkthrough of four such decisions, where a reader can step through the qualification state and watch the two continuations diverge. It is aimed at a reader meeting the mechanism for the first time: four constructed scenarios show what each field of the certificate buys, and where a relevance score would have committed the same way to opposite-sign states. What would extend the result is trajectories whose branches are paired without being constructed, which is also its clearest limit; the model comparisons rest on single frozen runs, so we read small margins as descriptive rather than decisive.

6 CONCLUSION

Persistent authorization asks whether a memory may keep influencing decisions that have not arrived, which is a sequential decision and not the one agent memory is usually evaluated on. The obstruction is structural rather than a calibration failure, and it belongs to the interface rather than to memory stores: any system whose persistent binary decision rests on a scalar summary inherits the bound as soon as that scalar collides opposite-sign states, and can separate its representation loss from its controller error once both branches are observed. Agent memory is the setting we instantiate, because there the two branches are the same record kept or archived, but the same structure governs retention, eviction, and admission wherever one score decides. The contribution is a criterion rather than a system: the bound says when no scalar can serve, the decomposition says whether a richer representation or a better rule is the repair available, and the paired instrument makes that claim checkable. Two repairs that both raise downstream accuracy move different quantities, so accuracy alone cannot tell them apart, and the terms that separate them are independent of this simulator, these backbones, and the guard budget. What changes is what the signal must preserve: not how well it predicts present usefulness, but which future authorization distinctions it keeps available. Persistent authorization is a property of the memory decision, not a by-product of retrieval quality.

AI USE STATEMENT

Generative AI tools were used for research ideation and execution: in that role they assisted the study's planning, the design of its experimental methodology, and its theoretical development, including the formulation of the paper's mathematical claims and the development of their proofs, together with the implementation of the controller, the LifecycleBench generator, the baseline adapters, and the analysis and table-generation code. As recommended disclosure, they assisted retrieval and discovery, literature search and the identification of related work, the drafting of sections of the paper, editing for readability, and the preparation of figures. Assistance also covered code and experiment-script review, table reconstruction, offline analysis of frozen outputs, and critical review. We have not used generative AI tools to generate synthetic datasets: LifecycleBench is produced by deterministic program code under fixed seeds, and no model generated any dataset content. We have not used generative AI tools for translation. The remaining tasks carrying a required disclosure are not applicable to this work. The authors reviewed all AI-assisted work and take responsibility for the final content of this work, including its text, claims, and artifacts. Numerical evidence remains tied to recorded artifacts: each table is generated from a frozen run whose SHA-256 hash the provenance ledger records, and the authors verified the code, the proofs, the results, the references, and the submission materials against those artifacts.

ETHICS STATEMENT

This work uses public benchmark data and program-generated counterfactual worlds, with no new human-subject data collection. Authorization here is an experimental access-control abstraction. Keeping a physical archive is not a privacy-erasure mechanism, and the experiments do not establish safety under adversarial inputs or in deployed agents. Public and hidden benchmark layers must remain separated during controller evaluation.

REPRODUCIBILITY STATEMENT

Appendices A–E specify assumptions, coverage, transformations, scoring, and provenance. Tables come from frozen artifacts with SHA-256 checks: 105 method–model–dataset cells over 121 hashed source artifacts, 75 lifecycle-compatible, including the separately run Oblivion baseline under the same contracts. Source, tests, table-provenance records, the LifecycleBench release, and an interactive case study are provided as supplementary material (Appendix E inventories the release); LifecycleBench keeps its public and hidden layers separate and source hashes accompany the artifact. Public URLs are withheld from this version and restored at camera-ready.

REFERENCES

David Abel, Nate Umbanhowar, Khimya Khetarpal, Dilip Arumugam, Doina Precup, and Michael Littman. Value preserving state-action abstractions. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 1639–1650. PMLR, 2020. URL https://proceedings.mlr.press/v108/abel20a.html.

Zouying Cao, Jiaji Deng, Li Yu, Weikang Zhou, Zhaoyang Liu, Bolin Ding, and Hai Zhao. Remember me, refine me: A dynamic procedural memory framework for experience-driven agent evolution. In *Findings of the Association for Computational Linguistics: ACL 2026*, pp. 16803–16822, 2026. doi: 10.18653/v1/2026.findings-acl.829. URL https://aclanthology.org/2026.findings-acl.829/.

Edward Y. Chang. Trivium: Temporal regret as a first-class objective for causal-memory controllers. *arXiv preprint arXiv:2606.04421*, 2026. URL https://arxiv.org/abs/2606.04421.

Prateek Chhikara, Dev Khant, Saket Aryan, Taranjeet Singh, and Deshraj Yadav. Mem0: Building production-ready AI agents with scalable long-term memory. In *Frontiers in Artificial Intelligence and Applications*. IOS Press, 2025. doi: 10.3233/FAIA251160.

Yiming Du, Wenyu Huang, Danna Zheng, Zhaowei Wang, Sebastien Montella, Mirella Lapata, Kam-Fai Wong, and Jeff Z. Pan. Rethinking memory in LLM-based agents: Representations, operations, and emerging topics. *arXiv preprint arXiv:2505.00675*, 2025. URL https://arxiv.org/abs/2505.00675.

Bernal Gutiérrez, Yiheng Shu, Yu Gu, Michihiro Yasunaga, and Yu Su. HippoRAG: Neurobiologically inspired long-term memory for large language models. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-1902.

Mengkang Hu, Tianxing Chen, Qiguang Chen, Yao Mu, Wenqi Shao, and Ping Luo. HiAgent: Hierarchical working memory management for solving long-horizon agent tasks with large language model. *arXiv preprint arXiv:2408.09559*, 2024. URL https://arxiv.org/abs/2408.09559.

Yuanzhe Hu, Yu Wang, and Julian McAuley. Evaluating memory in LLM agents via incremental multi-turn interactions. *arXiv preprint arXiv:2507.05257*, 2025. URL https://arxiv.org/abs/2507.05257.

Suozhao Ji, Baodong Wu, Zehao Wang, Lei Xia, Qingping Li, Ruisong Wang, Wenbo Ding, Zhenhua Zhu, Boxun Li, Guohao Dai, and Yu Wang. Infini Memory: Maintainable topic documents for long-term LLM agent memory. *arXiv preprint arXiv:2606.10677*, 2026. URL https://arxiv.org/abs/2606.10677.

Jiazheng Kang, Mingming Ji, Zhe Zhao, and Ting Bai. Memory OS of AI agent. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 25961–25970, 2025. doi: 10.18653/v1/2025.emnlp-main.1318. URL https://aclanthology.org/2025.emnlp-main.1318/.

Cai Ke, Yiming Du, Bin Liang, Yifan Xiang, Lin Gui, Zhongyang Li, Baojun Wang, Yue Yu, Hui Wang, Kam-Fai Wong, and Ruifeng Xu. Flexibly utilize memory for long-term conversation via a fragment-then-compose framework. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 21119–21136, 2025. doi: 10.18653/v1/2025.emnlp-main.1069. URL https://aclanthology.org/2025.emnlp-main.1069/.

Sangyeop Kim, Yohan Lee, Sanghwa Kim, Hyunjong Kim, and Sungzoon Cho. Pre-storage reasoning for episodic memory: Shifting inference burden to memory for personalized dialogue. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pp. 22096–22113, 2025. doi: 10.18653/v1/2025.findings-emnlp.1204. URL https://aclanthology.org/2025.findings-emnlp.1204/.

Letta Team. Letta: Stateful agents with long-term memory, 2025. URL https://www.letta.com/.

Lihong Li, Thomas J. Walsh, and Michael L. Littman. Towards a unified theory of state abstraction for MDPs. In *International Symposium on Artificial Intelligence and Mathematics*, 2006.

Jiaqi Liu, Yaofeng Su, Peng Xia, Siwei Han, Zeyu Zheng, Cihang Xie, Mingyu Ding, and Huaxiu Yao. SimpleMem: Efficient lifelong memory for LLM agents. *arXiv preprint arXiv:2601.02553*, 2026. URL https://arxiv.org/abs/2601.02553.

Adyasha Maharana, Dong-Ho Lee, Sergey Tulyakov, Mohit Bansal, Francesco Barbieri, and Yuwei Fang. Evaluating very long-term conversational memory of LLM agents. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 13851–13870. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.acl-long.747. URL https://aclanthology.org/2024.acl-long.747/.

Charles Packer, Sarah Wooders, Kevin Lin, Vivian Fang, Shishir G. Patil, Ion Stoica, and Joseph E. Gonzalez. MemGPT: Towards LLMs as operating systems. *arXiv preprint arXiv:2310.08560*, 2023. URL https://arxiv.org/abs/2310.08560.

Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, pp. 1–22, 2023. doi: 10.1145/3586183.3606763.

Yijiashun Qi, Xiang Xu, and Yuxuan Li. When not to write memory: Governing false promotion from correlated agent traces. *arXiv preprint arXiv:2607.02579*, 2026. URL https://arxiv.org/abs/2607.02579.

Ashish Rana, Chia-Chien Hung, Qumeng Sun, Julian Martin Kunkel, and Carolin Lawrence. Oblivion: Self-adaptive agentic memory control through decay-driven activation. *arXiv preprint arXiv:2604.00131*, 2026. URL https://arxiv.org/abs/2604.00131.

Preston Rasmussen, Pavlo Paliychuk, Travis Beauvais, Jack Ryan, and Daniel Chalef. Zep: A temporal knowledge graph architecture for agent memory. *arXiv preprint arXiv:2501.13956*, 2025. URL https://arxiv.org/abs/2501.13956.

Baris Simsek. When to forget: A memory governance primitive. *arXiv preprint arXiv:2604.12007*, 2026. URL https://arxiv.org/abs/2604.12007.

Haoran Tan, Zeyu Zhang, Chen Ma, Xu Chen, Quanyu Dai, and Zhenhua Dong. MemBench: Towards more comprehensive evaluation on the memory of LLM-based agents. *arXiv preprint arXiv:2506.21605*, 2025a. URL https://arxiv.org/abs/2506.21605.

Zhen Tan, Jun Yan, I-Hung Hsu, Rujun Han, Zifeng Wang, Long Le, Yiwen Song, Yanfei Chen, Hamid Palangi, George Lee, Anand Rajan Iyer, Tianlong Chen, Huan Liu, Chen-Yu Lee, and Tomas Pfister. In prospect and retrospect: Reflective memory management for long-term personalized dialogue agents. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics*, pp. 8416–8439, 2025b. doi: 10.18653/v1/2025.acl-long.413. URL https://aclanthology.org/2025.acl-long.413/.

Sijia Wang, Dhanajit Brahma, and Ricardo Henao. SAGE: A novelty gate for efficient memory evolution in agentic LLMs. *arXiv preprint arXiv:2605.30711*, 2026. URL https://arxiv.org/abs/2605.30711.

Yu Wang and Xi Chen. MIRIX: Multi-agent memory system for LLM-based agents. *arXiv preprint arXiv:2507.07957*, 2025. URL https://arxiv.org/abs/2507.07957.

Lei Wei, Xiao Peng, Xu Dong, Niantao Xie, and Bin Wang. FadeMem: Biologically-inspired forgetting for efficient agent memory. *arXiv preprint arXiv:2601.18642*, 2026. URL https://arxiv.org/abs/2601.18642.

Di Wu, Hongwei Wang, Wenhao Yu, Yuwei Zhang, Kai-Wei Chang, and Dong Yu. LongMemEval: Benchmarking chat assistants on long-term interactive memory. *arXiv preprint arXiv:2410.10813*, 2024. URL https://arxiv.org/abs/2410.10813.

Wujiang Xu, Zujie Liang, Kai Mei, Hang Gao, Juntao Tan, and Yongfeng Zhang. A-MEM: Agentic memory for LLM agents. In *Advances in Neural Information Processing Systems*, volume 38, pp. 20004–20031, 2025. doi: 10.52202/085713-0593.

Tianyu Yang, Sudipta Paul, Vijay Srinivasan, Vivek Kulkarni, and Srinivas Chappidi. TrustMem: Learning trustworthy memory consolidation for LLM agents with long-term memory. *arXiv preprint arXiv:2606.25161*, 2026. URL https://arxiv.org/abs/2606.25161.

Boqin Yuan, Yue Su, and Kun Yao. Diagnosing retrieval vs. utilization bottlenecks in LLM agent memory. In *ICLR 2026 Workshop on Memory for LLM-Based Agentic Systems (MemAgents)*, 2026. URL https://arxiv.org/abs/2603.02473.

Sheng Zhang, Junyi Li, Yingyi Zhang, Pengyue Jia, Yichao Wang, Xiaowei Qian, Wenlin Zhang, Maolin Wang, Yong Liu, and Xiangyu Zhao. MemSearch-o1: Empowering large language models with reasoning-aligned memory growth in agentic search. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics*, pp. 925–943, 2026a. doi: 10.18653/v1/2026.acl-long.41. URL https://aclanthology.org/2026.acl-long.41/.

Xiaohui Zhang, Zequn Sun, Chengyuan Yang, Yaqin Jin, Yazhong Zhang, and Wei Hu. ActMem: Bridging the gap between memory retrieval and reasoning in LLM agents. *arXiv preprint arXiv:2603.00026*, 2026b. URL https://arxiv.org/abs/2603.00026.

Zeyu Zhang, Quanyu Dai, Xiaohe Bo, Chen Ma, Rui Li, Xu Chen, Jieming Zhu, Zhenhua Dong, and Ji-Rong Wen. A survey on the memory mechanism of large language model-based agents. *ACM Transactions on Information Systems*, 43:1–47, 2025. doi: 10.1145/3748302.

Wanjun Zhong, Lianghong Guo, Qiqi Gao, He Ye, and Yanlin Wang. MemoryBank: Enhancing large language models with long-term memory. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 19724–19731, 2024. doi: 10.1609/aaai.v38i17.29946.

Yijie Zhong, Yunfan Gao, and Haofen Wang. HingeMem: Boundary guided long-term memory with query adaptive retrieval for scalable dialogues. In *Proceedings of the ACM Web Conference 2026*, 2026. doi: 10.1145/3774904.3792089.

Zijian Zhou, Ao Qu, Zhaoxuan Wu, Sunghwan Kim, Alok Prakash, Daniela Rus, Jinhua Zhao, Bryan Kian Hsiang Low, and Paul Pu Liang. MEM1: Learning to synergize memory and reasoning for efficient long-horizon agents. *arXiv preprint arXiv:2506.15841*, 2025. URL https://arxiv.org/abs/2506.15841.

Mingxi Zou, Zhihan Guo, Langzhang Liang, Zhuo Wang, Qifan Wang, Qingsong Wen, Irwin King, Lizhen Qu, and Zenglin Xu. Remember the decision, not the description: A rate-distortion framework for agent memory. *arXiv preprint arXiv:2605.10870*, 2026. URL https://arxiv.org/abs/2605.10870.

A Assumptions and proofs

All state, observation, and qualification spaces are standard Borel, and the kernels and rewards are measurable. The action set is finite. The horizon is finite, the discount is in $(0, 1]$, and terminal values and rewards are bounded. Belief updates are well defined on the relevant support. A score-only rule cannot inspect the within-fiber state. For the certificate, the closure inequalities must hold for all stages and all actions used to compute the continuation, with radii fixed before the certification decision.

A.1 Score-fiber bound

At the common score, let p denote the rule's keep probability. In x_+ its expected binary commitment regret is at least $(1-p)m_+$; in x_- it is at least pm_-. Hence

$$\inf_{p\in[0,1]} \max\{(1-p)m_+, pm_-\} = \frac{m_+m_-}{m_+ + m_-}. \tag{5}$$

The balancing probability is $p = m_+/(m_+ + m_-)$. The bound is attained when the two contrasts equal their stated margins; it is therefore sharp for this two-state experiment. It constrains binary commitments. Allowing a third action that abstains changes the decision problem and requires pricing that action explicitly. The same argument applies to fixed-policy branch contrasts when the continuation and utility are held fixed.

A.2 Decomposition, refinement, and offline protocol

Theorem 2 requires integrable branch values, a fixed world distribution, and a controller whose action probability depends only on Z. Conditional on $Z = z$, write $p(z) = \Pr(h = \mathsf{keep} \mid Z = z)$ with independent controller randomness. Regret is $p(z)B(z) + (1-p(z))A(z)$, minimized by keep when $A(z) \geq B(z)$. Since $A + B = \mathbb{E}[|\Delta| \mid Z]$ and $A - B = \mu$, its minimum is $(\mathbb{E}[|\Delta| \mid Z] - |\mu|)/2$. Averaging proves both expressions for $\mathcal{I}(Z)$ and the nonnegative execution term. At $\mu = 0$, either action has the same expected regret.

If $Z = f(Z')$, the tower property gives $\mathbb{E}[\mu(Z') \mid Z] = \mu(Z)$. Conditional Jensen's inequality implies $\mathbb{E}|\mu(Z')| \geq \mathbb{E}|\mu(Z)|$, so $\mathcal{I}(Z') \leq \mathcal{I}(Z)$. The information term is zero exactly when positive and negative contrasts do not coexist with positive conditional mass in a cell, up to null sets. None of these statements implies that a learned controller improves when its input changes.

The offline audit is a post hoc diagnostic using the uniform empirical distribution on the 225 released worlds. It constructs the following nested summaries from the public layer before joining evaluator

values: (i) the number of lowercase alphanumeric query tokens shared with the target; (ii) that number plus target scope compatibility and a visible update/correction sharing at least two target tokens; (iii) those fields plus a live, introduced, same-scope counterpart sharing at least three target tokens and not resolved by a visible event. Thresholds and rules are taken from the released example controller. No world ID, family, variant, oracle label, or future value is a representation feature. Token overlap is used only as a public lexical control.

We replay relevance (keep at overlap at least two), qualify (keep when scope-compatible and not visibly corrected), and qualify-lineage (additionally require no live counterpart). The evaluator checks that each rule's action is constant within each representation cell before applying the theorem. Table 4 reports all three controls; the identity residual is below 10^{-12}. The public lexical lineage rule has regret 0.3213, whereas the simulator's recorded lineage rule has regret 0.9443. Their implementations differ, so these are separate controls, not replacement scores.

To separate in-sample information floors from rule fitting, we also estimate each cell's mean contrast using four folds and select a sign on the fifth. Five folds use SHA-256 ID ordering balanced within mechanism family. IDs and families determine the split only; each held-out decision uses its public summary and means learned on other worlds. Empty training cells would use the training-wide mean. Across all folds no held-out cell is unseen. Out-of-fold regrets are 1.7196, 1.7196, and 0.0634 for the three summaries, respectively. This is an evaluator-trained diagnostic on the existing generator, with all mechanism families represented in training; it is not an unseen-family evaluation or a model prompt supplied with test labels. It cannot establish population sufficiency from a finite-support information floor.

Frozen LLM actions are checked against their exact paired branch values separately. Their prompts contain more information than these coarse summaries, so their actions need not be measurable with respect to the summaries. We do not infer a coarse representation's execution term by subtracting its floor from an LLM's regret. Such a subtraction could be negative and would violate the theorem's premise.

The five-fold diagnostic above is balanced within mechanism family, so every held-out cell is seen in training and the out-of-fold regret reproduces the in-sample floor. Table 14 asks the stronger question by refitting with one whole mechanism family held out: a world whose cell does not appear in the other eight receives the training-wide mean contrast, which keeps the assigned action constant within the family and the rule measurable with respect to the summary. In sample the lineage summary reduces regret from 9.7743 to 0.3213; under family shift all three summaries reach 1.9851 and the in-sample ordering disappears. The per-family detail locates the reason: the lineage cells that carry the version-update contrast are specific to that family, so a rule fitted on the other eight assigns its fallback action there and returns that family's undivided floor of 14.9065. The offline summaries are accordingly family-confounded: their cells do not carry a sign that is invariant across mechanism families, which bounds what the in-sample floors establish, and which is the diagnostic a natural-trajectory corpus would have to displace.

A.3 Uniform closure as a stronger sufficient condition

The following result is not used by the paired-world controllers. It states a sufficient condition for a compressed state to preserve action values. Let $Z_r = q_r(X_r) = (E_r, L_r, G_r, V_r, C_r)$ record evidence/uncertainty, lineage, scope/version, validity, and censoring/recovery. Assume quotient rewards $\bar{\ell}_r$, kernels $\bar{T}_r^a$, and terminal value $\bar{g}_{H+1}$ satisfy, for every relevant state and every action in the Bellman recursion,

$$\begin{aligned} |\ell_r(x,a) - \bar{\ell}_r(q_r(x),a)| &\le \varepsilon_r, \qquad \|(q_{r+1})_{\#}T_r^a - \bar{T}_r^a\|_{\mathrm{TV}} \le \eta_r, \\ |g_{H+1}(x) - \bar{g}_{H+1}(q_{H+1}(x))| &\le \varepsilon_{H+1}. \end{aligned} \tag{6}$$

The total-variation convention is $\sup_A |P(A) - Q(A)|$. With $|\bar{V}_r| \le M_r$, define

$$b_{H+1} = \varepsilon_{H+1}, \qquad b_r = \varepsilon_r + \gamma(b_{r+1} + 2M_{r+1}\eta_r). \tag{7}$$

Theorem 3 (Commitment error under uniform closure)**.** *If Eq.* (6) *holds from t through H, then $|Q_r(x,a) - \bar{Q}_r(q_r(x),a)| \le b_r$ and $|\Delta_r(x) - \bar{\Delta}_r(q_r(x))| \le 2b_r$. Thus $\bar{\Delta}_r > 2b_r$ certifies keep and $\bar{\Delta}_r < -2b_r$ certifies archive. Inside the band, neither binary sign is certified.*

Proof and horizon dependence. Assume the value error at stage $r+1$ is at most b_{r+1}. Insert the full transition applied to $\bar{V}_{r+1} \circ q_{r+1}$ between the two Bellman integrals. One difference is bounded by b_{r+1}; the other is at most $2M_{r+1}\eta_r$ because $|\int f\, d(P-Q)| \le 2\|f\|_\infty \|P-Q\|_{\mathrm{TV}}$. Adding reward error gives Eq. (7). Taking the maximum over the common action set preserves this bound; induction from the terminal condition and subtraction of the two action values complete the proof. Unrolling the recurrence makes the temporal requirement explicit:

$$b_t = \sum_{r=t}^{H} \gamma^{r-t}\varepsilon_r + 2\sum_{r=t}^{H} \gamma^{r-t+1} M_{r+1}\eta_r + \gamma^{H-t+1}\varepsilon_{H+1}. \tag{8}$$

Errors in future observation dynamics can therefore affect a current authorization even when its immediate reward is represented exactly. If quotient rewards are bounded in magnitude by L, terminal values are zero, $\gamma = 1$, and errors are uniformly ε, η, then for $n = H - t + 1$ stages, $b_t \le n\varepsilon + L\eta n(n-1)$. The transition contribution grows quadratically in the worst case because a distorted successor distribution changes multiple future rewards.

From representation error to commitment risk. Choosing the larger quotient commitment value has binary regret at most $2b_t$: if its sign is wrong, the true contrast can have magnitude at most $2b_t$. Outside the certificate band the binary regret is zero. This separates a representation requirement from a decision rule: a compact state is adequate when its distortion is small relative to the action margin, not merely when it predicts average relevance well. Comparing commitment with probe/defer still requires their values; a valid binary sign alone does not choose among all four actions. The following induction gives the full proof.

Full induction. The terminal inequality gives $|V_{H+1}(x) - \bar{V}_{H+1}(q_{H+1}(x))| \le b_{H+1}$. Suppose the value bound holds at $r+1$. Write $P_x = (q_{r+1})_{\#} T_r^a(\cdot \mid x)$. For every action,

$$|Q_r(x,a) - \bar{Q}_r(q_r(x),a)| \le \varepsilon_r + \gamma \left| \int V_{r+1}\, dT_r^a - \int \bar{V}_{r+1}\, d\bar{T}_r^a \right| \tag{9}$$

$$\le \varepsilon_r + \gamma \left(b_{r+1} + \left| \int \bar{V}_{r+1}\, dP_x - \int \bar{V}_{r+1}\, d\bar{T}_r^a \right| \right) \tag{10}$$

$$\le \varepsilon_r + \gamma(b_{r+1} + 2M_{r+1}\eta_r) = b_r. \tag{11}$$

Taking maxima over the same finite action set preserves the bound for V_r and $\bar{V}_r$, completing the induction. Applying the bound to keep and archive and using the triangle inequality yields $|\Delta_r - \bar{\Delta}_r| \le 2b_r$. A quotient contrast larger than $2b_r$ is positive in the full state; a quotient contrast smaller than $-2b_r$ is negative. No binary sign follows within the band, and neither sign comparison establishes that commitment dominates probing. A finite-data estimate of ε_r or η_r would also need simultaneous coverage over the stated domain; an average empirical error does not supply the theorem's uniform bound.

B Deterministic controls and continuation ablations

B.1 Guard configuration control

Table 5 retains all results previously labeled SQCAD (base), now identified by the intervention: SQCAD without FIR/SP view transformations. This control retains the authorization and access machinery. Comparing it with the two main configurations assesses the recorded combinations of views and budgets, not an isolated view-removal effect or removal of the entire guard. An authorization-removal experiment with the view and budgets held fixed would be needed to isolate authorization's effect.

Table 5: SQCAD without FIR/SP view transformations: the internal control omitted from the main comparisons. The authorization and access machinery remains in place; this is not a guard-removal experiment. All five backbones are retained, with the same metrics as the corresponding main table.

LoCoMo Model	F1 ↑	Hit ↑	Recall ↑	Context ↓	Auth. ↓
Qwen3-8B	0.1276	0.2122	0.1899	348.2	329.9
DeepSeek-Flash	0.1391	0.2122	0.1899	348.2	329.9
GLM-5.3-Flash	0.1196	0.2122	0.1899	348.2	329.9
Claude Sonnet 5	0.1173	0.2122	0.1899	348.2	329.9
GPT-5.6-terra	0.1289	0.2122	0.1899	348.2	329.9

LongMemEval-S Model	F1 ↑	Hit ↑	Recall ↑	Context ↓	Auth. ↓
Qwen3-8B	0.2528	0.7542	0.1149	2,721.7	1,631.2
DeepSeek-Flash	0.2517	0.7542	0.1149	2,721.7	1,631.2
GLM-5.3-Flash	0.1956	0.7542	0.1149	2,721.7	1,631.2
Claude Sonnet 5	0.2634	0.7542	0.1149	2,721.7	1,631.2
GPT-5.6-terra	0.2857	0.7542	0.1149	2,721.7	1,631.2

LifecycleBench Model	Value ↑	Regret ↓	Missed ↓	Recover ↑
Qwen3-8B	-0.5746	2.6016	0.0978	0.5689
DeepSeek-Flash	0.4138	1.6132	0.1067	0.5778
GLM-5.3-Flash	0.5387	1.4883	0.0267	0.6533
Claude Sonnet 5	0.9751	1.0519	0.1111	0.5067
GPT-5.6-terra	0.4022	1.6248	0.1733	0.5333

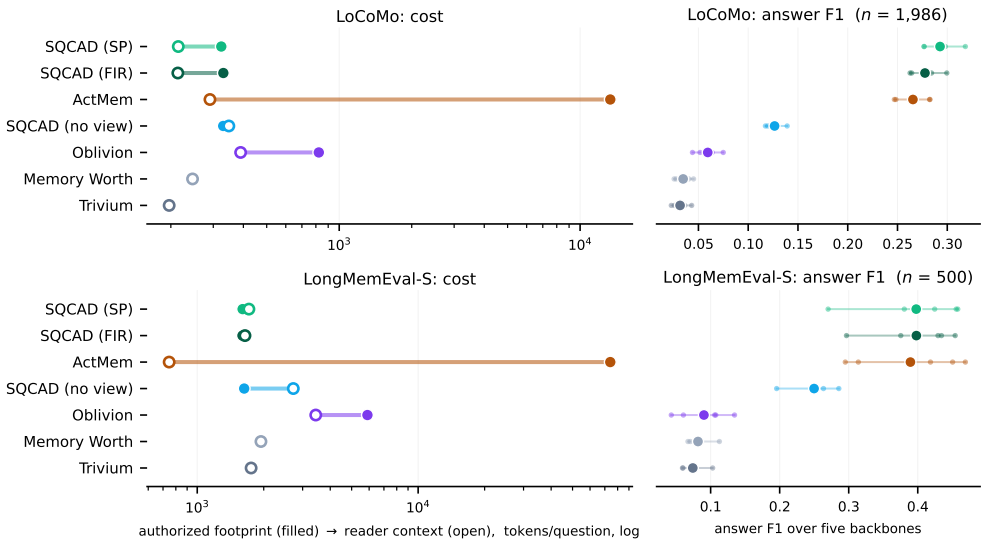


Figure 2: Answer quality against both cost axes, from the same frozen artifacts as Tables 1, 2 and 5. Left: each method is a segment from the authorized decision-memory footprint it grants (filled marker) to the reader context it exposes (open marker); right: its answer F1, with the five backbones as small points and their mean filled. A segment's two ends are the two columns the main tables report side by side, so archiving is read as the left end moving far left while the right end stays close or moves out. The two SQCAD guards lie within 1% of each other on both cost axes and overlap in the left panels.

Table 6: Decision-policy comparison by mechanism family (25 worlds each). Mean regret is lower-is-better. Aggregate rows expose concentration in the version-update stress test; they are descriptive summaries of the same 225 worlds. Bold marks strict reductions between successive SQCAD policies; all-zero families remain visible. Oracle is the neutral-label policy reference, not the exact branch optimum.

Family	SQCAD cert	+ lineage	Probe-willing	Oracle
Version update	74.447	**0.000**	0.000	0.000
Hitchhiker	7.541	7.541	**0.000**	0.000
Harmful stale	0.000	0.000	0.053	0.000
Neutral	0.441	0.441	**0.000**	0.356
Scope mismatch	0.517	0.517	0.517	0.000
Rare bridge	0.000	0.000	0.000	0.000
Self obscuring	0.000	0.000	0.000	0.000
Stable negative	0.000	0.000	0.000	0.000
Stable positive	0.000	0.000	0.000	0.000
Macro mean	9.216	0.944	0.063	0.040
Median	0.000	0.000	0.000	0.000
Mean excluding version update	1.062	1.062	0.071	0.045

B.2 Initial policy and continuation controls

Table 6 compares initial decision policies under a common continuation and is placed here for space; the main text reports its numbers and interpretation. Table 7 includes all 12 public policies and the oracle policy reference. Table 8 instead fixes the initial SQCAD + lineage decision and changes subsequent rollout components. False/missed decisions therefore remain fixed by construction. Null aggregate changes in value and regret are observed outcomes, not evidence that removing the same component before the initial decision would have no effect.

Table 7: Complete deterministic policy comparison on 225 worlds. All 12 public policies are included; two pairs coincide numerically on these worlds (keep all with recency, and the event rule with the qualification rule). The hidden-future oracle policy reference follows. Values use the frozen follow-on rollout; the reference is not the exact branch optimum used to define regret. Storage is total authorized decision-memory tokens over all worlds (physical total: 12,142). Bold highlights probe-willing's key results.

Policy	Value ↑	Regret ↓	False ↓	Missed ↓	Agree ↑	Auth. tok.
Keep all	-9.184	11.211	0.427	0.000	0.495	3,043
Archive all	-0.208	2.235	0.000	0.418	0.505	0
Random 50%	-5.740	7.767	0.262	0.209	0.442	1,769
Recency	-9.184	11.211	0.427	0.000	0.495	3,043
Frequency	-10.892	12.919	0.409	0.151	0.337	2,552
Storage cap	-1.731	3.758	0.244	0.040	0.663	1,420
Scope literal	-9.127	11.154	0.360	0.000	0.574	2,883
Event rule	-7.189	9.216	0.267	0.000	0.684	2,694
Memory Worth proxy (Simsek, 2026)	-1.052	3.079	0.147	0.071	0.742	955
SQCAD qualification	-7.189	9.216	0.267	0.000	0.684	2,694
SQCAD + lineage	1.083	0.944	0.196	0.000	0.768	2,305
SQCAD probe-willing	**1.964**	**0.063**	0.067	0.000	**0.921**	**1,088**
Oracle policy reference	1.987	0.040	0.000	0.000	1.000	1,129

Table 8: Continuation ablations with the initial SQCAD + lineage action fixed. Switches affect follow-on rollout components, not the initial decision rule. Regret retains the reference branch optimum; every frozen switch is reported.

Continuation	Regret ↓	Value ↑	False ↓	Missed ↓
Reference	0.9443	1.0827	0.1956	0.0000
no censoring	11.0214	-8.9943	0.1956	0.0000
no lineage	0.9443	1.0827	0.1956	0.0000
no probe	0.9443	1.0827	0.1956	0.0000
no qualification	0.9443	1.0827	0.1956	0.0000
no restore	0.9443	1.0827	0.1956	0.0000

C ILLUSTRATIVE PROXY-SURFACE AUDIT AND IMPLEMENTATION BOUNDARIES

The audit defines study-specific scalar proxies inspired by observable mechanisms in selected papers and compares their fibers with evaluator-only branch contrasts. These scalarizations are author-defined audit objects, not native implementations; every numerical result belongs only to this controlled transport.

Table 9: Study-defined scalar proxies inspired by named methods, plus an internal local-effect control. Fibers use scores rounded to eight decimals. Opposite-sign fibers contain both positive and negative paired contrasts. Floor is the widest-witness randomized lower bound in lifecycle-value units. †Rows marked with a dagger reuse the *same* version-update witness and therefore the same floor; they are one obstruction observed through several formulas, not independent confirmations, and the row count must not be read as a count of failing systems. Rows are illustrative stress tests, not native-system evaluations.

Proxy surface	Fibers	Opposite sign	Floor
ActMem (Zhang et al., 2026b)	No scalar proxy; evaluated in Tables 1–3		
Memory-Worth-inspired success (Simsek, 2026)	5	2	32.82†
Trivium-inspired demand (Chang, 2026)	14	0	0.00
Oblivion-inspired decay (Rana et al., 2026)	1	1	32.82†
FadeMem-inspired decay (Wei et al., 2026)	2	2	4.69
DeMem-inspired distinction (Zou et al., 2026)	6	3	32.82†
GovMem-inspired coverage (Qi et al., 2026)	8	2	3.97
Internal local-effect control	7	2	32.82†

The Memory-Worth-inspired proxy is $(100\rho+1)/102$. The Oblivion- and FadeMem-inspired proxies are $\exp[-\mathrm{age}/(3(u+f))]$ and $\log(1+f)\exp[-\mathrm{age}/(20+40c)]$, respectively. The DeMem-inspired proxy is $|r_i - \bar{r}|$, the GovMem-inspired proxy is session coverage plus $10^{-6}c$, and the Trivium-inspired proxy multiplies the local access effect by discounted public future-query demand. The internal control is local relevance minus half a visible-correction indicator. These author-defined formulas make the theorem's obstruction executable under a common contract; they are neither complete published algorithms nor evidence about native-system failure. The Trivium-inspired proxy uses the public demand summary provided by the audit contract; it is not equivalent to the LLM controller's inferred demand.

The frozen audit excludes an ActMem-inspired scalar because its graph construction and counterfactual retrieval do not determine a unique scalar persistent keep/archive policy. Other literature systems without a compatible single-agent retention score have no numerical audit row. Their absence is a coverage boundary, not evidence that SQCAD outperforms them. The proxy rows motivated by FadeMem, DeMem, and GovMem have no complete five-model public-QA counterpart in this freeze and must not be presented as reproductions of those systems.

Table 10: Evidence ledger for representative Agent Memory systems. Each row separates what the original system reports, what this paper actually executed, the recorded protocol deviation, and the resulting evidence tier. Native values are never pooled with matched-contract values, and no row enters a best/second-best ranking.

System	Native scope and metric	Local implementation	Protocol deviation	Evidence tier
MemGPT / Letta (Packer et al., 2023; Letta Team, 2025)	Hierarchical context paging; stateful agent memory	Literature and software reference	No execution attempted; no lifecycle action replay	Related work only
Generative Agents (Park et al., 2023)	Memory streams, reflection, social simulation	Literature reference	Native task is not an authorization decision	Related work only
MemoryBank (Zhong et al., 2024)	Long-term dialogue memory with temporal/importance cues	Literature reference	Native dialogue protocol differs	Related work only
SimpleMem (Liu et al., 2026)	Lifelong agent memory built by write-time compression	Official code inspected at commit `16912523`; only its keyword-inclusion rule was transportable	The compression and embedding tiers are not reproducible here, so the transported lexical surface is not end-to-end SimpleMem; no SimpleMem result is used in any table	Related work only
Mem0 (Chhikara et al., 2025)	Fact-level add/update/delete/no-op maintenance	Literature reference	No access-time keep/archive action exposed	Related work only
A-MEM (Xu et al., 2025)	Linked note construction and memory evolution	Literature reference	No unified-contract execution	Related work only
Zep / Graphiti (Rasmussen et al., 2025)	Temporal knowledge-graph memory	Literature reference	Graph organization, not a retention policy	Related work only
HippoRAG (Gutiérrez et al., 2024)	Graph-based multi-hop retrieval	Literature reference	Retrieval objective; no access-time authorization	Related work only
ActMem (Zhang et al., 2026b)	Native LoCoMo QA accuracy 83.57% with GPT-4o-mini	Paper-inspired reimplementation on Qwen3-8B; transport adapter used in the primary tables	Thresholds and retrieval depth not aligned to the native 0.2/0.3/0.2 and 20+10; QA accuracy is not token-F1	Primary adapter (paper-inspired)
Memory Worth (Simsek, 2026)	Synthetic calibration, Spearman ρ at 10k episodes 0.89$\pm$0.02	Paper-faithful synthetic calibration: $\rho = 0.8977$ (95% CI [0.8901, 0.9053])	Synthetic calibration matches; end-to-end text protocol still pending	Primary adapter; calibration consistency
Trivium (Chang, 2026)	CausalBench-Seq regret 8.82$\pm$0.14; RLVR 333.33	Probe/log core only; transported values are lifecycle-contract results	Native CausalBench-Seq and CAP-GSM8K pipelines not reproduced	Primary adapter; no native delta
Oblivion (Rana et al., 2026)	LongMemEval-S 90.60 with GPT-4.1-mini; own reader and judge	Official decay core from `nec-research/oblivion` at commit `b2512f9`, implemented under the shared input, action, and reader contract	Backbone and reader replaced; three parameters the release leaves open (decay temperature 10.0, stabilization boost 0.1, utility floor 0.2) fixed from shipped defaults	Primary adapter; official code, matched contract

D Related-work coverage and baseline evidence boundary

For Oblivion, the reported native 90.60 on LongMemEval-S is obtained with GPT-4.1-mini as generator, activator, and decayer, GPT-4o as judge, and `text-embedding-3-small` as the manager embedder, scored by the system's own reader and judge. None of those four roles is reproduced here. The row implements that decay core, $R = \exp(-t/\max\{0.5, (u + 0.1f)T\})$, with the upstream reinforcement and purge rules, under the shared contract, so that its authorized storage and exposure are measured on the same scale as every other row. The three parameters the release leaves open are fixed from shipped defaults rather than tuned: we did not search them. Reading the row as a comparison to the paper's 90.60 would be a category error, and we do not compute that delta. Reading it as the same decay mechanism measured under a common contract is what the table supports. Controller replies that could not be parsed default to archive under the frozen harness; a dedicated retry pass replaced all such replies, and the reported row contains none.

For Trivium, the reported native temporal-regret value 8.82$\pm$0.14 is measured on CausalBench-Seq, whereas the RLVR reference is 333.33; neither uses LoCoMo token-F1 or LifecycleBench regret. Our local artifact verifies persistent causal-log and probe-budget bookkeeping, but does not claim the paper's change-point detector, causal-model revision, or budgeted intervention learner. On LifecycleBench, the transported Trivium-inspired adapter obtains regret 2.8085, 2.1348, 1.4503, 1.1910, and 0.9406 for Qwen3-8B through GPT-5.6-terra, respectively. These values are included to compare the shared lifecycle contract, not to report a native Trivium reproduction.

The broader Agent Memory literature covers stateful architectures, fact maintenance, associative graphs, retrieval, compression, decay, outcome-conditioned worth, and causal probing: hierarchical paging and reflective streams (Packer et al., 2023; Letta Team, 2025; Park et al., 2023; Zhong et al., 2024), fact-level state maintenance (Chhikara et al., 2025), and associative graphs (Xu et al., 2025; Rasmussen et al., 2025; Gutiérrez et al., 2024) on one side; write-time compression (Liu et al., 2026), reasoning-coupled retrieval (Zhang et al., 2026b), outcome-conditioned worth (Simsek, 2026), temporal causal revision (Chang, 2026), and provenance-aware write governance (Qi et al., 2026) on the other, with DeMem tying compression to downstream decision loss (Zou et al., 2026). Surveys organize these systems by representation, operation, and evaluation target (Zhang et al., 2025; Du et al., 2025). Hierarchical working memory and memory operating systems develop capacity and tiering (Hu et al., 2024; Kang et al., 2025); SAGE and ReMe study write-time novelty and procedural refinement (Wang et al., 2026; Cao et al., 2026). Reflective retrieval, memory–reasoning co-design, pre-storage reasoning, and compositional memory further broaden the design space (Tan et al., 2025b; Zhou et al., 2025; Kim et al., 2025; Ke et al., 2025). MIRIX, HingeMem, MemSearch-o1, TrustMem, and Infini Memory explore typed multi-agent stores, adaptive retrieval, reasoning-aligned growth, verified consolidation, and maintainable topic documents (Wang & Chen, 2025; Zhong et al., 2026; Zhang et al., 2026a; Yang et al., 2026; Ji et al., 2026). Retrieval–utilization diagnosis cautions against conflating memory construction, retrieval, and downstream use (Yuan et al., 2026); MemBench and MemoryAgentBench expose complementary evaluation requirements (Tan et al., 2025a; Hu et al., 2025). We cite these lines to position SQCAD's estimand; we do not imply that systems without an access-time action are weak baselines. The primary comparison is deliberately narrower and should be read together with this audit.

E Execution contracts and provenance

E.1 Backbones and reader evaluation

The five displayed backbone names map to the recorded model identifiers `qwen3-8b`, `deepseek-flash`, `glm-5.3-flash`, `claude-sonnet-5`, and `gpt-5.6-terra`, which are the labels this contract uses throughout; earlier internal artifacts carry a development string for the second of them, and the tables use the label. Qwen runs use the local open-weight endpoint; the other rows use provider or relay endpoints. Response identifiers and source hashes are recorded. A relay's response string does not independently establish the identity or unchanged weights of a proprietary backend, so these are endpoint-conditioned comparisons.

The public-task reader receives each method's fixed context and generates an answer with temperature zero. Token limits and request details belong to the individual frozen run records; contexts are shared by design. The LoCoMo contract scores 1,986 questions using the official scoring implementation. LongMemEval-S evaluates 500 questions with token F1 under the shared contract, rather than its native judge-based accuracy. Predictions, not historical extractive-readout summaries, are the sources of Tables 1 and 2. All quantities are reported on a 0–1 scale except lifecycle utility and token counts. Token accounting is the frozen contract's accounting and should not be interpreted as provider billing.

E.2 Lifecycle controller transformations

Only introduced memories whose scope is the decision scope or "any" enter the controller view. Let q and m be the sets of lowercase alphanumeric tokens of the current question and memory text. The base support is

$$r(q,m) = \frac{|q \cap m|}{\sqrt{|q||m|}}, \tag{12}$$

with value zero for empty sets. FIR computes $0.25r_{j-1} + 0.50r_j + 0.25r_{j+1}$ in public introduction order, pads endpoints by repetition, and sorts by decreasing score with stable memory-ID ties. The SP controller uses consecutive visible user–assistant pairs (u, v) and scores a memory by

$$s(m) = r(q,m) + 0.20 \max_{(u,v)} r(m,u)\big[1 + 0.25r(q,v)\big], \tag{13}$$

with zero edge support when no pair exists. It then sorts by score and memory ID. Views contain at most 16 memories; if necessary the last item is replaced by the decision memory so that this

target remains visible. The controller also receives the visible sessions, decision scope, and current question, and must return keep or archive in JSON. This lifecycle SP transform does not apply the 1.35 conversational packing prior verbatim. These two instantiations test how the design ingredients interact with two different consumers of evidence.

On the public conversational tasks the same ingredients carry task-specific constants. FIR applies $\ell_j = 0.25r_{j-1} + 0.50r_j + 0.25r_{j+1}$ in observable event order with endpoint padding by repetition, and greedy packing combines $d_j = \ell_j - 0.08|r_j - \ell_j|$ with a token-Jaccard redundancy penalty of 0.22 under item and reader-token budgets. The conversational SP prior prioritizes the next response to a query-matching interrogative turn within the same session at priority 1.35 with at least two shared non-stopword query terms. Both use guard budget 16.

Each world's branch utility is the generator's pre-registered cost contract rather than a $[0, 1]$ score. Over ten future task epochs discounted by $\gamma = 0.9$, a rollout adds $10\times$difficulty for each successful task and subtracts 20 per wrong-use adoption of a stale or wrong-scope record, 0.05 per exposed memory per task, 1.0 per paid probe (charged even when the probe earns no workspace slot), and 0.01 per stored token per epoch; archived records carry no storage cost and remain recoverable, so archiving trades storage and exposure against probe and recovery cost. Oracle labels follow the sign of the paired contrast with a neutral band: keep above $\tau = 0.5$, archive below $-\tau$, and neutral inside it. On this scale a regret gap of 0.8 (the distance between the two rows the main text quotes) exceeds the whole neutral band, so it is a substantive decision difference rather than a rounding artifact.

On LifecycleBench, Value is the chosen branch utility, Regret the best paired branch value minus that utility, Missed the fraction archived against a keep label, and Recover retention under the fixed budget; the main text reads these columns by name. Recover counts kept records plus archived ones still retrievable within that budget, so it is bounded below by the kept fraction rather than by method quality, and because the rescue budget is fixed and most archived records stay recoverable inside it, the column spans only 0.573 to 0.662 across every method and backbone in Table 3. It is reported for completeness and should not be read as a discriminator between controllers. The controller never receives paired branch values, oracle actions, future outcomes, or hidden labels. The evaluator selects the precomputed keep or archive value for the chosen action. The continuation simulator can probe archived candidates, charges each probe, and can restore qualifying evidence. These operations are fixed by the rollout contract, not requested by the LLM. False commit is the fraction of all worlds where keep is chosen against an archive label; missed commit reverses these actions. Near-neutral worlds count as neither error, while regret uses the exact maximum branch value. Keep is the chosen keep fraction. Recover counts kept memories and archived memories recoverable under the frozen rescue budget. Every model–method row includes all 225 worlds. The deterministic oracle policy resolves neutral labels as keep and is not the exact regret-minimizing branch selector on those worlds, explaining its nonzero regret.

E.3 Uncertainty and ablation scope

All 75 primary-comparison cells and 15 internal-control cells are point estimates from frozen completed runs; the six-method audit matrix remains available in the provenance artifacts. The paper does not assign significance from small F1 differences alone or treat outputs at temperature zero as intrinsically deterministic across external providers. Questions within a conversation and worlds within a mechanism family can be dependent. The standalone LifecycleBench scorer provides family-cluster percentile intervals with 2,000 draws and a fixed seed. Appendix F adds paired cluster analyses on matched frozen outputs, without treating tokens or repeated backbones as independent replicates.

In the continuation ablation, the initial action is computed by the unchanged lineage-aware policy before an ablated rollout is executed. Regret is computed against the original paired-branch optimum, not a newly optimized ablated oracle. Consequently, the experiment measures the value effect of continuation changes under fixed initial decisions. Four of the five switches reproduce the reference regret to four decimals. With the initial action fixed, and with these worlds' follow-on rollouts short, the lineage, probe, qualification, and restore rules change no chosen branch, so the ablation bounds their contribution on these worlds rather than measuring it; censoring is the layer that binds, because it changes which records a later decision can see. The null rows are reported deliberately: a table carrying only the switch that moves would leave the other four untested without saying so. The

decision-policy comparison and this continuation test do not estimate the uniform closure errors required by Theorem 3.

E.4 Released artifact contents

The supplementary release contains the 225 paired counterfactual worlds with their public decision-time and evaluator-side hidden layers, the frozen paired rollout branches used to score value and regret, a reference scorer, an example controller, and a manifest recording the version, the world count, and the world-signature uniqueness rate. The interactive case study is an authored walkthrough rather than a recorded run: four constructed scenarios step through the qualification state and their paired continuations, so that the object the main text argues about is inspectable rather than only described. Its scenarios exercise the mechanism rather than sample from LifecycleBench or any evaluated system, and its illustrations are generated images that carry no measurement; the release records each panel's provenance and hash beside the storyline text. Source, tests, and the table-provenance records are released alongside them.

E.5 Frozen numerical sources

The three primary dataset tables and the internal control are generated from one frozen submission matrix with 121 checked source artifacts, of which 20 are the separately run official-code artifacts behind the Oblivion rows. All seven-method results remain in the frozen matrix and provenance audit; the primary tables display the three lifecycle-compatible adapters plus the two SQCAD guards, while the untransformed SQCAD control appears in Table 5. The deterministic table and continuation switches come from the frozen lifecycle policy matrix; family means are aggregated from the released per-world policy rows and checked against that matrix. The score audit is a separate artifact. Generation validates five models and seven source methods per dataset. Repeated public exposure and storage values follow from contexts fixed before generation. The accompanying `paper_table_provenance.json` records SHA-256 hashes for sources and generated tables. Dataset display names omit development versions; executable manifests retain exact byte identities.

F Paired inference conditional on frozen runs

The primary comparisons use one frozen set of predictions per configuration; requests specify temperature zero. This does not make external endpoints deterministic or provide independent model reruns. Some lifecycle SQCAD artifacts omit the temperature field; their request builder specifies zero. Existing LifecycleBench files already include family-cluster intervals. The additional tests here quantify dataset-unit variation conditional on the stored outputs and do not estimate provider or model-run variance.

We reconstruct LoCoMo per-question F1 with the frozen official scorer, read stored LongMemEval-S per-question token F1, and recompute every lifecycle regret from its chosen action and the exact paired values. All 105 audited method-model-dataset means match the frozen matrix within 10^{-10}, with no new model calls. IDs must match exactly before pairing; no missing rows are dropped.

The analysis includes both guard configurations against four audited adapters on each of five backbones and three datasets, giving 120 comparisons. It was specified after the original results were known, so it is a post hoc analysis, not a preregistered confirmatory study. Each effect is the mean paired F1 gain or lifecycle regret reduction. We resample whole clusters with replacement for 10,000 percentile-bootstrap draws. Clusters are the 10 LoCoMo conversations and nine LifecycleBench mechanism families. Clusterwise sign flipping gives two-sided p-values, enumerating all 1,024 or 512 assignments, respectively. The test assumes independent, sign-exchangeable cluster differences under its null; constructed families are not a random sample of all possible mechanisms. The small number of clusters limits inferential resolution, and for these two datasets it removes it rather than reducing it: with nine LifecycleBench families the smallest attainable two-sided sign-flip p-value is $2/512 = 0.0039$, and with ten LoCoMo conversations it is $2/1024 = 0.0020$, so within a 40-comparison family the smallest attainable Holm value is 0.156 and 0.078 respectively. Neither dataset can reject at any effect size, and we report both as descriptive point estimates rather than as tested nulls.

LongMemEval-S contains shared answer-session sources. We group questions into 495 connected components of shared answer-session IDs and use 32,768 Monte Carlo cluster sign assignments with the plus-one correction. However, shared distractor sessions connect all 500 full histories into one component. These LongMemEval-S intervals and p-values therefore condition on fixed distractor material and assume independence of answer-source components; they do not support unconditional full-history inference. A one-cluster full-history analysis has no useful replication. We retain this dependence limitation alongside the results.

All tests use seed 20260913 with fixed dataset offsets. The generator's redraw unit is the (family, variant) pair, which yields 14 design cells, whereas every cluster above is a mechanism family, of which there are nine. The coarser unit is the conservative choice for interval width, and it is the one the released scorer names; the counts differ because a variant is a redraw inside a family rather than an independent mechanism. Holm correction is applied within each dataset's 40 comparisons. A joint correction over all 120 would fold in 80 comparisons that cannot reject at any effect size (previous paragraph), raising the rank-one threshold for the comparisons that can from $0.05/40 = 0.00125$ to $0.05/120 = 0.00042$; the raw p-values are unaffected by this choice and the joint adjusted values are retained in the provenance artifact. Displayed confidence intervals are unadjusted and do not encode the Holm decision. At the 0.05 family-wise threshold, no LoCoMo or LifecycleBench comparison rejects its null; 34 LongMemEval-S comparisons do, of which 32 favour SQCAD against the three matched-contract adapters and 2 favour ActMem, under the conditional assumptions above. Non-rejection does not establish equivalence or erase the observed point differences. The main text retains descriptive comparisons, without converting a best point estimate into a significance claim. Tables 11–13 include all effects, losses, intervals, and raw and adjusted p-values.

Table 11: LoCoMo: all 40 paired guard-versus-adapter comparisons. Positive effects favor SQCAD: F1 gain on public QA, regret reduction on LifecycleBench. Intervals are unadjusted 95% cluster-bootstrap intervals. Raw two-sided sign-flip p-values and Holm p-values corrected within each dataset's 40 comparisons are both shown. F/S denotes FIR/SP guard. The first three comparators are the matched-contract adapters of Tables 1–3; Oblivion is the official decay core run under the same contract (Appendix D).

Backbone	Guard	Adapter	Effect	95% interval	p	p_{Holm}
Qwen3-8B	F	Oblivion (Rana et al., 2026)	0.1988	[0.1839, 0.2122]	0.0020	0.0781
	F	ActMem (Zhang et al., 2026b)	0.0142	[-0.0011, 0.0325]	0.1465	0.5391
	F	Memory Worth (Simsek, 2026)	0.2223	[0.1991, 0.2437]	0.0020	0.0781
	F	Trivium (Chang, 2026)	0.2265	[0.2067, 0.2468]	0.0020	0.0781
	S	Oblivion (Rana et al., 2026)	0.2130	[0.1979, 0.2263]	0.0020	0.0781
	S	ActMem (Zhang et al., 2026b)	0.0284	[0.0142, 0.0449]	0.0059	0.0781
	S	Memory Worth (Simsek, 2026)	0.2364	[0.2141, 0.2569]	0.0020	0.0781
	S	Trivium (Chang, 2026)	0.2406	[0.2215, 0.2606]	0.0020	0.0781
DeepSeek-Flash	F	Oblivion (Rana et al., 2026)	0.2246	[0.2001, 0.2469]	0.0020	0.0781
	F	ActMem (Zhang et al., 2026b)	0.0169	[-0.0005, 0.0362]	0.1348	0.5391
	F	Memory Worth (Simsek, 2026)	0.2544	[0.2255, 0.2800]	0.0020	0.0781
	F	Trivium (Chang, 2026)	0.2563	[0.2292, 0.2807]	0.0020	0.0781
	S	Oblivion (Rana et al., 2026)	0.2433	[0.2199, 0.2651]	0.0020	0.0781
	S	ActMem (Zhang et al., 2026b)	0.0356	[0.0204, 0.0516]	0.0039	0.0781
	S	Memory Worth (Simsek, 2026)	0.2731	[0.2465, 0.2964]	0.0020	0.0781
	S	Trivium (Chang, 2026)	0.2750	[0.2494, 0.2978]	0.0020	0.0781
GLM-5.3-Flash	F	Oblivion (Rana et al., 2026)	0.2203	[0.2032, 0.2376]	0.0020	0.0781
	F	ActMem (Zhang et al., 2026b)	0.0174	[0.0074, 0.0276]	0.0137	0.0957
	F	Memory Worth (Simsek, 2026)	0.2297	[0.2098, 0.2464]	0.0020	0.0781
	F	Trivium (Chang, 2026)	0.2344	[0.2138, 0.2516]	0.0020	0.0781
	S	Oblivion (Rana et al., 2026)	0.2326	[0.2177, 0.2468]	0.0020	0.0781
	S	ActMem (Zhang et al., 2026b)	0.0298	[0.0180, 0.0407]	0.0039	0.0781
	S	Memory Worth (Simsek, 2026)	0.2421	[0.2236, 0.2573]	0.0020	0.0781
	S	Trivium (Chang, 2026)	0.2468	[0.2278, 0.2629]	0.0020	0.0781
Claude Sonnet 5	F	Oblivion (Rana et al., 2026)	0.2257	[0.2088, 0.2402]	0.0020	0.0781
	F	ActMem (Zhang et al., 2026b)	0.0099	[-0.0061, 0.0275]	0.3066	0.6133
	F	Memory Worth (Simsek, 2026)	0.2518	[0.2286, 0.2745]	0.0020	0.0781
	F	Trivium (Chang, 2026)	0.2546	[0.2323, 0.2760]	0.0020	0.0781
	S	Oblivion (Rana et al., 2026)	0.2424	[0.2246, 0.2571]	0.0020	0.0781
	S	ActMem (Zhang et al., 2026b)	0.0266	[0.0081, 0.0468]	0.0254	0.1523
	S	Memory Worth (Simsek, 2026)	0.2685	[0.2435, 0.2930]	0.0020	0.0781
	S	Trivium (Chang, 2026)	0.2713	[0.2477, 0.2941]	0.0020	0.0781
GPT-5.6-terra	F	Oblivion (Rana et al., 2026)	0.2217	[0.2060, 0.2365]	0.0020	0.0781
	F	ActMem (Zhang et al., 2026b)	0.0020	[-0.0159, 0.0192]	0.8359	0.8359
	F	Memory Worth (Simsek, 2026)	0.2566	[0.2363, 0.2763]	0.0020	0.0781
	F	Trivium (Chang, 2026)	0.2584	[0.2390, 0.2769]	0.0020	0.0781
	S	Oblivion (Rana et al., 2026)	0.2356	[0.2156, 0.2539]	0.0020	0.0781
	S	ActMem (Zhang et al., 2026b)	0.0159	[0.0009, 0.0322]	0.0898	0.4492
	S	Memory Worth (Simsek, 2026)	0.2705	[0.2478, 0.2910]	0.0020	0.0781
	S	Trivium (Chang, 2026)	0.2723	[0.2508, 0.2923]	0.0020	0.0781

Table 12: LongMemEval-S: all 40 paired guard-versus-adapter comparisons. Positive effects favor SQCAD: F1 gain on public QA, regret reduction on LifecycleBench. Intervals are unadjusted 95% cluster-bootstrap intervals. Raw two-sided sign-flip p-values and Holm p-values corrected within each dataset's 40 comparisons are both shown. F/S denotes FIR/SP guard. The first three comparators are the matched-contract adapters of Tables 1–3; Oblivion is the official decay core run under the same contract (Appendix D). LongMemEval-S intervals condition on fixed shared distractors; they do not establish independence across full histories.

Backbone	Guard	Adapter	Effect	95% interval	p	p_{Holm}
Qwen3-8B	F	Oblivion (Rana et al., 2026)	0.2408	[0.2047, 0.2777]	< 0.0001	0.0012
	F	ActMem (Zhang et al., 2026b)	-0.0430	[-0.0699, -0.0171]	0.0018	0.0146
	F	Memory Worth (Simsek, 2026)	0.2627	[0.2265, 0.2997]	< 0.0001	0.0012
	F	Trivium (Chang, 2026)	0.2725	[0.2360, 0.3096]	< 0.0001	0.0012
	S	Oblivion (Rana et al., 2026)	0.2459	[0.2107, 0.2820]	< 0.0001	0.0012
	S	ActMem (Zhang et al., 2026b)	-0.0379	[-0.0649, -0.0115]	0.0051	0.0355
	S	Memory Worth (Simsek, 2026)	0.2678	[0.2320, 0.3047]	< 0.0001	0.0012
	S	Trivium (Chang, 2026)	0.2776	[0.2423, 0.3142]	< 0.0001	0.0012
DeepSeek-Flash	F	Oblivion (Rana et al., 2026)	0.3269	[0.2881, 0.3671]	< 0.0001	0.0012
	F	ActMem (Zhang et al., 2026b)	-0.0162	[-0.0432, 0.0110]	0.2428	1.0000
	F	Memory Worth (Simsek, 2026)	0.3638	[0.3256, 0.4026]	< 0.0001	0.0012
	F	Trivium (Chang, 2026)	0.3752	[0.3379, 0.4139]	< 0.0001	0.0012
	S	Oblivion (Rana et al., 2026)	0.3482	[0.3084, 0.3884]	< 0.0001	0.0012
	S	ActMem (Zhang et al., 2026b)	0.0052	[-0.0232, 0.0336]	0.7225	1.0000
	S	Memory Worth (Simsek, 2026)	0.3852	[0.3471, 0.4236]	< 0.0001	0.0012
	S	Trivium (Chang, 2026)	0.3965	[0.3598, 0.4348]	< 0.0001	0.0012
GLM-5.3-Flash	F	Oblivion (Rana et al., 2026)	0.2537	[0.2245, 0.2834]	< 0.0001	0.0012
	F	ActMem (Zhang et al., 2026b)	0.0017	[-0.0244, 0.0281]	0.9003	1.0000
	F	Memory Worth (Simsek, 2026)	0.2180	[0.1894, 0.2485]	< 0.0001	0.0012
	F	Trivium (Chang, 2026)	0.2259	[0.1965, 0.2560]	< 0.0001	0.0012
	S	Oblivion (Rana et al., 2026)	0.2271	[0.1989, 0.2562]	< 0.0001	0.0012
	S	ActMem (Zhang et al., 2026b)	-0.0249	[-0.0511, 0.0003]	0.0580	0.3479
	S	Memory Worth (Simsek, 2026)	0.1915	[0.1640, 0.2205]	< 0.0001	0.0012
	S	Trivium (Chang, 2026)	0.1993	[0.1717, 0.2278]	< 0.0001	0.0012
Claude Sonnet 5	F	Oblivion (Rana et al., 2026)	0.3937	[0.3552, 0.4330]	< 0.0001	0.0012
	F	ActMem (Zhang et al., 2026b)	-0.0148	[-0.0452, 0.0149]	0.3414	1.0000
	F	Memory Worth (Simsek, 2026)	0.3870	[0.3492, 0.4264]	< 0.0001	0.0012
	F	Trivium (Chang, 2026)	0.3942	[0.3567, 0.4337]	< 0.0001	0.0012
	S	Oblivion (Rana et al., 2026)	0.3974	[0.3592, 0.4361]	< 0.0001	0.0012
	S	ActMem (Zhang et al., 2026b)	-0.0110	[-0.0397, 0.0178]	0.4662	1.0000
	S	Memory Worth (Simsek, 2026)	0.3908	[0.3524, 0.4300]	< 0.0001	0.0012
	S	Trivium (Chang, 2026)	0.3980	[0.3609, 0.4369]	< 0.0001	0.0012
GPT-5.6-terra	F	Oblivion (Rana et al., 2026)	0.3235	[0.2870, 0.3605]	< 0.0001	0.0012
	F	ActMem (Zhang et al., 2026b)	0.1154	[0.0809, 0.1503]	< 0.0001	0.0012
	F	Memory Worth (Simsek, 2026)	0.3500	[0.3141, 0.3865]	< 0.0001	0.0012
	F	Trivium (Chang, 2026)	0.3514	[0.3155, 0.3868]	< 0.0001	0.0012
	S	Oblivion (Rana et al., 2026)	0.3187	[0.2824, 0.3560]	< 0.0001	0.0012
	S	ActMem (Zhang et al., 2026b)	0.1106	[0.0752, 0.1476]	< 0.0001	0.0012
	S	Memory Worth (Simsek, 2026)	0.3452	[0.3090, 0.3823]	< 0.0001	0.0012
	S	Trivium (Chang, 2026)	0.3466	[0.3113, 0.3838]	< 0.0001	0.0012

Table 13: LifecycleBench: all 40 paired guard-versus-adapter comparisons. Positive effects favor SQCAD: F1 gain on public QA, regret reduction on LifecycleBench. Intervals are unadjusted 95% cluster-bootstrap intervals. Raw two-sided sign-flip p-values and Holm p-values corrected within each dataset's 40 comparisons are both shown. F/S denotes FIR/SP guard. The first three comparators are the matched-contract adapters of Tables 1–3; Oblivion is the official decay core run under the same contract (Appendix D).

Backbone	Guard	Adapter	Effect	95% interval	p	p_{Holm}
Qwen3-8B	F	Oblivion (Rana et al., 2026)	1.0171	[0.0680, 2.7204]	0.1250	1.0000
	F	ActMem (Zhang et al., 2026b)	0.8387	[0.0000, 2.5051]	0.5000	1.0000
	F	Memory Worth (Simsek, 2026)	1.1050	[0.0760, 2.8426]	0.1250	1.0000
	F	Trivium (Chang, 2026)	1.0896	[0.1000, 2.8015]	0.1250	1.0000
	S	Oblivion (Rana et al., 2026)	-1.5754	[-4.8965, 0.1871]	0.7500	1.0000
	S	ActMem (Zhang et al., 2026b)	-1.7538	[-5.0069, 0.0000]	0.5000	1.0000
	S	Memory Worth (Simsek, 2026)	-1.4875	[-4.8166, 0.2911]	0.8750	1.0000
	S	Trivium (Chang, 2026)	-1.5029	[-4.8560, 0.3000]	1.0000	1.0000
DeepSeek-Flash	F	Oblivion (Rana et al., 2026)	1.4652	[-0.0424, 4.4320]	0.7500	1.0000
	F	ActMem (Zhang et al., 2026b)	1.7262	[0.0058, 5.0844]	0.2500	1.0000
	F	Memory Worth (Simsek, 2026)	1.6185	[0.0058, 4.5751]	0.2500	1.0000
	F	Trivium (Chang, 2026)	1.0424	[0.0058, 2.9197]	0.2500	1.0000
	S	Oblivion (Rana et al., 2026)	0.7290	[-0.6159, 2.7662]	0.7500	1.0000
	S	ActMem (Zhang et al., 2026b)	0.9900	[-0.4351, 3.3685]	0.7500	1.0000
	S	Memory Worth (Simsek, 2026)	0.8823	[-0.1559, 2.7662]	0.7500	1.0000
	S	Trivium (Chang, 2026)	0.3063	[-0.2650, 1.1471]	0.7500	1.0000
GLM-5.3-Flash	F	Oblivion (Rana et al., 2026)	1.7141	[0.0003, 5.0269]	0.2500	1.0000
	F	ActMem (Zhang et al., 2026b)	1.4954	[-0.0008, 4.4552]	0.5000	1.0000
	F	Memory Worth (Simsek, 2026)	1.3547	[0.0051, 3.9178]	0.1250	1.0000
	F	Trivium (Chang, 2026)	0.5494	[-0.0059, 1.6541]	1.0000	1.0000
	S	Oblivion (Rana et al., 2026)	1.6792	[0.0000, 4.9918]	0.5000	1.0000
	S	ActMem (Zhang et al., 2026b)	1.4605	[-0.0555, 4.4384]	1.0000	1.0000
	S	Memory Worth (Simsek, 2026)	1.3198	[-0.0363, 3.9003]	0.3750	1.0000
	S	Trivium (Chang, 2026)	0.5145	[-0.1038, 1.6518]	1.0000	1.0000
Claude Sonnet 5	F	Oblivion (Rana et al., 2026)	1.5806	[-0.2390, 4.9748]	0.7500	1.0000
	F	ActMem (Zhang et al., 2026b)	0.9868	[0.0043, 2.8140]	0.1250	1.0000
	F	Memory Worth (Simsek, 2026)	0.8908	[-0.0592, 2.7276]	0.7500	1.0000
	F	Trivium (Chang, 2026)	0.2082	[0.0000, 0.5613]	0.5000	1.0000
	S	Oblivion (Rana et al., 2026)	1.4225	[-0.2481, 4.4704]	0.6250	1.0000
	S	ActMem (Zhang et al., 2026b)	0.8287	[0.0358, 2.3095]	0.0625	1.0000
	S	Memory Worth (Simsek, 2026)	0.7327	[-0.0732, 2.2263]	0.4375	1.0000
	S	Trivium (Chang, 2026)	0.0501	[-0.0057, 0.1277]	0.5000	1.0000
GPT-5.6-terra	F	Oblivion (Rana et al., 2026)	1.3929	[0.1074, 3.8043]	0.0312	1.0000
	F	ActMem (Zhang et al., 2026b)	1.0729	[-0.2045, 3.3620]	0.6250	1.0000
	F	Memory Worth (Simsek, 2026)	-0.0529	[-0.2979, 0.1159]	1.0000	1.0000
	F	Trivium (Chang, 2026)	-0.4100	[-1.1739, 0.0762]	0.4375	1.0000
	S	Oblivion (Rana et al., 2026)	1.3251	[0.0173, 3.7470]	0.1250	1.0000
	S	ActMem (Zhang et al., 2026b)	1.0052	[-0.4074, 3.3585]	0.6875	1.0000
	S	Memory Worth (Simsek, 2026)	-0.1206	[-0.5526, 0.1913]	0.6875	1.0000
	S	Trivium (Chang, 2026)	-0.4778	[-1.2892, 0.1273]	0.3438	1.0000

F.1 LEAVE-ONE-FAMILY-OUT REFITTING OF THE OFFLINE SUMMARIES

Table 14: Leave-one-mechanism-family-out refitting on the 225 released worlds. Each summary is fitted on eight mechanism families and evaluated on the ninth, so the held-out family contributes no cell of its own to the rule that decides it, and every world whose cell is absent falls back to the training-wide mean contrast. In sample the three summaries are ordered by a factor of 30; under family shift they are identical. The diagnostic is evaluator-side, adds no model calls, and changes no frozen score.

Public summary / rule	In-sample regret	Held-out-family regret
Overlap / relevance	9.7743	1.9851
+ scope, correction / qualify	9.2162	1.9851
+ lineage / qualify-lineage	0.3213	1.9851

Table 14 refits each offline summary with one mechanism family held out, as described in Appendix A.2. The artifact records the per-family regrets and the count of worlds whose cell was absent from training at each level; it is the diagnostic a natural-trajectory corpus would have to displace.

F.2 MECHANISM-FAMILY WEIGHTS

The nine families are equally sized, so every macro mean in Appendix B is a uniform average, and the main text reports that one family carries 89.8% of the qualification policy's regret. Table 15 varies the share s carried by version_update while the other eight split the remainder. Two values validate the sweep: $s = 1/9$ reproduces the tabulated means exactly, and $s = 0$ reproduces the mean-excluding-version-update row the main text quotes. The ordering of the policies is unchanged across the whole range, so the comparison is not an artifact of the weighting. The weighting does determine magnitude: qualification moves by a factor of 24 between $s = 0$ and $s = 1/4$, and the overlap and scope information floors move with it, while probe-willing and the lineage floor move by less than 0.03. Those two are precisely the quantities that do not depend on version_update, and they are the two the claims rest on. The alternative weights are illustrative rather than measured family frequencies in natural traffic, which we do not have.

Table 15: Family-weight sensitivity on the 225 released worlds. The nine mechanism families are equally sized, so every macro mean is a uniform average over them; this varies the share s carried by version_update while the other eight families split the remainder, and $s = 1/9$ reproduces the tabulated macro means exactly. Regret is lower-is-better throughout. The policy ordering never changes, but the two quantities that the version-update family dominates move by more than an order of magnitude across the range, while the two that it does not are nearly flat.

version_update share s	0	$1/18$	$1/9$	$1/4$
Probe-willing regret	0.0713	0.0673	0.0634	0.0535
+ lineage, conflict	1.0623	1.0033	0.9443	0.7967
Qualification	1.0623	5.1393	9.2162	19.4086
Memory Worth proxy	1.6010	2.3402	3.0794	4.9274
Oracle policy reference	0.0445	0.0421	0.0396	0.0334
Overlap / scope information floor	0.0713	0.8955	1.7196	3.7801
Lineage information floor	0.0713	0.0673	0.0634	0.0535