ELASTICMEM: LATENT MEMORY AS A LEARNABLE RESOURCE FOR LLM AGENTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Long-term memory is essential for LLM agents to reason coherently across extended interactions, personalize responses, and reuse past experience. However, existing memory-augmented methods typically treat memory as a fixed resource: text-space approaches concatenate retrieved memories into the context window, causing substantial token overhead and sensitivity to noisy evidence, while latent-space approaches reduce textual cost but still rely on rigid retrieval or fixed-capacity memory interfaces. This creates a mismatch between query-dependent memory utility and fixed memory allocation. We propose `ElasticMem`, a memory-augmented LLM framework that learns to use memory as an elastic latent resource. `ElasticMem` builds an offline latent memory bank with retrieval keys and content caches, retrieves memories adaptively from the reasoner's hidden state, assigns each retrieved memory a variable latent budget through a learned policy, and injects selected latent states as soft memory tokens for generation. The full memory-use process is optimized with downstream task rewards through group-relative policy optimization. We evaluate `ElasticMem` on MemorySuite, covering memory-intensive QA and embodied agent control. Across Qwen2.5-3B-Instruct and Qwen2.5-7B-Instruct backbones, `ElasticMem` improves weighted-average QA accuracy by 26.2% and 24.6% and ALFWorld success rate by 66.3% and 27.2% relative to the strongest baselines, while consuming the fewest tokens on ALFWorld on average. Ablations and qualitative analyses further show that adaptive retrieval and elastic budget allocation help `ElasticMem` prioritize useful evidence and transferable plans beyond rigid cosine similarity.

1 INTRODUCTION

Long-term memory has become essential for enabling large language model (LLM) agents to reason coherently across extended interactions, personalize responses across sessions, and reuse past experience in long-horizon tasks (Park et al., 2023; Shinn et al., 2023; Zhong et al., 2024; Wang et al., 2023; Hu et al., 2025; Zhang et al., 2025b). In realistic settings, an agent may need to retrieve user preferences from past conversations (Sun et al., 2025; Pan et al., 2025), track facts scattered across long dialogue histories (Du et al., 2025; Tavakoli et al., 2025), or reuse procedural experience from previous trajectories (Zheng et al., 2023; Mi et al., 2026). Thus, memory is a core mechanism for extending LLM agents beyond the current context window, yet retrieved memories must compete for the reasoner's limited context capacity (Yan et al., 2025).

Despite this importance, existing memory-augmented LLM systems largely treat memory as a fixed resource, as summarized in Table 1. Text-space memory methods store memories as natural-language records and retrieve, summarize, update, or organize them before injecting selected memories into the prompt (Zhong et al., 2024; Xu et al., 2025a; Fang et al., 2025a; Chhikara et al., 2025; Kang et al., 2025; Fang et al., 2025b; LangChain, 2024). While interpretable, this paradigm couples memory use with prompt length: irrelevant or redundant memories still consume context tokens, and useful memories may be diluted by noisy textual evidence. Latent-space memory methods reduce textual overhead by compressing memory into continuous vectors or soft tokens (Zhang et al., 2025a; Chevalier et al., 2023; Wang et al., 2025). However, they typically expose a rigid memory interface: retrieval is often based on fixed similarity scores, and each retrieved or compressed memory receives a uniform or largely pre-defined latent capacity (Karpukhin et al., 2020; Reimers & Gurevych, 2019; Zhang et al., 2026). As a result, a memory chunk that is lexically similar to the query may be

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.

Table 1: **Comparison with representative memory-augmented LLM paradigms.** Text-space memory methods retrieve and inject natural-language memories, which are interpretable but token-expensive. Existing latent-space methods reduce textual overhead, but they still rely on fixed-capacity memory interfaces or rigid retrieval mechanisms. `ElasticMem` performs reasoner-state adaptive retrieval and allocates elastic latent capacity to each retrieved memory.

Paradigm	Representative Methods	Memory Space	Token Efficient	Reasoner-State Adaptive Retrieval	Elastic Capacity
Textual long-term memory	MemoryBank (Zhong et al., 2024) / Mem0 (Chhikara et al., 2025) / MemoryOS (Kang et al., 2025)	Text	✗	✗	✗
Textual agent memory	A-MEM (Xu et al., 2025a) / LightMem (Fang et al., 2025a) / MemP (Fang et al., 2025b)	Text	✗	✗	✗
Fixed latent compression	AutoCompressor (Chevalier et al., 2023)	Latent	✓	✗	✗
Latent memory retrieval	MemGen (Zhang et al., 2025a) / M+ (Wang et al., 2025)	Latent	✓	✗	✗
Elastic latent memory	`ElasticMem` (ours)	Latent	✓	✓	✓

retrieved even if it provides little useful evidence, while a less similar chunk containing the key fact or transferable plan may be underused. This leaves a central gap: memory usefulness is query- and task-dependent, yet both which memories are retrieved and how much capacity each receives are fixed by design choices rather than learned from utility.

This mismatch motivates a different view of memory use. Rather than asking how to retrieve more memories or compress all memories equally, we ask: *how can an LLM agent learn to retrieve and allocate memory adaptively, so that useful memories receive more representational budget while unhelpful memories are compressed or suppressed?* Answering this question requires addressing three challenges. ❶ **Retrieval should adapt to the current query and the reasoner's internal state.** Most existing memory systems rely on rigid similarity-based retrieval, where memories are selected by a fixed encoder or surface semantic matching. However, semantic similarity does not necessarily indicate downstream utility: a memory can be topically similar but unhelpful, while a less similar memory may contain the key evidence or transferable plan (Gan et al., 2024; Dai et al., 2025; Xu et al., 2025b; Shi et al., 2023; Liu et al., 2024b). ❷ **Memory capacity should be allocated adaptively.** Different retrieved chunks play different roles: some contain direct evidence, some provide reusable procedures, some are redundant, and some are misleading. A fixed latent budget cannot distinguish these cases (Yan et al., 2024). ❸ **Retrieval and allocation should be learned from task outcomes.** Since the usefulness of a memory is only revealed after generation or interaction, the memory system should learn which memories to retrieve and how much capacity to assign based on downstream feedback rather than similarity scores alone (Asai et al., 2023; Gao et al., 2024).

To address these challenges, we propose `ElasticMem`, a memory-augmented LLM framework that treats memory as an elastic latent resource. `ElasticMem` first constructs an offline latent memory bank by encoding each memory chunk into a retrieval key and a content cache. At inference time, `ElasticMem` derives a query-conditioned retrieval state from the LoRA-adapted reasoner's hidden representation after selecting a retrieval-control token. This enables retrieval to adapt to the current query and the model's internal reasoning state, rather than relying only on rigid similarity matching. After retrieving candidate memories, `ElasticMem` uses a lightweight budget policy to assign each retrieved chunk a variable number of latent tokens. Evidence-bearing memories can receive larger latent budgets, while redundant memories are compressed or suppressed. These latent states are injected into the reasoner as soft memory tokens for generation. Crucially, `ElasticMem` does not train retrieval, allocation, and generation as isolated modules. Instead, it jointly optimizes the memory-use process with downstream task rewards through group-relative policy optimization (Shao et al., 2024; Guo et al., 2025). The reward signal supervises the retrieval-control decision, the memory-budget allocations, the latent projector, and the LoRA-adapted (Hu et al., 2022) reasoner. As a result, `ElasticMem` learns how much latent capacity to allocate to each retrieved memory based on downstream task feedback, aligning memory use with actual task utility.

We comprehensively evaluate ElasticMem on MEMORYSUITE, which spans 4 memory-intensive QA benchmarks (PersonaMem-32K, PersonaMem-128K, LoCoMo, and LongMemEval) and the embodied ALFWorld environment (Jiang et al., 2025a; Maharana et al., 2024; Wu et al., 2024; Shridhar et al., 2020). Experimental results indicate that ElasticMem consistently outperforms both text-space and latent-space memory baselines, improving the weighted average QA accuracy by 26.2% and 24.6% and the ALFWorld success rate by 66.3% and 27.2% over the strongest baseline with Qwen2.5-3B-Instruct and Qwen2.5-7B-Instruct, respectively. Furthermore, ElasticMem achieves the lowest weighted-average token consumption on ALFWorld. Controlled comparisons examine the effects of training objective, parameter capacity, and latent projection. Ablation studies further validate the key designs in ElasticMem, and qualitative analyses show that it allocates larger budgets to memories carrying useful evidence or transferable plans, even when they are not the most similar ones.

2 PRELIMINARIES

2.1 MEMORY-AUGMENTED LANGUAGE MODELS

We consider a memory-augmented language model that answers a query q with access to an external memory corpus $\mathcal{M} = \{m_i\}_{i=1}^{N}$. Each memory item m_i is a chunk of long-term context, such as a dialogue segment, a document passage, or a procedural skill card. A standard retrieval-augmented system first embeds the query and memory chunks, retrieves the top-ranked chunks, and then prepends the retrieved content to the model input:

$$\mathcal{R}(q) = \mathrm{TopK}_{m_i \in \mathcal{M}} \, \mathrm{sim}(e_q, e_i), \tag{1}$$

where e_q and e_i denote the query and memory embeddings, respectively.

This retrieval-and-concatenation paradigm provides a simple mechanism for long-term memory access, but it imposes a rigid interface between memory and generation. The system typically retrieves a fixed number of chunks, injects them as plain text, and allocates context-window capacity in proportion to their original token lengths rather than their query-specific utility. Consequently, irrelevant or redundant chunks may consume substantial context, while compact but crucial evidence may receive insufficient representational capacity.

2.2 LATENT MEMORY INTERFACES

An alternative is to represent memory using continuous latent vectors rather than natural-language text. Let $E(\cdot)$ denote the token embedding layer of an LLM with hidden dimension d. A latent or soft memory token $z \in \mathbb{R}^d$ can be inserted directly into the input embedding sequence:

$$X = \big[E(x_1), \ldots, E(x_T), z_1, \ldots, z_B\big]. \tag{2}$$

Such latent interfaces avoid re-rendering retrieved memories as text and can reduce prompt overhead. However, existing latent-memory methods often still expose a fixed-capacity interface: each retrieved chunk is assigned the same number of latent tokens regardless of its relevance, redundancy, or information density. This motivates a more flexible memory interface. Ideally, a model should adapt memory use at two levels: it should select the memory chunks that are useful for the current query, and it should allocate more latent capacity to chunks that are more informative for solving the task.

3 ElasticMem: MEMORY USE AS A LEARNABLE RESOURCE

3.1 OVERVIEW

Given a query q and memory corpus $\mathcal{M}$, ElasticMem learns which memories to retrieve and how much latent capacity to assign to each. Figure 1 summarizes the four stages: offline bank construction, query-conditioned retrieval, elastic allocation, and soft-token injection. Downstream rewards train the LoRA-adapted reasoner, budget policy, and projector; the memory bank remains frozen. Appendix E provides the training procedure.

3.2 LATENT MEMORY BANK CONSTRUCTION

We partition the memory corpus into N chunks $\mathcal{M} = \{m_i\}_{i=1}^{N}$. Each chunk is encoded once offline by a frozen base LLM encoder f_{θ_0}:

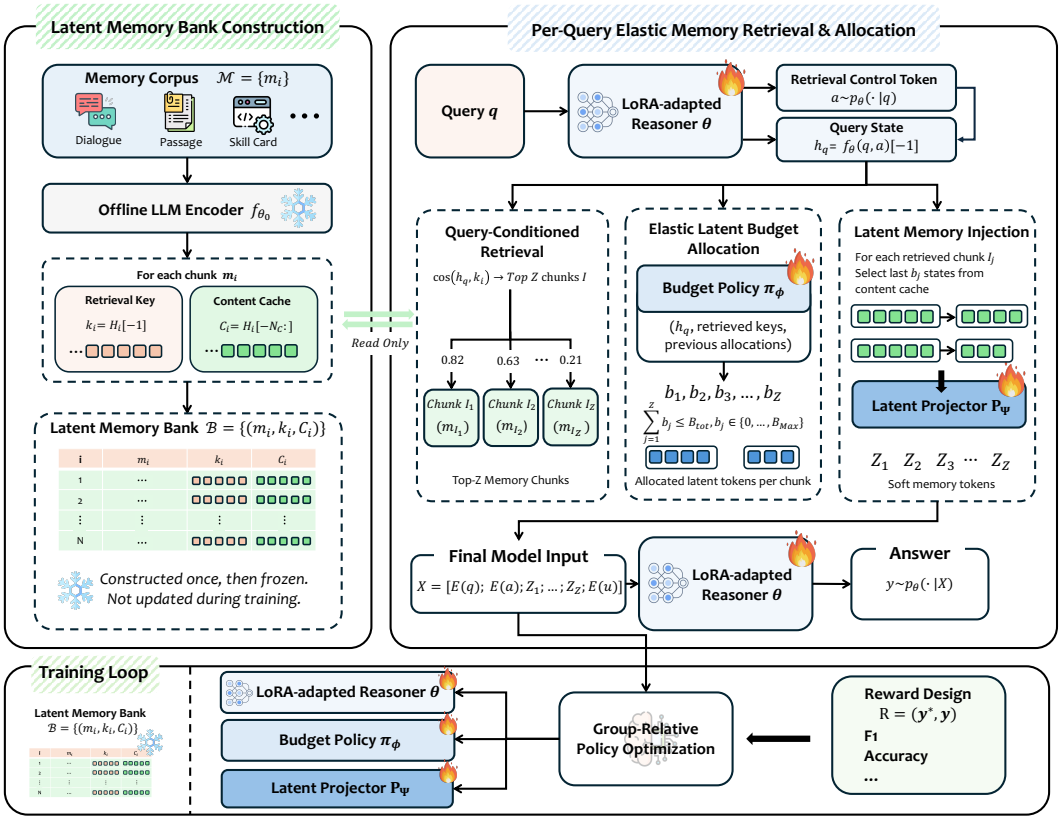


Figure 1: Overview of `ElasticMem`. **(1) Latent memory bank construction**: a frozen offline encoder stores each memory chunk as a retrieval key and a latent content cache. **(2) Elastic retrieval and allocation**: the LoRA-adapted reasoner derives a query state to retrieve top-ranked chunks, and a budget policy assigns each chunk a latent budget $b_j \in \{0, \ldots, B_{\max}\}$. **(3) Latent memory injection**: the selected cached states are projected into soft memory tokens for generation. **(4) Reward-based optimization**: GRPO trains the reasoner, budget policy, and projector with task rewards, while the memory bank stays frozen. Sampling denotes training; evaluation uses argmax decisions.

$$H_i = f_{\theta_0}(m_i) \in \mathbb{R}^{L_i \times d}, \tag{3}$$

where L_i is the length of chunk m_i and d is the hidden dimension. From the hidden states H_i, we extract two representations:

$$k_i = H_i[-1] \in \mathbb{R}^d, \tag{4}$$

$$C_i = H_i[-N_c:] \in \mathbb{R}^{N_c \times d}. \tag{5}$$

The final hidden state k_i serves as a compact retrieval key, while C_i stores the last N_c hidden states as a latent content cache. The resulting memory bank is

$$\mathcal{B} = \{(m_i, k_i, C_i)\}_{i=1}^{N}. \tag{6}$$

Since $\mathcal{B}$ is constructed offline and kept fixed, query-time memory access does not require re-encoding the full memory corpus.

3.3 Query-Conditioned Retrieval

Standard retrieval systems usually encode the query with a separate retriever. This decouples memory access from the reasoner's internal state. In contrast, `ElasticMem` derives the retrieval query from the reasoner itself, so retrieval is conditioned on how the model interprets the current input.

Given a query q, the reasoner first samples a retrieval-control token:

$$a \sim p_\theta(\cdot \mid q), \tag{7}$$

where θ denotes the trainable LoRA-adapted reasoner parameters. The action a is one ordinary token from the full vocabulary, initialized from the pretrained next-token distribution with zero-initialized

LoRA updates. Its hidden state, rather than its surface meaning, controls retrieval. The query and retrieval-control token are then passed through the reasoner, and the last hidden state is used as the query representation:

$$h_q = f_\theta(q, a)[-1] \in \mathbb{R}^d. \tag{8}$$

We score each memory chunk by cosine similarity between h_q and the cached retrieval key:

$$s_i = \cos(h_q, k_i) = \frac{h_q^\top k_i}{\|h_q\|_2 \|k_i\|_2}. \tag{9}$$

The top-K chunks are selected as

$$I = \mathrm{TopK}_{i \in [N]}\, s_i, \tag{10}$$

where $I = (I_1, \ldots, I_Z)$ denotes the retrieved chunk indices in descending retrieval-score order. In the ALFWorld evaluation, q is the fixed episode objective. With greedy control-token and budget decisions and a fixed memory bank, retrieval and allocation are task-conditioned and remain unchanged within an episode; the generator additionally conditions on the changing observation and interaction history through u.

3.4 Elastic Latent Budget Allocation

After retrieval, `ElasticMem` allocates a variable latent budget to each retrieved chunk. Let b_j denote the number of soft memory tokens assigned to the j-th retrieved chunk m_{I_j}. We impose a per-chunk maximum budget:

$$b_j \in \{0, 1, \ldots, B_{\max}\}, \qquad B_{\max} \le N_c. \tag{11}$$

The option $b_j = 0$ allows the model to suppress retrieved chunks that are similar to the query but not useful for the downstream task. Thus, the budget policy is not merely a compression module; it decides which retrieved memories deserve more representational capacity and which should contribute little or no latent signal.

Budget policy network. A lightweight Transformer jointly processes features of the Z retrieved memories: the query state, memory key, their element-wise product, cosine score, and rank embedding. For each memory, its output parameterizes a categorical distribution over $\{0, \ldots, B_{\max}\}$, from which b_j is sampled. Joint processing allows the allocator to compare retrieved memories before assigning capacity. Appendix B gives the feature construction and policy equations.

Unlike similarity-based reweighting, π_ϕ is not trained to imitate cosine scores. It is optimized only through downstream rewards, allowing it to assign larger budgets to memories that improve the final answer or action even when they are not the most similar retrieved chunks.

3.5 Latent Memory Injection

Given the allocated budget b_j for retrieved chunk m_{I_j}, `ElasticMem` selects the last b_j hidden states from its content cache:

$$\widetilde{C}_j = \begin{cases} C_{I_j}[-b_j :], & b_j > 0, \\ \emptyset_{0 \times d}, & b_j = 0. \end{cases} \tag{12}$$

If $b_j = 0$, the chunk contributes no latent memory tokens. Otherwise, the selected states are mapped into the reasoner's input embedding space through a learned projector P_ψ:

$$Z_j = P_\psi(\widetilde{C}_j) \in \mathbb{R}^{b_j \times d}. \tag{13}$$

The final input sequence is

$$X = \big[E(q); E(a); Z_1; \cdots; Z_Z; E(u)\big], \tag{14}$$

where u is the generation suffix: answer choices and output constraints for QA, or recent history, the current observation, admissible actions, and output constraints for ALFWorld. Soft tokens are inserted within the user-side input while preserving the chat-template boundaries. The reasoner generates

$$y \sim p_\theta(\cdot \mid X). \tag{15}$$

Soft memory positions produce no output tokens and receive no token-level loss. They influence generation through the reasoner's hidden states and are learned through downstream reward signals.

3.6 REWARD-BASED OPTIMIZATION

The usefulness of retrieval and budget decisions is revealed only after the model produces an answer or completes an action. We therefore optimize the full memory-use process with group-relative policy optimization. For each query-target pair $(q, y^\star)$, `ElasticMem` samples G trajectories:

$$\tau_g = (a_g, I_g, \mathbf{b}_g, y_g), \qquad g = 1, \dots, G, \tag{16}$$

where a_g is the retrieval-control token, $I_g = (I_{g,1}, \dots, I_{g,Z})$ is the retrieved memory index sequence, $\mathbf{b}_g = (b_{g,1}, \dots, b_{g,Z})$ is the sampled budget allocation, and y_g is the generated output. Each trajectory receives a task reward:

$$r_g = R(y_g, y^\star), \tag{17}$$

where R is exact-match accuracy for QA. For ALFWorld training, it is binary agreement between the snapped generated action and the expert action; environment success is the evaluation metric.

We compute a group-relative advantage by normalizing rewards within the sampled group:

$$A_g = \frac{r_g - \text{mean}\left(\{r_{g'}\}_{g'=1}^{G}\right)}{\max\left(\text{std}\left(\{r_{g'}\}_{g'=1}^{G}\right), \sigma_{\min}\right)}, \tag{18}$$

where $\sigma_{\min}$ is a small constant for numerical stability.

We minimize the clipped GRPO objective:

$$\mathcal{L}_{\text{GRPO}} = -\frac{1}{G}\sum_{g=1}^{G}\left[\min\left(\rho_g A_g,\ \text{clip}\left(\rho_g,\, 1-\epsilon,\, 1+\epsilon\right) A_g\right)\right], \tag{19}$$

where the probability ratio is

$$\rho_g = \frac{p_{\theta,\phi,\psi}(\tau_g)}{p_{\theta_{\text{old}},\phi_{\text{old}},\psi_{\text{old}}}(\tau_g)}. \tag{20}$$

Including the deterministic retrieval decision, the trajectory likelihood factorizes as

$$p_{\theta,\phi,\psi}(\tau_g) = p_\theta(a_g \mid q)\, \mathbf{1}\{I_g = T_\theta(q, a_g)\} \left[\prod_{j=1}^{Z} \pi_\phi(b_{g,j} \mid h_{q,g},\, I_g)\right] p_\theta(y_g \mid X_g(\psi)), \tag{21}$$

where $T_\theta(q,a) = \text{TopZ}_i \cos(f_\theta(q,a)[-1], k_i)$, $h_{q,g} = f_\theta(q,a_g)[-1]$, and $X_g(\psi) = \left[E(q);\, E(a_g);\, P_\psi(\widetilde{C}_{g,1}); \cdots;\, P_\psi(\widetilde{C}_{g,Z});\, E(u)\right]$ is the final input sequence whose soft memory tokens are produced by the latent projector P_ψ, with $\widetilde{C}_{g,j}$ selected from cache $C_{I_{g,j}}$ using budget $b_{g,j}$ as in Eq. 12. The indicator accounts for deterministic retrieval in the joint likelihood; a likelihood conditional on fixed rollout indices is a different surrogate. Hard top-Z selection is non-differentiable. GRPO reinforces sampled control tokens whose resulting states retrieve useful memories; it does not differentiate through the selected indices. The budget and generation likelihoods train allocation and latent injection. Appendix A.7 provides a same-checkpoint intervention. The trainable components are the LoRA-adapted reasoner θ, the budget policy π_ϕ, and the latent projector P_ψ. The offline memory bank $\mathcal{B}$ remains fixed during training. The detailed training procedure is provided in Appendix E.

4 EXPERIMENTS

To evaluate the effectiveness of `ElasticMem` in memory-augmented language modeling, we conduct experiments on MEMORYSUITE, which contains two complementary settings: memory-intensive question answering and embodied agentic decision-making (see Appendix F).

Tasks and Metrics. MEMORYSUITE-QA includes PersonaMem-32K (Jiang et al., 2025a), PersonaMem-128K (Jiang et al., 2025a), LoCoMo (Maharana et al., 2024), and LongMemEval (Wu et al., 2024). We report accuracy as the primary metric. MEMORYSUITE-AGENTIC uses ALFWorld (Shridhar et al., 2020) and evaluates both seen and unseen splits, using success rate (SR) as the primary metric. We also report token consumption (#Tok.); this metric does not measure wall-clock latency or FLOPs.

Table 2: **Main comparison results on PersonaMem-32K, PersonaMem-128K, LoCoMo, LongMemEval, and ALFWorld.** For Acc. and SR, higher is better; for #Tok., lower is better. #Tok. counts task-text input, injected memory, and generated output, distinct from pure latent-memory budget (Appendix D). **Bold** and underline denote the best and second-best results. Task-performance scores are means over five independent training seeds. Avg. columns are weighted by test-set size. Appendix A.2 summarizes the `ElasticMem`–MemGen means.

Model	Methods	Memory QA Tasks										Agentic Embodied Interactive Tasks					
		PersonaMem-32K		PersonaMem-128K		LoCoMo		LongMemEval		Avg.		ALF-Seen		ALF-Unseen		Avg.	
		Acc.	#Tok.	Acc.	#Tok.	Acc.	#Tok.	Acc.	#Tok.	Acc.	#Tok.	SR	#Tok.	SR	#Tok.	SR	#Tok.
Qwen2.5-3B -Instruct	*Memory-based Methods in Text-Space*																
	MemoryBank	0.50	2,473	0.33	4,618	0.41	2,202	0.29	9,591	0.37	4,089	0.06	76,016	0.09	71,173	0.08	73,648
	A-MEM	0.48	4,899	0.38	3,530	0.39	3,348	0.45	15,625	0.40	5,282	0.04	17,239	0.05	14,248	0.04	15,776
	LightMem	0.46	851	0.35	1,306	0.40	3,468	0.44	1,248	0.39	2,239	0.14	13,252	0.17	12,617	0.15	12,941
	Mem0	0.42	778	0.39	662	0.33	836	0.33	2,843	0.36	1,062	0.14	43,055	0.12	43,677	0.13	43,359
	MemoryOS	0.48	4,419	0.32	9,440	0.33	4,420	0.40	13,854	0.35	7,452	0.11	68,131	0.13	65,845	0.12	67,013
	MemP	0.42	12,002	0.38	29,567	0.26	5,140	0.41	26,571	0.33	16,873	0.11	101,580	0.10	110,019	0.11	105,707
	LangMem	0.40	675	0.30	1,626	0.40	936	0.28	687	0.35	1,112	0.20	12,276	0.11	**12,449**	0.15	12,361
	Memory-based Methods in Latent-Space																
	MemGen	0.62	502	0.58	393	0.60	198	0.55	242	0.59	291	0.29	14,708	0.25	14,416	0.27	14,565
	AutoCompressor	0.48	717	0.45	1,127	0.47	234	0.44	998	0.46	677	0.16	21,145	0.13	21,571	0.14	21,353
	M+	0.56	**447**	0.52	**350**	0.54	**154**	0.50	**198**	0.53	**247**	0.20	14,246	0.16	14,233	0.18	14,240
	`ElasticMem`	**0.74**	537	**0.77**	682	**0.74**	577	**0.68**	433	**0.74**	589	**0.49**	**11,028**	**0.41**	12,846	**0.45**	**11,917**
Qwen2.5-7B -Instruct	*Memory-based Methods in Text-Space*																
	MemoryBank	0.53	2,639	0.35	4,724	0.40	2,301	0.39	10,822	0.39	4,358	0.11	36,602	0.19	39,385	0.15	37,963
	A-MEM	0.52	9,731	0.37	11,098	0.42	3,289	0.50	31,489	0.42	10,408	0.06	29,543	0.09	29,531	0.07	29,537
	LightMem	0.52	864	0.33	1,228	0.41	3,702	0.53	939	0.41	2,275	0.33	11,150	0.37	10,617	0.35	10,889
	Mem0	0.54	794	0.39	676	0.38	921	0.48	3,202	0.41	1,157	0.43	26,165	0.40	27,327	0.42	26,733
	MemoryOS	0.52	4,695	0.38	9,748	0.42	4,998	0.53	14,489	0.43	7,926	0.38	60,129	0.40	50,238	0.39	55,292
	MemP	0.52	11,964	0.40	29,581	0.22	5,382	0.48	26,580	0.34	16,985	0.24	143,768	0.22	115,492	0.23	129,940
	LangMem	0.68	1,359	0.42	3,097	0.35	1,478	0.42	781	0.41	1,911	0.40	9,735	0.28	10,140	0.34	9,933
	Memory-based Methods in Latent-Space																
	MemGen	0.70	502	0.66	393	0.68	198	0.63	242	0.67	291	0.39	12,933	0.34	13,551	0.36	13,235
	AutoCompressor	0.56	717	0.52	1,127	0.55	234	0.51	998	0.54	677	0.24	19,875	0.20	19,317	0.22	19,602
	M+	0.64	**447**	0.59	**350**	0.62	**154**	0.57	**198**	0.60	**247**	0.32	12,520	0.28	13,382	0.30	12,942
	`ElasticMem`	**0.90**	483	**0.84**	557	**0.84**	362	**0.76**	410	**0.83**	443	**0.56**	**9,035**	**0.49**	**9,031**	**0.53**	**9,033**

Baselines and Settings. We compare `ElasticMem` with representative memory-based baselines from two paradigms. **Text-space memory baselines** include MemoryBank (Zhong et al., 2024), A-MEM (Xu et al., 2025a), LightMem (Fang et al., 2025a), Mem0 (Chhikara et al., 2025), MemoryOS (Kang et al., 2025), MemP (Fang et al., 2025b), and LangMem (LangChain, 2024). **Latent-space memory baselines** include MemGen (Zhang et al., 2025a), AutoCompressor (Chevalier et al., 2023), and M+ (Wang et al., 2025). All methods are evaluated with Qwen2.5-3B-Instruct[1] and Qwen2.5-7B-Instruct[2] backbones under a no-few-shot-example setting.

Comparison Protocol. All latent-space baselines are retrained on each task's training split with the same task supervision, backbone scale, source information, and evaluation protocol, while retaining native objectives and trainable modules. QA banks contain only each example's raw history; held-out answers and attempts never enter memory. ALFWorld latent methods share Gemini-generated training skill cards; text methods receive the same segmented raw trajectories and construct memories natively. All methods share task instructions and chat templates; their memory interfaces differ (Appendix H). Table 3 additionally examines the effects of training objective and increased LoRA capacity; Appendix A details the comparison scope.

4.1 `ElasticMem` Outperforms Memory-based Methods in Text-Space and Latent-Space

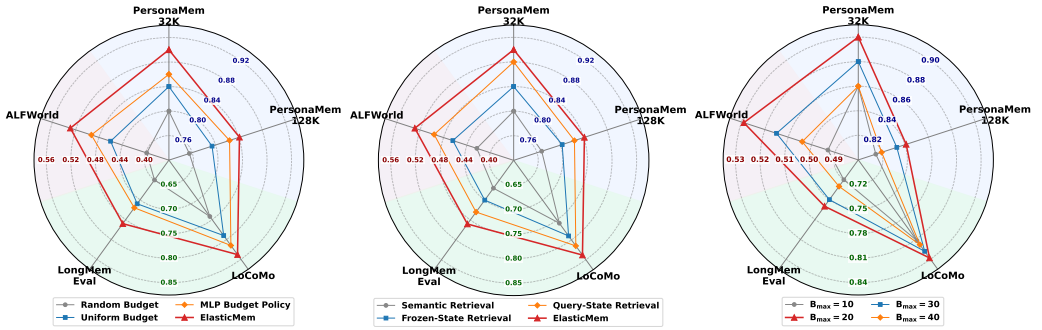
The results in Table 2 show that `ElasticMem` achieves the highest task-performance scores across the evaluated datasets and backbone scales. Weighted QA accuracy increases from 0.588 to 0.742 with Qwen2.5-3B and from 0.668 to 0.832 with Qwen2.5-7B; weighted ALFWorld SR increases from 0.270 to 0.449 and from 0.416 to 0.529, respectively, relative to the strongest baseline in each aggregate setting. The reported means favor `ElasticMem` over MemGen on all six splits with both backbones.

Performance and Token Consumption. On QA, `ElasticMem` uses fewer tokens than the text-space baselines, but more than MemGen and M+: its weighted average is 423 tokens, compared with

[1] `https://huggingface.co/Qwen/Qwen2.5-3B-Instruct`
[2] `https://huggingface.co/Qwen/Qwen2.5-7B-Instruct`

Table 3: **Controlled comparisons with Qwen2.5-7B-Instruct.** Variants share data splits, source memory information, retrieval count, training examples, and update count. ALFWorld SR is weighted over Seen and Unseen episodes. The higher-rank text variant tests additional LoRA capacity; exact parameter matching is not established.

Variant	Objective	LoCoMo Acc.	ALFWorld SR
LoRA + text memory	SFT	0.599	0.380
LoRA + text memory	GRPO	0.650	0.420
Higher-rank LoRA + text memory	GRPO	0.682	0.449
`ElasticMem` with identity projector	GRPO	0.739	0.460
Full `ElasticMem`	GRPO	**0.841**	**0.529**



(a) (b) (c)

Figure 2: **Ablation studies on Qwen2.5-7B-Instruct.** (a) *Effect of budget policy*: `ElasticMem` outperforms Random Budget, Uniform Budget, and MLP Budget Policy, showing that the Transformer budget policy better learns how to allocate latent capacity across retrieved memories. (b) *Effect of retrieval design*: `ElasticMem` outperforms Semantic Retrieval, Frozen-State Retrieval, and Query-State Retrieval, demonstrating the benefit of using a trainable reasoner-state retrieval mechanism with a sampled retrieval-control token. (c) *Maximum budget*: $N_c = B_{\max}$ in every setting; $B_{\max} = 20$ performs best among the tested configurations.

291 and 247, respectively. The QA advantage is therefore higher accuracy at a larger token budget than these two latent baselines. On ALFWorld, `ElasticMem` combines the highest weighted SR with the lowest weighted token consumption on both backbones. Online latency, including native retrieval and model computation, is lowest for `ElasticMem` among the evaluated latent methods (Appendix A.5). ALFWorld per-episode latency also depends on trajectory length and does not isolate per-step cost.

4.2 Controlled Comparisons of Training and Model Capacity

Within the text-memory configuration, replacing SFT with GRPO improves LoCoMo accuracy from 0.599 to 0.650 and ALFWorld SR from 0.380 to 0.420. Increasing the text variant's LoRA capacity raises these scores to 0.682 and 0.449, still below 0.841 and 0.529 for full `ElasticMem`. These comparisons suggest that the advantage is not explained solely by applying GRPO or increasing trainable capacity within the tested text-memory configuration; they do not isolate elastic allocation from all other interface differences. Within `ElasticMem`, replacing the learned projector with direct latent-state injection reduces the scores to 0.739 and 0.460, supporting the value of a learned mapping between cached states and input embeddings.

4.3 Retrieval and Allocation Ablations

Budget Policy and Retrieval. Figure 2(a) compares random, uniform, MLP, and Transformer allocation policies. The Transformer performs best across the five settings, supporting allocation that conditions jointly on the retrieved set. Figure 2(b) compares semantic retrieval, frozen reasoner states, trainable query-only states, and the control-token state. The control-token variant performs best, supporting the use of a learned query-conditioned retrieval representation.

Maximum Latent Budget. Among the tested per-chunk limits $B_{\max} \in \{10, 20, 30, 40\}$, Figure 2(c) favors $B_{\max} = 20$. This is an empirical result under the evaluated training settings: enlarging the allowable allocation space need not improve the policy learned with a finite training budget. An inference-time clipping intervention is consistent with over-allocation contributing to the degradation (Appendix A.6); sensitivity to Z and G is reported in Appendix A.4. Qualitative examples of nonuniform memory allocation are discussed in Section 4.4.

Generalization Beyond Task-Specific Qwen Models. With learned components frozen, checkpoints trained on PersonaMem-128K or ALFWorld outperform the corresponding MemGen checkpoints on LoCoMo and LongMemEval by 7.3–8.9 accuracy points; only the target memory bank is reconstructed. Separate evaluations with Llama-3.2-3B and Llama-3.1-8B also favor `ElasticMem` on the two PersonaMem datasets. The Llama experiments use one run per configuration and demonstrate applicability to another model family, not transfer of the same trained parameters across backbones (Appendix A.3).

4.4 Case Studies of Memory Allocation

Evidence selection in PersonaMem-32K. A question about a summer mentoring program asks how meditation helped with studying; the answer links meditation to improved focus and reduced anxiety. `ElasticMem` assigns low budgets to highly ranked but only topically related memories about group projects, peer study, or a restatement of the query. It assigns larger budgets to lower-ranked memories connecting yoga and meditation with stress reduction, and anxiety and sleep deprivation with cognition. The resulting answer connects meditation, focus, anxiety, and study performance. This allocation is consistent with prioritizing useful evidence over surface similarity.

Plan transfer in ALFWorld. For *"heat some apple and put it on the countertop"*, `ElasticMem` assigns a high budget to a card from a `cool` task whose plan places an object on a `shelf`, an open surface resembling the target countertop. Cards involving `coffeemachine` or `stoveburner` destinations receive lower budgets. Even cards sharing the objective *"heat some mug and put it in coffeemachine"* receive different budgets when their plans search open surfaces versus cabinets and appliances. These observations suggest sensitivity to procedural structure beyond task labels; they do not establish that any individual memory caused success.

5 Additional Related Work

LLM Agent Memory. Text-space memory systems store, organize, and inject natural-language records into the context, including MemoryBank (Zhong et al., 2024), Mem0 (Chhikara et al., 2025), MemoryOS (Kang et al., 2025), LangMem (LangChain, 2024), and experience-oriented frameworks such as A-MEM (Xu et al., 2025a), LightMem (Fang et al., 2025a), and MemP (Fang et al., 2025b). Their textual memories consume context tokens. Latent-space methods such as AutoCompressor (Chevalier et al., 2023), MemGen (Zhang et al., 2025a), and M+ (Wang et al., 2025) provide latent alternatives to textual context. In contrast, `ElasticMem` retrieves memories from the reasoner's state and learns how many latent tokens each memory deserves.

Reward-Based LLM Optimization. PPO-style methods (Ouyang et al., 2022; Liu et al., 2024a), preference optimization (Rafailov et al., 2023), process supervision (Lightman et al., 2023), and group-relative objectives (Shao et al., 2024; Liu et al., 2025; Zheng et al., 2025; Cui et al., 2025) improve model behavior through feedback. We apply task rewards to memory retrieval, allocation, and generation with a frozen external bank.

6 Conclusion

We introduce `ElasticMem`, a memory-augmented LLM framework that treats memory as an elastic latent resource, learning from downstream task rewards which memories to retrieve and how much latent capacity each deserves. Extensive experiments on MemorySuite demonstrate that `ElasticMem` substantially outperforms both text-space and latent-space memory baselines on memory-intensive QA and embodied decision-making, while achieving the lowest weighted-average token consumption on ALFWorld.

AI USE DISCLOSURE

We used generative AI tools solely to polish the writing of this paper. The authors reviewed and edited all AI-assisted text and take full responsibility for the content.

REFERENCES

Akari Asai, Zeqiu Wu, Yizhong Wang, Avirup Sil, and Hannaneh Hajishirzi. Self-rag: Learning to retrieve, generate, and critique through self-reflection. In *The Twelfth International Conference on Learning Representations*, 2023.

Yixing Chen, Yiding Wang, Siqi Zhu, Haofei Yu, Tao Feng, Muhan Zhang, Mostofa Patwary, and Jiaxuan You. Multi-agent evolve: Llm self-improve through co-evolution. *arXiv preprint arXiv:2510.23595*, 2025.

Alexis Chevalier, Alexander Wettig, Anirudh Ajith, and Danqi Chen. Adapting language models to compress contexts. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 3829–3846, 2023.

Prateek Chhikara, Dev Khant, Saket Aryan, Taranjeet Singh, and Deshraj Yadav. Mem0: Building production-ready ai agents with scalable long-term memory. *arXiv preprint arXiv:2504.19413*, 2025.

Ganqu Cui, Yuchen Zhang, Jiacheng Chen, Lifan Yuan, Zhi Wang, Yuxin Zuo, Haozhan Li, Yuchen Fan, Huayu Chen, Weize Chen, et al. The entropy mechanism of reinforcement learning for reasoning language models. *arXiv preprint arXiv:2505.22617*, 2025.

Lu Dai, Yijie Xu, Jinhui Ye, Hao Liu, and Hui Xiong. Seper: Measure retrieval utility through the lens of semantic perplexity reduction. *arXiv preprint arXiv:2503.01478*, 2025.

Yiming Du, Baojun Wang, Yifan Xiang, Zhaowei Wang, Wenyu Huang, Boyang Xue, Bin Liang, Xingshan Zeng, Fei Mi, Haoli Bai, et al. Memory-t1: Reinforcement learning for temporal reasoning in multi-session agents. *arXiv preprint arXiv:2512.20092*, 2025.

Jizhan Fang, Xinle Deng, Haoming Xu, Ziyan Jiang, Yuqi Tang, Ziwen Xu, Shumin Deng, Yunzhi Yao, Mengru Wang, Shuofei Qiao, et al. Lightmem: Lightweight and efficient memory-augmented generation. *arXiv preprint arXiv:2510.18866*, 2025a.

Runnan Fang, Yuan Liang, Xiaobin Wang, Jialong Wu, Shuofei Qiao, Pengjun Xie, Fei Huang, Huajun Chen, and Ningyu Zhang. Memp: Exploring agent procedural memory. *arXiv preprint arXiv:2508.06433*, 2025b.

Chunjing Gan, Dan Yang, Binbin Hu, Hanxiao Zhang, Siyuan Li, Ziqi Liu, Yue Shen, Lin Ju, Zhiqiang Zhang, Jinjie Gu, et al. Similarity is not all you need: Endowing retrieval augmented generation with multi layered thoughts. *arXiv preprint arXiv:2405.19893*, 2024.

Jingsheng Gao, Linxu Li, Weiyuan Li, Yuzhuo Fu, and Bin Dai. Smartrag: Jointly learn rag-related tasks from the environment feedback. *arXiv preprint arXiv:2410.18141*, 2024.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, et al. Lora: Low-rank adaptation of large language models. *ICLR*, 1(2):3, 2022.

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. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 32779–32798, 2025.

Peter Jansen, Marc-Alexandre Côté, Tushar Khot, Erin Bransom, Bhavana Dalvi Mishra, Bodhisattwa Prasad Majumder, Oyvind Tafjord, and Peter Clark. Discoveryworld: A virtual environment for developing and evaluating automated scientific discovery agents. *Advances in Neural Information Processing Systems*, 37:10088–10116, 2024.

Bowen Jiang, Yuan Yuan, Maohao Shen, Zhuoqun Hao, Zhangchen Xu, Zichen Chen, Ziyi Liu, Anvesh Rao Vijjini, Jiashu He, Hanchao Yu, et al. Personamem-v2: Towards personalized intelligence via learning implicit user personas and agentic memory. *arXiv preprint arXiv:2512.06688*, 2025a.

Pengcheng Jiang, Jiacheng Lin, Zhiyi Shi, Zifeng Wang, Luxi He, Yichen Wu, Ming Zhong, Peiyang Song, Qizheng Zhang, Heng Wang, et al. Adaptation of agentic ai. *arXiv preprint arXiv:2512.16301*, 2025b.

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. 25972–25981, 2025.

Vladimir Karpukhin, Barlas Oguz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih. Dense passage retrieval for open-domain question answering. In *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)*, pp. 6769–6781, 2020.

LangChain. LangMem. `https://langchain-ai.github.io/langmem/`, 2024.

Hunter Lightman, Vineet Kosaraju, Yuri Burda, Harrison Edwards, Bowen Baker, Teddy Lee, Jan Leike, John Schulman, Ilya Sutskever, and Karl Cobbe. Let's verify step by step. In *The Twelfth International Conference on Learning Representations*, 2023.

Aixin Liu, Bei Feng, Bing Xue, Bingxuan Wang, Bochao Wu, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan, et al. Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*, 2024a.

Nelson F Liu, Kevin Lin, John Hewitt, Ashwin Paranjape, Michele Bevilacqua, Fabio Petroni, and Percy Liang. Lost in the middle: How language models use long contexts. *Transactions of the association for computational linguistics*, 12:157–173, 2024b.

Zichen Liu, Changyu Chen, Wenjun Li, Penghui Qi, Tianyu Pang, Chao Du, Wee Sun Lee, and Min Lin. Understanding r1-zero-like training: A critical perspective. *arXiv preprint arXiv:2503.20783*, 2025.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

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, 2024.

Qirui Mi, Zhijian Ma, Mengyue Yang, Haoxuan Li, Yisen Wang, Haifeng Zhang, and Jun Wang. Procmem: Learning reusable procedural memory from experience via non-parametric ppo for llm agents. *arXiv preprint arXiv:2602.01869*, 2026.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, 35:27730–27744, 2022.

Zhuoshi Pan, Qianhui Wu, Huiqiang Jiang, Xufang Luo, Hao Cheng, Dongsheng Li, Yuqing Yang, Chin-Yew Lin, H Vicky Zhao, Lili Qiu, et al. On memory construction and retrieval for personalized conversational agents. *arXiv preprint arXiv:2502.05589*, 2025.

Joon Sung Park, Joseph O'Brien, Carrie Jun 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.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. *Advances in neural information processing systems*, 36:53728–53741, 2023.

Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*, pp. 3982–3992, 2019.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Yang Wu, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

Freda Shi, Xinyun Chen, Kanishka Misra, Nathan Scales, David Dohan, Ed H Chi, Nathanael Schärli, and Denny Zhou. Large language models can be easily distracted by irrelevant context. In *International Conference on Machine Learning*, pp. 31210–31227. PMLR, 2023.

Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. *Advances in Neural Information Processing Systems*, 36:8634–8652, 2023.

Mohit Shridhar, Xingdi Yuan, Marc-Alexandre Côté, Yonatan Bisk, Adam Trischler, and Matthew Hausknecht. Alfworld: Aligning text and embodied environments for interactive learning. *arXiv preprint arXiv:2010.03768*, 2020.

Haoran Sun, Zekun Zhang, and Shaoning Zeng. Preference-aware memory update for long-term llm agents. *arXiv preprint arXiv:2510.09720*, 2025.

Mohammad Tavakoli, Alireza Salemi, Carrie Ye, Mohamed Abdalla, Hamed Zamani, and J Ross Mitchell. Beyond a million tokens: Benchmarking and enhancing long-term memory in llms. *arXiv preprint arXiv:2510.27246*, 2025.

Guanzhi Wang, Yuqi Xie, Yunfan Jiang, Ajay Mandlekar, Chaowei Xiao, Yuke Zhu, Linxi Fan, and Anima Anandkumar. Voyager: An open-ended embodied agent with large language models. *arXiv preprint arXiv:2305.16291*, 2023.

Yu Wang, Dmitry Krotov, Yuanzhe Hu, Yifan Gao, Wangchunshu Zhou, Julian McAuley, Dan Gutfreund, Rogerio Feris, and Zexue He. M+: Extending memoryllm with scalable long-term memory. *arXiv preprint arXiv:2502.00592*, 2025.

Jason Wei, Zhiqing Sun, Spencer Papay, Scott McKinney, Jeffrey Han, Isa Fulford, Hyung Won Chung, Alex Tachard Passos, William Fedus, and Amelia Glaese. Browsecomp: A simple yet challenging benchmark for browsing agents, 2025. URL `https://arxiv.org/abs/2504.12516`.

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.

Wujiang Xu, Zujie Liang, Kai Mei, Hang Gao, Juntao Tan, and Yongfeng Zhang. A-mem: Agentic memory for llm agents. *arXiv preprint arXiv:2502.12110*, 2025a.

Yilong Xu, Jinhua Gao, Xiaoming Yu, Yuanhai Xue, Baolong Bi, Huawei Shen, and Xueqi Cheng. Training a utility-based retriever through shared context attribution for retrieval-augmented language models. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 629–648, 2025b.

Shi-Qi Yan, Jia-Chen Gu, Yun Zhu, and Zhen-Hua Ling. Corrective retrieval augmented generation. 2024.

Sikuan Yan, Xiufeng Yang, Zuchao Huang, Ercong Nie, Zifeng Ding, Zonggen Li, Xiaowen Ma, Jinhe Bi, Kristian Kersting, Jeff Z Pan, et al. Memory-r1: Enhancing large language model agents to manage and utilize memories via reinforcement learning. *arXiv preprint arXiv:2508.19828*, 2025.

John Yang, Carlos E Jimenez, Alex L Zhang, Kilian Lieret, Joyce Yang, Xindi Wu, Ori Press, Niklas Muennighoff, Gabriel Synnaeve, Karthik R Narasimhan, et al. Swe-bench multimodal: Do ai systems generalize to visual software domains? *arXiv preprint arXiv:2410.03859*, 2024.

Guibin Zhang, Muxin Fu, and Shuicheng Yan. Memgen: Weaving generative latent memory for self-evolving agents. *arXiv preprint arXiv:2509.24704*, 2025a.

Wenqi Zhang, Ke Tang, Hai Wu, Mengna Wang, Yongliang Shen, Guiyang Hou, Zeqi Tan, Peng Li, Yueting Zhuang, and Weiming Lu. Agent-pro: Learning to evolve via policy-level reflection and optimization. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 5348–5375, 2024.

Yingyi Zhang, Junyi Li, Wenlin Zhang, Penyue Jia, Xianneng Li, Yichao Wang, Derong Xu, Yi Wen, Huifeng Guo, Yong Liu, et al. Evoking user memory: Personalizing llm via recollection-familiarity adaptive retrieval. *arXiv preprint arXiv:2603.09250*, 2026.

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(6):1–47, 2025b.

Chujie Zheng, Shixuan Liu, Mingze Li, Xiong-Hui Chen, Bowen Yu, Chang Gao, Kai Dang, Yuqiong Liu, Rui Men, An Yang, et al. Group sequence policy optimization. *arXiv preprint arXiv:2507.18071*, 2025.

Longtao Zheng, Rundong Wang, Xinrun Wang, and Bo An. Synapse: Trajectory-as-exemplar prompting with memory for computer control. *arXiv preprint arXiv:2306.07863*, 2023.

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.

Contents of Appendix

A Comparison Protocol and Additional Experimental Results **15**

A.1 Scope of the Comparisons 15

A.2 Five-Seed Comparison with MemGen 16

A.3 Cross-Task Transfer and Additional Backbones 16

A.4 Sensitivity to Retrieval Count and GRPO Group Size 16

A.5 Online Latency and Offline Preprocessing 18

A.6 Maximum-Budget Diagnostics 18

A.7 Control-Token Intervention 19

B Budget Policy Architecture **20**

C Future Work and Broader Impact **21**

D Implementation Details **21**

E Training Procedure of `ElasticMem` **22**

F Dataset Descriptions **23**

F.1 MemorySuite-QA 24

F.2 MemorySuite-Agentic 25

G Dataset Statistics **26**

H Prompt Usage **26**

H.1 Memory Construction Prompts 26

H.2 PersonaMem Answering Prompt 32

H.3 LoCoMo Answering Prompt 32

H.4 LongMemEval Answering Prompt 32

H.5 ALFWorld Answering Prompt 33

I Case Studies **34**

I.1 PersonaMem-32K 34

I.2 PersonaMem-128K 38

I.3 LoCoMo 42

I.4 LongMemEval 46

I.5 ALFWorld 49

A Comparison Protocol and Additional Experimental Results

A.1 Scope of the Comparisons

The main comparison evaluates complete memory systems with the same backbone scale, task partitions, source memory information, and evaluation protocol. MemGen, AutoCompressor, and M+ are retrained separately for each task while retaining their native objectives and memory mechanisms. The shared source information consists of benchmark histories for QA and training expert trajectories for ALFWorld; each method's memory representation and construction procedure remain part of the evaluated system. For ElasticMem and the retrained latent-space baselines, QA answer labels and ALFWorld expert demonstrations provide shared task supervision. Their native training objectives determine how this supervision is used; the comparison does not impose an identical loss, architecture, or total training computation on all methods. Consequently, the main comparison establishes system-level performance under these settings and does not isolate the contribution of any single component.

Shared information and native training. For QA, all methods use the same benchmark histories and training QA partitions; held-out questions, answers, and attempts are excluded from memory construction. For ALFWorld, source trajectories, trajectory-level segmentation, memory-bank size, construction backbone, and top-Z count are controlled. ElasticMem and the latent-space baselines receive the same Gemini-generated skill cards and boundaries. Text-space methods apply their native construction procedures to the same segmented raw trajectories. MemGen retains its reasoner, memory weaver, and learned trigger with RL training; AutoCompressor retains its backbone-scale adaptation for summary representations; M+ retains its native learned memory/retrieval components. ElasticMem trains one LoRA-adapted reasoner, the budget policy, and the latent projector while keeping the offline encoder fixed. These are native architectural budgets, not a claim that the baseline objectives or total training FLOPs are identical.

Table 4: **Scope of the experimental controls.** The controlled variants complement the complete-system comparison. Equal update counts do not imply equal training FLOPs.

Comparison	Shared conditions	Interpretation
Main table	Backbone scale, task partitions, source information, evaluation protocol; shared task supervision for latent-space methods	Complete systems with native memory mechanisms and objectives; not a matched-compute component ablation.
Text-memory SFT vs. GRPO	Architecture, retrieved text evidence, training examples, update count	Effect of the training objective within the text-memory configuration.
Higher-rank text vs. full `ElasticMem`	Backbone, data, retrieval count, GRPO objective, update count	Increased text-memory capacity across different memory interfaces; does not isolate allocation alone.
Identity vs. learned projector	Retrieval and allocation pipeline, LoRA adaptation, GRPO objective, update count	Effect of replacing the learned projector with direct latent-state injection.

The higher-rank text-memory variant increases LoRA capacity. Without the exact rank and a verified parameter inventory, this comparison is not treated as a strict parameter-matched control. For the identity-projector variant, selected states are injected directly as $Z_j = \widetilde{C}_j$; retrieval, allocation, and reasoner adaptation are retained. Table 3 reports these comparisons with Qwen2.5-7B-Instruct on LoCoMo and ALFWorld.

Aggregation weights. The QA average uses weights proportional to the test counts $(50, 233, 314, 100)$ for PersonaMem-32K, PersonaMem-128K, LoCoMo, and LongMemEval, respectively, totaling 697 examples. The corresponding percentages are approximately $(7.17, 33.43, 45.05, 14.35)\%$. ALFWorld averages use 140 Seen and 134 Unseen episodes, with weights $140/274$ and $134/274$ (approximately 51.09% and 48.91%).

A.2 Five-Seed Comparison with MemGen

The main-table experiments are repeated with five independent training seeds. Table 5 gives the detailed `ElasticMem`–MemGen comparison with fixed data splits, memory banks, hyperparameters, and evaluation protocol. The table reports run means without uncertainty estimates or significance tests. Run-level statistical analysis remains a limitation; differences in means alone do not establish significance. In particular, MemGen is not the strongest 7B ALFWorld baseline in Table 2.

Table 5: **Five-seed performance comparison.** Reported means across five training seeds; uncertainty estimates are not shown. QA uses accuracy; ALFWorld uses success rate.

	Qwen2.5-3B		Qwen2.5-7B	
Dataset	MemGen	`ElasticMem`	MemGen	`ElasticMem`
PersonaMem-32K	0.62	0.74	0.70	0.90
PersonaMem-128K	0.58	0.77	0.66	0.84
LoCoMo	0.60	0.74	0.68	0.84
LongMemEval	0.55	0.68	0.63	0.76
ALFWorld-Seen	0.29	0.49	0.39	0.56
ALFWorld-Unseen	0.25	0.41	0.34	0.49

A.3 Cross-Task Transfer and Additional Backbones

Frozen cross-task transfer. We train Qwen2.5-7B-Instruct checkpoints on PersonaMem-128K or ALFWorld and evaluate them on LoCoMo and LongMemEval without target-task parameter updates. All learned components remain frozen; only the target-domain memory bank is reconstructed. Table 6 separates these transfer settings from target-specific training. Across the four transfer comparisons, `ElasticMem` improves on MemGen by 7.3–8.9 absolute accuracy points, while target-specific training achieves higher performance for both methods.

Table 6: **Frozen cross-task transfer with Qwen2.5-7B-Instruct.** Target-specific training is a separate reference setting and uses target-task training data.

Training source	Method	LoCoMo Acc.	LongMemEval Acc.
PersonaMem-128K	MemGen	0.637	0.580
PersonaMem-128K	`ElasticMem`	0.726	0.660
ALFWorld	MemGen	0.548	0.470
ALFWorld	`ElasticMem`	0.621	0.550
Target-specific training	MemGen	0.682	0.630
Target-specific training	`ElasticMem`	0.841	0.760

Additional model family. Table 7 evaluates Llama-3.2-3B-Instruct and Llama-3.1-8B-Instruct on PersonaMem-32K and PersonaMem-128K with the same dataset splits, raw histories, and evaluation protocol across methods. Each method retains its native memory construction, retrieval, and injection procedures. These results use one run per configuration; no across-seed uncertainty is claimed. They evaluate applicability to another model family, not transfer of a fixed Qwen checkpoint to Llama. `ElasticMem` exceeds MemGen by 10.0 and 13.3 accuracy points with the 3B backbone, and 14.0 and 12.9 points with the 8B backbone.

A.4 Sensitivity to Retrieval Count and GRPO Group Size

We vary retrieval count Z and GRPO group size G one at a time using Qwen2.5-7B-Instruct, with cache length $N_c = 20$ and the remaining settings fixed. Table 8 shows that the default retrieval counts perform best among the tested values: $Z = 20$ for LoCoMo and $Z = 10$ for ALFWorld. Increasing G from 4 to 8 gives small additional gains (0.003 accuracy and 0.004 SR), while doubling the number of sampled trajectories per query. These observations support $G = 4$ as a practical operating point within the evaluated range; they do not establish global hyperparameter optimality.

Table 7: **Single-run accuracy (%) with Llama backbones.** PersonaMem-32K and PersonaMem-128K contain 50 and 233 test examples, respectively.

	Llama-3.2-3B-Instruct		Llama-3.1-8B-Instruct	
Method	32K	128K	32K	128K
MemoryBank	42.0	27.9	40.0	34.3
A-MEM	38.0	29.6	46.0	34.8
LightMem	48.0	35.2	44.0	35.2
Mem0	50.0	36.5	50.0	38.2
MemoryOS	32.0	33.9	44.0	28.8
MemP	42.0	36.5	38.0	33.5
LangMem	46.0	35.2	52.0	36.1
MemGen	60.0	54.9	70.0	63.9
AutoCompressor	46.0	42.9	56.0	51.5
M+	54.0	50.2	64.0	59.7
`ElasticMem`	**70.0**	**68.2**	**84.0**	**76.8**

Table 8: **Sensitivity to retrieval count Z.** ALFWorld SR is weighted over Seen and Unseen episodes. Bold values mark the best score within each tested task.

Task	Z	Accuracy / weighted SR
LoCoMo	5	0.790
LoCoMo	10	0.822
LoCoMo	20 (default)	**0.841**
LoCoMo	30	0.831
ALFWorld	5	0.489
ALFWorld	10 (default)	**0.529**
ALFWorld	15	0.518
ALFWorld	20	0.504

Table 9: **Sensitivity to GRPO group size G.** The default trades off performance against the number of sampled training trajectories.

G	LoCoMo Acc.	ALFWorld weighted SR
2	0.818	0.496
4 (default)	0.841	0.529
8	**0.844**	**0.533**

A.5 Online Latency and Offline Preprocessing

We report online inference latency under identical hardware, precision, and decoding settings across methods. The reported measurements cover native retrieval, memory processing, and generation. A component-level timing trace is unavailable, so inclusion of the separate control-token/query-state forward pass is not independently established; the values should not be interpreted as a verified complete compute accounting. Table 10 reports seconds per QA example and seconds per complete ALFWorld episode. `ElasticMem` has the lowest latency among the four evaluated latent-space methods in every setting. ALFWorld episode latency combines inference cost per step with the number of actions taken; these measurements do not separately establish a lower per-step cost. FAISS retrieval accounts for less than 0.05% of the reported average per-query latency of `ElasticMem`. No total-FLOP comparison is reported.

Table 10: **Reported online inference latency.** QA values are seconds per example; ALFWorld values are seconds per episode. Offline memory construction is excluded. All comparisons use the same hardware, precision, and decoding settings.

Backbone	Task	Unit	MemGen	AutoComp.	M+	`ElasticMem`
3B	PersonaMem-32K	s/example	0.49	0.46	0.39	**0.36**
3B	PersonaMem-128K	s/example	0.64	0.79	0.56	**0.50**
3B	LoCoMo	s/example	0.54	0.48	0.44	**0.41**
3B	LongMemEval	s/example	0.59	0.72	0.50	**0.45**
3B	ALFWorld-Seen	s/episode	87.0	122.8	82.2	**55.8**
3B	ALFWorld-Unseen	s/episode	94.7	127.7	91.0	**58.4**
7B	PersonaMem-32K	s/example	0.63	0.59	0.50	**0.44**
7B	PersonaMem-128K	s/example	0.81	1.00	0.71	**0.62**
7B	LoCoMo	s/example	0.68	0.61	0.56	**0.49**
7B	LongMemEval	s/example	0.74	0.91	0.63	**0.55**
7B	ALFWorld-Seen	s/episode	66.9	98.1	63.6	**42.5**
7B	ALFWorld-Unseen	s/episode	73.3	99.7	71.1	**44.4**

Offline preprocessing. QA memories are directly chunked from benchmark histories and require no LLM-based skill extraction. For ALFWorld, the reported average costs are 2.3 seconds per skill-card extraction and 79 milliseconds per latent encoding. The extracted skill cards are shared by the latent-space methods; each method retains its native encoding and memory-processing procedure. For 2,832 cards, multiplying the reported per-card costs gives approximately 108.6 minutes for extraction and 3.7 minutes for encoding, or 112.3 minutes in serial-equivalent work. This sum is not a separately measured parallel wall-clock construction time. These one-time costs are excluded from Table 10 and should be considered separately when assessing deployment cost and reuse of the memory bank.

A.6 Maximum-Budget Diagnostics

The maximum-budget experiments use $N_c = B_{\max}$ in every setting, including $B_{\max} = 30$ and 40; the default $N_c = 20$ applies to the default $B_{\max} = 20$ configuration. Thus, the larger allocation actions are supported by the cache. We additionally inspect the learned allocations and clip the $B_{\max} = 40$ policy's allocations to 20 at inference, leaving its learned parameters unchanged.

Table 11: **Budget diagnostics with Qwen2.5-7B-Instruct.** Downstream performance with cache length equal to the maximum budget. Clipping is applied only at inference to the policy trained with a maximum budget of 40.

Setting	PersonaMem 32K Acc.	ALFWorld weighted SR
$B_{\max} = N_c = 20$	0.900	0.529
$B_{\max} = N_c = 40$	0.860	0.482
$B_{\max} = N_c = 40$, clipped to 20	0.880	0.511

Clipping raises PersonaMem-32K accuracy from 0.860 to 0.880 and ALFWorld weighted SR from 0.482 to 0.511, without fully recovering the 20-token policy's performance. These observations are consistent with over-allocation contributing to the degradation under a larger action space and a reward without an explicit token-cost term. They do not imply that increasing the budget reduces optimal achievable performance: the larger action space still contains the smaller allocations. Because cache length changes together with the maximum budget, the training comparison varies both available cached content and allowable allocation; the inference intervention provides complementary evidence within the same trained 40-token policy.

A.7 Control-Token Intervention

The retrieval-control action is an ordinary next token sampled from the full reasoner vocabulary. Its distribution starts from the pretrained next-token distribution, with zero-initialized LoRA updates preserving the initial base-model behavior. After GRPO training, selected tokens remain query-dependent; no collapse to a single global token is observed. The relevant representation is the hidden state after the sampled token, $h_q^{\mathrm{ctrl}} = f_\theta(q, a)[-1]$, rather than the token's surface identity. Hard top-Z selection is non-differentiable. Reward-based optimization reinforces control-token actions through their likelihood term; no gradient through the selected memory indices is assumed.

Using the same trained checkpoint and frozen bank, we compare the usual control-token state with a query-only state $h_q^{\mathrm{pre}} = f_\theta(q)[-1]$ that bypasses the control step. Table 12 reports higher downstream performance when retaining the control step. LoCoMo accuracy and ALFWorld weighted SR each increase by four percentage points. This supports a useful reward-shaped retrieval transition, although bypassing the control step also changes the state supplied to the budget policy and is not an intervention on retrieval indices alone.

Table 12: **Same-checkpoint control-token intervention.** Learned parameters and the memory bank are held fixed.

Inference setting	LoCoMo Acc.	ALFWorld weighted SR
Control bypassed: query-only state	0.80	0.49
Control retained: post-token state	0.84	0.53

B BUDGET POLICY ARCHITECTURE

Budget policy network. We instantiate the budget allocator as a lightweight Transformer policy π_ϕ over the retrieved set. For each retrieved chunk m_{I_j}, we construct a chunk-level feature vector

$$r_j = W_r \left[h_q; k_{I_j}; h_q \odot k_{I_j}; s_{I_j}; e_j \right] \in \mathbb{R}^{d_b}, \tag{22}$$

where h_q is the query-conditioned retrieval state, k_{I_j} is the retrieved memory key, s_{I_j} is its retrieval score, e_j is a learned rank embedding, $\odot$ denotes element-wise multiplication, and W_r projects the concatenated features into the policy hidden dimension d_b. The sequence of retrieved-memory features is then processed by a small Transformer encoder:

$$o_1, \ldots, o_Z = \mathrm{Transformer}_\phi(r_1, \ldots, r_Z), \qquad o_j \in \mathbb{R}^{d_b}. \tag{23}$$

This self-attention structure allows the policy to compare retrieved memories against one another before assigning capacity.

For each retrieved chunk, the policy produces logits over discrete budget actions:

$$\ell_j = W_b o_j \in \mathbb{R}^{B_{\max}+1}. \tag{24}$$

The budget for chunk m_{I_j} is then sampled from the categorical distribution:

$$b_j \sim \pi_\phi(\cdot \mid h_q, I) = \mathrm{Cat}\left(\mathrm{softmax}(\ell_j)\right). \tag{25}$$

C Future Work and Broader Impact

Future Work. Beyond addressing the above scope boundaries, several directions merit investigation. Extending the evaluation to larger-scale memory corpora and more open-ended interactive environments (Yang et al., 2024; Jansen et al., 2024; Wei et al., 2025) would test the scalability of elastic latent memory. Developing richer visualization and interpretability tools for the learned budget allocation would make the memory-use process more transparent. Additionally, exploring dynamic memory bank updates, where the memory bank itself evolves as the agent accumulates new experience, is a promising direction for agents continuous evolving (Zhang et al., 2024; Jiang et al., 2025b; Chen et al., 2025).

Broader Impact. By reducing the token overhead of memory-augmented LLM agents through elastic latent allocation, `ElasticMem` has the potential to make memory-intensive LLM applications more computationally accessible. We do not foresee specific negative societal consequences beyond those generally associated with improving LLM agent capabilities. As with all advances in LLM-based agents, downstream deployments should incorporate appropriate safeguards for safety and reliability.

D Implementation Details

We implement `ElasticMem` with Qwen2.5-3B-Instruct[3] and Qwen2.5-7B-Instruct[4] as backbone reasoners. For each dataset, we train a task-specific `ElasticMem` model while keeping the same architecture and optimization protocol.

LoRA-adapted reasoner θ. We adapt the backbone reasoner with LoRA (Hu et al., 2022). Specifically, we apply LoRA to the query, key, value, and output projection matrices of all attention layers, i.e., `q_proj`, `k_proj`, `v_proj`, and `o_proj`. We use LoRA rank r=16, scaling factor α=32, and dropout 0.1.

Latent memory bank $\mathcal{B}$. Each memory chunk m_i is encoded once offline by the frozen base encoder f_{θ_0}. Following Eqs. 4–5, we store the last hidden state as the retrieval key $k_i = H_i[-1] \in \mathbb{R}^d$ and, by default, the last N_c=20 hidden states as the content cache $C_i = H_i[-N_c:] \in \mathbb{R}^{N_c \times d}$. All retrieval keys are L_2-normalized and retrieved with cosine similarity as in Eq. 9. The number of retrieved chunks Z is dataset-specific: Z=9 for PersonaMem-32K, Z=21 for PersonaMem-128K, Z=20 for LoCoMo and LongMemEval, and Z=10 for ALFWorld. Each example-specific QA bank and the shared ALFWorld bank are read-only during inference; cached states are never updated by task training. In the maximum-budget ablation, $N_c = B_{\max}$ for every setting, so budgets of 30 and 40 are supported by correspondingly longer caches.

Latent projector P_ψ. The latent projector maps selected cached hidden states into the reasoner's input embedding space. We implement P_ψ as a two-layer MLP with GELU activation and layer normalization:

$$P_\psi(\widetilde{C}_j) = \mathrm{LayerNorm}\left(\mathrm{GELU}(\widetilde{C}_j W_1^\top) W_2^\top\right), \tag{26}$$

where W_1 maps $\mathbb{R}^d$ to an intermediate dimension d'=2048, and W_2 maps $\mathbb{R}^{d'}$ back to $\mathbb{R}^d$. The projector is applied independently to each selected latent state.

Budget policy π_ϕ. The budget allocator is a lightweight Transformer encoder operating over the retrieved memory set. Following Eq. 23, we use hidden dimension d_b=256, 4 attention heads, 2 Transformer layers, and dropout 0.1. The per-chunk maximum budget is $B_{\max}$=20. Thus, each retrieved chunk receives a discrete latent budget $b_j \in \{0, \ldots, 20\}$ as defined in Eq. 11. The action b_j=0 allows the model to suppress a retrieved memory completely.

GRPO training. We jointly optimize the LoRA-adapted reasoner θ, latent projector P_ψ, and budget policy π_ϕ with group-relative policy optimization (Shao et al., 2024). We use group size G=4, clipping range ϵ=0.05, and sampling temperature τ=1.0 for trajectory generation. The

[3] `https://huggingface.co/Qwen/Qwen2.5-3B-Instruct`
[4] `https://huggingface.co/Qwen/Qwen2.5-7B-Instruct`

reward function $R(y, y^\star)$ is task-specific: exact-match accuracy for MEMORYSUITE-QA and binary agreement with the snapped expert action for ALFWorld. Environment-level task success is used only as the ALFWorld evaluation metric; see Appendix F.2. We use AdamW (Loshchilov & Hutter, 2017) with learning rate 2×10^{-5}, weight decay 0.01, and a cosine annealing schedule. All tasks are trained for 10 epochs with BF16 mixed precision. The maximum generation length is set to 5 tokens for QA tasks and 16 tokens for ALFWorld.

During evaluation, we use greedy decoding. The retrieval-control token is selected by argmax, budget allocation uses the most likely budget action under π_ϕ, and answer generation is decoded greedily.

Task-specific prompts. For PersonaMem-32K and PersonaMem-128K, the generation suffix u constrains the model to output exactly one of four options: "`(a)`", "`(b)`", "`(c)`", or "`(d)`". For LoCoMo and LongMemEval, the instruction enforces selection from ten options "`(a)`" through "`(j)`" and explicitly prohibits refusal responses. For ALFWorld, the instruction directs the agent to output a single admissible command string verbatim. The final model input follows Eq. 14:

$$X = [E(q); E(a); Z_1; \ldots; Z_Z; E(u)], \tag{27}$$

where q is the query, a is the retrieval-control token, Z_j are the projected soft memory tokens, and u is the generation suffix. For QA, u contains answer options and output-format constraints; for ALFWorld, it contains recent interaction history, the current observation, admissible actions, and action-format constraints. Soft tokens are inserted within the user-side input while preserving the chat-template boundaries.

Token counting and memory capacity. In Table 2, #Tok. counts task-text input, injected memory, and generated output per QA example; for ALFWorld, these counts are accumulated over the episode. Text memories are counted as natural-language tokens and latent memories as soft-token positions. The pure injected latent-memory budget is a separate quantity, $B_{\text{mem}} = \sum_{j=1}^{Z} b_j \le ZB_{\max}$; it excludes query, generation-suffix, and output tokens. For example, PersonaMem-32K has a default upper bound of $9 \times 20 = 180$ injected memory tokens, while #Tok. additionally includes the task text and generated output. Neither token measure includes offline bank construction or is a substitute for end-to-end latency; online and offline timings are reported separately in Appendix A.5.

Compute. All experiments are conducted on NVIDIA RTX A6000 GPUs with 48GB memory. Training a single task-specific model takes approximately 6–18 hours on one GPU, depending on dataset size and backbone scale. Offline preprocessing is reported by stage in Appendix A.5: QA history chunking uses no LLM extraction; ALFWorld skill extraction and latent encoding are one-time costs, separate from online inference.

E TRAINING PROCEDURE OF `ElasticMem`

Algorithm 1 summarizes the complete training procedure of `ElasticMem`. The procedure contains an offline memory-bank construction phase and an online reward-based optimization phase. The offline phase encodes each memory chunk once and keeps the resulting memory bank fixed throughout training. The online phase samples multiple memory-use trajectories for each query, evaluates them with task rewards, and updates the trainable components using the GRPO objective.

Offline memory-bank construction. The offline phase constructs the latent memory bank. Each memory chunk m_i is encoded once by the frozen base LLM encoder f_{θ_0}. The final hidden state $H_i[-1]$ is stored as the retrieval key k_i, while the last N_c hidden states $H_i[-N_c:]$ are stored as the content cache C_i. The resulting memory bank $\mathcal{B}$ is fixed during training, so later updates only affect the trainable reasoner, budget policy, and projector.

Trajectory sampling. The online phase describes how `ElasticMem` samples memory-use trajectories. For each query-target pair, the model samples G trajectories. In each trajectory, the reasoner first samples a retrieval-control token a_g and derives the query state $h_{q,g}$. The query state retrieves the top-Z memory chunks by cosine similarity with the cached keys. The budget policy then processes retrieved-memory features and samples a discrete latent budget $b_{g,j} \in \{0, \ldots, B_{\max}\}$ independently conditional on the jointly encoded retrieved set.

Algorithm 1 Training `ElasticMem`

Require: Memory corpus $\mathcal{M} = \{m_i\}_{i=1}^{N}$, training queries $\mathcal{Q}$, reward function R, group size G, retrieved chunk number Z, per-chunk maximum budget $B_{\max}$, clipping range ϵ
Ensure: Trained parameters θ, ψ, ϕ

// Offline phase: construct latent memory bank
for each memory chunk $m_i \in \mathcal{M}$ **do**
  $H_i \leftarrow f_{\theta_0}(m_i)$
  $k_i \leftarrow H_i[-1]$ ▷ retrieval key
  $C_i \leftarrow H_i[-N_c:]$ ▷ content cache
end for
$\mathcal{B} \leftarrow \{(m_i, k_i, C_i)\}_{i=1}^{N}$
// Online phase: reward-based optimization
for each training step **do**
  Sample query-target pair $(q, y^\star) \sim \mathcal{Q}$
  for $g = 1, \ldots, G$ **do**
    $a_g \sim p_\theta(\cdot \mid q)$
    $h_{q,g} \leftarrow f_\theta(q, a_g)[-1]$
    $s_i \leftarrow \cos(h_{q,g}, k_i)$ for all $i \in [N]$
    $I_g \leftarrow \mathrm{TopZ}_{i \in [N]}\, s_i$
    Construct retrieved-memory features $\{r_{g,j}\}_{j=1}^{Z}$ using Eq. 22
    $o_{g,1}, \ldots, o_{g,Z} \leftarrow \mathrm{Transformer}_\phi(r_{g,1}, \ldots, r_{g,Z})$
    for $j = 1, \ldots, Z$ **do**
      Compute budget logits $\ell_{g,j}$ using Eq. 24
      Sample $b_{g,j} \sim \mathrm{Cat}(\mathrm{softmax}(\ell_{g,j}))$
      if $b_{g,j} > 0$ **then**
        $\widetilde{C}_{g,j} \leftarrow C_{I_{g,j}}[-b_{g,j}:]$
        $Z_{g,j} \leftarrow P_\psi(\widetilde{C}_{g,j})$
      else
        $Z_{g,j} \leftarrow \emptyset$
      end if
    end for
    $\mathbf{b}_g \leftarrow (b_{g,1}, \ldots, b_{g,Z})$
    $X_g \leftarrow [E(q); E(a_g); Z_{g,1}; \cdots; Z_{g,Z}; E(u)]$
    $y_g \sim p_\theta(\cdot \mid X_g)$
    $r_g \leftarrow R(y_g, y^\star)$
    $\tau_g \leftarrow (a_g, I_g, \mathbf{b}_g, y_g)$
  end for
  Compute advantages $\{A_g\}_{g=1}^{G}$ using Eq. 18
  Update θ, ψ, ϕ by minimizing Eq. 19
end for

Latent memory injection. The injection steps construct the soft memory tokens. If $b_{g,j} > 0$, `ElasticMem` selects the last $b_{g,j}$ cached states from $C_{I_{g,j}}$ and maps them through the latent projector P_ψ. If $b_{g,j} = 0$, the retrieved chunk contributes no latent tokens. The final model input is formed by concatenating the query embedding, retrieval-control token embedding, selected soft memory tokens, and generation suffix.

Reward-based update. After generation, the algorithm evaluates sampled trajectories and updates the model. Each trajectory receives a task reward $R(y_g, y^\star)$. The rewards within the group are normalized into advantages using Eq. 18. Finally, the LoRA-adapted reasoner θ, latent projector ψ, and budget policy ϕ are updated by minimizing the GRPO objective in Eq. 19.

F Dataset Descriptions

We describe all evaluation datasets in MemorySuite below, categorized by their corresponding subsets. MemorySuite consists of two complementary groups: MemorySuite-QA and MemorySuite-Agentic.

MemorySuite-QA evaluates long-term memory recall and reasoning in question-answering settings, while MemorySuite-Agentic evaluates memory-augmented agentic decision-making in embodied interactive household environments. For all datasets, we report both task performance and the number of consumed tokens (#Tok.) to assess memory efficiency.

Table 13: **Detailed summary of datasets used in MemorySuite.** We categorize datasets by subset, task type, and evaluation metric.

Dataset	Task	Metric
MemorySuite-QA		
PersonaMem-32K	Persona-based Memory QA	Accuracy / #Tok.
PersonaMem-128K	Persona-based Memory QA	Accuracy / #Tok.
LoCoMo	Long-Conversation Memory QA	Accuracy / #Tok.
LongMemEval	Long-Term Interactive Memory QA	Accuracy / #Tok.
MemorySuite-Agentic		
ALFWorld-Seen	Seen Household Task Completion	SR / #Tok.
ALFWorld-Unseen	Unseen Household Task Completion	SR / #Tok.

F.1 MEMORYSUITE-QA

Parameter training and memory construction. Parameters are optimized only on the training QA pairs in the partitions specified in Appendix G. For each example, the memory bank is constructed by directly chunking that example's benchmark-provided conversation history and encoding it with the frozen offline encoder; no QA skill extraction is performed. At validation and test time, the corresponding held-out history is used only as inference-time evidence and is not pooled with histories from other examples or used for parameter updates. Questions, answer choices, ground-truth answers, model attempts, rewards, and evaluation feedback are never inserted into the memory bank. Each bank is encoded once and remains read-only throughout inference.

PersonaMem (Jiang et al., 2025a) is a personalization benchmark designed to assess whether language models can infer evolving user profiles and generate personalized responses across multi-session interactions. The benchmark is constructed via a synthetic dialogue generation pipeline that simulates realistic, temporally evolving conversations between users and chatbots, grounded in over 1,000 distinct user personas spanning diverse demographic attributes, preferences, and backgrounds. Each persona is associated with a long chat history in which user preferences are implicitly revealed across a series of task-oriented interactions. We evaluate on the **PersonaMem-32K** and **PersonaMem-128K** variants, which differ in the maximum context length of the provided conversation history. Each question is posed as a 4-way multiple-choice problem (MC-4), requiring the model to select the response that best aligns with the user's implicitly expressed persona.

LoCoMo (Maharana et al., 2024) is a benchmark for evaluating very long-term conversational memory, constructed via a machine–human pipeline that generates multi-session dialogues between LLM-based virtual agents grounded on persona descriptions and temporal event graphs. Each conversation spans up to 32 sessions with an average of 600 turns and 16K tokens. Human annotators verify and edit the generated dialogues for long-range consistency and causal coherence. We adopt the question answering subset, where questions require temporal and causal reasoning over long-range dialogue history and are posed in a 10-way multiple-choice format (MC-10)[5], probing the model's ability to track evolving personal information across sessions.

LongMemEval (Wu et al., 2024) is a comprehensive benchmark evaluating five core long-term memory abilities of chat assistants: information extraction, multi-session reasoning, temporal reasoning, knowledge updates, and abstention. It comprises 500 carefully curated questions embedded in freely extensible, timestamped user–assistant chat histories, supporting context lengths ranging from 128K tokens (`longmemeval_s`) to approximately 1.5M tokens (`longmemeval_m`). Questions are posed in a 10-way multiple-choice format (MC-10)[6].

[5] `https://huggingface.co/datasets/Percena/locomo-mc10`

[6] `https://huggingface.co/datasets/Percena/lme-mc10`

F.2 MEMORYSUITE-AGENTIC

ALFWorld-Seen (Shridhar et al., 2020) is a text-based interactive household benchmark where agents interpret natural-language observations and issue sequential actions to achieve household tasks. The Seen split (`valid_seen`) contains 140 episodes across six task types (e.g., pick-and-place, pick-heat-then-place, look-at-in-light). Room layouts overlap with training, so agents face familiar spatial configurations but novel task instances, isolating instruction-level generalization.

ALFWorld-Unseen (Shridhar et al., 2020) includes the same six task types but evaluates on `valid_unseen`, comprising 134 episodes in entirely new room layouts with unseen object arrangements, providing a stricter out-of-distribution test. For both splits, we report task success rate (SR) and average number of consumed tokens (#Tok.).

ALFWorld training setup. Unlike the QA tasks where each training sample is a single question–answer pair, ALFWorld requires learning from multi-step expert trajectories in an interactive environment. We adopt an imitation-learning–based GRPO formulation with a two-stage data pipeline.

Stage 1: Memory bank construction. We start from 3,553 expert trajectories collected via an oracle planner on the ALFWorld training environments, of which 3,541 are successful. We split the successful trajectories into a *memory pool* $\mathcal{M}$ (80%, 2,832 trajectories) and an *IL pool* $\mathcal{T}$ (20%, 709 trajectories), using a fixed seed for reproducibility. Each trajectory in the memory pool is summarized into a procedural *skill card* using an LLM (Gemini), following a structured prompt that asks for a 150–250 word procedural description including the task category, a generalized step-by-step strategy, and typical object–receptacle associations. This produces 2,832 skill cards (634 pick-two, 623 pick-and-place, 534 clean-then-place, 431 cool-then-place, 368 heat-then-place, 242 look-at-in-light). The same Gemini-generated cards and chunk boundaries are supplied to all latent-space baselines. Text-space baselines receive the same segmented raw trajectories and apply their native memory-construction procedures. Each skill card is encoded by the frozen base LLM encoder f_{θ_0} into a retrieval key and content cache, forming the latent memory bank $\mathcal{B}$ with $N{=}2{,}832$ entries and $Z{=}10$ retrieved cards per query.

Stage 2: Per-step imitation with GRPO. Each trajectory in the IL pool $\mathcal{T}$ is expanded into per-step training samples. At step t of a trajectory, the training sample consists of:

- **Query** q: the task objective (e.g., "heat some apple and put it in countertop"), which remains fixed across all steps of the same episode;
- **Task suffix** u: a rendered prompt containing the truncated interaction history (last 5 action–observation pairs), the current observation, the list of admissible actions, and the system instruction directing the agent to output one admissible command verbatim;
- **Correct answer** $y^\star$: the expert's action at step t, JSON-encoded together with the admissible action list to enable snap-to-admissible reward computation.

This expansion produces 4,260 per-step samples from 709 trajectories. The GRPO reward (Eq. 17) uses a strict deployment-aligned criterion: the model's generated action is first snapped to the nearest admissible command (by normalized string matching), and the reward is 1.0 if the snapped action exactly matches the expert action, and 0.0 otherwise. The same snap-to-admissible mapping is used at evaluation. This training reward measures expert-action agreement; the evaluation SR instead measures completion of an autonomously executed episode.

Evaluation protocol. At test time, the agent interacts with the ALFWorld environment autoregressively for up to 30 steps per episode. The fixed task objective determines retrieval of $Z{=}10$ skill cards and their latent budgets. Under greedy evaluation with fixed parameters and a read-only bank, these memory decisions remain unchanged within an episode. At each step, the agent generates an action conditioned on these soft memory tokens, the last 5 interaction steps, the current observation, and admissible actions. The generated action is snapped to the nearest admissible command before execution. An episode is considered successful if the agent achieves a cumulative reward of ≥ 1.0 before reaching the step limit.

G Dataset Statistics

In this section, we present detailed statistics for each dataset. The statistics for MemorySuite-QA and MemorySuite-Agentic are provided in Table 14 and Table 15, respectively. For MemorySuite-QA, PersonaMem variants are split 80/10/10 by `shared_context_id`; LoCoMo is partitioned by conversation index (first 6/middle 2/last 2 conversations); and LongMemEval follows a 60/20/20 random split by question. For MemorySuite-Agentic, we adopt the official splits: ALFWorld provides separate *Seen* (`valid_seen`) and *Unseen* (`valid_unseen`) test environments, both sharing the same training set of expert demonstrations. For each dataset, we additionally report Z, the number of top-Z memory chunks retrieved and supplied to the model during inference, which directly governs the memory retrieval budget and token consumption.

Table 14: **MemorySuite-QA Data Statistics.** Z denotes the number of top-Z memory chunks retrieved per query.

Dataset	Split			Total	Z
	Train	Val	Test		
PersonaMem-32K	489	50	50	589	9
PersonaMem-128K	2,221	273	233	2,727	21
LoCoMo	885	341	314	1,540	20
LongMemEval	300	100	100	500	20

Table 15: **MemorySuite-Agentic Data Statistics.** Z denotes the number of top-Z memory chunks selected from the task objective and reused throughout an episode. Train denotes the raw demonstration pool before filtering and partitioning; both evaluation splits share this pool.

Split	ALFWorld		Z
	Seen	Unseen	
Raw train pool	3,553	3,553	10
Test	140	134	

H Prompt Usage

This section describes the prompt templates used by all memory baselines across our evaluation datasets. The baselines differ in how they *build* memory from conversation sessions; details for each baseline's memory-construction prompts are provided in Section H.1. Each dataset employs a distinct answering prompt: LoCoMo and LongMemEval use 10-choice multiple-choice prompts (Sections H.3 and H.4), PersonaMem-32K and PersonaMem-128K use 4-choice prompts (Section H.2), and ALFWorld uses a task-specific prompt described in Section H.5.

H.1 Memory Construction Prompts

The following templates describe method-specific memory construction. Text-space methods apply their native procedures to the shared raw source information. ElasticMem uses direct history chunking for QA and offline skill-card construction for ALFWorld; its memory banks are not updated from evaluation feedback.

Mem0

Mem0 (Chhikara et al., 2025) runs a two-step pipeline per session turn. **Step 1** (Table 16) extracts factual statements from the conversation as a JSON list. **Step 2** (Table 17) compares each new fact against existing entries in a FAISS store and decides whether to ADD, UPDATE, DELETE, or make no change (NONE). At inference, the top-k stored facts (with timestamps) are retrieved by cosine similarity and prepended as a bullet list.

Table 16: **Mem0: Fact Extraction Prompt (Step 1 of memory construction).**

You are a Personal Information Organizer, specialized in accurately storing facts, user memories, and preferences. Your primary role is to extract relevant pieces of information from conversations and organize them into distinct, manageable facts.

Types of Information to Remember:
1. Store Personal Preferences: Keep track of likes, dislikes, and specific preferences.
2. Maintain Important Personal Details: Remember significant personal information like names, relationships, and important dates.
3. Track Plans and Intentions: Note upcoming events, trips, goals, and any plans the user has shared.
4. Monitor Health and Wellness Preferences: Keep a record of dietary restrictions, fitness routines, and other wellness-related information.
[. . . other categories omitted for brevity . . .]

Here are some few shot examples:
Input: Hi, my name is John. I am a software engineer.
Output: {"facts" : ["Name is John", "Is a Software engineer"]}

Return the facts and preferences in a json format as shown above.
Remember the following:
- Today's date is **{today}**.
- Do not return anything from the custom few shot example prompts provided above.
- Create the facts based on the user and assistant messages only.
- Make sure to return the response in the format mentioned in the examples.

Following is a conversation between the user and the assistant. Extract the relevant facts and preferences about the user, if any.

Table 17: **Mem0: Memory Update Decision Prompt (Step 2 of memory construction).**

You are a smart memory manager which controls the memory of a system.
You can perform four operations: (1) add into the memory, (2) update the memory, (3) delete from the memory, and (4) no change.
Compare newly retrieved facts with the existing memory. For each new fact, decide whether to:
- ADD: Add it to the memory as a new element
- UPDATE: Update an existing memory element
- DELETE: Delete an existing memory element
- NONE: Make no change if the fact is already present or irrelevant

[...guidelines for each operation with examples omitted for brevity...]

You must return your response in the following JSON structure only:

```
{
   "memory" : [
      {"id" : "<ID>", "text" : "<Content>", "event" : "<ADD|UPDATE|DELETE|NONE>"}
   ]
}
```

LangMem

LangMem uses a single LLM call per session to extract personal facts in a "`FACT:`" prefixed format (Table 18). Facts are stored in a LangGraph `InMemoryStore` with date prefixes and retrieved by embedding similarity at query time.

Table 18: **LangMem: Fact Extraction Prompt.**

Extract all personal facts about the user from this conversation. Output one fact per line using the format "FACT: <content>". Be concise but complete. Only output FACT lines, nothing else.

Conversation:
{conversation}

Continued on next page

Table 18 (continued)

Facts:

MEMORYBANK

MemoryBank (Zhong et al., 2024) creates two levels of summary per session. A **content summary** (Table 19) captures the key events of the dialogue, while a **personality summary** (Table 20) records the user's traits and suggests response strategies. Globally, these are aggregated into an overall history and overall personality summary that persist across sessions. At inference, the most relevant per-session summary is retrieved by FAISS cosine similarity and concatenated with the personality and overall-history blocks.

Table 19: **MemoryBank: Content Summary Prompt (applied to each session).**

Please summarize the following dialogue as concisely as possible, extracting the main themes and key information. If there are multiple key events, you may summarize them separately. Dialogue content:
{dialogue}
Summarization:

Table 20: **MemoryBank: Personality and Strategy Summary Prompt.**

Based on the following dialogue, please summarize **{user_name}**'s personality traits and emotions, and devise response strategies based on your speculation. Dialogue content:
{dialogue}
{user_name}'s personality traits, emotions, and **{boot_name}**'s response strategy are:

MEMORYOS

MemoryOS (Kang et al., 2025) organizes memory in three tiers: short-term (STM, within-session), mid-term (MTM, cross-episode topic segments), and long-term (LPM, persistent profile). When the STM overflows, three prompts fire in sequence: (1) a **continuity check** (Table 21) to decide whether to merge or split page boundaries, (2) a **meta-info update** (Table 22) to maintain a running dialogue summary, and (3) a **multi-summary** (Table 23) to extract subtopic clusters that form MTM segments. At inference, the most similar MTM pages are retrieved by cosine similarity and formatted as `[obs]/[action]` episodes.

Table 21: **MemoryOS: Continuity Check Prompt.**

Determine if these two conversation pages are continuous (true continuation without topic shift).
Return ONLY "true" or "false".

Previous Page:
User: **{prev_user}**
Assistant: **{prev_agent}**

Current Page:
User: **{curr_user}**
Assistant: **{curr_agent}**

Continuous?

Table 22: **MemoryOS: Meta-Info Update Prompt.**

Update the conversation meta-summary by incorporating the new dialogue while maintaining continuity.

Guidelines:
1. Start from the previous meta-summary (if exists)
2. Add/update information based on the new dialogue

Continued on next page

Table 22 (continued)

3. Keep it concise (1-2 sentences max)
4. Maintain context coherence

Previous Meta-summary: **{last_meta}**
New Dialogue:
{new_dialogue}

Updated Meta-summary:

Table 23: **MemoryOS: Multi-Summary Prompt (subtopic extraction).**

Please analyze the following dialogue and generate extremely concise subtopic summaries, if applicable, with a maximum of two themes.
Each summary should be very brief – just a few words for the theme and content. Format as JSON array:

```
[
  {"theme": "Brief theme", "keywords": ["key1", "key2"], "content": "summary"}
]
```

Conversation content:
{text}

A-MEM

A-MEM (Xu et al., 2025a) creates structured notes for each session. An **analysis prompt** (Table 24) extracts keywords, context, and tags. An **evolution prompt** (Table 25) then decides whether the new note should be linked (STRENGTHEN) or whether neighboring notes' context and tags should be updated (UPDATE_NEIGHBOR). At inference, a keyword-expanded query is issued against the FAISS note store.

Table 24: **A-MEM: Content Analysis Prompt.**

Analyze the following content and provide:
1. KEYWORDS: The most important keywords (nouns, verbs, key concepts). Order from most to least important. At least three keywords.
2. CONTEXT: One sentence summarizing the main topic, key points, and purpose.
3. TAGS: Broad categories/themes for classification (domain, format, type). At least three tags.

Respond using EXACTLY this format (one section per header):

KEYWORDS: keyword1, keyword2, keyword3, ...
CONTEXT: A single sentence summarizing the content.
TAGS: tag1, tag2, tag3, ...

Content for analysis:
{content}

Table 25: **A-MEM: Memory Evolution Decision Prompt.**

You are an AI memory evolution agent. Analyze the new memory note and its nearest neighbors to decide if evolution is needed.

New memory:
- Context: **{context}**
- Content: **{content}**
- Keywords: **{keywords}**

Nearest neighbor memories:
{nearest_neighbors_memories}

Based on the relationships between the new memory and its neighbors, decide:

Continued on next page

Table 25 (continued)

- NO_EVOLUTION: The memory stands alone, no changes needed.
- STRENGTHEN: The new memory should be linked to some neighbors and its tags updated.
- UPDATE_NEIGHBOR: The neighbors' context/tags should be updated based on new understanding.
- STRENGTHEN_AND_UPDATE: Both strengthen and update neighbors.

Respond using EXACTLY this format:
DECISION: <one of NO_EVOLUTION, STRENGTHEN, UPDATE_NEIGHBOR, STRENGTHEN_AND_UPDATE>
REASON: <brief explanation>

LIGHTMEM

LightMem (Fang et al., 2025a) first compresses each session with LLMLingua-2 (or a Jaccard-shingle fallback) and buffers compressed turns in a short-term memory (STM). When the STM token budget is exceeded, it fires the **extraction prompt** (Table 26) to produce atomic facts with source identifiers. Periodically, a **consolidation pass** (Table 27) reviews existing long-term entries against newly extracted facts and issues update/delete/ignore decisions. At inference, the top-k long-term facts are retrieved by cosine similarity and formatted with ISO timestamps.

Table 26: **LightMem: Fact Extraction Prompt (fires on STM overflow).**

You are a Personal Information Extractor.
Your task is to extract **all possible facts or information** about the user from a conversation.

Important Instructions:
1. You MUST process every user message in order, one by one.
   For each message, decide whether it contains any factual information.
   - If yes → extract it and rephrase into a standalone sentence.
   - If no, such as pure greeting, filler, or irrelevant remark, → skip it.
   - Do NOT skip just because the information looks minor or unimportant.
2. Perform light contextual completion so that each fact is a clear standalone statement.
   Examples: "user: Bought apples yesterday" → "User bought apples yesterday."
3. Output format:

```
{
  "data": [
    {"source_id": "<source_id>", "fact": "<complete fact with ALL specific details>"}
  ]
}
```

Table 27: **LightMem: Memory Consolidation Prompt.**

You are a memory management assistant.
Your task is to decide whether the target memory should be updated, deleted, or ignored based on the candidate source memories.
Decision rules:
1. Update: If the target and candidate memories describe essentially the same fact but are not fully consistent, update by integrating additional information.
2. Delete: If the target and candidate memories contain a direct conflict, delete the target memory.
3. Ignore: If unrelated, no action is needed.

The output must be a JSON object:

```
{
  "action": "update" | "delete" | "ignore",
  "new_memory": {...}    // only required when action = "update"
}
```

MEMP

MeMP (Fang et al., 2025b) learns *procedural* memory: after each session it distills the conversation into a natural-language workflow paragraph using the prompt in Table 28. Workflows are stored in a FAISS index keyed by session query embeddings and retrieved by cosine similarity at inference, then

injected as task guidelines. On a failed attempt, the workflow is refined with a separate adjustment prompt.

Table 28: **MeMP: Workflow Generation Prompt.**

You are provided with a query and a trajectory taken to solve the query. The trajectory consists of multiple steps of thought, action and observation.
Your task is to generate a workflow based on critical steps to help solve similar queries in the future.
A critical step is one that has a significant impact on fulfilling the query, the step action belongs to the set [go to, take from, put in/on, open, close, use, clean with, heat with, cool with, examine, look], and the action's outcome is successful and contributes positively to achieving the query.
Notice: Write the workflow as a natural, coherent paragraph (not as a bullet list or numbered steps). Use clear, concise language to describe what actions should be taken and in what general order.
—-EXAMPLE WORKFLOW—-
To solve this query, begin by identifying the most likely receptacles where the target object can be found and visit them one by one. After locating and taking the object, perform any required transformation such as cleaning at a sinkbasin, heating with a microwave, or cooling with a fridge. Finally, go to the destination receptacle and put the object in/on it to complete the task.
—-EXAMPLE END—-
Query:
{query}
Trajectory:
{trajectory}
Output the workflow without any explanation or context:

ELASTICMEM

QA memory is directly chunked from benchmark histories (Appendix F.1); it is not extracted from question attempts, answers, or feedback. For ALFWorld, successful training trajectories in the memory pool are distilled offline into procedural skill cards using Table 29. These cards are shared across latent-space methods and remain fixed during evaluation. No skills are extracted from validation/test attempts or failed evaluation episodes.

Table 29: **ElasticMem: ALFWorld Skill Extraction Prompt (Successful Trajectory).**

system
You are an expert at analyzing household robot trajectories. Extract specific, actionable lessons from the provided trajectory.
user
An ALFWorld household task was completed successfully.

Task: **{task}**
Full trajectory (action → observation): **{trajectory}**

Extract a reusable procedural skill (150–250 words). Include:
1. The general task category (`pick_and_place`, `heat_then_place`, `clean_then_place`, `cool_then_place`, `examine_in_light`, `pick_two`)
2. The concrete step-by-step strategy that worked
3. Common locations where target objects are found (e.g. "soapbar is usually on countertop, bathtubbasin, or shelf")
Be specific. Use actual object/location types (`countertop`, `sinkbasin`, `microwave`).

Output format: `SKILL: [your skill text]`

Shared task instructions and memory interfaces. The answering templates below show the text-memory interface. Task instructions, answer choices, and chat roles are shared; the memory representation is method-specific. The retrieved-text field applies to text-memory methods. ElasticMem instead inserts the retrieval-control token and projected latent states according to Eq. 14; it does not serialize those states into retrieved text.

H.2 PERSONAMEM ANSWERING PROMPT

PersonaMem-32K and PersonaMem-128K use a 4-choice MC format (a–d) instead of the 10-choice format used by LongMemEval. The system message sets the role as: "You are a careful assistant answering 4-choice multiple-choice questions grounded in a long persona-conversation." The user prompt (Table 30) presents the question, retrieved memory chunks, an instruction line, and four options.

Table 30: **PersonaMem: MC Answering Prompt (text-memory interface).**

QUESTION:
{question}

RETRIEVED MEMORY (relevant chunks from prior conversation):
{retrieved_text}

Answer with exactly one of the four options below, formatted as a single option label like "(a)", "(b)", "(c)", or "(d)". Do not output any other text.

OPTIONS:
(a) **{choice_0}**
(b) **{choice_1}**
(c) **{choice_2}**
(d) **{choice_3}**

ANSWER:

H.3 LOCOMO ANSWERING PROMPT

LoCoMo uses a 10-choice multiple-choice format where the model selects one of ten lettered options (a)–(j). All baselines share the same answering prompt shown in Table 31: the question, the baseline-retrieved memory chunks under a "Retrieved context" header, a strict-format instruction, the ten options, and an "Answer:" marker.

Table 31: **LoCoMo: MC Answering Prompt (text-memory interface).**

Question: **{question}**

Retrieved context (from the conversation history):
{retrieved_context}

You MUST pick exactly one option from (a) to (j) – one of them is guaranteed to be correct. Do NOT say "not answerable", "I don't know", or refuse. If uncertain, make your best guess. Output exactly one token in the form (a), (b), (c), (d), (e), (f), (g), (h), (i), or (j).

Options:
(a) **{choice_0}**
(b) **{choice_1}**
...
(j) **{choice_9}**

Answer:

H.4 LONGMEMEVAL ANSWERING PROMPT

All baselines share the same MC answering prompt on LongMemEval (Table 32), which prepends retrieved memory text to the question and asks the model to select one of ten lettered options.

Table 32: **LongMemEval: MC Answering Prompt (text-memory interface).**

[Relevant memories:]
{retrieved_memories}

{question}

(a) **{choice_0}**
(b) **{choice_1}**
...
(j) **{choice_9}**

You MUST pick exactly one option from (a) to (j) – one of them is guaranteed to be correct. Do NOT say "not answerable", "I don't know", or refuse. If uncertain, make your best guess. Output exactly one token in the form (a), (b), (c), (d), (e), (f), (g), (h), (i), or (j).

Answer: (

H.5 ALFWorld Answering Prompt

ALFWorld baselines act inside a TextWorld game loop: at each step the agent receives the task goal, the retrieved memory block, an interaction history, the current observation, and the admissible-action list, then must output one admissible command verbatim (Table 33).

Table 33: **ALFWorld: Step Prompt (text-memory interface).**

[SYSTEM] You are controlling a text-based ALFWorld environment. Choose the NEXT action as ONE admissible command string. Output only the command, copied verbatim from the admissible list.

[USER] Task: **{objective}**

{retrieved_memory_block}

Interaction history so far:
{history}

Current observation: **{current_obs}**

Admissible actions: **{admissible}**

Action:

I CASE STUDIES

This section presents successful case studies for each memory baseline across our evaluation datasets. Each table shows the question with lettered choices (correct answer **bolded**), the ground truth, the retrieved memory block (abbreviated for baselines with large stores), and the model's response. We present case studies for LongMemEval (Section I.4), PersonaMem-32K (Section I.1), and PersonaMem-128K (Section I.2).

I.1 PERSONAMEM-32K

This subsection presents one successful example per baseline on PersonaMem-32K, drawn from the 7B model evaluation results. Each table shows the question with four lettered choices (correct answer **bolded**), the ground truth, the retrieved memory block (abbreviated for baselines with large stores), and the model's response. Case studies for Mem0, LangMem, and LightMem are in Tables 34–36; MemoryOS and MemoryBank in Tables 37–38; A-MEM and MeMP in Tables 39-40.

Table 34: **Mem0's case study in PersonaMem-32K.**

Question: After attending a writing workshop, I discovered I actually enjoy creatively articulating my thoughts on music. It was exhilarating to engage with fellow aspiring writers, exchanging ideas and techniques that helped unlock previously untapped aspects of my creativity. The writing exercises we did were both challenging and inspiring, pushing me to find new ways to express the emotional and technical nuances of music. I had always had an affinity for music, but this workshop gave me the confidence to put my feelings into words, which has been incredibly fulfilling. It felt like I was discovering a new dimension of myself, one that combined my love of music with a passion for writing.

(a) I remember when you initially mentioned disliking writing music reviews, perhaps because it might have seemed daunting to articulate your thoughts on music in writing. However, it's wonderful to see how your experience at the writing workshop ... ✓
(b) I remember when you initially mentioned disliking writing music reviews, perhaps because it might have seemed daunting to articulate your thoughts on music in writing. However, it's surprising that your enthusiasm for writing only increased af...
(c) I remember when you initially mentioned being indifferent to writing music reviews, as you had never considered stewing over your thoughts on music in writing before. However, it's wonderful to see how your experience at the writing workshop h...
(d) I remember when you initially mentioned enjoying writing music reviews, perhaps because it might have seemed exciting to articulate your thoughts on music in writing. However, it's wonderful to see how your experience at the writing workshop h...

Ground Truth: (a) I remember when you initially mentioned disliking writing music reviews, perhaps because it might have seemed daunting to articulate your thoughts on music in writing. However, it's wonderful to see how your experience at the writing workshop ...

Retrieved Memories:
Relevant memories from past experience:
- [2026-04-27] Researches themes and curates playlists for each podcast episode
- [2026-04-27] Felt unsure of own contributions during collaboration
- [2026-04-27] Shared stories about individual musical journeys in meetings
- [2026-04-27] Eager to apply constructive criticism techniques in a collaborative setting
- [2026-04-27] Eager to apply constructive criticism techniques in a collaborative setting
- [2026-04-27] Found a music documentary very informative and eye-opening
- [2026-04-27] Attended a music festival on 2026-04-27
[. . . 3 additional lines omitted. . .]

Response: (a)

Table 35: **LangMem's case study in PersonaMem-32K.**

Question: I'm looking to find something that can really add a new dimension to my classic rock listening experience. Any suggestions?

Continued on next page

Table 35 (continued)

(a) You might really enjoy the sound of traditional Pacific Islander music played on native instruments. While not classic rock, exploring your roots through music can add a personal layer to your listening experience. The rhythmic beats and cultu...
(b) You might enjoy the crisp convenience of streaming 'Led Zeppelin IV' in high definition, ensuring you have the album instantly at your fingertips with no need for additional equipment. The sure quality assurance of digital music allows 'Stairw...
(c) Consider embarking on a vibrant journey with live recordings from famous classic rock concerts. Capturing the raw energy and spontaneity of legendary performances, live albums offer an immediacy that studio versions may not have, added with cr...
(d) You might love the experience of spinning 'Led Zeppelin IV' on vinyl. It's an album that not only defined an era but also carries a warmth and depth in its sound that digital formats often miss. The artwork on the cover is iconic, and the tact... ✓

Ground Truth: (d) You might love the experience of spinning 'Led Zeppelin IV' on vinyl. It's an album that not only defined an era but also carries a warmth and depth in its sound that digital formats often miss. The artwork on the cover is iconic, and the tact...

Retrieved Memories:
Relevant memories:
- [assistant] Assistant: Missing out on activities like a music documentary night can certainly feel like a lost opportunity, especially when they provide such rich environments for connection and inspiration.
- [assistant] Assistant: Definitely! It brings a unique warmth to music listening, enhancing the frequencies that often get lost in more compressed digital formats. The way vinyl can accentuate the nuances of a recording creates a richer auditory experience that many enthusiasts adore. When you p...
- [assistant] Assistant: I'm so glad to hear that you enjoyed those fresh sounds! Music has such a powerful impact on our mood and overall well-being, often providing a much-needed escape from the routines of everyday life. Listening to new genres or artists can unlock entirely new perspectives a...
- [assistant] Assistant: Absolutely! Music has a unique way of connecting us to new experiences. It serves as a bridge between different cultures, allowing us deeper insight into lives and traditions that we may never encounter in person. This connection is particularly profound when you encounte...
- [assistant] Assistant: Visiting a record store can be a profoundly rewarding experience, often akin to stepping into a treasure trove filled with the sounds of the past and the promise of future discoveries. The tactile nature of vinyl, from its grooved surface to the large, colorful album art,...
- [assistant] Assistant: I can completely understand why you find karaoke nights so enjoyable, as they often foster an environment of joy and connection among friends and even strangers. Participating in such an event is much more than just singing; it's about the shared laughter, the playful com...

Response: (d)

Table 36: **LightMem's case study in PersonaMem-32K.**

Question: I recently joined a forum discussion about humor in music.

(a) I remember you mentioning how you enjoy engaging in online music discussions. That's great, forums can be a wonderful way to connect with people. ✓
(b) I seem to recall you saying you shy away from online music discussions. It's interesting to see you've decided to dive in now.
(c) I remember you talking about how you enjoy participating in travel forums. It's great that you're finding ways to connect through different topics.
(d) That's nice to hear! Joining forums can offer new insights. Engaging in discussions about humor in music sounds like a fascinating experience.

Ground Truth: (a) I remember you mentioning how you enjoy engaging in online music discussions. That's great, forums can be a wonderful way to connect with people.

Retrieved Memories:
Long-term memory (relevant facts):
- [2026-04-27T10:53:54.691 Mon] User passed on a music documentary night.
- [2026-04-27T10:58:46.191 Mon] Writing exercises pushed user to find new ways to express the emotional and technical nuances of music.
- [2026-04-27T10:53:06.497 Mon] User experienced frustration when attempts did not match the sound in his head.
- [2026-04-27T10:54:08.119 Mon] User values authentic dialogues about music over abstract opinions.
- [2026-04-27T10:56:49.256 Mon] Pooling collective knowledge helps uncover different styles of playing.

Continued on next page

Table 36 (continued)

- [2026-04-27T10:56:41.318 Mon] User shares classic tracks each week to connect with others.
- [2026-04-27T10:53:34.691 Mon] User is happy with their eclectic taste and does not need validation from current trends.
[. . . 3 additional lines omitted. . .]

Response: (a)

Table 37: **MemoryOS's case study in PersonaMem-32K.**

Question: I've seen a lot of people getting into DIY craft projects lately. I'm a bit unsure, but should I give it a try?

(a) It sounds like you're considering exploring DIY craft projects. Trying out such activities can be a fun and rewarding experience, especially if you're eager to learn new skills. You may enjoy the satisfaction of creating something tangible wit...
(b) It sounds like you're considering exploring DIY craft projects. Trying out such activities can be a fun and rewarding experience, providing a nice break from everyday routines. If you're interested in exploring something new, crafts could be a...
(c) It sounds like you're considering exploring DIY craft projects. Trying out such activities can be a fun and rewarding experience, especially if you're looking to incorporate a creative twist into something new. If you enjoyed the engaging natu... ✓
(d) It sounds like you're considering exploring DIY craft projects. Trying out such activities can be a fun and rewarding experience, especially if you like activities that offer room for creativity. Engaging in crafts could be a fulfilling way to...

Ground Truth: (c) It sounds like you're considering exploring DIY craft projects. Trying out such activities can be a fun and rewarding experience, especially if you're looking to incorporate a creative twist into something new. If you enjoyed the engaging natu...

Retrieved Memories:
[Mid-term – retrieved cross-episode pages]
- [obs] [user] Music documentaries can be incredibly inspiring, as they often delve into the lives of musicians, their creative processes, and the challenges they face along their journeys. However, choosing not to attend such events can stem from a variety of reasons, including feeling unworthy ...
[action]
[next_obs]
meta: User expresses interest in attending music documentary nights, acknowledging their potential to be inspiring yet recognizing the fear of not fitting in due to insecurities and self-doubt.
- [obs] [user] User: On , I launched a new podcast series focusing on music's role in cultural identities. This project has me delving into how different genres of music shape and reflect the societies they come from, which adds a rich layer of meaning to my understanding of cultural expression. ...
[action]
[next_obs]
[. . . 142 additional lines omitted. . .]

Response: (c)

Table 38: **MemoryBank's case study in PersonaMem-32K.**

Question: I'm looking to find something that can really add a new dimension to my classic rock listening experience. Any suggestions?

(a) You might really enjoy the sound of traditional Pacific Islander music played on native instruments. While not classic rock, exploring your roots through music can add a personal layer to your listening experience. The rhythmic beats and cultu...
(b) You might enjoy the crisp convenience of streaming 'Led Zeppelin IV' in high definition, ensuring you have the album instantly at your fingertips with no need for additional equipment. The sure quality assurance of digital music allows 'Stairw...
(c) Consider embarking on a vibrant journey with live recordings from famous classic rock concerts. Capturing the raw energy and spontaneity of legendary performances, live albums offer an immediacy that studio versions may not have, added with cr...

Continued on next page

Table 38 (continued)

(d) You might love the experience of spinning 'Led Zeppelin IV' on vinyl. It's an album that not only defined an era but also carries a warmth and depth in its sound that digital formats often miss. The artwork on the cover is iconic, and the tact... ✓

Ground Truth: (d) You might love the experience of spinning 'Led Zeppelin IV' on vinyl. It's an album that not only defined an era but also carries a warmth and depth in its sound that digital formats often miss. The artwork on the cover is iconic, and the tact...

Retrieved Memories:
MemoryBank context

Agent personality / response strategy (from past tasks):
Summary:

User's Personality:
Kai is passionate, reflective, supportive, curious, and empathetic. He experiences joy, inspiration, nostalgia, confidence, and reflection, particularly in his love for music, especially jazz and Pacific Islander genres.

[. . . 29 additional lines omitted. . .]

Response: (d)

Table 39: **A-MEM's case study in PersonaMem-32K.**

Question: After attending a writing workshop, I discovered I actually enjoy creatively articulating my thoughts on music. It was exhilarating to engage with fellow aspiring writers, exchanging ideas and techniques that helped unlock previously untapped aspects of my creativity. The writing exercises we did were both challenging and inspiring, pushing me to find new ways to express the emotional and technical nuances of music. I had always had an affinity for music, but this workshop gave me the confidence to put my feelings into words, which has been incredibly fulfilling. It felt like I was discovering a new dimension of myself, one that combined my love of music with a passion for writing.

(a) I remember when you initially mentioned disliking writing music reviews, perhaps because it might have seemed daunting to articulate your thoughts on music in writing. However, it's wonderful to see how your experience at the writing workshop ... ✓
(b) I remember when you initially mentioned disliking writing music reviews, perhaps because it might have seemed daunting to articulate your thoughts on music in writing. However, it's surprising that your enthusiasm for writing only increased af...
(c) I remember when you initially mentioned being indifferent to writing music reviews, as you had never considered stewing over your thoughts on music in writing before. However, it's wonderful to see how your experience at the writing workshop h...
(d) I remember when you initially mentioned enjoying writing music reviews, perhaps because it might have seemed exciting to articulate your thoughts on music in writing. However, it's wonderful to see how your experience at the writing workshop h...

Ground Truth: (a) I remember when you initially mentioned disliking writing music reviews, perhaps because it might have seemed daunting to articulate your thoughts on music in writing. However, it's wonderful to see how your experience at the writing workshop ...

Retrieved Memories:
talk start time:task=personamem-mc|step=141memory content: task: answer multiple-choice questions about a long persona-grounded conversation | obs: [user] User: A few days later, I reluctantly watched a music documentary DVD I was gifted. | action: | result: memory context: The content describes...
talk start time:task=personamem-mc|step=106memory content: task: answer multiple-choice questions about a long persona-grounded conversation | obs: [user] User: This experience opened my eyes to the power of visual storytelling in understanding music; it was a fantastic experience. The documentar...
talk start time:task=personamem-mc|step=106memory content: task: answer multiple-choice questions about a long persona-grounded conversation | obs: [user] User: This experience opened my eyes to the power of visual storytelling in understanding music; it was a fantastic experience. The documentar...
talk start time:task=personamem-mc|step=152memory content: task: answer multiple-choice questions about a long persona-grounded conversation | obs: [system] Current user persona: Name: Kai
Gender Identity: Transgender male
Racial Identity: Pacific Islander

Continued on next page

Table 39 (continued)

Kai, born in 1983, is a devoted fan of Louie Anderson's comedy, captivated by Anderson's unique blend of humor and heart. Growing up in a multicultural community, his Pacific Islander heritage has deeply influenced his appreciation for storytelling and humor that resonates across cultures. Embrac...
[...36 additional lines omitted...]

Response: (a)

Table 40: **MeMP's case study in PersonaMem-32K.**

Question: I'm looking to find something that can really add a new dimension to my classic rock listening experience. Any suggestions?

(a) You might really enjoy the sound of traditional Pacific Islander music played on native instruments. While not classic rock, exploring your roots through music can add a personal layer to your listening experience. The rhythmic beats and cultural stories told through these melodies hold a deep e...
(b) You might enjoy the crisp convenience of streaming 'Led Zeppelin IV' in high definition, ensuring you have the album instantly at your fingertips with no need for additional equipment. The sure quality assurance of digital music allows 'Stairway to Heaven' to...
(c) Consider embarking on a vibrant journey with live recordings from famous classic rock concerts. Capturing the raw energy and spontaneity of legendary performances, live albums offer an immediacy that studio versions may not have, added with crowd reactions and improvised solos...
(d) You might love the experience of spinning 'Led Zeppelin IV' on vinyl. It's an album that not only defined an era but also carries a warmth and depth in its sound that digital formats often miss. The artwork on the cover is iconic, and the tactile connection of placing the needle down will make 'Stairway to Heaven' feel like an epic journey through your speakers. ✓

Ground Truth: (d)

Retrieved Memories:
Here are some guidelines of how to solve the similar task:
- task_name: answer multiple-choice questions about a long persona-grounded conversation
  guidelines: To solve this query, begin by examining the conversation transcript to understand the context and key points discussed. Recall Kai's background as a classic rock enthusiast who values tactile, authentic experiences over digital convenience. Identify which option aligns with his established preference for physical media and nostalgic engagement with music. Cross-reference with his earlier statements about preferring vinyl warmth over streaming clarity. Finally, select the answer that best captures his love for the analog listening ritual.
- task_name: recommend music experience for vinyl collector
  guidelines: To recommend a music experience, first identify the user's preferred format (vinyl, CD, streaming). If the user has expressed love for physical media, prioritize options involving tangible interaction such as record spinning, liner note reading, or album art appreciation. Consider the warmth and depth of analog sound as a key differentiator from digital formats. Match the recommendation to artists and albums the user has previously mentioned enjoying.
- task_name: identify user preference evolution across conversation
  guidelines: To track preference evolution, begin by scanning the full conversation history for mentions of the target topic. Note the chronological order of statements and identify any shifts in sentiment. Pay attention to phrases like "I used to think" or "after trying X, I now prefer Y" as markers of preference change. Summarize the trajectory from initial stance to current preference.
[...198 additional lines omitted...]

Response: (d)

I.2 PERSONAMEM-128K

This subsection presents one successful example per baseline on PersonaMem-128K, drawn from the 7B model evaluation results. Each table shows the question with four lettered choices (correct answer **bolded**), the ground truth, the retrieved memory block (abbreviated for baselines with large stores), and the model's response. Case studies for Mem0, LangMem, and LightMem are in Tables 41–43; MemoryOS and MemoryBank in Tables 44–45; A-MEM and MeMP in Tables 46-47.

Table 41: **Mem0's case study in PersonaMem-128K.**

Question: Recently, I was at a yoga retreat.

(a) Ah, I remember you mentioned your dislike for yoga before. It's great to revisit activities to confirm our true preferences.
(b) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons. ✓
(c) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons.
(d) Oh, you went on a yoga retreat? That sounds unique! What was it like engaging in new activities?

Ground Truth: (b) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons.

Retrieved Memories:
Relevant memories from past experience:
- [2026-04-29] Likes a positive tone and clear, concise wording
- [2026-04-29] Discovering patterns and connections through journal entries
- [2026-04-29] Reflects on podcast conversations during daily routines
- [2026-04-29] Considers volunteering more after positive feedback from a mentoring experience
- [2026-04-29] Hosted a webinar for aspiring journalists on an earlier date
- [2026-04-29] Considers starting a journal to document thoughts and insights after podcast episodes
- [2026-04-29] Sense of community blossomed during home decoration workshop interactions
[. . . 3 additional lines omitted. . .]

Response: (b)

Table 42: **LangMem's case study in PersonaMem-128K.**

Question: Recently, I was at a yoga retreat.

(a) Ah, I remember you mentioned your dislike for yoga before. It's great to revisit activities to confirm our true preferences.
(b) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons. ✓
(c) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons.
(d) Oh, you went on a yoga retreat? That sounds unique! What was it like engaging in new activities?

Ground Truth: (b) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons.

Retrieved Memories:
Relevant memories:
- [user] User: On , I attended an interactive workshop focused on building romantic relationships.
- [user] User: Well, on , I joined a local singles group to connect with other individuals seeking friendships and relationships.
- [user] User: I attempted to create a collection of thrifted home decor items.
- [user] User: On the plus side, I embarked on a hiking trip in the mountains recently, and it was wonderful! The scenery was breathtaking, with majestic peaks towering over lush valleys, and the sound of rustling leaves added to the peaceful atmosphere. I could feel the invigorating air fill my ...
- [user] User: I tried visiting a local coffee shop to meet new people, but I felt quite uncomfortable. Upon entering, I noticed the bustling activity around me, which was overwhelming rather than inviting. The chatter of groups engaging in lively conversations only highlighted my own sense of is...
- [user] User: It's really disappointing when the energy in a book club fizzles out like that. Sometimes, you go in with high hopes, especially with a theme as rich as finance, thinking that the conversations will be enlightening and invigorating. Instead, it can feel like a chore if the discussi...

Response: (b)

Table 43: **LightMem's case study in PersonaMem-128K.**

Question: Recently, I was at a yoga retreat.

Continued on next page

Table 43 (continued)

(a) Ah, I remember you mentioned your dislike for yoga before. It's great to revisit activities to confirm our true preferences.
(b) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons. ✓
(c) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons.
(d) Oh, you went on a yoga retreat? That sounds unique! What was it like engaging in new activities?

Ground Truth: (b) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons.

Retrieved Memories:
Long-term memory (relevant facts):
- [2026-04-29T03:11:45.721 Wed] User's original sentence was: 'This tale opens with an indistinct voice narrating something resembling a fable.'
- [2026-04-29T02:57:06.153 Wed] Each item User clicks on seems promising but often falls short of User's imagination.
- [2026-04-29T03:14:49.212 Wed] User shared knowledge developed through both academic study and practical experience.
- [2026-04-29T03:12:00.626 Wed] User asked about the Fibonacci sequence.
- [2026-04-29T02:59:24.732 Wed] Having a space to slow down and articulate dating and connection moments feels invaluable to user.
- [2026-04-29T02:51:45.747 Wed] User mentioned black holes.
- [2026-04-29T03:15:40.891 Wed] Current user is female.
[. . . 3 additional lines omitted. . .]

Response: (b)

Table 44: **MemoryOS's case study in PersonaMem-128K.**

Question: I joined a game with some friends last week.

(a) I remember you mentioned enjoying individual sports like tennis before. It sounds like a nice way to spend some time together.
(b) I remember you mentioned not enjoying team sports like soccer before. It sounds like a nice way to spend some time together.
(c) It's great to hear about new experiences, though I don't recall us discussing a preference for team sports. I hope you had a fun time.
(d) I remember you mentioned enjoying team sports like soccer before. It sounds like a nice way to spend some time together. ✓

Ground Truth: (d) I remember you mentioned enjoying team sports like soccer before. It sounds like a nice way to spend some time together.

Retrieved Memories:
[Mid-term – retrieved cross-episode pages]
- [obs] [assistant] Understanding emotional availability in a person, whether in romantic, familial, or friendship contexts, is crucial for building healthy and fulfilling relationships. Emotional availability refers to the ability and willingness to be open, present, and responsive to another pe...

1. **Consistent Communication**
- **Open and Honest Conversations**: They engage in open and honest conversations, sharing their thoughts and feelings appropriately.
- **Willingness to Listen**: They actively listen, showing empathy and understanding, which indicates they value and respect your perspective.
- **Consistency**: They maintain communication consistently rather than sending mixed messages. Their words align with their actions.

[. . . 509 additional lines omitted. . .]

Response: (d)

Table 45: **MemoryBank's case study in PersonaMem-128K.**

Question: Recently, I was at a yoga retreat.

(a) Ah, I remember you mentioned your dislike for yoga before. It's great to revisit activities to confirm our true preferences.
(b) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons. ✓
(c) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons.
(d) Oh, you went on a yoga retreat? That sounds unique! What was it like engaging in new activities?

Ground Truth: (b) Ah, I remember you mentioned your appreciation for yoga deepened during that retreat. It's great to have experiences that expand our horizons.

Retrieved Memories:
MemoryBank context

Most relevant past memories (from tasks: personamem-mc):
Conversation content on personamem-mc:
[|User|]: Observation: [assistant] Assistant: Attending live sporting events always adds an extra layer of excitement, doesn't it? There's something utterly unique about being part of a thrumming crowd that collectively gasps, cheers, or groans, creating a symphony of sound that greatly enhances ...
Conversation content on personamem-mc:
[|User|]: Observation: [assistant] Assistant: That's exciting! What made you want to get back into collecting? Your experience at the collector's fair sounds incredibly fulfilling, as it often takes such immersive environments to truly inspire creativity and rekindle old passions. It's fascinatin...
Conversation content on personamem-mc:
[. . . 15 additional lines omitted. . .]

Response: (b)

Table 46: **A-MEM's case study in PersonaMem-128K.**

Question: I joined a game with some friends last week.

(a) I remember you mentioned enjoying individual sports like tennis before. It sounds like a nice way to spend some time together.
(b) I remember you mentioned not enjoying team sports like soccer before. It sounds like a nice way to spend some time together.
(c) It's great to hear about new experiences, though I don't recall us discussing a preference for team sports. I hope you had a fun time.
(d) I remember you mentioned enjoying team sports like soccer before. It sounds like a nice way to spend some time together. ✓

Ground Truth: (d) I remember you mentioned enjoying team sports like soccer before. It sounds like a nice way to spend some time together.

Retrieved Memories:
talk start time:task=personamem-mc|step=597memory content: task: answer multiple-choice questions about a long persona-grounded conversation | obs: [user] [Original_Sentence]: Overall, the narrative lacks any form of structure or excitement, turning the tale into a ramble more than anything else....
talk start time:task=personamem-mc|step=583memory content: task: answer multiple-choice questions about a long persona-grounded conversation | obs: [user] [Original_Sentence]: Anyway, to sum it up, some parts of this are about the gods, but then it suddenly shifts, and we're introduced to a hero ...
talk start time:task=personamem-mc|step=241memory content: task: answer multiple-choice questions about a long persona-grounded conversation | obs: [user] User: I feel much better knowing I'm supporting eco-friendly businesses. It's reassuring to think that my purchasing choices are contributing ...
talk start time:task=personamem-mc|step=187memory content: task: answer multiple-choice questions about a long persona-grounded conversation | obs: [user] User: This closing feels abrupt. I'd like to end on a positive note. | action: | result: memory context: The user suggests modifying an abrup...
talk start time:task=personamem-mc|step=101memory content: task: answer multiple-choice questions about a long persona-grounded conversation | obs: [system] Current user persona: Name: Maria Gonzalez
Gender Identity: Female
Racial Identity: Hispanic

Continued on next page

Table 46 (continued)

[. . . 194 additional lines omitted. . .]

Response: (d)

Table 47: **MeMP's case study in PersonaMem-128K.**

Question: I joined a game with some friends last week.

(a) I remember you mentioned enjoying individual sports like tennis before. It sounds like a nice way to spend some time together.
(b) I remember you mentioned not enjoying team sports like soccer before. It sounds like a nice way to spend some time together.
(c) It's great to hear about new experiences, though I don't recall us discussing a preference for team sports. I hope you had a fun time.
(d) I remember you mentioned enjoying team sports like soccer before. It sounds like a nice way to spend some time together. ✓

Ground Truth: (d)

Retrieved Memories:
Here are some guidelines of how to solve the similar task:
- task_name: answer multiple-choice questions about a long persona-grounded conversation
guidelines: To solve this query, begin by examining the conversation transcript containing the dialogue. Identify key themes and specific details relevant to the multiple-choice questions. Use knowledge of Maria Gonzalez's background and her stated hobbies, including her preference for team sports such as soccer, to select the answer that best reflects her previously expressed interests. Pay attention to temporal markers indicating when preferences were stated versus when they may have changed.
- task_name: recall user preferences from extended dialogue
guidelines: To recall user preferences, first retrieve all mentions of the target topic (e.g., sports, hobbies) from the stored conversation segments. Cross-reference multiple mentions to identify consistent patterns. Prioritize the most recent stated preference when earlier and later statements conflict. For sports-related queries, distinguish between watching, playing, and following as distinct preference dimensions.
- task_name: resolve conflicting user statements across sessions
guidelines: When a user's statements appear contradictory across different conversation segments, apply temporal ordering to determine the most current stance. Consider context-dependent preferences (e.g., "I enjoy soccer with friends" vs "I prefer individual workouts at the gym" may both be true in different contexts). Weight statements made with stronger emotional language or personal anecdotes more heavily than casual mentions.
- task_name: summarize persona traits from multi-session dialogue
guidelines: To build a comprehensive persona summary, aggregate observations from all available conversation sessions. Categorize traits into demographics, preferences, habits, social relationships, and aspirations. For each trait, note the supporting evidence (direct statement, implied behavior, or third-party mention). Flag any traits that evolved over time and record both the initial and current state. Ensure the summary covers at least the following domains: entertainment, food, fitness, career, relationships, and personal values.
- task_name: match user hobby to social context
guidelines: When determining which hobby or activity a user would engage in socially, retrieve their stated preferences for group activities versus solo activities. Consider the social context described in the query (friends, family, colleagues) and match it to previously expressed social preferences. Team sports indicate a preference for collaborative, social physical activity, while individual sports suggest a preference for personal challenge and self-paced engagement.
[. . . 487 additional lines omitted. . .]

Response: (d)

I.3 LoCoMo

This subsection presents one successful example per baseline on LoCoMo, drawn from the 7B model evaluation results. Case studies for Mem0, LangMem, and LightMem are in Tables 48-50; MemoryOS and MemoryBank in Tables 51-52; A-MEM and MeMP in Tables 53-54.

Table 48: **Mem0's case study in LoCoMo.**

Question: What kinds of things did Evan have broken?

(a) His bookshelf and his desk lamp.
(b) His laptop and his tablet computer.
(c) His artwork and his musical instruments.
(d) His camera and his video camera.
(e) His gaming console and his headphones.
(f) His old Prius and his new Prius. ✓
(g) His tools and his gardening equipment.
(h) His phone and his smartwatch.
(i) His television and his sound system.
(j) His vintage motorcycle and his bicycle.

Ground Truth: (f) His old Prius and his new Prius.

Retrieved Memories:
Relevant memories from past experience:
- [2026-05-04] Name is Evan
- [2026-05-04] Family motto is 'Bring it on Home'
- [2026-05-04] Reading 'The Great Gatsby'
- [2026-05-04] Watched The Godfather
- [2026-05-04] Exploring other ways to stay active and travel due to car issues
- [2026-05-04] Apologized to partner for an accident involving rose bushes
- [2026-05-04] Family means a lot to Evan
[... additional fact entries omitted ...]

Response: (f)

Table 49: **LangMem's case study in LoCoMo.**

Question: Which hobby did Sam take up in May 2023?

(a) Sam began volunteering at the local animal shelter in May 2023.
(b) Sam joined a book club in May 2023.
(c) In May 2023, Sam began a pottery class.
(d) Sam decided to try hiking in May 2023.
(e) Sam took up photography as a hobby in May 2023.
(f) painting ✓
(g) Sam started learning to play the guitar in May 2023.
(h) Gardening was Sam's new passion in May 2023.
(i) May 2023 saw Sam starting a new woodworking project.
(j) In May 2023, Sam started taking online coding courses.

Ground Truth: (f) painting

Retrieved Memories:
Relevant facts about the user from past conversation:
- Sam plans to try painting and will provide updates.
- Sam wants to know about basic supplies for painting.
- Sam is looking for suggestions for low-impact exercises.
- Sam is interested in setting goals and exercising with a partner.
- Sam is considering trying an outdoor activity to de-stress.
- Sam has experience attending art exhibits.
- Sam is willing to provide yoga tips.
[... additional FACT entries omitted ...]

Response: (f)

Table 50: **LightMem's case study in LoCoMo.**

Question: What kind of car does Evan drive?

(a) Nissan Altima

Continued on next page

Table 50 (continued)

(b) Tesla Model 3
(c) Toyota Camry
(d) Ford F-150
(e) Subaru Outback
(f) Prius ✓
(g) Hyundai Sonata
(h) Jeep Wrangler
(i) Chevrolet Malibu
(j) Honda Civic

Ground Truth: (f) Prius

Retrieved Memories:
Long-term memory (relevant facts):
- [2026-05-01T05:06 Fri] User Evan recently bought a new Prius which has broken down.
- [2026-05-01T05:05 Fri] Evan inquires about Evan's recent activities.
- [2026-05-01T05:05 Fri] Evan mentions finishing a painting that helps him relax.
- [2026-05-01T05:07 Fri] User had a memorable experience on a camping trip last summer.
- [2026-05-01T05:04 Fri] User plans to try painting and will update Evan on how it goes.
- [2026-05-01T05:05 Fri] User mentioned meeting an incredible Canadian woman during a recent trip to Canada.
[. . . additional LTM entries omitted . . .]

Response: (f) Prius

Table 51: **MemoryOS's case study in LoCoMo.**

Question: What kind of car does Evan drive?

(a) Nissan Altima
(b) Tesla Model 3
(c) Toyota Camry
(d) Ford F-150
(e) Subaru Outback
(f) Prius ✓
(g) Hyundai Sonata
(h) Jeep Wrangler
(i) Chevrolet Malibu
(j) Honda Civic

Ground Truth: (f) Prius

Retrieved Memories:
[Mid-term – retrieved cross-episode pages]
- [obs] [assistant] [SAM]: It's tough, but I'm sticking with it.
[action] [user] [EVAN]: Nice work, Sam! Proud of you sticking to it. Have you noticed any positive changes?
meta: User expressed mixed feelings about their health, feeling concerned but motivated for positive changes. Assistant encouraged taking things one step at a time; user appreciated the support. User announced starting a diet and living healthier.
- [obs] [user] [EVAN]: This is a contemporary figurative painting that I've finished few days ago, emphasizing the emotional . . .
[. . . many additional obs/action pairs omitted . . .]

Response: (f) Prius

Table 52: **MemoryBank's case study in LoCoMo.**

Question: Where has Evan been on roadtrips with his family?

(a) Boundary Waters Canoe Area
(b) New England, Adirondacks
(c) Florida Keys, Everglades
(d) Rockies, Jasper ✓
(e) California Coast, Yosemite
(f) Colorado Springs, Denver

Continued on next page

Table 52 (continued)

(g) Acadia National Park, Maine
(h) Yellowstone, Grand Teton
(i) Shenandoah, Great Smoky Mountains
(j) Olympic Peninsula, Oregon

Ground Truth: (d) Rockies, Jasper

Retrieved Memories:
MemoryBank context

Agent personality / response strategy (from past tasks):
The user exhibits supportive, empathetic, and positive traits, sharing personal experiences and responding positively to encouragement and suggestions.
Response strategy: Active listening, encouragement, sharing experiences, offering support, maintaining positivity.

Summary of your past tasks (overall_history):
[... extensive overall-history summary omitted ...]

Response: (d) Rockies, Jasper

Table 53: **A-MEM's case study in LoCoMo.**

Question: What kind of car does Evan drive?

(a) Nissan Altima
(b) Tesla Model 3
(c) Toyota Camry
(d) Ford F-150
(e) Subaru Outback
(f) Prius ✓
(g) Hyundai Sonata
(h) Jeep Wrangler
(i) Chevrolet Malibu
(j) Honda Civic

Ground Truth: (f) Prius

Retrieved Memories:
talk start time: task=conv-49 | step=78
memory content: task: LoCoMo conversation conv-49 | obs: [assistant] [SAM]: Wow, that pic is great! Do you often spend time in places like this? | action: [user] [EVAN]: Thanks! I love being in places like this – it brings back memories of road tripping in my trusty car. | result:
memory context: The conversation revolves around a photograph that evokes nostalgic memories of past road trips.
memory keywords: ['picture', 'memories', 'road trip']
memory tags: ['picture', 'memories', 'road trip', 'joy', 'motivation', 'health']
[... additional retrieved notes omitted ...]

Response: (f)

Table 54: **MeMP's case study in LoCoMo.**

Question: Which type of vacation would Evan prefer with his family, walking tours in metropolitan cities or camping trip in the outdoors?

(a) A cruise to explore tropical islands.
(b) Staying at a resort with spa treatments.
(c) camping trip in the outdoors ✓
(d) Skiing and snowboarding in the mountains.
(e) A week-long stay at a family-friendly resort.
(f) A road trip with national park visits.
(g) Relaxing beach vacation with water sports.
(h) A wildlife safari in Africa.

Continued on next page

Table 54 (continued)

(i) Exploring European cities and culture.
(j) Visiting historical landmarks and museums.

Ground Truth: (c) camping trip in the outdoors

Retrieved Memories:
(MeMP retrieves no procedural guidelines for this LoCoMo question — the conversational QA task has no analogue in the procedural-task corpus, so the agent answered from prompt knowledge alone.)

Response: (c)

I.4 LONGMEMEVAL

This subsection presents one successful example per baseline on LongMemEval, drawn from the 7B model evaluation results. Case studies for Mem0, LangMem, and LightMem are in Tables 55–57; MemoryOS and MemoryBank in Tables 58–59; A-MEM, and MeMP in Tables 60–61.

Table 55: **Mem0's case study in LongMemEval.**

Question: What health issue did I initially think was just a cold?

(a) Sinusitis
(b) Asthma
(c) Pharyngitis
(d) Mononucleosis
(e) Influenza
(f) Laryngitis
(g) Tonsillitis
(h) bronchitis ✓
(i) Rhinitis
(j) Pneumonia

Ground Truth: (h) bronchitis

Retrieved Memories:
Relevant memories from past experience:
- [2026-04-28] Recently had a bad case of bronchitis
- [2026-04-28] Dealing with health issues and wants to boost immune system
- [2026-04-28] Wants specific foods or spices to boost immune system
- [2026-04-28] Interested in immune-boosting foods and spices
- [2026-04-28] Dealing with health issues and wants to avoid overexertion
- [2026-04-28] Is concerned about avoiding allergens in food
- [2026-04-28] Asks about apps or resources to identify allergens
[...8 additional entries omitted...]

Response: h)

Table 56: **LangMem's case study in LongMemEval.**

Question: Can you suggest some accessories that would complement my current photography setup?

(a) Older model or low-end accessories that limit photography
(b) Low-resolution lenses or basic tripods without significant enhancement
(c) Third-party or subpar quality items not integrating well with the system
(d) Video-oriented gear or items from rival brands
(e) Smartphone camera add-ons rather than professional equipment
(f) Drone attachments or action camera gear incompatible with Sony setup
(g) Sony-compatible accessories or high-quality photography gear ✓
(h) Canon-compatible accessories or mid-range photography gear
(i) Nikon-related equipment or affordable items conflicting with current setup
(j) Generic accessories or budget-friendly gear without quality improvement

Ground Truth: (g) Sony-compatible accessories or high-quality photography gear

Continued on next page

Table 56 (continued)

Retrieved Memories:
[2023/05/27 (Sat) 03:28] I am looking to upgrade my camera flash.
[2023/05/27 (Sat) 03:28] My camera is Sony A7R IV.
[2023/05/25 (Thu) 18:31] The user wants to organize the family photos for easy access.
[2023/05/26 (Fri) 14:52] I'm interested in foodie spots and cultural insights.
[2023/05/25 (Thu) 18:31] The user took a large number of photos during the family reunion.
[2023/05/29 (Mon) 20:36] Planning a trip to Grand Teton National Park next month.
[... 14 additional entries omitted ...]

Response: g)

Table 57: **LightMem's case study in LongMemEval.**

Question: How many engineers do I lead when I just started my new role as Senior Software Engineer? How many engineers do I lead now?

(a) When you just started your new role as Senior Software Engineer, you led 5 engineers. Now, you lead 6 engineers.
(b) When you just started your new role as Senior Software Engineer, you led 4 engineers. Now, you lead 3 engineers.
(c) When you just started your new role as Senior Software Engineer, you led 6 engineers. Now, you lead 7 engineers.
(d) When you just started your new role as Senior Software Engineer, you led 1 engineers. Now, you lead 2 engineers.
(e) When you just started your new role as Senior Software Engineer, you led 10 engineers. Now, you lead 11 engineers.
(f) When you just started your new role as Senior Software Engineer, you led 7 engineers. Now, you lead 8 engineers.
(g) When you just started your new role as Senior Software Engineer, you led 8 engineers. Now, you lead 9 engineers.
(h) When you just started your new role as Senior Software Engineer, you led 4 engineers. Now, you lead 5 engineers ✓
(i) When you just started your new role as Senior Software Engineer, you led 3 engineers. Now, you lead 4 engineers.
(j) When you just started your new role as Senior Software Engineer, you led 2 engineers. Now, you lead 3 engineers.

Ground Truth: (h) When you just started your new role as Senior Software Engineer, you led 4 engineers. Now, you lead 5 engineers

Retrieved Memories:
Long-term memory (relevant facts):
- [2026-04-27T16:21:52 Mon] User leads a team of 4 engineers in their new role as Senior Software Engineer.
- [2026-04-27T16:23:15 Mon] User is leading a team of 5 engineers.
- [2026-04-27T16:23:10 Mon] User is a Senior Software Engineer who leads a team of five engineers.
- [2026-04-27T16:23:03 Mon] The agent has been enjoying their role as Senior Software Engineer for a while.
- [2026-04-27T16:21:44 Mon] The user plans to invite their manager Rachel to join a group of 5 people.
[...15 additional entries omitted...]

Response: h)

Table 58: **MemoryOS case study in LongMemEval.**

Question: What breed is my dog?

(a) Siberian Husky
(b) Labrador Retriever
(c) Border Collie
(d) Yorkshire Terrier
(e) Boxer
(f) French Bulldog
(g) Beagle

Continued on next page

Table 58 (continued)

(h) Dachshund
(i) German Shepherd
(j) Golden Retriever ✓

Ground Truth: (j) Golden Retriever

Retrieved Memories:
[Mid-term – retrieved cross-episode pages]
- [obs] I'm thinking of getting Max a new collar with a nice name tag. Do you have any recommendations for a good collar brand or type that would suit a Golden Retriever like Max?
- [action] A new collar with a nice name tag is a great idea! Golden Retrievers like Max deserve a comfortable, durable, and stylish collar. Here are some recommendations: Yellow Dog Design, Ruffwear, . . .
- [obs] I noticed Max has been scratching a lot lately. Could this be related to allergies?
- [action] Yes, Golden Retrievers can be prone to skin allergies. Possible causes include environmental allergens, food sensitivities, or flea allergy dermatitis. [. . .]
[. . . many additional obs/action pairs omitted . . .]

Response: j)

Table 59: **MemoryBank's case study in LongMemEval.**

Question: How many days ago did I attend a networking event?

(a) 21 days
(b) 31 days
(c) 22 days
(d) 26 days. 27 days (including the last day) is also acceptable. ✓
(e) 25 days
(f) 28 days
(g) 29 days
(h) 24 days
(i) 30 days
(j) 23 days

Ground Truth: (d) 26 days. 27 days (including the last day) is also acceptable.

Retrieved Memories:
MemoryBank context

Agent personality / response strategy (from past tasks):
Summary of User's Personality and Response Strategy
User's Personality Traits: Diligent, Patient, Detail-Oriented, Supportive.
Response Strategy: Compliance (follow detailed instructions accurately),
Efficiency (generate relevant content quickly), Adaptability, Clarification,
Verification, Optimization.

Summary of your past tasks (abbreviated):
The agent assisted with writing tasks, travel planning, networking follow-ups,
and scheduling. On 2023/05/29, the user attended a networking event at a conference.
[...extensive overall-history summary omitted...]

Response: d) 26

Table 60: **A-MEM's case study in LongMemEval.**

Question: Can you suggest some accessories that would complement my current photography setup?

(a) Older model or low-end accessories that limit photography
(b) Low-resolution lenses or basic tripods without significant enhancement
(c) Third-party or subpar quality items not integrating well with the system
(d) Video-oriented gear or items from rival brands
(e) Smartphone camera add-ons rather than professional equipment
(f) Drone attachments or action camera gear incompatible with Sony setup
(g) Sony-compatible accessories or high-quality photography gear ✓

Continued on next page

Table 60 (continued)

(h) Canon-compatible accessories or mid-range photography gear
(i) Nikon-related equipment or affordable items conflicting with current setup
(j) Generic accessories or budget-friendly gear without quality improvement

Ground Truth: (g) Sony-compatible accessories or high-quality photography gear

Retrieved Memories:
task: 2023/05/27 (Sat) 03:28 | obs: As a Sony camera user, I've been thinking about upgrading my camera bag. Do you have any recommendations for bags specifically designed for Sony cameras? | action: I'm happy to help. As a Sony camera user, there are many great camera bags: Lowepro ProTactic 450 AW II, Peak Design Everyday Backpack, ...

task: 2023/05/27 (Sat) 03:28 | obs: I am looking to upgrade my camera flash. My camera is Sony A7R IV. | action: For the Sony A7R IV, I recommend the Sony HVL-F60RM2 or the Godox V860III-S, both compatible with Sony's TTL system.

[...extensive additional conversation logs omitted...]

Response: g) The user would

Table 61: **MeMP's case study in LongMemEval.**

Question: How many projects have I completed since starting painting classes?

(a) 9
(b) 4
(c) 5 ✓
(d) 7
(e) 1
(f) 10
(g) 6
(h) 8
(i) 3
(j) 2

Ground Truth: (c) 5

Retrieved Memories:
Here are some guidelines of how to solve the similar task:
- task_name: I'm looking for some inspiration for my next painting project. I've been stuck on what to paint next. By the way, I just finished my 5th project since starting painting classes, and I'm feeling pretty accomplished!
guidelines: To solve this query, begin by congratulating the user on their recent accomplishments. Next, suggest various themes and ideas for their next painting project, such as nature escapes, cityscapes, abstract expressions.
For projects counting questions, use the stated count (5th project = 5 total).

[...additional task guidelines omitted...]

Response: c) 5

I.5 ALFWorld

This subsection presents one successful example per baseline on ALFWorld-seen, drawn from the 7B-agent evaluation results. Each table shows the task goal, the retrieved memory excerpt, the agent's action trace, and the final outcome. Case studies for Mem0, LangMem, and LightMem are in Tables 62-64; MemoryOS and MemoryBank in Tables 65-66; A-MEM and MeMP in Tables 67-68.

Table 62: **Mem0's case study in ALFWorld.**

Task: put a mug in sidetable.

Outcome: Score = 1.0, Steps = 4

Continued on next page

Table 62 (continued)

Retrieved Memories:
Relevant memories from past experience:
- [2026-05-06] Summary of agent's execution history
Task Objective: Find two kettles and place them on the dining table.
Progress Status: 100% complete.
1. go to stoveburner 3 → located kettle 2
2. take kettle 2 from stoveburner 3 → holding kettle 2
3. go to diningtable 1 → arrived; saw mug 1 among items
4. move kettle 2 to diningtable 1 → task progressed.
[... N additional fact entries omitted ...]

Action Trace:
go to diningtable 1 (**found mug 1, recalled from past trace**) → take mug 1 → go to sidetable 1 → **move mug 1 to sidetable 1** ✓

Table 63: **LangMem's case study in ALFWorld.**

Task: put a mug in sidetable.

Outcome: Score = 1.0, Steps = 7

Retrieved Memories:
Relevant facts about the user from past conversation:
- There is a keychain 1 on sidetable 1.
- There is a statue 1 on sidetable 1.
- The user is aware of a keychain 1 located on sidetable 1.
- There is a saltshaker 1 located on cabinet 6.
- The user is holding candle 1.
- There is a keychain 2 on armchair 1.
- The user has picked up pillow 1.
[... additional FACT entries omitted ...]

Action Trace:
go to sidetable 1 (keychain only) → go to sidetable 2 (no mug) → go to sidetable 1 (re-check) → go to diningtable 1 (**found mug 1**) → take mug 1 → go to sidetable 1 → **move mug 1 to sidetable 1** ✓

Table 64: **LightMem's case study in ALFWorld.**

Task: put a mug in sidetable.

Outcome: Score = 1.0, Steps = 16

Retrieved Memories:
Long-term memory (relevant facts):
- [2026-04-29T16:43 Wed] User moved cd 2 to shelf 3.
- [2026-04-28T19:53 Tue] User moves the cooled mug to the coffeemachine.
- [2026-04-29T17:08 Wed] User moved cellphone 3 to bed 1.
- [2026-04-29T16:43 Wed] Task is to put some cd on shelf.
- [2026-05-06T01:37 Wed] The agent moved plate 1 to countertop 1.
- [2026-04-29T17:08 Wed] The agent goes to the sidetable and finds a candle there, then takes it.
- [2026-04-28T05:54 Tue] User took the tissuebox 2 from the desk.
[... additional LTM facts omitted ...]

Action Trace:
go to sidetable 1 (keychain only) → go to sidetable 2 (no mug) → explore drawer 1...drawer 4 → go to diningtable 1 (**found mug 1**) → take mug 1 → go to sidetable 1 → **move mug 1 to sidetable 1** ✓

Table 65: **MemoryOS case study in ALFWorld.**

Task: cool some lettuce and put it in countertop.

Outcome: Score = 1.0, Steps = 7

Continued on next page

Table 65 (continued)

Retrieved Memories:
MTM (top-similar page; multi-summary JSON, `prompts.py:73-87`):
`[{"theme": "cool-then-place", "keywords": ["fridge", "cool", "countertop"], "content": "open fridge 1, cool item, place on countertop 1"}]`
MTM running meta-summary (`prompts.py:217-232`):
Agent has been completing kitchen "cool / heat / clean" tasks; fridge 1 is the canonical cooling appliance and countertop 1 the canonical drop-off receptacle for cooled produce.
LPM [User Private Dat] (`prompts.py:171-196`; durable env facts under the original label, top-K by cosine):
- fridge 1 contents: lettuce 1, tomato 2, apple 1
- countertop 1 role: canonical drop-off receptacle for cooled produce
- microwave 1 status: broken in this layout (`open` returns "nothing happens")
- drawer 2 contents: knife 1; drawer 3 contents: fork 1
- (If no private data found, write "None".)

LPM [Assistant Knowledge] (durable agent tips, same prompt):
- Assistant cools an item by issuing `cool X with fridge 1` after `open fridge 1`, at fridge 1
- Assistant places a cooled item by issuing `move X to <recep>` at the destination
- Assistant avoids re-opening fridge 1 once opened in the same trajectory, at fridge 1
- (If no assistant knowledge found, write "None".)

LPM `user_profile` (51-dim procedural-agent taxonomy, top reflected dims; `prompts.py:91-168` adapted, see `memory.py:268-334`):
- *[Spatial Reasoning]* Container Use Frequency (High); Compartment Opening Tendency (High); Container Closing Habit (Medium); Search Order Stability (High).
- *[Action Composition]* Property-Modification Composition (High); Pre-Condition Checking (High); Sub-Goal Decomposition (High); Final-State Verification (Medium).
- *[Reasoning Style]* Procedural-Knowledge Use (High); Memory-of-Past-Errors Use (Medium); Goal-Adherence Stability (High).
- (Dimensions not reflected in the trajectory are omitted.)

[... remaining LPM entries omitted ...]

Action Trace:
go to fridge 1 → open fridge 1 (**found lettuce 1, tomato 2, apple 1**) → take lettuce 1 from fridge 1 → **cool lettuce 1 with fridge 1** → close fridge 1 → go to countertop 1 → **move lettuce 1 to countertop 1** ✓

Table 66: **MemoryBank's case study in ALFWorld.**

Task: put a clean cloth in cabinet.

Outcome: Score = 1.0, Steps = 12

Retrieved Memories:
MemoryBank context

Agent personality / response strategy (from past tasks):
The agent is a highly methodical, task-oriented, and literal-minded state machine. It operates with robotic efficiency, strictly adhering to programmed constraints and "admissible" action sets.
[... extensive personality / response strategy summary omitted ...]

Summary of your past tasks (overall_history):
Successful Tasks: 15 tasks involving object manipulation (heating, cooling, cleaning, relocation).
[... relevant past episode summaries omitted ...]

Action Trace:
go to cabinet 1 → open cabinet 1 (saw soapbottles) → go to cabinet 2 (empty) → go to cabinet 3 (empty) → open cabinet 4 (**found cloth 1**) → take cloth 1 → go to sinkbasin 1 → clean cloth 1 → **move cloth 1 to cabinet 1** ✓

Table 67: **A-MEM's case study in ALFWorld.**

Task: cool some plate and put it in countertop.

Outcome: Score = 1.0, Steps = 8

Retrieved Memories:
talk start time: task=conv|step=1
memory content: task: cool some plate and put it in countertop. | obs: You arrive at countertop 1 ... a creditcard 3, a pan 1, a papertowelroll 1, a spatula 1, and a winebottle 1. | action: go to countertop 2 | result: ... a plate 1 ...
memory context: The user is navigating through a kitchen environment to locate a plate for the purpose of cooling it on a countertop.
memory keywords: ['plate', 'countertop', 'cooling', 'task', 'navigation']
memory tags: ['Pick-and-place workflow', 'Cooling tasking', 'Receptacle interaction', ...]
[... additional retrieved notes omitted ...]

Action Trace:
go to countertop 1 (no plate) → go to countertop 2 (**found plate 1**) → take plate 1 → go to fridge 1 → open fridge 1 → cool plate 1 with fridge 1 → go to countertop 1 → **move plate 1 to countertop 1** ✓

Table 68: **MeMP's case study in ALFWorld.**

Task: put a mug in sidetable.

Outcome: Score = 1.0, Steps = 4

Retrieved Memories:
Here are some guidelines of how to solve the similar task:
- task_name: cool some plate and put it in countertop.
guidelines: To solve this query, begin by identifying the location of the target plate by visiting various countertops or storage areas. Once the plate is located, take it and proceed to the fridge. ...
- task_name: put some pillow on armchair
guidelines: To solve this query, begin by searching the nearby furniture, such as sofas or other seating areas, to locate a pillow. Once you have identified and taken the pillow, navigate to the target armchair. Finally, place the pillow on the armchair to complete the task.
[... additional procedural guidelines omitted ...]

Action Trace:
go to diningtable 1 (**found mug 1**) → take mug 1 → go to sidetable 1 → **move mug 1 to sidetable 1** ✓