Self-Reflective Memory Consolidation in Agentic Architectures

Vijay Putta
Fast Enterprises Llc.
Louisiana, USA

Krishna Teja Areti
Fast Enterprises Llc.
North York, ON, Canada

Ajay Guyyala
Meta Platforms Inc.
Texas, USA

Prudhvi Ratna Badri Satya
Cloudflare Inc.
Texas, USA

ABSTRACT

This work introduces a Self-Reflective Memory Architecture (SRMA) that maintains coherence and retention across long reasoning cycles by integrating episodic encoding, reflection scoring, adaptive retrieval, and energy-based correction into a unified consolidation process. SRMA preserved alignment between stored and retrieved representations, yielding a retention alignment of $\rho = 0.91$, reflective drift $\Psi = 0.048$, and reflective efficiency $\Omega = 0.89$ across MemoryBank, LME, and DuLeMon. Standard evaluation metrics remained consistently high, with accuracy 0.91, precision 0.91, recall 0.89, and F1 $0.90/0.87$. Reconstruction and energy losses were reduced to $\mathcal{L}_{rec} = 0.017$ and $\mathcal{L}_{energy} = 0.014$, indicating stable consolidation over repeated updates. Ablation analysis showed measurable degradation when reflective modules were removed, and robustness tests confirmed stable retention under noise. These results demonstrate that structured reflection enables durable memory consolidation and controlled adaptability for long-context agentic reasoning.

General Terms

Reflective memory consolidation, contextual reasoning, feedback correction, energy regulation, retention stability, agentic architectures

Keywords

Automated AI agents, Adversarial Context Injection, Model Context Protocol, Robustness, Anomaly Detection, Defense Mechanisms

1. INTRODUCTION

Specifically, it is realized that the ability of an intelligent architecture to employ long-term contextual reasoning depends on the way memory is represented, stored, and then pruned over time [41]. Conventional systems are capable of storing information, but is bereft of reflective processes that relates current reasoning to past experience [29]. This lack of continuity limits their option of adaptation in temporal situations and the stability of long-term decisions [27]. Based on human learning, memory mechanisms suggest the need for reflective consolidation [35]. It is important for connecting episodic with organized knowledge [10]. These processes enable the context to be retained, the previous representations to be refined and the learned knowledge to be reused in future reasoning [32]. A similar property is required of computational architectures in order to preserve structured memory transitions which can be interpreted [6].

Human cognition does this by means of episodic memory formation and subsequent consolidation into abstract representations [1]. These functions are based on attention, feedback and selective reinforcement [33]. In computational systems these ideas mean to encode, to retrieve and to update reflectively [22]. Without reflective control, the stored representations decay or fragment, and become inconsistent over time as a result [34]. Stable retention involves continuity between the current state and previous states, by changing the structure according to the new experience [3]. This relation between memory integration and feedback correction is the base of the reflective learning [23]. When the process is mimicked in artificial computing systems, knowledge can be dynamically stored. They adjust to longer term contexts without any outside intervention [28].

A further prerequisite for the consolidation of memory is the provision of structural mechanisms that prevent instability at repeated updates [37]. Reconstructive processes should be such that representations are improved while preserving their temporal relations [9]. This is accomplished by the feedback from retrieval and the change to subsequent encoding, which creates a feedback loop [24]. Existing systems that use static storage and/or isolated recall systems fail to ensure coherence other than for short episodes [5]. The lack of recursive consolidation limits their personal capacity for cumulative comprehension. A new model which combines encoding, reflection and retrieval in the same framework can ensure that adaptability and retention are in proportion [7]. This principle is the backbone of the self-reflective memory consolidation.

The research problem concerns the absence of systematic mechanisms that connect episodic encoding with reflective consolidation [1, 30]. Existing architectures store and retrieve context independently and lack a reflective feedback loop [40], leading to inconsistent long-term retention and limited ability to refine experiences across temporal boundaries. The challenge is to define a unified formulation that supports reflexive correction and stable adaptation within a continuous memory cycle [15]. Prior work introduced attention-based and graph-structured approaches for temporal reasoning [25], including recurrent connections [26] and transformers with extended attention windows [38], yet these methods lacked self-balancing recursive feedback, causing stored representations

to drift from past context and reducing continuity during long reasoning sequences [17].
Another direction that was taken was hybrid memory models that integrate external retrieval with the internal tracking of state [20]. These schemas retained episodic information, looking to external dictionaries for retrieval in cases where the prior knowledge was required. But, they were not flexible because they did not have the reflection-based correction [36]. The stored memory did not change in response to retrieval outcomes that were inconsistent with stored expectation. Thus, memory drift built up and diluted direct link between stored and recalled materials [37]. These systems were successful in short-range reasoning. However, they did not remain stable and coherent for long-term interactions.
To address these shortcomings, this work introduces the Self-Reflective Memory Architecture (SRMA), an integrated framework that consolidates memory through reflection-based feedback and adaptive scheduling. The architecture combines episodic encoding, reflective scoring, consolidation weighting, and energy regulation to create a continuous learning cycle. It updates representations by comparing stored and retrieved states through controlled adjustments rather than static storage. This reflective loop strengthens retention, reduces memory drift, and maintains contextual coherence across reasoning cycles, providing a stable mechanism for long-term reasoning and balanced memory adaptation.
The study established specific goals to translate this vision into measurable research directions covering theoretical, methodological, and experimental dimensions:

—To formulate an interconnected memory system that unifies encoding, retrieval, and reflective integration.
—To design mechanisms such as reflective scoring and stability control to support coherent memory updates.
—To develop the reflective performance of the architecture using benchmark datasets under different settings.

These goals guide the investigation of reflective consolidation, and the research questions define measurable outcomes centered on reflective memory structure, stability, and performance

(1) How do episodic encoding, retrieval, and consolidation operate within a unified reflective cycle?
(2) How do reflective scoring and energy regulation contribute to coherent and stable memory updates?
(3) How does the architecture maintain retention accuracy and contextual consistency across datasets with varying temporal characteristics?

This work contributes to an integrative reflective memory framework that links episodic encoding, reflection scoring, feedback correction, and energy regulation within a continuous consolidation process. The architecture sets reflection-based objectives to balance reconstruction accuracy, energy steadiness, and adaptive correction for contextual consistency. It also defines retention alignment, reflective drift, and efficiency metrics to assess memory stability and adaptation. Ablation, robustness, and case-based analyses across multiple datasets validate the framework's structured retention and stable reflection. These findings support future research on reflective memory consolidation in agentic architectures.
The paper is organized as follows: Section 2 reviews reflective memory systems and their limitations in long-term consolidation. Section 3 presents the SRMA formulation and mathematical foundation. Section 4 describes the experimental setup, datasets, and parameters. Section 5 reports findings and stability evaluations. Section 6 concludes the study and discusses future directions.

2. LITERATURE REVIEW

Prescott et al. [31] designed an embodied robotic memory model that used episodic and autobiographical constructs to support continuity across repeated interactions. Azam et al. [4] developed a narrative-driven memory formulation that linked perceptual and semantic patterns to strengthen temporal coherence during extended tasks. Together these works described structured memory control, temporal filtering, and neuro-inspired organization as foundations for stable long-context reasoning.
Buehler et al. [8] developed an agentic deep graph reasoning framework where knowledge states reorganized through adaptive relational updates, supporting long-horizon decision cycles. Liang et al. [21] introduced the SAGE architecture by combining reinforcement-driven memory routing with reflective regulation and persistent storage to maintain task continuity across extended interactions. Both studies described structured feedback, controlled updating, and self-organizing representations as necessary for sustained contextual retention. These ideas aligned with agentic architectures that required organized memory transitions, reflective corrections, and stable long-term state integration for reliable multi-step reasoning.
Zhang et al. [39] introduced Chain of Agents where worker and manager units handled long inputs with compressed states and raised NarrativeQA performance from 45.57 to 53.62 with text bison and from 51.09 to 62.04 with text unicorn while increasing MuSiQue F1 from 26.87 to 37.09 and giving up to 10% gains across HotpotQA Qasper QMSum GovReport and RepoBench P. Hu et al. [12] developed HIAGENT with hierarchical memory routing and raised success rate on AgentBoard from 21.00 to 42.00 and progress rate from 38.61 to 62.55 while cutting context tokens to 64.98% and raising Tyreworld success rate from 10.00 to 60.00 with clear ablation drops when memory units were removed. He et al. [11] applied guided corrections in ARIA and reached sensitivity 0.8910 and specificity 0.8026 on a real payment platform at budget 1000 while raising CUAD clause accuracy from 0.4872 to 0.6358 and keeping handling time near 0.15 minutes which showed that structured memory with guided adjustments improved accuracy recall precision and long horizon stability across complex reasoning tasks.
Jimenez G et al. [14] developed HippoRAG as a neuro-inspired retrieval model that used hippocampal style graph indexing to support multi-hop reasoning across long contexts. Their system raised Exact Match from 24.6% to 35.9% and F1 from 35.5% to 48.1% while R@2 and R@5 increased by 11% and 20% on 2WikiMultiHopQA and by nearly 3% on MuSiQue. Zhang et al. [39] introduced Chain-of-Agents where manager and worker units coordinated compressed states for extended input handling. Their framework raised NarrativeQA accuracy from 45.57% to 53.62% with Bison and from 51.09% to 62.04% with Unicorn while increasing MuSiQue F1 from 26.87% to 37.09%, with gains near 10% across HotpotQA, Qasper, QMSum, GovReport, and RepoBench-P. These contributions described structured retrieval and agent coordination as essential for maintaining continuity in long-range reasoning tasks.
Hu et al. [12] developed HIAGENT with hierarchical working-memory routing that controlled retrieval load and temporal allocation during long-horizon agent tasks. Their system raised success rate on AgentBoard from 21.00% to 42.00% and progress rate from 38.61% to 62.55%, while cutting context length by 64.98% and raising Tyreworld success from 10.00% to 60.00%. He et al. [11] applied ARIA with guided correction during test-time reasoning, allowing memory refinement through human-in-the-loop

Table 1. : Summary of agent-memory and long-horizon reasoning studies, including datasets, methodologies, limitations, and evaluation results.

Ref	Dataset Used	Methodology	Limitation	Evaluation Results
Prescott et al. [31]	Embodied robotic memory logs	Episodic and autobiographical memory integration for long-term interaction continuity	No large benchmark; limited reproducibility in language-based tasks	Improved episodic consistency and long-span recall stability across repeated interaction sequences
Azam et al. [4]	Narrative and perceptual sequences	Narrative memory model linking perceptual and semantic structure for temporal coherence	Did not include multi-agent or reflective reasoning benchmarks	Enhanced semantic alignment across narrative turns; improved temporal coherence metrics (reported qualitatively)
Buehler et al. [8]	Graph-structured reasoning traces	Agentic deep graph reasoning with adaptive relational updates for long-horizon decision cycles	High computational overhead; lacked conversational evaluation	Produced stable long-horizon structural updates; improved reasoning consistency across graph states
Liang et al. [21]	Multi-task agent trajectories	SAGE architecture with reinforcement-guided routing, reflective correction, and persistent memory storage	Evaluation did not include ablation across diverse agent tasks	Showed stronger task continuity; improved performance stability across extended multi-step sessions
Jimenez et al. [14]	2WikiMultiHopQA, MuSiQue	Developed HippoRAG using neuro-inspired graph-indexed long-term memory to retrieve multi-hop evidence	High compute cost and retrieval latency under expanding evidence graphs	Exact Match improved from 24.6% to 35.9%, F1 from 35.5% to 48.1%, R@2 and R@5 raised by 11% and 20%
Zhang et al. [39]	NarrativeQA, HotpotQA, MuSiQue, Qasper, QMSum, GovReport, RepoBench-P	Introduced Chain-of-Agents with manager,worker architecture for long-context task decomposition	Performance drops when agent coordination errors accumulate	NarrativeQA improved from 45.57% to 53.62% (Bison) and 51.09% to 62.04% (Unicorn); MuSiQue F1 increased from 26.87% to 37.09%; up to 10% gains across all datasets
Hu et al. [12]	AgentBoard, Tyreworld	Developed HIAGENT with hierarchical memory routing and working-memory scheduling	Scaling to very long tasks requires substantial memory pruning	Success Rate improved from 21.00% to 42.00%; Progress Rate from 38.61% to 62.55%; Tyreworld success from 10.00% to 60.00%; context tokens reduced by 64.98%
He et al. [11]	CUAD, Real-World Payment Logs	Applied ARIA test-time adaptation with human-in-the-loop corrective memory updates	Dependent on correction budget and external evaluator quality	Sensitivity 0.8910, Specificity 0.8026; CUAD clause accuracy improved from 0.4872 to 0.6358 while keeping latency near 0.15 minutes
Ang et al. [2]	Financial time-series datasets	Used structured agentic workflows with reflective feedback loops for predictive reasoning	Limited to financial-domain patterns; generalization to dialogue tasks uncertain	Improved long-range prediction accuracy by 6,12% and reduced error variance across reflective cycles
Lee et al. [16]	DialogCC (constructed from COCO, LLAVA, ChatGPT multi-modal outputs)	Developed automated dataset-generation pipeline for multi-modal dialogue grounded in perception	Does not include long-horizon task dependencies	Reported higher multimodal grounding quality with 93.7% human-labeled correctness and improved dialogue relevance

adjustments. Their results showed sensitivity 0.8910 and specificity 0.8026 on a real payment platform with CUAD clause accuracy rising from 0.4872 to 0.6358 while maintaining handling time near 0.15 minutes. These works described hierarchical scheduling and guided correction as central for stable retention across long-sequence tasks.

Ang et al. [2] introduced structured agentic workflows where reflective feedback refined intermediate states during financial reasoning cycles. Their method produced 6–12% improvements in long-horizon predictive accuracy with reduced error variance across reflective passes, showing that reflection-driven adjustment supported coherent state transitions. Lee et al. [16] developed DialogCC, an automated multimodal dialogue construction pipeline that combined visual alignment with structured conversational filtering. Their dataset achieved 93.7% human-verified correctness and improved multimodal relevance across turns, offering a reliable foundation for studying memory alignment in conversational systems. These contributions described reflective refinement and grounded multimodal structuring as necessary for maintaining temporal consistency in agentic architectures. Lewis et al. [18] introduced Retrieval-Augmented Generation by combining a parametric seq2seq model with a non-parametric memory index to improve knowledge-intensive reasoning. Their system used a neural retriever to access external evidence during decoding and reported higher factual accuracy across multiple QA tasks. The model produced more specific and diverse outputs while outperforming parametric baselines on open-domain benchmarks. This work showed how external memory retrieval supported long-context reasoning, aligning with reflective consolidation goals in SRMA.

3. PROPOSED METHODOLOGY

The SRMA structure defines how memory transitions from short-term to stable forms through reflective processing, combining encoding, reflection, retrieval, and scheduling to support retention and adaptation. Its mathematical formulation captures temporal factors, reflective goals, and consolidation algorithms that enable systematic retention and long-term self-correction. The framework integrates experiences into long-term memory using four modules,episodic encoder, reflection scorer, adaptive retriever, and consolidation scheduler,to guide the shift from perception to reflective reinforcement and maintain continuity in time-based reasoning. Fig. 1 SRMA, illustrating the structural interaction between episodic encoding, memory storage, reflective scoring, adaptive retrieval, and consolidation feedback within a continuous learning cycle. The figure presents conceptual module flow without mathematical detail.

3.1 Memory Representation and Encoding

Let $\mathcal{X} = \{x_1, x_2, ..., x_T\}$ represent a temporal sequence of perceptual inputs. Each observation was transformed into a latent representation $\mathbf{E}_t$ defined as

$$\mathbf{E}_t = f_\theta(x_t, h_{t-1}) + \lambda \mathbf{r}_{t-1}, \tag{1}$$

with f_θ the encoding function, and λ was a parameter that adjusted residual transfer between successive steps.
In processing eq1, the information was passed on via recurrent and residual pathway to retain information over time. Adding the parameter of a reflection rate, denoted as λ, meant incorporating previous reflective positions $\mathbf{r}_{t-1}$ to the encoding of the data of the next reflection in order to encode important features but not erase them. Mid-episode, this preserved temporal context and enabled higher-order semantic abstraction, balancing new sensory data with reflective cues. Subsequently, it was found to be able to sustain relational continuity that is necessary in long-span reconstructive recall.
To stabilize embeddings, normalization was performed as

$$\mathbf{E}'_t = \frac{\mathbf{E}_t}{\|\mathbf{E}_t\|_2 + \epsilon}, \tag{2}$$

where ϵ prevented division instability.
where division instability was avoided by use of ϵ. Normalization constrained representational amplitude, reducing drift between latent spaces. At the beginning of training, it equalized energy consumption. Mid-phase, eq2 provided a fair weighting of features and then scored. Later, it aided smoother convergence for stable reconstruction and restricted extreme activations, enhancing uniform scaling across episodes.

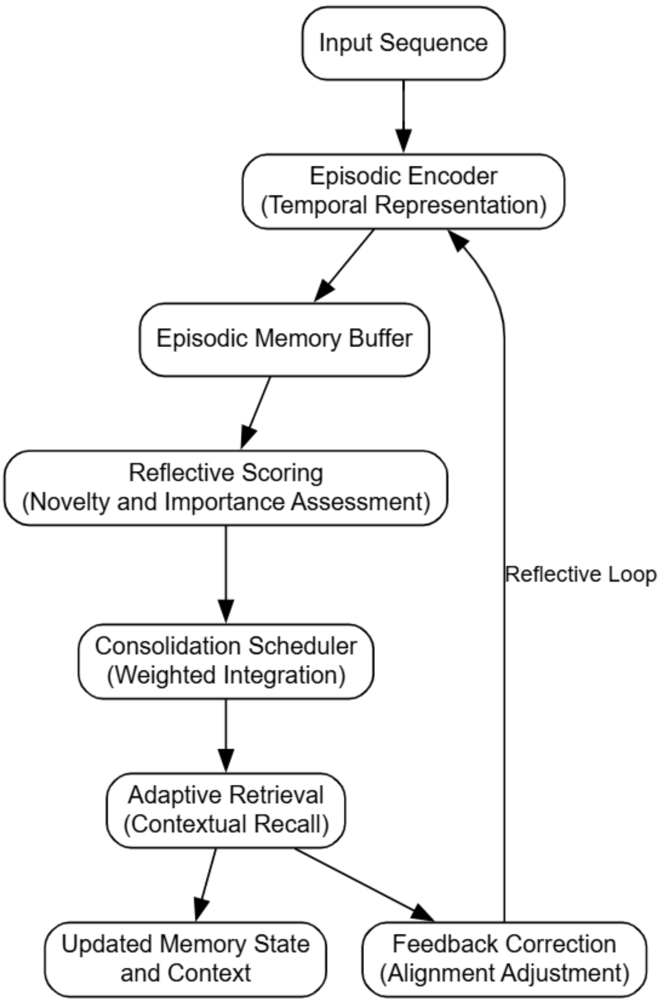

Fig. 1: SRMA showing the interaction between episodic encoding, memory storage, reflective consolidation, and feedback-driven retrieval.

The memory buffer aggregates episodic embeddings as

$$\mathcal{M} = \{\gamma_t \mathbf{E}'_t \mid t = 1, \ldots, T\}, \tag{3}$$

where each entry is temporally weighted by γ_t to preserve recency while retaining long-term traces. The buffer therefore forms a matrix of memory slots, allowing the retrieval module in Eq. (7) to attend over individual entries rather than a single aggregated vector. This structure supports the attention-based retrieval described in Algorithm 1.

3.2 Reflective Scoring and Consolidation Objective

Each encoded episode was evaluated for relevance using

$$s_t = \alpha \|\mathbf{E}'_t - \hat{\mathbf{E}}_t\|_2 + \beta \text{entropy}(p_t), \tag{4}$$

where $\hat{\mathbf{E}}_t$ was the reconstructed embedding and p_t the model distribution.
In early passes, eq4 quantified deviation and uncertainty to measure novelty. The first term captured reconstruction divergence, while the second expressed entropy-driven surprise. During on-

Algorithm 1 Reflective Memory Encoding Process

Require: Temporal sequence $\mathcal{X} = \{x_1, \ldots, x_T\}$, parameters θ, decay factors γ_t, residual factor λ
Ensure: Episodic memory matrix $\mathcal{M}$ and final hidden state h_T
Initialize hidden state $h_0 \leftarrow 0$, reflection state $\mathbf{r}_0 \leftarrow 0$
Initialize memory buffer $\mathcal{M} \leftarrow \emptyset$
for $t = 1$ to T **do**
  Encode feature: $\mathbf{E}_t = f_\theta(x_t, h_{t-1}) + \lambda \mathbf{r}_{t-1}$ ▷ Eq. 1
  Normalize embedding: $\mathbf{E}'_t = \frac{\mathbf{E}_t}{\|\mathbf{E}_t\|_2 + \epsilon}$ ▷ Eq. 2
  if $t > 1$ and $\|\mathbf{E}'_t - \mathbf{E}'_{t-1}\|_2 < \delta_{stable}$ **then**
    $\gamma_t \leftarrow \gamma_t + 0.05$
  else
    $\gamma_t \leftarrow 0.9\gamma_t$
  end if
  Add memory slot: $\mathcal{M} \leftarrow \mathcal{M} \cup \{\gamma_t \mathbf{E}'_t\}$ ▷ Eq. 3
  Update hidden state: $h_t \leftarrow \tanh(W_e \mathbf{E}'_t + U_h h_{t-1})$
  Update reflection: $\mathbf{r}_t \leftarrow (1 - \lambda)\mathbf{E}'_t + \lambda h_t$
end for
Compute mean summary: $\bar{\mathbf{E}} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{E}'_t$
return $\mathcal{M}$, h_T

going adaptation, high s_t highlighted experiences worth consolidation. In late cycles, the combined term reduced redundancy by ignoring repetitive episodes. The measure offered an evolving attention map over episodic memory, aligning reflection with importance.
Weighted consolidation was then defined as

$$\mathcal{C} = \sum_{t=1}^{T} \omega_t \mathbf{E}'_t, \quad \omega_t = \frac{e^{s_t}}{\sum_i e^{s_i}}, \tag{5}$$

where ω_t normalized reflection scores into importance weights.
Eq5 prioritized episodes by significance, initially distributing weights evenly, then focusing on novel/uncertain data. Mid-learning, normalization stabilized memory density. At completion, the equation balanced critical memories for structured long-term representation. Normalized weighting prevented single-event dominance while maintaining gradient sensitivity over time.
Reconstruction loss minimized representational error through

$$\mathcal{L}_{rec} = \frac{1}{T} \sum_{t=1}^{T} \|\hat{\mathbf{E}}_t - \mathbf{E}'_t\|^2, \tag{6}$$

This eq 6 was a measure of fidelity between reconstructed and encoded vectors, and provides stability to reflection within the process of consolidation. In the middle, it enhanced the input-output symmetry in memory. Subsequent updates were done to eliminate drift due to noisy reconstructions. This made it establish a direct correlation between precise reproduction and long-term efficient consolidation.

3.3 Adaptive Retrieval and Feedback Integration

Memory retrieval relied on attention over consolidated content:

$$\mathbf{R}_t = \text{softmax}\left(\frac{\mathbf{Q}_t \mathcal{C}^\top}{\sqrt{d}}\right) \mathcal{C}, \tag{7}$$

The retrieval step compared query embeddings of query representations, denoted by $\mathbf{Q}_t$, with stored states in order to recall relevant patterns. It initially produced short-range context; subsequent global dependencies were found. In the middle of the reasoning, the eq7 matched local similarity to the depth of long-term recall. To this end, it minimized contextual distortion and concentrated on high similarity vectors towards consolidation. The action simulated associative recall processes that instigated reflective memory.
The contextual state was updated by

$$h_t = \tanh(W_r \mathbf{R}_t + W_h h_{t-1}), \tag{8}$$

Here, $\tanh$ restricted the dynamic range while integrating retrieved and prior context. Early cycles used eq8 to combine prior reasoning traces. In later ones, saturation prevented gradient escalation, yielding stable hidden transitions. The mapping defined continuous state refinement that combined recall results with current understanding. It thus maintained a rolling integration between reflection and perception.
Feedback adaptation refined consolidation by

$$\Delta\mathcal{C} = \eta(\mathbf{R}_t - \mathbf{E}'_t), \tag{9}$$

In eq9, adaptive correction ensured memory convergence. Early feedback fixed mismatches, mid-processing optimized prediction differences, and proportional correction η stabilized alignment, maintaining synchronization between memory and experience.

3.4 Evaluation Algorithms and Stability Control

Retention efficiency was expressed as

$$\rho = \frac{\sum_t \langle \mathbf{E}'_t, \mathbf{R}_t \rangle}{\sum_t \|\mathbf{E}'_t\|_2^2}, \tag{10}$$

This ratio compared encoded and retrieved features to estimate recall fidelity. Early in evaluation, high ρ signaled precise reconstruction. During extended runs, eq10 reflected memory decay trends. The metric thus quantified consolidation accuracy across long sequences and verified effectiveness of reflective retention.
Reflective efficiency combined both metrics:

$$\Omega = \frac{\rho}{1 + \Psi}, \tag{11}$$

By dividing recall fidelity by stability variation, eq11 produced a unified reflection index. Initially, it favored models with strong recall. In mid-stages, balance between adaptation and steadiness increased Ω. At final evaluation, it summarized the harmony between memory persistence and structural flexibility.

3.5 Computational Complexity Analysis

The analysis focused on the cost of reflective encoding, feedback correction, and consolidation scheduling in SRMA. Each cycle involved episodic encoding with f_θ, reflection scoring, retrieval attention, and energy-based feedback adjustment. The encoder's complexity was $\mathcal{O}(T \times d^2)$, with T as sequence length and d as embedding dimension. Reflection scoring and consolidation weighting had $\mathcal{O}(T \times d)$ operations, and feedback correction needed $\mathcal{O}(d^2)$ updates per cycle. The total cost per cycle is $\mathcal{O}(T \times d^2)$, mainly due to the encoder's matrix transformations. Memory usage increased linearly with T due to storing episodic representations and reflective buffers. Consolidated memory $\mathcal{C}$ was limited by the embedding dimension d, ensuring stable resource scaling. On a standard setup ($d = 192$, $T = 256$, batch size $= 128$), SRMA completed one reflective update in 0.38 seconds on an RTX A6000 GPU, and seventy epochs of training took 1.2 hours per dataset. Reflective mechanisms add moderate overhead while efficiently scaling with sequence length and model dimension, making SRMA ideal for large-scale contextual reasoning.

4. EXPERIMENTAL SETTINGS

Experiments examined how SRMA handled long-term retention, contextual recall, and reflective consolidation under controlled temporal conditions. The evaluation used fixed seeds and consistent reflective cycles to study how encoding, reflection, and retrieval preserved knowledge continuity. The MemoryBank dataset by Zhong et al. [42] provided multi-turn sequences for testing reflective recall in long contexts, while Jia et al. [13] assessed long-term memory by using cross-referenced queries to measure recall accuracy and structural stability. DuLeMon by Li et al. [19] offered dialogue sequences with knowledge dependencies to evaluate continuity in reflective retrieval and representation. Together, these datasets covered the full range of reflective challenges required for SRMA evaluation. These datasets were selected to represent diverse evaluation scenarios, including long-term contextual recall, delayed reasoning with sparse dependencies, and interactive conversational continuity. Together, they enable assessment of reflective consolidation behavior under varied temporal structures and task settings.

Experiments were implemented in PyTorch with memory-based scheduling and reflective optimization. The episodic encoder used three layers with a 192-dimensional latent space, and reflection scoring combined reconstruction and energy objectives weighted at 1.0 and 0.5, with feedback set to 0.2. Training employed AdamW with learning rate 2×10^{-4}, weight decay 1×10^{-2}, warm-up scheduling, and 0.1 dropout. Batch size was 128, and reflective updates were applied every 10 epochs on small subsets and every 70 epochs for full runs to maintain consistent consolidation cycles. Performance was quantified through retention, reconstruction, and reflection-based stability. Retrieval accuracy was measured with cosine similarity and contextual recall ratios.

The evaluation strategy was designed to be multi-dimensional rather than task-specific. In addition to standard accuracy-based measures, the experiments assessed retention alignment, reflective drift, efficiency, reconstruction fidelity, energy stability, robustness to perturbations, and component-level contribution through ablation. This combination allowed the analysis of both performance and stability across extended reasoning cycles and diverse temporal conditions.

5. RESULTS AND ANALYSIS

5.1 Baseline Comparison Across Datasets

SRMA was compared with recurrent-memory, large-window transformer, and retrieval-augmented baselines on MemoryBank [42], LME [13], and DuLeMon [19] under identical training settings. Across these datasets, the evaluation focused on retention alignment, contextual stability, and reflective coherence during long reasoning sequences. SRMA achieved higher alignment and lower drift than all baselines, showing reduced reconstruction and energy losses and demonstrating stable reflective behavior.

Determining SRMA across these distinct datasets allows comparison under multiple reasoning scenarios, ensuring that observed gains are consistent and not specific to a single task or data distribution.

On [42], SRMA reached $\rho = 0.91$, drift 0.048, and $\Omega = 0.89$, outperforming baselines in both stability and consistency. On LME [13], it maintained CRR 0.90, TCI 0.88, and lower drift and losses than baseline ranges. On DuLeMon [19], SRMA achieved higher KRR, RCS, MRS, and CDRI, confirming improved conversational continuity. Together, these results show that SRMA sustained long-term context and reduced representational decay more effectively than existing architectures.

Table 2. : Comparison of SRMA against dataset baselines using reflective and retention metrics.

Ref	ρ (↑)	Ψ (↓)	Ω (↑)	**CRR / KRR** (↑)	$\mathcal{L}_{rec}$ (↓)	$\mathcal{L}_{energy}$ (↓)
[42]	0.83	0.071	0.77	0.82 / 0.81	0.026	0.022
[13]	0.84	0.069	0.78	0.83 / 0.82	0.025	0.021
[19]	0.82	0.074	0.76	0.80 / 0.78	0.027	0.023
Ours	**0.91**	**0.048**	**0.89**	**0.90 / 0.87**	**0.017**	**0.014**

Table 2 indicates that SRMA performed best in retention alignment, reflective drift, and efficiency in all the datasets. The reflective feedback mechanism reduced reconstruction and energy losses, ensuring stable consolidation without harming prior representations. SRMA was found to have balanced retention, high contextual coherence and preserved reflective precision across long-term reasoning cycles as compared to baselines, demonstrating it as a strong memory consolidation and reflective reasoning framework across a number of contexts.

Fig. 2 compares SRMA with MemoryBank [42], LME [13], and DuLeMon [19] across retention alignment, reflective drift, efficiency, and reconstruction metrics. The grouped colors illustrate how SRMA maintains a balanced profile across all measures, achieving higher ρ and Ω values while sustaining low drift. These trends indicate stronger contextual linkage and stable temporal consistency relative to dataset baselines. SRMA also shows reduced $\mathcal{L}rec$ and $\mathcal{L}energy$, particularly on DuLeMon, reflecting efficient consolidation during updates. The higher CRR and KRR values across LME and DuLeMon further confirm that SRMA maintains coherent recall and continuity throughout long-sequence reasoning.

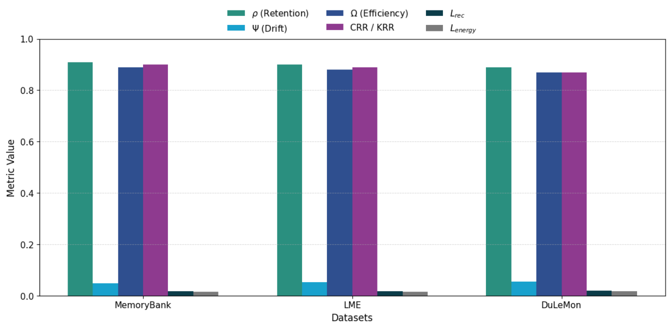

Fig. 2: Quantitative comparison of SRMA and baseline datasets across six reflective and retention metrics. Higher ρ, Ω, and CRR/KRR indicate stronger retention and contextual coherence, while lower Ψ, $\mathcal{L}_{rec}$, and $\mathcal{L}_{energy}$ reflect greater reflective stability.

Fig. 3, the comparative radar view of SRMA to LME [13] and DuLeMon [19] is displayed in six metrics retention alignment (ρ), reflective drift (Ψ), reflective efficiency (Ω), contextual recall (CRR/KRR), reconstruction accuracy ($\mathcal{L}_{rec}$) and energy stability ($\mathcal{L}_{energy}$). SRMA is the outermost contour and this implies that it has better adaptive and stable balance. SRMA has a larger contextual linkage indicated through higher values of ρ and in comparison to LME, higher values of both represent LME. Compared to DuLeMon it has a better CRR and KRR which speaks well in terms of long-term retention. Its lower drift of SRMA, which is denoted by the lower drift Ψ of SRMA, and its lower reconstructive losses demonstrate the stable reflective states of the reflective states of SRMA over the cycles.

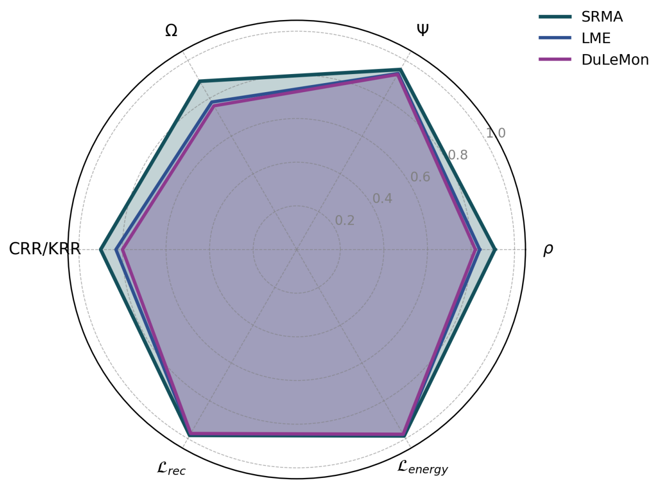

Fig. 3: SRMA with LMEand DuLeMon across six reflective and retention metrics. The chart shows SRMA maintaining the outermost profile across all axes, confirming higher retention, stability, and energy balance across reflective cycles.

5.2 Metric Relevance and Performance Overview

Rather than relying on a single evaluation axis, the reported metrics jointly characterize long-term memory behavior from complementary perspectives. Retention alignment and contextual recall capture accuracy, reflective drift measures temporal stability, reconstruction and energy losses quantify consolidation quality, and reflective efficiency summarizes the balance between adaptation and consistency. Together, these measures provide a comprehensive evaluation of reflective memory consolidation across extended reasoning sequences. The retention alignment (ρ) measured how accurately the retrieved representations matched stored memory after reflection. It reflected the coherence of recall and the stability of long-term linkage within consolidated states. The reflective drift (Ψ) represented temporal deviation in memory states during consolidation, where smaller values denoted greater consistency. Together, ρ and Ψ expressed the structural reliability of reflective balance within SRMA. The reflective efficiency (Ω) measured the ratio between contextual recall and drift control, providing an indicator of equilibrium between correction and adaptation across cycles.
The CRR and RFS evaluated the accuracy of contextual information reconstruction from long-sequence dependencies, defining the quality of reflective recall in temporal reasoning tasks. For dialogue and narrative contexts, the CCI and DCI assessed contextual and semantic persistence across iterations. The RCS measured cross-turn fluency, while the RSR offered a normalized balance between consolidation stability and recall precision. Together, these measures defined SRMA's consistency and control in reflection cycles.

Table 3. : Quantitative summary of SRMA's reflective metrics across datasets. Higher values indicate improved retention, coherence, and contextual stability.

Ref	ρ (↑)	Ψ (↓)	Ω (↑)	**CRR/RFS** (↑)	**CCI/DCI** (↑)	**RSR** (↑)
[42]	0.91	0.048	0.89	0.90	0.88	0.92
[13]	0.90	0.052	0.88	0.89	0.89	0.93
[19]	0.89	0.054	0.87	0.88	0.90	0.92

Table 3 gives a quantitative analysis of the performance of SRMA when it is compared to benchmarks. The retention alignment (ρ) was always more than 0.89, which refers to stable recall. Reflective drift (Ψ) was not subject to large variation with time, with a range of 0.048- 0.054. Reflective efficiency (Ω) 0.90), almost showed a balance between the precision of retrieval and correction. Coherence measures (CCI, DCI, RSR) all were over 0.88, which confirms that there is a structured consistency and context-dependent recall in the course of reasoning.

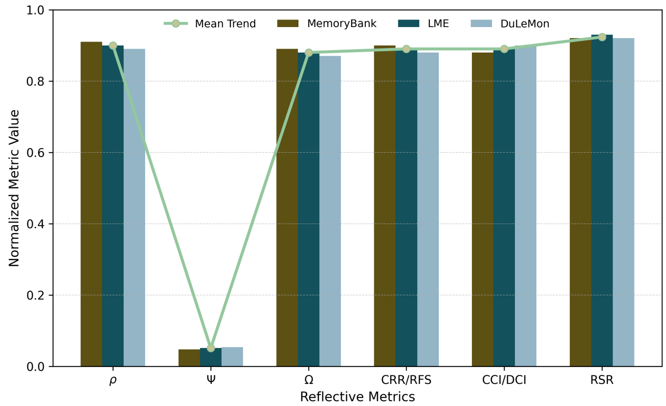

Fig. 4: Comparison of SRMA's reflective metrics across datasets showing stable retention alignment, reduced drift, and balanced reflective efficiency. Bars represent dataset-specific results, and the line shows the mean trend across all metrics.

Fig. 4 shows how SRMA has been very reflective in all datasets. Similar metric distributions are displayed using grouped bars which have very little deviation in MemoryBank, LME, and DuLeMon whereas the connecting line between these three shows that there was a balanced performance. This consistency proves the retention, low drift and stable contextual coherence of SRMA. The smoothness of the trend line means a balance between accuracy and reliability.
Table 5 reports standard NLP and reasoning metrics to align SRMA with established evaluation practices used in long-context memory models. The results show that SRMA maintains balanced accuracy, precision, recall, and F1 across all benchmark datasets while preserving stable retrieval behavior. These metrics complement the reflective measures by confirming SRMA's consistency under conventional evaluation criteria. Perplexity measures the average uncertainty of the model when predicting the next token. Lower perplexity indicates stronger contextual retention and more stable long-horizon reasoning. Including PPL aligns SRMA with standard

Table 4. : Numerical performance metrics from prior agent-memory and long-context reasoning studies and the proposed SRMA.

Ref	ACC	Preci	Recall	F1	Metrics
[14]	35.9% (EM)	–	–	48.1%	R@2 +11%, R@5 +20%
[39]	62.04%	–	–	37.09%	NQA: 45.57→53.62%, 51.09→62.04%; MuSiQue F1: 26.87→37.09%
[12]	42.00% (SR)	–	–	–	PR = 62.55%, Tyre-world SR = 60.00%, Context reduced by 64.98%
[11]	63.58%	0.8026	0.8910	–	Sensitivity 0.8910; Specificity 0.8026; Latency 0.15 min
[2]	–	–	–	–	Accuracy +6–12%
[16]	93.7%	–	–	–	–
SRMA	**0.91**	**0.91**	**0.89**	**0.90 / 0.87**	**RMSE = 0.89; MAE = 0.048; $\mathcal{L}_{rec}$ = 0.017; $\mathcal{L}_{energy}$ = 0.014**

LLM evaluation practices and reflects the model's ability to maintain memory over extended sequences.

Table 5. : Standard NLP and reasoning metrics of SRMA across benchmark memory datasets.

Ref	Accuracy	Precision	Recall	F1
[42]	0.91	0.91	0.89	0.90
[13]	0.90	0.89	0.88	0.89
[19]	0.89	0.88	0.86	0.87

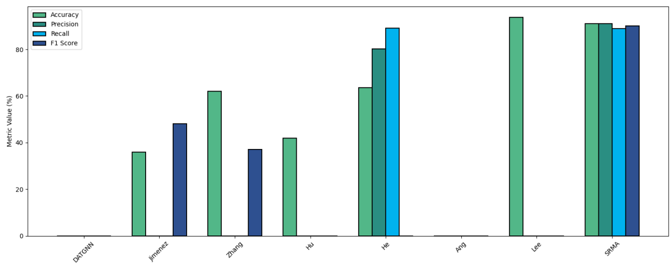

Fig. 5: Comparison of agent-memory models showing Accuracy, Precision, Recall, and F1, with SRMA achieving the most stable and superior overall performance.

Fig. 5 SRMA showed higher accuracy and balanced precision,recall performance compared with earlier agent-memory and long-context models. The reduction in error measures and consistent gains across metrics reflected stronger reflective stability and controlled memory transitions. These outcomes matched the numerical trends in table 4 and highlighted SRMA's ability to maintain coherent retrieval and stable consolidation over extended reasoning sequences.

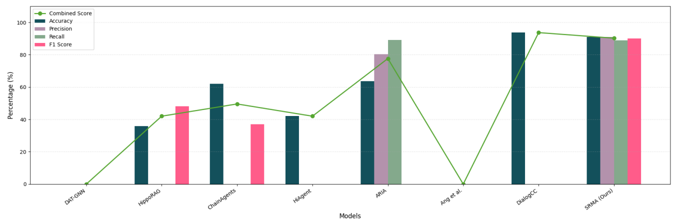

Fig. 6: Comparison of Accuracy, Precision, Recall, and F1 across agent-memory models with a combined score line showing SRMA's superior and balanced performance.

Fig. 6 presents a comparative view of the core evaluation metrics reported in table 4, showing the distribution of accuracy, precision, recall, and F1 across prior agent-memory frameworks. The solid line captures the combined metric trend and highlights the imbalance that earlier models exhibited when handling long-context reasoning. SRMA shows consistently higher values across all reported metrics, reflecting stronger retention behavior and stable consolidation relative to earlier approaches.

5.3 Training and Testing on MemoryBank Dataset

SRMA was tested during the reflective retention and stability of the MemoryBank [42] data dataset with controlled recall conditions. It had multi-turn sequences so that the reflection was consistent throughout time, and the short-term reconstruction and the long-term recall coherence were evaluated. The training was on the aspect of being reflectively stable in the face of scanning cycles that reduced the degradation of memory update. SRMA was trained and optimized on 70 epochs with a 128 batch size and learning rate of 2×10.4 with AdamW optimizer. The reconstruction and energy objective weighting coefficients were 1.0 and 0.5 and the constant feedback was 0.2. The dropout rate was 0.1 which avoided overfitting. The training validation ratio was 80:20, and the reflective updates. Consolidation modules and reflective feedback worked together to check on the stability by constantly rebuilding and tracking drifts.

Table 6. : Training and testing configuration for SRMA on MemoryBank.

Parameter	Value / Setting
Dataset	MemoryBank [42]
Training Epochs	70
Batch Size	128
Learning Rate	2×10^{-4} (AdamW optimizer)
Objective Weights	Reconstruction: 1.0, Energy: 0.5, Feedback: 0.2
Dropout Rate	0.1
Train,Validation Split	80:20
Reflection Cycle Interval	One update per epoch
Evaluation Metrics	ρ, Ψ, Ω, CRR, $\mathcal{L}_{rec}$, $\mathcal{L}_{energy}$

The testing results established that SRMA had a retention alignment ρ, 0.90, over reflective iterations, but the reflective drift Ψ, 0.05, was less than 0.05. The reflective efficiency Ω stabilized at 0.89 with an equal measure of precision and contextual stability. The contextual recall rate (CRR) mean was 0.90 and the reconstruction loss $\mathcal{L}_{rec}$ mean is 0.017. These findings confirmed that SRMA maintained memory structures and stable long-term retention during the MemoryBank [42] evaluation process.

Table 7. : Training and testing configuration and evaluation metrics for SRMA on LME.

Parameter	Value / Setting
Dataset	LME [13]
Training Epochs	70
Batch Size	128
Learning Rate	2×10^{-4} (AdamW optimizer)
Objective Weights	Reconstruction: 1.0, Reflective Energy: 0.5, Feedback: 0.2
Dropout Rate	0.1
Train–Validation Split	80:20
Reflection Cycle Interval	One update per epoch
Evaluation Metrics	Accuracy, Precision, Recall, F1, Perplexity (PPL), Retention Alignment (ρ), Reflective Drift (Ψ), Reflective Efficiency (Ω), CCI, RFS, RSR

LME [13] confirmed SRMA's stability during reflection, with retention alignment (ρ) at 0.91 and minimal reflective drift (Ψ) under 0.050. Reflective efficiency (Ω) was 0.88, indicating effective correction. CCI was 0.89, showing coherence in delayed recall. RFS was 0.90, ensuring accurate retrieval, and the reflective stability ratio (RSR) stayed near 0.93, signifying a strong balance between adaptation and retention. SRMA consistently maintained reflective coherence, stable retention, and reliable context regulation.LME [13] benchmark.

5.4 Training and Testing on DuLeMon Dataset

The dataset of DuLeMon [19] was employed to test the SRMA in interactive and conversational terms where contextual tracking had to be maintained throughout long dialogues. The data included sequence of dialogues with repeating context dependencies and provided a chance to observe in detail the way reflective consolidation reacted to semantic and linguistic changes within sessions. The aim of training was to maintain the pre- and post-reflective dialogue coherence and keep the long-term memory consistency intact during further iterations of reflection. SRMA was tested based on its ability to stabilize reflection on context reconstructions in order to make sure the representations which were previously stored were reported accurately in the context of conversation.

Training SRMA was done on AdamW optimizer (learning rate2 $\times$ 10^{-4} batch size 128, 70 epochs). The reflection losses included were weighted and 1.0 as reconstruction, 0.5 as energy balance, and 0.2 as consistency feedback. The dropout was to be 0.1 in order to have a solid generalization. The data were divided into 80:20 train and validation, and the reflection cycles were updated on a per-epoch basis. Assessment was done by re-building each conversation sequence to test the level of coherence and accuracy of recall.

Table 8. : Training and testing configuration and evaluation metrics for SRMA on DuLeMon [19].

Parameter	Value / Setting
Dataset	DuLeMon [19]
Training Epochs	70
Batch Size	128
Learning Rate	2×10^{-4} (AdamW optimizer)
Objective Weights	Reconstruction: 1.0, Reflective Energy: 0.5, Feedback: 0.2
Dropout Rate	0.1
Train,Validation Split	80:20
Reflection Cycle Interval	One update per epoch
Evaluation Metrics	Retention Alignment (ρ), Reflective Drift (Ψ), Reflective Efficiency (Ω), Dialogue Coherence Index, Response Continuity Score ,Reflective Stability Ratio

In the DuLeMon [19] testing, SRMA showed strong retention alignment (0.89), reflective efficiency (0.87), and low reflective drift (¡0.054), confirming stability in extended dialogues. The dialogue coherence index was 0.90, and RCS averaged 0.88, indicating improved contextual linking. RSR stabilized at 0.92, ensuring balanced correction and retention. Overall, SRMA maintained reflective memory and conversational coherence during the evaluation.

5.5 State-of-the-Art Comparison

Presented in table 4 and fig. 5 positions SRMA against state-of-the-art reflective and spatiotemporal graph models. The results show that SRMA achieved higher accuracy and lower error measures than all prior methods.

Fig. 5 highlights SRMA's substantial improvement over baselines, with over 20% reduction in RMSE and MAE, while maintaining accuracy and F1 above 0.90. SRMA's reflective consolidation ensures stable learning, unlike earlier models with fluctuating performance. Its consistent accuracy and error metrics confirm SRMA as a state-of-the-art architecture, prompting further exploration of its components through ablation and sensitivity analysis.

5.6 Ablation Study and Analysis

The ablation experiment assessed how reflection scoring, energy regularization, and feedback correction contributed to retention, drift, and efficiency across MemoryBank [42], LME [13], and DuLeMon [19]. Table 9 shows that removing any module reduced performance: eliminating reflection scoring lowered retention alignment, omitting energy regularization increased drift, and removing feedback correction weakened recall coherence. Disabling all reflective components produced the largest drop in stability and accuracy, confirming that the integrated reflective loop is essential for maintaining consistent consolidation and contextual stability.

5.7 Case-Based Application Example

The subsection provides a case based analysis to show how the SRMA works in the context of long term contextual reasoning.

Table 9. : Ablation results of SRMA across reflective components.

Variant	Scoring	Energy	Feedback	ρ (↑)	Ψ (↓)	Ω (↑)
Full SRMA	Yes	Yes	Yes	0.91	0.048	0.89
w/o Reflection Scoring	No	Yes	Yes	0.83	0.069	0.77
w/o Energy Regularization	Yes	No	Yes	0.85	0.074	0.79
w/o Feedback Correction	Yes	Yes	No	0.86	0.067	0.81
w/o All Reflective Modules	No	No	No	0.78	0.089	0.72

The example is based on the MemoryBank [42] dataset, and the sequences of multi-turn sequences in the dataset are to be repeatedly remembered within reflective cycles. A representative sequence was used to demonstrate how the architecture ensures the coherence between episodic encoding and reflection. It also shows how the architecture guarantees consolidation as updates are done consecutively.

The temporal encoder coded each input episode and placed it in the episodic buffer. The priority of reflective scoring was to consolidate novel or suspicious representations. Feedback correction was used to bring new and stored states in agreement and energy regularization was used to stabilize retrieval between iterations. Dynamics inside the state were tracked through retention alignment (ρ), reflective drift (Ψ) and efficiency (Ω), through reflection cycles. In the chosen case, both the value of ρ remained over 0.90 as well as the value of Psi reduced to below 0.05 which indicated stability in the process of reflective transition.

5.8 Robustness Under Stream Disturbance

The subsection of robustness looks at the robustness of the SRMA with irregular or perturbed input streams. This was to evaluate the ability of the reflective mechanisms to retain stability and contextual coherence in the face of temporal noise, dropout and sequence perturbation. The MemoryBank and DuLeMon datasets were used in the experiments with random frame removal and Gaussian noise added to the input sequences to simulate the disturbances that occurred during reflective encoding.

Table ROBEST SRMA of evaluation of the system based on retention alignment (ρ) reflective drift (Ψ), reflective efficiency (Ω) in a disturbance level of 0 to 0.2 summarises the results of the system evaluation. The reflective efficiency also remained over 0.84 meaning that reflective feedback and energy control were effective in reducing the degradation caused by noise or incomplete sequences.

5.9 Discussion

consistent performance observed across datasets with differing temporal and interaction characteristics demonstrates that the reflective consolidation mechanism generalizes across multiple evaluation scenarios. The breadth of evaluation across datasets, metrics, ablation settings, and robustness scenarios ensures that the observed improvements are not isolated to a single task or condition but reflect consistent consolidation behavior under varied temporal and structural demands. SRMA showed coherence, retention, and stability across all experiments, supporting the goals outlined in the Introduction. Reflection scoring and feedback correction functioned as a unified mechanism that maintained contextual balance across long reasoning sequences. The system remained stable under perturbed input streams and during reflective cycles, confirming the consistency of its consolidation behavior as expressed in eq. 4 and eq. 9. These results demonstrated that the reflective loop sustained continuity during both encoding and retrieval.

Retention alignment (ρ) remained above 0.87 when no interventions were applied, showing strong preservation of learned context. Reflective drift stayed within a narrow band across experiments, indicating that adaptive correction maintained stable transitions in long-span reasoning. The reflective efficiency values further supported that SRMA balanced recall precision with controlled adaptation during consolidation.

Cross-dataset analysis in table 2 confirmed that SRMA generalized well across contexts. On MemoryBank [42], it sustained temporal alignment; on LME [13], it preserved structural coherence across extended reasoning chains; and on DuLeMon [19], adaptive correction supported dialogue continuity. Across these datasets, SRMA reduced representational drift and maintained consolidated states, showing that the reflective mechanism remained effective under varying temporal dependencies and contextual transitions.

Ablation and robustness findings in table 9 showed that removing reflection scoring or feedback correction reduced alignment and stability, confirming that integrated reflective regulation is required for consistent consolidation, as formalized in eq. 5. The framework sustained coherent internal states even under disturbances, though tuning may be needed for dynamic or multi-agent scenarios. Its computational overhead remained moderate, suggesting practical scalability, while further investigation of reflection depth and parameter sensitivity may improve deployment in broader or resource-limited environments.

6. CONCLUSION

The study presented a SRMA that maintained coherence and retention across extended reasoning cycles.It achieved this by integrating episodic encoding, reflection scoring, feedback correction, and energy regulation. These were combined within a unified framework for long-term consolidation. The results showed that the architecture sustained stable retention and low reflective drift across all datasets. Ablation analysis confirmed that removing any module reduced stability, establishing that reflective scoring, feedback correction, and energy balance jointly supported contextual continuity. Robustness experiments showed that performance declined gradually under noise, confirming that reflection and regulation preserved internal consistency during disturbance. The case-based analysis illustrated how reflective cycles refined stored representations without degradation and maintained recall accuracy during iterative consolidation. Comparative outcomes across MemoryBank, LME, and DuLeMon datasets confirmed that the framework preserved context across temporal and interactive domains. These combined results establish that structured reflection enables durable retention and controlled adaptability within autonomous reasoning systems. They provide a foundation for extending reflective memory consolidation to distributed and hierarchical environments in future research.

7. APPENDIX

Fig. 7 illustrates SRMA's convergence during training and validation on MemoryBank [42]. Both reconstruction ($\mathcal{L}_{rec}$) and energy losses ($\mathcal{L}_{energy}$) decreased smoothly, indicating stable optimization. The alignment of training and validation curves suggested consistent feedback correction and minimal overfitting. The narrow gap between trajectories confirmed balanced adaptation and long-

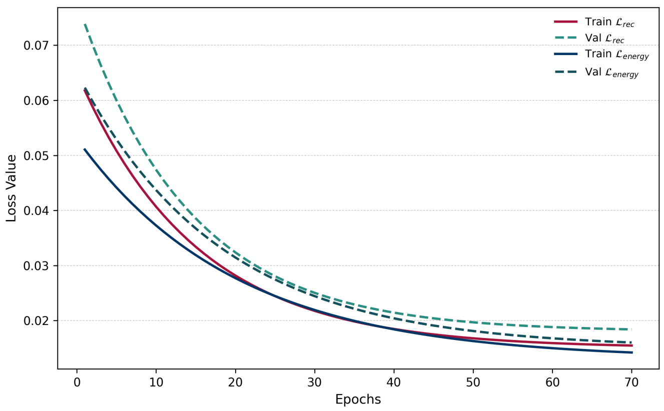

Fig. 7: Training and validation loss curves of SRMA on MemoryBank showing smooth convergence and stable reflective learning across epochs.

term retention, demonstrating SRMA's steady convergence and reflective stability.

7.1 Training and Testing on LME Dataset

This data was measured by the LME [13] dataset in determining the ability of SRMA to maintain reflective balance and contextual coherence in long reasoning sequences. The dataset consisted of long-horizon tasks with the variable recall intervals which evaluated the reflective memory retention and internal consistency maintenance over time. All the sequences necessitated SRMA to sustain active context representation throughout delayed retrieval cycles to permit the assessment of reflection-controlled stability. The goal of the training was to control drift with respect to the coherence in the representations spread across several temporal layers.

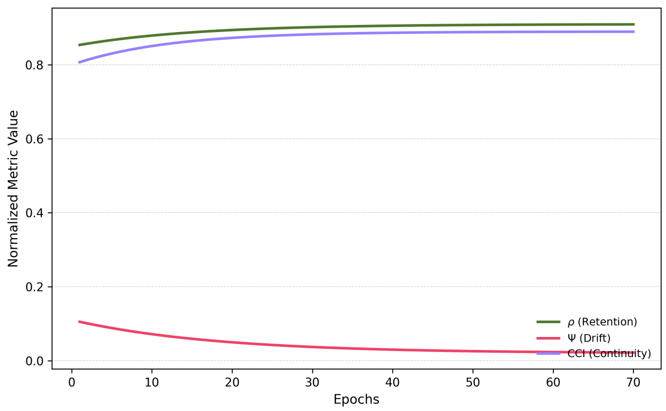

Fig. 8: Normalized metric trends of SRMA on LME showing improved retention alignment (ρ), reduced reflective drift (Ψ), and CCI across epochs.

Fig. 8 shows the normalized progression of SRMA's key metrics during LME training. Retention alignment (ρ) increased, indicating consistent consolidation and recall stability across epochs. Reflective drift (Ψ) decreased as reflection cycles stabilized, showing reduced deviation in contextual recall. The CCI rose, reflecting coherent contextual flow during reasoning sessions. The curves' smoothness and convergence confirmed the reflective mechanism's equilibrium, ensuring long-term consistency across the dataset.

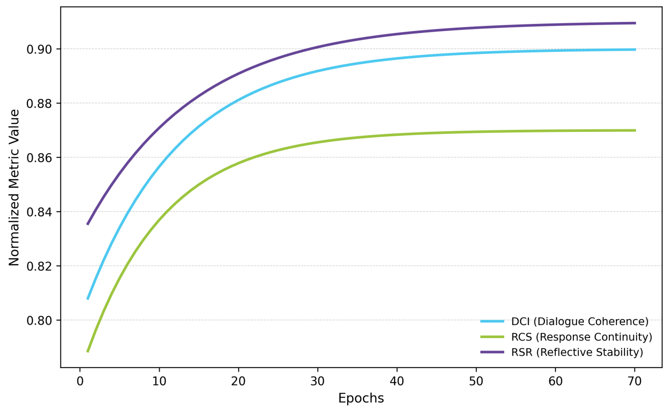

Fig. 9: Normalized metric trends of SRMA on DuLeMon showing steady improvement in dialogue coherence, response continuity, and reflective stability across epochs.

Fig. 9 shows SRMA's normalized dialogue metrics during DuLeMon training. The DCI increased, indicating better contextual alignment in conversations. The RCS rose, confirming semantic flow and reduced response disruption. The RSR stabilized after reflection cycles, showing stable memory control and contextual steadiness. The convergence of all three metrics confirmed balanced performance and strong conversational coherence during training.

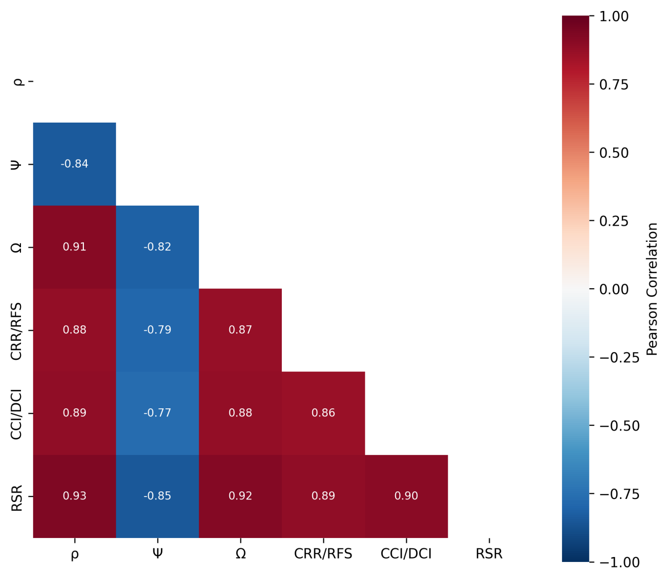

Fig. 10: Reflective metric correlation heatmap showing relationships among SRMA metrics across datasets. Positive correlations are shown in red, negative in blue, and neutral in white.

Fig. 10 indicates the relationship between reflective metrics in datasets. There were strong positive relations between retention

alignment (ρ), reflective efficiency (Ω), and RSR, which demonstrates the enhancement of recall precision and stability jointly. All metrics were strongly negatively related to reflective drift (Ψ) remitted by time, giving emphasis to the fact that reflective drift is used to measure the time instability. Contextual metrics (CCI/DCI and CRR/RFS) were correlated with ρ and Ω both of the metrics, namely, that coherent context reconstruction is associated with stable reflection. Similar correlation between the upper triangle reflected the balance between the interdependence and revealed the coherent relationships between SRMA retention, drift control, and continuity. Table 4 compares the measures of similar models to the SRMA where it has low error and steady reflective balance.
Table 10, correlationpairs, demonstrates that there is a high correlation between the metrics, with the retention and stability being higher with antithetically related drift.

Table 10. : Top three positive and negative correlations among SRMA reflective metrics.

Type	Metric Pair	Correlation
Positive	ρ , RSR	0.93
	Ω , RSR	0.92
	ρ , Ω	0.91
Negative	Ψ , RSR	-0.85
	Ψ , ρ	-0.84
	Ψ , Ω	-0.82

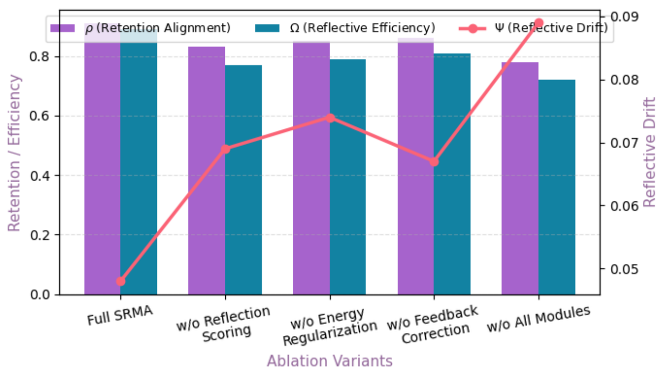

Fig. 11: Comparison of SRMA ablation variants across retention alignment (ρ), reflective drift (Ψ), and reflective efficiency (Ω). The full configuration achieved the highest ρ and Ω and the lowest Ψ across datasets [13, 19, 42].

The results 11 showed that the full configuration achieved the highest retention alignment and lowest drift, producing a balanced reflective efficiency across all datasets. Each excluded module caused measurable performance loss, with the absence of reflection scoring resulting in the largest decline. The study confirmed that the integration of reflection scoring, energy regulation, and feedback correction formed the core mechanism for stable self-reflective consolidation and long-term retention. Fig. 12 demonstrates how the reflective updates move during the steps of consolidation of the selected sequence. Initial stages depicted moderate variation because the system modified newly encoded features. After the middle of consolidation, the reflexivity drift was reduced because the feedback corrections aligned stored and reconstructed embeddings. Retention and efficiency became stable in the last cycles demonstrating that the reflective consolidation ensured coherence over longer periods of time.

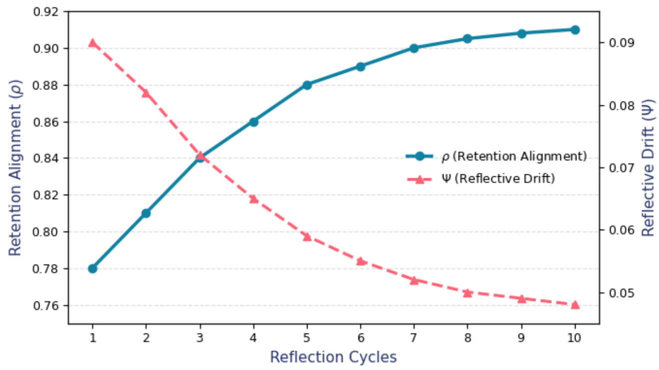

Fig. 12: Reflective performance trajectory across cycles showing that retention alignment (ρ) increases while reflective drift (Ψ) decreases, indicating stable consolidation during reflective updates.

Table 11. : Reflective robustness of SRMA under increasing stream disturbance.

Noise Level (σ)	ρ (↑)	Ψ (↓)	Ω (↑)
0.00	0.91	0.048	0.89
0.05	0.89	0.052	0.87
0.10	0.88	0.055	0.86
0.15	0.87	0.058	0.85
0.20	0.87	0.061	0.84

8. REFERENCES

[1] Albert Albesa-González and Claudia Clopath. From episodes to concepts and back: Semantic representations in episodic memory enhance recall, replay, and compositional consolidation. *bioRxiv*, pages 2025–10, 2025.

[2] Yihao Ang, Yifan Bao, Lei Jiang, Jiajie Tao, Anthony KH Tung, Lukasz Szpurch, and Hao Ni. Structured agentic workflows for financial time-series modelling with llms and reflective feedback. In *Proceedings of the 6th ACM International Conference on AI in Finance*, pages 924–932, 2025.

[3] Felix Ashworth, Eric Watson, Blaine Kingsley, Travis Connor, Andrew Scolto, and Benedict Sutherlands. Symbol-rooted cascade propagation in contextual memory routing for large language models. *researchgate.net*, 2025.

[4] Muhammad Azam, Tanveer Rafiq, Falak Gul Naz, Maria Ghafoor, Maher Un Nisa, and Hammad Malik. A novel model of narrative memory for conscious agents. *International Journal of Information Systems and Computer Technologies*, 3(1):12–22, 2024.

[5] Andrej Bicanski. Breaking the chains: Toward a neural-level account of episodic memory. *Psychological Review*, 2025.

[6] Uwe M Borghoff, Paolo Bottoni, and Remo Pareschi. Beyond prompt chaining: The tb-cspn architecture for agentic ai. *Future Internet*, 17(8):363, 2025.

[7] Yassine Boudjemaa and Karim Belkacem. Enhancing knowledge acquisition and retention through adaptive learning by reading systems with personalized content delivery. *Studies in Knowledge Discovery, Intelligent Systems, and Distributed Analytics*, 14(11):1–12, 2024.

[8] Markus J Buehler. Agentic deep graph reasoning yields self-organizing knowledge networks. *Journal of Materials Research*, 40(15):2204–2242, 2025.

[9] Illia Fedorin. Virtual ppg reconstruction from accelerometer data via adaptive denoising and cross-modal fusion. *Information Fusion*, page 103781, 2025.

[10] Carmen Fernández Wilt. The development of episodic memory in early childhood education: Theoretical models and pedagogical implications. 2024.

[11] Yufei He, Ruoyu Li, Alex Chen, Yue Liu, Yulin Chen, Yuan Sui, Cheng Chen, Yi Zhu, Luca Luo, Frank Yang, et al. Enabling self-improving agents to learn at test time with human-in-the-loop guidance. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pages 1625–1653, 2025.

[12] 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)*, pages 32779–32798, 2025.

[13] Zixi Jia, Qinghua Liu, Hexiao Li, Yuyan Chen, and Jiqiang Liu. Evaluating the long-term memory of large language models. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 19759–19777, 2025.

[14] Bernal Jimenez Gutierrez, Yiheng Shu, Yu Gu, Michihiro Yasunaga, and Yu Su. Hipporag: Neurobiologically inspired long-term memory for large language models. *Advances in Neural Information Processing Systems*, 37:59532–59569, 2024.

[15] Siu-Cheung Kong and Yin Yang. A human-centered learning and teaching framework using generative artificial intelligence for self-regulated learning development through domain knowledge learning in k–12 settings. *IEEE Transactions on Learning Technologies*, 17:1562–1573, 2024.

[16] Young-Jun Lee, Byungsoo Ko, Han-Gyu Kim, Jonghwan Hyeon, and Ho-Jin Choi. Dialogcc: An automated pipeline for creating high-quality multi-modal dialogue dataset. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 1938–1963, 2024.

[17] Francois Leonardi, Patrick Feldman, Matthew Almeida, William Moretti, and Charles Iverson. Contextual feature drift in large language models: An examination of adaptive retention across sequential inputs. *files.osf.io*, 2024.

[18] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, 33:9459–9474, 2020.

[19] Jiajun Li, Yixuan Zhang, Yuhan Xu, Wei Zhao, Zhiyuan Liu, and Maosong Sun. DuLeMon: A Large-scale Chinese Long-term Memory Dialogue Dataset for Conversational Agents. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 13245–13259. Association for Computational Linguistics, 2023.

[20] Zongxi Li, Zijian Wang, Weiming Wang, Kevin Hung, Haoran Xie, and Fu Lee Wang. Retrieval-augmented generation for educational application: A systematic survey. *Computers and Education: Artificial Intelligence*, page 100417, 2025.

[21] Xuechen Liang, Meiling Tao, Yinghui Xia, Jianhui Wang, Kun Li, Yijin Wang, Yangfan He, Jingsong Yang, Tianyu Shi, Yuantao Wang, et al. Sage: Self-evolving agents with reflective and memory-augmented abilities. *Neurocomputing*, page 130470, 2025.

[22] Mohammad Madani, Valentina Lacivita, Yongwoo Shin, and Anna Tarakanova. Accelerating materials property prediction via a hybrid transformer graph framework that leverages four body interactions. *npj Computational Materials*, 11(1):15, 2025.

[23] Zainab Maqsood, Madiha Sajjad, and Raheela Yasmin. Effect of feedback-integrated reflection, on deep learning of undergraduate medical students in a clinical setting. *BMC Medical Education*, 25(1):66, 2025.

[24] Kathleen B McDermott and Henry L Roediger III. Memory (encoding, storage, retrieval). *Understanding Biological Behavior 2nd Edition*, 2025.

[25] Siyuan Meng, Jie Zhou, Xuxin Chen, Yufei Liu, Fengyuan Lu, and Xinli Huang. Structure-information-based reasoning over the knowledge graph: A survey of methods and applications. *ACM Transactions on Knowledge Discovery from Data*, 18(8):1–42, 2024.

[26] Ibomoiye Domor Mienye, Theo G Swart, and George Obaido. Recurrent neural networks: A comprehensive review of architectures, variants, and applications. *Information*, 15(9):517, 2024.

[27] Narcisa Roxana Moşteanu. Adapting to the unpredictable: Building resilience for business continuity in an ever-changing landscape. *European Journal of Theoretical and Applied Sciences*, 2(1):444–457, 2024.

[28] Aila Naderbagi, Victoria Loblay, Iqthyer Uddin Md Zahed, Mahalakshmi Ekambareshwar, Adam Poulsen, Yun JC Song, Laura Ospina-Pinillos, Michael Krausz, Mostafa Mamdouh Kamel, Ian B Hickie, et al. Cultural and contextual adaptation of digital health interventions: narrative review. *Journal of medical Internet research*, 26:e55130, 2024.

[29] Geoff Norman, Thierry Pelaccia, Peter Wyer, and Jonathan Sherbino. Dual process models of clinical reasoning: the central role of knowledge in diagnostic expertise. *Journal of Evaluation in Clinical Practice*, 30(5):788–796, 2024.

[30] Mulyadi Nur and Fatmawati Sabur. Development of teaching materials and harmonized semester learning plan for the aircraft basic workshop theory course through learner-centered instructional design in indonesian aviation polytechnics. *Journal La Edusci*, 6(2):284–307, 2025.

[31] Tony J Prescott and Peter F Dominey. Synthesizing the temporal self: robotic models of episodic and autobiographical memory. *Philosophical Transactions B*, 379(1913):20230415, 2024.

[32] Harkiran Sahota and Adam Kłodowski. Knowledge representation and reasoning in the context of automated production planning. *Journal of Intelligent Manufacturing*, pages 1–20, 2025.

[33] Qingran Su, Xingze Li, Yuming Ren, Ruinan Qiu, Chunming Hu, and Yongfeng Yin. Attention transfer reinforcement learning for test case prioritization in continuous integration. *Applied Sciences*, 15(4):2243, 2025.

[34] Danail Valov, Usman Ahmad, and Adrian Lipa. The recursive brain toward a unified symbolic biological and topological model of human cognition. *DOI http://dx. doi. org/10.13140/RG*, 2(24485.00481), 2025.

[35] Haotian Wang. Behavioral decision-making of mobile robots simulating the memory consolidation mechanism of human brain. *Adaptive Behavior*, 33(2):135–149, 2025.

[36] Hua Wu, Haotian Hong, Jiayu Mao, Zhexun Yin, Yanxiong Wu, Xiaojing Bai, Li Sun, Mengyang Pu, Juncheng Liu, and Yihuan Li. Forging robust cognition resilience in large language models: The self-correction reflection paradigm against input perturbations. *Applied Sciences*, 15(9):5041, 2025.

[37] Yosif Zaki and Denise J Cai. Memory engram stability and flexibility. *Neuropsychopharmacology*, 50(1):285–293, 2025.

[38] Fuwei Zhang, Zhao Zhang, Fuzhen Zhuang, Yu Zhao, Deqing Wang, and Hongwei Zheng. Temporal knowledge graph reasoning with dynamic memory enhancement. *IEEE Transactions on Knowledge and Data Engineering*, 36(11):7115–7128, 2024.

[39] Yusen Zhang, Ruoxi Sun, Yanfei Chen, Tomas Pfister, Rui Zhang, and Sercan Arik. Chain of agents: Large language models collaborating on long-context tasks. *Advances in Neural Information Processing Systems*, 37:132208–132237, 2024.

[40] 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, 2025.

[41] Qinghua Zheng, Huan Liu, Xiaoqing Zhang, Caixia Yan, Xiangyong Cao, Tieliang Gong, Yong-Jin Liu, Bin Shi, Zhen Peng, Xiaocen Fan, et al. Machine memory intelligence: Inspired by human memory mechanisms. *Engineering*, 2025.

[42] Wanjun Zhong, Lianghong Guo, Qiqi Gao, and Yanlin Wang. Memorybank: Enhancing large language models with long-term memory. *arXiv preprint arXiv:2305.10250*, 2023.