Tailoring Memory Granularity for Multi-Hop Reasoning over Long Contexts

Peijun Qing **Xingjian Diao** **Chiyu Ma**
Saeed Hassanpour **Soroush Vosoughi**†
Department of Computer Science, Dartmouth College
{peijun.qing.gr, soroush.vosoughi}@dartmouth.edu

Abstract

Multi-hop reasoning over long contexts remains challenging, as it requires integrating relevant contexts scattered across distant sources while resisting semantic drift and noise from distracting content. While retrieval-augmented generation (RAG) has emerged as the prevailing solution, most RAG approaches encode and store context in monolithic memory representations, resulting in noisy retrieval and brittle reasoning. To overcome these limitations, we introduce TAG (Tailoring Memory Granularity), a framework that prestructures memory into diverse granularities and employs a reward-guided navigator to adaptively compose hybrid memory tailored to each query. The navigator is trained with a multi-objective Bradley–Terry loss that learns the relative utility of different memory types, enabling dynamic routing across granularities. This design allows RAG systems to balance fine-grained detail with high-level abstraction, yielding more reliable reasoning. Extensive experiments on long-context multi-hop question answering (QA) benchmarks show that TAG achieves state-of-the-art performance. With only 0.033% additional parameters, it remains lightweight, highlighting its practicality as a scalable and effective solution for enhancing language model agents in complex, real-world scenarios.

1 Introduction

Large Language Models (LLMs) are increasingly augmented with agentic capabilities and personas, enabling them to interact with environments, perform planning and reasoning, utilize tools, and autonomously accomplish complex goals via supervised finetuning and reinforcement learning (Wang et al., 2024c; Xi et al., 2025; Yuan et al., 2025b,a; Bi et al., 2025; Tian et al., 2025; Zhou et al., 2025b; Li et al., 2024a; Li and Deng, 2023). A cornerstone of such sophisticated agency is the memory module, which controls how agents process, store, and retrieve past information to inform future actions (Zhang et al., 2024; Lee et al., 2024; Diao et al., 2025c; Hu et al., 2026). This is particularly critical for long-context reasoning tasks, where relevant information may be distributed across vast textual passages, demanding nuanced understanding at multiple memory granularities—from fine-grained factual details to global semantics (Li et al., 2024b; Lee et al., 2024).

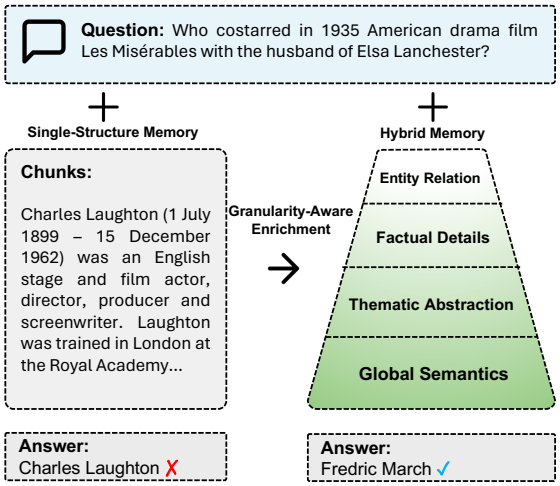

Figure 1: Hybrid memory system of LLM-based agents offers rich, granular text information.

Previous LLM-based agents often rely on singular memory structure, typically text chunks (Hu et al., 2024; Packer et al., 2023). While straightforward, such uni-modal memory can lead to inefficient retrieval and the inclusion of noisy or irrelevant information in the agent's context window, potentially impairing reasoning, a phenomenon referred to as being *"lost in the middle"* (Liu et al., 2023; Wu et al., 2025).

Cognitive Fit Theory, inspired by cognitive science (Vessey, 1991; Umanath and Vessey, 1994), suggests that human cognitive processes are optimized when the representation of information aligns with task requirements. For example, while

†Corresponding authors.

segmented text maintains local context, knowledge triples prove superior for tasks that necessitate clearly defined relationships (Anokhin et al., 2024). As a result, recent progress has utilized LLMs to create a range of knowledge structures (Li et al., 2023; Jain et al., 2024; Li et al., 2024d) and to implement a hybrid memory system (Zeng et al., 2024), aiming to find the best structural representations for a variety of tasks in complex real-world scenarios.

However, a critical challenge remains: different memory structures inherently convey different granularities of information, and existing approaches often fail to adaptively leverage the optimal blend of these varied structural representations based on the specific demands of an incoming query, as shown in Figure 1. In this work, we introduce TAG, a novel framework that integrates multiple, structurally diverse memory types and employs a reward-guided retrieval mechanism to adaptively build the optimal memory composition for each query. TAG constructs hybrid memory with stratified information granularity, including document chunks for local contextual details, atomic facts for precise factual units, knowledge triples for explicit entity relationships, and summaries for abstractive understanding.

The central component of TAG is a memory router module, based on a multi-objective reward model, which predicts the relative preference of each memory structure for a given query. To achieve this, we propose a multi-object Bradley-Terry loss, enabling the reward model to learn a memory blending coefficient that guides a weighted retrieval strategy. This allows our LLM agent to dynamically assemble a granularly balanced and task-relevant hybrid memory, enhancing retrieval-augmented generation. We evaluate TAG on three long-context multi-hop reasoning benchmarks. Our method achieves state-of-the-art performance and outperforms the uni-structure remarkably, by up to 7% on the HotPotQA benchmark. Additionally, our lightweight framework only adds 0.033% additional parameters, highlighting its practicality as a scalable and effective solution for enhancing LLM agents in complex, real-world scenarios.

2 Related Work

RAG for long-context multi-hop reasoning. Multi-hop QA over long contexts is commonly tackled with retrieval-augmented generation (RAG), but performance remains brittle due to dispersed evidence and distractors. Prior work improves robustness via iterative retrieval and reasoning—e.g., question decomposition and interleaving retrieval with intermediate reasoning steps (Press et al., 2023; Trivedi et al., 2023)—or via better indexing and retrieval units under long-context constraints, including hierarchical abstractions and coarse-to-fine retrieval (Lewis et al., 2020a; Zhao et al., 2024). Analyses further show that retrieval effectiveness is highly sensitive to *granularity*, with no single unit universally optimal (Chen et al., 2024).

Structured and hybrid memory for LLM agents. LLM agents typically store experiences as raw chunks in a vector store (Lewis et al., 2020b; Packer et al., 2023), which can introduce noise and exacerbate "lost-in-the-middle" failures (Liu et al., 2023). To increase information density and controllability, recent systems construct semi-structured memories (e.g., summaries, atomic facts) or structured memories (e.g., triples/graphs) (Xu et al., 2023; Min et al., 2023; Li et al., 2024b; Anokhin et al., 2024; Baek et al., 2023; Sun et al., 2024). Hybrid memory systems combine multiple forms to trade off precision and context (Zeng et al., 2024).

Existing methods largely *fix* the retrieval unit or operate within a single representation at a time (even when hierarchical), and hybrid systems often mix structures without *query-adaptive* allocation. We instead study *per-query composition* across heterogeneous memory granularities, training a lightweight reward-guided router to dynamically allocate retrieval budget over multiple memory structures. See Appendix A for a broader discussion.

3 Method

Figure 2 illustrates the overview of our method. Section 3.1 describes the generation of each memory structure. Section 3.2 explains the pipeline of our augmented retrieval process.

3.1 Memory Generation

Structural memory generation equips agents with the ability to transform raw textual documents D_q into structured representations M_q, thereby enhancing the storage, retrieval, and reasoning capabilities of LLM-based agents. Following prior work (Zeng et al., 2024), we build our memory system using four representative structures with increasing levels of abstraction and granularity: knowledge triples

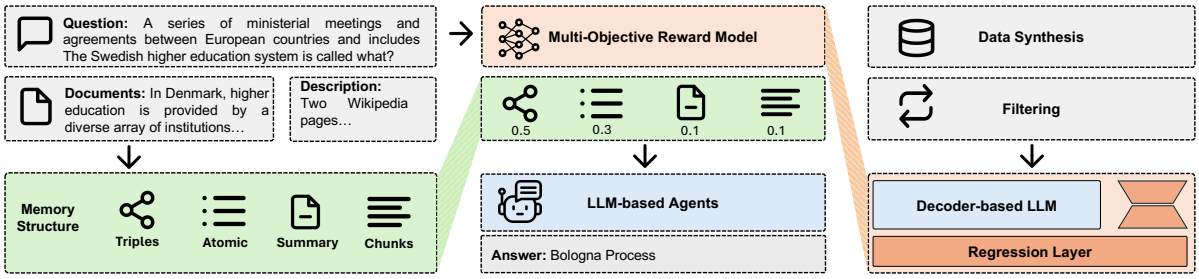

Figure 2: Overview of the augmented hybrid memory retrieval via reward-guided structuring. For each question and its corresponding document, raw information is transformed into various structural memories. The router then determines the optimal allocation strategy for the hybrid memory and orchestrates the retrieval of the most relevant memories to support precise and contextually enriched responses.

(T_q), atomic facts (A_q), summaries (S_q), and document chunks (C_q).

Knowledge Triples T_q encodes semantic relationships between entities in the form (h, r, t), where h, r, and t denote the *head*, *relation*, and *tail*, respectively. We follow prior work (Anokhin et al., 2024; Fang et al., 2024; Zeng et al., 2024) to generate such triples using a prompt-conditioned LLM. For example, from a document mentioning *Bologna Process*, we may extract: (*Bologna Process*, *under*, *Lisbon Recognition Convention*), or (*Bologna Process*, *named after*, *University of Bologna*). The prompt for generating triples is shown in Figure 6a.

Atomic Facts A_q are concise, standalone declarative statements that convey a single factual assertion, or verifiable pieces of information extracted from the source document D_q (Min et al., 2023; Li et al., 2024b). Each atomic fact is designed to represent a minimal unit of knowledge, facilitating precise retrieval and reasoning. For instance, from a document discussing the Bologna Process, we might extract: *The Bologna Process was opened to other countries in the European Cultural Convention of the Council of Europe.* Unlike knowledge triples, which strictly represent explicit semantic relationships between entities, atomic facts can express more abstract content, such as implicit relationships and conditions that are difficult to succinctly encode in a triple structure. The prompt for generating atomic facts is shown in Figure 7.

Summaries S_q are concise, high-level representations of documents D_q, capturing essential information while omitting extraneous details. This approach ensures that the summaries retain both global semantics and critical details pertinent to downstream tasks (Lee et al., 2024). The prompt for generating summaries is shown in Figure 6b.

Chunks C_q denote contiguous segments of text derived from document D_q, designed to preserve local coherence and facilitate efficient processing. Following typical chunking methods that employ fixed-length segmentation (Gao et al., 2023; Zeng et al., 2024; Dong et al., 2023), the chunked memory is represented as $C_q(D_q) = \{c_1, c_2, \ldots, c_i\}$, where each c_j is a chunk with at most L length.

3.2 Augmented Memory Composition via Reward-Guided Retrieval

After building up the hybrid memory for a given question q, we propose an augmented retrieval mechanism that adaptively fuses multiple structural memories. This process is guided by a multi-objective reward model trained to infer an optimal allocation for each memory type. The inference pipeline is illustrated in Figure 2. The design and training of the reward model are detailed in Section 4. Specifically, for a given question q associated with corresponding hybrid memory, we denote the full hybrid memory M_q as:

$$M_q = \{C_q, T_q, A_q, S_q\}. \tag{1}$$

The reward model R maps the input question q to a 4-dimensional weight vector:

$$w_q = R(q) = [w_C, w_T, w_A, w_S], \tag{2}$$

where each w_i denotes the importance weight of the corresponding memory type (Chunks, Triples, Atomic facts, and Summaries, respectively). To ensure these weights reflect probabilities summing to 1, we apply a softmax normalization:

$$w_i' = \frac{\exp(w_i/\tau)}{\sum_{j\in\{C,T,A,S\}} \exp(w_j/\tau)}, \tag{3}$$

where $i \in \{C, T, A, S\}$ and τ denotes a temperature parameter adjusting the distribution sharpness. Next, we quantize these normalized weights into discrete retrieval counts for the memory sets. Given a desired total retrieval budget K, we compute the quantized counts as:

$$n_i = \lfloor w_i' \cdot K \rfloor, \quad \text{for each } i \in C, T, A, S, \quad (4)$$

where $\lfloor \cdot \rfloor$ denotes the floor function. Typically, this results in $\sum_i n_i \leq K$, leaving a small remainder $R = K - \sum_i n_i$. To allocate this remainder, we distribute the leftover retrieval counts to memory types based on the largest fractional components from the product $w_i' \cdot K$. During retrieval:

$$\hat{M}_q = \cup_{M \in M_q} \text{Retrieve}(q, M, n_M), \quad (5)$$

where n_M denotes the quantized retrieval count for memory type M. The retrieval function $\text{Retrieve}(q, M, n_M)$ samples the top-n_M items from each memory type based on semantic similarity to q. Specifically, we utilize Qwen series embedding models for semantic retrieval (details in Appendix C).

4 Navigator Training

A central component of the TAG framework is the memory navigator, which determines the optimal composition of memory structures for each input query. However, the absence of annotated datasets specifying preferred memory allocations presents a significant challenge for supervised training. To overcome this, we propose a weakly supervised training pipeline that enables the learning of a robust multi-objective reward model navigator.

4.1 Data Construction

Currently, no benchmark datasets explicitly annotate optimal memory-structure allocations, which limits direct supervision for training models to learn query-specific memory selection strategies. Inspired by prior work that leverages large language models as evaluative judges (Li et al., 2024d; Ma et al., 2025), we adopt an in-context learning framework (Shi et al., 2024; He et al., 2025; Li et al., 2025e,d) to approximate such supervision. Specifically, for each query, we prompt the model to infer the most suitable memory structure from four candidate types, conditioned on the query and a minimal set of relevant document content. We then refine and balance the dataset by selecting 200 examples for each memory structure across all datasets. Each example is labeled with a preferred structure t_w, while other structures are marked as less relevant, resulting in a multi-pairwise preference sample:

$$D_{\text{synthetic}} = \{q^{(k)}, C^{(k)}, t_w^{(k)}, \}_{k=1}^{N}, \quad (6)$$

where $q^{(k)}$ denotes the query, $C^{(k)}$ is the associated document content, and $t_w^{(k)}$ are the preferred structure types, respectively.

4.2 Model Architecture

Drawing inspiration from typical multi-objective reward model design (Wang et al., 2024b, 2023b, 2024a; Zhou et al., 2025c), we leverage a pre-trained decoder-based LLM as the feature extractor. To repurpose this backbone for preference modeling, we freeze the original language modeling head and instead append a lightweight regression head tailored to the memory structure ranking task. To enable efficient fine-tuning, we adopt Parameter-Efficient Fine-Tuning (PEFT) via Low-Rank Adaptation (LoRA) (Hu et al., 2022; Qing et al., 2024; Zhang et al., 2025b,a; Wang et al., 2025b), which introduces trainable low-rank updates into selected layers of the LLM while keeping the majority of parameters frozen. This design allows the router to adapt quickly to the ranking task with minimal computational overhead. During the training phase of the router, only the LoRA parameters and the regression head are updated; the pre-trained LLM backbone remains frozen. For inference, the process involves two stages: first, the LoRA adapters and the regression head are activated to predict the memory allocation weights based on the input query (concatenated with specific instructions and system prompts tailored for weight prediction). Subsequently, these components are deactivated, and the LLM agent's original output head is used for generating the final answer, now informed by the retrieved memory. The input query is again concatenated with different instructions and system prompts suitable for answer generation.

4.3 Bradley-Terry Loss for Multi-Score Regression

The reward model R maps an input query to a four-dimensional score vector $\mathbf{s} = [s_1, s_2, s_3, s_4]$, where each component corresponds to the predicted utility of one memory type. Unlike traditional classification objectives, we aim to learn *relative* contributions, as multiple memory types may offer

Memory Structure	HotPotQA		2WikiQA		MuSiQue		Average	
	EM	F1	EM	F1	EM	F1	EM	F1
Uni-memory								
Chunks	52.50	69.12	35.50	47.44	16.50	33.02	34.83	49.86
Triples	29.00	44.89	29.00	39.85	9.00	19.76	22.33	34.83
Atomic Facts	37.50	50.05	27.00	38.14	12.50	23.61	25.67	37.27
Summaries	51.50	68.88	38.50	48.25	22.00	36.39	37.33	51.17
Hybrid-memory								
`Random`	53.00	71.17	37.00	48.66	20.50	36.66	36.83	52.16
`Best@1`	47.00	61.07	34.50	45.51	14.50	29.09	32.00	45.22
`Best@2`	48.00	65.42	36.00	47.01	15.50	30.88	33.83	47.77
`Best@3`	53.00	70.28	35.00	46.00	15.00	30.12	33.00	46.47
`Equal`	**58.00**	73.61	40.00	51.01	16.50	32.66	38.83	52.43
`TAG`	**58.00**	**75.73**	**40.50**	**51.22**	**22.50**	**37.92**	**40.33**	**54.96**

Table 1: Performance of single (uni) and hybrid memory structures using single-step retrieval across three datasets. Random refers to randomly retrieved memory. Best@k selects memory from the Top-K most suitable memory structures. Equal retrieves an equal number of items from each of the four available memory structures. The best scores are bolded.

complementary benefits. To this end, we adopt the Bradley-Terry (BT) loss (Bradley and Terry, 1952), a principled probabilistic framework for modeling pairwise preferences. For a given pair (i, j), the BT model defines the probability that memory type i is preferred over j as:

$$P(i \succ j) = \frac{\exp(s_i)}{\exp(s_i) + \exp(s_j)} = \sigma(s_i - s_j), \tag{7}$$

where $\sigma(\cdot)$ is the sigmoid function. Given a training instance with a preferred type g and all other types $j \neq g$, we define the loss as the average negative log-likelihood across all pairwise comparisons:

$$\mathcal{L}_{\text{BT}}(\mathbf{s}, g) = \frac{1}{K-1} \sum_{j \neq g} -\log\left(\sigma(s_g - s_j)\right), \tag{8}$$

where $K = 4$ is the total number of structure types. This loss navigates the model to assign higher scores to the preferred structure relative to the others while encouraging a *relative ranking* rather than an absolute classification. Section 7.2 conduct ablation study on different loss functions.

5 Experimental Results

5.1 Experimental Setting

Dataset Following prior work (Gutiérrez et al., 2024), we evaluate our method on three challenging long-context Multi-hop question answering (QA) datasets: HotPotQA (Yang et al., 2018), 2WikiMultihopQA (Ho et al., 2020), and MuSiQue (Trivedi et al., 2022). To reduce cost during memory construction, we adopt the evaluation set used in Zeng et al. (2024). Additional details about these datasets are provided in Appendix D.

Evaluation To evaluate QA performance, we follow previous work (Li et al., 2024b) and use standard metrics such as Exact Match (EM) score and F1 score for all the datasets.

Baseline We compare our method against several hybrid and adaptive memory selection strategies. `Random` follows Zeng et al. (2024) by retrieving a randomly selected mix of memory items. `Best@k` selects items from the top k most suitable memory structures. For instance, `Best@1` (an alternative to StructRAG (Li et al., 2024d)) uses only the most preferred structure, while `Best@2` and `Best@3` use the top two or three preferred structures, respectively. Equal retrieves an equal number of items from each of the four available memory structures.

Implementation Details We build our framework upon the Qwen2 (Bai et al., 2023) series (Appendix C), using default hyperparameter configurations. Specifically, Qwen2.5-7B-Instruct serves as the backbone for both multi-objective reward modeling (see Section 4 for details) and LLM-agent inference. To validate the robustness of our approach, we further conduct experiments based on Llama-3.1-8B-Instruct, under identical datasets and hyperparameter settings (Results are provided in Table 7). We use Qwen2.5-72B-Instruct for synthetic

Memory Structure	HotPotQA		2WikiQA		MuSiQue		Average	
	EM	F1	EM	F1	EM	F1	EM	F1
TAG	**58.00**	**75.73**	**40.50**	**51.22**	**22.50**	**37.92**	**40.33**	**54.96**
W/o Chunks	51.00	68.53	38.50	49.16	19.50	34.94	36.33	50.88
W/o Triples	56.50	72.48	37.00	49.51	22.50	37.70	38.67	53.23
W/o Atomic Facts	55.00	72.99	40.50	51.63	20.00	33.76	38.50	52.79
W/o Summaries	51.50	71.01	36.50	48.54	15.00	31.73	34.33	50.43

Table 2: Performance of removing each memory structure compared to keeping full memory.

data generation and serve the model via API using `vllm`[1]. We present the details of reward model training in Appendix E.

5.2 Main Results

We present the core performance of TAG in the standard single-step retrieval setting (Rubin et al., 2022). As summarized in Table 1, TAG consistently surpasses both uni-structured and hybrid-memory baselines across all evaluated benchmarks. It delivers the strongest overall results on every dataset, with substantial gains over the best-performing single-memory configuration. In particular, on HotPotQA, TAG attains an F1 score of 75.73, exceeding the strongest single-memory baseline (Chunks, 69.12) by 6.61 points. Likewise, it achieves F1 scores of 51.22 on 2WikiQA and 37.92 on MuSiQue, establishing new state-of-the-art performance on both benchmarks. These results highlight the effectiveness of query-adaptive, multi-granular memory composition in retrieving task-relevant evidence and strengthening long-context multi-hop reasoning for LLM-based agents.

5.3 Ablation Study

To better understand the contribution of each memory structure in our framework, we conduct an ablation study by systematically removing one structure at a time. The results are shown in Table 2. We observe that the TAG consistently outperforms all ablated variants across datasets, highlighting the complementary strengths of the four memory structures. Removing chunks and summaries leads to large performance drops, with an average F1 score decrease of 4.53 for summaries. This suggests that chunks and summaries, which offer high-level context and thematic overviews that aid global reasoning, are especially critical for grounding fine-grained reasoning in contextually rich passages. Removing triples and atomic facts causes a moderate drop in performance. Triples encode explicit entity-relation-entity structures that benefit factoid-style reasoning but may lack flexibility in capturing implicit connections.

Components	Parameters	Additional
Base Model	7.62B	100%
LoRA Block	2.52M	0.033%
Regression Head	0.014M	0.00019%

Table 3: Comparison of parameter size for router components relative to the base model.

Dataset	Predict Weights (s)	Generation (s)
2WikiQA	0.0451	0.4719
HotpotQA	0.0466	0.4582
MuSiQue	0.0468	0.4857

Table 4: Average per-call latency for weights prediction and answer generation across three datasets.

6 Latency Analysis

We assess the efficiency of the proposed reward-guided memory router by analyzing its parameter overhead and inference latency. As described in Section 4.2, the router comprises a LoRA-adapted LLM and a lightweight regression head that predicts optimal memory composition for each query. Table 3 reports the parameter size of these components. The LoRA module introduces only 0.033% additional parameters relative to the base LLM, while the regression head contributes a negligible 0.00019%. This minimal overhead highlights the scalability and practicality of our router for integration into large-scale agent systems. To evaluate runtime efficiency, we measure the average latency introduced by the router on three multi-hop QA benchmarks. As shown in Table 4, the memory

[1] https://pypi.org/project/vllm/

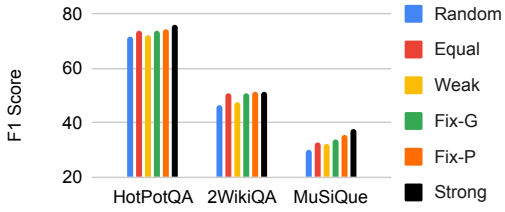

Figure 3: F1 score of different routing Strategies.

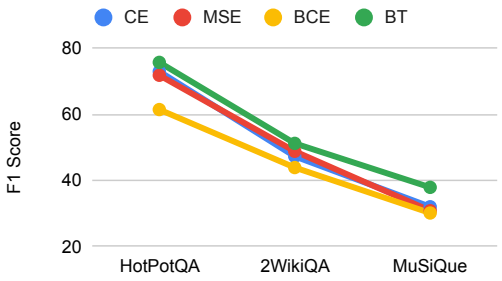

Figure 4: F1 score of different training loss functions.

routing step requires only 0.0451 to 0.0468 seconds per query, which is marginal compared to the total answer generation time (approximately 0.46 to 0.49 seconds). These results demonstrate that our reward-guided routing mechanism significantly improves the flexibility and task adaptiveness of memory retrieval, while incurring minimal computational cost, making it a highly viable solution for real-world deployment.

7 Analysis of Navigating Strategy

In this section, we aim to answer two questions: *1) How does the quality of the router impact the performance? 2) Is it necessary to employ a multi-objective reward model for memory routing?*

7.1 Routing Strategy

To assess the impact of the reward model's quality on TAG, we conducted a comparative analysis using routers of varying routing capabilities. We evaluated four distinct router configurations: (1) *Strong Router*: trained on the complete synthetic dataset (2400 examples), representing our best-performing reward model. (2) *Weak Router*: trained on a randomly selected 50% subset of the training data (1200 examples), simulating a less informed model. (3) *Average Router*: a heuristic baseline that assigns equal importance ($w_i = 0.25$) to all four memory structures. (4) *Bad Router*: a baseline that assigns random weights to memory structures. (5) *Fix-G (global)*: a single global weight vector $w = [w_C, w_T, w_A, w_S]$ applied to all queries, where each w_i is proportional to the single-structure F1 score of the corresponding memory type reported in Table 1, reflecting their relative importance. (6) *Fix-P (per-dataset)*: similar to (5), but a separate w is computed for each dataset (HotpotQA, 2Wiki, MuSiQue) using their respective single-structure F1 scores from Table 1.

Figure 3 demonstrates that the effectiveness of TAG heavily depends on the router's quality. The *Strong Router* achieves the highest F1 across all datasets, while the *Weak Router* can underperform the *Average Router*, indicating that insufficiently trained reward models may misallocate memory. Notably, both fixed-weight schemes (*Fix-G* and *Fix-P*) lag behind the query-adaptive *Strong Router*, underscoring the necessity of *per-query* routing over corpus-level weighting. Complete results are reported in Table 10.

7.2 Loss Function Design

As discussed in Section 4.3, our objective is to model the relative contributions of multiple memory structures to a given question, rather than selecting a single best structure. To this end, the model outputs a four-dimensional score vector, where each dimension reflects the predicted utility of a corresponding memory structure. We compare our proposed multi-object Bradley-Terry (BT) loss against three alternative loss functions: Cross-Entropy (CE)(Shannon, 1948), Mean Squared Error (MSE)(Bishop, 2006), and Binary Cross-Entropy (BCE)(Cox, 1958) (see Appendix E for implementation details). Each loss function introduces distinct inductive biases. CE treats the task as multi-class classification, assuming a single optimal structure. MSE models the task as regression, minimizing the squared L2 distance between predicted scores and the preference labels. BCE allows independent estimation of each structure's relevance, making it more flexible for multi-label scenarios. In contrast, our multi-object BT loss is explicitly designed to encourage pairwise ranking consistency. It encourages the model to score the preferred structure higher than the others, aligning directly with our goal of capturing relative usefulness. As illustrated in Figure 4, the BT loss achieves the highest average F1 score among all the loss functions, indicating that its ranking-based formulation is more effective for modeling nuanced structural preferences. Full results are reported in Table 9.

8 Case Study

To better understand the importance of hybrid memory composition, we analyze a failure case involving multi-hop reasoning from HotPotQA dataset. The query asks: *"Who costarred in the 1935 American drama film Les Misérables with the husband of Elsa Lanchester?"*.

Mix Memory

Triples:
• Charles Laughton; nationality; English
...
Atomic Facts:
• Charles Laughton was trained at the Royal Academy of Dramatic Art in London.
• Fredric March and Charles Laughton star in Les Misérables.→`Factual Details`
...
Chunks:
• Les Misérables is a 1935 American drama film starring Fredric March and Charles Laughton...→`Global Semantics`
...
Answer: Fredric March ✓

Single chunk structure memory

Chunks:
• Les Misérables is a 1935 American drama film starring Fredric March and Charles Laughton... →`Global Semantics`
• Charles Laughton...lived and worked with Elsa Lanchester... →`Noise Content`
...
Answer: Charles Laughton ✗

When relying on a single memory type such as document chunks, the agent retrieves one passage chunk including statement that Les Misérables stars Charles Laughton and Fredric March, and another chunk describe biographical information including that Charles Laughton was the husband of Elsa Lanchester. However, without structured representations to guide relevance, the agent selects Charles Laughton as the answer—incorrectly identifying Lanchester's husband instead of his co-star. This mistake highlights the "semantic drift" problem (Spataru et al., 2024) typical of long chunks, where textual noise or proximity bias misleads the model. Under Hybrid memory configuration, the agent correctly identifies the answer as **Fredric March**. This memory context explicitly links Elsa Lanchester to her husband, Charles Laughton, and co-lists both Laughton and March in the cast of *Les Misérables*. The model successfully performs two-hop reasoning: first inferring that Charles Laughton is Lanchester's husband, and then identifying Fredric March as Laughton's co-star.

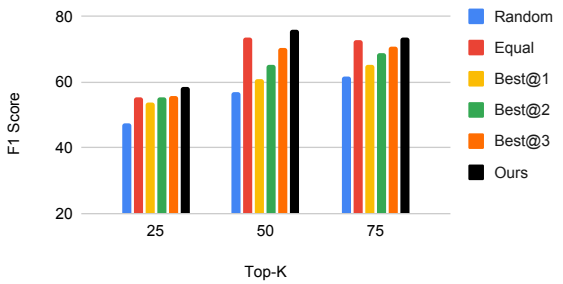

Figure 5: F1 score of different routing strategies under varying Top-K.

9 Hyperparameter Sensitivity

To further validate the robustness of `TAG` framework, we analyze its sensitivity to the hyperparameter Top-K, which controls the number of retrieved memory units during inference. We evaluate performance across three values of $K \in 25, 50, 75$ using the HotPotQA dataset. As shown in Figure 5, `TAG` achieves optimal performance at K = 50, with an F1 score of 75.73, outperforming other configurations. Overall, these results underscore the importance of balancing retrieval breadth and precision. While larger K values provide more opportunities to gather relevant facts, they may also introduce noise. Our adaptive routing mechanism helps mitigate this trade-off by learning to allocate retrieval resources more effectively across different memory types. The full results is provided in Table 8.

10 Conclusion

This work introduces `TAG`, a novel framework designed to overcome the limitations of uni-structural memory in large language model agents for complex, long-context reasoning. By integrating diverse memory structures and employing a reward-guided retrieval mechanism trained with a multi-object Bradley-Terry loss, `TAG` adaptively composes an optimal memory set tailored to each query. Our extensive experiments underscore the critical role of adaptive, granular memory composition in enhancing the long-context reasoning capabilities of LLM agents.

Limitations

TAG has several limitations. First, it relies on automatically generated structured memories produced by a large teacher model. While this follows common offline index-enrichment practice in modern RAG pipelines, any generation errors, omissions, or inconsistencies can propagate into retrieval and reasoning, and TAG does not include an explicit mechanism to validate, denoise, or correct individual memory units. Second, the navigator is trained with weak synthetic preference supervision rather than human annotations or end-task rewards. Although our router-quality analyses indicate performance degrades predictably as supervision quality weakens, we do not fully characterize the synthetic dataset's properties or the teacher's selection rationale, and future work should study robustness to label noise more directly or explore alternative supervision, such as partial human feedback or end-to-end RL. Third, our experiments are confined to open-source models, which might not be representative of the broader landscape of LLMs, particularly those that are closed-source and potentially optimized for proprietary datasets.

Ethics Statement

We have not identified any ethical concerns directly related to this study.

References

Petr Anokhin, Nikita Semenov, Artyom Sorokin, Dmitry Evseev, Mikhail Burtsev, and Evgeny Burnaev. 2024. Arigraph: Learning knowledge graph world models with episodic memory for llm agents. *arXiv preprint arXiv:2407.04363*.

Jihun Baek, Soyeong Lee, Sungmin Kim, Kangwook Lee, and Sungju Kim. 2023. Knowledge graph-enhanced large language models for instruction following. *arXiv preprint arXiv:2310.04172*.

Jinze Bai, Shuai Bai, Yunfei Chu, Zeyu Cui, Kai Dang, Xiaodong Deng, Yang Fan, Wenbin Ge, Yu Han, Fei Huang, et al. 2023. Qwen technical report. *arXiv preprint arXiv:2309.16609*.

Jinhe Bi, Danqi Yan, Yifan Wang, Wenke Huang, Haokun Chen, Guancheng Wan, Mang Ye, Xun Xiao, Hinrich Schuetze, Volker Tresp, et al. 2025. Cot-kinetics: A theoretical modeling assessing lrm reasoning process. *arXiv preprint arXiv:2505.13408*.

Christopher M. Bishop. 2006. *Pattern Recognition and Machine Learning*. Springer, New York.

Ralph Allan Bradley and Milton E. Terry. 1952. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345.

Hanning Chen, Yang Ni, Wenjun Huang, Hyunwoo Oh, Yezi Liu, Tamoghno Das, and Mohsen Imani. 2025. Lvlm_csp: Accelerating large vision language models via clustering, scattering, and pruning for reasoning segmentation. In *Proceedings of the 33rd ACM International Conference on Multimedia*, pages 3932–3941.

Tong Chen, Hongwei Wang, Sihao Chen, Wenhao Yu, Kaixin Ma, Xinran Zhao, Hongming Zhang, and Dong Yu. 2024. Dense X retrieval: What retrieval granularity should we use? In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*.

Ming Cheng, Xingjian Diao, Shitong Cheng, and Wenjun Liu. 2024a. Saic: Integration of speech anonymization and identity classification. In *AI for Health Equity and Fairness: Leveraging AI to Address Social Determinants of Health*, pages 295–306. Springer.

Ming Cheng, Ziyi Zhou, Bowen Zhang, Ziyu Wang, Jiaqi Gan, Ziang Ren, Weiqi Feng, Yi Lyu, Hefan Zhang, and Xingjian Diao. 2024b. Efflex: Efficient and flexible pipeline for spatio-temporal trajectory graph modeling and representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2546–2555.

David R. Cox. 1958. The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2):215–232.

Xingjian Diao, Zheyuan Liu, Chunhui Zhang, Weiyi Wu, Keyi Kong, Lin Shi, Kaize Ding, Soroush Vosoughi, and Jiang Gui. 2026. Addressing overthinking in large vision-language models via gated perception-reasoning optimization. *arXiv preprint arXiv:2601.04442*.

Xingjian Diao, Weiyi Wu, Keyi Kong, Peijun Qing, Xinwen Xu, Ming Cheng, Soroush Vosoughi, and Jiang Gui. 2025a. Protovqa: An adaptable prototypical framework for explainable fine-grained visual question answering. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 1040–1057.

Xingjian Diao, Chunhui Zhang, Keyi Kong, Weiyi Wu, Chiyu Ma, Zhongyu Ouyang, Peijun Qing, Soroush Vosoughi, and Jiang Gui. 2025b. Soundmind: Rl-incentivized logic reasoning for audio-language models. *arXiv preprint arXiv:2506.12935*.

Xingjian Diao, Chunhui Zhang, Weiyi Wu, Zhongyu Ouyang, Peijun Qing, Ming Cheng, Soroush Vosoughi, and Jiang Gui. 2025c. Temporal working memory: Query-guided segment refinement for enhanced multimodal understanding. *arXiv preprint arXiv:2502.06020*.

Kuicai Dong, Yujing Chang, Derrick Goh Xin Deik, Dexun Li, Ruiming Tang, and Yong Liu. 2025a. MMDocIR: Benchmarking multimodal retrieval for long documents. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 30971–31005, Suzhou, China. Association for Computational Linguistics.

Kuicai Dong, Yujing Chang, Shijie Huang, Yasheng Wang, Ruiming Tang, and Yong Liu. 2025b. Benchmarking retrieval-augmented multimodal generation for document question answering. *Preprint*, arXiv:2505.16470.

Kuicai Dong, Derrick Goh Xin Deik, Yi Quan Lee, Hao Zhang, Xiangyang Li, Cong Zhang, and Yong Liu. 2024. MC-indexing: Effective long document retrieval via multi-view content-aware indexing. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 2673–2691, Miami, Florida, USA. Association for Computational Linguistics.

Kuicai Dong, Shurui Huang, Fangda Ye, Wei Han, Zhi Zhang, Dexun Li, Wenjun Li, Qu Yang, Gang Wang, Yichao Wang, Chen Zhang, and Yong Liu. 2025c. Doc-researcher: A unified system for multimodal document parsing and deep research. *Preprint*, arXiv:2510.21603.

Kuicai Dong, Aixin Sun, Jung-jae Kim, and Xiaoli Li. 2023. Open information extraction via chunks. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 15390–15404, Singapore. Association for Computational Linguistics.

Kuicai Dong, Zhao Yilin, Aixin Sun, Jung-Jae Kim, and Xiaoli Li. 2021. DocOIE: A document-level context-aware dataset for OpenIE. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 2377–2389, Online. Association for Computational Linguistics.

Jinyuan Fang, Zaiqiao Meng, and Craig MacDonald. 2024. TRACE the evidence: Constructing knowledge-grounded reasoning chains for retrieval-augmented generation. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 8472–8494, Miami, Florida, USA. Association for Computational Linguistics.

Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Meng Wang, and Haofen Wang. 2023. Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*.

Bernal Jiménez Gutiérrez, Yiheng Shu, Yu Gu, Michihiro Yasunaga, and Yu Su. 2024. Hipporag: Neurobiologically inspired long-term memory for large language models. *arXiv preprint arXiv:2405.14831*.

Bernal Jiménez Gutiérrez, Yiheng Shu, Yu Gu, Michihiro Yasunaga, and Yu Su. 2024. Hipporag: Neurobiologically inspired long-term memory for large language models. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.

Yangfan He, Jianhui Wang, Yijin Wang, Yan Zhong, Xinyuan Song, Junjiang Lin, Xinhang Yuan, Jingqun Tang, Yi Xin, Hao Zhang, et al. 2025. Enhancing intent understanding for ambiguous prompt: A human-machine co-adaption strategy. *arXiv preprint arXiv:2501.15167*.

Xanh Ho, Anh-Khoa Duong Nguyen, Saku Sugawara, and Akiko Aizawa. 2020. Constructing a multi-hop qa dataset for comprehensive evaluation of reasoning steps. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 6609–6625.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. Lora: Low-rank adaptation of large language models. In *International Conference on Learning Representations*.

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

Sen Hu, Yuxiang Wei, Jiaxin Ran, Zhiyuan Yao, and Lei Zou. 2026. Does memory need graphs? a unified framework and empirical analysis for long-term dialog memory. *arXiv preprint arXiv:2601.01280*.

Yucheng Hu and Yuxing Lu. 2024. Rag and rau: A survey on retrieval-augmented language model in natural language processing. *arXiv preprint arXiv:2404.19543*.

Haiduo Huang, Jiangcheng Song, Yadong Zhang, and Pengju Ren. 2025a. Selectkd: Selective token-weighted knowledge distillation for llms. *arXiv preprint arXiv:2510.24021*.

Haiduo Huang, Jiangcheng Song, Wenzhe Zhao, and Pengju Ren. 2025b. Fasteagle: Cascaded drafting for accelerating speculative decoding. *arXiv preprint arXiv:2509.20416*.

Haiduo Huang, Fuwei Yang, Zhenhua Liu, Yixing Xu, Jinze Li, Yang Liu, Xuanwu Yin, Dong Li, Pengju Ren, and Emad Barsoum. 2025c. Jakiro: Boosting speculative decoding with decoupled multi-head via moe. *arXiv preprint arXiv:2502.06282*.

Haiduo Huang, Fuwei Yang, Zhenhua Liu, Xuanwu Yin, Dong Li, Pengju Ren, and Emad Barsoum. 2025d. Specvlm: Fast speculative decoding in vision-language models. *arXiv preprint arXiv:2509.11815*.

Wenjun Huang, Ziteng Cui, Yinqiang Zheng, Yirui He, Tatsuya Harada, and Mohsen Imani. 2025e. Dr. raw: Towards general high-level vision from raw with efficient task conditioning. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

Wenjun Huang, Yang Ni, Arghavan Rezvani, SungHeon Jeong, Hanning Chen, Yezi Liu, Fei Wen, and Mohsen Imani. 2025f. Recoverable anonymization for pose estimation: A privacy-enhancing approach. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 5239–5249. IEEE.

Liu Huanshuo, Hao Zhang, Zhijiang Guo, Jing Wang, Kuicai Dong, Xiangyang Li, Yi Quan Lee, Cong Zhang, and Yong Liu. 2025. CtrlA: Adaptive retrieval-augmented generation via inherent control. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 12592–12618, Vienna, Austria. Association for Computational Linguistics.

Priya Jain et al. 2024. Knowledge structure generation and utilization by llms: Emerging methods. *arXiv preprint arXiv:2402.03546*.

Kuang-Huei Lee, Xinyun Chen, Hiroki Furuta, John F. Canny, and Ian Fischer. 2024. A human-inspired reading agent with gist memory of very long contexts. In *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*. OpenReview.net.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020a. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems*.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020b. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33:9459–9474.

Bin Li and Hanjun Deng. 2023. Bilateral personalized dialogue generation with contrastive learning. *Soft Computing*, 27(6):3115–3132.

Bin Li, Bin Sun, Shutao Li, Encheng Chen, Hongru Liu, Yixuan Weng, Yongping Bai, and Meiling Hu. 2024a. Distinct but correct: generating diversified and entity-revised medical response. *Science China Information Sciences*, 67(3):132106.

Jinze Li, Yixing Xu, Haiduo Huang, Xuanwu Yin, Dong Li, Edith CH Ngai, and Emad Barsoum. 2025a. Gumiho: A hybrid architecture to prioritize early tokens in speculative decoding. *arXiv preprint arXiv:2503.10135*.

Shilong Li, Yancheng He, Hangyu Guo, Xingyuan Bu, Ge Bai, Jie Liu, Jiaheng Liu, Xingwei Qu, Yangguang Li, Wanli Ouyang, et al. 2024b. Graphreader: Building graph-based agent to enhance long-context abilities of large language models. *arXiv preprint arXiv:2406.14550*.

Shutao Li, Bin Li, Bin Sun, and Yixuan Weng. 2024c. Towards visual-prompt temporal answer grounding in instructional video. *IEEE transactions on pattern analysis and machine intelligence*, 46(12):8836–8853.

Xiang Li et al. 2023. Structured knowledge generation by large language models. *arXiv preprint arXiv:2310.08547*.

Xiangyang Li, Kuicai Dong, Yi Quan Lee, Wei Xia, Hao Zhang, Xinyi Dai, Yasheng Wang, and Ruiming Tang. 2025b. CoIR: A comprehensive benchmark for code information retrieval models. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 22074–22091, Vienna, Austria. Association for Computational Linguistics.

Xianhang Li, Yanqing Liu, Haoqin Tu, and Cihang Xie. 2025c. Openvision: A fully-open, cost-effective family of advanced vision encoders for multimodal learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3977–3987.

Yanshu Li, Yi Cao, Hongyang He, Qisen Cheng, Xiang Fu, Xi Xiao, Tianyang Wang, and Ruixiang Tang. 2025d. M²IV: Towards efficient and fine-grained multimodal in-context learning via representation engineering. In *Second Conference on Language Modeling*.

Yanshu Li, Jianjiang Yang, Tian Yun, Pinyuan Feng, Jinfa Huang, and Ruixiang Tang. 2025e. Taco: Enhancing multimodal in-context learning via task mapping-guided sequence configuration. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 736–763.

Zhuoqun Li, Xuanang Chen, Haiyang Yu, Hongyu Lin, Yaojie Lu, Qiaoyu Tang, Fei Huang, Xianpei Han, Le Sun, and Yongbin Li. 2024d. Structrag: Boosting knowledge intensive reasoning of llms via inference-time hybrid information structurization.

Nelson F Liu, Kevin Lin, John Hewitt, Ashwin Paranjape, Michele Bevilacqua, Fabio Petroni, and Percy Liang. 2023. Lost in the middle: How language models use long contexts. *arXiv preprint arXiv:2307.03172*.

Yanqing Liu, Xianhang Li, Zeyu Wang, Bingchen Zhao, and Cihang Xie. 2024. Clips: An enhanced clip framework for learning with synthetic captions. *arXiv preprint arXiv:2411.16828*.

Yanqing Liu, Xianhang Li, Letian Zhang, Zirui Wang, Zeyu Zheng, Yuyin Zhou, and Cihang Xie. 2025. Openvision 2: A family of generative pretrained visual encoders for multimodal learning. *arXiv preprint arXiv:2509.01644*.

Chiyu Ma, Enpei Zhang, Yilun Zhao, Wenjun Liu, Yaning Jia, Peijun Qing, Lin Shi, Arman Cohan, Yujun Yan, and Soroush Vosoughi. 2025. Judging with many minds: Do more perspectives mean less prejudice? *arXiv preprint arXiv:2505.19477*.

Sewon Min, Kalpesh Krishna, Xinxi Lyu, Mike Lewis, Wen-tau Yih, Pang Wei Koh, Mohit Iyyer, Luke Zettlemoyer, and Hannaneh Hajishirzi. 2023. FActScore: Fine-grained atomic evaluation of factual precision in long form text generation. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*.

Charles Packer, Sarah Wooders, Kevin Lin, Vivian Fang, Shishir G Patil, Ion Stoica, and Joseph E Gonzalez. 2023. Memgpt: Towards llms as operating systems. *arXiv preprint arXiv:2310.08560*.

Ofir Press, Muru Zhang, Sewon Min, Ludwig Schmidt, Noah Smith, and Mike Lewis. 2023. Measuring and narrowing the compositionality gap in language models. In *Findings of the Association for Computational Linguistics: EMNLP 2023*.

Peijun Qing, Chongyang Gao, Yefan Zhou, Xingjian Diao, Yaoqing Yang, and Soroush Vosoughi. 2024. Alphalora: Assigning lora experts based on layer training quality.

Ohad Rubin, Jonathan Herzig, and Jonathan Berant. 2022. Learning to retrieve prompts for in-context learning. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2655–2671.

Claude E. Shannon. 1948. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423.

Fobo Shi, Peijun Qing, Dong Yang, Nan Wang, Youbo Lei, Haonan Lu, Xiaodong Lin, and Duantengchuan Li. 2024. Prompt space optimizing few-shot reasoning success with large language models. In *Findings of the Association for Computational Linguistics: NAACL 2024*.

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

Ava Spataru, Eric Hambro, Elena Voita, and Nicola Cancedda. 2024. Know when to stop: A study of semantic drift in text generation. *arXiv preprint arXiv:2404.05411*.

Zhen Sun, Runjin Wang, Siyuan Chen, Kexuan Wang, Kaisheng Feng, Chonggang Wang, Shimin Shi, Yejin Zhang, Xin Huang, Yu Zhang, et al. 2024. Structgpt: A general framework for large language model to reason over structured data. *arXiv preprint arXiv:2402.00818*.

Yijun Tian, Shaoyu Chen, Zhichao Xu, Yawei Wang, Jinhe Bi, Peng Han, and Wei Wang. 2025. Reinforcement mid-training. *Preprint*, arXiv:2509.24375.

Harsh Trivedi, Niranjan Balasubramanian, Tushar Khot, and Ashish Sabharwal. 2022. Musique: Multi-hop questions via single-hop question composition. *Transactions of the Association for Computational Linguistics*, 10:539–554.

Harsh Trivedi, Niranjan Balasubramanian, Tushar Khot, and Ashish Sabharwal. 2023. Interleaving retrieval with chain-of-thought reasoning for knowledge-intensive multi-step questions. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.

Narayan S Umanath and Iris Vessey. 1994. Multiattribute data presentation and human judgment: A cognitive fit perspective. *Decision Sciences*, 25(5-6):795–824.

Iris Vessey. 1991. Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences*, 22(2):219–240.

Haoxiang Wang, Yong Lin, Wei Xiong, Rui Yang, Shizhe Diao, Shuang Qiu, Han Zhao, and Tong Zhang. 2024a. Arithmetic control of llms for diverse user preferences: Directional preference alignment with multi-objective rewards. In *Association for Computational Linguistics*.

Haoxiang Wang, Wei Xiong, Tengyang Xie, Han Zhao, and Tong Zhang. 2024b. Interpretable preferences via multi-objective reward modeling and mixture-of-experts. *arXiv preprint arXiv:2406.12845*.

Junxi Wang, Jize Liu, Na Zhang, and Yaxiong Wang. 2025a. Consistency-aware fake videos detection on short video platforms. In *International Conference on Intelligent Computing*, pages 200–210. Springer.

Junxi Wang, Yaxiong Wang, Lechao Cheng, and Zhun Zhong. 2025b. Fakesv-vlm: Taming vlm for detecting fake short-video news via progressive mixture-of-experts adapter. *arXiv preprint arXiv:2508.19639*.

Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, et al. 2024c. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6):186345.

Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Luo, Wayne Xin Tian, Miaohua Niu, Le Wu, Keyu Wang, Shucheng Wang, and Lining Wang. 2023a. A survey on large language model based autonomous agents. *arXiv preprint arXiv:2308.11432*.

Zhilin Wang, Yi Dong, Jiaqi Zeng, Virginia Adams, Makesh Narsimhan Sreedhar, Daniel Egert, Olivier Delalleau, Jane Polak Scowcroft, Neel Kant, Aidan Swope, and Oleksii Kuchaiev. 2023b. Helpsteer: Multi-attribute helpfulness dataset for steerlm. *Preprint*, arXiv:2311.09528.

Yuxiang Wei, Yanteng Zhang, Xi Xiao, Tianyang Wang, Xiao Wang, and Vince D Calhoun. 2025a. 4d multimodal co-attention fusion network with latent contrastive alignment for alzheimer's diagnosis. *arXiv preprint arXiv:2504.16798*.

Yuxiang Wei, Yanteng Zhang, Xi Xiao, Tianyang Wang, Xiao Wang, and Vince D Calhoun. 2025b. Morebrain: Routed mixture of experts for interpretable and generalizable cross-subject fmri visual decoding. *arXiv preprint arXiv:2505.15946*.

Weiyi Wu, Xinwen Xu, Chongyang Gao, Xingjian Diao, Siting Li, Lucas A Salas, and Jiang Gui. 2025. Assessing and mitigating medical knowledge drift and conflicts in large language models. *arXiv preprint arXiv:2505.07968*.

Zhiheng Xi, Wenxiang Chen, Xin Guo, Wei He, Yiwen Ding, Boyang Hong, Ming Zhang, Junzhe Wang, Senjie Jin, Enyu Zhou, et al. 2023. The rise and potential of large language model based agents: A survey. *arXiv preprint arXiv:2309.07864*.

Zhiheng Xi, Wenxiang Chen, Xin Guo, Wei He, Yiwen Ding, Boyang Hong, Ming Zhang, Junzhe Wang, Senjie Jin, Enyu Zhou, et al. 2025. The rise and potential of large language model based agents: A survey. *Science China Information Sciences*, 68(2):121101.

Wulin Xie, Xiaohuan Lu, Yadong Liu, Jiang Long, Bob Zhang, Shuping Zhao, and Jie Wen. 2024. Uncertainty-aware pseudo-labeling and dual graph driven network for incomplete multi-view multi-label classification. In *Proceedings of the 32nd ACM international conference on multimedia*, pages 6656–6665.

Wulin Xie, Yi-Fan Zhang, Chaoyou Fu, Yang Shi, Bingyan Nie, Hongkai Chen, Zhang Zhang, Liang Wang, and Tieniu Tan. 2025a. Mme-unify: A comprehensive benchmark for unified multimodal understanding and generation models. *arXiv preprint arXiv:2504.03641*.

Wulin Xie, Lian Zhao, Jiang Long, Xiaohuan Lu, and Bingyan Nie. 2025b. Multi-view factorizing and disentangling: A novel framework for incomplete multi-view multi-label classification. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 1914–1923. IEEE.

Yikang Xu, Shuyuan Li, J D Choi, and M Coyle. 2023. Memorysandbox: A unified framework for diverse agent memory. *arXiv preprint arXiv:2312.10272*.

Sikuan Yan, Xiufeng Yang, Zuchao Huang, Ercong Nie, Zifeng Ding, Zonggen Li, Xiaowen Ma, Jinhe Bi, Kristian Kersting, Jeff Z. Pan, Hinrich Schütze, Volker Tresp, and Yunpu Ma. 2026. Memory-r1: Enhancing large language model agents to manage and utilize memories via reinforcement learning. *Preprint*, arXiv:2508.19828.

Dong Yang, Peijun Qing, Yang Li, Haonan Lu, and Xiaodong Lin. 2022. Gammae: Gamma embeddings for logical queries on knowledge graphs.

Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William Cohen, Ruslan Salakhutdinov, and Christopher D Manning. 2018. Hotpotqa: A dataset for diverse, explainable multi-hop question answering. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2369–2380.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. 2022. React: Synergizing reasoning and acting in language models. *arXiv preprint arXiv:2210.03629*.

Xiangchi Yuan, Xiang Chen, Tong Yu, Dachuan Shi, Can Jin, Wenke Lee, and Saayan Mitra. 2025a. Mitigating forgetting between supervised and reinforcement learning yields stronger reasoners. *arXiv preprint arXiv:2510.04454*.

Xiangchi Yuan, Chunhui Zhang, Zheyuan Liu, Dachuan Shi, Leyan Pan, Soroush Vosoughi, and Wenke Lee. 2025b. Superficial self-improved reasoners benefit from model merging. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 5912–5932.

Ruihong Zeng, Jinyuan Fang, Siwei Liu, and Zaiqiao Meng. 2024. On the structural memory of llm agents. *arXiv preprint arXiv:2412.15266*.

Hao Zhang, Bo Huang, Zhenjia Li, Xi Xiao, Hui Yi Leong, Zumeng Zhang, Xinwei Long, Tianyang Wang, and Hao Xu. 2025a. Sensitivity-lora: Low-load sensitivity-based fine-tuning for large language models. *arXiv preprint arXiv:2509.09119*.

Hao Zhang, Zhenjia Li, Runfeng Bao, Yifan Gao, Xi Xiao, Heng Zhang, Shuyang Zhang, Bo Huang, Yuhang Wu, Tianyang Wang, et al. 2025b. Hyper-adalora: Accelerating lora rank allocation during training via hypernetworks without sacrificing performance. *arXiv preprint arXiv:2510.02630*.

Hao Zhang, Mengsi Lyu, Bo Huang, Yulong Ao, and Yonghua Lin. 2025c. Trimtokenator-lc: Towards adaptive visual token pruning for large multimodal models with long contexts. *arXiv preprint arXiv:2512.22748*.

Heng Zhang, Tianyi Zhang, Yuling Shi, Xiaodong Gu, Yaomin Shen, Zijian Zhang, Yilei Yuan, Hao Zhang, and Jin Huang. 2025d. Can representation gaps be the key to enhancing robustness in graph-text alignment? *arXiv preprint arXiv:2510.12087*.

Wenlin Zhang, Xiangyang Li, Kuicai Dong, Yichao Wang, Pengyue Jia, Xiaopeng Li, Yingyi Zhang, Derong Xu, Zhaocheng Du, Huifeng Guo, Ruiming Tang, and Xiangyu Zhao. 2025e. Process vs. outcome reward: Which is better for agentic rag reinforcement learning. *Preprint*, arXiv:2505.14069.

Zeyu Zhang, Xiaohe Bo, Chen Ma, Rui Li, Xu Chen, Quanyu Dai, Jieming Zhu, Zhenhua Dong, and Ji-Rong Wen. 2024. A survey on the memory mechanism of large language model based agents. *arXiv preprint arXiv:2404.13501*.

Qingfei Zhao, Ruobing Wang, Yukuo Cen, Daren Zha, Shicheng Tan, Yuxiao Dong, and Jie Tang. 2024. LongRAG: A dual-perspective retrieval-augmented generation paradigm for long-context question answering. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*.

Changshi Zhou, Feng Luan, Jiarui Hu, Shaoqiang Meng, Zhipeng Wang, Yanchao Dong, Yanmin Zhou, and Bin He. 2025a. Learning efficient robotic garment manipulation with standardization. *arXiv preprint arXiv:2506.22769*.

Changshi Zhou, Haichuan Xu, Ningquan Gu, Zhipeng Wang, Bin Cheng, Pengpeng Zhang, Yanchao Dong, Mitsuhiro Hayashibe, Yanmin Zhou, and Bin He. 2025b. Language-guided long horizon manipulation with llm-based planning and visual perception. *arXiv preprint arXiv:2509.02324*.

Yixiao Zhou, Ziyu Zhao, Dongzhou Cheng, Zhiliang Wu, Jie Gui, Yi Yang, Fei Wu, Yu Cheng, and Hehe Fan. 2025c. Dropping experts, recombining neurons: Retraining-free pruning for sparse mixture-of-experts llms. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 15169–15186.

A Extended Related Work

A.1 LLM Agents and Memory

LLM-based agents extend language models with planning, tool use, and memory components, enabling long-horizon interaction and learning from experience (Xi et al., 2023; Wang et al., 2023a; Yao et al., 2022; Shinn et al., 2023). Memory plays a critical role in grounding reasoning and maintaining state across extended contexts (Zhang et al., 2024; Lee et al., 2024; Diao et al., 2025c; Yan et al., 2026). A prevalent design stores past observations as chunked text indexed in vector databases (Lewis et al., 2020b; Packer et al., 2023). Despite its scalability, chunk-level memory is often noisy and susceptible to attention failures in long inputs, including the "lost-in-the-middle" phenomenon (Liu et al., 2023). These limitations motivate structured and multi-view memory representations that more explicitly expose salient content and relationships, especially for long and heterogeneous documents.

A.2 Long-Context Multi-Hop QA with RAG

Retrieval-augmented generation (RAG) (Lewis et al., 2020a; Hu and Lu, 2024; Li et al., 2025b; Zhang et al., 2025e) is the dominant paradigm for grounding LLMs in external corpora. However, multi-hop reasoning over long contexts remains challenging due to dispersed evidence and abundant distractors (Zhang et al., 2025c). One line of work interleaves retrieval with intermediate reasoning, decomposing queries or issuing sub-queries to iteratively gather missing evidence (Press et al., 2023; Trivedi et al., 2023). While effective, such pipelines often increase latency and are sensitive to early-stage retrieval errors.

Other approaches improve long-context retrieval by modifying retrieval units, indexing schemes, or access strategies. Multi-view and content-aware indexing encodes complementary global and local document signals to improve robustness on long documents (Dong et al., 2024). Hierarchical abstractions enable coarse-to-fine retrieval (Lewis et al., 2020a), while dual-view methods combine document-level representations with localized passages (Zhao et al., 2024). Recent benchmarks further show that retrieval effectiveness varies substantially with document length, modality, and evidence dispersion, revealing persistent challenges in multi-hop and multimodal settings (Dong et al., 2025a,b).

Recent analyses highlight the importance of *memory granularity*—document-, passage-, or proposition-level units—with no single granularity performing optimally across tasks (Chen et al., 2024). This motivates adaptive mechanisms that dynamically select or combine granularities conditioned on the query. Related work explores adaptive control in RAG to modulate retrieval and generation behaviors (Huanshuo et al., 2025), but typically assumes a fixed underlying memory representation.

Graph-based methods offer an alternative by explicitly modeling cross-passage associations (Yang et al., 2022; Zhang et al., 2025d; Cheng et al., 2024b). Automatically induced or schemaless graphs support multi-hop retrieval via propagation, complementing dense retrieval when lexical or semantic signals are weak (Gutiérrez et al., 2024). In contrast to approaches that primarily modify retrievers or indices, our work focuses on *query-adaptive composition across heterogeneous memory representations*.

A.3 Structured Memory Representations

To improve information density and retrieval controllability, recent work extracts semi-structured and structured representations from text. Summaries capture global semantics and support high-level retrieval and planning (Xu et al., 2023; Lee et al., 2024). Atomic facts or proposition-like units provide fine-grained, verifiable information that improves precision and reduces irrelevant context (Min et al., 2023; Li et al., 2024b). Document-level information extraction further emphasizes preserving long-range context when constructing structured knowledge (Dong et al., 2021).

Triples and graph representations explicitly encode entity relations, facilitating relational reasoning and multi-hop inference (Anokhin et al., 2024; Baek et al., 2023; Sun et al., 2024). In multimodal settings, unified document parsing systems extract structured representations across text, layout, and visual modalities to support downstream reasoning (Dong et al., 2025c). These representations involve trade-offs between contextual richness and precision: highly structured units may lose discourse context, while less structured units can introduce noise.

Hybrid memory systems combine multiple representations to leverage complementary strengths (Zeng et al., 2024). However, existing approaches often rely on fixed or heuristic mixture policies rather than *query-adaptive* selection. Our

method addresses this limitation by learning a lightweight router that produces query-conditioned allocations over memory types, enabling dynamic routing across granularities.

B Loss Function

We consider the following three alternatives to the Bradley-Terry (BT) loss:

B.1 Cross-Entropy Loss (CE)

Cross-Entropy Loss (CE): This loss treats the task as a standard multi-class classification problem. The model outputs a score vector $\mathbf{s} = [s_1, s_2, s_3, s_4] \in \mathbb{R}^4$, and the loss encourages the score corresponding to the gold method g to be highest. The softmax function is applied to the scores to produce a probability distribution:

$$p_i = \frac{\exp(s_i)}{\sum_{j=1}^{4} \exp(s_j)}, \tag{9}$$

and the loss is defined as:

$$\mathcal{L}_{\text{CE}} = -\log(p_g) = -\log\left(\frac{\exp(s_g)}{\sum_{j=1}^{4} \exp(s_j)}\right). \tag{10}$$

While effective for hard classification, this loss enforces mutual exclusivity among the methods and discourages overlapping or partial contributions.

B.2 Mean Squared Error (MSE)

Mean Squared Error (MSE) Loss: This regression-based loss minimizes the squared distance between the predicted score vector $\mathbf{s}$ and the one-hot ground-truth label vector $\mathbf{y} = [y_1, y_2, y_3, y_4]$, where $y_g = 1$ and $y_j = 0$ for $j \neq g$. The loss is defined as:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{4}\sum_{i=1}^{4}(s_i - y_i)^2. \tag{11}$$

This formulation supports soft contributions but lacks a mechanism to enforce comparative ranking or preference ordering.

B.3 Binary Cross-Entropy(BCE)

Binary Cross-Entropy with Logits Loss (BCE): This loss models each score independently using binary classification. The raw scores $\mathbf{s}$ are passed through the sigmoid function $\sigma(s_i) = 1/(1+e^{-s_i})$, and the loss compares each prediction to its corresponding binary label $y_i \in \{0, 1\}$:

$$\mathcal{L}_{\text{BCE}} = -\tfrac{1}{4}\sum_{i=1}^{4}\left[y_i \log \sigma(s_i) + (1 - y_i)\log(1 - \sigma(s_i))\right] \tag{12}$$

Unlike CE, BCE does not assume exclusivity and can assign high confidence to multiple methods. It also provides a natural extension path to soft supervision with fractional labels.

C Open-sourced models

We follow previous works (Li et al., 2024d) and use the Qwen2 series model in our work, as shown in Table 5.

Section	textbfModel Name
Retriever	`gte-Qwen2-1.5B-instruct`
Agent Backbone	`Qwen2.5-7B-Instruct`
Reward Model	`Qwen2.5-7B-Instruct`
Data Synthesis	`Qwen2.5-72B-Instruct`

Table 5: Open-sourced models used this work.

D Datasets

We evaluate our method on three challenging English multi-hop QA datasets, adapted for long-context reasoning by utilizing full Wikipedia passages. HotpotQA features 2-hop questions authored by native speakers, derived from two related Wikipedia paragraphs. 2WikiMultihopQA consists of questions requiring up to 5 reasoning hops, which are synthesized using manually designed templates to ensure true multi-hop reasoning and prevent shortcut solutions. Questions in MuSiQue are composed from simpler questions to involve up to 4 reasoning hops. They are subsequently paraphrased by human annotators to enhance linguistic naturalness and guard against superficial shortcuts. For our long-context setting, we used the complete Wikipedia passages from which the original supporting and distracting paragraphs were sourced. The statistical information of datasets is provided in Table 6.

Dataset	Avg. # Tokens	# Samples
HotpotQA	1,362	200
2WikiMultihopQA	985	200
MuSiQue	2,558	200

Table 6: The statistics and example of datasets.

E Reward Model training

We construct a dataset comprising 2,400 examples. We adopt a standard LoRA setup with rank 8 on the `q_proj` and `v_proj` layers of the decoder-based backbone LLM. We training the model in 1

Memory Structure	HotPotQA		2WikiQA		MuSiQue		Average	
	EM	F1	EM	F1	EM	F1	EM	F1
Summary	52.00	69.07	34.00	46.24	16.50	32.85	34.17	49.39
Chunks	52.00	68.48	33.50	**47.90**	17.00	32.30	34.17	49.56
Triples	33.50	45.75	24.50	36.98	7.50	17.35	21.83	33.36
Atomic Facts	38.50	51.12	29.50	43.10	10.00	18.04	26.00	37.42
Equal	52.00	70.24	34.00	47.19	11.00	26.17	32.33	47.87
TAG	**53.50**	**71.01**	**35.00**	47.27	**17.50**	**32.42**	**35.33**	**50.23**

Table 7: Results of TAG and individual memory structures on Llama-3.1-8B-Instruct across three multi-hop QA benchmarks.

Memory	Top-k = 25		Top-k = 50		Top-k = 75	
	EM	F1	EM	F1	EM	F1
Random	35.00	47.43	53.00	71.17	52.50	69.13
Best@1	36.50	53.93	47.00	61.07	35.50	55.31
Best@2	37.50	55.17	37.50	50.05	45.50	60.43
Best@3	38.00	55.93	53.00	70.28	51.50	68.88
Equal	40.00	55.48	**58.00**	73.61	54.50	72.63
TAG	**44.00**	**58.34**	**58.00**	**75.73**	**55.50**	**73.71**

Table 8: Hybrid memory performance under different Top-k retrieval settings.

Loss Function	HotPotQA	2WikiQA	MuSiQue
CE	72.99	47.29	32.04
MSE	71.83	48.83	30.84
BCE	61.45	43.91	30.16
BT	**75.73**	**51.22**	**37.92**

Table 9: F1 scores of different training loss functions across datasets.

NVIDIA RTX A6000 GPU using a learning rate of 2×10^{-5}, 3 training epochs, and a batch size of 8.

F Heterogeneous Memory Granularity

Long-context multi-hop reasoning inherently involves information at different levels of granularity. Depending on the query, relevant evidence may take the form of fine-grained facts, explicit relations, temporal segments, or high-level abstractions, making a single fixed representation insufficient. This heterogeneity has been repeatedly observed across prior work. Temporal grounding, multimodal fusion, and consistency modeling rely on intermediate representations at different semantic and temporal scales (Li et al., 2024c; Wei et al., 2025a; Wang et al., 2025a), while uncertainty-aware and multi-view frameworks show that separating heterogeneous evidence is preferable to collapsing it into a monolithic form (Xie et al., 2024, 2025b; Diao et al., 2025b). Routing and modularization mechanisms further exploit such diversity by selecting different computation paths or experts conditioned on the input (Wei et al., 2025b; Chen et al., 2025; Diao et al., 2025a). At the representation level, modern multimodal encoders naturally produce intermediate features with varying abstraction and structure (Li et al., 2025c; Liu et al., 2025, 2024; Huang et al., 2025e; Zhou et al., 2025a), and task constraints such as recoverability motivate distinct intermediate forms (Huang et al., 2025f). Across different system designs, a recurring theme is that intermediate information is processed in a non-uniform manner, with certain signals being prioritized, deferred, or selectively engaged depending on their utility (Huang et al., 2025d; Cheng et al., 2024a; Huang et al., 2025c,b; Li et al., 2025a; Huang et al., 2025a). Such observations are consistent with broader analyses showing that uniformly processing all intermediate signals can lead to inefficient or brittle reasoning (Xie et al., 2025a; Diao et al., 2026). Motivated by these observations, our work treats memory granularity as a first-class design choice and learns a query-conditioned strategy to compose heterogeneous memories for multi-hop reasoning.

Router	HotPotQA	2WikiQA	MuSiQue
Random	71.54	46.21	30.22
Equal	73.61	51.01	32.66
Weak	72.35	47.38	32.06
Fix-G	73.92	50.58	33.68
Fix-P	74.22	**51.36**	35.26
Strong	**75.73**	51.22	**37.92**

Table 10: F1 scores of different router designs across datasets.

You are now an intelligent assistant tasked with meticulously extracting both key elements and triples from a long text.

1. Key Elements: The essential nouns (e.g., characters, times, events, places, numbers), verbs (e.g., actions), and adjectives (e.g., states, feelings) that are pivotal to the text's narrative.
2. Triples: Structured triplets in the format of "subject, relation, object". Each triple should represent a clear and concise fact, relation, or interaction within the observation. You should aim for simplicity and clarity, ensuring that each triplet has no more than 7 words.

Requirements:
#####
1. Ensure that all identified key elements are reflected within the corresponding atomic facts.
2. You should extract key elements and atomic facts comprehensively, especially those that are important and potentially query-worthy and do not leave out details.
3. Whenever applicable, replace pronouns with their specific noun counterparts (e.g., change I, He, She to actual names).
4. Ensure that the key elements and triples you extract are presented in the same language as the original text (e.g., English or Chinese).
5. Avoid Redundant Triples: Do not include irrelevant information like the current location of the agent (e.g., "you, are in, location") or placeholder entities such as "none."
6. Your answer format for each line should be: [Serial Number], [Atomic Facts], [List of Key Elements, separated with '|']
#####

Example:
#####
User: One day, a father and his little son .....
#
Assistant:
1. Father, went to, home | father | went to | home
2. Son, went to, home | son | went to | home
3. Father, accompanied by, son | father | accompanied by | son
4. ...
#####
#
Please strictly follow the above format. Let's begin.

Context:
{context}

(a) Prompt for generating knowledge triples.

You are a helpful assistant responsible for generating a comprehensive summary of the data provided below.
Given one or two atomic facts, and its original descriptions, all related to the atomic facts.
Please concatenate all of these into a single, comprehensive description. Make sure to include information collected from all the descriptions.
If the provided descriptions are contradictory, please resolve the contradictions and provide a single, coherent summary.
Make sure it is written in third person, and include the names so we have the full context.

#######
-Data-
Atomic facts:
{elements}

Original Description List:
{description_list}
#######

Output:

(b) Prompt for generating summaries.

Figure 6: Prompts for generating knowledge triples and summaries.

You are now an intelligent assistant tasked with meticulously extracting both key elements and atomic facts from a conversation history..

1. Key Elements: The essential nouns (e.g., characters, times, events, places, numbers), verbs (e.g., actions), and adjectives (e.g., states, feelings) that are pivotal to the text's narrative.
2. Atomic Facts: The smallest, indivisible facts, presented as concise sentences. These include propositions, theories, existences, concepts, and implicit elements like logic, causality, event sequences, interpersonal relationships, timelines, etc.

Requirements: #####
1. Ensure that all the atomic facts contain full and complete information, reflecting the entire context of the sentence without omitting any key details.
2. Ensure that all identified key elements are reflected within the corresponding atomic facts.
3. You should extract key elements and atomic facts comprehensively, especially those that are important and potentially query-worthy and do not leave out details.
4. Whenever applicable, replace pronouns with their specific noun counterparts (e.g., change I, He, She to actual names).
5. Ensure that the key elements and atomic facts you extract are presented in the same language as the original text (e.g., English or Chinese).
6. You should output a total of key elements and atomic facts that do not exceed 1024 tokens.
7. Your answer format for each line should be: [Serial Number], [Atomic Facts], [List of Key Elements, separated with '|']
#####

Example:
#####
Conversation:
1. Caroline said, "Woohoo Melanie! I passed the adoption agency interviews last Friday! I'm so excited and thankful. This is a big move towards my goal of having a family."
2. Melanie said, "Congrats, Caroline! Adoption sounds awesome. These figurines I bought yesterday remind me of family love. Tell me, what's your vision for the future?"
and shared a photo of a couple of wooden dolls sitting on top of a table.

Atomic Facts and Key Elements:
1. Caroline passed the adoption agency interviews last Friday. | Caroline | adoption agency interviews | last Friday
2. Caroline is excited and thankful for passing the adoption agency interviews. | Caroline | excited | thankful | adoption agency interviews
3. Passing the adoption agency interviews is a big move towards Caroline's goal of having a family. | Caroline | adoption agency interviews | goal | having a family
4. Melanie congratulated Caroline on passing the adoption agency interviews. | Melanie | Caroline | adoption agency interviews | Congratulations
5. Melanie thinks that adoption sounds awesome. | Melanie | Adoption | awesome
6. Melanie bought figurines yesterday. | Melanie | figurines | yesterday
7. The figurines Melanie bought remind her of family love. | Melanie | figurines | family love
8. Melanie asked Caroline about her vision for the future. | Melanie | Caroline | vision for the future
9. Melanie shared a photo of wooden dolls sitting on a table. | Melanie | wooden dolls | table | photo
#####
#
Please strictly follow the above format. Let's begin.

Conversation:
{conversation}

Atomic Facts and Key Elements:

Figure 7: Prompt for generating atomic facts.