Position Paper: Large Language Models need Episodic Memory

Guangfu Hao[1,2], Yuhan Zhang[1,2], Guoqing Ma[1], Yang Chen[1], Frederic Alexandre[3], and Shan Yu[1,*]
[1]*Institute of Automation, Chinese Academy of Sciences, Beijing, China*
[2]*School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China*
[3]*Inria centre at the university of Bordeaux, Neurodegenerative Diseases Institute, Bordeaux, France*
{haoguangfu2021, zhangyuhan2022, maguoqing2022, yang.chen}@ia.ac.cn, Frederic.Alexandre@inria.fr

2025 International Joint Conference on Neural Networks (IJCNN) | 979-8-3315-1042-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/IJCNN64981.2025.11229266

***Abstract*—Large Language Models (LLMs) have made significant strides in natural language processing, yet they still struggle with maintaining consistent personas, tracking long-term dependencies, and adapting to new information during conversations. This paper argues that integrating episodic memory, a crucial component of human cognition, into LLMs is essential for addressing these limitations and advancing towards more human-like artificial intelligence. We first review the current progress and limitations of LLMs in memory retention and processing. Then, we explore recent insights from cognitive neuroscience regarding episodic memory, highlighting its role in human cognition. The paper demonstrates why LLMs need episodic memory, focusing on potential improvements in temporal cognition, autobiographical continuity, reality monitoring, and adaptive learning. We propose theoretical approaches for implementing episodic memory in LLMs, discussing challenges such as efficient storage and retrieval of episodic information, and suggest potential architectures inspired by the hippocampal-cortical memory system. This research aims to bridge the gap between LLMs and human-like cognitive abilities, potentially revolutionizing both LLMs development and our understanding of human memory processes.**

***Index Terms*—Episodic Memory, Large Language Models, Cognitive Architecture**

I. Introduction

LLMs have achieved remarkable feats in natural language processing, demonstrating human-like performance in tasks ranging from text generation to complex reasoning [1]. These models, trained on vast corpora of text, have shown an unprecedented ability to understand and generate contextually appropriate language. However, despite their impressive capabilities, LLMs face fundamental limitations in several critical areas: LLMs struggle to distinguish between factual and fabricated information, leading to the "hallucination" problem [2]; they lack a coherent sense of self across time, impeding their ability to maintain consistent personas or adapt behavior based on past interactions [3]; and unlike humans who can rapidly learn from single experiences, LLMs require extensive retraining to incorporate new knowledge, highlighting their inefficiency in experience-based learning [4]. These limitations stem from the absence of a crucial component of human cognition: episodic memory (see Figure 1).

Episodic memory, first proposed by Endel Tulving [5], allows for the encoding, storage, and retrieval of specific experiences and events. It enables us to mentally time travel, reliving past experiences and projecting ourselves into hypothetical future scenarios [6]. This cognitive function is not merely a record-keeping mechanism; it fundamentally shapes our ability to learn from past experiences, make informed decisions, and adapt our behavior in novel situations [7]. Unlike the flexible and adaptive episodic memory system in humans, LLMs currently lack a comparable mechanism for rapidly integrating new information. As LLMs grow in size and complexity, they become increasingly difficult and expensive to retrain, making frequent updates to their vast stores of factual knowledge impractical [8]. This challenge underscores the need for a more dynamic and flexible memory system in LLMs.

This study first investigates the current capabilities and limitations of LLMs in memory processing, then explores cognitive neuroscience insights on episodic memory and their relevance to LLMs. We subsequently propose theoretical frameworks and potential architectures for implementing episodic memory in LLMs. Finally, we discuss the broader implications of this integration for LLMs development and our understanding of human cognition.

II. Memory Limitations in LLMs

LLMs operate on a principle of statistical pattern recognition, leveraging vast amounts of training data to generate contextually appropriate responses. This approach, while powerful, inherently lacks the dynamic, adaptive memory systems characteristic of human cognition. The memory limitations of LLMs manifest in various ways, such as limited context retention [9], lack of persistent memory across sessions [10], absence of temporal understanding [11], and inability to learn from individual interactions [12]. Such limitations not only affect the models' performance in complex tasks requiring long-term memory and contextual understanding, but also hinder their capacity for personalized interactions and experiential learning.

As the field of LLMs continues to evolve, addressing these memory-related challenges has become a focal point

This work was supported in part by the Strategic Priority Research Program of the Chinese Academy of Sciences (CAS)(XDB1010302), CAS Project for Young Scientists in Basic Research, Grant No. YSBR-041 and the International Partnership Program of the Chinese Academy of Sciences (CAS) (173211KYSB20200021). *Corresponding author: shan.yu@nlpr.ia.ac.cn

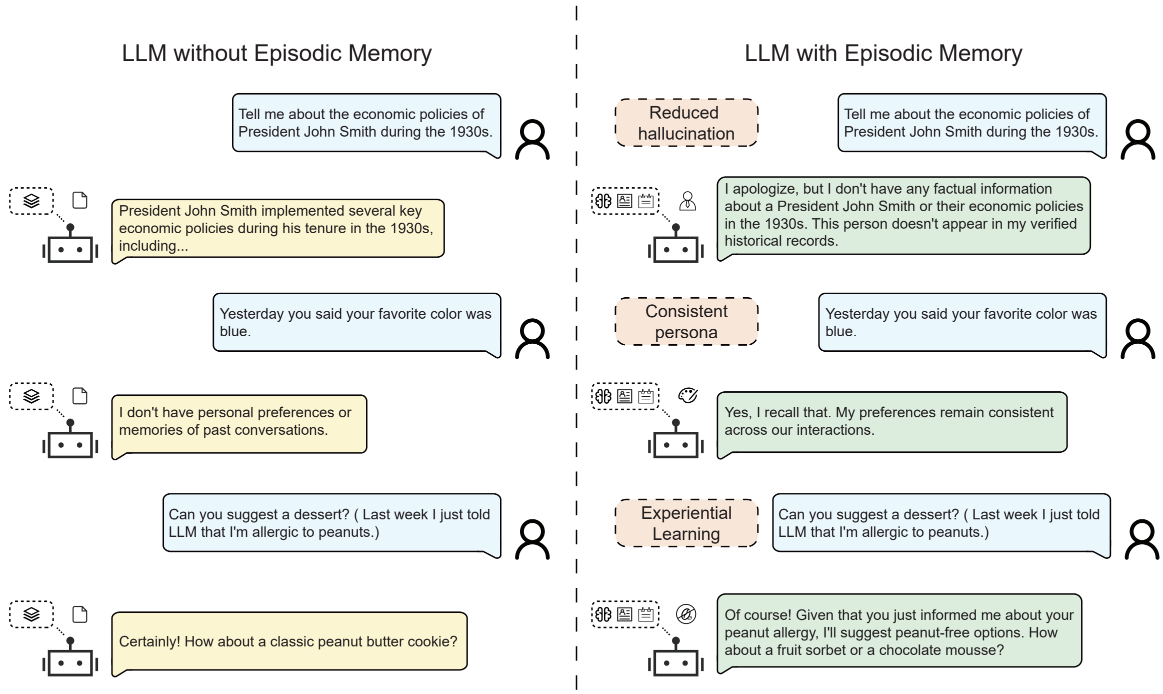

Fig. 1. Comparative illustration of LLM capabilities with and without episodic memory.

of research and development [13]. Efforts to enhance LLMs' memory capabilities have led to significant progress, yet substantial gaps remain between current LLMs and the nuanced, flexible memory processes observed in human cognition. This section examines the current state of memory capabilities in LLMs, explores the limitations of existing approaches, and underscores the critical need for more advanced memory systems, particularly episodic memory, in these models.

A. Current progress in LLM memory capabilities

The concept of memory in LLMs can be broadly divided into short-term and long-term components, analogous to human memory systems.

Short-term memory (STM) in LLMs is primarily represented by the context window, which serves as a form of working memory [14]. Recent advancements have significantly expanded these context windows, with models like GPT-4o [15], Claude 3.5 Sonnet [16], and Gemini 1.5 Pro [17] processing up to 128k, 200k, and 2M tokens respectively. However, larger contexts don't address the fundamental need for persistent, structured long-term memory akin to human cognition.

Long-term memory in LLMs (LTM) is implemented through various external mechanisms and architectural modifications, enabling access, utilization, and potential updating of information beyond the immediate context window. LTM in LLMs encompasses parameterized knowledge encoded in the model's weights, external knowledge bases [18], persistent storage mechanisms [19], and techniques for fine-tuning and continual learning [20].

To enhance the memory capabilities of LLMs, researchers have developed both offline and online methods (see Figure 2). Offline methods focus on improving the model's inherent ability to retain and utilize information through various training and architectural modifications. Online methods, on the other hand, focus on augmenting LLMs' memory capabilities during inference time.

1) *Offline Methods*:

Retraining and Fine-tuning involves updating the entire model with new data or refining pre-trained models on task-specific datasets [21]. Full fine-tuning updates all model weights but risks catastrophic forgetting. Parameter-efficient fine-tuning (PEFT) methods like LoRA [22] selectively update a subset of parameters to mitigate this risk and reduce computational demands. Task-specific, multi-task, and sequential fine-tuning offer different strategies for adapting models to specific domains or multiple tasks simultaneously.

Reinforcement Learning with Human Feedback (RLHF) trains models through interactions with human feedback, indirectly improving the model's ability to utilize relevant information in context [23]. It encompasses techniques such as reward modeling [24], where human evaluators rank model-generated outputs; Proximal Policy Optimization (PPO) [25], which iteratively updates the model's policy while maintaining stability; and comparative ranking and preference learning, which provide nuanced feedback through relative evaluations.

Knowledge-editing directly modifies the model's internal representations to update specific knowledge [26]. Two primary strategies have been developed: locate-then-edit and meta-learning. The locate-then-edit approach, such as Knowledge Neuron [27] and ROME [28], first identifies specific parameters or neurons associated with the target knowledge, then directly modifies these elements. Meta-learning techniques, like MEND [29] and MALMEN [30], employ a secondary model to predict and implement appropriate parameter updates for the main model.

*2) **Online Methods**:*

Retrieval Augmented Generation (RAG) enables LLMs to access and incorporate information from external knowledge bases during inference [31]. The process involves indexing documents into vector representations, retrieving relevant information based on query similarity, and generating responses by synthesizing retrieved context with the input query. Advanced RAG implementations incorporate techniques such as index optimization [32], query expansion [33], and post-retrieval re-ranking [34] to improve relevance and accuracy.

Dynamic Knowledge Graphs (DKGs) represent structured information that can be updated over time [35]. DKGs employ mechanisms for updating nodes and edges, allowing for the integration of new information and the modification of existing relationships. Temporal Knowledge Graphs (TKGs) [36], a specialized subset of DKGs, explicitly encode temporal information, facilitating advanced reasoning tasks such as entity, relation, and time prediction.

Modified-attention enhances LLMs' ability to handle long-range dependencies. The Memorizing Transformer [37] incorporates a non-differentiable memory of recent (key, value) pairs. LongMem [38] utilizes a frozen backbone LLM as a memory encoder and an adaptive residual side-network as a memory retriever and reader. Infini-attention [39] introduces a compressive memory into the standard attention mechanism, enabling efficient processing of infinitely long inputs with bounded memory and computation.

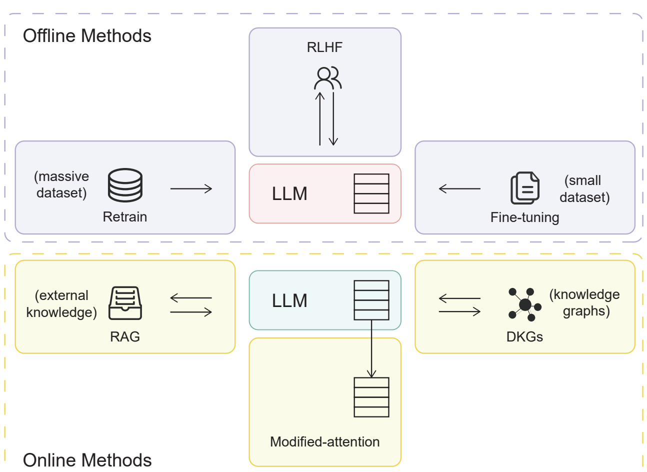

Fig. 2. Implementation of Memory Mechanisms in LLMs.

As research in this field progresses, the focus is shifting towards developing more sophisticated memory architectures that can emulate the nuanced and adaptive nature of human memory and cognition. These advancements aim to address many of the current limitations of LLMs, potentially leading to models capable of more natural, context-aware, and personalized interactions. The ongoing exploration of memory enhancement techniques in LLMs represents a crucial step towards developing more sophisticated systems that can engage in increasingly human-like dialogue and reasoning processes.

B. Limitations of existing memory approaches

Despite these advancements, LLMs continue to face significant memory-related challenges. While expanded context windows and RAG have improved the immediate recall capabilities of LLMs, they fall short of addressing the core issues that separate LLM memory from human-like episodic memory.

*1) **Temporal Understanding**:* Unlike human memory, which naturally distinguishes between recent and distant experiences, LLMs treat all information within their context window with equal temporal significance. This absence of temporal grounding leads to issues in language generation and comprehension. LLMs struggle to accurately track the progression of events, often conflating past, present, and future in their responses [40]. This temporal ambiguity results in inconsistencies in narrative construction, where the model may inadvertently mix historical facts with current events or fail to maintain a coherent timeline in storytelling [41]. The inability to prioritize temporally relevant information hampers the model's capacity to evolve its understanding based on chronological order, particularly problematic in scenarios requiring long-term context maintenance.

*2) **Sense of Self**:* The development of a coherent sense of self over time is another critical area where current LLM architectures fall short. Human cognition is deeply rooted in autobiographical memory, which allows for the maintenance of a consistent persona and the ability to adapt behavior based on cumulative past experiences. LLMs, operating on a session-by-session basis, lack persistent self-awareness [42], leading to inconsistent personality traits across interactions and an inability to form and maintain relationships with users over time. This limitation severely hampers the potential for deep, meaningful, and evolving LLM-human relationships and prevents LLMs from engaging in genuine introspection or self-reflection [43].

*3) **Hallucination**:* The issue of hallucination in LLMs, while not directly a memory problem, is exacerbated by the lack of episodic memory. This manifests in the model's tendency to generate plausible but factually incorrect information, undermining reliability in real-world applications [44]. This problem is particularly acute in scenarios requiring factual accuracy, when providing information in professional fields like healthcare or law [45]. Despite the current progress in the RAG system, researches have shown that the retrievers are not perfect, since noisy retrieval can negatively affect LLM performance [46]. This is especially crucial in multi-hop reasoning scenarios, as irrelevant evidence can trigger a chain of errors. This leads to the generation of plausible but incorrect information, undermining the reliability of the model in critical applications such as healthcare or legal domains.

*4) **Learn from Experiences**:* LLMs lack mechanisms for real-time learning or knowledge accumulation through interaction [47]. This limitation manifests in several critical ways. Unlike humans, who can rapidly learn from single, significant events, LLMs cannot genuinely improve or refine their responses based on feedback or corrections provided during interactions [48]. This limitation is particularly evident in tasks requiring cultural sensitivity, understanding of social dynamics, or adaptation to individual user preferences. The lack of experience-based learning also hampers the LLM's ability to engage in creative problem-solving or generate truly novel ideas [49]. Unlike humans, who can combine disparate experiences to form new insights, LLMs are limited to recombining existing information within their training data, without the ability to incorporate new experiences or external inputs into their knowledge base in a meaningful, persistent way.

Addressing these limitations involves complex considerations of computational efficiency, scalability, and the fundamental architecture of neural networks. As research progresses, the focus must shift towards developing memory architectures that can emulate the nuanced and adaptive nature of human episodic memory, drawing insights from cognitive science, neuroscience, and computer science to bridge the gap between artificial and human-like memory systems.

C. The need for episodic memory in LLMs

The need for episodic memory in LLMs stems from the fundamental role it plays in human cognition, enabling the encoding, storage, and retrieval of specific experiences and events. This capability is essential for overcoming the persistent challenges faced by current LLM architectures, particularly in areas of temporal understanding, sense of self, factual consistency, and experience-based learning.

Episodic memory would provide LLMs with a temporal framework, allowing differentiation between past, present, and future events, and understanding chronological progression. This temporal grounding is vital for maintaining coherence in extended narratives and engaging in future-oriented tasks such as planning and prediction. Moreover, episodic memory implementation would foster a more consistent and evolving sense of self within LLMs, enabling the development of a stable persona over time and fostering more meaningful, personalized interactions.

The issue of hallucination could be significantly mitigated through the integration of episodic memory. By providing a mechanism to distinguish between factual knowledge and generated content, episodic memory could enhance the model's ability to verify information against past experiences, reducing the likelihood of producing inconsistent or false information. This improvement in factual reliability is particularly crucial in applications requiring high accuracy, such as in educational, legal, or healthcare contexts.

Furthermore, episodic memory would enable LLMs to learn from experiences in a manner more akin to human cognition. The ability to encode and retrieve specific interactions would allow these models to refine their responses based on feedback, accumulate context-specific understanding, and engage in more nuanced problem-solving. This capability for experience-based learning is fundamental to developing LLMs that can truly improve and adapt through use, rather than remaining static after initial training.

The implementation of episodic memory in LLMs also holds profound implications for advancing our understanding of human cognition. By attempting to replicate aspects of human memory systems in artificial neural networks, researchers can gain insights into the underlying mechanisms of memory formation, storage, and retrieval.

However, integrating episodic memory into LLMs presents significant technical challenges. It requires developing architectures that can efficiently store, update, and retrieve large volumes of episodic information without compromising the model's performance or generalization capabilities. Addressing these challenges necessitates interdisciplinary approaches, drawing insights from neuroscience, cognitive psychology, and computer science to design memory systems that can seamlessly interface with existing LLM frameworks.

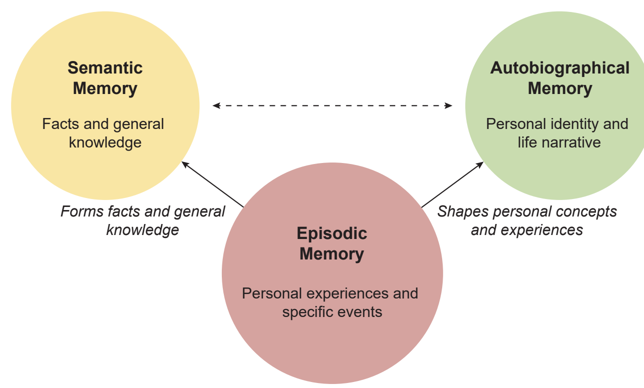

Fig. 3. Relationship between Episodic Memory, Semantic Memory, and Autobiographical Memory.

III. INTEGRATING EPISODIC MEMORY

The integration of episodic memory into LLMs represents a critical frontier in addressing the current limitations of these systems and advancing towards more human-like intelligence. By examining insights from cognitive neuroscience, we can identify key features of episodic memory that are essential for enhancing LLMs' capabilities. In this section, we will explore the theoretical foundations of episodic memory from a cognitive neuroscience perspective, identifying key features that are essential for enhancing LLMs' capabilities. We will then examine the technical approaches, challenges, and potential solutions for implementing episodic memory-like systems within the constraints of current LLMs architectures. This comprehensive analysis aims to provide a roadmap for future research and development in this critical area, potentially transforming the way machines interact with and process information in a manner more akin to human cognitive processes.

A. Insights from cognitive neuroscience

*1) **Definition and characteristics of episodic memory**:* Episodic memory, a sophisticated cognitive system, enables individuals to mentally traverse subjective time, reliving specific past experiences with profound self-awareness [50]. This memory system is intricately connected with semantic memory (general knowledge about the world) and autobiographical memory (personal life story and identity) [51]. The dynamic interplay between these systems is crucial for learning from experiences, applying knowledge in diverse contexts, and maintaining a coherent sense of self across time.

The relationship between episodic, semantic, and autobiographical memory is complex and interdependent [52], as illustrated in Figure 3. Episodic memories often serve as building blocks for semantic knowledge acquisition through abstraction and generalization. [53]. Conversely, semantic knowledge provides a framework for organizing and interpreting new episodic experiences [54]. This bidirectional relationship underlies our ability to learn from experiences and apply that knowledge flexibly. Moreover, episodic memory plays a vital role in forming and maintaining autobiographical memory, contributing to a broader autobiographical knowledge structure that allows individuals to construct a coherent sense of self [55]. Key features of episodic memory include:

Autonoetic Consciousness: A crucial aspect of episodic memory is autonoetic consciousness, which refers to the awareness of oneself in subjective time [56]. This self-awareness enables individuals to not only recall past events but also to project themselves into imagined future scenarios [57], a process known as episodic future thinking. This self-awareness is central to episodic memory's role in self-identity and personal narrative construction.

Temporal-Spatial Contextualization: Unlike semantic memory, episodic memories are firmly anchored to specific times and places, providing a rich contextual framework for the remembered event [58]. This contextual binding allows individuals to differentiate between similar experiences and place them accurately within their personal timeline. The combination of temporal and spatial information aids in the reconstruction of complex events, helping to provide a more complete memory experience.

Constructive Nature: It is imperative to recognize that episodic memory is not a verbatim replay of past events but rather a reconstructive process [59]. During recall, memories are reconstructed based on current goals, knowledge, and context. This constructive nature explains the malleability of memories and their susceptibility to distortion over time [60]. While this characteristic might seem disadvantageous at first glance, it actually serves an adaptive function, allowing for the flexible use of past experiences in novel situations.

Adaptive Function: The functions of episodic memory extend far beyond simple recollection. It provides a basis for memory-based predictions, supporting adaptive behavior in the present and immediate future [61]. By allowing access to specific past experiences, episodic memory enables individuals to make informed decisions and quickly adapt to changing circumstances. Furthermore, it supports novel problem-solving and creative thinking by allowing flexible expression of inferred relationships between elements that were never directly experienced together [62].

These interconnected features collectively enable the vivid recollection and mental re-experiencing of personal events, playing a vital role in shaping individual identity, supporting complex decision-making processes, and enabling adaptive behavior across various temporal contexts.

*2) **Neural Substrates of Episodic Memory**:* The neural substrates of episodic memory form a complex, distributed network involving multiple brain regions, with the hippocampal formation serving as the central hub [63]. This intricate system, which includes the hippocampus proper (CA1-CA4), dentate gyrus (DG), subiculum, and entorhinal cortex (EC), works in concert with various neocortical areas to support the encoding, consolidation, and retrieval of rich, contextually-bound experiences (see Figure 4). Recent advances in neuroimaging, electrophysiology, and computational modeling have provided unprecedented insights into the functional roles and interactions of these structures, revealing a sophisticated interplay that underlies our ability to mentally travel through time, reliving past experiences and imagining future scenarios.

The process of episodic memory formation begins with sensory information processing in primary sensory areas, progressing through a hierarchy of association cortical areas to form multimodal representations. The parahippocampal region, comprising the perirhinal and parahippocampal cortices, serves as a critical intermediary between the neocortex and hippocampus [64], integrating various sensory and associative information. The hippocampal circuit itself follows a largely unidirectional path known as the trisynaptic circuit. The EC projects to the DG via the perforant path, which in turn sends mossy fibers to CA3. CA3 neurons then project to CA1 via Schaffer collaterals, and notably, they also connect to CA1 via an alternative pathway through CA2. Finally, CA1 neurons send their axons to the subiculum and back to the EC, completing the hippocampal loop [65]. This circuit architecture supports the hippocampal indexing theory, wherein the hippocampus stores compressed representations of neocortical activation patterns associated with specific experiences [66]. During memory recall, the hippocampus uses these stored patterns to reactivate the appropriate neocortical areas, resulting in the subjective experience of remembering.

Each subregion of the hippocampal formation contributes unique computational properties to the process of memory encoding and retrieval. The EC, particularly its medial portion, contains grid cells that, along with place cells in the hippocampus and border cells in the subiculum, form a neural circuit for spatial navigation and context representation [67]. The DG, with its sparse firing patterns and powerful mossy fiber synapses, is crucial for pattern separation [68], ensuring that similar experiences are encoded as distinct neural representations. CA3, with its extensive recurrent connections, is thought to act as an autoassociative network, supporting

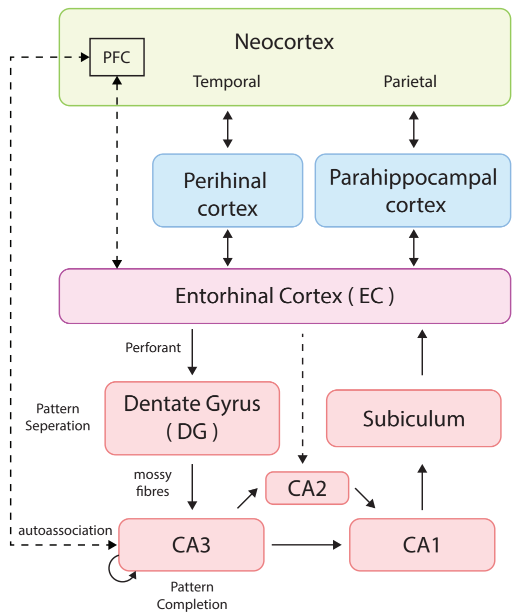

Fig. 4. Neural architecture involved in Episodic Memory, illustrating the interconnections between neocortical regions and the hippocampal formation.

pattern completion [69]. This allows for the retrieval of complete memories from partial cues, a hallmark of episodic recall. CA1, receiving inputs from both CA3 and the EC, is positioned to compare current sensory inputs with stored memories, potentially signaling novelty or facilitating memory updating [70]. Recent research has identified specialized ”time cells” in CA1 [71], elucidating how the hippocampus encodes the temporal aspects of experiences, essential for the sequential nature of episodic memories. CA2, receiving inputs directly from EC, provides an alternative pathway for quick spatial memory recall, bypassing the traditional DG-CA3-CA1 circuit [72]. The subiculum is crucial for recall-associated processes like memory updating and retrieval-driven responses [73].

Beyond the medial temporal lobe, the prefrontal cortex (PFC) plays a multifaceted role in episodic memory processes. The medial PFC, in particular, interacts closely with the hippocampus to integrate new information with existing knowledge structures or schemas [74], facilitating more efficient encoding and retrieval of congruent information. The lateral PFC contributes to strategic encoding and retrieval processes [75], maintaining the temporal order of events within a memory and aiding in the selection and inhibition of memories during retrieval. The neural substrates of episodic memory also extend to regions beyond the hippocampal-prefrontal circuit. The posterior parietal cortex, for instance, has been implicated in the subjective experience of remembering and in the manipulation of retrieved information [76]. The retrosplenial cortex plays a role in contextual memory and in translating between allocentric (world-centered) and egocentric (self-centered) spatial representations [77]. Additionally, the amygdala contributes to the emotional salience of memories [78], influencing both encoding and retrieval processes for emotionally charged experiences.

Recent research has highlighted the importance of oscillatory synchrony between the hippocampus and other brain regions [79], particularly in the theta frequency range (4-12 Hz), facilitating information transfer and memory consolidation [80]. The direction of communication between the PFC and hippocampus appears to be task-dependent, with information flowing from the ventral hippocampus to the PFC during memory delays or contextual cuing, and from the PFC to the hippocampus during memory retrieval and decision-making [81]. Studies have further elucidated the dynamic nature of memory representations across different stages of memory processing [82]. The content of memory representations has been found to include low-level sensory representations, semantic representations, spatiotemporal representations, and schema representations, which are distributed across various regions of the brain. During encoding, maintenance, storage, and retrieval, memory representations involve communication and transformation between the lower-level sensory cortex and higher-level cortex, showing a decrease in the intensity of lower-level sensory representations and an increase in the intensity of higher-level abstract representations [83].

This distributed network supports the encoding, consolidation, and retrieval of rich, contextually bound memories through coordinated activity of multiple specialized neural systems. The dynamic nature of memory representations, from encoding to retrieval, underscores the reconstructive nature of episodic memory and its crucial role in supporting flexible cognition and adaptive behavior.

*3) **Role of Episodic Memory in Cognitive Processes**:* Episodic memory plays a crucial role in various cognitive processes, extending far beyond simple recollection of past events. Its influence permeates multiple aspects of human cognition, shaping our understanding of time, self, reality, and learning [84]. This section explores the multifaceted role of episodic memory in four key cognitive domains.

In temporal cognition, episodic memory enables the mental reconstruction of past experiences and the imagination of future scenarios. Unlike semantic memory, which allows us to know what might happen based on general knowledge, episodic memory enables us to vividly relive past events and project ourselves into hypothetical future situations [85], involving the flexible recombination of elements from past experiences to construct potential future events. This ability is crucial for planning, decision-making, and goal-directed behavior [86]. Recent neuroimaging studies have revealed substantial overlap in the neural substrates supporting episodic memory retrieval and future simulation [87].

Episodic memory also contributes significantly to autobiographical continuity and self-concept. By allowing individuals to mentally revisit past experiences, episodic memory facilitates the construction of a coherent narrative identity that integrates past events with present circumstances and future aspirations [88]. This process of autobiographical reasoning enables individuals to derive meaning from their experiences and understand their own development over time [89]. Moreover, the sharing of episodic memories forms the basis of social bonding and collective memory formation, contributing to the development of shared identities and cultural narratives.

Furthermore, episodic memory plays a critical role in reality monitoring and source attribution processes. The rich contextual details encoded in episodic memories, including perceptual, spatial, and temporal information, provide a basis for distinguishing between events that were personally experienced, information learned from others, and internally generated ideas [90]. This ability, known as source monitoring, is fundamental to maintaining a coherent understanding of reality and one's place within it. In cognitive psychology research, reality monitoring paradigms have demonstrated that the vivid, multi-sensory nature of episodic memories contributes to the ability to differentiate between true memories and false or imagined events [91].

Lastly, episodic memory serves as a cornerstone of experiential learning and adaptive behavior [92]. This capacity is particularly evident in tasks requiring the flexible application of past experiences to novel situations. In analogical problem-solving, individuals draw upon specific episodic memories to identify relevant similarities between past and present situations, facilitating the transfer of solutions across contexts. The detailed nature of episodic memories, including information about the outcomes of past actions, supports counterfactual thinking - the ability to mentally simulate alternative scenarios based on past experiences [93]. This process is crucial for learning from mistakes, optimizing decision-making, and adapting behavior in response to changing environments.

In conclusion, the role of episodic memory in cognitive processes is multifaceted and pervasive. Its contributions to mental time travel, autobiographical reasoning, reality monitoring, and adaptive learning underscore its fundamental importance in human cognition. As research in this field progresses, a deeper understanding of these processes may not only advance our knowledge of human memory but also inform the development of systems capable of human-like flexibility, self-awareness, and experiential learning.

B. Implementing episodic memory in LLMs

The integration of episodic memory into LLMs represents a crucial step towards enhancing their cognitive capabilities and bridging the gap between artificial and human-like intelligence. This section explores the theoretical approaches, technical challenges, and proposed architectures for implementing episodic memory in LLMs.

1) ***Theoretical Approaches amd technical Challenges****:* The implementation of episodic memory in LLMs draws inspiration from cognitive neuroscience and human memory systems. A prominent theoretical framework is the dual-store memory model, distinguishing between STM and LTM. In LLMs, the context window analogizes to STM, while external storage systems represent LTM. This distinction allows for efficient processing of immediate information while maintaining a vast repository of past experiences. The finite context window must be managed through dynamic retrieval and suppression of information from episodic memory, mirroring the human brain's ability to selectively activate relevant memories while inhibiting irrelevant ones. The theoretical approach to implementing episodic memory in LLMs encompasses three primary components: encoding, storage, and retrieval.

Encoding: Involves capturing rich contextual details of each interaction. Key aspects include developing temporal context embeddings, which encode the temporal aspects of information and allow the model to differentiate between recent and older data. Furthermore, multimodal encoding techniques can be implemented to create a unified representation of various types of information, mirroring the rich, multisensory nature of human episodic memories. Algorithms for segmenting continuous input into discrete ”episodes” and mechanisms for tagging memories with emotional salience can enhance the model's ability to prioritize and recall significant events.

Storage: Involves sophisticated systems for maintaining vast amounts of contextual information while ensuring quick access. Hierarchical memory structures promise efficient storage of detailed episodic traces in a lower-level memory store while facilitating the gradual integration into more generalized knowledge structures, mirroring human memory consolidation. Addressing catastrophic forgetting remains a significant challenge, requiring techniques like elastic weight consolidation or algorithms simulating hippocampal-cortical dialogue. These approaches would allow LLMs to acquire new knowledge while preserving critical aspects of their existing memory.

Retrieval: Is crucial for leveraging stored episodic memories. Extending attention mechanisms to query the episodic memory store allows focus on relevant past experiences when processing new inputs. Similarity-based retrieval techniques access memories based on their similarity to the current context or query. Associative retrieval techniques, mimicking the associative nature of human recall, allow the model to access related memories through indirect connections. Implementing both deliberate and spontaneous retrieval mechanisms enhances the model's flexibility and adaptive behavior, with deliberate retrieval initiated by specific cues and spontaneous retrieval occurring automatically based on current context.

In conclusion, the integration of episodic memory into LLMs can enhance these models' ability to maintain context over extended periods, learn from individual experiences, and apply past knowledge to new situations. By addressing the technical challenges associated with encoding, storage, and retrieval of episodic information, researchers can develop LLMs capable of more nuanced understanding of temporal relationships, improved consistency in long-term interactions, and more adaptive responses to novel situations.

2) ***Proposed Architecture****:* Drawing upon the insights from cognitive neuroscience and the theoretical approaches discussed, we propose a novel architecture for integrating episodic memory into LLMs. This architecture, which we term the Episodic-Augmented Large Language Model (EALLM), consists of three main components: the Core Language Model, the Episodic Memory Module, and the Memory-Model Interface. (see Figure 5).

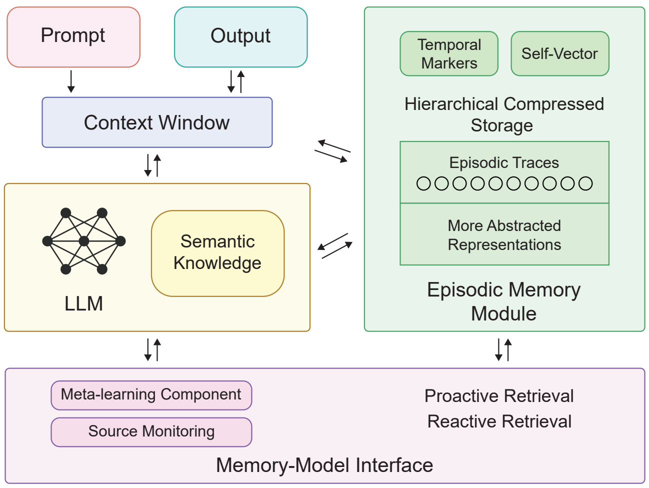

Fig. 5. Episodic-Augmented Large Language Model (EALLM) Architecture.

The Episodic Memory Module stores detailed episodic traces, including rich contextual information such as temporal markers and emotional valence. The module employs a dynamic compression algorithm that gradually consolidates and abstracts these detailed traces into higher-level representations, capturing general patterns and schemas derived from multiple related episodes over time. This structure allows for efficient storage and retrieval while supporting the gradual integration of episodic experiences into semantic-like knowledge structures. The module incorporates a forgetting mechanism, dynamically adjusting the retention of memories based on factors such as recency, emotional salience, and frequency of access. To develop a stable sense of self, the EALLM incorporates a persistent ”self-vector” within the Episodic Memory Module, serving as an anchor point for the model's persona and allowing for the maintenance of consistent traits and preferences over time.

The Memory-Model Interface bridges the Core Language Model and the Episodic Memory Module, extending attention mechanisms to include queries to the Episodic Memory Module. It performs both proactive and reactive memory retrieval, integrating past experiences into current language processing. A meta-learning component allows the model to adapt its memory encoding and retrieval strategies based on task performance and feedback. The Interface also incorporates a source monitoring mechanism, tagging memories with metadata indicating their origin to address the hallucination problem.

To address computational challenges, we propose a tiered storage system that keeps frequently accessed or highly salient memories in faster, more readily available storage, while less critical memories are stored in deeper, more compressed formats. This approach, combined with intelligent caching mechanisms, helps to balance the trade-off between comprehensive memory retention and computational efficiency.

In conclusion, the proposed EALLM architecture represents a significant step towards endowing LLMs with human-like episodic memory capabilities. By combining insights from cognitive neuroscience with advanced machine learning techniques, this architecture has the potential to address many of the current limitations of LLMs, including improved temporal understanding, enhanced consistency in long-term interactions, and more adaptive responses to novel situations.

IV. Conclusions

This paper has explored the critical need for integrating episodic memory into LLMs, highlighting the potential for significant advancements in artificial intelligence that more closely mimics human cognitive abilities. Our examination of current LLMs memory approaches revealed that while beneficial, these methods fall short of replicating the nuanced and adaptive nature of human episodic memory. The identified limitations – inadequate temporal understanding, lack of a consistent sense of self, proneness to hallucination, and inability to learn efficiently from individual experiences – all point to the absence of crucial components that episodic memory provides in human cognition.

By delving into cognitive neuroscience research on episodic memory, we have illuminated key features that could significantly enhance LLM capabilities. The autonoetic consciousness, temporal-spatial contextualization, and adaptive functions of episodic memory offer a roadmap for developing more intelligent LLMs. Our exploration of the neural substrates of episodic memory, particularly the intricate workings of the hippocampal formation and its interactions with neocortical regions, provides valuable insights for designing artificial memory systems that can more faithfully emulate human cognitive processes.

The proposed EALLM architecture represents a significant step towards addressing these limitations. By incorporating a hierarchical, associative memory structure inspired by the hippocampal-cortical memory system, along with mechanisms for dynamic encoding, storage, and retrieval of episodic information, the EALLM has the potential to overcome current constraints in LLMs. This architecture promises improvements in temporal reasoning, maintenance of a consistent persona, mitigation of hallucination issues, and enhancement of experiential learning capabilities.

However, significant challenges remain in implementing episodic memory in LLMs, including issues of computational efficiency, scalability, and ethical implications. Future research should focus on refining the proposed architecture, developing more efficient algorithms for memory consolidation and retrieval, and exploring potential applications across various domains. The integration of episodic memory into LLMs represents a paradigm shift in LLMs development, potentially yielding systems capable of more natural, context-aware, and adaptive interactions, while also deepening our understanding of human cognitive processes.

REFERENCES

[1] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat *et al.*, "Gpt-4 technical report," *arXiv preprint arXiv:2303.08774*, 2023.

[2] G. Perković, A. Drobnjak, and I. Botički, "Hallucinations in llms: Understanding and addressing challenges," in *2024 47th MIPRO ICT and Electronics Convention (MIPRO)*. IEEE, 2024, pp. 2084–2088.

[3] J. Chen, Z. Liu, X. Huang, C. Wu, Q. Liu, G. Jiang, Y. Pu, Y. Lei, X. Chen, X. Wang *et al.*, "When large language models meet personalization: Perspectives of challenges and opportunities," *World Wide Web*, vol. 27, no. 4, p. 42, 2024.

[4] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray *et al.*, "Training language models to follow instructions with human feedback," *Advances in neural information processing systems*, vol. 35, pp. 27 730–27 744, 2022.

[5] E. Tulving *et al.*, "Episodic and semantic memory," *Organization of memory*, vol. 1, no. 381-403, p. 1, 1972.

[6] M. A. Conway, "Episodic memories," *Neuropsychologia*, vol. 47, no. 11, pp. 2305–2313, Sep. 2009.

[7] D. C. Rubin, "The Basic-Systems Model of Episodic Memory," *Perspectives on Psychological Science*, vol. 1, no. 4, pp. 277–311, Dec. 2006.

[8] D. Li, A. S. Rawat, M. Zaheer, X. Wang, M. Lukasik, A. Veit, F. X. Yu, and S. Kumar, "Large language models with controllable working memory," *ArXiv*, vol. abs/2211.05110, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:253420654

[9] S. Chen, S. Wong, L. Chen, and Y. Tian, "Extending context window of large language models via positional interpolation," *arXiv preprint arXiv:2306.15595*, 2023.

[10] C. Packer, V. Fang, S. G. Patil, K. Lin, S. Wooders, and J. E. Gonzalez, "Memgpt: Towards llms as operating systems," *arXiv preprint arXiv:2310.08560*, 2023.

[11] B. Fatemi, M. Kazemi, A. Tsitsulin, K. Malkan, J. Yim, J. Palowitch, S. Seo, J. Halcrow, and B. Perozzi, "Test of time: A benchmark for evaluating llms on temporal reasoning," *arXiv preprint arXiv:2406.09170*, 2024.

[12] J. Liang, Z. Wang, Z. Ma, J. Li, Z. Zhang, X. Wu, and B. Wang, "Online training of large language models: Learn while chatting," *arXiv preprint arXiv:2403.04790*, 2024.

[13] J. Li and J. Li, "Memory, consciousness and large language model," *arXiv preprint arXiv:2401.02509*, 2024.

[14] S. Pawar, S. Tonmoy, S. Zaman, V. Jain, A. Chadha, and A. Das, "The what, why, and how of context length extension techniques in large language models–a detailed survey," *arXiv preprint arXiv:2401.07872*, 2024.

[15] OpenAI, "Hello gpt-4o," May 2024, accessed 10-06-2024. [Online]. Available: https://openai.com/index/hellogpt-4o/

[16] A. Anthropic, "Claude 3.5 sonnet," Anthropic, Claude-3.5 Model Card, 2024.

[17] M. Reid, N. Savinov, D. Teplyashin, D. Lepikhin, T. Lillicrap, J.-b. Alayrac, R. Soricut, A. Lazaridou, O. Firat, J. Schrittwieser *et al.*, "Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context," *arXiv preprint arXiv:2403.05530*, 2024.

[18] J. Cui, Z. Li, Y. Yan, B. Chen, and L. Yuan, "Chatlaw: Open-source legal large language model with integrated external knowledge bases," *arXiv preprint arXiv:2306.16092*, 2023.

[19] O. Topsakal and T. C. Akinci, "Creating large language model applications utilizing langchain: A primer on developing llm apps fast," in *International Conference on Applied Engineering and Natural Sciences*, 2023, pp. 1050–1056.

[20] T. Wu, L. Luo, Y.-F. Li, S. Pan, T.-T. Vu, and G. Haffari, "Continual learning for large language models: A survey," *arXiv preprint arXiv:2402.01364*, 2024.

[21] X. Lin, W. Wang, Y. Li, S. Yang, F. Feng, Y. Wei, and T.-S. Chua, "Data-efficient fine-tuning for llm-based recommendation," *arXiv preprint arXiv:2401.17197*, 2024.

[22] Z. Han, C. Gao, J. Liu, S. Q. Zhang *et al.*, "Parameter-efficient fine-tuning for large models: A comprehensive survey," *arXiv preprint arXiv:2403.14608*, 2024.

[23] G. K.-M. Liu, "Transforming human interactions with ai via reinforcement learning with human feedback (rlhf)," *Massachusetts Institute of Technology*, 2023.

[24] Y. Bai, A. Jones, K. Ndousse, A. Askell, A. Chen, N. DasSarma, D. Drain, S. Fort, D. Ganguli, T. Henighan *et al.*, "Training a helpful and harmless assistant with reinforcement learning from human feedback," *arXiv preprint arXiv:2204.05862*, 2022.

[25] T. Wu, B. Zhu, R. Zhang, Z. Wen, K. Ramchandran, and J. Jiao, "Pairwise proximal policy optimization: Harnessing relative feedback for llm alignment," 2023. [Online]. Available: https://arxiv.org/abs/2310.00212

[26] N. Zhang, Y. Yao, B. Tian, P. Wang, S. Deng, M. Wang, Z. Xi, S. Mao, J. Zhang, Y. Ni *et al.*, "A comprehensive study of knowledge editing for large language models," *arXiv preprint arXiv:2401.01286*, 2024.

[27] D. Dai, L. Dong, Y. Hao, Z. Sui, B. Chang, and F. Wei, "Knowledge neurons in pretrained transformers," *arXiv preprint arXiv:2104.08696*, 2021.

[28] K. Meng, D. Bau, A. Andonian, and Y. Belinkov, "Locating and editing factual associations in gpt," *Advances in Neural Information Processing Systems*, vol. 35, pp. 17 359–17 372, 2022.

[29] E. Mitchell, C. Lin, A. Bosselut, C. Finn, and C. D. Manning, "Fast model editing at scale," *arXiv preprint arXiv:2110.11309*, 2021.

[30] C. Tan, G. Zhang, and J. Fu, "Massive editing for large language models via meta learning," *arXiv preprint arXiv:2311.04661*, 2023.

[31] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel *et al.*, "Retrieval-augmented generation for knowledge-intensive nlp tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020.

[32] J. Chen, H. Lin, X. Han, and L. Sun, "Benchmarking large language models in retrieval-augmented generation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 16, 2024, pp. 17 754–17 762.

[33] X. Ma, Y. Gong, P. He, H. Zhao, and N. Duan, "Query rewriting for retrieval-augmented large language models," *arXiv preprint arXiv:2305.14283*, 2023.

[34] R. Nogueira, Z. Jiang, and J. Lin, "Document ranking with a pretrained sequence-to-sequence model," *arXiv preprint arXiv:2003.06713*, 2020.

[35] M. Farghaly, M. Mounir, M. Aref, and S. M. Moussa, "Investigating the challenges and prospects of construction models for dynamic knowledge graphs," *IEEE Access*, vol. 12, pp. 40 973–40 988, 2024.

[36] J. Wang, K. Sun, L. Luo, W. Wei, Y. Hu, A. W.-C. Liew, S. Pan, and B. Yin, "Large language models-guided dynamic adaptation for temporal knowledge graph reasoning," *arXiv preprint arXiv:2405.14170*, 2024.

[37] Y. Wu, M. N. Rabe, D. Hutchins, and C. Szegedy, "Memorizing transformers," 2022. [Online]. Available: https://arxiv.org/abs/2203.08913

[38] W. Wang, L. Dong, H. Cheng, X. Liu, X. Yan, J. Gao, and F. Wei, "Augmenting language models with long-term memory," in *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., vol. 36. Curran Associates, Inc., 2023, pp. 74 530–74 543.

[39] T. Munkhdalai, M. Faruqui, and S. Gopal, "Leave no context behind: Efficient infinite context transformers with infini-attention," *arXiv preprint arXiv:2404.07143*, 2024.

[40] Y. Wang and Y. Zhao, "Tram: Benchmarking temporal reasoning for large language models," *arXiv preprint arXiv:2310.00835*, 2023.

[41] C. Cornell, S. Jin, and Q. Zhang, "The role of episodic memory in storytelling: Comparing large language models with humans," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 46, 2023.

[42] M. Shanahan, K. McDonell, and L. Reynolds, "Role play with large language models," *Nature*, vol. 623, no. 7987, pp. 493–498, 2023.

[43] Y. Liang, Z. Song, H. Wang, and J. Zhang, "Learning to trust your feelings: Leveraging self-awareness in llms for hallucination mitigation," 2024. [Online]. Available: https://arxiv.org/abs/2401.15449

[44] Z. Xu, S. Jain, and M. Kankanhalli, "Hallucination is inevitable: An innate limitation of large language models," 2024. [Online]. Available: https://arxiv.org/abs/2401.11817

[45] R. Azamfirei, S. R. Kudchadkar, and J. Fackler, "Large language models and the perils of their hallucinations," *Critical Care*, vol. 27, no. 1, p. 120, 2023.

[46] Y. Wang, P. Li, M. Sun, and Y. Liu, "Self-knowledge guided retrieval augmentation for large language models," in *Conference on Empirical Methods in Natural Language Processing*, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:263828724

[47] J. Zheng, S. Qiu, C. Shi, and Q. Ma, "Towards lifelong learning of large language models: A survey," 2024. [Online]. Available: https://arxiv.org/abs/2406.06391

[48] T. Carta, C. Romac, T. Wolf, S. Lamprier, O. Sigaud, and P.-Y. Oudeyer, "Grounding large language models in interactive environments with online reinforcement learning," in *International Conference on Machine Learning*. PMLR, 2023, pp. 3676–3713.
[49] P. Mehrotra, A. Parab, and S. Gulwani, "Enhancing creativity in large language models through associative thinking strategies," *arXiv preprint arXiv:2405.06715*, 2024.
[50] E. Tulving, "Episodic memory: From mind to brain," *Annual review of psychology*, vol. 53, no. 1, pp. 1–25, 2002.
[51] M. A. Conway and C. W. Pleydell-Pearce, "The construction of autobiographical memories in the self-memory system." *Psychological review*, vol. 107, no. 2, p. 261, 2000.
[52] C. P. Múnera, C. Lomlomdjian, B. Gori, V. Terpiluk, N. Medel, P. Solís, and S. Kochen, "Episodic and semantic autobiographical memory in temporal lobe epilepsy," *Epilepsy Research and Treatment*, vol. 2014, no. 1, p. 157452, 2014.
[53] A. Boyle, "The mnemonic functions of episodic memory," *Philosophical Psychology*, vol. 35, no. 3, pp. 327–349, Apr. 2022.
[54] D. L. Greenberg and M. Verfaellie, "Interdependence of episodic and semantic memory: Evidence from neuropsychology," *Journal of the International Neuropsychological society*, vol. 16, no. 5, pp. 748–753, 2010.
[55] M. J. Kahana, M. W. Howard, and S. M. Polyn, "Associative retrieval processes in episodic memory," 2008.
[56] M. A. Wheeler, D. T. Stuss, and E. Tulving, "Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness." *Psychological bulletin*, vol. 121, no. 3, p. 331, 1997.
[57] S. B. Klein, "Autonoetic consciousness: Reconsidering the role of episodic memory in future-oriented self-projection," *Quarterly Journal of Experimental Psychology*, vol. 69, no. 2, pp. 381–401, 2016.
[58] B. Subagdja and A.-H. Tan, "Neural modeling of sequential inferences and learning over episodic memory," *Neurocomputing*, vol. 161, pp. 229–242, 2015.
[59] J. B. Mahr and G. Csibra, "Why do we remember? the communicative function of episodic memory," *Behavioral and brain sciences*, vol. 41, p. e1, 2018.
[60] D. L. Schacter, S. A. Guerin, and P. L. S. Jacques, "Memory distortion: An adaptive perspective," *Trends in cognitive sciences*, vol. 15, no. 10, pp. 467–474, 2011.
[61] K. Duncan, A. Sadanand, and L. Davachi, "Memory's penumbra: episodic memory decisions induce lingering mnemonic biases," *Science*, vol. 337, no. 6093, pp. 485–487, 2012.
[62] K. P. Madore, D. R. Addis, and D. L. Schacter, "Creativity and memory: Effects of an episodic-specificity induction on divergent thinking," *Psychological science*, vol. 26, no. 9, pp. 1461–1468, 2015.
[63] T. A. Allen and N. J. Fortin, "The evolution of episodic memory," *Proceedings of the National Academy of Sciences*, vol. 110, no. supplement_2, pp. 10 379–10 386, 2013.
[64] E. M. Aminoff, K. Kveraga, and M. Bar, "The role of the parahippocampal cortex in cognition," *Trends in cognitive sciences*, vol. 17, no. 8, pp. 379–390, 2013.
[65] T. Klausberger and P. Somogyi, "Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations," *Science*, vol. 321, no. 5885, pp. 53–57, 2008.
[66] T. J. Teyler and J. W. Rudy, "The hippocampal indexing theory and episodic memory: updating the index," *Hippocampus*, vol. 17, no. 12, pp. 1158–1169, 2007.
[67] E. T. Rolls, S. M. Stringer, and T. Elliot, "Entorhinal cortex grid cells can map to hippocampal place cells by competitive learning," *Network: Computation in Neural Systems*, vol. 17, no. 4, pp. 447–465, 2006.
[68] R. Kassab and F. Alexandre, "Pattern separation in the hippocampus: distinct circuits under different conditions," *Brain Structure and Function*, vol. 223, no. 6, pp. 2785–2808, 2018.
[69] N. C. Hindy, F. Y. Ng, and N. B. Turk-Browne, "Linking pattern completion in the hippocampus to predictive coding in visual cortex," *Nature neuroscience*, vol. 19, no. 5, pp. 665–667, 2016.
[70] J. E. Lisman and N. A. Otmakhova, "Storage, recall, and novelty detection of sequences by the hippocampus: elaborating on the socratic model to account for normal and aberrant effects of dopamine," *Hippocampus*, vol. 11, no. 5, pp. 551–568, 2001.
[71] W. Mau, D. W. Sullivan, N. R. Kinsky, M. E. Hasselmo, M. W. Howard, and H. Eichenbaum, "The same hippocampal ca1 population simultaneously codes temporal information over multiple timescales," *Current Biology*, vol. 28, no. 10, pp. 1499–1508, 2018.
[72] M. W. Jones and T. J. McHugh, "Updating hippocampal representations: Ca2 joins the circuit," *Trends in neurosciences*, vol. 34, no. 10, pp. 526–535, 2011.
[73] D. S. Roy, T. Kitamura, T. Okuyama, S. K. Ogawa, C. Sun, Y. Obata, A. Yoshiki, and S. Tonegawa, "Distinct neural circuits for the formation and retrieval of episodic memories," *Cell*, vol. 170, no. 5, pp. 1000–1012, 2017.
[74] M. T. Van Kesteren, S. F. Beul, A. Takashima, R. N. Henson, D. J. Ruiter, and G. Fernández, "Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: from congruent to incongruent," *Neuropsychologia*, vol. 51, no. 12, pp. 2352–2359, 2013.
[75] M. Sandrini, N. Censor, J. Mishoe, and L. G. Cohen, "Causal role of prefrontal cortex in strengthening of episodic memories through reconsolidation," *Current Biology*, vol. 23, no. 21, pp. 2181–2184, 2013.
[76] C. Sestieri, G. L. Shulman, and M. Corbetta, "The contribution of the human posterior parietal cortex to episodic memory," *Nature Reviews Neuroscience*, vol. 18, no. 3, pp. 183–192, 2017.
[77] A. S. Alexander and D. A. Nitz, "Retrosplenial cortex maps the conjunction of internal and external spaces," *Nature neuroscience*, vol. 18, no. 8, pp. 1143–1151, 2015.
[78] S. B. Hamann, T. D. Ely, S. T. Grafton, and C. D. Kilts, "Amygdala activity related to enhanced memory for pleasant and aversive stimuli," *Nature neuroscience*, vol. 2, no. 3, pp. 289–293, 1999.
[79] L. L. Colgin, "Oscillations and hippocampal–prefrontal synchrony," *Current opinion in neurobiology*, vol. 21, no. 3, pp. 467–474, 2011.
[80] M. Geva-Sagiv, E. A. Mankin, D. Eliashiv, S. Epstein, N. Cherry, G. Kalender, N. Tchemodanov, Y. Nir, and I. Fried, "Augmenting hippocampal–prefrontal neuronal synchrony during sleep enhances memory consolidation in humans," *Nature neuroscience*, vol. 26, no. 6, pp. 1100–1110, 2023.
[81] M. D. Brockmann, B. Pöschel, N. Cichon, and I. L. Hanganu-Opatz, "Coupled oscillations mediate directed interactions between prefrontal cortex and hippocampus of the neonatal rat," *Neuron*, vol. 71, no. 2, pp. 332–347, 2011.
[82] G. Xue, "From remembering to reconstruction: The transformative neural representation of episodic memory," *Progress in Neurobiology*, vol. 219, p. 102351, 2022.
[83] X. Xiao, Q. Dong, J. Gao, W. Men, R. A. Poldrack, and G. Xue, "Transformed neural pattern reinstatement during episodic memory retrieval," *Journal of Neuroscience*, vol. 37, no. 11, pp. 2986–2998, 2017.
[84] A. M. Nuxoll and J. E. Laird, "Extending Cognitive Architecture with Episodic Memory," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2007, pp. 1560–1564.
[85] R. E. Beaty, P. Seli, and D. L. Schacter, "Thinking about the past and future in daily life: an experience sampling study of individual differences in mental time travel," *Psychological Research*, vol. 83, pp. 805–816, 2019.
[86] K. K. Szpunar, D. R. Addis, V. C. McLelland, and D. L. Schacter, "Memories of the future: New insights into the adaptive value of episodic memory," *Frontiers in behavioral neuroscience*, vol. 7, p. 47, 2013.
[87] D. R. Addis, A. T. Wong, and D. L. Schacter, "Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration," *Neuropsychologia*, vol. 45, no. 7, pp. 1363–1377, 2007.
[88] M. E. Berryhill, L. Phuong, L. Picasso, R. Cabeza, and I. R. Olson, "Parietal lobe and episodic memory: bilateral damage causes impaired free recall of autobiographical memory," *Journal of Neuroscience*, vol. 27, no. 52, pp. 14 415–14 423, 2007.
[89] S. B. Klein and S. Nichols, "Memory and the sense of personal identity," *Mind*, vol. 121, no. 483, pp. 677–702, 2012.
[90] B. G. Kuhlmann, N. Symeonidou, H. Tanyas, and L. Wulff, "Remembering and reconstructing episodic context: An overview of source monitoring methods and behavioral findings," *Psychology of learning and motivation*, vol. 75, pp. 79–124, 2021.
[91] J. W. Sherman and G. R. Bessenoff, "Stereotypes as source-monitoring cues: On the interaction between episodic and semantic memory," *Psychological Science*, vol. 10, no. 2, pp. 106–110, 1999.
[92] T. Bast, "Toward an integrative perspective on hippocampal function: from the rapid encoding of experience to adaptive behavior," *Reviews in the Neurosciences*, vol. 18, no. 3-4, pp. 253–282, 2007.
[93] D. Hassabis and E. A. Maguire, "Deconstructing episodic memory with construction," *Trends in cognitive sciences*, vol. 11, no. 7, pp. 299–306, 2007.