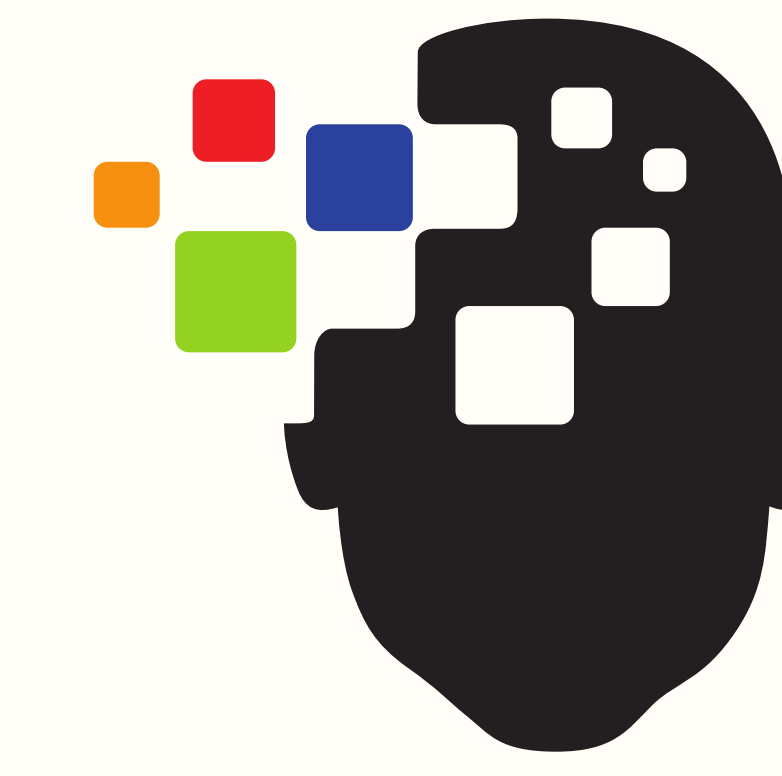


Memoria: Resolving Fateful Forgetting Problem through Human-Inspired Memory Architecture

Sangjun Park cosmoquester@gmail.com, JinYeong Bak



ICML
International Conference
On Machine Learning

Key Contributions

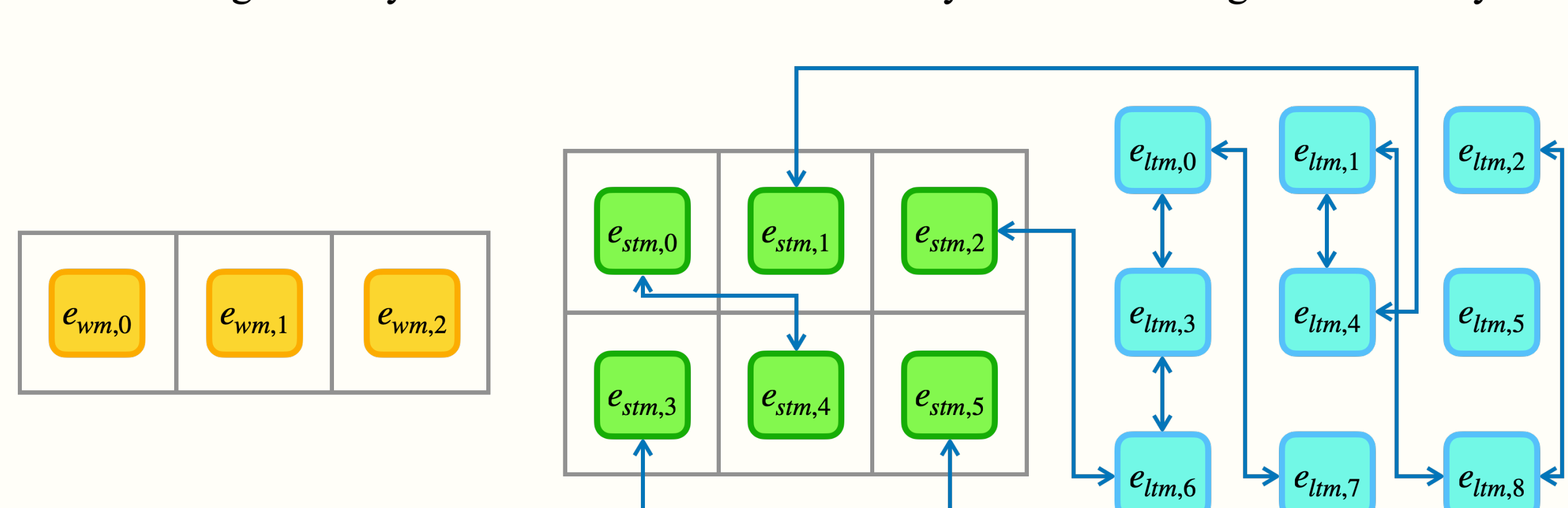
- Designed **Memoria** as an **external memory framework** for neural networks, incorporating various **theories of human memory**.
- Proved that Memoria **successfully addresses fateful forgetting, representing its superior performance** across the sorting, language modeling, and text classification tasks.
- Discovered the **similarity of long-term memory between Memoria and humans** by showing that Memoria reproduces the three well-known effects of human memory.

Motivation

- Traditional external memory approaches focus on retaining recent information, leading to **fateful forgetting after certain time steps, regardless of how much they are important**.
- We address the issue of fateful forgetting by **solving critical problems in memory system: long-term importance, selective preservation, cue-based activation, and memory searching**.

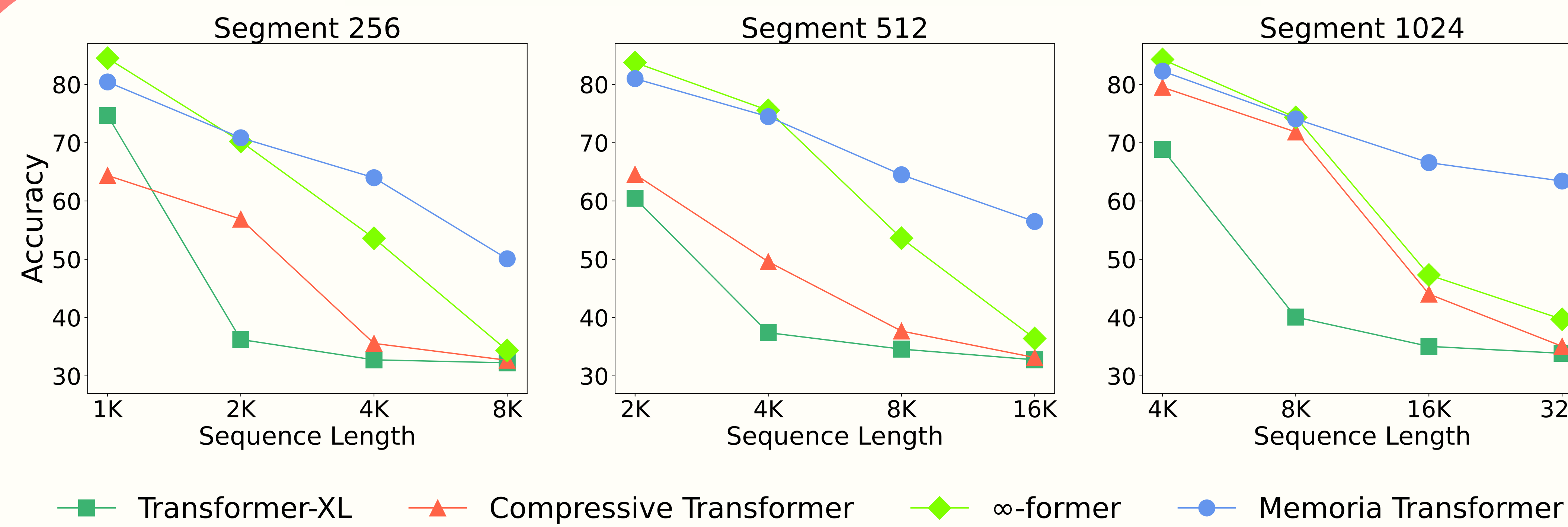
Structure

Working Memory Short-term Memory Long-term Memory



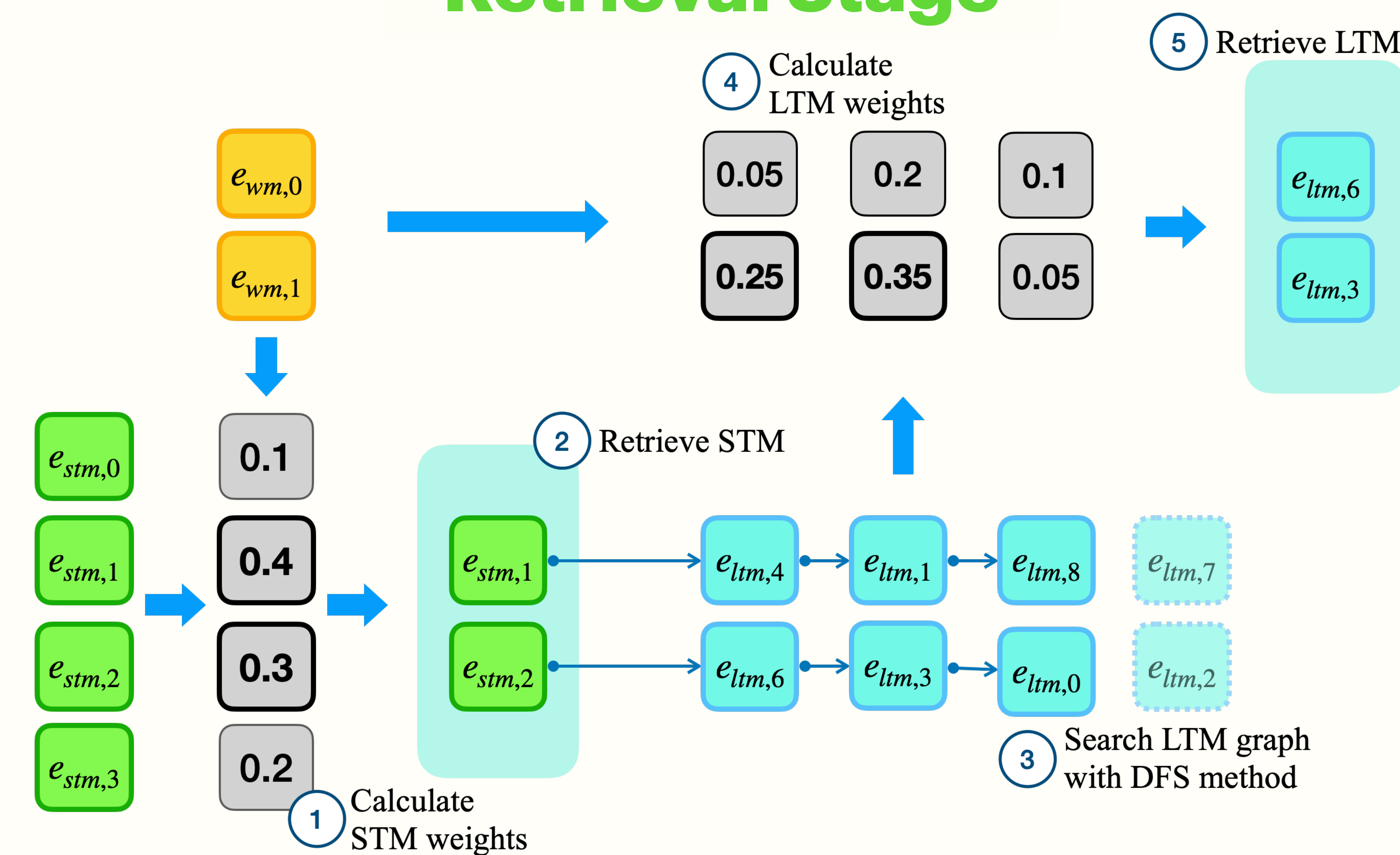
- Memoria stores engrams in three categories: working, short-term, and long-term memory.

Performance Evaluation



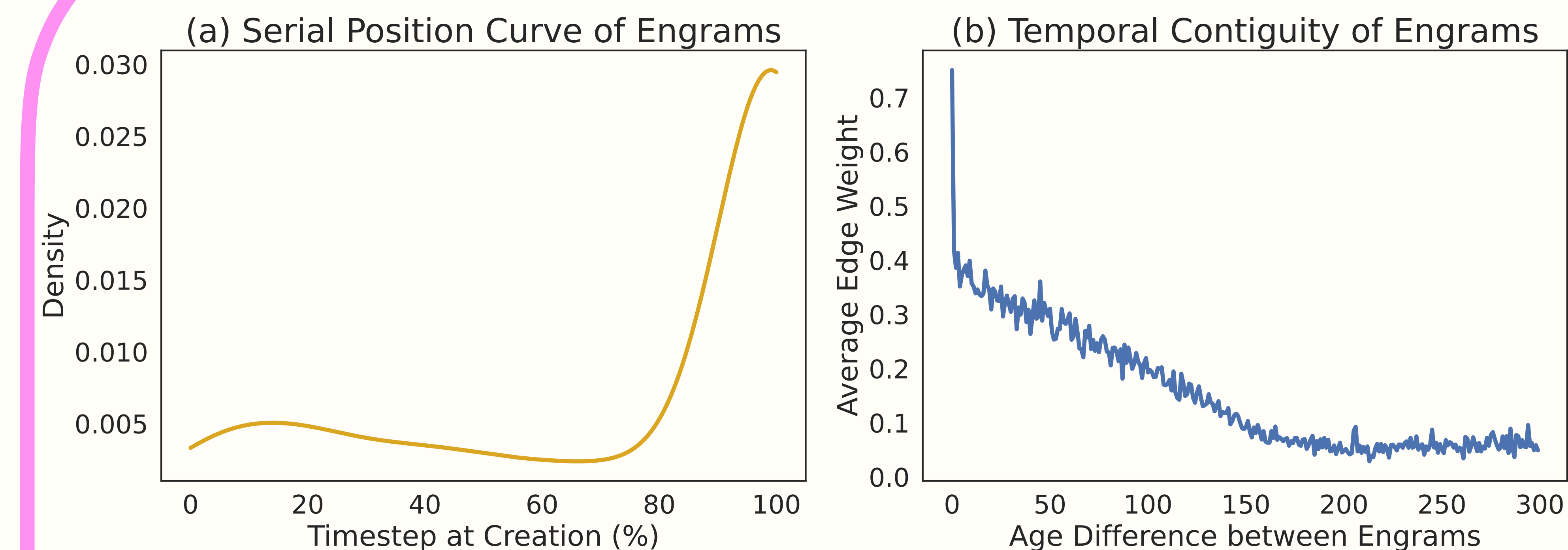
- In the sorting task, Memoria exhibits greater robustness compared to other baselines as the input sequence length increases. The result demonstrates Memoria Transformer successfully handles the issue of **fateful forgetting**.
- The ablation study shows the **complementary functions of the three memories** where the impact of STM/LTM grows with longer inputs.

Retrieval Stage



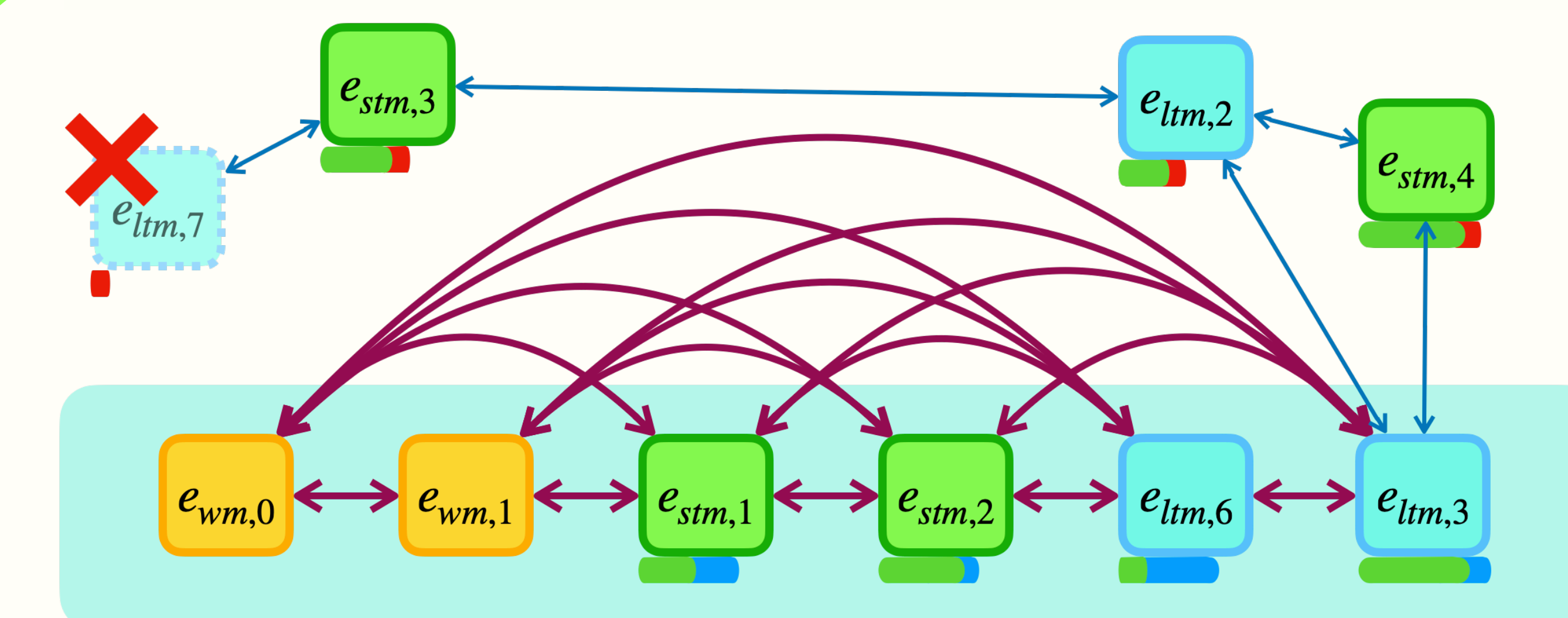
- Memoria utilizes working memory to identify relevant engrams in short-term and long-term memory based on embedding distances to working memory.
- The weights between the engrams and working memory denote the levels of association with the current context, enabling **cue-based activation**.
 - The internal weights of Memoria are utilized for **memory searching**.

Psychological Memory Effects



- In Figure (a), the fact that **engrams are not just concentrated in recent timesteps but are evenly preserved during the entire period** illustrates the **effective mitigation of fateful forgetting**.
- The distribution of engram creation time shows that initial and recent information is better maintained than information in the middle stages after passing all time steps. This tendency displays the patterns of **the primacy effect and recency effect**.
- Figure (b) illustrates that when the age difference is small, the edge weight is high, indicating **the temporal contiguity effect**.

Memorize & Forget Stage



- The connections between activated engrams are strengthened across all pairs, raising the probabilities for them to be retrieved together afterward.
- Differential allocation of lifespans depending on each engram's contribution enables **selective preservation** based on **long-term importance**.

Links

