Implementing Long-Term Memory in Conversational Agents

By Arsalan Khan

Abstract

This paper explores the architectural components required for implementing long-term memory in conversational agents. Drawing from practical examples such as the Lovable agent platform, it outlines how persistent memory can be achieved using semantic embeddings, episodic logs, and structured rules. The same principles are mapped to a proposed system, SophieAI, which adopts a modular execution model designed to operate across domains like physiology, finance, and AI tooling.

1. Introduction

Persistent memory is an essential component for conversational agents that aim to maintain context, personalization, and continuity over time. This paper evaluates how systems like Lovable simulate memory, and proposes a structured memory model for SophieAI.

2. Memory Layers in Intelligent Agents

Four primary memory layers are required: semantic (contextual recall via vector embeddings), declarative (user facts and preferences), episodic (event logging), and procedural (if/then behavior rules).

3. Memory Architecture in Lovable

Lovable uses embedding models to convert conversations into vector space and stores them in a vector database. User traits are maintained in structured profiles, and important moments are distilled into episodic logs. Rules guide actions based on context.

4. Proposed System: SophieAI

SophieAI replicates this architecture using open components: ChromaDB for vector memory, SQLite for episodes, YAML files for rules, and flat files for traits. A prompt builder assembles context based on query domain.

5. Real-World Use Cases

SophieAI can track physical states like fascia tension, recall previous successful protocols, or act upon hormonal states by referencing its memory structure. It can also provide technical answers in domains like documentation or trading based on retrieved knowledge.

6. Conclusion

Long-term memory is achievable through structured, modular architecture. Systems like Lovable prove the viability, and SophieAI extends the concept to multi-domain reasoning and execution.