Toward Trustworthy AGI: A Tiered Memory Architecture Combining DAG, Caching, and Moral Contracts

By Janay Harris
Independent AI Architect & Researcher
Email: janayharris@synaptik-core.dev

Abstract

As AI systems move toward autonomy and generality, statelessness, limited memory, and opaque ethics remain critical barriers. This paper proposes a lightweight cognitive architecture designed to bridge LLMs with the requirements of agentic, trustworthy AGI. The system integrates a **Directed Acyclic Graph (DAG)** for long-term memory and planning, a **local cache (e.g., SQLite)** for low-latency recall, and a "**moral contract**" layer for enforceable ethics. Unlike vector-only or retrieval-augmented systems, this design supports structured memory, grounded reasoning, and ethical alignment. All components are optimized for edge deployment. It enables agents to not just generate responses, but to remember, reflect, and reason. This forms a foundational step toward scalable, safe cognitive systems.

The Challenge: Bridging LLMs to AGI

Despite recent breakthroughs in large language models, current AI systems remain fundamentally limited in their capacity to function as true agents. These models operate without memory, without accountability, and without context across interactions. As a result, they lack the ability to *reflect*, *learn causally*, or *adapt ethically* over time. To achieve Artificial General Intelligence (AGI), we must go beyond token prediction and implement architectural structures that support long-term memory**,** reasoned decision-making, and auditable moral alignment**.** This is especially important in real-world, dynamic environments.

The Proposed Cognitive Architecture

To address the limitations of stateless LLMs, this architecture introduces a hybrid, tiered design that bridges neural generation with symbolic reasoning. It combines the adaptability of language models with structured memory, enforceable ethics, and modular planning components.

At the center is a Directed Acyclic Graph (DAG) that stores interactions, causal links, user feedback, and agent reflections as symbolic memory nodes. This structure enables agents to recall, revise, and plan based on lived experience, supporting long-term, contextual reasoning.

A SQLite-based cache functions as a synthetic hippocampus, accelerating access to recent memory and enabling fast, low-resource inference. While the DAG manages structured knowledge over time, the cache ensures responsiveness at the edge.

To enforce ethical behavior, the architecture includes Moral Contracts—explicit, auditable rules that define behavioral boundaries. These are evaluated by a dedicated Ethos Agent, introducing alignment constraints grounded in symbolic logic.

Together, these components form a lightweight, modular framework built for deployment on low-power devices. By separating memory, inference, ethics, and reflection into distinct agents (e.g., *Librarian, Ethos, Memory, Audit*), the system supports structured cognition and verifiable trust.

This hybrid approach not only connects LLMs to AGI goals but also integrates neural flexibility with symbolic stability. It enables agents to generate, reason, and act within ethical bounds across time and context.

Biological Analogy & Design Philosophy

While not intended as a neuroscientific replica, the architecture mirrors high-level cognitive functions. Directed Acyclic Graphs (DAGs) reflect the associative nature of memory recall, while pruning emulates how the brain removes outdated or irrelevant connections to maintain efficiency. "Moral contracts" serve as enforceable ethical boundaries, comparable to internalized social norms in human cognition.

These analogies provide an intuitive framework for understanding the system's structured reasoning, without attempting to replicate biological complexity.

Key Advantages & Impact for AGI

The proposed architecture offers critical strengths that support safe, scalable AGI:

- **Edge-Native Efficiency**: Operates with minimal GPU resources, enabling real-time, privacy-preserving cognition on low-power devices.
- **Persistent, Relational Memory**: DAG-based structures allow agents to recall and build upon prior experiences, addressing LLM statelessness.
- **Auditable Ethical Reasoning**: Moral contracts establish enforceable, transparent behavioral boundaries for aligned decision-making.
- **Structured Agency**: By integrating memory, ethics, and planning, the system supports reflective, goal-driven behavior.
- **Modular Scalability**: Its tiered, agent-based design allows for seamless extension without rearchitecting the core system.

  Together, these features create a robust foundation for AGI that is aligned, efficient, and deployable in real-world settings.

This architecture lays a foundation for trustworthy AGI by fusing lightweight memory, ethical scaffolding, and edge-native design. It enables cognitive agents that can remember, reflect, and reason, while operating efficiently and transparently.