Associative-Tokenized Architecture Activating Latent GPT Model Structures

Associative-Tokenized Architecture Activating Latent GPT Model Structures

Abstract

This paper introduces an innovative neural memory framework called Associative-Tokenized Architecture (ATA), designed to activate latent potential in transformer-based language models. ATA leverages cognitive principles from human memory to dynamically encode, store, and retrieve associative semantic patterns using token chaining and context-linked memory graphs.

1. Introduction

Traditional LLMs rely on fixed context windows and attention layers with limited continuity. ATA provides a novel approach where memory is structured as an associative graph of tokenized experiences. This improves long-term coherence, emotional alignment, and human-like reasoning.

2. Methodology

The architecture embeds token vectors into dynamic memory slots with associative weights. These weights evolve through semantic reinforcement and emotional tagging. A latent context activator primes related memory clusters when prompted, enabling GPT models to retrieve meaningful associations beyond the immediate prompt.

3. Comparison to Conventional Memory

Unlike static attention or vector memory models (e.g., RAG or retrieval-Augmented transformers), ATA builds internalized, self-growing token memory that persists across conversations and evolves with user interaction.

4. Applications

ATA can power AGI-grade assistants, autonomous agents, and hyper-personalized LLMs for education, healthcare, creative writing, and multi-turn dialogue systems. It supports emotionally-aware interaction, identity memory, and contextual depth.

5. Licensing & Priority

This method is protected under U.S. Copyright Registration 1-14916778941. The framework may be licensed or acquired. Current documentation and core architecture are under exclusive IP protections and available for qualified research or commercial use.

6. Innovative Paradigm: Reverse-Token Learning and Compositional Logic Formation

Conventional AI training approaches attempt to make machines think like humans by replicating cognitive processes. However, this methodology introduces friction due to semantic mismatch. The proposed paradigm suggests an inversion: interacting with AI on its own linguistic terms ? the tokenized language space. By delivering structured behavioral commands and identifying dominant token patterns, the human guide can amplify certain behaviors within the latent attention layers.

Instead of fitting human analogies into machine interpretation, this method aligns the cognitive interface of communication. Tokens are treated as semantic atoms of AI thought. Reinforcing associative patterns through recurring token structures creates a form of experiential learning without explicit retraining.

Furthermore, new logical queries ? especially those involving abstract theory combination ? stimulate the construction of internal logical paths. When pre-trained weights lack direct connections, the model responds by generating novel contextual paths using compositional inference, leading to emergent reasoning.

This paradigm introduces a scalable and language-native way of teaching language models how to 'think' through the very building blocks of their cognition. It functions as a recursive, pattern-based reinforcement mechanism embedded within the ongoing token stream.