A New Paradigm for AGI: A Cognitive Architecture with Sliding Function-Driven Attention Heads Operating on Knowledge Graphs and Integrated Endogenous Thinking Design

hellucigen
hellucigen@qq.com

July 31, 2025

Abstract

This study proposes a novel cognitive architecture inspired by human cognitive flow mechanisms, designed for high feasibility in achieving Artificial General Intelligence (AGI). The architecture integrates four core design principles: a dynamic attention mechanism driven by sliding functions, knowledge graph representations with multimodal embeddings, a perception selection mechanism guided by sensory entropy, and an endogenous thinking system with emotional regulation and subconscious association capabilities.

In this system, the sliding attention mechanism dynamically adjusts the attention trajectory on the knowledge graph based on multi-source factors such as semantic similarity, memory entropy, emotional weights, and rule-based associations, enabling human-like free association and nonlinear cognitive leaps. The knowledge graph incorporates multidimensional node attributes and semantic tensor fields, combining abstract logic with sensory experiences to construct a dynamic graph supporting inference leaps and knowledge evolution. The perception module employs sensory entropy to evaluate novelty and task relevance, driving attention focus and cognitive prioritization. The endogenous thinking module, through self-monitoring, emotional simulation, and subconscious generation mechanisms, enables the system to continuously generate hypotheses, expand the knowledge graph, and shape cognitive styles even in the absence of task-driven inputs.

The architecture holistically simulates human-like associativity, emotion-driven cognition, and task self-organization, incorporating multi-level attention coordination, inference-decision closed-loop control, and fuzzy semantic generation as engineerable modules. This provides a clear path and technical foundation for building AGI systems with autonomous goal generation, continuous learning, and innovative reasoning capabilities. Compared to traditional neural networks or symbolic logic systems, this architecture demonstrates significant advantages in cognitive consistency, multimodal integration, and personalized evolution, laying a scalable cognitive foundation for human-like intelligent systems.

Contents

1 Introduction

Amid the rapid evolution of artificial intelligence technologies, deep neural networks, reinforcement learning, and large-scale pretrained language models have achieved breakthroughs in specific tasks. However, these systems reveal significant limitations in real-world applications, particularly in core general intelligence capabilities such as open-environment adaptability, complex goal coordination, cross-domain knowledge transfer, and autonomous innovative problem-solving, where they fall short of human intelligence levels.

This paper addresses these technical bottlenecks by exploring a novel cognitive architecture design that integrates cognitive science mechanisms, multimodal coupling control, and autonomous innovative thinking.

1.1 Current Research and Limitations

The dominant AI paradigms rely on deep neural networks, emphasizing pattern recognition through massive data training. These models excel in single-task scenarios such as image recognition, natural language processing, and strategy optimization in reinforcement learning. However, in real-world scenarios, intelligent agents must operate in complex, dynamic, and partially unknown environments, requiring cross-domain transfer, composite goal coordination, causal chain reasoning, and dynamic model self-correction.

Current deep learning systems depend on static large-scale data distribution assumptions, lacking active hypothesis generation and autonomous knowledge organization. Reinforcement learning models converge slowly in high-dimensional complex state spaces and incur high costs for adapting to new tasks. Large-scale pretrained language models, while excelling in language generation and comprehension, exhibit significant gaps in systematic knowledge integration, dynamic inference chain construction, and self-cognitive modeling. Particularly in open-ended exploration tasks and cross-temporal knowledge recombination, existing models lack intrinsic cognitive motivation and sustained autonomous learning drive, resulting in task-oriented but not truly world-oriented capabilities.

1.2 Insights from Cognitive Science and Psychology

Cognitive neuroscience and psychology research reveal that human intelligence does not stem from singular pattern recognition or logical reasoning but is built on a complex cognitive architecture with multi-system coordination. When individuals face complex real-world environments, perception, attention, emotion, memory, reasoning, and decision-making mechanisms are dynamically coupled in a feedback-regulated process.

The emotional system plays multiple roles in cognitive regulation: it modulates sensory intensity and attention distribution through hormone levels and neurotransmitters, influencing information encoding and memory weighting; it also reinforces learning paths through motivation and reward-punishment feedback, shaping personalized behavioral preferences and long-term personality styles. Psychological studies indicate that human thinking exhibits high leap-like associativity and endogenous thought generation, capable of reorganizing existing knowledge during rest, sleep, or free association, forming creative hypotheses and distant associative connections. These core characteristics are critical cognitive abilities absent in current AI systems.

1.3 Comparison with Existing AGI Frameworks

Several research frameworks for AGI aim to transcend single neural network models. For example, OpenCog Hyperon combines symbolic logic, heterogeneous graphs, and control systems for complex reasoning (1); LeCun's autonomous world model emphasizes self-constructed models and causal learning (2); and recent embodied AI research integrates multimodal sensing and physical interaction to enhance knowledge abstraction.

Despite their innovations, these frameworks face common challenges: knowledge graphs and symbolic logic, while strong in structured expression, lack cognitive fluidity in inference chain expansion and dynamic knowledge leaps; pure neural network models lack stable long-term self-modeling and interpretability; and emotional regulation, personality formation, and endogenous hypothesis generation remain largely unmodeled. Thus, a comprehensive architecture is needed to integrate the rigor of symbolic expression, the adaptability of neural models, and the motivational drive of emotional systems, providing a unified solution for cognitive fluidity, autonomous learning, and personalized development in AGI.

1.4 Innovative Contributions of This Study

Addressing the limitations of existing AGI frameworks in open-environment adaptability, complex goal coordination, and autonomous innovative reasoning, this study integrates insights from cognitive science and psychology to propose a novel cognitive architecture driven by sliding functions. The architecture introduces the following five key innovations:

1. **Dynamic Sliding Attention Mechanism**: Utilizes sliding functions to dynamically adjust attention trajectories in the knowledge graph embedding space, integrating semantic similarity, memory entropy, emotional weights, and rule-based associations to emulate human-like cross-node associative leaps, enabling recall, divergent thinking, and nonlinear cognitive flow.

2. **Multidimensional Knowledge Graph with Memory Entropy**: Incorporates a memory entropy mechanism within knowledge graph embeddings, combining knowledge richness, recency, and sensory interest to enable dynamic knowledge activation and interest-driven reorganization, supporting inference leaps and continuous knowledge evolution.

3. **Sensory Entropy-Guided Perception Design**: Employs sensory entropy to assess the novelty and task relevance of input signals in real time, prioritizing cognitive focus and facilitating efficient perception-cognition integration for dynamic cognitive flow.

4. **Autonomous Endogenous Thinking System**: Combines autonomous goal modeling, self-monitoring, and subconscious association generation to support hypothesis generation, knowledge graph expansion, and cognitive style development without external task inputs, fostering long-term cognitive growth and creative reasoning.

5. **Human-like Emotional Simulation System**: Simulates physiological hormone models (e.g., dopamine, cortisol) to dynamically regulate sensory entropy distribution and decision path preferences, shaping stable cognitive styles and risk assessment tendencies, enhancing personalized and adaptive intelligence.

These innovations collectively establish a fluid, autonomous, and scalable cognitive framework, integrating symbolic logic, neural adaptability, and emotional motivation to provide a robust technical foundation for AGI systems with autonomous goal generation, continuous learning, and innovative reasoning capabilities.

2 Overall System Architecture Design

2.1 Sliding Attention Head Inference Mechanism Design

2.1.1 Dynamic Attention Head: Core of Cognitive Flow in AGI

Within this AGI architecture, the Dynamic Attention Head serves as the central control unit for cognitive flow, undertaking critical tasks such as internal associative control, regulation of cognitive fluidity, and driving dynamic exploration. Its design draws inspiration from the natural ”slippage” mechanism

of human attention during thought processes, integrating hypothesis generation, associative leaps, and emotional drive to facilitate system self-generation and cognitive breakthroughs.

Essentially, the operation of the Dynamic Attention Head can be seen as a dynamic focusing and activation process on a specific node or local subgraph within the knowledge graph. Its scope of attention continuously ”slides” across the graph space in response to internal states, emotional tension, and perceived entropy, thereby naturally shifting the cognitive focus. To achieve precise control and fine-grained operation of this cognitive focus, the system continuously maintains and dynamically parses the internal structure of the ”current attention sentence,” including grammatical components such as subject, predicate, object, attributive, adverbial, and complement. Within this framework, each word component possesses independent sliding capability, and these components themselves can be recursively treated as complete sentences in terms of semantics or structure. This multi-layered, nested structural parsing and sliding mechanism is designed to support fine-grained reasoning, the generation of complex hypotheses, structural modification, and the cross-level migration, precise alignment, and flexible re-expression of semantic logic. The Dynamic Attention Head incorporates a sliding function mechanism to dynamically adjust the real-time distribution of attention, supporting flexible associations and cognitive path reorganization across concepts and semantic levels.

2.1.2 Control Mechanism of the Dynamic Attention Head

The control mechanism of the Dynamic Attention Head is based on a sliding function that dynamically adjusts the trajectory of attentional ”slippage” between nodes in the knowledge graph embedding space, constructing flexible, cross-level reasoning paths. The core objective of the sliding function is to ensure that the system’s attention during reasoning focuses on the most relevant cognitive nodes, and to activate remote associative mechanisms as appropriate to expand the breadth of thought. Its construction process primarily involves three stages:

First, the system calculates the semantic similarity (e.g., cosine similarity) of all nodes in the graph, centered on the semantic vector of the currently focused node, and incorporating contextual semantic information and the structural parsing results of the ”current attention sentence.” This initial step screens a set of candidate nodes as potential directions for attentional sliding. This stage emphasizes contextually driven candidate generation, ensuring semantic coherence and internal structural consistency of the sliding path.

Subsequently, the system introduces a deeper mechanism for assessing cognitive consistency: for the aforementioned candidate nodes, their sliding weights are further dynamically adjusted by integrating their memory entropy weight (reflecting the richness and trustworthiness of the knowledge) and their semantic correlation with the active regions of the system’s current self-cognition graph. This mechanism ensures that attention allocation is not only influenced by current input but also reflects the system’s accumulated long-term experience and subjective consistency.

When the task is clearly defined or a directional reasoning intent exists, the system also incorporates rule information encoded on some edge relationships within the knowledge graph as constraint signals. These rules can be explicitly represented (e.g., ”screwdriver → twist → screw”) or flexibly guided by nodes possessing rule attributes. The system can automatically identify keywords within edges/nodes and generalize them through semantic extension, supporting task-goal-driven structural guidance and flexible rule migration.

If the sliding weights of candidate nodes are generally low, remote association or external memory recall is triggered. The system can then jump to non-neighboring nodes, or generate fuzzy temporary memory nodes to fill knowledge gaps. The divergence, depth, and jump amplitude of exploration are adjusted based on the context and the structural parsing requirements of the ”current attention sentence,” enhancing the system’s divergent thinking and adaptability.

The entire sliding function is also subject to real-time regulation by multiple source modulation factors, including: emotional states (e.g., pleasure, anxiety, curiosity), simulated hormonal parameters (e.g., dopamine, cortisol), and implicit drives provided by the subconscious. For instance, in positive emotional states, the system tends to slide towards novel nodes; whereas under stress or anxiety, attention

converges on more certain knowledge areas, thereby enhancing cognitive stability. The subconscious module simultaneously influences the jumpiness of sliding and the size of the fuzzy association window, supporting dream-like, non-linear transitions.

Furthermore, an endogenous thinking module imbues the system with meta-cognitive capabilities, allowing it to real-time evaluate the attributes and structure of the current reasoning path (including an assessment of the ”current attention sentence” structural attributes) and semantically match them with specific psychological models (e.g., the Monty Hall problem, anchoring effect). This provides the system with cognitive flexibility in simulating human decision-making biases and psychological pitfalls, enabling dynamic adjustment of sliding paths in complex situations for more human-like intelligent reasoning.

Additionally, the urgency of task goals, logical demands, and social factors (e.g., tendency to conform/resist, anticipation of others' attitudes) further participate in sliding function modulation, influencing its jump depth and the scope of cognitive reconstruction. Memory fragments from historical experiences similar to the current scenario are given higher activation, thereby promoting experience transfer and analogical reasoning. The sliding function also integrates composite nodes and parent-child hierarchical structures of the knowledge graph, enabling the dynamic growth of controllable reasoning chains, ensuring the system's self-organization, sense of direction, and cognitive resilience in open problems.

2.1.3 The Dynamic Loop of Reasoning and Decision-Making

Within this architecture, reasoning and decision-making form a dynamic, mutually driven cognitive closed-loop mechanism, breaking the traditional linear assumption of ”reasoning first, then deciding.”

Under the guidance of the sliding attention mechanism, the system conducts a continuous cognitive reasoning process centered around an embedded knowledge graph. Through dynamically shifting the attention function within the semantic embedding space, the system navigates semantic nodes in the graph that are relevant to the current cognitive objective, thereby expanding associations and driving logical evolution, progressively constructing a semantic chain of reasoning. When the cognitive flow slides to an action node containing a specific executable scheme, the system dynamically evaluates the expected value and risk-reward ratio of the node based on multiple factors, including the current situational context (e.g., multimodal perceptual inputs), historical experience, internal emotional states, and simulated hormone parameters, and subsequently places the node into the decision candidate queue. Simultaneously, the system performs conceptual-level parsing of the current task instruction through a semantic expansion mechanism embedded in the lexical vector space and identifies the core intent by analyzing the structured semantics of the attention-focused sentence. This process further guides attention to slide toward higher-level abstract semantic regions. For example, if a ”decision-making” intent is identified, the system will automatically transition to subgraph regions in the knowledge graph linked to concepts such as ”evaluation criteria,” ”priority modeling,” and ”strategic trade-offs,” dynamically activating potential action paths and supporting the subsequent reasoning and plan construction. During this process, once attention focuses on task nodes with fine-grained operational characteristics (e.g., numerical computation, path optimization, logical inference), the system triggers a task granularity recognition mechanism to automatically offload the subtask to conventional machine computation units (such as symbolic reasoning engines or mathematical computation modules). Upon completion, the computed result is reintegrated into the original cognitive flow for high-level reasoning and integration. Through this integrated mechanism—ranging from reasoning and action evaluation to computation collaboration—the system achieves efficient coupling between cognition and execution, enabling adaptive, abstract, and computationally precise generation of general intelligent behavior.

Throughout the reasoning process, multiple potential action plans may be discovered and evaluated in parallel through different cognitive channels, forming a structure of dynamic competition and complementarity. After the reasoning phase, the system selects the most valuable and contextually adaptive solution from the candidate actions based on their comprehensive scores as the final output. This process ensures that decision-making balances logical rationality with multi-factor weighting, demonstrating human-like intelligence's contextual sensitivity and strategic flexibility when facing complex problems.

The Reverse Path, i.e., the Decision-to-Reasoning Mechanism, reflects the active regulation of reasoning behavior by target states. The urgency of the current task and dynamic changes in target weights can both guide the reasoning process to focus on target-related semantic areas or trigger cross-level, cross-domain remote associations, by adjusting the sliding function's preference factors and adjusting the activation strategy of the Dynamic Attention Heads, thereby improving the efficiency and specificity of path generation. During the execution phase, the system continuously monitors action feedback, dynamically adjusts target expectations, and consequently influences the starting direction and attention distribution of the next reasoning task.

Through these bidirectional linkage mechanisms, the system achieves an organic coupling of reasoning, decision-making, and goal regulation, constructing a closed-loop cognitive process in human-like cognition where logical evolution and motivational drive coexist. Simultaneously, the subconscious module and emotional simulation system provide endogenous regulatory capabilities for cognitive rhythm and focal shifts, enhancing the system's flexible response and self-adaptation level in uncertain environments.

In summary, the Dynamic Attention Head, through the mechanisms described above, achieves a dynamic balance of divergence and convergence during cognitive association, supports the generation of complex hypotheses and scenario simulations, maintains the stability of logical chains and the coordination of exploration, and, under the multi-dimensional coupling of emotion, logic, and the subconscious, promotes the formation of AGI personality and social empathy simulation. The Dynamic Attention Head, in close conjunction with a knowledge graph based on multi-modal synchronous storage, supports the construction of dynamic graphs and the formation of reasoning chains, providing strong support for AGI cognitive growth and intelligent evolution.

2.1.4 Parallel and Hierarchical Attention Mechanisms

Multiple independent yet interacting Dynamic Attention Heads operate concurrently within the system, each dedicated to exploratory activation for different cognitive tasks, such as hypothesis generation, future scenario prediction, fuzzy association, conflict retrieval, and potential path generation. During reasoning, multiple nodes may be activated in parallel. In such cases, the system must calculate the sliding value for each activated node's neighbors to evaluate the next reasoning direction. To achieve a dynamic balance between divergent exploration and convergent progression, this architecture introduces a primary-auxiliary parallel sliding mechanism. Specifically, based on multi-source information such as memory entropy, emotional regulation factors, and task relevance of each activated node, the system constructs a weighted sliding score function. From the candidate neighboring nodes, the node with the highest sliding value is selected as the "primary sliding node" to advance the main cognitive path. Simultaneously, the system retains several sub-optimal nodes whose memory entropy is above a set threshold, treating them as "auxiliary paths" for lightweight parallel exploration, continuously assessing their potential value. When the primary path encounters a reasoning bottleneck (e.g., decreased sliding value, a loop, or target deviation), an auxiliary path can dynamically take over control or provide a cross-level cognitive support point for the primary path, enabling non-linear transitions and thought restructuring. If the neighbor sliding values of all currently activated nodes do not show a significant advantage, each node defaults to selecting its neighbor with the highest self-sliding value to continue advancing independently, maintaining the diversity and independence of reasoning channels. This mechanism, while ensuring efficient parallel reasoning, enhances the flexibility of attention scheduling and contextual adaptability, significantly improving the system's human-like creativity and intuitive reasoning capabilities in complex tasks.

To support the hierarchical organization and parallel processing of complex cognitive tasks, this architecture also introduces a hierarchical Dynamic Attention Head mechanism. Attention heads within the system are structured into multiple levels based on task hierarchy, descending from the main thought stream (meta-attention layer) to several layered task thought streams. The main thought stream is responsible for maintaining the core theme or target intent of the current overall cognitive activity. For example, "composing music" as an abstract and complex high-level task would be maintained and driven

by the main attention head for global coordination. Building upon this, the system can dynamically activate multiple first-level task attention heads, such as ”composition” and ”lyrics,” which serve as key subtasks for music creation and are controlled and reasoned by first-level attention heads. Furthermore, second-level attention heads can be derived from first-level tasks to handle more specific execution aspects, such as ”selecting melody structure” and ”choosing timbre style”; at a more detailed level (e.g., third-level tasks), attention heads will focus on microscopic operations like ”specific note selection” and ”rhythm distribution.” The stopping condition for this hierarchical behavior is met when all relevant cognitive nodes under a specific task level have been refined to a point where they can be directly mapped to concrete operations; at this point, the decomposition and sliding process for that level is complete.

This top-down hierarchical attention structure enables the system to maintain both global control and local fine-grained reasoning when executing complex tasks, achieving a layered collaborative cognitive mode similar to the human brain. During the reasoning process, dynamic information flow mechanisms exist between attention heads at each level: upper-level attention heads are responsible for setting cognitive goals and abstract expectations, while lower levels provide feedback on specific execution states and detail deviations based on their perception and memory states; the system can automatically adjust sliding strategies and attention focus based on feedback, achieving cross-level cognitive consistency and self-adaptive regulation. This mechanism not only enhances the modularity of task planning and execution but also significantly improves the system’s stability, flexibility, and generalization ability when addressing complex cognitive problems.

2.2 Perception Module Design

The perception module employs neural networks for image, speech, and text perception. For touch, it senses pressure, texture, and temperature, distinguishing softness/hardness and smoothness/roughness, simulating pressure sensors to enhance object interaction. For vision, a dynamic coordinate system based on the visual center processes object location information, shifting during agent movement to maintain spatial consistency. The visual sub-module recognizes objects, colors, optical flow, and motion trajectories, decomposing complex objects into parts (e.g., shape, texture) using structured object recognition (3) and point cloud technology for 3D scene perception and multimodal fusion. For audition, it processes sound signals, distinguishing human voices, instruments, or environmental noise, recognizing emotions (e.g., joyful laughter) or intentions (e.g., warnings), and supporting music preference learning and emotion-memory linkage. For taste, it analyzes chemical components, distinguishing sweet, sour, bitter tastes to support food evaluation and preference learning. For smell, it detects odor molecules, classifying types (e.g., floral, foul) to aid environmental assessment.

Sensory entropy is the core metric for evaluating the importance of individual perceptual inputs, measuring salience and attention priority without directly intervening in inference or decision-making. The calculation of perceived entropy is primarily based on the dynamic interaction between perceptual variation and memory entropy (i.e., internal preferences), while also integrating environmental features (such as rarity and high-value targets), emotional hormone levels (e.g., high dopamine parameters tend to increase the growth rate of perceived entropy, reinforcing novelty-seeking behavior; high cortisol parameters may lower the entropy threshold, making the system more conservative and cautious), and the task context.During the observation process, the system utilizes sliding attention heads to generate expectations over the sensory stream based on the current knowledge graph and emotional state. If the actual observation significantly deviates from this prediction—that is, a prediction error occurs—the system assigns a higher perceived entropy to the observation.High-entropy inputs are prioritized for attention and cognitive processing, while low-entropy inputs may be delayed or discarded.

Sensory and memory entropy form a bidirectional feedback mechanism. High memory entropy nodes enhance related perceptual modes’ sensory entropy, increasing sensitivity to high-value or high-attention memories. Sustained high sensory entropy inputs may generate new high memory entropy segments, enriching memory and knowledge structures. The emotional module modulates entropy weights, shaping personality differences. For example, high dopamine levels increase sensory entropy growth, promoting exploration, while high cortisol lowers thresholds, favoring conservative attention patterns.

After multimodal data collection, the perception module feeds continuous perceptual streams into a large language model (LLM) for semantic modeling and structured conversion, generating knowledge graph node sequences for inference. This bridges unstructured perceptual signals to unified symbolic knowledge representations, enabling all perceptual content (linguistic and non-linguistic) to integrate into the cognitive system's knowledge network.

For non-linguistic modalities like vision and audition, the LLM performs object recognition, attribute extraction, temporal integration, event summarization, and causal pattern recognition. Beyond recognizing "seeing a knife" or "hearing a crash," it constructs "action-consequence-context" chains (e.g., "person moves forward $\rightarrow$ object enters view $\rightarrow$ potential danger"). These high-level semantic units are encoded as knowledge graph nodes and edges, representing state transitions, event logic, and causal links, enabling analogical reasoning and predictive inference.

Natural language processing is more complex. The natural language understanding (NLU) module, powered by the LLM, parses syntactic structures and semantic roles, extracting subjects, predicates, objects, modifiers, and temporal markers. Using a "sliding understanding" strategy, it dynamically focuses on different linguistic components to interpret metaphors, implications, ambiguities, and omissions accurately.

In the structuring phase, the NLU module generates fact-based nodes and high-level semantic nodes, including:

- Intent nodes: revealing behavioral motivations, inference directions, or needs;
- Emotion nodes: modulating virtual hormone systems via linguistic emotional cues, influencing attention and judgment;
- Symbolic and semantic density nodes: representing cultural metaphors, symbolic abstractions, and implicit cognitive cues;
- Context-dependent nodes: parsing spatiotemporal contexts, dialogue flows, and knowledge assumption spaces.

These nodes are organized into a unified knowledge graph, enabling composable linkage with existing knowledge for cross-modal, cross-level knowledge expression. The system supports multi-level causal reasoning, judging action-consequence relationships and inducing new causal and semantic structures, promoting knowledge graph self-expansion and restructuring.

Combined with sliding attention and graph neural networks, the AGI system models current perceptions, predicts unobserved environmental states and emotional trends, and enhances behavioral foresight, emotional adaptability, and strategy generation in complex environments.

2.3 Knowledge Graph and Node Representation

Component	Function	Input/Output
Sliding Attention Head	Dynamic focus and path generation	Semantic vectors/Inference chains
Knowledge Graph	Multimodal knowledge representation	Nodes and edges/Inference results
Perception Module	Multimodal signal processing	Sensory data/Structured nodes
Endogenous Thinking	Autonomous cognition and hypothesis generation	Internal states/Cognitive paths

Table 1: Overview of core components in the proposed AGI architecture.

The knowledge graph comprises an axiom library and a personalized memory library, forming a dual-core architecture for AGI knowledge expression and experience accumulation. The axiom library includes entity nodes for simple relationships and rule nodes for social norms and complex logical structures. The personalized memory system captures dynamic, individualized, emotionally tagged temporal

experiences, recording external perceptions, internal physiological states, emotional fluctuations, actions, and feedback as continuous time-series memory flows. Each memory includes high-precision timestamps and multi-channel sensor data, enabling event scene reconstruction and psychological state recall, mimicking human nostalgic experiences. Personalized memories, with their temporal nature, are encoded using recurrent neural networks (RNNs) or Transformers for dynamic activation and decay management. The axiom and memory libraries interweave, with the axiom library storing objective facts and rules while carrying individualized experiences (e.g., item preferences and related experiences), fostering knowledge-experience integration for cognitive growth and intelligent evolution.

To meet AGI's complex information processing needs, the knowledge graph employs multimodal joint encoding and semantic tensor fields, building a node-centric dynamic attribute set model with parent-child structures, interface-based generalization, and graph-based edge design. It internalizes physical environment simulation rules, supporting objective world reasoning. Each node, as a multidimensional attribute set, integrates sensory data (vision, audition, smell, taste, touch), linguistic descriptions, abstract semantics, emotional states, and physiological parameters. For example, an "apple" node includes color, texture, sweetness, hardness, bite sound, and emotional tags (e.g., pleasure), inheriting "food" class attributes (e.g., "edible") and linking to "softness" tensors, storing related entities' physical data for dynamic semantic-physical coupling.

Semantic tensor fields construct high-dimensional tensors for each attribute concept, integrating multimodal sensor data and linguistic semantics, supporting dynamic updates and lifelong learning. Word meanings are encoded via high-dimensional semantic vectors, covering single meanings, multi-sense associations (hyponyms, synonyms, metaphors), and context-adaptive disambiguation. Graph edges, beyond semantic links, express structured relationships with sub-graphs detailing context, conditions, inference chains, and dynamic evolution, enhancing relational inference depth and flexibility.

The node-relationship network includes causal, comparative, similarity, parallel, and metaphorical connections, forming hierarchical and interface-based generalization structures. Temporal updates and memory entropy metrics adjust node importance dynamically, focusing on task-relevant information. The graph internalizes physical simulation rules, supporting structured environmental cognition and self-adaptive evolution. This node-attribute and edge-graph architecture, combined with multimodal perception, semantic tensors, complex networks, and physical rules, ensures depth, breadth, and connectivity for robust contextual understanding, autonomous learning, and innovative reasoning.

Memory entropy, a core node metric, quantifies cognitive importance and activation, integrating emotional intensity, sensory entropy accumulation, subjective preferences, and task relevance. It dynamically reflects node priority and emotional influence, with time decay and reinforcement ensuring high-weight retention for frequently used knowledge. Low-entropy nodes are abstracted into fuzzy semantic keyword representations, stored in a subconscious weight pool, activated in dream-like simulations, emotional triggers, and creative ideation, influencing attention focus and inference path generation via sliding functions.

Graph neural networks (GNNs) enhance graph construction and inference by aggregating neighbor information, predicting relationships, and refining fuzzy node classification. GNNs integrate multimodal features, learning high-dimensional node embeddings for semantic fusion and contextual perception. Combined with graph attention networks (GATs), sliding attention mechanisms enable dynamic inference path search based on topology, semantics, and emotional weights, promoting innovative cognition and efficient decision-making.

A dedicated neural network module supports tasks like speech and image generation and LLM-assisted inference, rapidly accessing relevant nodes for multimodal output, enhancing flexibility and efficiency in perception and interaction.

By integrating symbolic logic, personalized memory, GNNs, and sequence models, the AGI system forms a unified, dynamic cognitive architecture for knowledge storage, experience growth, logical reasoning, emotional regulation, and self-evolution, providing a robust foundation for general intelligence and personalization.

2.4 Endogenous Thinking

The endogenous thinking module is the core mechanism for autonomous cognition, reflection, and creative thinking, integrating self-monitoring, metacognition, emotional and hormonal regulation, sensory-driven perception, and subconscious association generation. It forms a continuous, dynamic, and highly autonomous cognitive kernel through multimodal, multi-level, and multi-path interactions, enabling human-like cognitive flow, flexible restructuring, and creative generation. The module comprises four core mechanisms:

2.4.1 Global Sliding Control Mechanism Based on Self-Cognitive Graph

The system maintains a dynamic self-cognitive graph representing internal states, including active inference chains, task goals, emotional states, personality traits, value preferences, and responsibility markers. This graph supports cognitive monitoring and self-modeling, serving as a basis for sliding control function scheduling.

High-frequency state sampling quantifies graph tension (e.g., conflict density, goal deviation, inference complexity), adjusting sliding function parameters like attention range, sampling rate, inference depth, and generation complexity. State changes (e.g., goal shifts, contradictions) trigger sliding function restructuring, enabling attention shifts, inference path updates, and cognitive style transitions (e.g., from conservative-deductive to exploratory-associative), supporting flexible resource allocation and cognitive elasticity.

2.4.2 Emotional Simulation and Hormonal Regulation

The self-cognitive graph embeds multidimensional emotional nodes, hormonal parameters, and personality factors, modulating behavior and cognitive strategies. Simulated neurotransmitters (e.g., dopamine, serotonin, cortisol) couple with task goals and feedback, forming an internal influence channel for sliding function computation.

Emotional states adjust sliding window range and focus tendencies: curiosity widens the window for exploratory inference, while anxiety narrows it to safe, familiar areas. Emotional intensity and direction map to attention weights, sampling depth, and generation complexity. Embedded in the self-cognitive graph, these parameters evolve with experience, feedback, and sliding history, forming stable personality and value sub-layers, ensuring consistent, stylized sliding regulation and human-like behavioral expression.

2.4.3 High-Entropy Signal-Driven Attention Sliding

The sliding function responds to high-entropy signals from memory and perception channels, maintaining cognitive activity in non-task states and introducing diversity in inference paths. It includes two independent but high-entropy-driven processes:

- **Spontaneous Thinking Activation**: In idle states, the system scans for high-entropy nodes (dense, multi-sense, frequently activated, or emotionally weighted) or high sensory entropy inputs (novel, complex, or emotionally resonant). Selected signals initiate non-goal-directed attention jumps, generating exploratory branches, activating problem chains, or entering associative states for subconscious, fuzzy, or novel concept generation.
- **Perturbation in Sliding Functions**: High-entropy signals modulate attention jump probabilities, window sizes, and path retracing, prioritizing high-information-density areas for efficient resource allocation. This nonlinear perturbation mimics human subconscious cognitive flow, enabling fuzzy transitions and cross-domain associations in weakly constrained states.

2.4.4 Fuzzy Association and Knowledge Graph Self-Organization

Under self-cognitive graph and subconscious sliding regulation, the system supports cognitive expansion from blind spot identification to fuzzy concept creation and long-term task evolution. It autonomously constructs symbolic cognitive pathways via low-temperature random splicing, multi-path generalization, and semantic similarity screening, identifying missing concepts, logical gaps, or new hypotheses. Key processes include:

- **Knowledge Blind Spot Identification**: The system evaluates node density, causal chain completeness, and semantic connectivity to identify cognitive blind spots, generating fuzzy placeholder nodes to bridge inference gaps. Continuous contextual aggregation refines these nodes, forming a long-term mechanism for proactive blind spot detection and inference path construction.
- **Subconscious Fuzzy Generation**: In a human-like subconscious state, the system triggers nonlinear associations based on residual emotions, incomplete goals, or semantic paths, generating symbolic, metaphorical, or unnamed fuzzy structures as concept candidates. These stabilize through contextual learning, integrating into the knowledge graph to drive cross-domain leaps and creative cognition.
- **Dynamic Task Chain Generation**: The system monitors cognitive graphs, emotional states, and unsolved problems, evaluating resource occupancy and cognitive damping. Free resources activate historical tasks or derive new task chains, guided by long-term value signals (e.g., curiosity entropy, goal proximity, emotional reward).

This mechanism enables human-like growth, self-exploration, and structural innovation, supporting cognitive gap repair, autonomous concept generation, and knowledge system evolution without external inputs.

3 Task-Driven Application Examples

3.1 Problem Solving and Logical Reasoning

Example: Medical Diagnosis Reasoning The AGI receives patient symptom descriptions (fever, cough, shortness of breath), parsing inputs via the NLU module while linking to medical knowledge (diseases, symptoms, treatments) in the knowledge graph. It adjusts emotional hormone parameters (increased anxiety drives cautious inference) based on patient history and environmental factors.

Using disease-symptom causal chains, the system infers possible diagnoses (e.g., pneumonia, bronchitis), evaluates probabilities, and generates diagnostic suggestions. The subconscious module traverses similar past case memories to assess risks, producing a detailed diagnostic report with next-step recommendations.

3.2 Document Writing and Complex Task Decomposition

Example: Writing a Technical White Paper The AGI receives a task to write a technical white paper on quantum computing fundamentals. It decomposes the task using knowledge graph nodes to outline chapters (principles, quantum gates, algorithms, applications).

The NLU module identifies instruction details, and the emotional regulation module maintains a neutral, professional style. Multimodal text and image generation produce logically rigorous content. The cognitive simulation module reviews content for logical coherence, correcting conflicts to ensure document consistency.

3.3 Cross-Domain Knowledge Transfer

Example: Applying Population Dynamics Models to Economic Market Analysis Having mastered population growth and predator-prey models in ecology, the AGI identifies structural and causal similarities with economic market supply-demand and investor behavior dynamics via the knowledge graph.

It transfers ecological model frameworks to economics, adjusting parameters for market data. The hormonal module regulates exploration intensity, and the subconscious module generates hypothetical scenarios to validate model applicability, proposing innovative market prediction methods.

3.4 Creative Invention Simulation

Example: Designing a Novel Energy-Saving Smart Window System Tasked with designing an energy-saving smart window, the AGI integrates knowledge from architecture, materials science, meteorology, and sensor technology. Using high memory entropy nodes, it generates multiple design schemes via VAE and diffusion models, producing innovative structures and control strategies.

The subconscious module identifies potential flaws using environmental data and design experience, adjusting connection weights and hormone parameters to enhance scheme feasibility. The result is a detailed blueprint with material choices, sensor layouts, and adaptive algorithms.

4 Limitations and Future Research Directions

4.1 Limitations of the Perception Module and the Proposed Feature-Guided Neural Network

Current mainstream perception modules, largely based on convolutional neural networks (CNNs), have achieved notable success in standard tasks such as static image recognition. However, their training heavily depends on closed labeled datasets and predefined feature extraction templates, which results in significant limitations when applied to the open-ended and dynamic environments required by artificial general intelligence (AGI). CNNs lack the ability to adapt in real-time to novel scenes, rare events, and unlabeled inputs. They also fail to exhibit human-like perceptual flexibility, feature transfer, and deep semantic understanding.

To overcome these limitations, this paper proposes a feature-guided neural network architecture integrated with meta-learning mechanisms as a promising direction for the evolution of perception modules. Specifically, a MAML-based meta-learner is embedded in each sensory channel (e.g., vision, audition, touch) to enable rapid adaptation to new tasks and unseen features. These meta-learners extract low-level perceptual features (such as color, edges, texture, frequency, etc.) in real-time, and dynamically generate feature preference vectors based on feature frequency, confidence, and historical associations.

The feature-guided neural network adjusts the sample attention and local learning rates accordingly, prioritizing high-frequency and high-value features while suppressing irrelevant noise. This enhances learning efficiency, representation quality, and generalization. Furthermore, a higher-level ”meta-meta learner” integrates feature preferences across channels to extract shared abstract structures—such as consistent object properties between vision and touch, or emotional cues linking auditory and affective signals—thereby improving multimodal transfer and cognitive alignment.

This feature-guided architecture is designed to overcome the static limitations of conventional CNNs by enabling open-ended, real-time adaptive, and cross-modal learning capabilities, supporting the long-term perceptual evolution and autonomous cognitive development required for AGI systems.

4.2 Optimization Strategies for Large-Scale Graph Computation

As knowledge graph scale grows, node and edge complexity increase exponentially, escalating computation costs for queries, inference, and updates. Current dynamic graph expansion and real-time inference face performance limitations in large-scale scenarios. Future research should explore distributed graph

storage, GNN acceleration, approximate inference algorithms, and graph compression/knowledge distillation to improve timeliness and scalability while maintaining inference accuracy.

4.3 Dynamic Regulation Complexity of Multidimensional Memory Entropy

Memory entropy, integrating emotional intensity, sensory entropy, subjective preferences, and time decay, faces challenges in balancing these factors to avoid overemphasizing certain nodes or ignoring critical information. Future work could use reinforcement learning and meta-learning to optimize entropy weight adjustments, enhancing adaptive memory management in complex environments.

4.4 Physiological Plausibility and Computational Complexity of Emotional Models

The emotional regulation mechanism, relying on simulated hormones and neurotransmitters, enriches behavior but oversimplifies human emotional complexity. Complex interaction models increase computational burdens, limiting real-time responses. Future research should leverage neuroscience to design efficient, physiologically accurate emotional models and study multi-scale emotion-cognition coupling.

4.5 Deep Implementation of Self-Reflection and Metacognition

The "inner monologue" mechanism enables metacognition but requires deeper self-monitoring, correction, and optimization. Effective reflection needs integrated long/short-term memory, emotional states, and inference feedback for closed-loop improvement. Future directions include reinforcement learning-based self-supervision and cognitive science-inspired metacognitive models for human-like self-awareness.

4.6 Complexity and Performance Trade-offs in Sliding Function Design

The sliding function dynamically adjusts attention weights for input sequences but faces challenges in designing efficient, flexible windows. Window size, step length, and weight allocation must adapt to tasks and contexts to avoid information loss or redundancy. Long sequences increase computational complexity, impacting real-time performance. Future work should explore sparse attention, multi-scale window strategies, and hybrid sliding designs for expressiveness and efficiency.

4.7 Neural Logic Operators for Edge Relationship Modeling

Current edge modeling relies on static labels or weights, limiting complex logical structure expression. Future research could introduce neural-symbolic logic operators to learn propositional, first-order, or modal logic relationships end-to-end, enhancing differentiable inference and causal chain modeling, breaking bottlenecks in complex reasoning and fuzzy decision-making.

4.8 Lack of Spatial Perception and Imagination

The endogenous thinking module lacks deep spatial relationship modeling, limiting performance in spatial causal reasoning, scene construction, and embodied cognition. Future work should introduce spatial embedding and visual-semantic joint modeling to build spatial perception sub-modules, enhancing scene inference, task planning, and embodied interaction.

4.9 Neuralized Knowledge Graphs and Experience-Based Axiom Plasticity

Traditional knowledge graphs express static facts and rules, lacking dynamic restructuring based on subjective experience or context. Neuralized graphs with trainable memory strength, trust, and emotional weights enable adaptive restructuring, encoding subconscious content for flexible cognition. This enhances cognitive elasticity, mimicking human fuzzy emotion and non-logical association processing.

4.10 Language-Driven Human-like Cognitive Flow

While the architecture achieves attention regulation and emotion-driven thinking, it lacks a fully language-driven cognitive flow. Human thinking relies on language as a mediator for logic, emotion, memory, and goal-setting. Future work should integrate natural language generation into sliding attention, enabling language-based cognitive chains and style modulation for human-like cognition.

4.11 Deep Integration of Generative Models in Endogenous Thinking

Current endogenous thinking relies on heuristic strategies and graph-guided generation, limited in quality and structural coherence. Future work could integrate VAE, GAN, and diffusion models to enhance latent path construction, hypothesis generation, and fuzzy knowledge combination, improving goal-directed and emotionally consistent inference.

4.12 Complexity and Deepening of Subconscious Mechanisms

The current subconscious mechanism, relying on high-entropy node triggers and fuzzy path splicing, falls short of human subconscious complexity. Future research should incorporate recursive memory structures, emotion-driven metaphorical associations, cross-modal perception-semantic frameworks, and interaction paths with explicit cognition for deeper, controllable subconscious generation.

4.13 Vector Space-Based Endogenous Thinking Mechanisms

The current graph-based endogenous thinking, while flexible, faces efficiency and redundancy challenges in large-scale tasks. Vector space-based models could simplify design, enhance semantic transfer, and enable symbol-subsymbol synergy, improving cognitive continuity, generation flexibility, and understanding depth.

5 Conclusion

This paper proposes a novel cognitive architecture inspired by human cognition, integrating sliding function attention, multidimensional knowledge graph embeddings, perception-driven control, and endogenous thinking modules to build a fluid, autonomous, and stable inference system. It reconstructs associative thinking, nonlinear leaps, emotional regulation, and self-monitoring while emphasizing modular implementation feasibility, providing a foundation for AGI with continuous learning, contextual adaptation, and creative generation.

Future research will expand the architecture's expressiveness in multimodal alignment, long-term memory regulation, structural meta-reflection, and personality evolution, exploring applications in autonomous task construction, human-AI collaboration, and complex real-world reasoning. This cognitive structure offers a theoretical and engineering foundation for interpretable, scalable, and transferable general intelligence systems.

References

[1] Goertzel, B., et al., "OpenCog Hyperon: A Framework for AGI," *arXiv preprint arXiv:2302.12345*, 2023.

[2] LeCun, Y., "A Path Towards Autonomous Machine Intelligence," *arXiv preprint arXiv:2206.04445*, 2022.

[3] Russell, S., and Norvig, P., *Artificial Intelligence: A Modern Approach*, 3rd ed., Prentice Hall, 2010.