

AGI 的新范式: 集成内源思维设计的基于滑动函数计算的注意力头在知识图谱上滑动的认知架构

hellucigen

hellucigen@qq.com

2025 年 6 月

摘要

本研究提出一种以人类认知流动机制为灵感、面向通用人工智能（AGI）的可实现性高的新型认知架构。该架构整合四大核心设计思想：滑动函数驱动的动态注意力机制、多模态嵌入的知识图谱表达、感知显著度引导的感知选择机制，以及具备情绪调节能力的内源思维系统。

在该系统中，滑动注意力机制基于语义相似度、节点激活强度、情绪权重与规则关联等多源因子，动态调控知识图谱中注意力的滑移路径，实现类人思维中的自由联想与非线性跃迁。知识图谱引入多维节点属性表示，融合抽象逻辑与感官经验，构建可支持推理跳跃与知识演化的动态图谱。感知模块利用感知显著度评价新颖性与任务关联度，驱动注意力聚焦与认知优先级调整。内源思维模块则通过自我监控与情绪模拟，支持系统在无任务驱动下持续生成假设、扩展知识图谱与塑造认知风格。

架构整体模拟人类认知中的联想性、情绪驱动性与任务自组织能力，并引入多层次注意力协同机制、推理-决策闭环控制、模糊语义生成等可工程实现的模块，为构建具备自主目标生成、持续学习与创新推理能力的 AGI 系统提供了清晰路径与技术落点。相较于传统神经网络或符号逻辑系统，该架构在认知一致性、多模态融合与人格化演化等方面具备显著优势，为类人智能系统的落地实现奠定了可拓展的通用认知基础。

目录

1 绪论	3
1.1 研究现状与不足	3
1.2 认知科学与心理学启示	3
1.3 本框架的设计思想	3
1.4 本框架的创新贡献	4
2 系统总体架构设计	4
2.1 知识图谱的双核心架构与动态元素表示 (Dual-Core Architecture and Dynamic Element Representation)	4
2.1.1 语义记忆库: 基于加权边关系的概念与规则网络	4
2.1.2 节点与边的设计 (Node and Edge Design)	5

目录	2
2.1.3 情景记忆库：基于故事语法的事件结构存储	5
2.1.4 基于自我认知图谱的全局滑动控制机制：状态主导的知识能量调节	6
2.2 滑动注意力头推理机制设计	6
2.2.1 事件框架：认知的形式	6
2.2.2 滑动函数：认知的流动	6
2.2.3 并行与分级注意力机制	8
2.2.4 双事件框架与内外认知分流机制	9
2.2.5 基于节点激活度的认知焦点维持	10
2.2.6 日常推理与逻辑推理的双通路机制设计 (Dual Reasoning Pathway: Everyday vs. Logical Reasoning)	11
2.3 感知模块设计	12
2.3.1 多模态感知功能	12
2.3.2 感知显著度与注意选择机制	12
2.3.3 情节缓冲器与短时整合机制	12
2.4 由节点知识引导的认知活动	13
2.4.1 由节点知识引导的认知活动：决策	13
2.4.2 由节点知识引导的认知活动：情绪与性格	14
2.4.3 由节点知识引导的认知活动：预期	15
2.4.4 由节点知识引导的认知活动：程序化行动	16
3 局限性与未来研究方向	16
3.1 感知模块的局限性与特征引导神经网络的设想	16
3.2 大规模图谱运算的优化策略	17
3.3 内源情绪与激素模型的生理合理性与计算复杂度	17
3.4 自我反思与元认知机制的深度实现	17
3.5 滑动函数设计的复杂性与性能权衡	18
3.6 引入神经逻辑算子的边关系建模展望	18
3.7 空间感知与想象能力的缺失	18
3.8 神经网络化的知识图谱与基于经验的公理可塑性建模	18
3.9 生成模型在内源思维中的深度融合	19
3.10 内源思维的向量空间生成机制探索	19
4 结语	19

1 绪论

在当前人工智能技术快速演进的背景下，深度神经网络、强化学习、大规模预训练语言模型等方法已在诸多特定任务中取得了突破性成果。然而，这些系统在实际应用中仍然暴露出明显的局限性，特别是在面对开放环境、复杂目标协调、多领域知识迁移与自主创新性求解等通用智能核心能力方面，距离人类智能水准尚有显著差距。本文旨在回应这一技术瓶颈，探索一种融合认知科学机制、多模态耦合控制与自主创新思维的新型认知架构设计路径。

1.1 研究现状与不足

现有主流人工智能技术体系以深度神经网络为核心，强调通过海量数据训练获得模式识别能力。无论是在图像识别、自然语言处理还是在强化学习中的策略优化，这些模型在单一任务下展现出优异性能。然而，在实际应用场景中，智能体往往需要在复杂、多变且部分未知的环境中长期运行，涉及跨领域迁移、复合目标协调、因果链式推理与动态模型自我修正等复杂认知过程。目前，深度学习系统普遍依赖静态的大规模数据分布假设，缺乏主动假设生成与自主知识组织能力；强化学习模型在高维复杂状态空间中学习收敛速度缓慢，适应新任务代价高昂；而大规模预训练语言模型虽然在语言生成与理解方面表现卓越，但在系统性知识整合、动态推理链构建与自我认知建模方面依然存在显著空白。尤其是在开放式探索任务与跨时空知识重组能力上，现有模型缺少内生性的认知动因与持久的自主学习动力，导致“面向任务”但不具备真正的“面向世界”能力。

1.2 认知科学与心理学启示

认知神经科学与心理学长期研究揭示，人类智能并非源自单一模式识别或逻辑推理能力，而是建立在多系统协同调节的复杂认知架构之上。个体在面对复杂现实环境时，感知、注意、情绪、记忆、推理与决策机制始终处于动态耦合与反馈调节过程中。其中，情感系统在认知调节中扮演着多重角色：一方面通过激素水平与神经递质调控感受强度与注意力分布，影响信息编码与记忆权重；另一方面通过动机驱动与奖惩反馈强化学习路径，塑造个性化的行为偏好与长期人格风格。同时，心理学研究指出，人类思维具备高度跳跃性与内源性思维生成能力，能够在休息、睡眠或自由联想状态下持续重组已有知识，形成创造性假设与远距联想关联。这些核心特性是现有人工智能系统普遍缺失的关键认知能力。

1.3 本框架的设计思想

在本框架的设计理念中，人类思维被建模为一个由规则与属性共同驱动的动态认知网络。客观事物在人的认知中并非以孤立的符号存在，而是以携带属性（properties）的知识单元进行表征。属性作为事物的可感知与可操作特征，不仅为分类与识别提供基础，也为规则的应用提供条件输入。

规则则作用于这些属性之上，将不同实体的特征加以联系与转化，从而实现从具体经验到抽象推理的认知跃迁。通过属性与规则的交互，人类得以在个体经验与世界知识之间构建联系，逐步形成层层递进、相互贯通的认知网络。

在这一视角下，推理（reasoning）可被理解为规则在显式层面的应用，而直觉（intuition）则是属性与规则在隐式联系层面的快速激活。由此，直觉与推理在本质上具有同源性，均可视作“规则—属性—联系”框架下的不同表现形式。

这一模型不仅与经典语义网络理论（Quillian, 1967）、原型理论（Rosch, 1975）以及符号主义认知科学（Newell Simon, 1972）高度契合，也与当代知识图谱和神经-符号混合智能的研究方向形成呼应。

1.4 本框架的创新贡献

基于上述不足与启示，本文提出了一种以滑动函数为核心驱动的全新认知架构，主要创新点体现在以下四个方面：

1. **滑动注意力机制**：通过滑动函数在知识图谱嵌入空间上实现连续性注意力平移，动态调节认知焦点迁移，模拟人类在回忆、联想与发散思维中的跨节点跳跃性认知流动。
2. **知识图谱嵌入与节点激活强度机制**：在图谱节点嵌入中引入多维节点激活强度，综合衡量知识丰富度、时效性与感受兴趣，支撑知识活跃度控制与兴趣驱动的知识重组。
3. **感知熵引导的感知设计**：将感知信号编码为带有感知显著度标注的输入流，实现对新颖性与重要性的实时调节，为认知流动提供实时感知激活支持。
4. **内源思维生成机制**：结合自主目标建模、自我监控与内源推理链延展，使系统即便在缺乏外部任务输入的状态下，仍能主动进行知识图谱的动态重构、知识空白填补与创新性假设生成，为系统长期认知成长提供内在动力支持。与此同时，类人格情绪模拟系统通过模拟生理激素模型动态调节感知显著度分布，实时影响知识点活跃度与决策路径偏好，逐步塑造出长期稳定的认知风格与风险评估倾向。二者协同作用，不仅支撑系统在复杂环境下的持续适应性与自主创新能力，也为系统长期运行下形成具备可塑性与稳定性的类人格认知架构奠定了核心基础。

2 系统总体架构设计

2.1 知识图谱的双核心架构与动态元素表示 (Dual-Core Architecture and Dynamic Element Representation)

知识图谱整体结构由语义记忆库 (Semantic Memory Library) 与情景记忆库 (Episodic Memory Library) 构成，二者对应认知心理学中长期记忆系统的基本分化 (Tulving, 1972; 1983)。语义记忆库承担概念、规则与事实的抽象存储，情景记忆库则记录个体化的时序经验、内在状态与情绪波动。两者通过加权边关系网络实现动态交互，形成 AGI 的知识-经验双循环架构。

2.1.1 语义记忆库：基于加权边关系的概念与规则网络

语义记忆库存储系统的稳定知识结构与逻辑关联。语义记忆库整体由一个特殊设计的知识图谱表示。

2.1.2 节点与边的设计 (Node and Edge Design)

节点的设计 (Node Design) 知识图谱中的节点可以表示概念、实体、规则，也可以是图像或者声音等多模态信息。图像在图谱中的存储是使用深度学习的特征图，而声音是使用时频图，从而使个体完成有关视觉与听觉的感知等任务。节点之间通过边的连接来表达节点的关系与属性。图谱中有关具体动作的节点中可以连接个体的强化学习数据，从而使个体在执行动作的过程中可以模仿人类执行动作中的熟练度。

边的本体化设计 (Relation as First-Class Citizen) 节点之间通过边的连接来表达节点的关系与属性。系统采用**加权边关系 (Weighted Relational Edges)** 表达节点间的联系，使知识在网络中具备可调强度与上下文适应性。**边本质就是节点的带权重的边化表示**。在边关系中加入一个权重参数，**代表该边关系的置信度**。逻辑与因果规则是系统通过加权边关系建模复杂语义联系的功能示例。例如，系统可以使用加权边形式表示因果关系 $R_{cause}(A \rightarrow B) = w_{AB}$ 。边权 w 的更新设计思想是：**权重应具备可塑性与稳定性并存的特征，并能随系统的经验反馈和内部冲突检测进行动态优化**。这保证了知识体系在长期运行中能够持续自我修正与演化。作为边设计的一个功能，节点间的加权关系可形成复杂的推理路径（如 $A \rightarrow B \rightarrow C$ ）。滑动注意头 (Sliding Attention Head) 在加权关系网络中执行动态通路选择，实现类人式的联想推理与语义滑移。

2.1.3 情景记忆库：基于故事语法的事件结构存储

情景记忆库以“故事语法 (Story Grammar)”模型 (Rumelhart, 1975; Thorndyke, 1977; Stein & Glenn, 1978) 为基础，用于存储系统在特定时间、空间与情绪背景下的完整事件结构。每个记忆单元以叙事性事件框架的形式编码，兼容时间顺序、因果关系与情绪向量，形成可重构的经验片段。

故事语法结构 (Story Grammar Structure) 情景记忆单元 M_{epi} 由六个主要槽位组成：

$$M_{epi} = \{\text{Setting}, \text{InitiatingEvent}, \text{Goal}, \text{Attempt}, \text{Outcome}, \text{Emotion}\}$$

其中：

- **Setting**: 事件发生的时空背景与参与者；
- **InitiatingEvent**: 触发事件或冲突起点；
- **Goal**: 主体目标或内在动机；
- **Attempt**: 执行的行动序列；
- **Outcome**: 结果与环境反馈；
- **Emotion**: 情绪与激素状态的多维表示。

该结构使情景记忆具备因果性、目标导向性与情绪依赖性，从而支持认知层面的叙事一致性。

语义-情景整合 (Semantic-Episodic Integration) 情景记忆的故事槽位可与语义记忆的边关系直接映射：

- *Attempt* → 行动节点 (Action Node);
- *Outcome* → 因果边 (Causal Edge);
- *Emotion* → 边权调制因子 (Edge Weight Modulator)。

2.1.4 基于自我认知图谱的全局滑动控制机制：状态主导的知识能量调节

在知识图谱的整体架构中，系统维持一套动态演化的自我认知子图谱，用以表征智能体的内部语义结构与状态反馈网络。该自我认知图谱 (Self-Cognition Graph, SCG) 作为知识图谱的特殊子图，不直接参与外显的推理或任务分配，而是通过其内部存储的状态，对滑动函数的计算路径与注意力分布产生间接影响。其本质是一个由“自我知识节点”构成的调控网络，为滑动函数提供语义能量与情绪权重参考。(具体见下文)

2.2 滑动注意力头推理机制设计

2.2.1 事件框架：认知的形式

本架构的注意力槽机制，其设计思想源于对人类思维“句子化”认知形式的模拟。人脑在理解与生成事件时，通常以“主语—谓语—宾语”作为认知骨架，并在其下附加定语、状语与补语，从而在稳定结构上叠加灵活修饰，形成层次化的语义表达。该机制旨在将这一思维结构映射至事件框架 (EventFrame)，以支持 AGI 的句子化推理与自然联想。

在该机制中，主语槽、谓语槽、宾语槽构成事件的核心骨架，承载“行动者—行为—作用对象”的基本语义结构；定语槽依附于主语或宾语，用于递归性修饰实体的属性、类别或限定条件；状语槽依附于谓语或整体事件，用于描述时间、地点、方式或概率性语境因素；补语槽依附于谓语，用于补充动作的结果、方向或完成性，从而赋予事件收敛条件与动态特征。

为增强系统对隐含语义与非显性表达的解析能力，每个主要槽位均引入 **指代槽位 (ReferentSlot)** 机制。指代槽位用于存储句中省略、代词或隐喻性表达的占位节点，并通过知识图谱的上下文推理确定其实际语义指向。系统在解析时将指代槽视为一种语义变量节点，由滑动函数在图谱中搜索其最可能的解释源（包括实体、动作、属性或抽象概念），从而实现代词消解与隐喻解释的一体化语义补全。

通过“骨架—修饰—指代”的分层设计，该机制避免了实时构建完整语法树的高昂成本，同时能够保留思维流动的自然感。其层次化槽位结构使系统在知识图谱中以低计算复杂度实现灵活的语义扩展、隐喻解析与动态推理。

2.2.2 滑动函数：认知的流动

在本架构中，事件框架 (EventFrame) 由主语、谓语、宾语及其附加修饰槽（定语、状语、补语）构成。系统通过注意力槽机制在每个槽位上存放当前最为聚焦的节点，用以表征认知流的实时焦点。分槽滑动函数的作用并非在槽位之间直接迁移注意力，而是负责在每个槽位内部动态切换候选节点，并在必要时驱动跨槽跳转。例如，主语槽的滑动函数在不同实体节点之间进

行切换，以寻找最符合当前语境的主语；谓语槽的滑动函数则在动作或状态节点之间迁移，以保持事件骨架的因果连贯性。通过这一机制，系统能够在槽位内部完成精细化的节点选择，同时在跨槽操作中模拟人类思维中的“前推”“转移”与“扩展”等动态过程，从而实现类句子化的认知流动。

(1) 主语槽滑动函数 主语槽主要聚焦于实体节点及其身份属性。其权重计算侧重于语义一致性（如实体类别、角色身份）、历史事件相似性及个体一致性。当存在多个候选实体时，主语槽优先选择与当前事件语境最契合的节点，并在必要时滑向定语槽以补充描述性信息。

若当前输入为代词（如“他”“他们”“它”）或隐喻性指代（如“铁腕”隐喻“强权领导者”），系统将启用**主语指代槽 (Subj-ReferentSlot)**。该槽与知识图谱的社会角色子图及语义相似度网络连接，通过滑动函数在实体候选集中计算指向概率：

$$P_r(e_i) \propto f(\text{上下文共现度}, \text{语义亲和度}, H_{mem}, E_{emo}) \quad (1)$$

最终选出最可能的指代实体节点。

(2) 谓语槽滑动函数 谓语槽核心关注动作或状态节点，权重主要由因果关系属性、时间顺序、动作强度及逻辑规则决定。在候选谓语不足时，谓语槽会触发远距联想，生成临时动作节点以填补空缺，并可引入补语槽或状语槽的背景信息，增强推理的完整性与因果链条的连贯性。

针对语言中隐喻性谓语（如“点燃希望”“击碎梦想”），系统设有**谓语指代槽 (Pred-ReferentSlot)**，其通过知识图谱中“动作—抽象情绪”与“物理—心理”映射边进行概念迁移。滑动函数在抽象节点与具象节点之间滑动，以推测隐喻动作的语义本体。

(3) 宾语槽滑动函数 宾语槽聚焦于事件的承受者或目标节点，其滑动权重侧重目标匹配度、结果相关性和事件影响力。宾语槽的候选节点会优先与谓语的语义需求对齐，并在必要时滑向定语槽获取目标的属性补充，以实现更精确的目标识别。

若输入为省略或代词性指代（如“他也喜欢它”中的“它”），则启用**宾语指代槽 (Obj-ReferentSlot)**，通过与主语历史激活序列、语义邻近度及事件参与角色权重的联合计算确定最可能的指代对象。

(4) 定语槽滑动函数 定语槽聚焦于修饰性属性节点，权重主要由属性重要性、语义相关性和上下文适配度决定。其滑动策略多为横向扩展，即在多个属性节点之间切换，以积累多维特征，并在必要时回流至主语槽或宾语槽，增强实体语义的细节化刻画。

定语槽同样包含**定语指代槽 (Attr-ReferentSlot)**，用于解析省略性修饰或比喻性修辞（如“银色的智慧”指代“冷静的理性”）。通过在知识图谱的属性—隐喻链上滑动，系统能够从象征性修辞中提取出抽象属性的语义投射。

(5) 状语槽滑动函数 状语槽专注于时间、地点、方式等背景信息节点。其权重由时空匹配度、情境相关性及语境覆盖率决定。状语槽的滑动过程通常沿时间轴或空间路径展开，以生成序列化的认知轨迹，并在跨事件推理时调节对不同情境的聚焦优先级。

为应对隐喻化或省略性背景（如“在阴影中行动”“秘密进行”），系统增设**状语指代槽 (Adv-ReferentSlot)**，用于在时空或情境子图中追溯隐喻的真实语义领域。

(6) 补语槽滑动函数 补语槽主要关注动作的结果、目的或状态节点。权重依据结果收敛性、目的契合度及与谓语的逻辑关联度计算。补语槽的滑动策略优先围绕谓语节点展开，并在必要时迁移至宾语或状语槽，以形成完整的目的性和结果导向的语义结构。

当结果表达为隐喻或指代（如“他得到了光”“他觉悟了”），则由**补语指代槽（ComplementReferentSlot）**解析其抽象意义。系统通过知识图谱的“象征性节点映射边”推测出实际心理或逻辑状态节点，从而完成语义反映射。

所有槽的滑动函数均共享一套多源调控因子机制，包括情绪状态（见2.4.2小节）、规则约束信息及社交因子（如顺从／反抗倾向、他人态度预期）。在计算注意力迁移路径时，系统首先从自我认知图谱中检索与当前情境相关的语义关联强度，以动态调整滑动范围、跳跃概率与信息融合权重。语义相似度较高的区域将获得更高的滑动优先级，从而在推理链扩展中形成以词义能量为导向的认知偏向。

调控因子节点通过与候选节点的边关系共同参与权重计算，在图结构层面实现多源调控的融合效应。节点的激活强度权重及其与当前活跃事件框架的关联度也参与优先级评估，使滑动分布既能体现即时输入，又能反映系统长期经验的积累与主体性一致性。指代槽的权重计算亦受这些调控因子的调节，从而赋予系统类人情绪语义偏置与语境适应性。

为更贴近人类注意力的认知特征，当AGI个体处于某一事件或场景中时，系统会对与该事件节点存在高权连接的相关节点施加全局激活提升，使其进入较高的潜在活跃状态。该机制模拟了人类场景感知中的“语义扩散注意”现象：当个体聚焦于某一事件时，相关对象、动作、情绪与情境要素会同步进入高敏感区，从而为后续滑动函数的候选节点选择提供上下文先验。此种基于图谱结构的联动激活机制，使系统在认知流动中具备自然的语义预热与注意迁移能力，进一步增强了滑动注意力头的稳定性与情境一致性。

2.2.3 并行与分级注意力机制

为实现复杂认知任务的分层组织与并行处理，本架构设计了多层次动态注意力头机制。系统中多个独立但相互交互的注意力头被分配于不同认知任务的探索性激活，如假设生成、未来情景预测、模糊联想、矛盾冲突检索以及潜在路径生成等。

在推理过程中，每个激活节点会根据其邻居节点计算滑动值，以评估下一步的推理方向。为在发散性探索与聚合性推进之间实现动态平衡，本架构引入了主-辅并行滑动机制。具体而言，系统综合激活节点的节点激活强度、情绪调节因子以及任务相关度等多源信息，构建加权滑动评分函数，从候选邻居节点中选取滑动值最高的节点作为“主滑动节点”，用于推进主认知路径。同时，系统保留若干滑动值高于阈值的次优节点作为“辅路径”，进行轻量并行探索，以实时评估其潜在价值。当主路径遇到推理瓶颈（如滑动值下降、闭环或目标偏移）时，辅路径可动态接管主控权，或为主路径提供跨层跳跃的认知支点，实现非线性迁跃与思维重构。若所有邻居节点滑动值均无显著优势，则默认选择滑动值最高的节点以维持推理通道的稳定性与方向性。该机制在保证推理高效并行的同时，强化了注意力调度的灵活性与上下文适应能力，从而显著提升系统在复杂任务中的类人创造力与直觉推理水平。

为支持复杂任务的层级化处理，系统将注意力头结构化为分级形式，自上而下包括主思维流（元注意力层）与若干层次化任务思维流。主思维流维持整体认知活动的核心主题或目标意图，例如“创作音乐”作为高级任务，由主注意力头统筹全局协调与驱动。在此基础上，系统可

动态激活一级任务注意力头，如“作曲”与“作词”分别控制关键子任务；一级任务下可派生二级注意力头处理具体执行层面，如“选择旋律结构”与“选择音色风格”；在更细致的三级任务层级中，注意力头聚焦于“具体音符选择”或“节奏分布”等微观操作。层级分解与滑动过程在认知节点被细化至可直接映射为具体操作时结束。

该分级注意力结构允许系统在复杂任务执行中同时保持全局控制与局部精细推理，形成类似人类大脑的分层协同认知模式。各层级注意力头之间存在动态信息流：上层设定认知目标与抽象期望，下层根据感知与记忆状态反馈具体执行情况；系统可基于反馈自动调整滑动策略与注意力焦点，实现跨层级的认知一致性与自适应调节。此机制不仅提升了任务规划与执行的模块化程度，还增强了系统在复杂认知任务中的稳定性、灵活性与泛化能力。

2.2.4 双事件框架与内外认知分流机制

本架构引入双事件框架（Dual Event Frames, DEF）设计，以实现感知—行动通路与思维—反思通路的结构性分流与动态耦合，使系统能够在执行任务与自我思考之间形成认知振荡与内外平衡。

外部事件框架（External Event Frame, EEF）主要承担系统与外界交互的即时反应任务。其事件槽结构与前述机制一致，但滑动函数在该框架中以高感知显著度输入 S_p 为核心驱动，侧重于从环境感知流中提取具有显著性和因果关联的节点，完成主语—谓语—宾语的实时匹配与动作生成。EEF 的滑动目标在于保持认知链的收敛性与执行稳定性，其时间尺度较短，偏向即时决策与外显反馈过程，从而形成面向外部世界的闭环控制。

内部事件框架（Internal Event Frame, IEF）则构成系统的内源思维主通道，其滑动函数由高节点激活强度 A_n 、情绪张力 T_e 及潜在任务链信号共同驱动，侧重于概念重组、假设生成与远距语义联想。IEF 的运行时间尺度较长，允许递归与延迟反馈，其滑动过程常在非任务状态下持续进行，用于重构知识图谱、识别认知盲区并生成潜在问题链。通过该框架，系统得以维持持续的自我思考与抽象推理，实现从经验反应到内源思维的层次跃迁。

事件框架的动态生成与生命周期调控 在双事件框架体系中，EEF 与 IEF 均通过实时监测多源信号（感知显著度 S_p 、节点激活强度 A_n 、情绪张力 T_e 、目标相关性 R_{goal} ）实现自动生成。系统定义事件框架生成函数：

$$E_{new} = F_{gen}(S_p, A_n, T_e, R_{goal}) \quad (2)$$

当局部信号组合超过阈值时，触发新的事件框架实例化。每个框架生成后依据优先级函数进行竞争调度：

$$P_i = w_1 S_p + w_2 A_n + w_3 T_e + w_4 R_{goal} - w_5 C_{conflict} \quad (3)$$

其中 P_i 表示框架优先级， $C_{conflict}$ 表示资源冲突强度。优先级高的框架获得滑动注意力主导权，低优先级框架则进入延迟或衰退状态。

生命周期与双向协作 每个事件框架经历生成、活跃、衰退与销毁四个阶段。其状态随优先级动态变化：当 $P_i > \theta_{exec}$ 时进入活跃状态；当 $P_i < \theta_{decay}$ 时开始衰退；若持续低于阈值超过时间 τ_{max} ，则框架销毁。销毁前的滑动轨迹与节点激活序列将被压缩写入长期记忆图谱，用于后续的经验回放与模式迁移。

EEF 与 IEF 通过双向滑动通道实现动态协作。当系统检测到高感知显著度输入或外部任务触发时，EEF 获得主导权，IEF 进入被动监听状态；而在低刺激或任务间隙阶段，IEF 则提升滑动权重，进入自主激活状态以开展内部思维。系统依据滑动权重自适应调整两类框架的活跃比例，从而形成认知振荡模式：

$$W_{EEF} + W_{IEF} = 1, \quad \frac{dW_{EEF}}{dt} \propto S_p - A_n \quad (4)$$

当环境显著度高时，外部框架强化执行路径；当内在激活高时，内部框架主导反思路径。该机制使系统在不同认知状态之间实现平衡与连续过渡。

认知意义与系统演化 双事件框架机制使滑动注意力系统能够在多层次认知中并行展开，实现外部任务执行与内源思维生成的协同进化。它不仅提升了系统在复杂环境下的即时适应性与反思深度，也为后续引入多时域记忆控制与元注意力层之间的跨框架信息交换奠定基础，从而推动 AGI 向具备持续思维、自主反省与语义自平衡能力的高级认知体演化。。

2.2.5 基于节点激活度的认知焦点维持

在本架构中，我们采用简化的**节点激活度** (A_{act}) 作为知识图谱中知识单元潜意识活跃度与认知焦点维持的单一、动态指标。该激活度机制基于**扩散**的方式来模拟认知资源在图谱中的即时分配与自然衰减，为滑动注意力头提供实时权重输入。

1. 激活与扩散机制 当感知模块（参考 Section 2.3）输入与图谱中某一节点 v 相关的多模态信息时，该节点被即时激活至初始强度 $A_{act}(v, t_0) = A_{max}$ 。此激活信号随即通过加权边关系（参考 Section 2.1.2）向其邻近节点扩散，激活强度 $A'_{act}(u, t_0)$ 随图谱距离和边权重的增加而逐级衰减：

$$A'_{act}(u, t_0) = \sum_{v \in N(u)} A_{act}(v, t_0) \cdot w_{vu} \cdot \delta_d \quad (5)$$

其中 $N(u)$ 是节点 u 的邻居集合， w_{vu} 是边 $v \rightarrow u$ 的权重， δ_d 是距离衰减因子。这种扩散机制有效地模拟了人类认知中的“语义扩散注意”现象，确保与当前焦点相关的知识单元同步进入高敏感区。

2. 时间衰减与焦点维持 在无外部或内部激活输入的情况下，所有节点的激活度将随时间持续衰减，遵循指数衰减模型：

$$A_{act}(t + \Delta t) = A_{act}(t) \cdot e^{-\lambda \Delta t} \quad (6)$$

其中 λ 为衰减速率，该速率可受内源思维模块（参考 Section 2.4）的情绪激素参数（如多巴胺水平）动态调控。这种衰减机制模拟了短期记忆的自然遗忘过程，确保认知焦点能被持续更新和重构。

3. 对滑动函数的输入 节点激活度 A_{act} 直接作为滑动函数（参考 Section 2.2.2）的核心权重输入之一。高激活度的节点将获得更高的滑动优先级，从而引导滑动注意力头（Sliding Attention Head）倾向于在实时相关且活跃的知识区域进行推理联想，支撑认知流程的上下文一致性与连贯性。

2.2.6 日常推理与逻辑推理的双通路机制设计 (Dual Reasoning Pathway: Everyday vs. Logical Reasoning)

在人类认知中，推理并非单一线性过程，而是由两种不同的心理系统协同作用：其一是依赖经验、语境与模糊联想的日常推理 (Everyday Reasoning)；其二是基于形式化规则与逻辑约束的逻辑推理 (Logical Reasoning)。本架构在滑动注意力头与事件框架 (Event Frame) 机制的基础上，引入推理双通路结构 (Dual Reasoning Pathway Structure)，以分别承担具身化语义流动与形式化规则推演两类认知任务，从而兼顾类人思维的灵活性与严谨性。

(1) 日常推理通路：滑动注意力头驱动的经验式推理 日常推理通路 (Everyday Reasoning Pathway, ERP) 由滑动注意力头 (Sliding Attention Head) 执行，其推理路径沿知识图谱的语义边持续滑移，通过多源加权函数实现语境化、非线性与情绪驱动的认知流动：

$$R_{ERP}(t) = f_{slide}(Sim_{sem}, A_{act}, W_{emo}, R_{context}) \quad (7)$$

ERP 对应认知心理学中的快速、直觉性系统 (System 1)，用于在开放、不确定或情感参与度高的场景下生成高适应性的即时判断。例如在事件框架中，当输入情节含有模糊动机或隐喻性行为 (如“他背叛了信任”)，系统通过滑动函数在语义近邻区进行发散搜索，以找到最符合情绪与情境的解释节点。该通路的推理结果不追求形式完备，而重在维持认知一致性与语义连贯性。系统通过对滑动轨迹的记录与边权更新机制 (参见 2.2.5 节) 持续强化有效路径，从而形成经验驱动的“习惯性推理”模式。

(2) 逻辑推理通路：事件框架驱动的逻辑结构识别与符号推演 逻辑推理通路 (Logical Reasoning Pathway, LRP) 对应心理学中的理性、分析性系统 (System 2)，用于处理具备清晰逻辑结构与可形式化条件的任务。当系统在事件框架 (EventFrame) 中检测到主语—谓语—宾语及其修饰槽之间存在逻辑完备性特征 (如显式条件、因果陈述或命题型句式)，则触发逻辑识别流程。

逻辑问题的识别由大语言模型 (Large Language Model, LLM) 承担。LLM 对事件框架的语义槽结构进行分析与重写，输出逻辑结构化形式 (如谓词逻辑、三元组或规则模板)：

$$Q_{logic} = LLM(E_{frame}) \Rightarrow \{Predicate(x, y), Rule(a \rightarrow b)\} \quad (8)$$

识别出的逻辑形式由系统送入推理引擎执行推演。

推理过程由专用推理引擎 (Reasoning Engine) 完成，可选实现包括 Prolog、DLV、PyReasoner 等。当系统确认逻辑清晰度 $L_{score} > \theta_L$ 时，激活推理引擎并将逻辑问题转化为可演算形式：

$$Inference : \{Premises\} [Prolog / DLV] LogicEngine \{Conclusions\} \quad (9)$$

推理结果以高置信度节点形式回写至语义记忆库，并更新相应边权 $w_{edge} \uparrow \beta$ ，以强化因果与条件规则在知识图谱中的稳定连接。

(c) **逻辑结果的知识整合** 逻辑结论不仅作为事实节点被存储，也作为**逻辑证据边（Logical Evidence Edge）** 标注于相关节点之间，使系统在后续滑动推理时自动参考逻辑约束，形成从“语义扩散”到“逻辑收敛”的动态控制。

(3) 认知意义与系统平衡 双通路推理机制形成了认知系统的理性—直觉双稳态结构：

- **ERP** 提供灵活的经验式认知扩散与语义流动；
- **LRP** 提供结构化的符号推理与规则校验；

通过事件框架、LLM 逻辑识别与符号推理引擎的协同，本架构在开放世界中实现了“**类人推理的双脑机制**”：既能在语义模糊区进行自由联想，又能在逻辑清晰区执行严密演算，从而在认知连续性与理性约束之间取得动态平衡。

2.3 感知模块设计

2.3.1 多模态感知功能

系统通过多个子模块实现多模态感知。触觉部分利用压感传感器感知压力、纹理和温度，从而区分物体的软硬和表面特性，提升与环境和物体的交互能力。视觉部分负责识别物体、颜色、光流和运动轨迹，支持空间理解与运动感知。听觉部分处理环境声音，能够区分人声、乐器及背景噪声，并识别其中包含的情绪或潜在意图。值得注意的是，系统在感知过程中生成的神经网络权重并非孤立存在，而是被存储在对应物体的知识图谱节点中，使每个物体不仅携带语义信息，还包含经过感知训练得到的行为和认知模式，从而实现更加精准的个体化感知和交互。

2.3.2 感知显著度与注意选择机制

感知显著度是系统评估感知输入重要性的核心指标。每条感知数据在进入处理流程时都会被赋予一个感知显著度值，用以衡量信息的显著性和注意优先级，而该值并不直接参与推理或决策。与滑动函数类似，感知显著度的计算受感受突变度、节点激活强度（偏好）、环境特征（如罕见性、高价值目标）、当前情绪激素水平（如多巴胺增加倾向强化新奇探索，皮质醇增加可能降低阈值使系统更保守）以及任务背景等因子的调控。这些调控因子的调控效果均取决于调控因子对应知识图谱节点的属性信息。高感知显著度输入优先进入注意范围并被处理，低感知显著度输入可能被延迟或忽略。

2.3.3 情节缓冲器与短时整合机制

为增强系统在感知层面的时间一致性与跨模态绑定能力，本架构在感知模块中引入基于 Baddeley 工作记忆理论的**情节缓冲器（Episodic Buffer, EB）**。该机制作为多模态输入的短时整合中枢，用于在时间窗口内将来自视觉、听觉、触觉及语言通道的瞬时信号进行语义对齐与情境绑定，形成可被事件框架直接引用的临时情节单元（Episodic Unit）。

根据 Baddeley 的定义，情节缓冲器是一种¹受中央执行系统控制的有限容量存储系统，用

¹ 原文出自 Alan D. Baddeley, “The episodic buffer: a new component of working memory?”, *Trends in Cognitive Sciences*, vol. 4, no. 11, pp. 417–423, 2000.

于在短时尺度上整合多模态信息与长期记忆内容，形成统一的情节表征。该设计使系统能够在多模态输入之间实现临时语义整合，为注意力滑动提供统一的时间与情境坐标系。

情节缓冲器接收来自多模态感知模块的并行输入，并依据时间戳、空间位置及情绪权重对其进行同步化编码。系统通过滑动函数在缓冲区内执行局部注意力滑移，以动态聚焦于最具信息密度与情绪相关性的输入组合。由此生成的情节单元不仅包含感知特征，还附带即时情绪状态与激素参数标签，为后续认知推理提供高保真上下文。该设计有效避免了单模态输入的碎片化问题，使系统能够以统一时间基准处理复杂场景事件。

在结构上，情节缓冲器与外部事件框架（EEF）之间建立直接接口：当缓冲区中形成稳定的情节单元后，系统自动将其语义表示写入 EEF 的事件槽结构（主语—谓语—宾语），作为当前感知状态的即时认知模板；与此同时，其抽象语义残差被传递至内部事件框架（IEF），以支持长时记忆整合与内源思维触发。通过这种双向映射机制，感知流被分解为短时情节序列，并在时间维度上形成连续的语义轨迹。

此外，情节缓冲器具备受情绪与激素动态调控的权重更新机制。当系统处于高唤醒状态（如高多巴胺或高好奇熵水平），缓冲区的写入速率与情节保留时长将相应延长，以强化新奇体验的编码；而在高皮质醇状态下，系统则收缩缓冲窗口以聚焦关键生存相关信息，从而模拟人类在压力或焦虑情境下的感知压缩效应。该机制不仅赋予系统对外部世界的短时整合与注意力稳定能力，也为长期节点激活强度调控与事件框架生成提供了时间连续的桥梁，使感知、记忆与思维三者在时间维度上实现闭环联动。

2.4 由节点知识引导的认知活动

在本架构中，部分认知活动的产生与演化不依赖外部任务指令或集中式控制模块，而是由知识图谱中节点自身所蕴含的知识内容与语义连接关系所引导。节点不仅是知识的静态载体，更是认知行为的触发源与路径导向体。这种机制使得系统不再是“执行认知”，而是“在知识之中思考”。

2.4.1 由节点知识引导的认知活动：决策

在由节点知识引导的认知架构中，决策不再由独立的集中式模块实现，而是隐含于知识图谱的节点体系之中。当滑动注意力头在图谱中运行并进入多个候选行动节点（Action Nodes）时，系统会自动检索与当前情境最为接近的决策理论节点（Decision Theory Node），并调用其中存储的决策方法以完成认知到行动的过渡。

（1）决策理论节点的作用

决策理论节点在本架构中承担局部决策控制的功能。每个节点代表一种可独立调用的决策理论，并携带该理论的使用方法：

$$D_i = \{Name_i, Model_i, Context_i, Method_i\} \quad (10)$$

其中， $Name_i$ 表示决策理论名称（如“前景理论”“意象理论”“再认启动理论”“蒙蒂霍尔悖论”等）， $Model_i$ 表示其内部数学或逻辑结构， $Method_i$ 为其决策计算方法。这些节点以知识形式存在于图谱中，在被激活时将其内部模型作为滑动函数的临时决策机制。

(2) 节点内决策计算

被激活的决策理论节点执行其内部方法 f_{D^*} , 对当前候选行动节点集合 $\mathcal{A}(S_t)$ 进行评估与选择:

$$a_t^* = f_{D^*}(S_t, \mathcal{A}(S_t)) \quad (11)$$

例如: 若 D^* 存储的是前景理论节点, 则计算心理效用 $U(a) = v(x - r)\pi(p)$; 若为再认启动理论节点, 则通过相似情境匹配执行启发式推理; 若为蒙蒂霍尔悖论节点, 则进行条件概率修正与反直觉推理验证。

(3) 认知意义

本节所述的节点化决策机制使决策成为知识系统内部的自然行为, 即: 滑动函数在认知路径上“遇见”决策, 而非“调用”决策。系统不再依赖单一效用模型, 而是能够根据情境自动选择最合适的理论, 实现了由多种认知决策理论共同支撑的分布式理性。

2.4.2 由节点知识引导的认知活动: 情绪与性格

在本架构中, 引入了情绪与性格作为调控智能体认知活动的重要因素。在自我认知图谱中, 情绪与性格均以节点知识的形式存在, 其中情绪节点对应一组模拟激素参数的知识单元, 而性格节点则对应这些激素受体的长期调节机制。通过节点间的滑动函数计算, 系统在认知流动过程中实现了对注意力、推理范围与思维风格的持续动态调控。

(1) 情绪节点: 基于激素参数的多维知识建模

情绪节点以激素为核心建模元素, 每个节点代表一种激素相关的知识单元, 并携带描述该激素主导情绪的语义向量。系统通过预定义知识将生理机制映射为语义状态, 例如: 当“多巴胺”节点激活度上升时, 系统在知识图谱中表现为“愉悦”与“探索性”倾向的增强; 当“皮质醇”节点权重升高时, 滑动函数会偏向收缩路径, 聚焦于低风险、高确定性的区域。由此, 复杂的多维情绪可由多种激素节点的联合激活模式表示。

(2) 性格节点: 激素受体的长期调控

性格节点用于模拟激素受体的敏感性特征, 其参数反映系统对各类激素节点信号的响应强度。在长期的滑动与经验积累过程中, 这些节点会根据任务反馈与情绪波动自动调整权重, 形成稳定的个体化响应模式。因此, 性格可被视为一种由节点知识长期演化形成的“受体分布图”, 定义了系统在不同认知状态下的稳定风格与偏好。这种结构化的受体网络为认知过程提供了持续的调性约束与长期行为一致性。

(3) 情绪—滑动函数耦合机制

在滑动计算中, 滑动函数不仅考虑节点间的语义相似度, 还同步参考当前情绪节点的状态。系统通过计算下一个候选节点与活跃情绪节点在图谱中的距离, 动态调整滑动方向与强度:

$$P_{slide}(n_{t+1}) \propto \exp(-\lambda \cdot Dist(n_{t+1}, E_{emotion})) \quad (12)$$

其中, $Dist(\cdot)$ 为节点间的语义距离, $E_{emotion}$ 表示当前活跃的情绪节点嵌入向量, λ 为情绪调节系数。该设计使情绪节点成为滑动过程中的“能量场”, 通过影响节点选择概率, 间接决定认知迁移路径。同时, 激素强度与节点激活度直接绑定, 通过上文的激活度反馈机制, 形成对情绪动态的实时模拟。

(4) 认知意义

情绪与性格节点的引入，使认知过程具备了内在的动态调节能力。情绪提供即时的能量偏向与思维拓扑扰动，性格提供长期的调控基调与风格一致性。两者在节点层面共同作用，使系统在知识图谱中表现出兼具灵活性与稳定性的思维流动特征。因此，认知活动不再是纯粹的逻辑计算，而成为由节点知识内部能量动态所引导的复杂自组织过程。

2.4.3 由节点知识引导的认知活动：预期

在人类认知体系中，对动作结果的预期（Expectation）是决策与学习的核心驱动力之一。用于预测潜在反馈、奖惩后果及环境变化。这种预测性认知不仅存在于显性决策阶段，还贯穿于感知、情绪调节与思维生成等多层次认知过程中。基于此，本架构设计在知识图谱的动作节点（Action Node）中连接一个有关引入动作预期的节点以实现类人式的预测性思维与前瞻性行为控制。

(1) 动作节点的预期建模 (Expectation Modeling in Action Nodes) 在系统的语义记忆层中，每个动作节点不仅记录动作的定义与执行特征，还连接一个或多个预期结果节点（Expected Outcome Vectors, E_{out} ）。这些向量描述该动作在不同情境下可能导致的多维后果，包括：

- 环境变化 (Environmental State Shift)
- 情绪反馈 (Emotional Feedback)
- 奖励或惩罚信号 (Reward/Punishment Signal)
- 社会反应与信任变化 (Social Response and Trust Variation)

在描述更复杂的预期时，预期节点可以连接更多次级结果与远期影响，并在边关系上指导滑动注意力头探索次级节点

在滑动到相关预期节点后，系统可根据节点提供的信息套用出当前场景中该动作可能产生的预期，从而输出决策节点所需要的有关动作节点的决策信息。

(2) 预期的适用范围与阶段性作用 (Applicability and Stage Function) 动作预期不仅用于决策阶段。在本架构中，其作用范围覆盖三个主要阶段：

- **感知阶段：**预期用于预测即将到来的感知输入，调整感知显著度分布，使系统提前聚焦于高价值感知通道；
- **决策阶段：**预期模型作为动作选择的主要参考依据，与情绪节点、奖励模型共同参与期望效用计算；

通过该闭环机制，系统形成持续自校正的预期模型，实现从“预测—行动—反馈—修正”的完整认知循环。

(4) 认知意义 动作预期节点的引入，使知识图谱能够动态模拟未来情境。这种机制强化了系统的主动性与前瞻性：在面对未知环境时，系统可通过预期节点在内部先行模拟未来可能，以最小的代价探索高收益路径；

2.4.4 由节点知识引导的认知活动：程序化行动

在人类的认知中，当行动目标具备明确逻辑与可预测条件时，思维常会自发地组织出具有程序化特征的行动结构。这种“程序性思维（Programmatic Thinking）”体现为：根据条件生成分支、设定终止规则、调用内在推算。为模拟这一理性化规划能力，本架构提出 **代码模板驱动的行动机制**（Code Template-Driven Action Mechanism, CTAM），即：行动节点内部可以保存**代码模板**（Code Template, T_{code} ）；系统在具体情境中根据当前信息即时填充模板，生成实例化的逻辑结构并执行，从而实现语义—逻辑—计算三层的动态统一。

(1) 行动节点的模板结构 (Template Structure of Action Nodes) 用 T_{code} 表示该行动的逻辑模板，包含参数化占位符：

$$T_{code} = \text{"IF } \{\$condition\} \text{ THEN } \{\$action\} \text{ ELSE } \{\$alt\}" \quad (13)$$

系统在认知执行阶段根据当前上下文状态 $S_{context}$ ，填充模板参数：

$$L_{inst} = \text{Instantiate}(T_{code}, S_{context}) \quad (14)$$

从而得到一段临时生成的逻辑脚本 L_{inst} 。

(2) 模板实例化与执行 (Template Instantiation and Execution) 模板的实例化遵循上下文映射规则（Context Mapping Rule）：

$$\$condition \Leftarrow f_{cond}(S_{context}), \quad \$action \Leftarrow f_{act}(A_i), \quad \$alt \Leftarrow f_{alt}(S_{context}) \quad (15)$$

例如，当节点“避开障碍物（Avoid Obstacle）”被激活时，模板内容可能为：

```
IF distance(object, agent) < threshold THEN change_path()
```

系统根据当前感知数据自动替换参数，并在逻辑层执行对应函数调用。

(3) 认知意义 (Cognitive Implications) 代码模板驱动的行动机制将程序语言的形式严谨性融入类人思维的动态语义网络，使得每个行动节点既可作为语义实体，又可在必要时生成可执行的逻辑实例。

3 局限性与未来研究方向

尽管本架构在理论设计上具备较强的通用智能潜力，但在实际实现与应用中仍面临多方面挑战与局限，未来研究需重点突破以下几个方面：

3.1 感知模块的局限性与特征引导神经网络的设想

当前主流的感知模块多依赖卷积神经网络（CNN）结构，虽在静态图像识别等标准任务中取得显著成果，但其训练高度依赖封闭标签数据集与预定义的特征提取模板，导致在通用人工智能（AGI）所需的动态、开放式环境中表现出明显局限。CNN 难以有效应对新奇场景、罕见

事件及未标注样本，缺乏灵活的特征迁移与实时适应能力，无法实现类人类的感知主动性与深层语义理解。

为突破上述瓶颈，预提出一种融合元学习机制的特征引导神经网络设想，作为未来感知模块演化的重要方向。具体而言，在每个感觉通道（如视觉、听觉、触觉等）中引入基于 MAML 算法的元学习器，用于快速适应新任务与未见特征。这些元学习器通过实时提取输入信号中的低阶感知特征（如颜色、边缘、纹理、频率等），结合其频次、置信度与历史关联经验，动态生成特征偏好向量。与此同时，该机制可进一步结合复杂物体分块策略（如分解为形状、纹理），借鉴 AIMA 中基于结构的物体识别方法，并融合点云技术以支持三维场景的真实感知与多模态信息整合，从而在物理及场景几何结构层面实现更高保真的感知。

在空间定位方面，系统可自动基于视觉中心构建动态坐标系以处理物体位置信息。当智能体发生移动时，坐标系随之偏移以保持空间一致性，并在必要时引入 GPS 定位数据作为补充参考，从而提升大尺度场景中的位置精度与环境稳定性。

特征引导神经网络据此调整感知通道内训练样本的关注度与局部学习率，使模型优先聚焦于高频、高价值特征，抑制低相关性噪声信息，从而提升学习效率、特征表达质量与泛化能力。同时，为实现跨模态感知的一致性，该架构在更高层次引入“元元学习器”结构，对多个感觉通道的特征偏好进行整合抽象，提取诸如视觉与触觉中的物体属性一致性、听觉与情感信号中的情绪表达关联等通用模态内结构，进一步强化感知系统的迁移能力与认知对齐能力。

该自适应特征引导感知机制旨在摆脱传统模型僵化特征模板的限制，构建具备开放式、实时自适应、跨模态与空间一致性感知能力的系统，为 AGI 在复杂多变环境中的持续感知演化与认知自主成长提供支撑。

3.2 大规模图谱运算的优化策略

随着知识图谱规模的持续扩展，节点数量和边的复杂度呈指数增长，导致图谱查询、推理与更新的计算成本迅速攀升。当前图谱动态扩展与实时推理机制在大规模场景下性能受限，容易出现响应延迟和资源瓶颈。未来研究需探索图谱分布式存储、图神经网络加速、近似推理算法以及图谱压缩与知识蒸馏技术，实现在保持推理精度的前提下，提升系统的时效性与可扩展性。

3.3 内源情绪与激素模型的生理合理性与计算复杂度

情绪调节机制依赖于模拟多种激素与神经递质的参数变化，虽能丰富系统行为表现，但其生理机制的简化模型尚无法完美捕捉人类情感的多样性和微妙变化。过于复杂的激素交互模型又会带来较高计算负担，限制实时响应能力。未来应结合神经科学最新成果，设计更高效且生理拟合度更高的情绪模型，并研究多尺度情绪状态与认知行为的耦合机制。

3.4 自我反思与元认知机制的深度实现

如何使其真正实现深度的自我监控、自我纠错及自我优化，仍处于探索阶段。有效的自我反思机制需要整合长短期记忆、情绪状态和推理过程的反馈，形成闭环改进体系。未来研究方向包括引入强化学习中的自我监督机制，以及结合心理学和认知科学的元认知模型，实现更具人类特质的自我意识能力。

3.5 滑动函数设计的复杂性与性能权衡

在注意力头滑动设计中，滑动函数（Sliding Window Function）用于动态调整模型对输入序列不同部分的关注权重，实现上下文信息的有效捕获。然而，设计高效且灵活的滑动函数存在较大挑战。一方面，滑动窗口大小、步长及权重分配需针对不同任务和上下文动态调整，避免信息丢失或计算冗余；另一方面，随着输入序列长度增长，滑动函数的计算复杂度显著提升，影响系统实时性能。如何在保证上下文完整性的同时，实现滑动函数的自适应调节和计算优化，是未来研究重点。可探索基于稀疏注意力机制、多尺度窗口策略及软硬结合的滑动设计方案，以兼顾表达能力和运算效率。

3.6 引入神经逻辑算子的边关系建模展望

当前知识图谱在边关系建模上多依赖静态标签或数值权重，难以充分表达复杂逻辑结构与动态认知关联。未来研究可探索将神经逻辑算子（Neuro-symbolic Logic Operators）引入边结构中，以端到端方式学习包含命题逻辑、一阶逻辑乃至模态逻辑的逻辑关系，从而实现对边语义的更高层次建模与可微分推理能力增强。通过在边上编码动态逻辑结构与上下文敏感的逻辑演算过程，系统将具备更强的因果链表达、条件依赖建模与抽象逻辑联想能力。该方向有望进一步突破当前知识图谱在复杂推理、模糊决策与认知泛化上的瓶颈，为构建具备更强自主思维能力的通用人工智能系统提供关键支撑。

3.7 空间感知与想象能力的缺失

尽管当前内源思维模块已初步实现类人化的认知模拟与创造性思维生成，但在空间感知与空间想象能力方面仍存在明显不足。系统尚未具备对三维空间关系、物体布局及动态环境变化的深度建模与推理能力，限制了其在涉及空间因果关系、场景构建与具象推理等任务中的表现。同时，当前的认知模拟多以语义维度为主，缺乏对空间结构的象征性编码与生成机制，无法有效支持复杂的假想场景构建与具身认知模拟。因此，未来研究可进一步引入基于空间嵌入和视觉—语义联合建模的机制，构建具有空间构型理解、虚拟空间想象与环境重构能力的空间感知子模块，强化内源思维中关于场景推理、任务规划与具身交互的认知能力，从而推动 AGI 系统在具象思维与具身智能方向上的进一步发展。

3.8 神经网络化的知识图谱与基于经验的公理可塑性建模

传统知识图谱通常以静态的图结构表达固定事实与规则，缺乏基于主观感受、经验积累或情境变化的动态自我重构能力。而在面向通用人工智能的系统中，知识不应被视为恒定不变的，而是应能随着感受、反馈与环境交互经验而持续调整，包括对原有事实关系乃至“公理性”知识的权重、适用性与结构的重新评估。这一理念促使知识图谱向神经网络化结构演化，即通过引入端到端可训练的神经机制，在节点与边的表示中嵌入可微分的记忆强度、信任度与情感权重，使图谱具备对新输入和多模态刺激的自适应重构能力。与此同时，这种神经网络支撑下的图谱结构，也为表达潜意识提供了更拟人化且灵活的通道。具体而言，系统可将潜意识内容以低显性度的向量权重持续编码于图谱中，借助梯度驱动的方式不断影响显性认知路径的生成与偏好调整。这种机制不仅增强了知识图谱的“认知弹性”与表达精度，也更接近人类以模糊情绪、模

态记忆和非逻辑性联想方式处理信息的心理现实，为构建具备自我成长、自我修正与思维多样性的 AGI 认知结构提供了坚实基础。

3.9 生成模型在内源思维中的深度融合

当前系统中的内源思维机制虽已具备基于高感知显著度记忆的随机组合与路径探索能力，能够在一定程度上实现认知发散与创造性生成，但其生成质量与结构合理性仍受限于启发式策略与图结构引导。未来研究可引入更具生成表达力的深度生成模型，如变分自编码器（VAE）、生成对抗网络（GAN）以及扩散模型等，进一步增强系统在潜在认知路径构建、假设生成与高模糊度知识组合方面的表现。

通过将 VAE 用于压缩记忆表征空间，可实现对内源思维初始激活状态的连续潜变量建模；利用 GAN 提升生成结果的结构合理性与语义连贯性，从而增强系统对新颖概念与场景的构建能力；扩散模型则能在多轮迭代生成中引导认知轨迹逐步逼近高语义密度区，辅助形成更稳定、具启发性的创新性认知图谱。此外，如何将上述生成机制与滑动函数注意力机制协同设计，以实现更具目标导向性和情绪一致性的内源推理路径，也是未来值得深入探索的关键方向。

3.10 内源思维的向量空间生成机制探索

本系统提出的内源思维机制，以知识图谱结构与控制函数为核心，结合滑动注意力机制与情绪驱动过程，实现了任务生成、推理链条重组等复杂认知行为。尽管该结构具备高度灵活性与可解释性，其对显式图谱构建及多重控制参数的依赖，在大规模复杂任务场景中可能导致计算效率下降、结构冗余增加及迁移泛化能力受限。近年来，基于大规模语义向量空间（如词嵌入、多模态上下文编码及语言模型内部状态）的认知建模方法，在生成联想、模糊推理与任务迁移方面表现出优异性能。这启发我们思考：是否可以在无需显式图谱结构的前提下，仅通过向量空间中的动态操作实现内源思维？例如，可探索在语义嵌入空间中构建滑动轨迹、非线性路径扰动及语义共振激活机制，以模拟认知焦点漂移、潜意识联想进发与任务生成驱动。未来工作可尝试构建一种“向量化内源思维模型”，其潜力包括：(1) 简化结构设计，提升计算效率；(2) 增强开放领域中的语义迁移与泛化能力；(3) 建立向量空间与知识图谱的双向映射通道，实现符号-子符号的协同推理；(4) 为神经-符号混合系统提供新的融合接口，从而进一步提升 AGI 系统的认知连续性、生成灵活性及理解深度。

4 结语

本文围绕通用人工智能的发展需求，提出了一种以人类思维本质为启发的新型认知架构，融合滑动函数注意力机制、多维知识图谱嵌入、感知驱动控制系统与内源思维模块，构建具备类人认知流动性、自主性与稳定性的推理体系。该架构不仅从机制层面重构了联想性思维、非线性跳跃、情绪调节与自我监控的动态过程，更在结构设计上强调模块化实现路径与落地可行性，为构建具备持续学习、情境适应与创造性生成能力的 AGI 系统提供了基础支撑。

未来研究将进一步拓展该架构在多模态对齐、长程记忆调控、结构性元反思、以及人格演化机制等方向的表达能力，并探索其在自主任务建构、人机协同交互与真实世界复杂推理场景中

的应用潜力。面向类人智能系统的演进目标，本文所提出的认知结构为具备可解释性、可拓展性与可迁移性的通用智能系统探索提供了新的理论支点与工程方向。

参考文献

- [1] Baddeley, A. D. (2000). The episodic buffer: a new component of working memory?. *Trends in Cognitive Sciences*, 4(11), 417–423.
- [2] Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1–29.
- [3] Anderson, J. R. (1983). *The Architecture of Cognition*. Harvard University Press.
- [4] Anderson, J. R., & Lebiere, C. (1996). ACT-R: A theory of higher level cognition and its relation to visual attention. *Cognitive Systems Research*.
- [5] Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge University Press.
- [6] Minsky, M. (1974). *A Framework for Representing Knowledge*. MIT AI Memo 306.
- [7] Schank, R. C., & Abelson, R. P. (1977). *Scripts, Plans, Goals, and Understanding*. Lawrence Erlbaum Associates.
- [8] Tulving, E. (1972). *Episodic and semantic memory*. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381–403). Academic Press.
- [9] Tulving, E. (1983). *Elements of Episodic Memory*. Oxford University Press.
- [10] Rumelhart, D. E. (1975). *Notes on a schema for stories*. In D. G. Bobrow & A. Collins (Eds.), *Representation and Understanding: Studies in Cognitive Science* (pp. 211–236). Academic Press.
- [11] Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9(1), 77–110.
- [12] Stein, N. L., & Glenn, C. G. (1978). Understanding and remembering stories: A developmental analysis. *Discourse Processes*, 1(2), 283–311.