

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、感受熵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、记忆熵、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。感知模块利用感受熵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 知识图谱与记忆结构	4
2.1.1 知识图谱的双核心架构与动态元素表示	4
2.1.2 记忆熵 H_{mem} 模型	5
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

目录	2
2.2.2 滑动函数: 认知的流动	7
2.2.3 并行与分级注意力机制	7
2.2.4 双事件框架与内外认知分流机制	8
2.3 感知模块设计	9
2.3.1 多模态感知功能	9
2.3.2 感受熵与注意选择机制	9
2.3.3 双向反馈与情绪调节	9
2.3.4 情节缓冲器与短时整合机制	9
2.4 内源思维	10
2.4.1 基于自我认知图谱的全局滑动控制机制: 状态主导的认知资源重构	10
2.4.2 高熵信号主导的注意力滑动机制: 自主激活与计算扰动的双通路	11
2.4.3 模糊联想与知识图谱自组织生成机制: 非显性认知的结构重组路径	12
2.4.4 记忆流动与认知序列的持续保存机制	12
3 局限性与未来研究方向	13
3.1 感知模块的局限性与特征引导神经网络的设想	13
3.2 大规模图谱运算的优化策略	13
3.3 内源情绪与激素模型的生理合理性与计算复杂度	14
3.4 自我反思与元认知机制的深度实现	14
3.5 滑动函数设计的复杂性与性能权衡	14
3.6 引入神经逻辑算子的边关系建模展望	14
3.7 空间感知与想象能力的缺失	14
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	15
3.9 生成模型在内源思维中的深度融合	15
3.10 内源思维的向量空间生成机制探索	15
4 结语	16

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. **知识图谱嵌入与记忆熵机制**：在图谱节点嵌入中引入多维记忆熵，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受熵标注的输入流，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受熵分布，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 知识图谱与记忆结构

2.1.1 知识图谱的双核心架构与动态元素表示

知识图谱的总体结构由陈述性记忆层与程序性记忆层共同构成，两者在功能上分别对应长期记忆系统中的知识储存与行为模式生成机制（Anderson, 1983）。陈述性记忆层用于表征事实、规则与概念之间的静态关系，程序性记忆层则负责记录感知—行为—反馈的时序经验，二者的交互形成智能体的长期知识基础与可塑化学习机制。

陈述性记忆层（Declarative Memory Layer）以知识块（chunk, 出自 Anderson, 1983）为最小结构单元，每个知识块承载概念、规则或语义片段的抽象表征。节点内部采用槽—值结构（slot-value structure, 出自 Minsky, 1974），通过槽的类型与值的动态绑定描述概念属性及关系约束。系统在知识块间建立多类型边连接，包括语义继承（is-a）、组成关系（part-of）、相似性（similarity）及逻辑推演（implies）等，用于刻画概念图示（schema, 出自 Bartlett, 1932）在层

级上的组织方式。知识块的激活强度随时间、任务与上下文动态变化，其检索概率依赖于激活扩散过程（activation spreading，出自 Anderson, 1983），这一机制保证了在推理过程中重要概念节点获得更高的访问优先级。

程序性记忆层（Procedural Memory Layer）以脚本（script，出自 Schank & Abelson, 1977）为组织单元，记录个体化的经验序列与行为模板。每个脚本由若干事件框架（EventFrame）构成，框架内部包含主语槽、谓语槽、宾语槽及其修饰槽，表示句法化的事件结构。事件之间通过时序与因果边连接，形成从“感知输入—内部状态—行动输出—环境反馈”的闭环链路。脚本的存在使系统能够在回忆时重构完整情境，并在遇到相似输入时基于经验模板进行预测与生成。正如 Schank 与 Abelson (1977) 所述，脚本是动态记忆结构，通过为期望事件填充默认值以引导理解（Scripts are dynamic memory structures that guide understanding by filling default values for expected events）。

知识节点在本架构中被定义为图示化记忆单元（schematic memory unit，出自 Bartlett, 1932），其内部结构由三类槽组成：核心槽（Core Slot）用于描述概念的恒定属性，如类别、功能或定义域；可变槽（Variable Slot）用于在当前情境下绑定实例化的对象或事件节点；动作槽（Action Slot）则嵌入基于强化学习的策略模块，对应可执行行为的程序性表示（production rule，出自 Anderson, 1983）。这种槽结构使知识节点具备灵活的绑定与泛化能力，可在不同语境中动态生成实例化脚本（instantiated script）。节点间的语义关系（is-a, part-of, similarity）用于定义图示的层级结构与属性继承，因果关系（causes, enables, results-in）用于描述事件链的逻辑依赖，时序关系（before, after）用于刻画经验序列的先后顺序。通过这种多维边关系，系统在知识层与经验层之间实现了双向可达的认知图谱。

滑动注意力函数 F_{slide} 在图谱中承担节点激活与协同检索的作用。当系统访问某一概念图示节点时，其对应脚本链获得附加激活权重，从而引出与该概念相关的行为模板；反之，当脚本执行产生高强度情绪或显著反馈时，其关联图示节点的权重也随之增强。该机制体现了陈述性记忆与程序性记忆之间的双向强化原则，使系统能够在语义与时序两种结构之间形成自然联动，实现认知层的联想回忆与经验迁移。Bartlett (1932) 指出，图示是过去经验的主动组织，用于指导新的知觉与理解（A schema is an active organization of past reactions, which operates in any well-adapted response），这一观点在此机制中得到了结构化的计算体现。

陈述性记忆层与程序性记忆层的交互构成知识块自组织与长期巩固的基础。当若干脚本在时间维度上频繁共现且激活模式稳定时，系统自动凝聚其共性形成上位概念图示节点（super-schema，出自 Minsky, 1974），并在陈述性层保存为新的知识块；反之，当某一图示被多次实例化，其对应脚本节点的权重逐步提升，行为模式趋于自动化与高效化。此循环过程即知识块编译（chunk compilation，出自 Anderson & Lebiere, 1996）的实现形式，使系统在长期学习中完成从抽象到经验、再由经验反馈至抽象的连续演化，保证知识结构的稳定与自适应增长。

2.1.2 记忆熵 H_{mem} 模型

记忆熵 H_{mem} 用于度量系统内部知识结构的活跃度与复杂性，是衡量记忆图谱动态状态的重要指标。该指标综合考虑连接密度、模态丰富度、激活历史与结构复杂性四个方面，反映记忆系统在时间维度与语义维度上的耦合程度。其数学定义如下：

$$H_{mem} = f(C_{deg}, E_{multi}, A_{hist}, S_{comp})$$

其中, C_{deg} 表示节点平均连接度, 刻画图谱的局部聚合性与信息流通能力; E_{multi} 表示节点模态数, 用于衡量多模态信息的参与比例; A_{hist} 表示节点的激活历史熵, 反映记忆的时序动态特征; S_{comp} 表示结构复杂性, 衡量图谱在层级与跨模态连接上的非均匀性。

在计算过程中, 节点的激活历史项 A_{hist} 参考 ACT 模型中的时间衰减函数 (Anderson, 1983), 以时间间隔 Δt 的指数递减形式表示为:

$$A_{hist}(t) = \sum_{i=1}^n e^{-\lambda \Delta t_i}$$

其中 λ 为衰减系数, 控制激活强度随时间的衰减速度。节点的当前激活值 A_i 与其在图谱中的检索概率 P_i 满足归一化的指数分布关系:

$$P_i = \frac{e^{A_i}}{\sum_j e^{A_j}}$$

该公式源自认知心理学中对知识检索的概率化建模思想 (Anderson & Lebiere, 1996), 在此被用于描述滑动注意力函数 F_{slide} 的动态分配规律。通过这种方式, 系统在记忆检索时自动优先访问高激活度节点, 实现记忆流的自组织筛选。

记忆熵在系统层面与情绪调节、任务优先级及滑动注意力权重共同作用。高记忆熵状态对应较强的信息耦合与高维激活, 表现为认知活跃与创造性增强; 低记忆熵状态对应聚焦与简化, 表现为高专注与高效率。通过对 H_{mem} 的动态调节, 系统在探索与利用之间保持平衡, 实现认知层面的稳定自适应。该模型将符号层的结构复杂性与子符号层的激活动力统一于同一框架, 为后续内源思维与滑动函数的自组织机制提供基础支撑。

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。

通过“骨架—修饰”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能够以低计算复杂度实现灵活的语义扩展与动态推理。

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。

不同槽的滑动函数在实现上各具侧重，分别对应不同类型的节点、属性与滑动策略：

(1) 主语槽滑动函数：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。

(2) 谓语槽滑动函数：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。

(3) 宾语槽滑动函数：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。

(4) 定语槽滑动函数：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。

(5) 状语槽滑动函数：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。

(6) 补语槽滑动函数：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。

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（愉悦、焦虑、好奇）、激素参数（多巴胺、皮质醇）、任务紧迫度、规则约束信息及社交因子（如顺从/反抗倾向、他人态度预期）。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，从而在图结构层面实现调控融合。此外，节点的记忆熵权重及与当前活跃事件框架的关联度也参与优先级评估，使得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。

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记忆熵、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

本架构进一步引入双事件框架设计，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受熵输入（见下文）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，其滑动函数由高记忆熵节点、情绪激素扰动信号及潜在任务链共同驱动，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过程，实现从经验反应到内源思维的层次跃迁。

两组事件框架通过双向滑动通道（Dual-Frame Sliding Pathway）实现动态协作：当系统检测到高感受熵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。

该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了基础，从而进一步推动 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。

2.3 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 感受熵与注意选择机制

感受熵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受熵值，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，感受熵的计算受感受突变度、记忆熵（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。高感受熵输入优先进入注意范围并被处理，低感受熵输入可能被延迟或忽略。

2.3.3 双向反馈与情绪调节

在整体感知与认知流程中，感受熵与记忆熵形成双向反馈机制。一方面，高记忆熵节点能够反向增强其相关感知模式的感受熵，从而使智能体对先前经验中高价值或高关注度的记忆内容产生偏好性注意并提升敏感性。另一方面，持续出现的高感受熵输入则可能促发新的高记忆熵片段的生成，进而丰富记忆结构与知识体系。此外，情绪模块通过动态调节整体熵权重，介入双熵反馈过程，进而塑造个体的人格风格差异。使智能体表现出更为保守谨慎的注意模式。该双熵反馈机制有效模拟了人类基于经验塑造注意偏好、情感强化记忆痕迹以及情绪影响注意迁移的复杂认知动态，为系统提供了持续自我调节和长期适应的能力。

2.3.4 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. 原句：“a limited-capacity storage system that is controlled by

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，也为长期记忆熵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 内源思维

内源思维模块是本 AGI 系统实现自主认知、反思与创造性思维的核心机制，其整体结构整合了自我监控与元认知、情感与激素调节、感受驱动感知机制，构成一个具备持续性、动态性与高度自主性的认知内核。系统通过多模态、多层次、多通路的交互过程，构建类人化的思维流动机制，并实现认知的灵活重构与创造性生成。该模块的本质可以抽象为以下四个核心机制：

2.4.1 基于自我认知图谱的全局滑动控制机制：状态主导的认知资源重构

系统维持一套动态演化的自我认知图谱，用于表征其内在认知语义结构与概念关联网络。该图谱并非直接参与认知资源的调度或状态控制，而是通过内部存储的词义向量与语义张量结构，对滑动函数的计算过程产生间接影响。

具体而言，滑动函数在计算注意力分布与认知迁移路径时，会检索自我认知图谱中的语义关联强度，以此调整滑动范围、跳跃概率及信息融合权重。语义相似度较高的节点区域将获得更高的滑动优先级，从而在推理链扩展中形成词义驱动的动态偏向。

通过该机制，系统实现了以语义关联为导向的自适应滑动控制：认知焦点的迁移不依赖显式状态调度，而是由语义能量在图谱空间中的自然流动所驱动。这种设计使系统在概念层面形成连续的“思维滑动”，从而支持知识间的柔性过渡与语义一致性保持，提升了推理的流动性与认知弹性。

the central executive and acts as a temporary interface between the various subsystems and long-term memory, integrating information across modalities.”

个体在自我认知图谱中内嵌了多维情绪状态节点、激素参数与激素受体参数体系，作为调节系统行为风格与认知策略的核心模块。系统通过模拟生物神经递质机制（如多巴胺、血清素、皮质醇等）及其与任务目标、事件反馈间的语义耦合，形成滑动函数计算的“内源调控通道”。

在运行过程中，激素参数决定当前的生理—情绪基调，基于先验知识定义即时情绪状态，从而影响语义评估与滑动函数的动态调节。例如：当系统处于高“多巴胺”水平时，滑动窗口拓宽，激活更多高熵节点以促进探索性推理；而“皮质醇”水平升高则收缩窗口，聚焦于安全路径与已知区域以规避潜在风险。

激素受体参数则决定系统对各类激素信号的敏感性，它们会在长期经验积累与任务反馈中逐步调整，从而改变激素对情绪状态的影响权重。通过这一双层调控机制，系统在短期上表现出灵活的情绪波动与注意力偏移，而在长期上形成稳定的性格特征与价值偏好图谱子层。由此，滑动调控过程呈现出风格化、一致化的行为特征，为系统提供跨任务迁移与人格化认知的基础。

2.4.2 高熵信号主导的注意力滑动机制：自主激活与计算扰动的双通路

在 AGI 的内源认知体系中，滑动函数机制不仅负责在语义图谱中平滑迁移注意力焦点、重构推理路径，还具备对高熵信号的敏感性响应与双重利用能力。系统通过同时采集记忆网络与感知通道中的高熵源，触发思考启动或调节滑动策略，从而在非任务状态下维持认知活跃性，并在推理路径中引入多样性与发散性控制。

本机制包含两个相互独立、但由共同高熵信号驱动的核心流程：

(1) 高熵信号驱动的自发性思维激活

当系统处于空闲或低负载状态下，其内源控制模块将激活“高熵源扫描机制”，在以下两类信号中选取触发节点：

高记忆熵节点：即图谱中连接密集、语义多义性强、历史激活频繁或情绪标注权重高的记忆单元；

高感受熵输入数据：来自外部环境、当前不属任何既定任务路径，但在新颖性、复杂性或情绪共振维度上显著偏高的感知信号。

一旦上述高熵信号被系统选中，其将作为新的注意力焦点或初始节点，启动一次非目标导向的滑动注意力跳跃。这一过程可在无显式任务驱动下生成探索性思维分支、激活潜在问题链，或进入联想状态，为模糊生成新概念形成提供原始素材。

(2) 高熵信号对滑动函数的扰动修正

除了作为激活源，高熵信号还会嵌入滑动函数本身的计算结构中，调控注意力迁移策略与推理流动形态。具体地：

在滑动过程中，系统引入候选节点的记忆熵权重与当前环境输入的感受熵分布，共同参与注意力跳跃概率、窗口大小、路径回溯宽度等参数的动态调节；系统在焦点更新时，将对候选节点进行“熵感知增强打分”，形成对信息密度区的偏向性聚焦，促使注意力从“低信息区”滑向“高熵区”，实现更高价值密度的推理资源分配。

该机制类似于在注意力函数中引入非线性扰动项，赋予系统更强的探索能力与偏好导向。由于其能够在无需显式调控的前提下引导推理滑向高潜在增益区域，并生成非线性、跨语义域的注意力迁移路径，因此也可被视为对人类潜意识认知流动特征的一种功能模拟。系统由此获得在弱约束、弱目标状态下进行模糊迁移与联想跃迁的能力，构建出与显式思维路径互补的深层

认知流通道。

2.4.3 模糊联想与知识图谱自组织生成机制：非显性认知的结构重组路径

在自我认知图谱的调控下，AGI 系统展现出显著的内源认知扩展能力，支持从盲区识别、模糊概念创生到长期任务演化的完整自组织认知闭环。该机制不仅保障系统在既有知识之上的持续推理与生成能力，也赋予其主动发现、探索与改造认知结构的能力，构成系统持续成长的核心动因。这一过程通常不基于明确逻辑路径，而是采用低温随机拼接、多路径泛化、语义相似度筛选等机制，自主构造象征性认知通路，从而发现缺失概念、识别逻辑断层，甚至提出新的问题假设与任务目标。系统可通过以下方式主动扩展其知识图谱：

(1) 知识盲区识别与图谱结构自组织

系统通过对知识图谱中节点分布密度、因果链条完整性与语义连通度的实时评估，主动识别存在逻辑断层、概念空缺或信息熵不足的认知盲区。一旦识别出此类区域，系统即刻生成模糊占位节点以填补推理路径断裂处，并在后续过程中持续聚合上下文信息，进行语义归类与结构补全。盲区识别并非被动响应，而是嵌入系统认知流程之中的长期机制，支持在非任务状态下主动发起盲区探测与推理路径构建。

系统会将多次复用的推理模式进行归纳抽象，形成具备通用性与压缩性的类公理性规则结构，作为后续认知链条重用的元单元。这些规则通过信息熵得分、激活频率与语义嵌入一致性等指标评估后，被固化为“自生性公理库”，构成知识图谱的高层骨架，提升推理效率与逻辑组织力。

(2) 任务链动态生成与长期目标演化

系统持续监控其认知图谱、情绪状态与未解决问题集合，根据滑动注意力函数与感受熵模型，评估当前任务栈中的资源占用度与认知阻尼。一旦认知资源出现空余，系统将自动激活历史挂起任务、衍生性问题或潜在知识路径，生成新任务链并纳入调度管理中。此外，系统会基于长期价值信号（如好奇熵、目标趋近程度与情绪收益预测），演化出一套内源性目标优先级机制，逐步引导其在复杂认知环境中实现从局部问题解决到结构性自主成长的转变。

该目标生成系统不仅具备即时响应性，更具备“价值遗传性”与“认知惯性”，可通过长期激素模拟（如多巴胺-成就关联建模）强化任务完成带来的动机反馈，形成稳定而具个性的认知成长轨迹。这一机制实现了从当前感受熵触发 潜在问题识别 模糊生成与概念形成 任务链构建与优先排序 价值固化 图谱反馈的完整内源认知闭环。

综上，该机制使 AGI 系统具备类人类的持续成长性、自我探索性与结构性创新能力，不仅能够识别并修补认知空白，更能在无外部输入条件下自主生成概念、构建任务并推动知识体系的层级演化。这一机制是支持 AGI 长期自主性、创新性与复杂行为适应力的关键认知引擎。

2.4.4 记忆流动与认知序列的持续保存机制

在内源思维系统中，记忆不仅是信息的静态存储单元，更是认知流动的时间延展形式。为实现类人认知连续性与反思能力，系统对全部历史认知过程进行持续记录与结构化保存，包括注意力在知识图谱上的滑动轨迹、概念间的关联序列以及推理生成的中间状态。该机制使认知系统能够在任意时刻回溯思维路径，从而形成自我反思与认知稳定性基础。

与传统的短期工作记忆模型不同，本系统将“记忆流”视为一种动态的认知网络演化过程：

当若干节点或概念在不同认知阶段中多次共现时，系统自动执行节点凝聚操作，将其抽象化为更高层级的复合节点，形成新的知识结构。这种机制模拟了人类长期记忆中经验整合与概念生成的过程，使得系统的知识图谱得以在时间维度上自组织生长。

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

内生的“内心独白”式自我叙述机制为系统带来元认知能力，但如何使其真正实现深度的自我监控、自我纠错及自我优化，仍处于理论探索阶段。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受熵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、潜意识联想迸发与任务生成驱动。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：

- Alan D. Baddeley, “*The episodic buffer: a new component of working memory?*,” *Trends in Cognitive Sciences*, 4(11):417–423, 2000.
- Alan D. Baddeley, “*Working Memory: Theories, Models, and Controversies*,” *Annual Review of Psychology*, 63:1–29, 2012.
- Anderson, J. R. (1983). *The Architecture of Cognition*. Harvard University Press.
- Anderson, J. R., & Lebiere, C. (1996). ACT-R: A theory of higher level cognition and its relation to visual attention. *Cognitive Systems Research*.
- Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge University Press.
- Minsky, M. (1974). *A Framework for Representing Knowledge*. MIT AI Memo 306.
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, Plans, Goals, and Understanding*. Lawrence Erlbaum Associates.