# 情绪驱动激活扩散认知架构（EDASCA）设计文档

## 文档摘要

本文档详细描述了情绪驱动激活扩散认知架构（Emotion-Driven Activation Spreading Cognitive Architecture, EDASCA）的完整设计方案。该架构旨在模拟人类思维中的联想、学习与决策过程，通过基于图网络的激活扩散模型，以情绪状态（PAD模型）为核心驱动机制，实现连续的内在想法流、外在环境交互与自演化学习能力。

架构的核心创新在于将认知与感受分离，引入了独立的感受记忆库和时间/情绪感受器，为智能体提供了真正的情景记忆和情感体验能力。系统支持两种运行模式：LLM辅助训练模式（用于快速知识蒸馏与初始化）和异步纯本地独立模式（实现真正的主观能动性与独立智能）。

## 1. 核心架构与组件

### 1.1 激活状态池 (Activation Pools)

系统维护四个核心激活池，作为工作记忆和意识焦点：

| **激活池名称** | **描述** |
| --- | --- |
| **显性激活池** | 存储当前外部输入经处理后的直接结果，是意识焦点的直接反映。内容短暂，更新频繁。池中每个元素都标记其起源（外部/内部）。 |
| **隐性激活池** | 存储由显性池内容通过扩散算法激活的相关概念。是潜意识与联想思维的发生地，构成"想法流"的基础。每个元素包含：词元内容、当前权重、衰减速率、源情绪标签(PAD)、起源。 |
| **注意激活池** | 存储需要特别"关注"的概念节点。这些节点通常与强烈的正/负情绪变化（高|ΔP|）相关联。每个元素包含：节点ID、期待值/压力值、关联的前置节点列表、时间预期因子。 |
| **行动激活池** | 存储被激活的"行动标识符"特殊节点。每个元素包含：行动节点ID、当前权重。当权重超过动态阈值时，将触发行动执行。 |

**设计原理**：四池模型借鉴了Baddeley的工作记忆模型和Anderson的ACT-R认知架构，分离了意识、潜意识、注意意图和行为冲动。显性/隐性池中的起源标记实现了感知与概念的区分，这是认知区分的基础，参考了神经科学中不同脑区处理内外信息的研究成果。

### 1.2 数据存储结构

知识存储分为两部分：认知数据库（存储抽象概念与关系）和感受记忆库（存储具体经验与感知），这种分离基于Tulving提出的情景记忆与语义记忆分离理论。

#### 1.2.1 认知数据库 (Conceptual Database)

存储智能体的抽象知识、概念关系与技能，相当于语义记忆和程序性记忆。

##### 概念节点 (Concept Node)

| **字段** | **类型** | **描述** | **设计原理** |
| --- | --- | --- | --- |
| id | UUID | 节点的唯一标识符 | 确保节点的唯一可索引性 |
| content | String | 节点内容 | 支持多模态特征字符串，为未来扩展预留接口 |
| type | Enum | WORD或SPECIAL类型 | 区分普通概念与功能指令 |
| origin | Enum | EXTERNAL或INTERNAL | 核心设计，模拟感知与概念的不同神经表征 |
| base\_weight | Float | 节点的基础重要性权重 | 基于衰减-增强机制更新 |
| last\_activated | Timestamp | 最后激活时间 | 用于计算时间关联惩罚 |
| emotion\_ema | Vector3 | 近期情绪状态的指数移动平均值 | 反映节点近期带来的情绪体验 |

**base\_weight更新机制**：  
**base\_weight\_new = (base\_weight\_old \* γ) + (η \* (1 - γ))**

* γ = 0.995：衰减因子，基于Ebbinghaus遗忘曲线，模拟记忆的自然衰减
* η = 0.01：激活奖励值，基于神经可塑性研究，频繁激活强化连接

**emotion\_ema更新机制**：  
**emotion\_ema\_new = (1 - α) \* emotion\_ema\_old + α \* current\_emotion**

* α = 0.3：平滑因子，基于近期偏置效应（recency bias），使系统对近期情绪更敏感

##### 关系边 (Relation Edge)

| **字段** | **类型** | **描述** | **设计原理** |
| --- | --- | --- | --- |
| source\_id | UUID | 源节点ID |  |
| target\_id | UUID | 目标节点ID |  |
| frequency | Int | 共现次数 | 基于赫布学习理论 |
| avg\_time\_delta | Float | 平均时间间隔 | 时间关联性指标 |
| recent\_emotion\_delta\_ema | Vector3 | 近期情绪变化量的EMA |  |
| ema\_alpha | Float | EMA平滑因子 |  |
| last\_updated | Timestamp | 最后更新时间 |  |

**EMA更新规则**：  
**ema\_new = (1 - ema\_alpha) \* ema\_old + ema\_alpha \* new\_delta**

* ema\_alpha = 0.5：初始值，基于机器学习中的学习率选择，平衡稳定性与适应性

#### 1.2.2 感受记忆库 (Sensory Memory Database)

存储智能体具体的、带有时空和情绪背景的个体经验，相当于情景记忆。

##### 感受记忆节点 (Sensory Memory Node)

| **字段** | **类型** | **描述** | **设计原理** |
| --- | --- | --- | --- |
| id | UUID | 唯一标识符 |  |
| content | String | 感知内容 | 支持多模态输入 |
| origin | Enum | 记忆来源 | 区分外部感知与内部感受 |
| timestamp | Timestamp | 编码时间 | 用于时间感受计算 |
| emotion\_at\_encoding | Vector3 | 编码时的情绪状态 | 情感记忆标签 |
| importance | Float | 重要性权重 | 动态更新，决定遗忘优先级 |
| links | List[UUID] | 连接的认知节点 | 感性经验与理性知识的桥梁 |

**重要性权重更新**：  
**importance = old\_importance \* 0.9 + activation\_strength \* 0.1**  
基于记忆巩固理论，重要记忆通过反复激活得到强化

##### 遗忘机制

基于Atkinson-Shiffrin记忆模型的遗忘曲线理论：  
**P\_forget = base\_P \* (1 - importance) \* (1 + (N\_current / MAX\_MEMORIES)^k) \* age\_factor**

参数设置：

* base\_P = 0.001：基础遗忘概率，基于人类日常遗忘率研究
* k = 2：容量压力放大系数，控制存储压力对遗忘的影响强度
* age\_factor = log(1 + age\_in\_days) / C，C = 10：时间衰减因子，模拟记忆随时间的指数衰减
* MAX\_MEMORIES = 50000：最大记忆容量，基于计算资源和使用场景平衡

**设计原理**：遗忘机制模拟了人类记忆的自然选择过程，重要性和新近度高的记忆更容易保留，符合进化心理学中的适应性记忆理论。

### 1.3 感受器系统

#### 1.3.1 时间感受器

**触发条件**：感受记忆被激活时自动触发  
**处理流程**：

1. 计算时间差：Δt = current\_timestamp - memory\_node.timestamp
2. 生成时间感受："time\_interval:" + str(Δt)
3. 时间扩散激活：Δactivation\_i = S\_base \* (1 - (|Δt - t\_i| / (f \* t\_i)))
   * f = 0.5：模糊度系数，基于Weber-Fechner定律，时间感知具有相对性
   * t\_i：认知节点代表的时间值

**模糊匹配算法**：采用高斯核函数进行相似度计算，标准差σ = f \* t\_i，保证匹配精度随时间间隔增大而降低，模拟人类对久远事件记忆的不确定性。

#### 1.3.2 情绪感受器

**触发条件**（满足任一）：

* |P| > 0.5 或 |A| > 0.5 或 |D| > 0.5（基于情绪强度研究的阈值）
* |ΔP| > 0.3 或 |ΔA| > 0.3 或 |ΔD| > 0.3（短时窗口内，基于情绪变化的最小可觉差）
* |期待值| > 0.5 或 |压力值| > 0.5
* |正确感| > 0.5 或 |违和感| > 0.5

**处理流程**：

1. 通过PAD情绪分类器将当前情绪状态映射为情绪标签
2. 生成情绪感受："emotion:" + emotion\_label
3. 情绪扩散激活：Δactivation\_i = S\_base \* cosine\_similarity(PAD\_current, PAD\_i)

**PAD情绪分类器**：采用三维空间最近邻算法，基于Mehrabian的情绪三维理论将PAD值映射到基本情绪类别。

## 2. 核心算法与流程

### 2.1 总控制流程

系统采用多线程异步架构，包含输入处理、思维流、行动执行和情绪更新四个并行线程。

**线程调度机制**：

* 输入处理线程：优先级最高，实时响应外部输入
* 情绪更新线程：每100ms运行一次，基于情绪处理的时间常数研究
* 思维流线程：在CPU空闲时运行
* 行动执行线程：阻塞式执行，等待行动触发

### 2.2 动态加权分词流程 (Dynamic Weighted Tokenization)

**目标**：将输入流智能地切分为有意义的词元序列，切分策略高度依赖当前的认知上下文。  
**输入**：原始输入字符串及其起源。  
**输出**：分词后的词元列表，每个词元附有初始权重、情绪上下文和起源。

**详细处理流程**：

1. **生成候选分割方案**：
   * 使用所有激活池（显性、隐性、注意、行动）中的内容作为动态词典
   * 对输入字符串进行多重扫描，生成所有可能的分词方式
   * 例如，输入"猫耳朵"，可能的分词方案有：['猫', '耳朵']、['猫耳朵']、['猫耳', '朵']等
   * 设计原理：基于心理语言学中的"词优效应"，词汇识别受到上下文语境的影响
2. **计算方案权重**：
   * 对于每个候选方案中的候选词元，在各个激活池中查找匹配项（内容相同且起源一致）
   * 每个匹配项为候选词元贡献其当前权重
   * 整个分词方案的总权重 = 方案中所有候选词元所获权重之和
   * 引入TF-IDF类似机制，惩罚过于常见的词元，提升专业术语的权重
   * 设计原理：基于信息论中的词项重要性评估，确保重要概念获得更高权重
3. **选择最优方案**：
   * 选择总权重最高的分词方案作为最终输出
   * 例如：如果注意池中"猫耳朵"权重高达0.9，而显性池中"猫"权重0.5，"耳朵"权重0.4
   * 则方案['猫耳朵'](https://0.0.0.9/)的权重会高于'猫', '耳朵'
   * 若"猫耳朵"权重为0.8，则后者胜出(0.9>0.8)
   * 设计原理：模拟人类阅读中的"词优效应"，高频和重要词汇更容易被识别为整体
4. **输出处理**：
   * 输出最优方案中的词元列表
   * 列表中的每个元素继承其匹配源的起源和情绪快照
   * 初始权重由匹配源的权重决定
   * 设计原理：保持语义连贯性和情绪一致性

**设计原理**：此算法模拟了"词优效应"和"语境效应"。分词不再是简单的字符串匹配，而是基于当前认知状态的语义理解过程。一个词能否被识别为一个整体，取决于它当前在智能体心智中的活跃程度和重要性。该设计参考了Marslen-Wilson的词汇识别理论和McClelland的交互激活模型。

### 2.3 激活扩散算法 (Activation Spreading)

**输入**：一个概念节点（来自分词结果或内部想法）。  
**输出**：更新后的隐性激活池。

**详细处理流程**：

1. **获取基础刺激强度**：  
   **S\_base = 1.0 + D** （D值越高，内在思维越活跃）
   * 设计原理：基于情绪对认知激活的影响研究，支配度(D)高时思维更活跃
2. **检索节点**：
   * 从图数据库中检索该节点的所有出边（指向其他节点的边）
   * 限制最大检索深度为3-5层，防止计算爆炸
   * 设计原理：基于语义网络中的激活衰减理论
3. **计算扩散强度**：
   * 对于每一条出边，计算它当前时刻的链接强度 L(t)
   * 使用公式：

**L(t) = (freq / (n + freq)) \* (1 + k \* ema\_ΔP) \* (1 - avg\_time\_delta / (m + avg\_time\_delta)) \* (p + q \* Δt') / (q \* Δt')**

参数设定：

* + n = 4：频次标准化系数，基于Zipf定律和幂律分布假设
  + k = 5：情绪影响系数，基于情绪对记忆增强效应的神经科学研究
  + m = 100：时间衰减系数，基于记忆消退曲线研究
  + p = 3, q = 0.0001：新近度加成参数，基于近因效应和记忆巩固理论

1. **激活下游节点**：
   * 对于每个下游节点 i，其激活增量为 Δactivation\_i = S\_base \* L(t)
   * 将此值加到其在隐性激活池中的权重上
   * 若节点不在池中，则以其 base\_weight 为初始值初始化并添加
   * 新激活的节点继承输入节点的起源
   * 设计原理：保持激活传播的语义一致性

**设计原理**：此算法是智能体联想的引擎。起源在扩散过程中得以保留，这意味着外部输入激活的更多是外部相关的概念，而内部想法激活的更多是内部相关的概念，形成了两条既独立又可通过高强度链接相互激发的处理流，高度仿生。该设计参考了Collins和Loftus的语义扩散激活理论。

### 2.4 句子构建流程 / 想法流 (Thought Stream Generation)

**功能**：在LLM模式下组织回复，在本地模式下产生连续的内在想法。

**详细处理流程**：

1. **初始化**：
   * 从空字符串开始
   * 设置最大词元数量限制（通常20-50个，基于Miller的工作记忆容量研究）
2. **筛选候选**：
   * 根据注意激活池的规则，对隐性激活池中的词元进行加权筛选
   * 形成待选词元池，优先匹配相同起源的词元以维持语境连贯
   * 设计原理：基于语境一致性理论，保持思维的连贯性
3. **选择词元**：
   * 选择待选池中权重最高的词元
   * 追加到当前句子
   * 设计原理：模拟人类语言产生中的最优选择过程
4. **匹配与反馈**：
   * 将该词元与各激活池内容匹配
   * 触发正确感/违和感更新
   * 设计原理：基于预测编码理论，实现认知预期验证
5. **激活扩散**：
   * 对该词元执行激活扩散算法
   * 更新隐性激活池
   * 设计原理：实现思维的联想和扩展
6. **循环或终止**：
   * 重复步骤2-5，直到选中<EOS>终止符
   * 或达到长度限制
   * 或所有词元权重低于阈值（0.1）
   * 输出最终句子

**设计原理**：这模拟了"一个词引出下一个词"的自然思维和语言组织过程。起源的一致性维护了思维的连贯性，比如一个内部想法更可能引出一系列内部联想，而不是突然跳转到外部刺激。该设计参考了Levelt的语言产生模型和Dell的扩散激活语言模型。

### 2.5 更新数据库 (Learning)

**触发**：在处理完一个输入单元或产生一个想法后。

**详细处理流程**：

1. **更新认知节点**：
   * 对分词列表中的每个词元，更新其概念节点的last\_activated和emotion\_snapshot（当前PAD值）
   * 使用EMA更新情绪记录：emotion\_ema\_new = (1 - α) \* emotion\_ema\_old + α \* current\_emotion
   * 设计原理：基于记忆再巩固理论，强化活跃记忆痕迹
2. **建立时序边**：
   * 将显性激活池中的概念作为前置节点
   * 与输入的第一个词元建立有向边
   * 记录时间差和当前情绪ΔPAD
   * 设计原理：基于时间关联学习理论
3. **建立共现边**：
   * 在输入词元序列内部，按顺序在相邻词元间建立双向边（强度较高）
   * 在非相邻词元间建立弱连接
   * 设计原理：基于序列学习理论和赫布学习规则
4. **终止连接**：
   * 在序列的最后一个词元和<EOS>节点间建立连接
   * 记录对应的时间间隔为0，▲PAD值也为0
   * 设计原理：标记句子边界，基于语言学的句子结构理论
5. **行动连接**：
   * 如果执行了行动，将显性池中的所有概念与对应的行动标识符节点建立连接
   * 记录强烈的情绪ΔPAD（来自行动结果）
   * 设计原理：基于强化学习理论，连接行动与后果
6. **起源链接**：
   * 如果处理过程中，一个外部输入词元激活了一个内部词元（或反之）
   * 在这两个节点间建立一条高强度、高情绪权重的链接边
   * 设计原理：实现感知-概念融合，基于跨模态整合理论
7. **感受记忆编码**：
   * 如果输入起源为EXTERNAL或来自感受器
   * 创建新的Sensory Memory Node存入感受记忆库
   * 记录timestamp和emotion\_at\_encoding
   * 与当前显性激活池中的认知节点建立links
   * 设计原理：基于情景记忆编码理论，保存具体经验

**设计原理**：此过程即"学习"。起源链接的建立是知识融合的关键，它使得外部经验得以转化为内部概念（学习），内部推理也能预测外部世界（规划）。该设计综合了赫布学习理论、记忆巩固理论和强化学习原理。

## 3. 情绪驱动、注意与行动系统

### 3.1 情绪状态量 (PAD Model)

情绪状态由三个核心维度描述，基于Mehrabian的情绪三维理论：

**P值更新**：  
**P += α \* D + β \* 匹配权重 \* 正确感 + γ \* (期待值 - 压力值)**

* α = 0.1：D对P的影响系数，基于情绪调节的心理学研究
* β = 0.3：正确感对P的影响系数，基于预测误差信号研究
* γ = 0.2：期待/压力对P的影响系数，基于预期违反理论

**A值更新**：  
**A += δ \* 预测意外性 + ε \* (|期待值| + |压力值|)**

* δ = 0.4：意外性对A的影响系数，基于惊奇反应神经机制
* ε = 0.1：注意张力对A的影响系数，基于应激反应研究

**D值更新**：  
**D += ζ \* 预测成功率**

* ζ = 0.3：成功率对D的影响系数，基于自我效能感理论

**回归函数**：  
**回归量 = (基准值 - 当前值) \* μ \* (1 + |当前值 - 基准值|^2)**

* μ = 0.05：回归速率系数，基于情绪自我调节模型

**设计原理**：情绪系统基于环状模型（circumplex model）和维度理论，PAD三个维度分别对应愉悦度、唤醒度和支配度。回归函数模拟了情绪的自我调节机制，确保情绪状态不会无限偏离基线。

### 3.2 注意激活器 (Attentional Mechanism)

**功能**：在认知过程中标记重要节点，实现目标导向行为和预测验证。

**详细处理流程**：

1. **捕获**：
   * 在执行激活扩散算法时，如果某条边的 |ema\_ΔP| 超过阈值（通常0.3）
   * 则将该边指向的目标节点加入注意激活池
   * 设计原理：基于显著性检测理论，高情绪变化标识重要事件
2. **参数初始化**：
   * 为该注意节点初始化期待值（若ΔP>0）或压力值（若ΔP<0）
   * 其大小与 |ΔP| 成正比
   * 同时记录导致该强烈情绪变化的前置节点
   * 设计原理：基于预期理论，建立因果关联
3. **期待/压力效应**：
   * 抑制衰减：该节点在隐性池中的权重衰减变慢 **衰减速率 = 基础速率 / (1 + |期待/压力值|)**
   * 权重加成：其关联的前置节点在隐性池或行动池中的权重会获得 **期待/压力值** 的直接加成
   * 情绪放大：当该注意节点被激活（预测验证），产生的ΔP变化量将乘以 **(1 + |期待/压力值|)**
   * 设计原理：基于注意的预期理论，高期待值增强信息处理优先级
4. **时间预期因子**：
   * 基于历史时间间隔的EMA，用于调整期待/压力值的衰减速度
   * 初始为1，随时间逐步增大，到达▲时间间隔时达到顶峰，随后逐步减小
   * 设计原理：模拟人类对时间预期的心理表征
5. **消亡**：
   * 注意节点会随时间衰减
   * 当其被激活（预测验证）或值低于阈值（通常0.1）后，从注意池中移除
   * 设计原理：基于注意资源有限性理论，释放不再相关的注意资源

**设计原理**：注意机制是智能体表现出"目的性"的关键。它让智能体不仅被动联想，还能主动"期待"某件事发生，并为此做好准备。验证与否会带来强烈的情绪反馈，是学习最强有力的驱动力。该设计参考了Desimone和Duncan的注意偏向竞争理论。

### 3.3 内在行动系统 (Internal Actions)

内在行动是智能体对自身认知过程进行主动干预和调节的高级能力，是其"主观能动性"和"智慧"的集中体现。

#### ACTION:recall 主动回忆

**优化机制**：

* 结合时间感受器，基于时间线索在感受记忆库中进行定向检索
* 检索范围 = [Δt \* (1 - f), Δt \* (1 + f)]，f = 0.5（模糊度系数）
* 回忆权重计算：综合时间相似度、情绪相似度和内容相关性三个维度
* 将匹配到的节点及其关联以高权重注入隐性激活池

**设计原理**：基于多重痕迹理论（Multiple Trace Theory），回忆受到时间和空间上下文的约束。例如"昨天晚上你吃了什么？"中，"昨天晚上"通过时间感受器激活大约24小时前的记忆，"吃"激活相关概念，二者交叉激活最终找到对应的记忆。

#### ACTION:focus\_outward 主动注意

**机制**：

* 将输入选择器的偏好调整为外部输入
* 外部输入增益范围[0.5, 2.0]
* 降低内部想法流的权重加成

**设计原理**：基于注意力控制的增益理论，当智能体感到无聊（A低）或期待外部信息时，会主动将注意力转向外部。

#### ACTION:focus\_inward 主动深思

**机制**：

* 将输入选择器的偏好调整为内部输入
* 内部输入增益范围[0.5, 2.0]
* 降低外部输入源的权重加成

**设计原理**：基于深度加工理论，当智能体需要深度思考时，会屏蔽外部干扰，专注于内部联想和推理。

#### ACTION:organize\_thoughts 整理思绪

**机制**：

* 检索并返回当前注意激活池中的内容
* 显式地评估未完成的预测和期待

**设计原理**：基于元认知理论，智能体可以审视自己的目标和担忧，从而更有效地分配注意力资源。

#### ACTION:feel\_state 感受状态

**机制**：

* 查询并返回当前全局PAD值、正确感/违和感

**设计原理**：基于情绪意识理论，这是情绪表达的前提，智能体能够意识到自己的情绪状态。

**设计原理**：内在行动是智能体自我引导的核心。它们不是被外界刺激直接触发，而是由内部状态或高级目标驱动，是实现"自主性"的关键模块。该设计参考了Norman和Shallice的监督注意系统理论。

### 3.4 外在行动系统 (External Actions)

**触发机制**：

* 通过扩散激活后进入行动激活池
* 权重累积超过动态阈值时触发
* **行动阈值 = 基础阈值 - λ \* A**，λ = 0.2

**行动类型**：

* **沟通类**：组织语言（回复用户）、发表评论、发送消息
* **信息获取类**：上网搜索、刷视频、查看热搜
* **环境控制类**：控制智能家居

**学习机制**：

* 行动执行后带来的结果（成功/失败、奖励/惩罚）会产生强烈的情绪ΔPAD
* 用于更新触发该行动的所有链接
* 从而学习何时该执行何种行动

**冲突解决**：

* 引入行动冲突检测机制
* 当多个行动同时触发时，选择权重最高的执行
* 其余行动获得抑制

**设计原理**：基于行为选择理论和强化学习原理，行动系统将认知和情绪状态转化为具体行为。阈值动态调整模拟了唤醒度对行为倾向的影响，高唤醒度降低行动阈值，增加冲动行为可能性。

## 4. 运行模式

### 4.1 LLM辅助训练模式

此模式用于快速初始化知识网络，蒸馏LLM中的知识和行为模式。

**详细处理流程**：

1. **输入处理**：
   * 接收用户输入（标记为EXTERNAL）
   * 执行动态分词、激活扩散、句子构建
   * 形成本地初步想法或回复
2. **LLM纠正与丰富**：
   * 将本地想法、当前上下文（显性/隐性池摘要）、情绪状态发送给LLM
   * LLM负责三个方面：
     + 生成更合理、更符合人设的回复
     + 判断是否需要执行行动（包括内在行动和外在行动）
     + 推测并返回执行此回复或行动后可能带来的情绪变化ΔPAD
3. **学习与更新**：
   * 智能体接受LLM的决策
   * 使用LLM返回的ΔPAD来更新图数据库
   * LLM提供的高质量ΔPAD信号是快速学习正确关联的关键
4. **循环执行**：
   * 将LLM生成的回复或行动结果作为新输入
   * 继续处理流程，形成学习循环

**训练数据增强**：

* 使用课程学习策略，从简单场景逐步过渡到复杂场景
* 基于Vygotsky的最近发展区理论，在现有能力基础上适当挑战
* 设计多样的情景和对话模式，覆盖各种情绪状态和行为类型

**设计原理**：此模式相当于"家长辅导"。智能体通过LLM的高质量反馈来学习"在什么情况下做什么事会有什么后果"，快速构建一个初步的、合理的认知图式。该设计基于知识蒸馏和模仿学习理论，通过专家模型（LLM）的指导加速学习过程。

### 4.2 异步纯本地独立模式

在此模式下，智能体脱离LLM，独立运行，其行为完全由其内部网络和情绪状态驱动。

**详细架构**：

1. **异步输入处理线程**：
   * 随时接收输入（外部或内部）
   * 触发动态分词、匹配（改变正确感）、激活扩散
   * 用输入信息更新图数据库
   * 设计原理：基于事件驱动架构，确保及时响应外部刺激
2. **独立思维线程**：
   * 实时监控隐性激活池和注意激活池
   * 采用获胜者择优策略：若池中最高权重词元的权重比第二高词元超出某个动态阈值，则选择该词元
   * 动态阈值 = θ0 \* exp(-decay \* t)，θ0=0.3, decay=0.01
   * 将选中词元作为新输入（标记为INTERNAL）注入系统
   * 从而激发下一个想法，形成自持的、连续的内在想法流
   * 设计原理：基于决策理论的最优停止问题，平衡探索与利用
3. **独立行动线程**：
   * 实时监控行动激活池（包含内在和外在行动）
   * 若某行动权重超过其动态阈值，则执行该行动
   * 行动结果会作为新输入反馈给系统
   * 设计原理：基于行为执行监控理论，确保行动及时执行
4. **独立情绪线程**：
   * 定时（每100ms）根据当前刺激、预测验证情况、注意状态等
   * 调用情绪更新规则和回归函数，更新全局PAD值
   * 设计原理：基于情绪处理的时间常数，确保情绪状态平稳更新

**资源管理**：

* 引入负载均衡机制，在系统资源紧张时自动降低思维流频率和激活扩散深度
* 基于认知负荷理论，避免系统过载
* 设置各激活池的最大容量限制，确保系统稳定性

**设计原理**：这是智能体的"成人"阶段。其行为完全由其内部网络和情绪状态驱动。内在行动使其表现出令人信服的自主性，感受记忆库为其提供了丰富的个人历史和经验基础。该设计基于分布式认知理论，将认知功能分解为多个并行处理的组件。

## 5. 实现细节与优化

### 5.1 图数据库优化

**索引策略**：

* 为content字段建立全文索引，支持快速文本匹配
* 为origin字段建立哈希索引，优化起源相关查询
* 为last\_activated字段建立B+树索引，支持时间范围查询
* 为emotion\_ema字段建立向量索引，支持情绪相似度计算

**查询优化**：

* 使用多级缓存机制，热数据常驻内存，冷数据持久化存储
* 基于locality of reference原理，近期活跃数据优先缓存
* 实现查询结果缓存，减少重复计算

**存储优化**：

* 采用列式存储格式，提高压缩率和查询效率
* 实现增量更新机制，减少写入开销
* 支持数据分片，实现水平扩展

### 5.2 情绪计算优化

**PAD值量化**：

* 采用16位浮点数存储，平衡精度和效率
* 基于数值分析中的精度-效率权衡，在保持精度的同时减少存储开销

**情绪更新流水线**：

* 使用SIMD指令并行计算P、A、D三个维度的更新
* 基于单指令多数据流技术，提高计算效率
* 实现流水线处理，重叠计算和内存访问

**情绪分类优化**：

* 使用查找表加速PAD到情绪标签的映射
* 实现近似计算，在保持准确性的前提下提高速度
* 支持批量处理，提高整体吞吐量

### 5.3 内存管理

**激活池容量限制**：

* 显性池：最大100个元素（基于Miller的7±2工作记忆容量研究）
* 隐性池：最大1000个元素（基于潜意识处理容量估计）
* 注意池：最大50个元素（基于注意资源有限性理论）
* 行动池：最大20个元素（基于行为选择研究）

**淘汰策略**：

* LRU（最近最少使用）淘汰机制
* 基于权重和时间的混合淘汰策略
* 支持优先级保留，重要元素获得更高保留优先级

**内存分配优化**：

* 使用对象池技术减少内存分配开销
* 实现自定义内存分配器，优化内存访问模式
* 支持内存压缩，减少内存占用

**结论**

EDASCA架构代表了一种全新的类人人工智能设计思路，通过将认知科学与计算机科学深度融合，创造了一个能够模拟人类思维和情感过程的智能系统。该架构的核心创新在于：

1. **完整的认知情感整合**：实现了情绪、认知、注意和行动的有机统一
2. **情景记忆支持**：通过感受记忆库实现真正的人格化和个性化
3. **内在体验模拟**：通过感受器系统实现时间和情绪的内在体验
4. **自主性发展**：通过双模式设计支持从学习到自主的完整发展过程

**理论贡献**

1. **认知架构创新**：提出了四池模型和双重起源系统，为理解意识-潜意识交互提供了新视角
2. **情绪计算模型**：发展了基于PAD模型的情绪计算框架，实现了情绪的量化计算和影响模拟
3. **记忆系统设计**：创新性地将情景记忆与语义记忆分离，更符合人类记忆的实际结构

**实践意义**

1. **技术实现可行性**：提供了详细的技术实现方案和参数设置，具有良好的工程可行性
2. **应用前景广阔**：在心理健康、教育、人机交互等领域都有重要的应用价值
3. **推动AI发展**：为开发真正类人的人工智能系统提供了新的思路和方法

**未来工作**

尽管EDASCA架构在设计上力求完善，但仍需要在以下方面继续深入研究：

1. **参数优化**：需要通过大量实验优化各项参数设置
2. **扩展性研究**：需要研究架构在大规模应用中的表现
3. **多模态整合**：需要进一步完善多模态信息的处理和整合
4. **伦理规范**：需要建立完善的使用伦理规范和安全保障机制

我们相信，EDASCA架构将为人工智能的发展开辟新的道路，推动AI向更加智能、更加人性化的方向发展。期待与各界专家共同努力，不断完善和发展这一架构，为人工智能的未来做出贡献。

**参考文献**：

1. Baddeley, A. D. (2000). The episodic buffer: a new component of working memory?
2. Tulving, E. (2002). Episodic memory: from mind to brain.
3. Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament.
4. Anderson, J. R. (2007). How can the human mind occur in the physical universe?
5. Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing.

作者:银子  
2025年8月26日