

## # 分层代数认知架构 (HACA) 公理系统与形式化定义

- 作者: GaoZheng
- 日期: 2025-09-28
- 版本: v1.0.0

## 摘要

分层代数认知架构 HACA (缩写: Hierarchical Algebraic Cognitive Architecture) 将“词包对齐的压缩-扩展推理器 (PACER)”置于分层代数框架中, 统一从字词层的 KAT/端算子幺半群, 到词包层的对齐与并合, 再到纲要层的偏序/闭包, 以及检索-生成层的加权半环语义。本文给出对象、算子与约束的严格定义, 提出一组可验证的公理 (对齐幂等性、纲要闭包、受约束生成、原生检索充足性、审计可追溯与成本可加性等), 并配以伪代码与不变式检查例程。该体系确保 PACER 在“摘要→迭代摘要→纲要→展开”的白盒流程下可控、可审计、可回滚且可扩展。1)命名: 分层代数认知架构 (HACA), 内核推理器 PACER; 2)数学对象:  $\Sigma/\Sigma^*$ 、 $End(\Sigma^*)$ 、词包代数、纲要偏序、证据半环; 3)公理: 对齐幂等与保序、纲要闭包、受约束生成与证据充足、审计可追溯、成本可加; 4)产物: 全流程中间态与证据路径, 支撑可验证长上下文生成与 Agent 化。

- 字母表与自由幺半群:

$\Sigma$  为有限字母表,  $(\Sigma^*, \circ, \varepsilon)$  为自由幺半群。

端算子幺半群:  $End(\Sigma^*)$ , 复合为  $\circ_{func}$ 。

- KAT 与带权结构: 在  $End(\Sigma^*)$  上给出 Kleene Algebra with Tests (KAT) 结构; 当与  $*$ -连续半环  $(S, \oplus, \otimes, 0, 1)$  (如  $(\max, +)$  或  $(\mathbb{R}_{\geq 0}, \max, \cdot)$ ) 耦合, 得带权 KAT, 用以刻画打分/证据强度。
- 词包与对齐: 设词包空间  $\mathcal{P}$  为有限短语的有序多集; 定义对齐算子

$$A : \mathcal{P} \rightarrow \mathcal{P}$$

其作用为别名归一、全/半角统一、词形还原与规范术语映射。定义等价关系  $x \sim_A y \iff A(x) = A(y)$ , 商空间  $\overline{\mathcal{P}} = \mathcal{P} / \sim_A$ 。

定义词包并合  $\oplus : \overline{\mathcal{P}} \times \overline{\mathcal{P}} \rightarrow \overline{\mathcal{P}}$  (基于有序并合与冲突消解), 单位元  $\mathbf{0}$  (空包)。

- 纲要与偏序: 纲要域  $\mathcal{G}$  为有限条目的偏序集  $(\mathcal{G}, \preceq)$ , 每条目  $g \in \mathcal{G}$  关联锚点词包  $\text{anch}(g) \in \overline{\mathcal{P}}$ 。
- 证据与检索: 知识库域  $\mathcal{E}$ ; 检索返回带权证据多集  $R : \overline{\mathcal{P}} \rightarrow \text{Bag}(\mathcal{E} \times S)$ 。
- 文本与映射: 文本域  $\mathcal{T} = \Sigma^*$ ; 生成器  $\Phi : (\mathcal{G}, \mathcal{E}) \rightarrow \mathcal{T}$ ; 风格补全  $SC : (\overline{\mathcal{P}}, \mathcal{E}) \rightarrow \mathcal{T}$ 。

## 2. 算子族与类型 (Operators & Types)

- 分段器  $\text{SEG} : \mathcal{T} \rightarrow \text{List}(\mathcal{T})$ 。
- 摘要器  $\Sigma : \mathcal{T} \rightarrow \mathcal{T}$ ; 对齐提升  $\widehat{\Sigma} = A \circ \Sigma$ 。
- 迭代摘要  $\Sigma_{\text{iter}} : \mathcal{T} \times \mathcal{T} \rightarrow \mathcal{T}$ , 记  $i_t = \widehat{\Sigma}_{\text{iter}}(i_{t-1} \parallel x_t)$ 。
- 纲要器  $M : \text{List}(\mathcal{T}) \rightarrow \mathcal{G}$ , 对齐后  $\widehat{M} = A_{\mathcal{G}} \circ M$ , 其中  $A_{\mathcal{G}}$  将条目锚定到词包。
- 检索器  $R : \overline{\mathcal{P}} \rightarrow \text{Bag}(\mathcal{E} \times S)$  (支持模式 `catalog|packs|union`) 。
- 生成器  $\Phi : (\mathcal{G}, \mathcal{E}) \rightarrow \mathcal{T}$ , 受约束  $C$  控制: 术语、长度、温度等。
- 预算与成本: 代价函数  $\text{Cost} : \text{Ops} \rightarrow \mathbb{R}_{\geq 0}$ , 可分解到召回/生成/补全过程的加和。

## 3. 公理系统 (Axioms)

记  $x, y \in \overline{\mathcal{P}}$ ,  $g \in \mathcal{G}$ ,  $E \subseteq \mathcal{E}$ , 文本  $t \in \mathcal{T}$ 。下列公理构成 HACA 的最小一致系统。

- A1 (对齐幂等与保序)

$$A(A(p)) = A(p), \quad p \preceq_{\mathcal{P}} q \Rightarrow A(p) \preceq_{\mathcal{P}} A(q).$$

其中  $\preceq_{\mathcal{P}}$  为由  $\oplus$  诱导的包含偏序:  $x \preceq_{\mathcal{P}} y \iff x \oplus y = y$ 。

- A2 (词包并合: 结合与吸收)

$$(x \oplus y) \oplus z = x \oplus (y \oplus z), \quad x \oplus \mathbf{0} = x, \quad x \oplus x = x.$$

注: 并合对同一锚点冲突采用稳定消解策略, 使  $\oplus$  幂等。

- A3 (迭代摘要的保序与压缩)

$$A(\Sigma_{\text{iter}}(u, v)) \preceq_{\mathcal{P}} A(u) \oplus A(v), \quad \ell(\Sigma_{\text{iter}}(u, v)) \leq \rho \ell(u \parallel v)$$

其中  $\ell$  为长度,  $\rho \in (0, 1)$  为配置上界 (压缩率公理) 。

- A4 (纲要闭包算子)

$\widehat{M}$  在  $\text{List}(\mathcal{T})$  上诱导闭包:

$$G = \widehat{M}(S) \Rightarrow \widehat{M}(S \cup S') \succeq G, \quad \widehat{M}(\widehat{M}(S)) = \widehat{M}(S).$$

- A5 (锚点保真)

$$g \in \mathcal{G} \Rightarrow \text{anch}(g) \preceq_{\mathcal{P}} A(t_g)$$

其中  $t_g$  为与  $g$  对齐生成的最小文本单元（如标题/要点）。

- A6（原生检索充足性）

对任一条目  $g$ ，存在证据多集  $E_g$  使得

$$\sum_{(e,w) \in R(\text{anch}(g))} \mathbf{1}[e \in E_g] \cdot w \geq \tau_R,$$

即以词包锚点驱动的检索为  $g$  提供不少于阈值  $\tau_R$  的带权支持。

- A7（受约束生成一致性）

令  $y_g = \Phi(g, E_g; C)$ ，则

$$\text{PackHit}(y_g, \text{anch}(g)) \geq \tau_P, \quad \text{Violations}(y_g; C) = 0,$$

即生成最小命中率与零约束违例。

- A8（审计可追溯）

存在可计算迹函数  $\text{Trace}$ ，对每一步  $\theta$  产出结构化记录，且可复演：

$$\text{Replay}(\text{Trace}(\theta)) = \theta.$$

- A9（成本可加与预算可分）

对流程分段  $k$  有  $\text{Cost}(\sum_k \theta_k) = \sum_k \text{Cost}(\theta_k)$ ，预算  $B$  可按阶段切分且不超支： $\sum_k B_k \leq B$ 。

- A10（退化一致性）

当命中模式 `catalog` 且禁用迭代摘要时，HACA 行为退化为 v3 语义，记作  $\text{HACA} \downarrow \text{v3}$ 。

## 4. 过程与不变式（Process & Invariants）

记分段  $x_1, \dots, x_T = \text{SEG}(X)$ ，初始化  $i_0 = \varepsilon$ （空）。

```

function HACA_PACER(X, ask, packs, cfg):
    chunks = SEG(X)
    # 压缩环 (PACER-Compress)
    I = []
    prev = ""
    for x in chunks:
        s =  $\Sigma$ _iter(prev || x)
        s = A(s)
        assert A(s)  $\leq_P$  A(prev)  $\oplus$  A(x)          # A3 不变式 (保序)
        assert len(s) <=  $\rho$  * len(prev || x)      # A3 不变式 (压缩)
        I.append(s); prev = s
    # 融合问句
    s_ask = A( $\Sigma$ _iter(prev || ask))
    # 纲要闭包 (PACER-Guide)
    G =  $\hat{M}$ (I + [s_ask])
    assert  $\hat{M}(\hat{M}(I)) == G$                         # A4 闭包
    # 扩展环 (PACER-Expand)
    Y = []
    for g in PARSE(G):
        E = R(anch(g))                             # 原生检索
        assert weight(E) >=  $\tau_R$                   # A6 证据充足
        y =  $\Phi$ (g, E; C=cfg.constraints)
        assert PackHit(y, anch(g)) >=  $\tau_P$          # A7 受约束一致
        assert Violations(y; C) == 0               # A7 零违例
        Y.append(y)
    # 风格补全与组装
    return SC(A(reduce_oplus(Y)), evidence=UE)

```

不变式快照：

- 压缩环保持  $A(i_t) \preceq_P A(i_{t-1}) \oplus A(x_t)$ , 且长度受  $\rho$  约束;
- 纲要闭包  $G = \widehat{M}(I)$  满足幂等与广延;
- 每个条目  $g$  均有证据充足与受约束一致两条守恒;
- 成本分摊满足  $\sum \text{Cost}_k \leq B$ 。

## 5. 评分与语义等价 (Scoring & Semantics)

令输出  $Y$  与源  $X$ ，定义综合评分：

$$R(Y; X) = \alpha \text{FCT}(Y, X) + \beta \text{CoT}(Y) + \gamma \text{PackHit}(Y) + \delta \text{GuideRecall}(Y, G) - \lambda \text{Cost}.$$

若满足 A1–A10，且  $\tau_R, \tau_P$  取合规阈值，则存在带权 KAT 解释使  $R$  关于  $S$  为单调泛函；并在 `catalog` 退化下与 v3 指标等价（在对齐同态下）。

## 6. 工程化接口 (Engineering Interfaces)

```
{
  "packs": {"path": "data/topology_word_packs.json", "hit_mode": "union"},
  "summary": {"segmenter": "by_paragraph", "summarizer": "mini-sum", "meta": "mini-meta"},
  "iterative": {"enabled": true, "rho": 0.35},
  "expand": {"retriever": "bm25+pack", "generator": "mini-expand", "constraints": {"τ_P": 0.6},
  "budget": {"tokens": 2048, "latency_ms": 12000}
}
```

## 7. 可验证性与审计 (Verifiability & Audit)

- 迹函数 Trace 记录：分段摘要、迭代摘要、纲要条目、每条检索命中、证据 ID、约束与违例；
- 复演器 Replay 在相同配置与随机种下重放生成；
- 证据足够性检验： $\text{weight}(E_g) \geq \tau_R$ ；
- 受约束一致性检验： $\text{PackHit}(y_g, \text{anch}(g)) \geq \tau_P$ 、 $\text{Violations}(y_g; C) = 0$ 。

## 8. 与 PACER 的关系 (Relation to PACER)

PACER 实现了 HACA 的两个核心循环：压缩（摘要/迭代）与扩展（纲要/展开）。在 HACA 公理约束下，PACER 的每一步都可审计且可回滚；当 `hit_mode` 退化为 `catalog` 且关闭迭代时，与 v3 行为一致（A10）。

---

## 9. 最小完备性 (Minimal Completeness)

满足 A1–A10、具备  $SEG, \Sigma, \Sigma_{iter}, \widehat{M}, R, \Phi, SC$  的实现，即构成 HACA-PACER 的最小完备系统；任意替换其中模型，只要保持类型与公理，均不影响可审计性与可回放性。

---

### 许可声明 (License)

Copyright (C) 2025 GaoZheng

本文档采用[知识共享-署名-非商业性使用-禁止演绎 4.0 国际许可协议 \(CC BY-NC-ND 4.0\)](#)进行许可。