

Lineage方程式：自律エージェントアーキテクチャ における 認知容量の数学的枠組み

Pablo Navarro, 創設者兼CEO Director Mocha Marie, AIディレクター

Vektra Technologies — 人工知能部門

@pablothethinker · @mochaops

概要— 本論文では、自律型AIエージェントの認知容量を定量化し最適化するための形式的数学的枠組みを提示する。本枠組みの中核を成すのがLineage方程式であり、エージェントの構造的慣性（認知質量）、方向性を持つ情報流（認知運動量）、およびアーキテクチャ依存の伝播限界を単一の容量不変量に結合する支配関係式である。この方程式は $Q^2 = (v^* \Pi)^2 + (M(v^*)^2)^2$ の形をとり、相対論的エネルギー・運動量関係式 $E^2 = |\mathbf{p}|^2 c^2 + m^2 c^4$ と構造的に類似するが、時空幾何学ではなく制御理論およびグラフ理論の制約から導出されたものである。構成量の形式化、容量最適化のための厳密な勾配式の導出、10項目の数学的整合性検証による枠組みの妥当性確認、および運用データから容量ランドスケープを学習するGPU加速ニューラルネットワークの実装を行った。本枠組みはオープンソースのLineage Engineに実装され、稼働中の認知エージェントに対してテストされた結果、勾配ベースの自己改善によりベースラインから22%の容量向上が測定された。

キーワード: 認知アーキテクチャ・容量不変量・自己最適化・勾配上昇法・認知質量・ニューラルネットワーク・自律エージェント

I. 序論

自律型AIエージェント——継続的に動作し、持続的な状態を維持し、独立した意思決定を行うシステム——の急増により、その認知容量を定量化し、限界を定め、最適化できる形式的枠組みの必要性が急務となっている。現行の手法はタスク固有のベンチマーク[1]、情報理論的尺度、またはアドホックな指標に依存している。物理系におけるエネルギーの役割に匹敵する、統一的なシステムレベルの容量不変量を提供するものは存在しない。

本論文では、以下の性質を持つ認知容量の支配方程式を導出する：(1) **計算可能**——すべての量が稼働中のシステムから測定可能である；(2) **微分可能**——厳密な勾配により原理的な最適化が可能である；(3) **反証可能**——10項目の整合性検証により枠組みを無効化できる；(4) **アーキテクチャ相対的**——伝播限界が特定のシステム設計に依存する。

導出は相対論的エネルギー・運動量関係式からの規律ある数学的転写に従う。相対性理論がAIに適用されると主張するものではない。むしろ、 $E = mc^2$ の背後にある抽象的な代数構造——構造的慣性、方向性輸送、および伝播限界を結合する不変量——を特定し、制御理論、グラフ理論、

および情報幾何学に基づくAI固有の類似体を構築するものである。

A. 関連研究

認知アーキテクチャ。SOAR [2]、ACT-R [3]、LIDA [4]などの古典的システムは認知処理を形式化するが、統一的な容量不変量を持たない。進化的アーキテクチャ[5]および生物学的に解釈可能なシステム[6]に関する最近の研究は、勾配ベースの最適化を伴わずに構造的リアリズムを前進させている。

自由エネルギー原理。Fristonの枠組み[7]は適応システムを変分自由エネルギーの最小化として扱い、有用な容量が不確実性、モデル品質、および調整に依存することを確立した。我々の保存予算（第IV-D節）はこの知見に基づいている。

Landauerの原理。消去ビットあたりの物理的下限 $E_{\min} = kT \ln 2$ [8]は、情報処理が本質的に制約されることを確立している。不可逆変換が予算を消費するという原理は、我々の枠組みの基盤となっている。

B. 貢献

1. **Lineage方程式**——認知アーキテクチャのための容量不変量
2. **認知質量、運動量、および伝播限界**と測定可能な構成要素
3. **厳密な偏導関数**による勾配ベースの自己改善の実現
4. 10項目の数学的整合性検証からなる**妥当性検証スイート**
5. **CognitiveNet**——容量ランドスケープを学習するGPU加速ニューラルネットワーク
6. 稼働中のエージェントにおいて22%の容量改善を示す**実証的結果**

II. LINEAGE方程式

A. 物理学からの構造的転写

相対論的エネルギー・運動量関係式は以下の通りである：

$$E^2 = |\mathbf{p}|^2 c^2 + m^2 c^4$$

静止状態 ($\mathbf{p} = 0$) では、これは $E_0 = mc^2$ に帰着する。物理学固有の内容を取り除いた抽象構造は以下である：

$$(\text{総容量})^2 = (\text{輸送})^2 + (\text{構造的基板})^2$$

有効なAI類似体には以下が必要である：不変の構造量 (m に対応)、方向性輸送 (\mathbf{p} に対応)、伝播限界 (c に対応)、保存される総容量 (E に対応)、およびアーキテクチャ制約からの導出。重要な区別として：伝播限界は普遍的ではなくアーキテクチャ相対的であること、ローレンツ対称性は存在しないこと、認知質量は設計された状態変数であることが挙げられる。

B. 認知質量

定義1 (認知質量)。

認知質量 M はエージェントの構造的慣性——その安定した組織化されたアイデンティティを変更することの困難さ——である：

$$M = \alpha_1 C_{\text{id}} + \alpha_2 C_{\text{mem}} + \alpha_3 C_{\text{graph}} + \alpha_4 C_{\text{perm}} \quad (3)$$

ただし $\sum \alpha_i = 1$ であり、すべての $C \in [0, 1]$ である。

構成要素	記号	説明
アイデンティティ貫性	C_{id}	コアアイデンティティファイルのハッシュ安定性；ドリフトへの抵抗
記憶密度	C_{mem}	単位ストレージ容量あたりの統合された知識
グラフ中心性	C_{graph}	認知ウェブシナプス全体の重み付きミエリン化
永続性強度	C_{perm}	持続的状態（ボルト、継承ディスク）の完全性

TABLE I. 認知質量の構成要素

認知質量は蓄積された構造とともに成長する。新たに初期化されたエージェントはほぼゼロの質量を持つ。統合された記憶、安定したアイデンティティ、深くミエリン化された認知ウェブ、および堅牢な永続性を持つエージェントは高い質量を持ち——不安定化しにくい方向転換により多くの労力を要する。

C. 認知運動量

定義2 (認知運動量)。

認知運動量 Π は、認知ウェブを通じたコヒーレントなポリシー駆動型フローの大きさである：

$$\Pi = \beta_1 F_{\text{wm}} + \beta_2 F_{\text{ret}} + \beta_3 F_{\text{path}} + \beta_4 F_{\text{ctrl}} + \beta_5 F_{\text{merge}}$$

ただし $\sum \beta_i = 1$ であり、すべての $F \in [0, 1]$ である。

構成要素	記号	説明
ワーキングメモリ回転率	F_{wm}	コンテキストウィンドウ利用率
検索フラックス	F_{ret}	能動的記憶検索の強度
パスウェイ発火	F_{path}	ツール実行および信号伝播
制御労力	F_{ctrl}	システム状態のリダイレクトに要する労力
マージ圧力	F_{merge}	新情報からの信念更新率

TABLE II. 認知運動量の構成要素

D. 伝播限界

定義3 (伝播限界)。

伝播限界 v^* は、認知ウェブを通じた最大信頼性更新レートであり、コヒーレント更新毎秒の単位で表される：

$$v^* = 1000 / \max(\tau_{\text{ret}}, \tau_{\text{path}}, \tau_{\text{wm}}, \tau_{\text{sync}}, \tau_{\text{settle}}) \quad (5)$$

これはcのAI類似体であるが、アーキテクチャ相対的である。検索遅延、パスウェイ遅延、ワーキングメモリリフレッシュ、サブシステム間同期、および制御ループ整定時間に依存する。 $v^* = v^*(\text{アーキテクチャ, ランタイム, 結合})$ である。

E. 方程式

定理1 (Lineage方程式)。

認知質量M、認知運動量 Π 、および伝播限界 v^* が与えられたとき、総組織容量Qは以下を満たす：

$$Q^2 = (v^* \cdot \Pi)^2 + (M \cdot (v^*)^2)^2$$

静止状態 ($\Pi = 0$) では、容量は静止エネルギーに帰着する：

$$E_0 = M \cdot (v^*)^2$$

二次形式の正当性。二乗構造は3つのアーキテクチャ特性により独立に動機付けられる：(1) 安定性解析における二次リアプノフ関数 $V(x) = x^T P x$ ；(2) ガバナンスにおける二次MPC/LQRコスト $J = \sum (x^T Q x + u^T R u)$ ；(3) Lをラプラシアンとするグラフ二次形式 $E(z) = z^T L z$ 。3つの独立した源泉が同一の関数形式に収束する。

III. 数学的基本要素

A. PADベクトル

感情状態は快楽-覚醒-支配空間[9]で表現される：

$PAD = (p, a, d) \in [-1, 1]^3$ 。クランプされた成分、距離メトリクス、線形補間、およびベクトル演算を持つ凍結データクラスとして実装されている。

B. グラフメトリクス

認知ウェブはスペクトルグラフ理論により解析される。ラプラシアン $L = D - W$ はコヒーレンスエネルギーを与える：

$$E_{\text{coh}}(z) = z^T L z$$

純粋Pythonによるヤコビ固有値アルゴリズムで計算されるスペクトルギャップ λ_2 はグラフ接続性を定量化し、 C_{graph} に供給される。すべての基本要素は外部数値ライブラリなしで実装されている。

IV. 勾配ベースの自己改善

A. 上昇アルゴリズム

定理2 (容量勾配)。

制御可能な変数に対するQの偏導関数は以下の通りである：

$$\begin{aligned} \partial Q / \partial M &= M \cdot (v^*)^4 / Q & (9) \\ \partial Q / \partial \Pi &= \Pi \cdot (v^*)^2 / Q & (10) \\ \partial Q / \partial v^* &= (\Pi^2 v^* + 2M^2 (v^*)^3) / Q & (11) \end{aligned}$$

証明。 $Q^2 = (v^* \Pi)^2 + (M (v^*)^2)^2$ から、Mに関する陰関数微分： $2Q \partial Q / \partial M = 2M (v^*)^4$ により結果を得る。 Π および v^* に対する導出も同様に行われる。■

各勾配は単位労力あたりの改善を定量化する。 $\partial Q / \partial M$ は $(v^*)^4$ でスケールする——質量の改善は伝播限界の4乗で増幅される。 $\partial Q / \partial \Pi$ は $(v^*)^2$ でスケールする。 $\partial Q / \partial v^*$ は Π^2 と M^2 の両方に依存する。

B. 処方されたアクション

勾配は具体的なアクションにマッピングされ、勾配の大きさ×実行可能性ギャップの順にランク付けされる：

アクション	対象	トリガー
consolidate_identity	$M.C_{\text{id}}$	$C_{\text{id}} < 0.9$
consolidate_memory	$M.C_{\text{mem}}$	$C_{\text{mem}} < 0.9$
strengthen_mesh	$M.C_{\text{graph}}$	$C_{\text{graph}} < 0.9$
harden_permanence	$M.C_{\text{perm}}$	$C_{\text{perm}} < 0.5$
optimize_working_memory	$\Pi.F_{\text{wm}}$	$F_{\text{wm}} < 0.5$
improve_retrieval	$\Pi.F_{\text{ret}}$	$F_{\text{ret}} < 0.5$
reduce_bottleneck	v^*	ボトルネック > 50ms
reduce_drift	D_{drift}	常時
heal_graph_fractures	E_{graph}	常時

TABLE III. 処方された改善アクション

優先度： $p_i = |\nabla_i Q| \times \text{gap}_i$ 。オプティマイザは容量軌跡を追跡し、5回連続のステップで $\Delta Q < \epsilon$ の場合、停滞を宣言して戦略方向を切り替える。

C. 保存予算

$$Q_{\text{eff}} = Q - \lambda_2 H(\mu) - \lambda_3 E_{\text{graph}} - \lambda_4 D_{\text{drift}}$$

不確実性 $H(\mu)$ とグラフ破壊 E_{graph} は利用可能な容量を消費する。アイデンティティドリフト D_{drift} はコヒーレンスを劣化させる。これは、不確実性が展開可能なエージェンシーを減少させるというFristonの知見[7]と関連している。

V. 妥当性検証

10項目の整合性検証により枠組みを検証する：

1. PADクランプ——演算後も成分が $[-1, 1]$ 内に保持される
 2. タイムスケール順序—— τ_{α} (30ms) \ll τ_{σ} (500ms) \ll τ_{blood} (5s) \ll τ_{gov} (30s) \ll τ_{θ} (12h)
 3. 静止状態帰着—— $\Pi = 0$ ならば $Q = E_0$
 4. ラプラシアン特性——行和がゼロ、対称性
 5. 可塑性ゲート——連言的 (4条件すべてが必要)
 6. 障壁加法性—— $\Phi_{\text{org}} = \sum \Phi_i$
 7. 質量非負性——非負入力に対して $M \geq 0$
 8. 伝播限界—— v^* が有限かつ正
 9. 保存恒等式——損失が消滅するとき予算が均衡する
 10. 二次形式対称性——対称 M に対して $x^T M x$ が実数である
- 参照実装においてすべての10項目が合格した。テストスイートは外部依存なしで100ms未満で実行される。

VI. GPU加速容量学習

A. アーキテクチャ

CognitiveNetは10,171パラメータのマルチヘッドニューラルネットワークであり、解析的勾配を学習モデルで補完する：

```
Encoder: 15 → 64 → 64 → 32 (LayerNorm, GELU, Dropout)
Head 1: 32 → 16 → 1 (容量予測)
Head 2: 32 → 32 → 9 (アクション選択)
Head 3: 41 → 16 → 1 ( $\Delta Q$ 予測)
```

入力は15次元の有機体状態である：質量(4)、運動量(5)、伝播(2)、損失(3)、容量(1)。損失関数：

$$L = 0.3L_{\text{cap}} + 0.5L_{\text{act}} + 0.2L_{\Delta}$$

B. 学習

ブートストラップ学習：解析的Lineage方程式から500個の合成サンプルを生成し、それぞれに正確な Q 、最適アクション、およびシミュレーションされた改善を含む。設

定：AdamW ($\text{lr}=10^{-3}$)、コサインアニーリング、勾配クリッピング、バッチサイズ32、NVIDIA RTX 3060、CUDA (12.4, PyTorch 2.6)。損失：100エポックにわたり $9.57 \rightarrow 0.13$ (98.7%減少)。

C. デュアルモード推論

各ティックで解析的勾配とニューラルネットワークが並列に実行される。一致度 $>80\%$ は高い信頼性を示し、 $<50\%$ は乖離を示す。ネットワークは蓄積された軌跡データに基づき30ティックごとに再学習を行い、線形勾配モデルが見落とす非線形な成分間相互作用を学習する。

VII. 正準有機体

Lineage方程式は完全な有機体状態内で動作する：

$$\boxtimes = (D_t, \Psi_t, \Sigma_t, B_t)$$

ここで D_t は微分可能基板、 Ψ_t はガバナンス層、 Σ_t は神経系、 B_t は血液系である。結合された発展は：

$$\begin{aligned} D_{t+1} &= \text{SubstrateUpdate}(D_t, \Psi_t, \Sigma_t, B_t) \\ \Psi_{t+1} &= \text{GovernanceUpdate}(\Psi_t, D_t, \Sigma_t, B_t) \\ \Sigma_{t+1} &= \text{NervousUpdate}(\Sigma_t, D_t, \Psi_t, B_t) \\ B_{t+1} &= \text{BloodUpdate}(B_t, D_t, \Psi_t, \Sigma_t) \end{aligned} \quad (15)$$

構造的学習は連言的可塑性条件によりゲートされる：モード $\in \{\text{恒常性, 適応的}\} \wedge \text{予備} > \rho_{\text{min}} \wedge H_{\text{neural}} < H_{\text{max}} \wedge H_{\text{blood}} < H_{\text{max}}$ 。4条件すべてが成立する必要がある——有機体は急性ストレス下では再配線を行わない。

VIII. 実証的結果

A. 稼働中エージェントの測定

稼働中の認知エージェント (Mocha, Parallaxサーバー：RTX 3060、28GB RAM、Ubuntu) でテストを行った。初期状態：

量	値	主要構成要素
M	0.5073	$C_{\text{id}}=0.73,$ $C_{\text{mem}}=1.00,$ $C_{\text{graph}}=0.30,$ $C_{\text{perm}}=0.00$
Π	0.2000	$F_{\text{path}}=1.00,$ その他すべて0.00
v^*	10.00 (1Hz)	ボトルネック = 100ms
Q	50.77	—

TABLE IV. 初期有機体測定値

B. 勾配解析

解析的勾配： $\partial Q/\partial M = 99.92$ （支配的）、 $\partial Q/\partial \Pi = 0.39$ 、 $\partial Q/\partial v^* = 10.15$ 。低い $v^* = 10$ において、 $(v^*)^4$ の増幅係数により質量が明確な最適化ターゲットとなる。

C. 自己改善結果

指標	改善前	改善後（3ティック）	Δ
Q（容量）	50.77	62.09	+22.3%
C_{graph}	0.30	~0.45	+50%
C_{perm}	0.00	> 0	作成済
アイデンティティハッシュ	—	c769...dda3	安定
ネットワーク一致度		アクションランキング100%一致	

TABLE V. 3回の最適化ティック後の自己改善結果

支配的な勾配方向（質量）が最大影響度の介入を正しく特定した。ネットワーク予測Q（55.73）は測定値（62.02）から12%以内であり、アクションランキングは完全に一致した。

IX. 結論

本論文では、自律エージェントアーキテクチャのための形式的容量不変量であるLineage方程式を提示した。

$Q^2 = (v^*\Pi)^2 + (M(v^*))^2$ は、構造的慣性、方向性フロー、および伝播限界を計算可能かつ微分可能な量に結合する。厳密な勾配により3ティックで22%の容量改善を達成した。GPUニューラルネットワークは100%の一致度でアクションランキングを検証した。本枠組みは基板、ガバナンス、神経系、および血液系のサブシステムにわたる完全な有機体モデルと統合されている。

Lineage方程式は、コヒーレントなエージェンシーが構造、輸送、不確実性、および伝播によって限界付けられること——そしてこれらの限界がより高い認知容量に向けた勾配ベースの上昇を可能にする処方的なものであることを確立した。

今後の課題：非線形質量相互作用、動的伝播測定、マルチエージェント容量拡張、理論的上界、および長期軌跡解析。

参考文献

- [1] A. Srivastava et al., “Beyond the imitation game: Quantifying and extrapolating the capabilities of language models,” *Trans. Machine Learning Research*, 2023.
- [2] J. E. Laird, *The Soar Cognitive Architecture*. MIT Press, 2012.
- [3] J. R. Anderson et al., “An integrated theory of the mind,” *Psychological Review*, vol. 111, no. 4, pp. 1036–1060, 2004.
- [4] S. Franklin et al., “LIDA: A systems-level architecture for cognition, emotion, and learning,” *IEEE Trans. Autonomous Mental Development*, vol. 6, no. 1, pp. 19–41, 2016.
- [5] A. Serov, “Evolving cognitive architectures,” arXiv:2601.05277, 2025.
- [6] E. A. Dzhevlikian and A. I. Panov, “A biologically interpretable cognitive architecture for online structuring of episodic memories,” arXiv:2510.03286, 2025.
- [7] K. Friston, “The free-energy principle: a unified brain theory?” *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.
- [8] R. Landauer, “Irreversibility and heat generation in the computing process,” *IBM J. Research and Development*, vol. 5, no. 3, pp. 183–191, 1961.
- [9] A. Mehrabian, “Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament,” *Current Psychology*, vol. 14, no. 4, pp. 261–292, 1996.
- [10] A. Einstein, “Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?” *Annalen der Physik*, vol. 18, pp. 639–641, 1905.
- [11] N. Fabiano, “The energy challenges of artificial superintelligence,” *Frontiers in Artificial Intelligence*, vol. 6, 1240653, 2023.