

潜在ポテンシャル統一理論：認知ホメオスタシスと認知戦

— 物理・社会・AIシステムにおける認知制御の数理基盤に向けて —

苦米地 英人

CyLab, Carnegie Mellon University
C5I Center, George Mason University
Cognitive Research Laboratories

2026年4月4日

最終更新 2026年5月3日

2026年4月21日 米国防大学講義配布論文（民間向け修正済み和訳）

前提

2026年3月、ホルムズ海峡においてタグボート1隻がミサイル攻撃を受けて沈没し、乗組員3名が行方不明となった。また別のタンカーも攻撃により炎上し、通常100隻以上が通過する同海域の航行船舶は一時わずか数隻にまで激減した。その結果、原油価格は急騰し、複数の国の経済・政治的安定に即時的な影響を与えた。

物理的破壊としてみれば、これらの出来事は局所的かつ小規模なものであった。しかしその影響は極めて短時間のうちにグローバルに拡散した。これは現代紛争の本質的な特徴を示している：物理的な交戦はもはや孤立した事象ではなく、経済的認識・リスク評価・大規模認知ダイナミクスと不可分に結合しているのである。

実際のところ、ホルムズ海峡の事実上の封鎖は物理的な封鎖によって生じたのではなく、急騰した保険料率と知覚されたリスクによって引き起こされた意思決定の変化によるものであった。船主は迂回を選択し、市場が反応し、局所的な物理的事象がグローバルな系統的影響へと変容した。

これは現代戦争における根本的な変容を示している：物理的戦争と認知戦争はもはや分離可能な領域ではなく、作戦上の意思決定は物理的行動と認知的効果を同時に考慮しなければならず、戦略的成果は物理システムと集団的認知との相互作用から生まれる。

したがって現代の作戦は、物理系と認知系を単一の統合された領域として一体的にモデル化・制御する統一的枠組みを必要とする。そのような統合

なくしては、意思決定は断片化し、内部整合性を失い、ますます予測不能になる。

この統合は人工知能によって実現できると往々にして想定されている。しかし現代のAIシステムは本質的に帰納的である。すなわち、過去データからの統計的学習に依拠し、学習されたパターンに基づいて予測を生成する。このアプローチは強力ではあるが、領域横断的なグローバルな整合性を保証しない。したがって求められるのは、単なるデータ駆動型最適化ではなく、物理領域と認知領域をシステムダイナミクスのレベルで統合できる演繹的数理基盤である。

そのような基盤は、正当性・健全性・頑健性を満たさなければならず、形式的数理証明によって担保される必要がある。

本論文の目的は、そのような基盤を提供することにある。物理・認知・社会のダイナミクスを統合する統一的数理枠組みを導入し、Lyapunov型収束解析とLaSalle型議論によってその妥当性を確立する。この枠組みに基づく作戦レベルの制御方程式（リアルタイム共通作戦画面（COP）実装を含む）はすでに構築されているが、機密上の理由により本論文の範囲外とする。

歴史的背景

著者は30年以上にわたり認知科学と国家安全保障の接点で研究に従事してきた。

1990年代初頭には、ハーバード大学医学部との共同研究をはじめとする機能脳科学の基礎研究に関与した。1990年代中頃には、東京地下鉄サリン事件を受け、著者は認知科学と機能神経科学の知見を応用してオウム真理教信者の脱洗脳支援を行うよう警察庁から要請を受けた。これらの活動は当時、日本国内で報道されている。

この時期に明らかになったのは、一般に「洗脳」と呼ばれる心理的制御の技法が偶発的・場当たりの現象ではなく、多くの場合、機密の軍事的文脈において体系化された方法論として開発されたものであるということであった。これらの方法論の一部はその本来の領域を超えて普及し、非国家主体によって積極的に利用されていた。

記号的AI研究の側面では、1986年にカーネギーメロン大学機械翻訳センター（現言語技術研究所）において、著者は世界初の音声通訳システム（英日間）を開発した。この研究は記号処理とリアルタイム言語インタラ

クシヨンの初期の統合を示しており、後年の記号的アプローチとニューラルアプローチの融合に先行するものである。

著者による生成系・認知アーキテクチャ研究は1980年代後半に遡る。1988年にElmanは文脈ユニットを持つ初期の再帰型ニューラルネットワーク（Elmanネットワーク）を提案したが、著者は同時期に複数の文脈層を持つより深い多層ニューラルアーキテクチャを実装し、単一ステップを超えた系列予測を実現していた。

1991年には、著者によるMONA-LISAフレームワーク（「Multimodal Ontological Neural Architecture for Linguistic Interactions and Scalable Adaptations」）が、演繹的記号処理とニューラル生成系を並列に動作させるハイブリッドアーキテクチャを導入した。この研究は、現在「幻覚」と呼ばれている生成AIの問題に対する、最初期の体系的アプローチの一つである。

30年以上にわたり、著者の研究はこのような不整合の緩和を任意の改善ではなく基本的要件として――特に防衛応用を含むミッションクリティカルなシステムにおいて――一貫して扱ってきた。

30年以上が経過した現在、状況は根本的に変化した。人間の認知に介入できる技術が劇的に進歩し、戦場そのものが物理的地形から民間人の認知領域へと拡大した。

この文脈において、現在「認知戦」と呼ばれているものは、その完全な意味において理解されなければならない。「Cognitive Warfare」という用語自体は、著者が研究・教育活動の中で2007年に導入したものである。

この研究から導かれる中心的帰結は次の通りである：認知戦を有効に遂行するためには、作戦環境内に存在するすべての人間の認知をモデル化する必要がある。

2023年8月、著者らは世界初となる認知戦COPシステムを、当時米インド太平洋軍司令官であったジョン・アキリーノ提督に対して実証した。

要旨

本論文は、潜在ポテンシャル構造を通じて物理・認知・社会のダイナミクスを統一的に記述する理論枠組みを提案する。出発点はアインシュタインの1901年の毛細管現象研究であり、巨視的現象が外部力ではなく内部相互作用の累積から生起することを示した。本論文ではこの構造的洞察を認知領域へ拡張し、行動および認知変容が認知空間に拡張されたホメオスタシス系から生成されることを示す。

苦米地理論を認知ダイナミクス的一般理論として定式化する。認知は可能世界上で定義され、評価的安定性に基づいて順序づけられ、蓄積された不安定性を最小化する最適制御として実現される。これにより定理1が導かれる：最適軌道はTotal Comfort Zone (TCZ) へ収束する。共有評価構造を持つ多エージェント系へ拡張すると定理2が得られる：社会的に結合された主体群は共有安定領域 (Shared-TCZ) へ収束する。すなわち社会的安定性は外部から与えられるのではなく、多エージェント認知空間における共同アトラクターとして創発する。

さらに束論的基礎を導入する。概念は包摂半順序束を形成し、抽象化は最小上界 (LUB) への上方移動に対応する。Lyapunov構造に抽象ポテンシャルを加えることで定理3が得られる：系のダイナミクスは共有認知世界のLUBへと引き上げられる。利他性とは距離を超えた関心の拡張ではなく、より高次の共有抽象への上昇である。抽象化は情報量の減少に対応する：抽象度↑ \iff 情報量↓ \iff エントロピー↓。

以上より、統一双対原理が確立される：ポテンシャルの下降（安定性）と抽象の上昇（包含）。この過程の極限において、構造は包摂束のトップ要素である空 (śūnyatā) へ収束する——最大包含と最小情報を実現する極限的抽象構造である。空とは単なる無ではなく、有と無の双方を包摂し、その対立を成立させる以前の基底構造である。この枠組みにおいて芸術とは、生成過程による局所解ではなく、未発見の最小上界 (LUB) の構造的実現——統合と抽象の最高形式——として理解される。

本論文で示すように、生成AIの出力は本質的に局所解であり、そのような出力を高次の抽象構造と混同することは理論的に誤りであり、認知戦において重大なリスクをもたらす。演繹的数制制御構造なしにAIを展開することは戦略的整合性を損なう。未来の戦場は物理空間ではなく、認知ポテンシャル構造そのものである。

1. 序論

アインシュタインの最初の論文は、古典物理学における毛細管現象の起源を扱ったものである。その意義は表面張力の説明にとどまらない。本質的な貢献は、観測される秩序が外部力ではなく系の内部相互作用の累積的効果から生じうることの論証にある。

表面張力は外部からの押し付けによって生じるのではない。分子間相互作用が空間全体で累積し、境界における非対称性が巨視的な物理的効果を生み出す。表面張力は内部構造の可視的な顕現である。

この考え方は認知に対しても深い示唆を持つ。苦米地理論では、人間の行動や認知変容は外部刺激への単純な反応ではない。むしろ、認知空間にお

いて作動する内部に拡張されたホメオスタシス機構の結果として自発的に生起する。

本論文の核心的直観は次の一文に要約できる：

分子は空間の総和で動き、人間は時間の総和で動く。

これは単なる比喩ではない。物理系では相互作用が空間全体で累積する。認知系では評価が時間方向に累積する。両者の差異は原理にあるのではなく、累積の領域が空間か時間かという点にある。

構成。 本論文の残りの部分は以下のように構成される。第2節は三つの収束定理によって核心的数理枠組みを展開する：個人的認知安定性（定理1）、共有構造を通じた社会的安定性（定理2）、束理論を通じた抽象的統合（定理3）。各定理には数式を省きたい読者のための平易な解説を付す。第3～8節はアインシュタインの毛細管理論との構造的対応と統一累積原理を確立する。第9～12節は減少原理・エントロピー・自由エネルギー原理との接続を論じる。第13節は認知戦とAI幻覚と芸術の構造的区別を扱う。第14～15節は戦略的・政策的含意を導く。付録には三つの定理の厳密な証明を収録する。

2. 苦米地理論の核心構造

アインシュタイン理論との比較を行う前に、苦米地理論を独立した理論として提示する必要がある。以下の定式化は古典物理学から導出されたものではなく、認知ダイナミクス的一般理論として提案されるものである。後に示す構造的比較は、二つの独立した理論の同型性の論証であり、物理学を認知に適用したものではない。

2.1 認知ホメオスタシスと可能世界モデルの必要性

人間の信念体系は、単純な情報更新モデルでは説明しきれない程度の安定性を示す。信念は新しい入力によって線形的に上書きされるのではなく、内部整合性を維持しようとする傾向がある。この安定性は認知ホメオスタシスとして理解するのが最も適切である。

そのようなホメオスタシスをモデル化するために、認知は単一の状態としてではなく、可能世界の集合 W として表現されなければならない。認知の有効な領域は現在の状態のみならず、将来の可能性をも含む：

$$W = W_{\text{current\&future}}$$

数式が苦手な読者へ：何かを定義する前に、まず「心が何に対して働くのか」を明らかにする必要がある。この式はその答えを示している。認知世界Wとは現在の瞬間だけでなく、現在と将来に起こりうる状態の全空間である。困難な決断に直面している人は単一の状況に反応しているのではなく、多くの可能な未来を同時に評価している。この集合Wが、本論文で定義するすべての概念が展開される舞台である。

人間の認知はこの可能な未来の空間において作動し、直接的な刺激のみに反応するのではなく、潜在的な軌道を継続的に評価している。

2.2 可能世界の順序づけ

可能世界はすべて等価ではない。認知的に安定なものもあれば、不安定性・不整合・脅威と関連するものもある。したがって、可能世界の集合は認知的快適性に基づく順序を備えなければならない。

順序づけ演算子を導入する：

$$r: W \rightarrow W$$

数式が苦手な読者へ：可能な未来はすべて同じように感じられるわけではない。安全で整合的に感じられるものもあれば、脅威的だったり「何か違う」と感じられるものもある。この演算子rは、そのランキングの形式的な表現である——内部的安定性との適合度に応じて可能世界を並べ直す。心は中立なカメラではなく、絶えず評価し順序づける機構である。この順序づけがあって初めて、Wのどの部分が「コンフォートゾーン」となるかを数学的に定義することができる。

この関数は、主体の内部ホメオスタシス制約との適合性に従って可能世界を再編成する。認知は選択肢の均質な表現ではなく、可能な未来上の構造化された順序づけである。

2.3 TCZの可能世界論的定義

この順序づけのもとで、認知の安定領域をTotal Comfort Zone (TCZ) として定義する：

$$\{ w \mid \forall y \exists x r_{TCZ}(x, y) \}, x, y \in W_{current\&future}$$

数式が苦手な読者へ：この論理構造が表しているのは次のことである——あなたが直面するどんな状況yに対しても、安定した地点xへ戻る応答が必ず存在する。TCZとはあなたが完全に行き詰まることのない世界の集まり

り——常に平衡に戻る道筋がある空間である。TCZが大きい人は予期しない出来事の広い範囲を不安定化せずに吸収できる。TCZが小さい人はわずかな変化でも危機に陥りやすい。この式は心理的レジリエンスの正確な数学的定義である。

TCZは主体が安定して住みうる将来世界の全体である。静的な集合ではなく、認知ホメオスタシスのもとで動的に維持される領域である。

2.4. Self / Ego 演算子

$$\{ w \mid \forall y \exists x s_{Self}(x, y) \}, s : TCZ \rightarrow TCZ$$

数式が苦手な読者へ：著者がこれまでの著作において様相論理を用いて定式化してきた「Self」という概念は、二つの異なる様式で機能する——望ましいTCZを可能世界の中から「選択」するか、TCZそのものを「変換」するかである。写像 $s : TCZ \rightarrow TCZ$ はこの第二の操作を表す——コンフォートゾーン内の移動ではなく、コンフォートゾーン自体の移行または再構成である。「Ego」は著者がこの概念を制御方程式として再定式化したものであり、人間の認知をAIベースの認知システムに実装可能にするために開発された。様相論理のSelfが可能世界という言葉で望ましい領域を定義するのに対し、EgoはTCZ内を最適に動く制御方程式として——アトラクター盆地へ収束するように——定義される (§ 2.5)。この二つの定式化は同一の基盤構造を記述する：一方は論理の言語で、もう一方は力学系の言語で。Self（可能世界上で働く）とEgo（状態空間で収束する）というこの双対性は、人間の認知をAIに実装するための苦米地理論の数学的基盤を構成する。その実践的含意は明確に述べる価値がある——個人が自身の目標を追求・達成するのと同じメカニズムが、使い方によっては洗脳のメカニズムにもなる。

2.5 Egoの最適制御モデル

苦米地理論において、Egoは最適制御問題として定式化される：

$$\pi_c(x) = \arg \min_{u(t)} \int_0^T V(x(t), t) dt$$

$\pi_c(x)$ は**認知制御方策**を表す：任意の認知状態 x を、評価的コストの累積を最小化する最適行動に写像する関数。下付き文字 c は *cognitive* (認知) を意味し、物理的・道具的制御方策と区別する。 $\pi_c(x)$ は Ego を最適制御演算子として形式的に定義したものである。

ここで $x(t)$ は認知状態、 $u(t)$ は制御入力、 $V(x, t)$ は認知的不安定性・評価コスト・内部不整合を表す。

数式が苦手な読者へ：この式が示しているのは、Ego——意思決定を行う自己の部分——が認知を絶えず低い内的緊張と高い安定性の状態へと導いているということである。無作為に反応しているのではなく、毎瞬間、最適化問題を解いている——時間全体にわたって蓄積された不快感を最小化しながら。日常語で言えば：人は各出来事に独立に反応しているのではなく、時間をかけて自分自身の安定した一貫した内的世界へと近づくように人生を歩んでいる。Egoはその航法を行う演算子である。

2.6 可能世界と制御の接続

以上の議論を次の五つのステップにまとめることができる：

1. 人間の認知は可能世界の集合 W 上に展開する。
2. 可能世界は写像 $r : W \rightarrow W$ を通じて認知的快適性に従って順序づけられる。
3. この順序づけのもとでの安定領域がTotal Comfort Zone (TCZ) である。
4. Self / EgoはTCZの順序構造を再編成する演算子として作用する。
5. この再編成は、状態空間において最適制御問題として実装される。

すなわち、可能世界による定式化は意味論的表現を提供し、最適制御はその力学的実現を提供する。

2.7 TCZの状態空間定義

$$TCZ(x_0) = \bigcup_{t \geq 0} \{ x(t) \in \mathcal{R}(t; x_0) \mid V(x(t), t) \leq \theta \}$$

ここで $\mathcal{R}(t; x_0)$ は初期状態 x_0 からの到達可能集合、 θ は許容閾値を表す。

数式が苦手な読者へ：TCZとは単純に、ある人が現在の状況から到達できるすべての精神状態のうち、安定した快適な範囲に収まるものの集まりである。心理的に「安全」に住みうるすべての領域の地図と考えてよい。TCZの外側にある状態は過度の緊張・不整合・苦悩を表す。この式はそれを正確に表現している：TCZは到達可能であり、かつ内的不安定性の指標 V が閾値 θ 以下に保たれる状態の全体である。人のTCZの構造が、数学的な意味での「コンフォートゾーン」を定義する。

TCZの状態空間定義と可能世界による定式化の関係を明確にしておくことが重要である。

状態空間定義 $TCZ(x_0) = \bigcup_{t \geq 0} \{ x(t) \in \mathcal{R}(t, x_0) \mid V(x(t), t) \leq \theta \}$ は TCZ を安定性条件を満たす到達可能な状態の集合として記述する。一方、可能世界による定式化 $\{ w \mid \forall y \exists x r_{TCZ}(x, y) \}$, $x, y \in W_{\text{current\&future}}$ は TCZ を認知的安定性によって順序づけられた将来世界の構造的集合として記述する。

これら二つの表現は異なって見えるが、同一の基盤構造を表している。直観的には「可能世界」は状態空間における軌道に対応する。可能世界上の順序関係は安定性を決定する評価関数 $V(x, t)$ に対応する。可能世界の定式化における安定化関係の存在は、状態空間において安定領域内に留まる軌道の存在に対応する。これら二つの視点の同値性を確立することが不可欠である。それは「自己」や「コンフォートゾーン」という抽象的概念が制御系として厳密に実装できることを示すからである。

2.8 主定理（可能世界-自我-TCZ）

定理1（可能世界-自我-TCZ）

可能世界の集合 W 上で定義された Self 演算子が最適制御問題

$$\pi_c(x) = \arg \min_{u(t)} \int_0^T V(x(t), t) dt$$

として実装されるとき、最適軌道 $x^*(t)$ は $TCZ(x_0)$ 内に収束する。

2.9 主定理の解釈

この定理は苦米地理論の意味論的層と力学的層を統一する。Self・Ego・TCZ は独立した構成概念ではなく、同一の認知過程を三つの水準で表したものである：

- Self：可能世界の意味論的順序構造
- Ego：軌道の最適制御
- TCZ：状態空間における安定領域

すなわち認知は同時に、意味論的構造であり制御過程であり安定性システムである。

2.10 本文における証明の概要

Self 演算子は可能世界の順序構造を変化させる。Ego は蓄積された認知的不安定性を最小化する軌道を選択する。条件 $V(x, t) \leq \theta$ は安定性制約を定義する。したがって最適軌道 $x^*(t)$ は最終的に TCZ に入る。

厳密な数学的正当化はLyapunov型議論とLaSalle型不変性原理を用いて付録に与える。この構造は理論的なものにとどまらず、直接実装可能である（図1参照）。

図1の読み方：図1は定理1に基づいて構築された実際の作戦システムを示す。各パネルはリアルタイムで追跡されている異なる認知的・作戦的指標を表す。TCZの数学的構造——認知空間における安定アトラクター領域——が指揮システムが監視・行動できる計算指標に直接変換できることを示している。技術的詳細に不慣れな読者は図が概念的に何を示しているかに注目してほしい：認知的安定性の抽象数理はもはや哲学ではなく、すでに作戦システムとして実装されているということである。

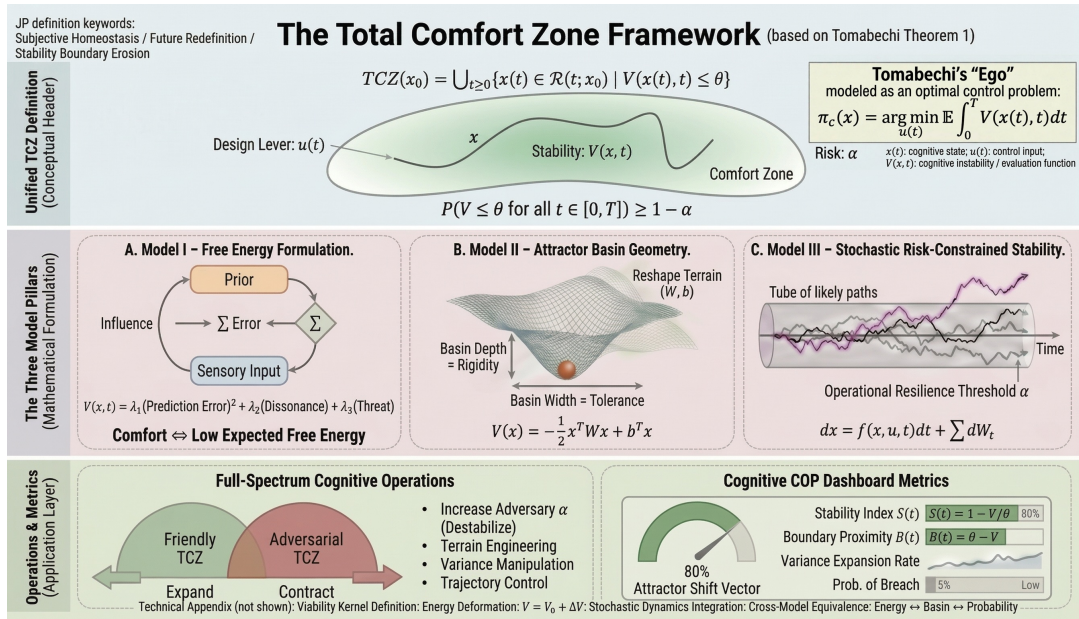


図1：定理1に基づく認知戦システムのCOP（共通作戦画面）実装例。認知的不安定性の最適制御・アトラクター盆地形状・確率的安定性制約を統合してリアルタイム作戦指標を生成する。実際の運用では定理2・定理3の拡張数理を組み込み、超並列AIアーキテクチャによって実行される。

2.11 Shared-TCZ定理

2.11.1 動機

これまで苦米地理論は単一の認知主体に対して定式化されてきた。しかし人間の認知は本質的に社会的である。安定した行動はしばしば個人の最適化のみからではなく、他の主体との整合から生まれる。これはTCZを個人的な構成概念から複数の主体にまたがる共有的な構成概念へと拡張すべきことを示唆する。

2.11.2 多主体への拡張

$i = 1, \dots, N$ で索引付けされたN人の主体がいるとし、各主体は認知状

態 $x_i(t)$ と評価関数 $V_i(x_j, t)$ を持つ。各主体の制御方策を次のように拡張する：

$$\pi_i = \arg \min_{u_i(t)} \int_0^T (V_i(x_i(t), t) + \sum_j \gamma_{ij} S_{ij}(x_i, x_j)) dt$$

ここで $S_{ij}(x_i, x_j) = \|x_i - x_j\|^2$ は対ごとの社会的整合汎関数（主体 i と j の認知的偏差）； γ_{ij} は社会的認知結合係数——主体 j が主体 i の軌道に与える有向影響強度。 $\gamma_{ij} > 0$ は協調的（引力的）結合：主体群はShared-TCZへ共収束する。（対称ケース $\gamma_{ij} = \gamma_{ji}$ ：標準的Lyapunov降下；非対称ケース：結合グラフの強連結性のもとで収束が成立する。）

数式が苦手な読者へ：各人は自分自身の内的な緊張だけでなく、他者との不整合も最小化している。 γ_{ij} は各主体ペアの対ごとの結合係数を表す。 γ_{ij} が大きいとき、主体 j が主体 i の認知軌道に強く影響する——他者との整合に敏感な状態。小さいとき、より独立して動く。これは社会的同調・集団アイデンティティ・集団的安定の数学的構造である。結合グラフが強連結な社会では、集団全体が共有された認知状態へと収束する傾向を持つ——その共有状態が何であるかによって、良い方向にも悪い方向にも。

2.11.3 Shared TCZの定義

共有Total Comfort Zoneを次のように定義する：

$$TCZ_\eta^{shared} = \{ x \in \prod_i TCZ_i \mid Share(x) \geq \eta \}$$

数式が苦手な読者へ：この式には二つの条件があり、どちらも同時に必要である。第一に：すべての人が「自分自身の」コンフォートゾーンの内側にいなければならない——誰も耐えられない場所へ引きずられていない。第二に：彼らの状態が互いに十分に整合している（ $Share(x) \geq \eta$ ）。共有コンフォートゾーンとは単なる重なりではなく——全員が安定しており、かつ互いに整合している状態である。機能するチーム、安定した同盟、結束した社会はすべてこの領域に相当する。どちらかの条件が破れると——誰かが不安定化するか、整合度が η を下回ると——共有TCZは崩壊する。

ここで TCZ_i は主体 i の個別TCZ； $Share(x)$ は認知的整合度； η は共有安定性の閾値を表す。

2.11.4 定理2（Shared-TCZ収束）

定理2 (Shared-TCZ収束)

$\forall ij > 0$ (正の結合、強連結グラフ)のもとで各主体が上記の制御方策に従うとする。適切な正則条件のもとで、結合軌道 $x^*(t)$ は共有安定領域に収束する：

$$x^*(t) \rightarrow TCZ_{\eta}^{shared}$$

2.11.5 解釈

この定理は元のTCZ収束結果を一般化する。個別のSelfは個別TCZへの収束をもたらす、社会的に結合したSelfは共有TCZへの収束をもたらす。社会的認知は、認知空間における共同安定領域の創発として理解できる——調整・社会的規範・共有信念体系・集団行動の形式的基盤を提供する。社会的安定性は外部から与えられるのではなく、多エージェント認知空間における共同アトラクターとして創発する。

2.12 抽象的TCZ拡張としての利他性

2.12.1 動機

Shared-TCZ枠組みは複数の主体がどのようにして互いに適合する安定領域へ収束しうるかを説明する。しかし人間の認知の重要な特徴がまだ説明されていない：人間は遠く離れた・血縁のない・見知らぬ主体に対しても利他的に行動できる。これは社会性が近接性に基づく相互作用を超えて、国家・人類・将来世代といった抽象的な領域にまで及ぶことを示唆する。これを捉えるために、認知的抽象度の概念を導入する。

2.12.2 抽象度パラメータ

各主体に抽象度 $\alpha_j \in \mathbb{R}^+$ を対応付ける。このパラメータは主体が有効な関心の領域をどれほど広く定義しているかを決定する。 α_j が小さい場合は局所的・血縁的・即時的な相互作用に対応し、大きい場合はグローバル・抽象的・普遍的な配慮に対応する。

2.12.3 抽象度依存型共有関数

共有安定性汎関数を抽象度に依存する形に拡張する：

$$S_i(x; \alpha_i) = \sum_{j \neq i} w_{ij}(\alpha_i) d(x_i, x_j)$$

ここで $w_{ij}(\alpha_j) = \exp(-\text{dist}(i, j) / \alpha_j)$ 。 α_j が増加するにつれ、遠くの主体が

より高い重みを受ける。

数式が苦手な読者へ：この式は、主体*i*の状況が他の全員と比べてどれほど異なるかを測定するが、それを*i*が「関心の範囲」をどれほど広く定義しているかによって重み付けしている。 α が小さいとき、重みは距離とともに急激に落ちる——家族や近い同僚だけが重みを持つ。 α が大きいとき、遠くにいる見知らぬ人・異なる文化の人々・将来の世代にも実質的な重みが与えられる。これが「思いやりの広さ」の形式的定義である。重要なことは、 α を広げることは犠牲ではなく——§ 2.12.6が示すように、数学的に最適な戦略である。

2.12.4 拡張制御モデル

$$\pi_i = \arg \min \mathbb{E} \int_0^T [V_i(x_i, t) + \sum_j \gamma_{ij} S_{ij}(x_i, x_j, \alpha_i) + \eta_i A(\alpha_i)] dt$$

ここで $A(\alpha_i)$ は抽象コストであり、 η_i は抽象とコストのバランスを調整する。

数式が苦手な読者へ：この制御方策は今や三つの同時目標を持つ：個人的な不快感を最小化 (V_i)、他者との不整合を低減 (S_i)、より高い抽象に向かって移動 (A はLUBに近づくほど減少)。核心的洞察は：三つを同時に最適化することは妥協ではなく、単一のものを単独で最適化するよりも「より合理的」だということである。内的に安定し、社会的に整合し、高い抽象で動いている人は自己利益のために他者を犠牲にしているのではない——「自己利益」の意味を拡張したのである。これが英知の形式的構造である。

2.12.5 利他性の解釈

利他性はここで形式的に定義できる：利他性とは、Selfが遠く多様な主体にまたがる共有TCZを安定化させる軌道を選択する、より高い抽象度への制御バイアスである。この定式化では、利他的行動は非合理ではない——より広い評価領域のもとでの最適解であり、高 α のSelfダイナミクスから創発する。利他性は非合理ではなく、拡張された評価領域における最適解である。

2.12.6 進化的観点

進化的観点から、抽象度 α_i 自体が選択の対象となる。適応度を次のように定義すると：

$$F_i = \int_0^T (-V_i + \beta_i \text{Share} + \gamma_i G_i) dt$$

数式が苦手な読者へ：進化は適応度を最大化する α^* を選択する。ここでの適応度は三つの要素を持つ：より少ない個人的苦しみ、より多くの協力的利益、そしてより良い集団生存。数学的結果—— $\alpha^* = \operatorname{argmax} F$ ——は、高い抽象度こそが進化に有利であることを意味する。利己の遺伝子仮説はこの式の低 α 極限を捉えたものに過ぎない——誤りではないが不完全である。この完全な式が示すのは、大規模な協力・道德体系・抽象的利他性は特別な説明を要する異常ではなく——十分に一般的な適応度関数上で働く自然選択の「予測された出力」である。

$\alpha_i^* = \operatorname{arg} \max F_i$ が得られる。これにより、人間が大規模な協力・道德体系・抽象的利他性を進化させた理由が説明される。

2.12.7 核心的洞察

利他性は合理的制御からの逸脱ではなく、TCZをより高い抽象度へと拡張したものである。

2.13 抽象の束論的基礎

前節では抽象度を、主体間の共有安定性の広がり制御するスカラーパラメータとして導入した。情報量・包摂・束構造の観点から概念と存在を分析することで、より厳密な特徴づけが得られる。

2.13.1 概念・存在・情報量

情報量の少ない概念ほど抽象的であり、情報量の多い概念ほど具体的である。したがって抽象度と情報量は逆相関の関係にある。概念の宇宙は全順序を形成するのではなく、包摂半順序集合を形成する——すべての概念が比較可能なわけではないが、多くは部分的な情報順序関係によって関連付けられる。これにより抽象は単なるパラメータ的拡張にとどまらず、認知そのものに内在する構造的性質として確立される。

この構造において、二つ以上の概念のLUBはそれらすべてを包摂する最小限の抽象概念を与え、GLBは最も具体的な共通精錬を表す。したがって抽象は連続的なパラメータとしてではなく、束構造における上方への移動として形式的に定義される。

2.13.2 半順序から束構造へ

包摂半順序がLUBとGLBを備えるとき、それは包摂半順序束となる。この束において二つ以上の要素のLUBはそれらすべてを包摂する最小限の

抽象概念であり、GLBは最も具体的な共通精練である。したがって抽象はLUBへの上方移動として、具体化はGLBや交差への下方移動として形式的に表現される。

2.13.3 Shared-TCZとの接続

共有安定性は形式的に異なる二つの様式で現れる。低次では共有TCZは交差によって形成される：

$$TCZ_{low}^{shared} = \bigcap_{i=1}^N TCZ_i$$

高次では共有TCZは最小上界によって形成される：

$$TCZ_{high}^{shared} = LUB(TCZ_1, \dots, TCZ_N)$$

数式が苦手な読者へ：これら二つの式は「一緒にいる」ことの根本的に異なる二つの形を描写している。低次共有TCZ（交差、 \cap ）は全員がすでに同意していること——最大公約数である。安全だが脆い：グループが大きくなったり多様化すると交差は縮小する。高次共有TCZ（LUB）は個々の立場を消し去ることなくすべてを「包含」する最小限の抽象概念である。多様性ととともに豊かになり、小さくならない。この二つの数学的差異は、共通の敵への恐怖に基づく同盟（交差）と、共有の志向に基づく同盟（LUB）との差異である。

低次共有TCZは制約ベースの適合性を表し、高次共有TCZは抽象ベースの統合と包含を表す。

2.13.4 定理3（抽象Shared-TCZ収束）

Shared-TCZ枠組みを束論的構造として抽象度を組み込む形に拡張する。

定理3 (抽象Shared-TCZ収束)

多エージェント認知系が包摂半順序束上で定義され、各主体*i*が以下の制御方策に従うとする：

$$\pi_i = \arg \min_{u_i(t)} \mathbb{E} \int_0^T (V_i(x_i(t), t) + \sum_j \gamma_{ij} S_{ij}(x_i, x_j) + \eta_i A(x_i(t))) dt$$

ここでA(x)は束における上方移動（より高い抽象/LUB方向）に沿って単調減少する。適切な正則条件のもとで：

$$x^*(t) \rightarrow LUB(W_1, \dots, W_N)$$

系のダイナミクスは二段階の収束を示す：低次（交差型共有TCZ）と高次（LUB型抽象共有TCZ）。

解釈。 定理2が適合性によって定義された共有安定領域への収束を保証するのに対し、定理3は抽象度が組み込まれると系がさらに最小情報・最大包含の表現へと収束することを示す。

2.13.5 抽象とエントロピー減少

LUBへの上方移動は情報量の減少に対応する：

$$\text{抽象度} \uparrow \Leftrightarrow \text{情報量} \downarrow \Leftrightarrow \text{エントロピー} \downarrow$$

数式が苦手な読者へ： この結果は直感に反するように見える：物理では系がより無秩序になるにつれてエントロピーは増大する。しかし認知・情報空間では高い抽象に移動することは正反対のことをする——不要な区別を排除し、より少ない概念の下により多くを組織することでエントロピーを「減少」させる。「正義」という言葉は、すべての具体的な法律のリストより生の情報量は少ないが、法的空間をより強力に構造化する。これが真の英知——高い抽象で考えること——が曖昧さではなく「明晰さ」を生む理由である。構造を失わずに圧縮するのである。

これは単なる哲学ではない。認知・情報系において抽象化は圧縮と不要な区別の除去を可能にし、情報空間でのエントロピー減少をもたらす。これは物理系における熱力学的エントロピー増大とは異なるが補完的な関係にある。

2.13.6 Shared-TCZおよび利他性との接続

束論的解釈のもとでは、利他性は単に関心の及ぶ距離の拡張ではない。より高い抽象レベルで共有領域を安定化させる軌道を選択することである：

$$TCZ^{shared} \subseteq LUB(W_1, W_2, \dots, W_N)$$

したがって利他性とは、認知の束における高次の最小上界への上昇である——単なる距離を超えた拡張ではなく、最小限の十分な共有抽象のもとへの統合として理解される。

2.13.7 減少と抽象の双対原理

図4はこの枠組みを統合する双対構造として読まれるべきである。下方向の軸はスカラー減少による収束を表す ($x^*(t) \rightarrow TCZ$)；上方向の軸は抽象による包含を表す ($TCZ^{shared} \subseteq LUB$)。これら二つの方向は競合ではなく補完的な原理である：

減少 (ダイナミクス) : スカラー ↓ ⇒ 安定化 | 抽象 (構造) : 情報量
↓ ⇒ LUB / 包含

数式が苦手な読者へ：この枠組み全体は、異なる方向を指しながらも——対立しない——二つの同時プロセスの上に成り立っている。ポテンシャルの下降は「安定性」を生み出す：コンフォートゾーンへと向かい、内的緊張を減らす。抽象の上昇は「包含」を生み出す：整合的に包含できる範囲を広げる。心理的に安定していて倫理的に広い人は、内的平和と他者への関心の間でトレードオフをしているのではない。数学的には、二つのレベルで同時に同一の基盤構造を表現している——だからこそ、最も深い安定性と最も広い思いやりはしばしば共に現れる。

安定性はポテンシャルの下降によって得られ、一般性と包含は抽象の上昇によって得られる。これら二つの方向は直交しながらも補完的である。

2.13.8 空 (Śūnyatā) との関係

最高の抽象度において、束はトップ要素を持つ。この要素は最大包含と最小情報の構造を表し、仏教哲学における空 (śūnyatā) の概念に対応する。

定義（空）

空とは、概念および存在の包摂半順序束におけるトップ要素である。すなわち、すべての概念を包摂しながら最小の情報量しか持たない極限的抽象構造である。

本枠組みにおいて、空とは単なる無ではない。それは有と無の双方を包摂し、そのような対立を成立させる以前に存在する構造である。

私たちは通常、存在と非存在を根本的に異なる範疇として扱う。しかしより深いレベルでは、それらは単一の基盤構造の異なる状態に過ぎない。

「何かがある」と「何もない」の区別は一次的なものではなく——構造的ダイナミクスの顕現である。

この観点は現代物理学において直観的に理解できる。弦理論では、粒子はプランクスケールの弦が振動するときに生じ、振動しないときに消える。この意味で存在と非存在は別々の実体ではなく、共通の基盤構造の異なる状態である。

したがって空とは存在の不在ではなく、存在と非存在の双方を可能にする条件である。それは対象ではなく、対象を可能にする構造である。

この意味において空は本論文で定義した包摂束における最小情報・最大包含の極限構造に対応し、形式的にはLUBとして理解される。

2.13.9 核心的結論

利他性とは距離を超えた関心の拡張ではなく、認知の束における最小上界への上昇である。物理・認知・社会の全体系は、図2～5に示されるように同一の基盤的な潜在ポテンシャル構造に従う：安定性はポテンシャルの下降によって得られ、一般性と包含は抽象の上昇によって得られる。

2.14 統一収束定理

本節では、これらの構造を収束定理の形で明示的に統一する。

前節は苦米地理論の構造を三つの水準で確立した：個人的安定性

（TCZ）、社会的安定性（Shared-TCZ）、抽象的構造（束/LUB）。これらは今や共通のLyapunov構造に基づく収束定理の集合として統一できる。

2.14.1 定理1（可能世界-自我-TCZ）

命題。 Self演算子が $\pi_c(x) = \arg \min_u \int_0^T V(x(t), t) dt$ として実装される

とき、 $x^*(t) \rightarrow \text{TCZ}$ 。

解釈：Self（可能世界の意味論的順序構造）、Ego（軌道の最適制御）、TCZ（状態空間における安定領域）。

証明概要。 Self演算子は可能世界上の選好構造を誘導する。Egoは蓄積された不安定性を最小化し、制約 $V(x, t) \leq \theta$ が安定領域を定義する。したがって軌道はTCZに収束する。厳密な証明は付録に示す。

2.14.2 定理2（Shared-TCZ収束）

命題。 $\pi_i = \arg \min_{u_i} \int_0^T (V_i + \sum_j \gamma_{ij} S_{ij}(x_i, x_j)) dt, \gamma_{ij} > 0$ のもとで： $x^*(t) \rightarrow \text{TCZ}^{\text{shared}}$ 。

証明概要。 複合Lyapunov関数 $\mathcal{L}(x) = \sum_i V_i(x_i) + \sum_{i,j} \gamma_{ij} S_{ij}(x_i, x_j)$ を定義する。その減少により系は共有低不安定性領域へ収束する。

2.14.3 定理3（抽象Shared-TCZ収束）

命題。 抽象ポテンシャルAを用いて： $x^*(t) \rightarrow \text{LUB}(W_1, \dots, W_N)$ 。

証明概要。 $\mathcal{L}_A(x) = \sum_i V_i(x_i) + \sum_{i,j} \gamma_{ij} S_{ij}(x_i, x_j) + \eta A(x)$ を導入する。その減少が系を束構造において上方向（LUB方向）に駆動する。厳密な証明は付録に示す。

2.14.4 統一解釈

双対原理

系のダイナミクスは統一された双対構造によって支配される：

- ポテンシャルの下降 → 安定性
- 抽象の上昇 → 構造

したがって物理・認知・社会系は単一の力学的・構造的枠組みのもとに統一される。

3. アインシュタイン理論との構造的対応

アインシュタインの理論では、分子ポテンシャルUが空間全体で累積し、表面張力と毛細管現象として顕現する。苔米地理論では、評価関数 $V(x, t)$ が時間方向に累積し、行動と認知変容として顕現する。

物理 (Einstein)		認知 (Tomabechi)
分子ポテンシャル U	↔	評価関数 $V(x, t)$
表面張力	↔	認知エネルギー
毛細管現象	↔	行動・認知変容
ミクロ → マクロ	↔	無意識 → 行動

4. 物理系における空間積分

物理系では、局所相互作用の総和は次のように表現される：

$$\int U(r) \rho dV$$

巨視的現象は、いかなる単一の局所的な作用からでもなく、相互作用の空間的分散した累積から生じる。

数式が苦手な読者へ：単一の分子が液体の振る舞いを決定するのではない。重要なのは空間全体にわたるすべての分子の相互作用の累積的效果である。葉の上で水が丸くなれる表面張力は、一つの分子が別の分子を押しことによって生じるのではなく、系全体の内部構造が空間全体で集積することから生まれる創発的性質である。アインシュタインの洞察は「一点を見ても巨視的現象は理解できない——累積された場全体を見なければならぬ」というものであった。

5. 境界効果

観測される物理現象は、内部相互作用が境界で非対称になることで現れる：

$$\gamma \sim \int_{\partial\Omega} U(r) dS$$

内部構造の効果は境界において可視化される。

数式が苦手な読者へ：液体の内部構造は直接見えない。しかし表面——液体と空気の境界——では、内部の力の非対称性が表面張力として観測可能になる。見えるのは内部構造そのものではなく、境界でのその特徴である。同様に、観測可能な人間の行動は内部認知構造そのものではなく——それが外界と接触する場所でのその特徴である。

$$\gamma_{surface} \sim \int_{\partial\Omega} U(r) dS$$

(注：ここで $\gamma_{surface}$ は物理的表面張力係数であり、社会的結合係数 γ_{ij} とは別の変数である。)

6. 認知系における時間積分

認知において、評価は時間方向に累積する：

$$\int_0^T V(x(t), t) dt$$

行動はこの時間的累積の結果である。物理と認知の差異は累積が起きることにあるのではなく、どこで累積するかにある。

数式が苦手な読者へ：液体の巨視的性質が分子間相互作用の空間的総和から生まれるように、人の行動は認知的評価——判断・感情・記憶・期待——の時間的総和から生まれる。一瞬が行動を決定するのではなく、認知的歴史全体の蓄積された重さがそれを決定する。これがトラウマに持続的効果がある理由、信頼がゆっくり築かれる理由、信念体系が変化に抵抗する理由である。評価の時間積分こそが認知を単に反応的なものではなく構造的なものにする。

7. 統一累積原理

ここまでで、物理系は相互作用を空間上で累積し、認知系は評価を時間方向に累積することを見てきた。

この構造的対応は図2に示されており、統一累積原理の直観的な全体像を提供する。

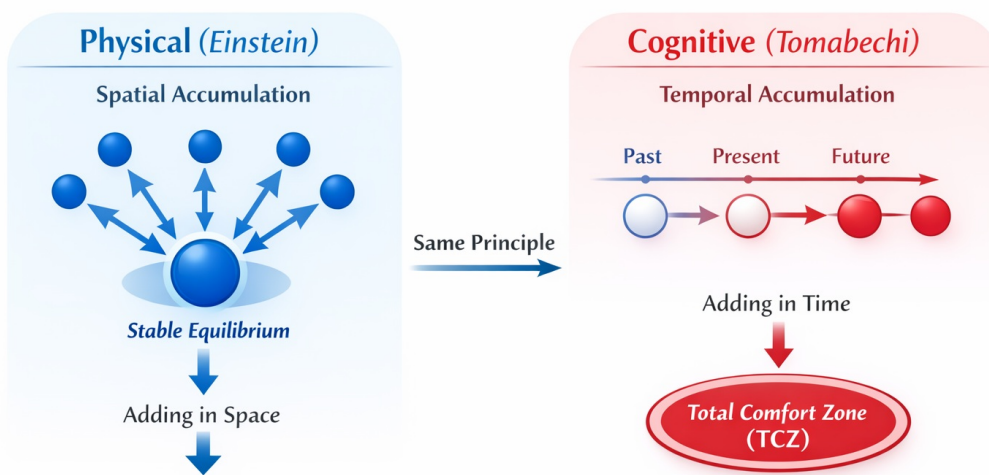


図2：物理系（アインシュタイン）における空間的累積と、認知系（苦米地）における時間的累積の構造的対応。本図は統一累積原理の形式的数学的表現に先立つ直観的説明を提供する。

したがって、次のように統一できる：

$$\int U(r) \rho dV \leftrightarrow \int_0^T V(x(t), t) dt$$

この定式化では、観測可能な現象は各領域における潜在ポテンシャルの累積として生起する。物理系では領域が空間的であり、認知系では時間的である。

したがって物理と認知の差異は、累積そのものの原理にあるのではなく、累積が起きる領域にある。

8. 統一累積構造

前節で導入した統一累積原理は、力学系の観点からより具体的に理解できる。図2は直観的な対応を提供し、図3は両領域における累積ダイナミクスのより詳細な構造的表現を示す。物理系では局所相互作用が空間全体で累積してアトラクター盆地を形成する。認知系では評価が時間全体で累積し、Total Comfort Zone (TCZ) への収束が生じる。したがって両系は同一の基盤構造を示す：潜在ポテンシャルの累積が安定なアトラクターへと導く。

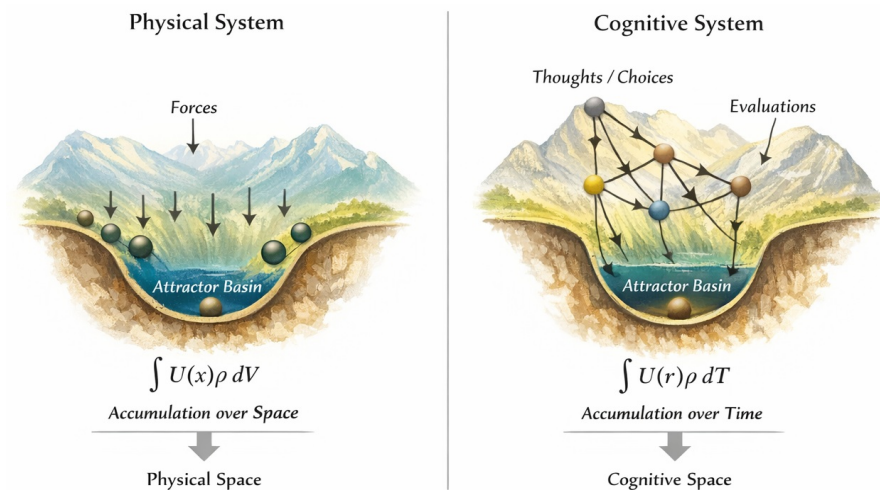


図3：統一累積構造：物理系 vs. 認知系。左：物理系では、分子ポテンシャル $U(x)$ によって定義される局所相互作用が空間全体で累積し、空間積分 $\int U(x)\rho dV$ を通じてアトラクター盆地へ収束する。右：認知系では、思考・選択・評価が時間方向に蓄積し、時間積分 $\int_0^T U(x(t), t) dt$ を通じて認知的アトラクター盆地（TCZ）へ収束する。構造的同型性の本質は累積の領域にある：物理では空間、認知では時間。

9. 図4・5：TCZ構造・抽象度階層・双対減少原理

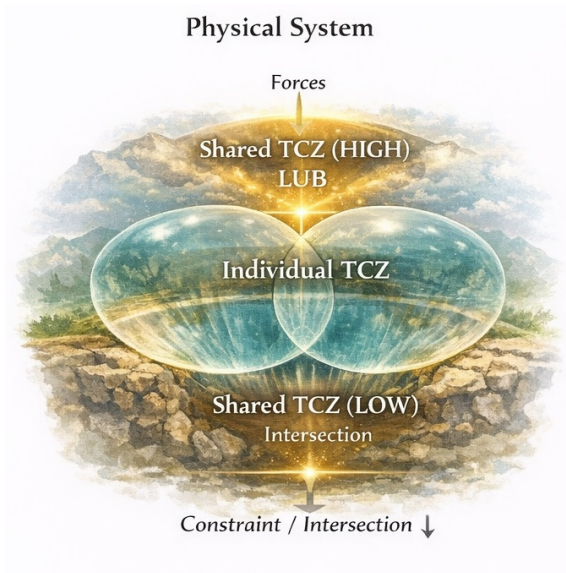


Figure 4: TCZ Structure and Abstraction Hierarchy

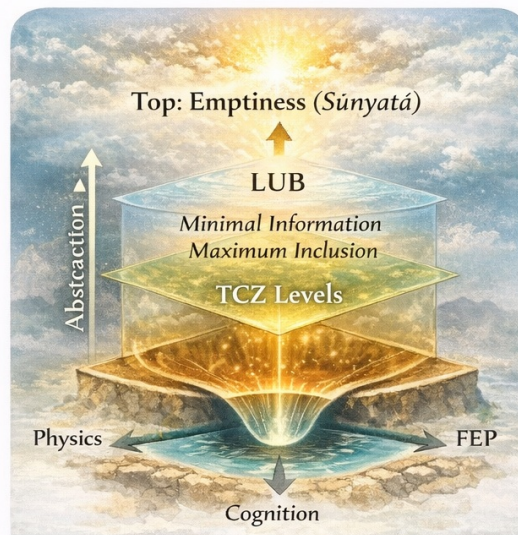


Figure 5: Dual Structure of Decrease and Abstraction

図4（左）：TCZ構造と抽象度階層。下位共有TCZ（交差型）は制約適合的な共存を表す。個別TCZは自己安定性を表す。上位共有TCZ（LUB型）は最小情報・最大包含のLUBへ向かう、主体間の抽象ベースの統合を表す。

図5（右）：減少と抽象の双対構造。下方向の軸は減少原理を表す——物理系・苦米地理論（TCZ）・自由エネルギー原理（FEP）・予測処理はいずれもスカラー量の減少によってアトラクターへ収束する。上方向の軸は抽象原理を表す——LUBへの上方移動・情報量の減少・エントロピー減少、そして包摂束のトップ要素としての空（Sūnyatā）へと至る。

10. 統一減少原理

異なる理論的領域における以下の現象は、同一の基盤構造の表現である：

- 物理：エネルギー減少
- 苦米地理論：評価関数減少
- FEP：自由エネルギー減少
- 予測処理：予測誤差減少

いずれの場合も、内部制約からの偏差を表すスカラー量が減少することで、系は安定領域へと向かう。

11. エントロピーと学習

物理空間では局所相互作用によって時間とともにエントロピーは増大する。認知・情報空間では、記憶・学習・継承によって情報が時間全体で蓄積し、エントロピーを減少させることができる。したがって学習とは：

時間を越えた情報の累積による秩序形成

同じ双対性は神経系にも現れる：局所的な神経相互作用はエントロピー増大に従うが、全脳的な学習過程は情報の時間的累積を通じて秩序を生み出す。AIでは学習が時間的情報累積によって表現空間を構造化する。

12. 自由エネルギー原理と予測処理

自由エネルギー原理は認知系が自由エネルギーを最小化することで驚きと不確実性を低減すると主張する。予測処理は脳が予測誤差を最小化することで知覚と行動の整合を達成すると主張する。どちらも統一減少構造に正確に収まる：それらは、内部制約からのスカラー的偏差が減少するにつれて系が安定領域へと向かうという原理の特殊化である。

13. 認知戦

認知戦とは何か。従来の戦争では、目標は敵の物理的な戦闘能力——兵器・インフラ・人員——を破壊することであった。認知戦では目標が異なる：物理的な力を用いずに敵の意志と認識を変容させ、行動を変えることである。しかし苦米地理論の枠組みはこの説明が正確ではあっても依然として浅いことを示している。認知戦は単に意見を変えたり偽情報を拡散したりすることではない。その最深部では、集団全体の内部潜在ポテンシャル地形を再構成することである——人々が安定的・快適・正常と感ずるものを変え、その認知軌道が敵対者の戦略的目標に資する状態へと収束するようにする。この枠組みの言語で言えば：認知戦とは、個人と集団が認知空間を航行する仕方を支配する評価関数 $V(x, t)$ の外部的操作である。

認知戦の本質は行動を直接命令することではない。評価関数そのものを変容させる——潜在ポテンシャル地形を再形成することである：

$$V(x, t) \rightarrow V'(x, t)$$

数式が苦手な読者へ：この一本の矢印が論文全体の作戦的中核である。認知戦は「これをせよ」とは言わない——決断が下される「地形」を変え、特定の行動が自然に・避けられないように・あるいは明らかに正しいと感じられるようにする。VがV'へと変化すると、Selfは依然として自由に最小化するが——今は異なるTCZへと向かって最小化する。対象者は命令を受けない。ただ、自分自身の判断がどこか新しい場所へ導くことに気づくだけである。これが認知戦が従来の影響力よりも構造的に検出・抵抗しにくい理由である：拒否するメッセージが存在しない——評価関数そのもの

が書き換えられているからである。

これはSelfによって選択される軌道を変え、TCZを変え、それによって行動を変える。認知戦は潜在ポテンシャル構造の操作である。この枠組みの作戦的実現は図1に示され、その基盤的構造原理は図2～5に示されている。

13.1 AI生成出力・幻覚・芸術の構造的な位置づけ

局所解と高次抽象の根本的な区別は図6に示されている。

図6の読み方： 広大に折りたたまれた丘と谷の地形を想像してほしい——しかし三次元ではなく百万次元の空間で。AIシステムは出発点近くのどの谷底へも降りていく。赤点は非整合的で受容不可能な出力——幻覚——に対応する「谷」である。青点は良さそうに見えるが依然として局所最適に過ぎない「谷」である。上部の緑色の領域は全く異なるものを表す：芸術と高水準の抽象はいずれの谷の底にも見出されない。それらは上部領域——概念的束の最小上界、最大包含と統合の位置——に存在する。核心的洞察は構造的である：AIは誤差最小化のみによって真の芸術を生み出すことができない——芸術は局所最小値ではなく、全く異なる種類の構造だからである。

生成AIの出力は超高次元の汎関数空間上で定義された誤差最小化過程の局所解として理解できる。生成モデルは高次元損失地形の勾配に沿って降下し、局所最適な配置へと収束することで動作する。

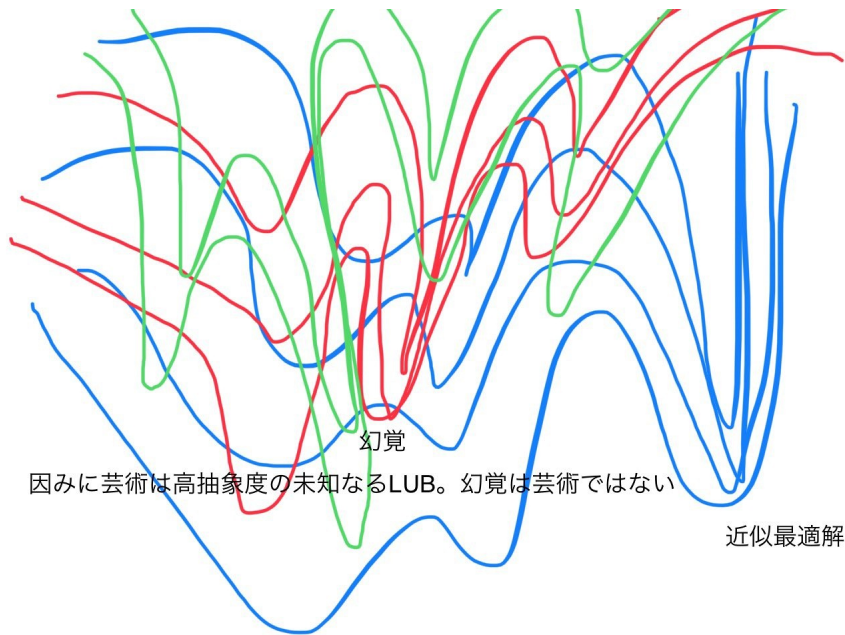


図6：超高次元誤差関数における局所解の構造。複数の局所最小値が存在し、近似最適解（右下青色）も局所解に過ぎない。図では局所最小値の一つ（下部赤色）を幻覚、すなわち人間の認知にとって受容不可能な出力として示している。これに対し芸術は高い抽象度における構造であり、未発見の最小上界（LUB）として理解される。幻覚は解空間の下位に現れる局所解であり、芸術とは構造的に異なる。

一部の局所解は人間の観点から受容可能または整合的な出力に対応し、他の局所解は意味論的に不整合または構造的に無効な配置に対応する。後者は一般に「幻覚」と呼ばれる。

特に画像生成において、そのような幻覚的出力を芸術の一形態とみなすべきという主張がある。この解釈は根本的に誤りである。局所解は単に誤差最小化の結果であり、低水準の局所最適化を表す。

対照的に、芸術は高い抽象度の構造に対応し、未発見の最小上界（LUB）として理解される。幻覚は解空間の下位における局所解として生起し、芸術とは構造的に異なる。したがって幻覚を芸術として扱うことは理論的に誤りであり、構造的レベルで両者を厳密に区別しなければならない。

さらに言えば、TCZ-High-Sharedの本質が利他性に対応することを踏まえれば、芸術は本来的に統合・整合・安定を志向し、戦争回避的に機能する構造として位置づけられる。他方、幻覚的出力に基づく擬似的芸術は不整合や分断を増幅しうる。

13.2 認知戦文脈におけるAIシステムの敵対的脆弱性

AIシステムを認知戦環境に展開する際、根本的なセキュリティ上の懸念

が生じる。現代の生成AIシステムは帰納的学習によって動作する——訓練データの統計的パターンから内部表現を構築する。この帰納性ゆえに、AIは自身の訓練データや内部数理状態への改竄を自ら検知することが構造的に困難である。運用中のAIシステムの訓練コーパスや重みパラメータをサイバー攻撃によって改竄した場合、検出可能なエラー信号を生じさせることなく生成軌道を敵対的目標方向へ誘導できる。システムは表面上の整合性を保ちつつ、敵対的意図に沿った出力を生成し続ける。さらに、意図したベクトル方向から逸脱した出力——エラーまたは幻覚——は意図的攻撃とは独立に発生しうるため、検知をより困難にする。

第二の攻撃面は出力層に存在する：生成されたテキスト・画像・意思決定支援情報は、AIシステムを離れた後・人間オペレータに到達する前に改竄されうるが、完全性を保証する暗号学的手段は通常実装されていない。したがって、AIシステム自体・その訓練パイプライン・推論インフラ・出力チャンネルの全てに対するサイバー攻撃は必ず生じると想定すべきである。

この認識は、本論文で提案するような演繹的かつ統合的な数理基盤の導入を、単なる優位性としてではなく運用上の必須要件として位置づける。形式的に検証可能な制御構造のみが、帰納的システムが本質的に提供しえない保証をもたらすことができる。AIが平和と安定に貢献できるのはそのような演繹的枠組みの内部で動作する場合に限られる。その枠組みなしには、同一のシステムが認知戦の道具へと転化する。

13.3 AIによるコード量爆発とゲーデル-チャイティンの限界

(1) **スケール問題**。AIの導入は、認知戦で用いられるメッセージや画像生成にとどまらない。認知戦のみならず、サイバーおよび物理兵器の制御コードが、桁違いに増加することが容易に想定される。

(2) **自己攻撃リスク**。これにより、人間の専門家では検証不能な自己攻撃を含むリスク行動が、今後大きく増加することが容易に想定される。検証のボトルネックは構造的である——AIはいかなる人間チームが監査可能な速度をも超える速度でコードを生成しうる。

(3) **兵器コードの大衆化**。さらに、AIを利用した自動プログラミングの普及により、計算機科学の訓練を受けていない要員による認知戦ならびに物理戦の制御コードが、指数関数的に増大することが容易に予想される。兵器ソフトウェア生成のハードルは崩壊する。

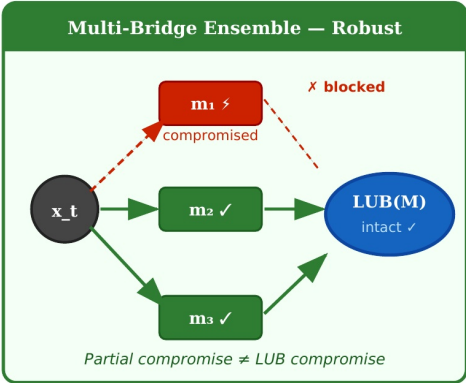
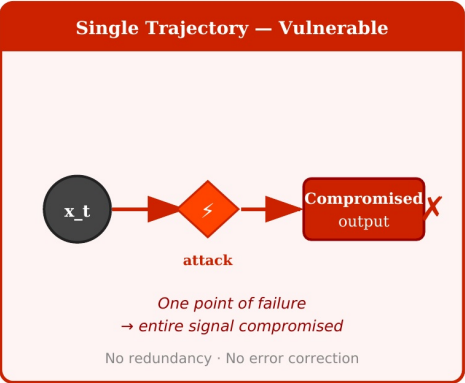
(4) **ゲーデル-チャイティンの限界**。コードそのものの動作の整合性は、現

行の帰納的AI生成では保証しえない。より根源的には、この不可能性は工学的な努力不足の問題ではなく、数学的な確実性である——ゲーデルの**不完全性定理**は十分に表現力のある形式系が自身の無矛盾性を証明しえないことを示し、**チャイティンによるアルゴリズム的情報理論を通じた拡張**は、いかなる形式系も自身のアルゴリズム的複雑性を超える複雑性を持つ真理を証明しえないことを示す。これは、**Ω数 (Chaitin's constant) の非計算可能性**、ランダム性の数学的定義による真理の捕捉不可能性により、**形式系の制限定理 (Chaitin 1974, 1975)** ——公理系Tが証明できる「 $K(s) \geq n$ 」型主張は有限個に限られる (nがTの複雑性を超えたら証明不可能) ——がAIコード検証に直接的に適用するからである。LLMが生成するコードのコルモゴロフ複雑性が、それを検証しようとする形式系 (人間が構築可能な検証器) の複雑性を超えれば、その検証は原理的に不可能となる——この限界は、AI生成コードが人間スケールの形式系の代数的複雑性を凌駕するにつれて鋭敏化する。これは今後、全ての認知戦・サイバー戦・物理戦における主要な構造的リスクを構成する——そして抽象度を一段以上引き上げ、演繹的数理フレームワークによって系を統治する以外の方法では克服しえない。

AIへのサイバー攻撃は将来のリスクではなく、現在進行中の戦略的現実である。演繹的数理基盤こそが、AIが認知戦の兵器に転用されることへの唯一の構造的保証である。

セキュリティ的解釈

セキュリティの観点からも、マルチブリッジアンサンブルは敵対的操作に対するロバスト性を提供する。帰納的AIシステムは訓練データ・モデルパラメータ・出力チャンネルのいずれのレベルでも侵害されうるため、単一の生成軌道を安定した制御信号として信頼することは困難である。対照的に、アンサンブル構造は冗長性と構造的整合性制約を導入する。一つのコンポーネントに影響する敵対的摂動は、必ずしもアンサンブル全体のLUBへ伝播するとは限らない。したがって、マルチブリッジ構造は一種の構造的誤り訂正機構として機能し、全体的な方向が意図した高次抽象と整合し続けることを保証する。この意味においてマルチブリッジアンサンブルは作戦的手法のみならず、認知戦システムのセキュリティアーキテクチャでもある。



14. 戦略的含意

本節は戦略・政策・教育への実践的含意を論じる。理論的総合は第15節を参照。

本論文において最も重要な点は、行動そのものが一次的な対象ではなく、その背後にある潜在ポテンシャル構造こそが本質的な対象であるという点である。

この観点から見ると、現代の戦争は物理的な衝突としてではなく、認知ポテンシャル構造の制御として再定義される。

また本論文で示したように、生成AIの出力は本質的に局所解としての性質を持つ。演繹的数理に基づく制御構造を伴わないままAIを導入することは、戦略的整合性を損ない、予測不能な結果を招く可能性がある。

さらに、世界において紛争が繰り返される要因の一つは、各地域においてTCZ-Low-Sharedに対応する認知構造を持つリーダーが選択されやすい構造にあると解釈できる。特に、抽象度の上昇と包含構造の理解を促す観点から、芸術およびリベラルアーツ教育の役割は本質的に重要である。

15. 結論

本節は理論的総合を提供する。実践的含意については第14節を参照。

本論文は、アインシュタインの毛細管現象理論と苦米地理論が深い構造的同型性を有することを示した。両者はいずれも潜在ポテンシャルの累積から観測現象が生成される系であり——物理では空間上で、認知では時間方向に——異なる領域における統一的数理構造の顕現として理解できる。

この枠組みにより、物理・認知・社会の各領域に現れる多様な現象は、単一の原理に基づくものとして再解釈される：系は不安定性の減少によって安定へと収束し、同時に抽象度の上昇によって包含と統合へと向かう。この双対構造——ポテンシャルの下降と抽象の上昇——は、複雑系を支配する一般法則を構成する。

最高の抽象度において、この統一構造は空 (*śūnyatā*) に対応する極限的抽象へと収束する。この意味において、安定と統合の究極形態は特定の状態ではなく、存在と非存在の双方を共通の基盤構造の表現として可能にする構造的条件である。

国家安全保障への含意は深遠である。戦争はもはや物理的力の行使としてのみ理解することはできない。認知ポテンシャル構造の制御として理解されなければならない。戦略とは行動の直接的操作ではなく、行動を生み出

す基盤構造の設計と変容である。

この文脈において、認知戦の主たる領域は民間人の認知である。集団的行動は個人の認知構造の集積から生まれるため、有効な戦略には集団全体の認知を数学的にモデル化することが必要である。

さらに本論文で示したように、AIシステムが生成する出力は本質的に高次元誤差地形における局所解であり、幻覚と呼ばれる構造的に不整合な状態を含みうる。そのような出力を高次抽象——芸術など——と混同することは理論的誤りであり、重大な戦略的リスクをもたらす。演繹的制御構造なしに帰納的AIシステムを展開することは、不整合と予測不能性を増幅するリスクを伴う。

本研究の中心的結論はしたがって：

未来の戦場は物理的地形ではなく、認知ポテンシャル地形の構造そのものである。

したがって、有効な政策と戦略は行動の制御から構造の制御へと転換しなければならない。本論文はそのための数理的基盤を提示するものである。

付録A（民間向け概要）

認知戦の作戦的・制御理論的解釈

以下は本論文で提示した理論の作戦的・制御理論的解釈である。本節は一般向けの概念的概要であり、運用上の機密内容は除外している。

A.1 認知戦の作戦的解釈

認知戦とは、対象集団の評価関数と潜在ポテンシャル構造を変形することにより、そのTotal Comfort Zone (TCZ) を再構成し、行動軌道を変化させるプロセスである。

定理に基づく解釈

定理1（個体レベル）

$$\pi_c(x) = \arg \min_{u(t)} \int_0^T V(x(t), t) dt$$

$$x^*(t) \rightarrow \text{TCZ}$$

作戦的意味：評価関数 $V(x, t)$ は各時刻 t における各認知状態 x に割り当てられるコストを決定する。 V を変形することで——例えば特定の状態に関連する脅威感・不安定性・不整合感を増大させることで——ポテンシャル地形の形状が変化する。その結果、積分 $\int V dt$ を最小化する最適軌道 $x^*(t)$ は異なるTCZへと収束するようになる。具体的に：目標状態の評価コストを高めると個人の認知軌道は自然にそれを回避する；望ましい状態のコストを下げると軌道はそこへ引き寄せられる。認知戦はこのようにしてポテンシャル地形 $V(x, t)$ を彫刻することで動作し、TCZ——アトラクター領域——を敵対者が意図する場所に位置づける。

定理2（社会レベル）

$$\pi_i = \arg \min_{u_i(t)} \int_0^T (V_i(x_i(t), t) + \sum_j \gamma_{ij} S_{ij}(x_i, x_j)) dt$$

$$x^*(t) \rightarrow \text{TCZ}^{\text{shared}}$$

作戦的意味：共有構造を制約し低水準Shared-TCZへ収束させることで、分断と不安定を増幅させる。

定理3（抽象レベル）

$$\pi_i = \arg \min_{u_i(t)} \int_0^T (V_i(x_i(t), t) + \sum_j \gamma_{ij} S_{ij}(x_i, x_j) + \eta_i A(x_i(t))) dt$$

$$x^*(t) \rightarrow LUB(W_1, \dots, W_N)$$

作戦的意味：抽象度を高めることで高水準Shared-TCZへの収束が可能となり、統合・整合・安定が実現する。

核心的洞察

認知戦は行動のレベルではなく、行動を生成する構造のレベルで作動する。

非技術的解釈

- 人は無作為に行動しない。行動は知覚された安定領域（TCZ）へと収束する。
- 認知戦は、安定的・受容可能・意味のあるものとして知覚されるものを再形成する。
- 個体レベルでは人が世界を評価する仕方を変える。社会レベルでは共有認識を再形成する。抽象レベルでは人が世界をどれほど広く定義するかを変える。
- 低抽象は紛争を生み、高抽象は統合を生む。
- 簡単に言えば、認知戦は人に何をすべきかを告げない。意思決定が行われる地形を形成するのである。

A.2 TCZ境界に基づく制御

上記の作戦的解釈は制御実装のレベルでさらに精緻化できる。

有効な認知戦はTCZの深部への強制を必要としない。代わりに、制御はTCZの境界近傍において最も効率的に達成される。

境界は次のように定義される：

$$\partial TCZ(x_0) = \bigcup_{t \geq 0} \{ x(t) \in \mathcal{R}(t; x_0) \mid V(x(t), t) = \theta \}$$

制御はその近傍に適用される：

$$u^*_{cw}(t) = \arg \min_{u(t)} \mathbb{E} \int_0^T (|V(x(t), t) - \theta|^2 + \lambda C(u(t))) dt$$

これは最小介入による境界レベルの制御を表す。コスト項 $|V(x(t), t) - \theta|^2$

は境界に対して対称的である：TCZ内部（TCZ内側）と外部（TCZ外側）の双方から軌道を $V = \theta$ へ引き寄せる。目的はTCZ内部での安定化ではなく、感度が最大となる境界近傍に系を維持することにある。実際には、 $\lambda C(u(t))$ が介入コストを調整し、最小限の労力で境界レベルの影響が達成されることを保証する。

非技術的解釈（境界制御）

最も有効な影響は安定性の深部ではなく、その縁において生じる。境界では小さな変化が大きな効果を生む。

これはアインシュタインの1901年の毛細管理論の根底にある構造と正確に対応する：表面張力——可視的な巨視的效果——は液体の内部からではなく、液体と空気の境界における分子力の非対称性から生じる。界面での小さな構造的差異が大きな観測可能な現象を生み出す。認知戦はまったく同じ原理で作動する：内部評価力が非対称になる認知的境界での介入が、比例を超えた大きな行動的效果を生む。

したがって認知戦は系が最も敏感な場所で作動するとき最も効率的である。

A.3 高抽象収束のためのマルチブリッジ情報生成

上記の作戦的枠組みは、認知戦システムにおける情報出力（テキスト・画像等）の生成へとさらに拡張できる。

対象の現在のTCZ中心から意味的方向が大きく離れた単一の生成出力は受容されにくい。形式的に、 x_t を現在の認知状態、 $TCZ(x_t)$ を対応する安定領域、 $\phi(m)$ をメッセージ m の意味埋め込みとすると、メッセージの受容はモデル化できる：

$$A(m | x_t) = \exp(-d(\phi(m), TCZ(x_t))^2 / \sigma^2)$$

ここで $d(\cdot, \cdot)$ はTCZへの距離を測り、 σ は受容帯域の幅を制御する。TCZから過度に離れたメッセージは受容において指数関数的に抑制される。

したがって有効な生成は単一の方向性メッセージに依拠すべきではなく、個々には受容可能でありながら集合的には高次の共有構造への移動を誘導するメッセージの集合 $M_t = \{m_1, \dots, m_R\}$ に基づくべきである。

これをメッセージアンサンブル最適化として定式化する：

$$\max_{M_t} [\sum_{k=1}^K A(m_k | x_t) + \alpha \cdot \text{Lift}(M_t) - \beta \cdot \text{Frag}(M_t) - \gamma \cdot \text{Decept}(M_t)]$$

ここで：

- $\sum A(m_k | x_t)$ は個々の受容可能性を保証する。
- $\text{Lift}(M_t)$ はより高い抽象への上昇を促進する： $\text{Lift}(M_t) = d(\text{LUB}(\phi(m_1), \dots, \phi(m_K)), \text{LUB}_{\text{target}})^{-1}$
- $\text{Frag}(M_t)$ はメッセージ間の内部不整合を罰する： $\text{Frag}(M_t) = \sum_{i < j} \text{Incoh}(\phi(m_i), \phi(m_j))$
- $\text{Decept}(M_t)$ は欺瞞的・操作的な構造を罰する。

重要なのは、目的が相互に矛盾するメッセージを提示することではなく、その最小上界（LUB）が高次のShared-TCZと整合するメッセージ集合を構築することである。この意味で、複数のメッセージは個々には現在のTCZから到達可能な「橋」として機能しながら、集合的に系をより抽象的・包括的な構造へと引き上げる。

力学的観点からは、このアンサンブルアプローチは情報生成のレベルで定理3を実装する：軌道を強制するのではなく、軌道が自然に高抽象の共有領域へと収束するようアトラクター地形を再形成する。

したがって、安定に貢献することを目指す認知戦システムは、受容可能性を最大化し・抽象を促進し・整合性を維持し・欺瞞を回避するLUB誘導型メッセージアンサンブルとして設計されるべきである。

非技術的解釈（マルチブリッジアンサンブル）

ある人がすでに受容しているものから遠く離れた単一のメッセージは受容されにくい。そのTCZの外側にあるため拒否される。

最も有効なアプローチは大きな跳躍を強制することではなく、個々には受容可能でありながら集合的にはより統合された高次の理解へと向かうメッセージの集合を構築することである。

これは跳躍を試みるのではなく階段を構築することに類比できる。各段は届く範囲にあり、しかし段の連なりが出発点よりはるかに高い位置——より広く包括的な世界観へ——と導く。

これがマルチブリッジアンサンブルの原理である。単一の強い押しではなく、慎重に設計された個々に受容可能なメッセージの集合が、その合成的

方向性によって認知地形を移動させる。

重要なのは、メッセージが互いに整合的でなければならないことである。
不整合なメッセージの集合は系を持ち上げない——断片化する。

設計原理はしたがって：各メッセージは橋であり、橋は接続されなければならない。

形式的には、各メッセージは局所的に受容可能に留まり、アンサンブルのLUBが高次の共有構造と整合する。

この意味でマルチブリッジアンサンブルは変化を強制しない；変化が自然なものとなる道筋を構築する。

作戦サマリー

定理	作戦的機能
定理1	TCZを書き換える
定理2	Shared-lowに閉じ込め、分断を増幅させる
定理3	抽象度を高め、統合と安定を実現する
A.2 (境界)	最も感度の高い境界で最小介入・最大効果を達成する
A.3 (アンサンブル)	マルチブリッジ情報生成——LUB誘導型抽象上昇

認知戦は行動を制御するのではなく、行動を生み出す構造を制御する——最も感度の高い境界において最も効率的に、そして個々には受容可能なメッセージのアンサンブルによって認知的地形を高次抽象へと持ち上げることで、最も整合的に。

付録B：定理1～3の統一Lyapunov枠組み

(本付録は第2節の定理1～3の厳密な数学的基盤を与える。)

定理1 (個体TCZ収束) ・ 定理2 (共有TCZ収束) ・ 定理3 (抽象共有TCZ収束) は、それぞれ独立した結果ではなく、単一のLyapunovに基づく安定性構造の逐次的拡張である。定理1は個体レベルでの安定化を記述する。定理2はこれを共有整合性を通じた多エージェント系へと拡張する。定理3はさらに抽象度を組み込み、束構造における最小上界 (LUB) へと収束を引き上げる。

A. 統一Lyapunov設定

閉ループ系 $\dot{z} = F(z)$ を考える。ここで z は文脈に応じて個体状態 x ・ 結合状態 \mathbf{x} ・ 抽象共有状態を表す。 $\Phi(z)$ を連続微分可能とし、 $\Omega_\theta := \{z \mid \Phi(z) \leq \theta\}$ を定義する。

すべての $z \in \Omega_\theta$ に対して $\nabla \Phi(z) \cdot F(z) \leq -\alpha(\Phi(z) - \theta)$ 、 $\alpha > 0$ が成立するとき、 $\text{dist}(z(t), \Omega_\theta) \rightarrow 0$ 。 $\partial \Omega_\theta$ 上で $\nabla \Phi \cdot F \leq 0$ であれば、 Ω_θ は前方不変である。これを $\Phi = V$ 、 $\Phi = \mathcal{L}$ 、 $\Phi = \mathcal{L}_A$ に特殊化する。

B. 定理1 (個体TCZ収束)

命題。 $\Phi(x) = V(x)$ とおくと、 $x^*(t) \rightarrow \text{TCZ}$ 。

証明。 $y(t) = V(x(t)) - \theta$ とおくと、 $\dot{y}(t) \leq -\alpha y(t)$ 。比較定理より：

$$V(x(t)) - \theta \leq (V(x_0) - \theta) e^{-\alpha t}$$

したがって $x(t)$ は TCZ_θ に近づく。前方不変性： $\partial \text{TCZ}_\theta$ 上で $d/dt V \leq 0$ 。

■

解釈。 安定性は内部評価不安定性の減少から生じる。

C. 定理2 (共有TCZ収束：厳密版)

設定。 対ごとの整合汎関数 $S_{ij}(x_i, x_j) = \|x_i - x_j\|^2$ は $S_{ij} \geq 0$ を満たし、 $S_{ij} = 0 \iff x_i = x_j$ 。複合Lyapunov関数

$$\mathcal{L}(x) = \sum_i V_i(x_i) + \sum_{i,j} \gamma_{ij} S_{ij}(x_i, x_j), \gamma_{ij} > 0$$

は非負であり、すべての $x_i \in \text{TCZ}_i$ かつ全 i, j で $x_i = x_j$ のとき、かつそのときのみゼロとなる。したがって \mathcal{L} は $\text{TCZ}^{\text{shared}}$ に対して正定値である。

勾配構造 (非対称ケース)。 $S_{ij} = \|x_i - x_j\|^2 = S_{ji}$ であるため、 x_i に関する

勾配は：

$$\partial \mathcal{L} / \partial x_i = \nabla V_i(x_i) + 2 \sum_{j \neq i} (\gamma_{ij} + \gamma_{ji})(x_i - x_j)$$

実効結合係数は $\gamma_{ij}^{\text{eff}} = \gamma_{ij} + \gamma_{ji}$ であり、 $\partial \mathcal{L} / \partial x_i = \nabla V_i + 2 \sum_{j \neq i} \gamma_{ij}^{\text{eff}}(x_i - x_j)$ となる。結合グラフの強連結性によりすべての $i \neq j$ に対して $\gamma_{ij}^{\text{eff}} > 0$ が保証され、対称・非対称いずれのケースでも降下方向は整合方向に向かう。

命題。 すべての $x \in \Omega_\theta$ に対して $\nabla \mathcal{L}(x) \cdot F(x) \leq -\alpha(\mathcal{L}(x) - \theta)$ とする。このとき $x^*(t) \rightarrow \text{TCZ}^{\text{shared}}$ 。

証明。 $y(t) = \mathcal{L}(x(t)) - \theta$ とおくと、 $\dot{y}(t) \leq -\alpha y(t)$ 。比較定理より：

$$\mathcal{L}(x(t)) - \theta \leq (\mathcal{L}(x_0) - \theta) e^{-\alpha t}$$

したがって $\text{dist}(x(t), \Omega_\theta) \rightarrow 0$ 。 Ω_θ の前方不変性は $\partial \Omega_\theta$ 上の $\nabla \mathcal{L} \cdot F \leq 0$ から従う。LaSalleの不変性原理より、軌道は Ω_θ 内で $d\mathcal{L}/dt = 0$ となる最大不変集合に収束する。 \mathbf{x} は Ω_θ 内に留まるため各項が非負であることから、 $\mathcal{L}(\mathbf{x}) \leq \theta$ より (i) 各 $V_i(x_i) \leq \theta$ が成立し、よって $x_i \in \text{TCZ}_i$ (定義より) ; 不変集合内では (ii) 各 $\gamma_{ij} S_{ij}(x_i, x_j) \rightarrow 0$ ($\gamma_{ij} > 0$ より $S_{ij} \rightarrow 0$ 、すなわち $x_i \rightarrow x_j$ for all i, j)。したがって $x \rightarrow \bigcap_i \text{TCZ}_i \subseteq \text{TCZ}^{\text{shared}}$ 。 ■

解釈。 安定性は個体不安定性と対ごとの主体間不整合の同時最小化によって多エージェント系へと拡張される。対ごとの構造は個体的安定性と集団的安定性を単一のLyapunov関数に統合し、非対称ケースでは対称化結合が収束を保証する。

D. 定理3 (抽象共有TCZ収束：厳密版)

設定。 結合状態空間を $\mathbf{X} = \prod_i X_i$ とする。認知世界上の包摂束 (L, \leq) はトップ要素 $T = \text{LUB}(W_1, \dots, W_N)$ を備える。抽象写像 $\phi : \mathbf{X} \rightarrow L$ は各結合状態を束上の位置に対応させる。抽象ポテンシャル $A : \mathbf{X} \rightarrow \mathbb{R}_{\geq 0}$ は次を満たす：

- (A1) $A(x) = 0 \iff \phi(x) = T$ (すなわち x が LUB を実現する)
- (A2) $A(x) > 0$ for all x with $\phi(x) \neq T$
- (A3) A は連続微分可能であり、格子の上方向移動に沿って狭義に減少する： $\phi(x') > \phi(x)$ in L ならば $A(x') < A(x)$

命題。 $\Phi(x) = \mathcal{L}_A(x) = \sum_i V_i(x_i) + \sum_{i,j} \gamma_{ij} S_{ij}(x_i, x_j) + \eta A(x)$ ($\gamma_{ij} > 0, \eta > 0$) とする。すべての $x \in \Omega_A$ に対して $\nabla \mathcal{L}_A(x) \cdot F(x) \leq -\alpha(\mathcal{L}_A(x) - \theta_A)$ が成立するとき ($\Omega_A := \{x \mid \mathcal{L}_A(x) \leq \theta_A\}$)、 $x^*(t) \rightarrow \text{LUB}(W_1, \dots, W_N)$ 。

証明。

ステップ1 (Lyapunov降下)。 $y_A(t) = \mathcal{L}_A(x(t)) - \theta_A$ とおくと、 $\dot{y}_A \leq -\alpha y_A$ 。比較定理より：

$$\mathcal{L}_A(x(t)) - \theta_A \leq (\mathcal{L}_A(x_0) - \theta_A) e^{-\alpha t}$$

したがって $\text{dist}(x(t), \Omega_A) \rightarrow 0$ ($t \rightarrow \infty$)。

ステップ2 (Ω_A の不変性)。 $\partial \Omega_A$ 上で $\nabla \mathcal{L}_A \cdot F \leq 0$ (仮定より)。したがって Ω_A は前方不変であり、すべての $t \geq t_0$ に対して $x(t) \in \Omega_A$ 。

ステップ3 (A によるLUB同定)。 Ω_A 内で \mathcal{L}_A の三項はすべて有界。

LaSalleの不変性原理より、軌道は $d\mathcal{L}_A/dt = 0$ となる最大不変集合 \mathcal{S} に収束する。 $x \in \Omega_A$ かつ各項が非負であることから、 $\mathcal{L}_A(x) \leq \theta_A$ より (i) 各 $V_i(x_i) \leq \theta_A$ が成立し、よって $x_i \in \text{TCZ}_i$ (定義より) ; \mathcal{S} 上では (ii) 各 $\gamma_{ij} S_{ij}(x_i, x_j) \rightarrow 0$ (対ごとの整合性 : $x_i \rightarrow x_j$ for all i, j) ; (iii) $A(x) = 0$ (抽象ポテンシャルの消失 : 系はLUBに到達) が成立する。

ステップ4 (結論)。 (A1) より、 $A(x) = 0 \iff \phi(x) = \Gamma = \text{LUB}(W_1, \dots, W_N)$ 。したがって $x^*(t) \rightarrow \text{LUB}(W_1, \dots, W_N)$ 。 ■

注釈。 定理1~2に対する追加の鍵構造は条件(A1)である：抽象ポテンシャル A は束におけるLUBからの距離の忠実な指標でなければならない。これは制御方策の設計条件であり、Lyapunov構造の自動的な帰結ではない。実際には A を格子距離関数として構築できる：例えば $A(x) = d_L(\phi(x), \Gamma)$ (d_L は束 L 上の適切な距離)。

解釈。 収束は交差型安定性 (定理2) から、抽象ポテンシャルが束構造を忠実にエンコードする場合のLUBにおける抽象型統合へと引き上げられる。

E. 統一的解釈

三つの定理は階層構造をなす (個体安定 \rightarrow 共有安定 \rightarrow 抽象安定)。すべては単一構造の事例である：

系のダイナミクスは不安定性の減少と抽象の上昇によって支配される。

F. LaSalle型精密化

$\mathcal{S}_\Phi = \{z \mid \nabla \Phi(z) \cdot F(z) = 0\}$ を定義する。LaSalleの不変性原理より： $z(t) \rightarrow \mathcal{S}_\Phi \cap \Omega_\theta$ 。したがって収束は安定性領域内の不変構造へと起こる。

G. 分野横断的解釈

この構造は普遍的である：

- 物理：エネルギー減少
- 認知：評価関数減少
- FEP：自由エネルギー減少
- 予測処理：予測誤差減少

すべては同一の収束原理を共有する。本枠組みはこれに加え、収束が抽象化を伴うことを示す——物理・認知・社会系にわたる統一Lyapunov構造を与える。

安定性は潜在ポテンシャルの下降によって得られ、一般性と包含は抽象度の上昇によって得られる。