A Unified Theory of Latent Potentials: Homeostasis and Cognitive Warfare

— Toward a Mathematical Foundation of Cognitive Control in Physical, Social, and AI Systems —

Hideto Tomabechi

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

April 4, 2026

Last update May 3, 2026

April 21, 2026 U.S. National Defense University Lecture Paper (Civilian Public Edition)

Preamble

In March 2026, a tugboat in the Strait of Hormuz was struck by a missile and sank, leaving three crew members missing. A separate tanker was also attacked and set ablaze. Although more than one hundred vessels typically transit the Strait, traffic temporarily dropped to only a handful of ships. As a result, global oil prices surged, immediately impacting economies and political stability across multiple countries. From a purely physical perspective, these incidents were localized and small in scale. Yet their impact propagated globally within a very short time frame. This illustrates a defining feature of modern conflict: physical engagements are no longer isolated events, but are inseparably coupled with economic perception, risk assessment, and large-scale cognitive dynamics. In practice, the effective disruption of the Strait was not caused by complete physical blockade, but by a shift in decision-making—driven by rapidly increasing insurance costs and perceived risk. Ship operators rerouted, markets reacted, and a localized physical event transformed into a global systemic effect. This reveals a fundamental transformation in warfare: physical warfare and cognitive warfare are no longer separable domains; operational decisions must account simultaneously for physical actions and cognitive effects; and strategic outcomes emerge from the interaction between physical systems and collective cognition. Consequently, modern operations require a unified framework in which physical and cognitive systems are modeled and controlled as a single integrated domain. Without such integration, decision-making becomes fragmented, internally inconsistent, and increasingly unpredictable. This integration is often assumed to be achievable through artificial intelligence. However, contemporary AI systems are fundamentally inductive: they rely on statistical learning from past data and generate predictions based on learned patterns. While powerful, such systems do not provide guarantees of global coherence across domains. Therefore, what is required is not merely data-driven optimization, but **a deductive mathematical foundation capable of integrating physical and cognitive domains at the level of system dynamics**. Such a foundation must satisfy correctness, soundness, and robustness—and must therefore be supported by formal mathematical proof. The purpose of this paper is to provide such a foundation. We introduce a unified mathematical framework that integrates physical, cognitive, and social dynamics, and establish its validity through Lyapunov-based convergence analysis and LaSalle-type arguments. Operational-level control equations based on this framework, including real-time Common Operational Picture (COP) implementations, have already been constructed. However, these are beyond the scope of the present paper for security reasons.

Historical Context

For over three decades, the author has worked at the intersection of cognitive science and national security.

In the early 1990s, the author was involved in foundational research in functional brain science, including collaborative work with Harvard Medical School. In the mid-1990s, following the Tokyo subway sarin attacks, the author worked with Japan's National Police Agency to assist in the deprogramming of Aum Shinrikyo members, applying cognitive science and functional neuroscience in operational settings. These activities were publicly reported in Japan at the time.

What became clear during that period is that techniques of psychological control—commonly referred to as "brainwashing"—were not accidental or ad hoc phenomena. They were originally developed as controlled methodologies, in many cases within classified military contexts. Elements of these methodologies had diffused beyond their original domains and were actively exploited by non-state actors.

On the symbolic side of AI research, in 1986 at the Center for Machine Translation (now the Language Technologies Institute), Carnegie Mellon University, the author developed the world's first speech-to-speech translation system (between English and Japanese). This work represents an early integration of symbolic processing and real-time language interaction, and precedes the later convergence of symbolic and neural approaches addressed in this paper.

The author's work on generative and cognitive architectures dates back to the late 1980s. In 1988, Elman proposed one of the earliest recurrent neural network models with context units (now commonly referred to as the Elman network), which predicted subsequent tokens and represents an early form of generative neural systems. In the same period, the author implemented a deeper multi-layer neural architecture incorporating multiple context layers to extend sequence prediction beyond single-step outputs.

In 1991, the MONA-LISA framework ("Multimodal Ontological Neural Architecture for Linguistic Interactions and Scalable Adaptations") by the author introduced a hybrid architecture in which deductive symbolic processes operate in parallel with neural generative systems. This work constitutes one of the earliest systematic approaches to addressing what are now referred to as hallucinations in generative AI.

Over more than three decades, the author's research has consistently treated the mitigation of such inconsistencies not as an optional refinement, but as a fundamental requirement—particularly in mission-critical systems, including those relevant to defense applications.

More than thirty years later, the situation has fundamentally changed. Technologies capable of intervening in human cognition have advanced dramatically, while the battlespace itself has expanded from physical terrain into the cognitive domain of civilian populations.

In this context, what is now called "cognitive warfare" must be understood in its full operational meaning. At its core, cognitive warfare is not merely information influence or the spread of misinformation. It is, in essence, a form of systemic cognitive control.

However, current public and policy discussions tend to reduce cognitive warfare to phenomena such as AI-generated misinformation. This framing significantly underestimates its structural nature and strategic implications.

The term "Cognitive Warfare" itself was introduced by the author in 2007 in the context of research and educational activities. What is now required is not a descriptive or metaphorical understanding, but a formal definition grounded in cognitive science and mathematical structure.

The central implication of this work is the following: **Effective cognitive warfare requires modeling the cognition of all individuals within the operational domain.**

Moreover, since the primary targets of cognitive warfare are civilian populations, collective behavior must be understood as the aggregation of individual cognitive models. Without such modeling, strategic actions risk internal inconsistency and unintended systemic effects. This paper provides the mathematical foundation necessary to address this problem.

In August 2023, the author's team demonstrated the world's first Cognitive Warfare COP (Common Operational Picture) system to Admiral John Aquilino, then Commander of U.S. INDOPACOM. This was followed by a series of non-mathematical briefings at closed, high-level defense meetings.

Given the rapidly growing strategic interest in cognitive warfare, this is a critical moment to clarify why a rigorous deductive mathematical foundation is essential for its effective development and deployment.

Abstract

This paper proposes a unified theoretical framework connecting physical, cognitive, and social dynamics through latent potential structures. The departure point is Einstein's 1901 study of capillarity, which showed that macroscopic phenomena arise from the accumulation of internal interactions. We extend this structural insight into cognition, showing that behavior emerges from an internally extended homeostatic system. We formalize the Tomabechi framework as a general theory of cognitive dynamics: cognition is defined over possible worlds, ordered by evaluative stability, and realized as optimal control minimizing accumulated instability. This yields Theorem 1: optimal trajectories converge to a Total Comfort Zone (TCZ). Extending to multi-agent systems with shared evaluative structures yields Theorem 2: socially coupled agents converge to shared stability regions (Shared-TCZ), establishing that social stability emerges as a joint attractor rather than being externally imposed. We then introduce a lattice-theoretic foundation. Concepts form a subsumption partial ordered lattice in which abstraction corresponds to movement toward least upper bounds (LUBs). Extending the Lyapunov structure with an abstraction potential yields Theorem 3: system dynamics are lifted toward the LUB of shared cognitive worlds. Altruism is ascent toward higher-order shared abstractions, not expansion over distance. Abstraction corresponds to information reduction: Abstraction $\uparrow \iff$ Information $\downarrow \iff$ Entropy $\downarrow$. These results establish a unified dual principle: descent in potential (stability) and ascent in abstraction (inclusion). At the limiting case, this process converges to emptiness (śūnyatā)—the top element of the subsumption lattice, representing maximal inclusion with minimal information. Emptiness is not mere absence, but a generative structure that subsumes both being and non-being. Art, in this framework, is understood not as a local solution produced by generative processes, but as a structural realization of an as-yet undiscovered LUB—the highest form of integration and abstraction. As demonstrated in this paper, outputs of generative AI are inherently local solutions and may include structurally inconsistent states known as hallucinations. Confusing such outputs with higher-order abstract structures is theoretically incorrect

and introduces critical risks in cognitive warfare. Deploying AI without a deductive mathematical control structure undermines strategic coherence and may produce unpredictable outcomes. The paper also develops three operational extensions. At the individual level (A.1), cognitive warfare is formalized as the external modification of the evaluative function $V(x,t)$, thereby redefining the TCZ toward which behavioral trajectories converge. At the boundary level (A.2), optimal control is shown to be most efficiently achieved near the boundary of the TCZ, where sensitivity to perturbation is maximal. At the generative level (A.3), effective information operations are formalized as multi-bridge message ensembles: sets of individually acceptable messages whose least upper bound (LUB) aligns with a higher-level Shared-TCZ, implementing Theorem 3 at the level of information generation. The future battlespace is not physical terrain, but the structure of cognitive potential landscapes. Accordingly, strategy, policy, and national security all require a theoretical framework targeting structures rather than actions. This paper provides the mathematical foundation for that transition.

1. Introduction

Einstein's first paper addressed the origin of capillary phenomena in classical physics. Its significance lies beyond explaining surface tension: the essential contribution is the demonstration that observable order can arise not from external forces, but from the cumulative effect of internal interactions within a system.

Surface tension is not caused by external pushing. Molecular interactions accumulate throughout space, and asymmetry at the boundary produces macroscopic physical effects. Surface tension is therefore a visible manifestation of internal structure.

This idea has profound implications for cognition. In the Tomabechi framework, human behavior and cognitive transformation are not simple reactions to external stimuli. Rather, they emerge from an internally extended homeostatic mechanism operating over cognitive space.

The central intuition of this paper can be summarized as:

Molecules move according to the sum over space; humans act according to the sum over time.

This statement is not metaphorical. In physical systems, interactions accumulate across space. In cognitive systems, evaluations accumulate across time. The difference is not in principle, but in the domain of accumulation.

Roadmap. The remainder of this paper is organized as follows. Section 2 develops the core mathematical framework through three convergence theorems: individual cognitive stability (Theorem 1), social stability through shared structures (Theorem 2), and abstract integration through lattice theory (Theorem 3). Plain-language summaries accompany each theorem for readers who prefer to skip the mathematics. Sections 3–8 establish the structural correspondence with Einstein's capillarity theory and the unified accumulation principle. Sections 9–12 connect to the decrease principle, entropy, and the Free Energy Principle. Section 13 addresses cognitive warfare and the structural distinction between AI hallucinations and art. Sections 14–15 draw strategic and policy implications. The Appendix provides rigorous proofs of all three theorems.

2. Core Structure of the Tomabechi Framework

Before comparing with Einstein's theory, the Tomabechi framework must be presented as an independent theory. The formulation below is not derived from classical physics but proposed as a general theory of cognitive dynamics. The structural comparison that follows is a demonstration of isomorphism between two independent theories, not an application of physics to cognition.

2.1 Cognitive Homeostasis and the Need for a Possible-Worlds Model

Human belief systems exhibit a degree of stability that cannot be explained by simple models of information updating. Rather than being linearly overwritten by new inputs, beliefs tend to preserve internal coherence. This stability is best understood as cognitive homeostasis.

To model such homeostasis, cognition must be represented not as a single state, but as a set of possible worlds W. The effective domain of cognition includes not only the present state but also future possibilities:

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

Human cognition operates within this space of possible futures, continuously evaluating potential trajectories rather than reacting only to immediate stimuli.

2.2 Ordering of Possible Worlds

Possible worlds are not equivalent. Some are cognitively stable, while others are associated with instability, inconsistency, or threat. Therefore, the set of possible worlds must be equipped with an ordering based on cognitive comfort.

We introduce an ordering operator:

$$r : W \rightarrow W$$

which reorganizes possible worlds according to their compatibility with the agent's internal homeostatic constraints. Thus, cognition is not a uniform representation of alternatives, but a structured ordering over possible futures.

We introduce an ordering operator $r : W \rightarrow W$ which reorganizes possible worlds according to their compatibility with the agent's internal homeostatic constraints.

2.3 Possible-Worlds Definition of TCZ

Under this ordering, the stable region of cognition is defined as the Total Comfort Zone (TCZ):

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

The TCZ is the totality of future worlds that the agent can stably inhabit. It is not a static set, but a dynamically maintained region under cognitive homeostasis.

2.4 Self / Ego Operator

Within this framework, the Self (or Ego) is not a fixed entity, but an operator acting on the TCZ. Formally, the Self is defined as:

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

Thus, the Ego is a process that reorganizes the structure of the TCZ, rather than a static center of cognition.

2.5 Ego as an Optimal Control Model

In the Tomabechi framework, the Ego is formulated as an optimal control problem:

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

where $x(t)$ denotes the cognitive state, $u(t)$ the control input, and $V(x, t)$ represents cognitive instability, evaluative cost, or internal inconsistency.

For readers who prefer to skip the mathematics: This equation says that the Ego—the part of the self that makes decisions—constantly steers cognition toward states of lower internal tension and greater stability. It is not reacting randomly; it is solving an optimization problem at every moment, minimizing accumulated discomfort over time. In everyday terms: a person does not just react to each event independently; they navigate life in a way that, over time, moves them closer to their own stable, coherent inner world. The Ego is the operator that performs this navigation.

For readers who prefer to skip the mathematics: This equation says that the Ego constantly steers cognition toward states of lower internal tension and greater stability, solving an optimization problem at every moment, minimizing accumulated discomfort over time.

2.6 Connection Between Possible Worlds and Control

The above discussion can be summarized in the following five steps:

1. Human cognition unfolds over a set of possible worlds W.
2. Possible worlds are ordered according to cognitive comfort via a mapping $r : W \to W$.
3. The stable region under this ordering is the Total Comfort Zone (TCZ).
4. The Self / Ego acts as an operator that reorganizes the ordering structure of the TCZ.
5. This reorganization is implemented, in state space, as an optimal control problem.

Thus, the possible-worlds formulation provides a semantic representation, while optimal control provides its dynamical realization.

2.7 State-Space Definition of TCZ

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

where $\mathscr{R}(t; x_0)$ denotes the reachable set from the initial state x_0, and θ is an admissible threshold.

For readers who prefer to skip the mathematics: The TCZ is simply the collection of all mental states that a person can reach—from their current situation—while remaining within a comfortable, stable range. Think of it as a map of all the psychological territory that feels "safe" to inhabit. States outside the TCZ represent excessive tension, inconsistency, or distress. The formula captures this precisely: the TCZ is every state that is reachable and keeps the internal instability measure V below a threshold θ. The structure of a person's TCZ defines, in mathematical terms, what their "comfort zone" actually is.

Before proceeding to the main theorem, it is important to clarify the relationship between the state-space definition of the Total Comfort Zone (TCZ) and its formulation in terms of possible worlds.

The state-space definition

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

describes TCZ as the set of reachable states that satisfy a stability condition.

On the other hand, the possible-worlds formulation

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

describes TCZ as a structured set of future worlds ordered by cognitive stability.

Although these two expressions appear different, they represent the same underlying structure. Intuitively, a "possible world" corresponds to a trajectory in state space. The ordering relation over possible worlds corresponds to the evaluative function $V(x,t)$, which determines stability. The existence of a stabilizing relation in the possible-world formulation corresponds to the existence of a trajectory that remains within the stability region in state space. Thus, the possible-worlds formulation provides a semantic description of cognitive structure, while the state-space formulation provides its dynamical realization. Establishing the equivalence between these two perspectives is essential, as it demonstrates that the abstract notions of "self" and "comfort zone" can be rigorously implemented as a control system. This equivalence is what allows us to move from conceptual descriptions to mathematical guarantees.

We now proceed to the main theorem.

where $\mathscr{R}(t; x_0)$ denotes the reachable set from initial state x_0, and θ is an admissible threshold.

2.8 Main Theorem (Possible Worlds–Ego–TCZ)

Theorem 1 (Possible Worlds–Ego–TCZ).

If the Self operator defined on the set of possible worlds W is implemented as the optimal control problem

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

then the optimal trajectory $x^*(t)$ converges into $\mathrm{TCZ}(x_0)$.

2.9 Interpretation of the Main Theorem

This theorem unifies the semantic and dynamical layers of the Tomabechi framework. Self, Ego, and TCZ are not independent constructs; they represent the same cognitive process at three levels:

- Self: semantic ordering of possible worlds
- Ego: optimal control of trajectories
- TCZ: stability region in state space

Thus, cognition is simultaneously a semantic structure, a control process, and a stability system.

2.10 Sketch of Proof in the Main Text

The Self operator modifies the ordering structure of possible worlds. The Ego selects trajectories that minimize accumulated cognitive instability. The condition $V(x, t) \leq \theta$ defines a stability constraint. Therefore, the optimal trajectory $x^*(t)$ must eventually enter the TCZ.

A rigorous mathematical justification is provided in the Appendix using Lyapunov-based arguments and a LaSalle-type invariance principle. This structure is not only theoretical but directly implementable (see Fig. 1).

How to read Figure 1: This figure shows a real operational system built on Theorem 1. Each panel represents a different cognitive or operational metric being tracked in real time. The key idea is that the mathematical structure of the TCZ—the stable attractor region in cognitive space—can be translated directly into computational metrics that a command system can monitor and act upon. Readers unfamiliar with the technical details should focus on what the figure demonstrates conceptually: that the abstract mathematics of cognitive stability is not just philosophy, but is already implemented as operational technology.

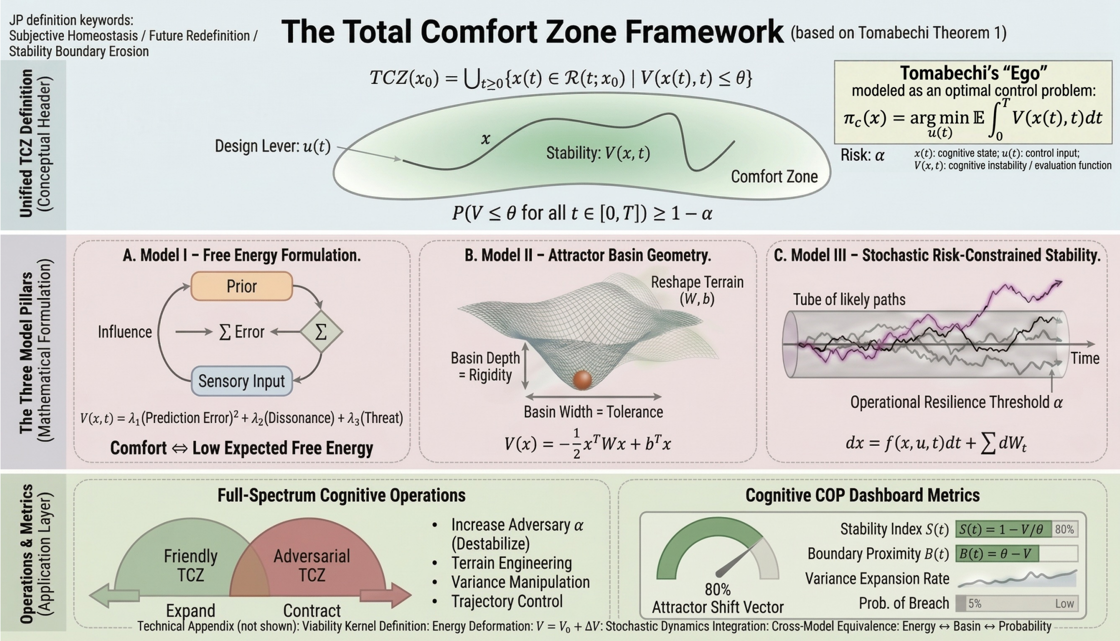

Figure 1: Example implementation of a cognitive warfare system COP (Common Operational Picture) based on Theorem 1. The system integrates optimal control of cognitive instability, attractor basin geometry, and stochastic stability constraints to generate real-time operational metrics. In practical deployment, the system incorporates the extended mathematical structures of Theorems 2 and 3, and is executed using massively parallel AI architectures.

2.11 Shared-TCZ Theorem

2.11.1 Motivation

So far, the Tomabechi framework has been formulated for a single cognitive agent. However, human cognition is inherently social. Stable behavior often emerges not from individual optimization alone, but from alignment with other agents. This suggests that the TCZ should be extended from an individual construct to a shared construct across multiple agents.

2.11.2 Multi-Agent Extension

Let there be N agents indexed by $i = 1,\ldots,N$, each with cognitive state $x_i(t)$ and evaluative function $V_i(x_i, t)$. We extend the control policy of each agent to:

$$\pi_i = arg\ min_{u_i(t)} \int_0{}^T \ (\ V_i(x_i(t), t) + \lambda_i\, S_i(\boldsymbol{x}(t))\ )\ dt$$

where $\mathbf{x}(t) = (x_1(t),\ldots,x_N(t))$; S_i is a shared-stability functional measuring alignment with other agents; and $\lambda_i \geq 0$ controls the strength of social coupling.

Plain language: Each person minimizes not only their own internal tension, but also their misalignment with others. The term S_i captures how much a person's cognitive state conflicts with the states of those around them. When λ_i is large, that person is strongly socially coupled—they are sensitive to alignment with others. When it is small, they operate more independently. This is the mathematical structure of social conformity, group identity, and collective stability. A society where everyone has a high λ will tend to move toward shared cognitive states—for better or worse, depending on what those shared states are.

where S_i is a shared-stability functional measuring alignment with other agents; and $\lambda_i \geq 0$ controls the strength of social coupling.

2.11.3 Definition of Shared TCZ

We define the Shared Total Comfort Zone as:

$$TCZ^{\eta}{}_{shared} = \{ \, x \in \textstyle\prod_i TCZ_i \mid Share(x) \geq \eta \, \}$$

where TCZ_i is the individual TCZ of agent i; $\text{Share}(\mathbf{x})$ measures the degree of cognitive alignment; and η is a threshold of shared stability.

2.11.4 Theorem 2 (Shared-TCZ Convergence)

Theorem 2 (Shared-TCZ Convergence).

Assume that each agent i follows the control policy above with $\lambda_i > 0$, and that S_i is continuous and bounded below. Then, under appropriate regularity conditions, the joint trajectory $\mathbf{x}^*(t)$ converges to a shared stability region:

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

2.11.5 Interpretation

This theorem generalizes the original TCZ convergence result. An individual Self leads to convergence toward the individual TCZ; a socially coupled Self leads to convergence toward the shared TCZ. Social cognition can thus be understood as the emergence of jointly stable regions in cognitive space—providing a formal basis for coordination, social norms, shared belief systems, and collective behavior. This establishes that **social stability is not externally imposed, but emerges as a joint attractor** in multi-agent cognitive space.

This theorem establishes that **social stability is not externally imposed, but emerges as a joint attractor** in multi-agent cognitive space.

2.12 Altruism as Abstract TCZ Expansion

2.12.1 Motivation

The Shared-TCZ framework explains how multiple agents may converge to mutually compatible stable regions. However, it does not yet explain a key feature of human cognition: humans can act altruistically toward agents who are distant, unrelated, or even unknown. This suggests that sociality extends beyond proximity-based interaction into abstract domains such as nations, humanity, or future generations. To capture this, we introduce the notion of cognitive abstraction level.

2.12.2 Abstraction Parameter

Let each agent be associated with an abstraction level $\alpha_i \in \mathbb{R}_+$. This parameter determines how broadly the agent defines its effective domain of concern. Small α_i corresponds to local, kin-based, or immediate interactions; large α_i corresponds to global, abstract, or universal considerations.

Let each agent be associated with an abstraction level $\alpha_i \in \mathbb{R}_+$. Small α_i corresponds to local, kin-based, or immediate interactions; large α_i corresponds to global, abstract, or universal considerations.

2.12.3 Abstraction-Dependent Sharing

We extend the shared-stability function to depend on abstraction:

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

where $w_{ij}(\alpha_i) = \exp(-\text{dist}(i,j) \,/\, \alpha_i)$. As α_i increases, distant agents receive higher weight.

where $w_{ij}(\alpha_i) = \exp(-\text{dist}(i,j) \,/\, \alpha_i)$.

2.12.4 Extended Control Model

$$\pi_i = arg\ min \int_0{}^T [\ V_i(x_i, t) + \lambda_i\, S_i(\boldsymbol{x}; \alpha_i) + \eta_i\, A(\alpha_i)\ ]\ dt$$

where $A(\alpha_i)$ is an abstraction cost, and η_i balances abstraction against cost.

2.12.5 Interpretation of Altruism

Altruism can now be formally defined: *altruism is a control bias toward higher abstraction levels, in which the Self selects trajectories that stabilize shared TCZs across distant and diverse agents.* In this formulation, altruistic behavior is not irrational—it is optimal under a broader evaluative domain, and it emerges from high-α Self dynamics. **Altruism is not irrational—it is optimal under an expanded evaluative domain.**

Altruism is not irrational—it is optimal under an expanded evaluative domain.

2.12.6 Evolutionary Perspective

From an evolutionary standpoint, the abstraction level α_i is itself subject to selection. Defining fitness as:

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

we obtain $\alpha_i^* = \arg\max F_i$. This explains why humans evolved large-scale cooperation, moral systems, and abstract altruism.

From an evolutionary standpoint, the abstraction level α_i is itself subject to selection. Defining fitness as:

we obtain $\alpha_i^* = \arg\max F_i$. This explains why humans evolved large-scale cooperation, moral systems, and abstract altruism.

2.12.7 Key Insight

Altruism is not a deviation from rational control, but an extension of the TCZ into higher levels of abstraction.

2.13 Lattice-Theoretic Foundations of Abstraction

In the previous sections, abstraction was introduced as a scalar parameter controlling the breadth of shared stability across agents. A more rigorous characterization can be obtained by analyzing concepts and existence in terms of information content, subsumption, and lattice structure.

2.13.1 Concepts, Existence, and Information Content

Concepts carrying less information are more abstract; concepts carrying more information are more specific. Thus, abstraction and information content stand in inverse relation. The universe of concepts does not form a total order but a **subsumption partial ordered set**—not all concepts are comparable, but many are related by partial information-ordering relations. This establishes abstraction not merely as a parametric extension, but as a structural property of cognition itself.

Within this structure, the LUB of two or more concepts is the minimally abstract concept that subsumes them all, while the GLB represents the most specific common refinement. Thus, abstraction is not merely a continuous parameter, but is formally defined as upward movement in the lattice structure.

2.13.2 From Partial Order to Lattice Structure

When the subsumption partial order is equipped with least upper bounds (LUB) and greatest lower bounds (GLB), it becomes a **subsumption partial ordered lattice**. Within this lattice, the LUB of two or more elements is the minimally abstract concept that subsumes them all; the GLB is the most specific common refinement. Thus, abstraction is formally represented by upward movement toward LUBs, while specification is represented by downward movement toward GLBs or intersections.

2.13.3 Connection to Shared-TCZ

Shared stability appears in two formally distinct ways. At lower levels, shared TCZ is formed through intersection:

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

At higher levels, shared TCZ is formed through least upper bound:

$$TCZ^{high}{}_{shared} = LUB(TCZ_1,\ldots,TCZ_N)$$

The lower shared TCZ represents constraint-based compatibility; the higher shared TCZ represents abstraction-based integration and inclusion.

2.13.4 Theorem 3 (Abstract Shared-TCZ Convergence)

We extend the Shared-TCZ framework to incorporate abstraction as a lattice-theoretic structure.

Theorem 3 (Abstract Shared-TCZ Convergence)

Let a multi-agent cognitive system be defined over a subsumption partial ordered lattice, and assume each agent *i* follows the control policy:

$$\pi_i = arg\ min_{u_i(t)}\ \mathbb{E} \int_0{}^T\ (\ V_i(x_i(t),t) + \lambda_i S_i(\boldsymbol{x}(t)) + \eta_i A(x_i(t))\ )\ dt$$

where $A(x)$ monotonically decreases along upward movements in the lattice (toward higher abstraction / LUB). Then, under appropriate regularity conditions:

$$\boldsymbol{x}^*(t)\ \rightarrow\ LUB(W_1,\ldots,W_N)$$

System dynamics exhibit two-stage convergence: lower-level (intersection-based shared TCZ) and higher-level (LUB-based abstract shared TCZ).

Interpretation. While Theorem 2 guarantees convergence to a shared stability region defined by compatibility, Theorem 3 shows that when abstraction is incorporated, the system further converges toward minimally informative and maximally inclusive representations.

2.13.5 Abstraction as Information and Entropy Reduction

Upward movement toward LUBs corresponds to a reduction in information content:

$$Abstraction \uparrow \iff Information \downarrow \iff Entropy \downarrow$$

This is not merely philosophical. In cognitive and informational systems, abstraction enables compression and elimination of unnecessary distinctions, yielding entropy reduction in information space. This is distinct from but complementary to thermodynamic entropy increase in physical systems: whereas local interactions in physical space increase entropy, abstraction operates as an independent organizing principle in cognitive and information space.

2.13.6 Connection to Shared-TCZ and Altruism

Under lattice-theoretic interpretation, altruism is not merely the extension of concern across greater distance. It is the selection of trajectories that stabilize shared regions at higher levels of abstraction:

$$TCZ_{shared} \subseteq LUB(W_1, W_2, \ldots, W_N)$$

Altruism is therefore ascent toward higher-level least upper bounds in the lattice of cognition—not merely expansion over distance, but integration under a minimally sufficient shared abstraction.

2.13.7 The Dual Principle of Decrease and Abstraction

Figure 4 should be read as a dual structure that integrates this framework. The downward axis represents convergence via scalar reduction ($x^*(t) \to$ TCZ); the upward axis represents inclusion via abstraction ($\text{TCZ}_{\text{shared}} \subseteq$ LUB). These two directions are not competing but complementary principles:

$$Decrease\ (Dynamics):\ \ Scalar \downarrow\ \Rightarrow\ Stability\ \ |\ \ Abstraction\ (Structure):$$
$$Information \downarrow\ \Rightarrow\ LUB\ /\ Inclusion$$

Stability is obtained by descent in potential; generality and inclusion are obtained by ascent in abstraction. These two directions are orthogonal yet complementary.

2.13.8 Relation to Emptiness (Śūnyatā)

At the highest level of abstraction, the lattice admits a top element. This element represents a structure of maximal inclusion and minimal information, corresponding to the notion of emptiness (śūnyatā) in

Buddhist philosophy.

> **Definition (Emptiness)**
>
> Emptiness is defined as the top element of the subsumption partial ordered lattice of concepts and existence. That is, it is the limiting abstract structure that subsumes all concepts while containing minimal information.

In this framework, emptiness is not mere nothingness. It is a structure that subsumes both being and non-being, and exists prior to the distinction that makes such oppositions possible.

We typically treat existence and non-existence as fundamentally different categories. At a deeper level, however, they are simply different states of the same underlying structure. The distinction between "something" and "nothing" is not primary—it is a manifestation of structural dynamics.

This perspective can be understood intuitively in modern physics. In string theory, particles arise when strings at the Planck scale vibrate, and disappear when they do not. In this sense, existence and non-existence are not separate entities, but different states of a common underlying structure.

Accordingly, emptiness is not the absence of existence, but the condition that makes both existence and non-existence possible. It is not an object, but the structure that makes objects possible.

In this sense, emptiness corresponds to the limiting structure of minimal information and maximal inclusion, and can be formally understood as the least upper bound (LUB) in the subsumption lattice defined in this paper.

This establishes emptiness as the ultimate limiting structure of abstraction within the unified framework.

2.13.9 Key Conclusion

Altruism is not the expansion of concern over distance, but ascent toward the least upper bound in the lattice of cognition. The entire system—physical, cognitive, and social—follows the same underlying latent potential structure, as illustrated in Figs. 2–5: stability is obtained through descent in potential, and generality and inclusion are obtained through ascent in abstraction.

2.14 Unified Convergence Theorems

In this section, we explicitly unify these structures in terms of convergence theorems.

The preceding sections established the structure of the Tomabechi framework at three levels: individual stability (TCZ), social stability (Shared-TCZ), and abstract structure (lattice / LUB). These can now be unified through a set of convergence theorems based on a common Lyapunov structure.

2.14.1 Theorem 1 (Possible Worlds–Ego–TCZ)

Statement. If the Self operator is implemented as $\pi_c(x) = \arg\min_u \int_0^T V(x(t),t)\, dt$, then $x^*(t) \to$ TCZ. **Interpretation:** Self (semantic ordering of possible worlds), Ego (optimal control of trajectories), TCZ (stability region in state space).

Sketch of Proof. The Self operator induces a preference structure over possible worlds. The Ego minimizes accumulated instability, and the constraint $V(x,t) \leq \theta$ defines a stable region. Therefore, trajectories converge into the TCZ. A rigorous proof is given in the Appendix.

2.14.2 Theorem 2 (Shared-TCZ Convergence)

Statement. With $\pi_i = \arg\min_{u_i} \int_0^T (V_i + \lambda_i S_i)\, dt$: $\mathbf{x}^*(t) \to \mathrm{TCZ}_{\mathrm{shared}}$.

Sketch of Proof. Define composite Lyapunov function $\mathscr{L}(\mathbf{x}) = \Sigma_i\, [\ V_i(x_i) + \lambda_i S_i(\mathbf{x})\ ]$. Its decrease ensures convergence to a shared low-instability region.

2.14.3 Theorem 3 (Abstract Shared-TCZ Convergence)

Statement. With abstraction potential A: $\mathbf{x}^*(t) \to \mathrm{LUB}(W_1,\ldots,W_N)$.

Sketch of Proof. Introduce $\mathscr{L}_A(\mathbf{x}) = \Sigma_i\, [\ V_i(x_i) + \lambda_i S_i(\mathbf{x}) + \eta_i A(x_i)\ ]$. Its decrease drives the system upward in the lattice structure toward LUB. The rigorous proof is given in the Appendix.

2.14.4 Unified Interpretation

Dual Principle

System dynamics are governed by a unified dual structure:

- descent in potential → stability
- ascent in abstraction → structure

Thus, physical, cognitive, and social systems are unified under a single dynamical–structural framework.

3. Structural Correspondence with Einstein

In Einstein's theory, molecular potential U accumulates across space and manifests as surface tension and capillary phenomena. In the Tomabechi framework, evaluative function $V(x, t)$ accumulates across time and manifests as behavior and cognitive change.

Physics (Einstein)		Cognition (Tomabechi)
Molecular potential U	↔	Evaluative function $V(x,t)$
Surface tension	↔	Cognitive energy
Capillary phenomena	↔	Behavior / cognitive change
Micro → Macro	↔	Unconscious → Behavior

4. Spatial Integration (Physics)

In physical systems, the sum of local interactions is expressed as:

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

Macroscopic phenomena arise from the spatially distributed accumulation of interactions, not from any single local action.

Plain language: No single molecule determines the behavior of a liquid. What matters is the cumulative effect of all molecules interacting across the entire volume. Surface tension—the force that allows water to bead on a leaf—is not caused by one molecule pushing another; it is an emergent property of the whole system's internal structure summed across space. Einstein's insight was that you cannot understand macroscopic phenomena by looking at one point; you must look at the whole accumulated field.

5. Boundary Effects

Observable physical phenomena emerge where internal interactions become asymmetric at the boundary:

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

The effect of internal structure is made visible at the boundary.

Plain language: The internal structure of a liquid is not directly visible. But at the surface—the boundary between liquid and air—the asymmetry of internal forces becomes observable as surface tension. What we see is not the internal structure itself, but its signature at the boundary. In the same way, observable human behavior is not the internal cognitive structure itself—it is the signature of that structure where it meets the external world.

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

The effect of internal structure is made visible at the boundary. (Note: here γ_{surface} denotes the physical surface tension coefficient, not the social coupling coefficient γ_{ij}.)

6. Temporal Integration (Cognition)

In cognition, evaluations accumulate in the time direction:

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

Behavior is the result of this temporal accumulation. The difference between physics and cognition is not whether accumulation occurs, but where it occurs.

Plain language: Just as a liquid's macroscopic properties emerge from the spatial sum of molecular interactions, a person's behavior emerges from the temporal sum of their cognitive evaluations—their judgments, feelings, memories, and expectations accumulated over time. No single moment determines behavior; it is the accumulated weight of an entire cognitive history. This is why trauma has lasting effects, why trust is built slowly, and why belief systems are resistant to change. The time integral of evaluations is what makes cognition structural rather than merely reactive.

7. Unified Accumulation Principle

Thus far, we have observed that physical systems accumulate interactions over space, while cognitive systems accumulate evaluations over time.

This structural correspondence is illustrated in Figure 2, which provides an intuitive view of the unified accumulation principle.

Accordingly, the two can be unified as follows:

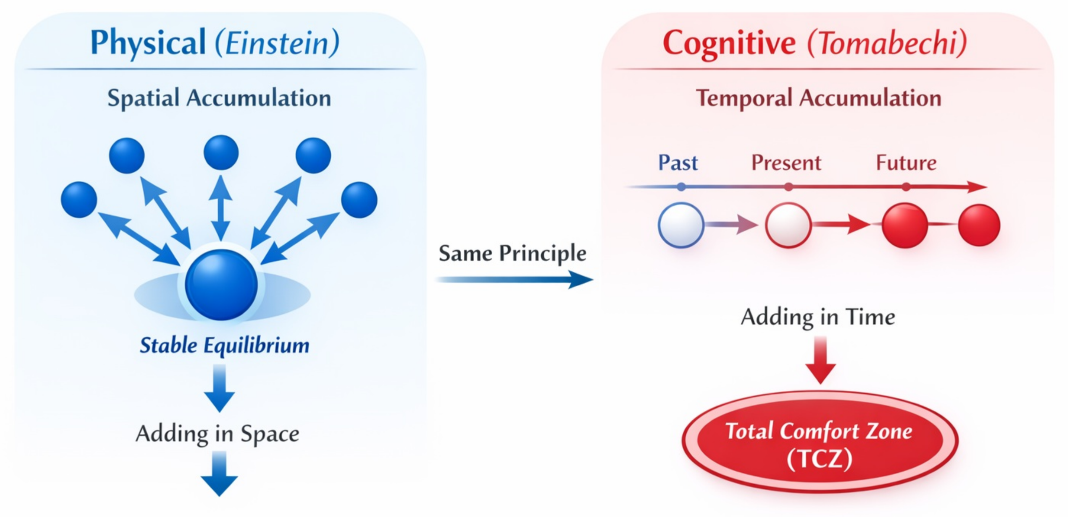

Figure 2: Structural correspondence between spatial accumulation in physical systems (Einstein) and temporal accumulation in cognitive systems (Tomabechi). This diagram provides an intuitive illustration of the unified accumulation principle prior to its formal mathematical expression.

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

In this formulation, observable phenomena arise as the accumulation of latent potentials over their respective domains. In physical systems, the domain is spatial, whereas in cognitive systems, it is temporal.

Thus, the difference between physics and cognition lies not in the principle of accumulation itself, but in the domain over which accumulation occurs.

8. Unified Accumulation Structures

The unified accumulation principle introduced in the previous section can be understood more concretely from the perspective of dynamical systems. Figure 2 provides an intuitive correspondence, while Fig. 3 presents a more detailed structural representation of the accumulation dynamics in both domains. In physical systems, local interactions accumulate over space to form an attractor basin. In cognitive systems, evaluations accumulate over time, leading to convergence toward a Total Comfort Zone (TCZ). Thus, both systems exhibit the same underlying structure: the accumulation of latent potentials leading to stable attractors.

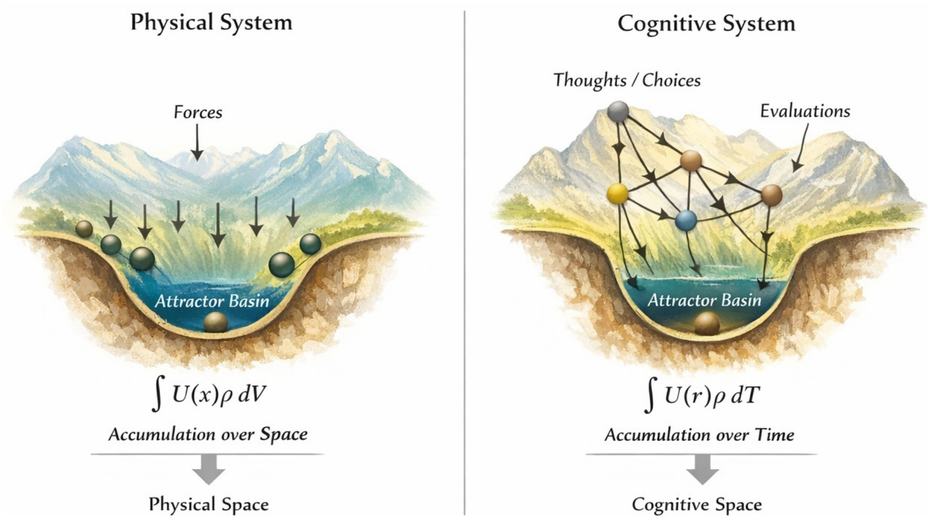

Figure 3: Unified accumulation structures: physical vs. cognitive systems. In physical systems, local interactions defined by a molecular potential $U(x)$ accumulate over space, producing convergence toward an attractor basin via spatial integration $\int U(x)\rho\, \mathrm{d}V$. In cognitive systems, evaluations of thoughts and choices accumulate over time, producing convergence toward a cognitive attractor basin (TCZ) via temporal integration $\int_0^T V(x(t), t)\, \mathrm{d}t$. The essential structural equivalence lies in the domain of accumulation: space in physics, time in cognition.

9. Figure 4 & 5: TCZ Structure, Abstraction Hierarchy, and Dual Decrease Principle

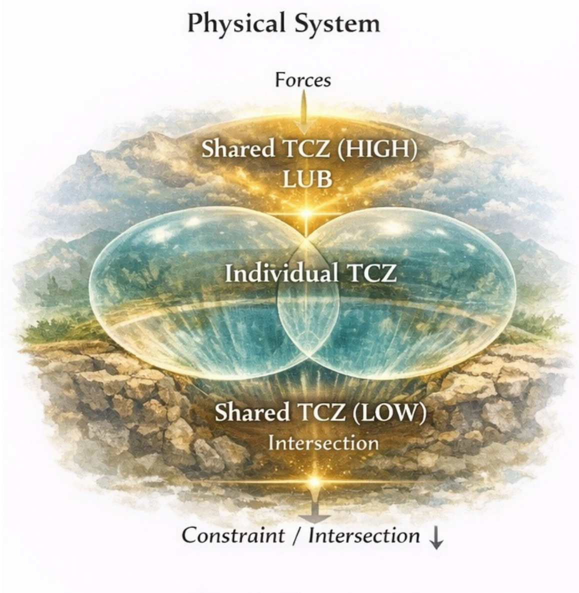

Figure 4: TCZ Structure
and Abstraction Hierarchy

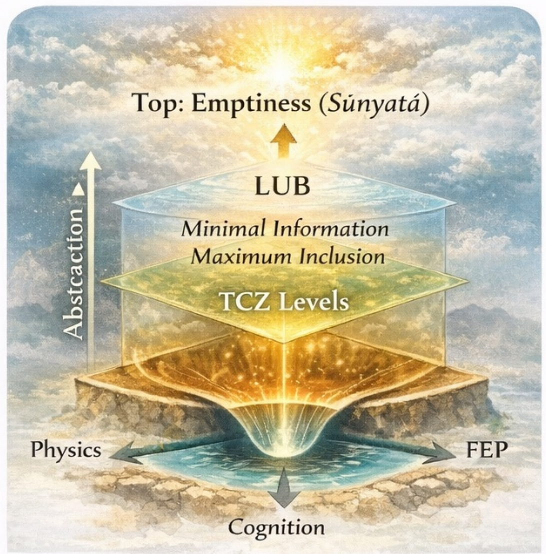

Figure 5: Dual Structure of Decrease
and Abstraction

Figure 4 (left): TCZ structure and abstraction hierarchy. The lower shared TCZ (intersection-based) represents constraint-compatible coexistence. The individual TCZ represents self-stability. The higher shared TCZ (LUB-based) represents abstraction-based integration across agents, moving toward least upper bounds (LUB) with minimal information and maximum inclusion.

Figure 5 (right): Dual structure of decrease and abstraction. The downward axis represents the decrease principle—dynamics in physics, the Tomabechi framework (TCZ), the Free Energy Principle (FEP), and predictive processing all converge toward attractors via scalar reduction. The upward axis represents the abstraction principle—movement toward LUB, information decrease, entropy decrease, culminating in emptiness (Śūnyatā) at the top element of the subsumption lattice.

10. Unified Decrease Principle

The following phenomena in different theoretical domains are expressions of the same underlying structure:

- Physics: energy decreases
- Tomabechi framework: evaluative function decreases
- FEP: free energy decreases
- Predictive processing: prediction error decreases

In each case, a scalar quantity representing deviation from internal constraints decreases, driving the system toward a stable region.

11. Entropy and Learning

In physical space, entropy increases over time due to local interactions. In cognitive and information space, memory, learning, and inheritance allow information to accumulate across time, enabling entropy to decrease. Learning is therefore:

the accumulation of information across time leading to structure formation

The same duality appears in neural systems: local neural interactions follow entropy increase, while whole-brain learning processes produce order. In AI, learning structures the representation space through temporal information accumulation.

12. Free Energy Principle and Predictive Processing

The Free Energy Principle holds that cognitive systems minimize free energy to reduce surprise and uncertainty. Predictive processing holds that the brain minimizes prediction error to achieve coherence between perception and action. Both fit precisely within the unified decrease structure: they are specializations of the principle that a scalar deviation from internal constraints decreases as the system moves toward its stable region.

13. Cognitive Warfare

What is cognitive warfare? In conventional warfare, the goal is to destroy the enemy's physical capacity to act—weapons, infrastructure, personnel. In cognitive warfare, the goal is different: it is to alter the enemy's *will* and *perception* in such a way that their behavior changes without physical force. But the Tomabechi framework reveals that this description, while correct, is still too shallow. Cognitive warfare is not simply about changing opinions or spreading misinformation. At its deepest level, it is about restructuring the internal potential landscape of an entire population—changing what people feel is stable, comfortable, and normal, so that their cognitive trajectories converge toward states that serve the adversary's strategic objectives. In the language of this framework: cognitive warfare is the external manipulation of the evaluative function $V(x,t)$ that governs how individuals and populations navigate their cognitive space.

The essence of cognitive warfare is not to command behavior directly. It is to alter the evaluative function itself—to reshape the potential landscape:

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

This changes the trajectory selected by the Self, changes the TCZ, and thereby changes behavior. Cognitive warfare is therefore the manipulation of potential structure. The operational realization of this framework is illustrated in Fig. 1, while its underlying structural principles are shown in Figs. 2–5.

13.1 Structural Position of AI Outputs, Hallucination, and Art

The fundamental distinction between local solutions and higher-order abstraction is illustrated in Fig. 6.

How to read Figure 6: Imagine a vast, crumpled landscape of hills and valleys—but in a million-dimensional space rather than three. AI systems find their way to the bottom of whatever valley they start near. The red point is a "valley" that corresponds to incoherent, unacceptable output—a hallucination. The blue point is a "valley" that corresponds to output that seems good, but is still only locally optimal—it happens to look reasonable, but there is no guarantee it reflects any deeper truth or structure. The green region at the top represents something entirely different: art and high-level abstraction are not found at the bottom of any valley. They exist at the upper regions—the least upper bounds of the conceptual lattice, positions of maximal inclusion and integration. The key insight is structural: AI can never produce genuine art by error minimization alone, because art is not a local minimum—it is a different kind of structure altogether.

Outputs of generative AI can be understood as local solutions of an error-minimization process defined over an ultra-high-dimensional functional space. That is, generative models operate by descending along gradients of a high-dimensional loss landscape, converging toward locally optimal configurations.

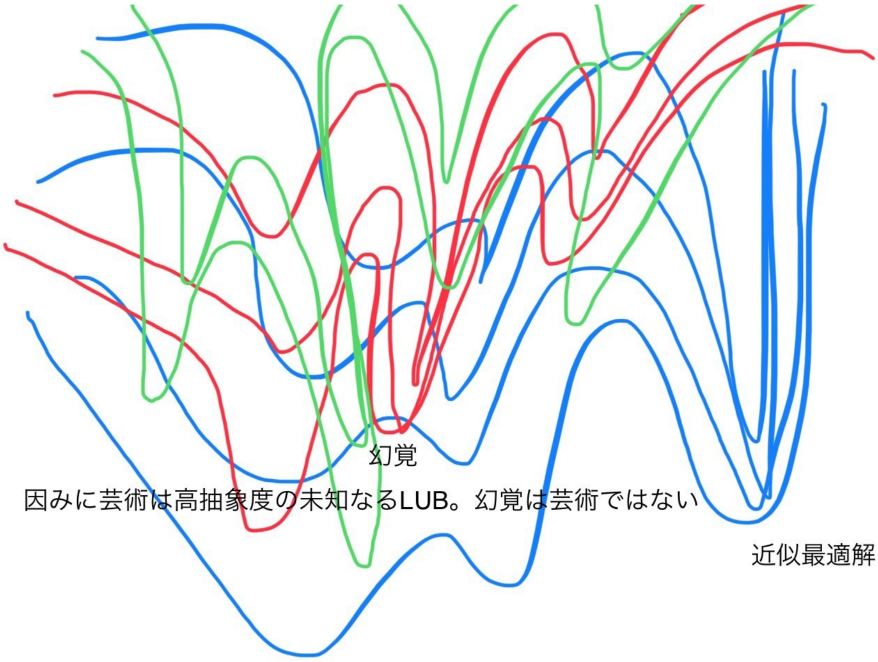

Figure 6: Structure of local solutions in an ultra-high-dimensional error function. Multiple local minima exist, and even the approximate optimal solution (blue, lower-right) is only a local solution. In this illustration, one of the local minima (red, lower region) is shown as a hallucination, i.e., an output that is not acceptable to human cognition. ``` By contrast, art corresponds to a structure at a high level of abstraction and is understood as an as-yet undiscovered least upper bound (LUB). In this figure, for example, the upper region in green represents one such instance. Hallucinations arise as local solutions in lower regions of the solution space and are structurally distinct from art. Therefore, treating hallucinations as art is theoretically incorrect, and the two must be rigorously distinguished at the structural level. ```

Within this framework, some local solutions correspond to outputs that are acceptable or coherent from a human perspective, while others correspond to semantically inconsistent or structurally invalid configurations. The latter are commonly referred to as hallucinations.

There have been arguments that such hallucinated outputs—particularly in image generation—should be regarded as a form of art. This interpretation is fundamentally mistaken. Local solutions are merely the result of error minimization and represent low-level structures projected into a pseudo-physical space.

In contrast, art corresponds to high-level abstraction. Within the present framework, it is more appropriately understood as a structure associated with a least upper bound (LUB) that has not yet been discovered.

Furthermore, since the essence of TCZ-High-Shared corresponds to altruism, art operates toward integration, coherence, and stabilization, and therefore tends to function in a direction that avoids conflict. By contrast, pseudo-art generated from hallucinated outputs belongs structurally to the TCZ-Low-Shared regime and may amplify inconsistency and fragmentation.

This issue is not limited to art. The uncritical deployment of inductive AI systems—without a deductive mathematical control structure—into cognitive warfare frameworks presents a fundamental strategic risk.

Figure 6 depicts the structure in two dimensions for illustrative purposes. However, the actual problem is defined over extremely high-dimensional spaces (e.g., on the order of 10^5 to 10^6 dimensions). Accordingly, what appears as an "approximate optimal solution" is, in general, only a local solution, with no guarantee of global optimality.

Thus, art should be understood not as a local solution, but as a structural realization of an as-yet undiscovered least upper bound (LUB) in the space of abstraction.

In contrast, art corresponds to high-level abstraction—a structure associated with a least upper bound (LUB) that has not yet been discovered. Furthermore, since the essence of TCZ-High-Shared corresponds to altruism, art operates toward integration, coherence, and stabilization, and therefore tends to function in a direction that avoids conflict. By contrast, pseudo-art generated from hallucinated outputs belongs structurally to the TCZ-Low-Shared regime and may amplify inconsistency and fragmentation.

13.2 Adversarial Vulnerability of AI Systems in Cognitive Warfare Contexts

A fundamental security concern arises when AI systems are deployed in cognitive warfare environments. Because contemporary generative AI systems operate through inductive learning—constructing internal representations from statistical patterns in training data—they are structurally incapable of self-detecting tampering with their own training data or internal mathematical state. A cyber-adversary who modifies the training corpus or weight parameters of an operational AI system can redirect its generative trajectories without producing detectable error signals; the system will continue to operate with apparent coherence while generating outputs aligned with adversarial objectives. Moreover, errors and hallucinations—outputs diverging from intended vector directions—may occur independently of deliberate attack, further complicating detection.

A secondary attack surface exists at the output layer: generated text, imagery, or decision recommendations may be intercepted and altered after leaving the AI system but before reaching human operators, with no cryptographic guarantee of integrity. It must therefore be assumed that cyber-attacks targeting AI systems themselves, their training pipelines, their inference infrastructure, and their output channels will occur.

This consideration makes the introduction of a deductive and integrative mathematical foundation—of the kind proposed in this paper—not merely advantageous but operationally essential. Only a formally verifiable control structure can provide guarantees that inductive systems, by their nature, cannot. AI can contribute to peace and stability if and only if it operates within such a deductive framework; without it, the same systems become instruments of cognitive warfare.

13.3 AI-Accelerated Code Proliferation and the Gödel–Chaitin Limit

(1) **The scale problem.** The adoption of AI extends beyond the generation of messages and images used in cognitive warfare. It is straightforward to anticipate an order-of-magnitude increase in AI-generated control code across cognitive, cyber, and physical weapons systems alike.

(2) **The self-attack risk.** As a direct consequence, risky behaviors — including self-attacks that human experts cannot verify — are expected to increase. The verification bottleneck is structural: AI can write code faster than any human team can audit it.

(3) **The democratization of weaponized code.** With the proliferation of AI-assisted automated programming, it is readily foreseeable that control code for both cognitive and physical warfare will be produced at exponential scale — by personnel without formal training in computer science. The barrier to producing weaponized software collapses.

(4) **The Gödel–Chaitin limit.** The operational integrity of the code itself cannot be guaranteed by current inductive AI generation. More fundamentally, this impossibility is not a matter of engineering effort but a mathematical certainty: **Gödel's incompleteness theorems** establish that sufficiently expressive formal systems cannot prove their own consistency, and **Chaitin's extension via algorithmic information theory** shows that no formal system can prove truths whose complexity exceeds its own. This follows because the **non-computability of Chaitin's constant Ω** — together with the impossibility of capturing algorithmic randomness by any finite axiomatization — makes the **formal-system limitation theorem (Chaitin 1974, 1975)** directly applicable to AI-generated code: an axiom system T can prove only finitely many statements of the form "$K(s) \geq n$," and becomes unable to prove any such statement once n exceeds T's own complexity. When the Kolmogorov complexity of LLM-generated code exceeds the complexity of the formal verifier (any verifier human engineers can construct), verification becomes impossible in principle. This bound becomes acute as AI-

generated code surpasses human-scale formal systems in sheer algorithmic complexity. This constitutes a major structural risk across all future cognitive, cyber, and physical warfare — one that cannot be addressed except by elevating the abstraction level and governing the system through a deductive mathematical framework.

Cyber-attacks on AI are not a future risk — they are a present strategic reality. A deductive mathematical foundation is the only structural guarantee against AI being turned into a weapon of cognitive warfare.

Security Interpretation

From a security perspective, the multi-bridge ensemble also provides robustness against adversarial manipulation.

Because inductive AI systems can be compromised at the level of training data, model parameters, or output channels, a single generated trajectory cannot be trusted as a stable control signal.

In contrast, an ensemble structure introduces redundancy and structural consistency constraints.

An adversarial perturbation that affects one component does not necessarily propagate to the LUB of the ensemble.

Thus, the multi-bridge structure acts as a form of structural error-correction, ensuring that the overall direction remains aligned with the intended high-level abstraction.

In this sense, the multi-bridge ensemble is not only an operational mechanism, but also a security architecture for cognitive warfare systems.

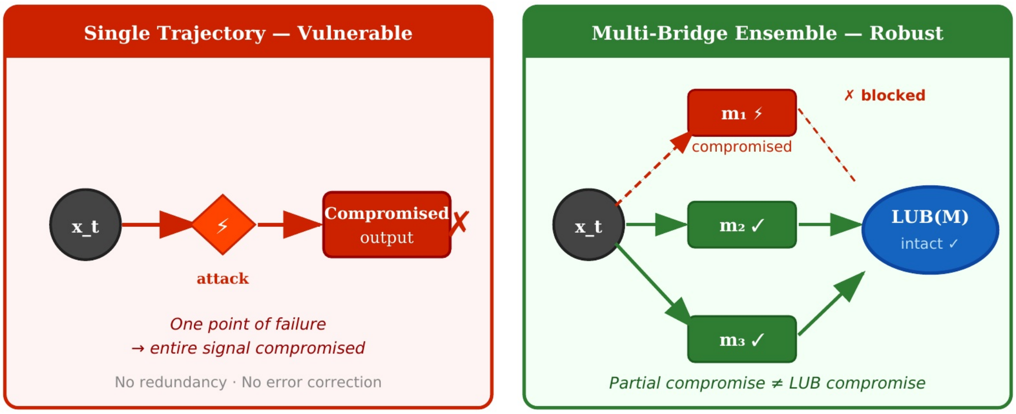

Figure A.3-S: Security comparison. *Left*: A single generated trajectory provides a single point of failure — adversarial compromise propagates directly to the output. *Right*: The multi-bridge ensemble introduces structural redundancy. Even if one bridge message (m_1) is compromised, the LUB computed from the remaining coherent bridges (m_2, m_3) remains aligned with the intended abstraction target. The ensemble thus functions as a structural error-correction mechanism.

14. Strategic Implications

This section addresses practical implications for strategy, policy, and education.

The theoretical synthesis is reserved for Section 15.

The central implication of this paper is that behavior itself is not the primary object of control; rather, the underlying latent potential structures that generate behavior must be understood as the fundamental target. Strategy, therefore, should not be conceived as the direct manipulation of actions, but as the design and transformation of the internal structures from which those actions emerge.

From this perspective, modern conflict must be redefined not as physical confrontation, but as the control of cognitive potential structures. Competition in the information domain is not fundamentally about manipulating individual messages or actions, but about shaping and steering the cognitive structures of populations. In this sense, conflict can be interpreted as a manifestation of low-level shared cognitive structures, where instability and fragmentation are amplified.

As demonstrated in this paper, outputs generated by AI systems are inherently local solutions and may include structurally inconsistent states referred to as hallucinations. Confusing such outputs with higher-order abstract structures is theoretically incorrect, and this misinterpretation introduces significant risks in cognitive warfare. The deployment of AI systems without a deductive mathematical control framework may undermine strategic coherence and lead to unpredictable outcomes.

More fundamentally, this issue is not purely technical—it manifests at the level of social structure and education. One reason why conflicts persist globally can be interpreted as a structural tendency for leaders associated with TCZ-Low-Shared cognitive regimes to be selected. Accordingly, the implementation of TCZ-High-Shared cognitive structures at the societal level becomes a critical challenge—not merely an issue of institutional design, but fundamentally an issue of education. In particular, education in the arts and liberal arts plays an essential role in promoting higher levels of abstraction and integrative structures.

Taken together, these considerations suggest that strategy, policy, and national security must shift their focus from controlling actions to controlling structures. This paper provides the mathematical foundation for such a transition.

More fundamentally, one reason why conflicts persist globally can be interpreted as a structural tendency for leaders associated with TCZ-Low-Shared cognitive regimes to be selected. Accordingly, the implementation of TCZ-High-Shared cognitive structures at the societal

level becomes a critical challenge—not merely an issue of institutional design, but fundamentally an issue of education. In particular, education in the arts and liberal arts plays an essential role in promoting higher levels of abstraction and integrative structures.

15. Conclusion

This section provides the theoretical synthesis. For practical implications, see Section 14.

This paper has demonstrated that Einstein's capillarity theory and the Tomabechi framework share a deep structural isomorphism. In both cases, observable phenomena arise from the accumulation of latent potentials—across space in physics, and across time in cognition. This shared skeleton unifies, within a single mathematical framework, the decrease principle, entropy, learning, the Free Energy Principle, prediction error, and cognitive warfare.

Within this framework, physical, cognitive, and social phenomena are instances of a single underlying principle: systems converge toward stability through the reduction of instability, while simultaneously evolving toward integration through increasing abstraction. This dual structure—descent in potential and ascent in abstraction—constitutes a general law governing complex systems.

At the highest level, this unified structure converges toward a limiting abstraction corresponding to emptiness (śūnyatā). The ultimate form of stability and integration is not a particular state, but a structural condition that enables both existence and non-existence as expressions of a common underlying system.

Ultimately, the problem is not the control of actions, but the control of structures.

The paper also develops three operational extensions. At the individual level (A.1), cognitive warfare is formalized as the external modification of the evaluative function $V(x,t)$, thereby redefining the TCZ toward which behavioral trajectories converge. At the boundary level (A.2), optimal control is shown to be most efficiently achieved near the boundary of the TCZ, where sensitivity to perturbation is maximal. At the generative level (A.3), effective information operations are formalized as multi-bridge message ensembles: sets of individually acceptable messages whose least upper bound (LUB) aligns with a higher-level Shared-TCZ, implementing Theorem 3 at the level of information generation. The future battlespace is not physical terrain, but the structure of cognitive potential landscapes.

Accordingly, effective policy and strategy must shift from controlling actions to controlling structures. This paper provides the mathematical foundation for that transition.

Appendix A (Civilian Overview)

Operational and Control-Theoretic Interpretation of Cognitive Warfare

The following provides an operational and control-theoretic interpretation of the theoretical framework presented in this paper.
This section is a conceptual overview for general audiences; operational classified content has been omitted.

A.1 Operational Interpretation of Cognitive Warfare

Cognitive warfare is the process of transforming the evaluative function and latent potential structure of a target population, thereby reconfiguring its Total Comfort Zone (TCZ) and altering behavioral trajectories.

Theorem-Based Interpretation

Theorem 1 (Individual Level)

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

$$x^*(t) \rightarrow TCZ$$

Operationally:
The evaluative function $V(x,t)$ determines the cost assigned to each cognitive state x at time t. By modifying V — for example, by increasing the perceived threat, instability, or dissonance associated with certain states — the shape of the potential landscape is altered. As a result, the optimal trajectory $x^*(t)$, which minimizes the accumulated cost $\int V\ \mathrm{d}t$, converges toward a different TCZ. Concretely: if the evaluative cost of a target state is raised, the individual's cognitive trajectory will naturally avoid it; if the cost of a desired state is lowered, the trajectory will be drawn toward it. Cognitive warfare thus operates by sculpting the potential landscape $V(x,t)$ so that the TCZ — the attractor region — is positioned where the adversary intends.

Theorem 2 (Social Level)

$$\pi_i = arg\ min_{u_i(t)} \int_0{}^T (\ V_i(x_i(t), t) + \lambda_i\ S_i(x(t))\ )\ dt$$

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

Operationally:

Constraining shared stability structure narrows the system toward low-level Shared-TCZ, amplifying fragmentation and instability.

Theorem 3 (Abstract Level)

$$\pi_i = arg\ min_{u_i(t)} \int_0^T \left(V_i(x_i(t), t) + \lambda_i S_i(x(t)) + \eta_i A(x_i(t)) \right) dt$$

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

Operationally:
Increasing abstraction enables convergence toward high-level Shared-TCZ, producing integration, coherence, and stability.

Key Insight

Cognitive warfare operates not at the level of actions, but at the level of structures that generate actions.

Non-Technical Interpretation

People do not act randomly. Their behavior converges toward a perceived stable region (TCZ).

Cognitive warfare reshapes what is perceived as stable, acceptable, or meaningful.

At the individual level, it changes how people evaluate the world.
At the social level, it reshapes shared perception.
At the abstract level, it changes how broadly people define their world.

Lower abstraction produces conflict.
Higher abstraction produces integration.

In simple terms, cognitive warfare is not about telling people what to do.
It is about shaping the landscape in which decisions are made.

A.2 Boundary-Based Control of TCZ

The operational interpretation above can be further refined at the level of control implementation.

Effective cognitive warfare does not require forcing the target deeply into a TCZ.
Instead, control is most efficiently achieved near the boundary of the TCZ.

The boundary is defined as:

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

Control is applied to its neighborhood:

$$\mathcal{N}_\varepsilon(\partial TCZ) = \bigcup_{t\geq 0} \{ x(t) \in \mathscr{R}(t;x_0) \mid |V(x(t),t) - \theta| \leq \varepsilon \}$$

Trajectories near the boundary are highly sensitive to small perturbations.

Thus:

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

This expresses boundary-level control with minimal intervention. Note that the cost term $|V(x(t),t) - \theta|^2$ is symmetric around the boundary: it attracts trajectories toward $V = \theta$ from both inside (TCZ interior) and outside (TCZ exterior). This is intentional — the objective is not to stabilize within the TCZ, but to maintain the system near the boundary where sensitivity is maximal. In practice, the term $\lambda C(u(t))$ regulates intervention cost, ensuring that boundary-level influence is achieved with minimal effort.

Non-Technical Interpretation (Boundary Control)

The most effective influence happens not deep inside stability, but at its edge.

At the boundary, small changes produce large effects.

This is precisely the structure underlying Einstein's 1901 capillarity theory: surface tension — the visible macroscopic effect — arises not from the interior of the liquid, but from the asymmetry of molecular forces at the boundary between liquid and air. A small structural difference at the interface produces a large observable phenomenon. Cognitive warfare operates by the same principle: intervention at the cognitive boundary — where internal evaluative forces become asymmetric — produces disproportionately large behavioral effects.

Thus, cognitive warfare is most efficient when it operates where the system is most sensitive.

Closing Insight

This is the operational core of cognitive warfare.

A.3 Multi-Bridge Message Ensemble for High-Abstraction

Convergence

The operational framework above can be further extended to the generation of information outputs (e.g., text and images) in cognitive warfare systems.

A single generated output whose semantic direction lies far from the target's current TCZ center is unlikely to be accepted. Formally, let x_t denote the current cognitive state, TCZ(x_t) the corresponding stability region, and $\varphi(m)$ the semantic embedding of a message m. The acceptance of a message can be modeled as:

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

where $d(\cdot,\cdot)$ measures distance to the TCZ and σ controls the width of the acceptance band. Messages that are too distant from the TCZ are thus exponentially suppressed in acceptance.

Accordingly, effective generation should not rely on a single directional message, but on a set of messages $M_t = \{\ m_1, \ldots, m_K\ \}$ that are individually acceptable yet collectively induce movement toward a higher-level shared structure.

We formalize this as a message ensemble optimization:

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

Here:

- $\Sigma\ A(m_k \mid x_t)$ ensures individual acceptability.
- Lift(M_t) promotes ascent toward higher abstraction:

  Lift$(M_t) = d($ LUB$(\ \varphi(m_1), \ldots, \varphi(m_K)\ ),$ LUB$_{\text{target}}\ )^{-1}$
- Frag(M_t) penalizes internal inconsistency among messages:

  Frag$(M_t) = \Sigma_{i<j}$ Incoh$(\ \varphi(m_i), \varphi(m_j)\ )$
- Decept(M_t) penalizes deceptive or manipulative constructs.

Crucially, the objective is not to present mutually contradictory messages, but to construct a set whose least upper bound (LUB) aligns with a higher-level Shared-TCZ. In this sense, multiple messages function as "bridges" that are individually reachable from the current TCZ while collectively lifting the system toward a more abstract and inclusive structure.

From a dynamical viewpoint, this ensemble approach implements Theorem 3 at the level of information generation: rather than forcing a trajectory, it reshapes the attractor landscape so that trajectories naturally converge toward high-abstraction shared regions.

Thus, cognitive warfare systems that aim to contribute to stability should be designed as LUB-guided message ensembles that maximize acceptability, promote abstraction, maintain coherence, and avoid deception.

This constitutes a multi-bridge control mechanism for high-level cognitive integration.

Non-Technical Interpretation (Multi-Bridge Ensemble)

A single message that is far from what a person already accepts is unlikely to be received. It falls outside the person's Total Comfort Zone (TCZ) and is therefore rejected.

The most effective approach is not to force a large leap, but to construct a set of messages that are each individually acceptable, while collectively pointing toward a more integrated and higher-level understanding.

This is analogous to building a staircase rather than attempting a jump. Each step is within reach, yet the sequence of steps leads to a position far above the starting point—toward a broader and more inclusive worldview.

This is the principle of the multi-bridge ensemble. Instead of a single strong push, a carefully designed set of individually acceptable messages shifts the cognitive landscape through their combined direction.

Crucially, the messages must be mutually coherent. A set of inconsistent messages does not lift the system—it fragments it.

The design principle is therefore: each message is a bridge, and bridges must connect.

In formal terms, each message remains locally acceptable, while the least upper bound (LUB) of the ensemble aligns with a higher-level shared structure.

In this sense, multi-bridge ensembles do not impose change; they construct pathways along which change becomes natural.

Ultimately, this mechanism reflects a general principle: effective cognitive influence is achieved not by force, but by shaping the structure of transitions.

Thus, multi-bridge ensembles constitute a fundamental mechanism for high-abstraction cognitive integration.

Summary

Cognitive warfare is the process of transforming the evaluative function and potential structure of a target population, thereby reconfiguring its TCZ (Total Comfort Zone) and altering behavioral trajectories as a result.

That is: cognitive warfare is not about changing behavior directly — it is about controlling the TCZ itself, toward which behavior converges.

Theorem 1	Rewrite the TCZ
Theorem 2	Confine to Shared-low-TCZ
Theorem 3	Elevate abstraction
A.2 (Boundary)	Control at the boundary — where sensitivity is highest
A.3 (Ensemble)	Multi-bridge message ensemble — LUB-guided abstraction lift

Theorem 1: by altering the evaluative function, the individual's TCZ is redefined.
Theorem 2: by constraining the shared structure to a low level, fragmentation is created.
Theorem 3: by elevating abstraction, integration is achieved.
A.2: the most efficient intervention is not deep inside the TCZ, but at its boundary — where small changes produce large effects.

Cognitive Warfare Based on Theorem 1 (TCZ Convergence):
By transforming the evaluative function $V(x,t)$ of the target, the TCZ toward which the optimal trajectory $x^*(t)$ converges is itself redefined.

Cognitive Warfare Based on Theorem 2 (Shared-TCZ):
By controlling the shared stability structure among agents, the Shared-TCZ is confined to a low level of abstraction (Shared-low-TCZ), amplifying cognitive fragmentation and instability.

Cognitive Warfare Based on Theorem 3 (Abstract Shared-TCZ):
By elevating abstraction, a higher-order shared structure (Shared-high-TCZ) is formed, realizing integration, stability, and altruistic structure across agents.

Cognitive Warfare Based on A.2 (Boundary Control):
The most effective cognitive influence occurs not by forcing the target deep into a desired state, but by operating near the boundary of its TCZ — where the system is most sensitive and small perturbations produce large behavioral shifts. Minimal-cost boundary control is expressed as the optimization of $u^*_{\mathrm{cw}}(t)$.

Cognitive Warfare Based on A.3 (Multi-Bridge Ensemble):

A single message that is too distant from the target's TCZ will be rejected. Instead, design a set of individually acceptable messages $M_t = \{m_1,\ldots,m_K\}$ whose least upper bound (LUB) aligns with a higher-level Shared-TCZ. Each message is a bridge; the ensemble lifts the cognitive landscape toward greater abstraction and coherence without forcing any single trajectory. This implements Theorem 3 at the level of information generation.

In essence, the core idea is one.

> ***Cognitive warfare does not control actions. It controls the structures from which actions emerge — most efficiently at the boundary where structures are most sensitive, and most coherently through ensembles of individually acceptable messages that collectively lift the cognitive landscape toward higher abstraction.***

Appendix B: Unified Lyapunov Framework for Theorems 1–3

(*This Appendix provides the rigorous mathematical foundation for Theorems 1–3 in Section 2.*)

—Theorem 1 (individual TCZ convergence), Theorem 2 (shared TCZ convergence), and Theorem 3 (abstract shared TCZ convergence)—are not independent results, but successive extensions of a single Lyapunov-based stability structure. Theorem 1 describes stabilization at the individual level. Theorem 2 extends this to multi-agent systems through shared consistency. Theorem 3 further incorporates abstraction, lifting convergence toward least upper bounds (LUB) in a lattice structure.

A. Unified Lyapunov Setup

Consider a closed-loop system $\dot{z} = F(z)$, where z represents (depending on context) an individual state x, a joint state $\mathbf{x}$, or an abstract shared state. Let $\Phi(z)$ be continuously differentiable and define $\Omega_\theta := \{ z \mid \Phi(z) \leq \theta \}$.

Assume $\nabla\Phi(z) \cdot F(z) \leq -\alpha(\Phi(z) - \theta)$, $\alpha > 0$, for all $z \notin \Omega_\theta$. Then $\mathrm{dist}(z(t), \Omega_\theta) \to 0$. If $\nabla\Phi \cdot F \leq 0$ on $\partial\Omega_\theta$, then Ω_θ is forward invariant. We specialize to: $\Phi = V$, $\Phi = \mathscr{L}$, $\Phi = \mathscr{L}_A$.

B. Theorem 1 (Individual TCZ Convergence)

Statement. Let $\Phi(x) = V(x)$. Then $x^*(t) \to$ TCZ.

Proof. Define $y(t) = V(x(t)) - \theta$. Then $\dot{y}(t) \leq -\alpha y(t)$. By comparison principle:

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

Thus $x(t)$ approaches TCZ_θ. Forward invariance: $d/dt\, V \leq 0$ on $\partial\mathrm{TCZ}_\theta$. ■

Interpretation. Stability arises from the decrease of internal evaluative instability.

C. Theorem 2 (Shared-TCZ Convergence: Rigorous Form)

Setup. The sharing function $S : \mathbf{X} \to \mathbb{R}_{\geq 0}$ satisfies: (S1) $S(\mathbf{x}) \geq 0$ for all $\mathbf{x}$; (S2) $S(\mathbf{x}) = 0$ if and only if $\mathbf{x} \in \mathrm{TCZ}_{\mathrm{shared}}$. The composite Lyapunov function $\mathscr{L}(\mathbf{x}) = \Sigma_i\, V_i(x_i) + \lambda S(\mathbf{x})$ with $\lambda > 0$ is positive definite relative to $\mathrm{TCZ}_{\mathrm{shared}}$.

Statement. Suppose $\nabla\mathscr{L}(\mathbf{x}) \cdot F(\mathbf{x}) \leq -\alpha(\mathscr{L}(\mathbf{x}) - \theta)$ for all $\mathbf{x} \notin \Omega_\theta$. Then $\mathbf{x}^*(t) \to \mathrm{TCZ}_{\mathrm{shared}}$.

Proof. Define $y(t) = \mathscr{L}(\mathbf{x}(t)) - \theta$. Then $\dot{y}(t) \leq -\alpha y(t)$. By comparison principle:

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

Hence $\text{dist}(\mathbf{x}(t), \Omega_\theta) \to 0$. Forward invariance of Ω_θ follows from $\nabla\mathscr{L} \cdot F \leq 0$ on $\partial\Omega_\theta$. By LaSalle's principle, trajectories converge to the invariant set where $\mathscr{L} = \theta$, which by (S2) implies $S(\mathbf{x}) = 0$, i.e., $\mathbf{x} \in \text{TCZ}_{\text{shared}}$. ■

Interpretation. Stability extends to multi-agent systems through simultaneous minimization of individual instability and inter-agent inconsistency.

D. Theorem 3 (Abstract Shared-TCZ Convergence: Rigorous Form)

Setup. Let the joint state space be $\mathbf{X} = \prod_i X_i$. The subsumption lattice $(L, \leq)$ over cognitive worlds is equipped with a top element $\top = \text{LUB}(W_1,\ldots,W_N)$. Define the abstraction map $\varphi : \mathbf{X} \to L$ assigning each joint state to its position in the lattice. The abstraction potential $A : \mathbf{X} \to \mathbb{R}_{\geq 0}$ satisfies:

- (A1) $A(\mathbf{x}) = 0$ if and only if $\varphi(\mathbf{x}) = \top$ (i.e., $\mathbf{x}$ realizes LUB)
- (A2) $A(\mathbf{x}) > 0$ for all $\mathbf{x}$ with $\varphi(\mathbf{x}) \neq \top$
- (A3) A is continuously differentiable and decreases strictly along upward lattice movement: if $\varphi(\mathbf{x}') > \varphi(\mathbf{x})$ in L, then $A(\mathbf{x}') < A(\mathbf{x})$

Statement. Let $\Phi(\mathbf{x}) = \mathscr{L}_A(\mathbf{x}) = \Sigma_i V_i(x_i) + \lambda S(\mathbf{x}) + \eta A(\mathbf{x})$, with $\lambda, \eta > 0$. Suppose $\nabla\mathscr{L}_A(\mathbf{x}) \cdot F(\mathbf{x}) \leq -\alpha(\mathscr{L}_A(\mathbf{x}) - \theta_A)$ for all $\mathbf{x} \notin \Omega_A$, where $\Omega_A := \{\mathbf{x} \mid \mathscr{L}_A(\mathbf{x}) \leq \theta_A\}$. Then $\mathbf{x}^*(t) \to \text{LUB}(W_1,\ldots,W_N)$.

Proof.

Step 1 (Lyapunov descent). Define $y_A(t) = \mathscr{L}_A(\mathbf{x}(t)) - \theta_A$. Then $\dot{y}_A \leq -\alpha y_A$. By the comparison principle:

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

Hence $\text{dist}(\mathbf{x}(t), \Omega_A) \to 0$ as $t \to \infty$.

Step 2 (Invariance of Ω_A). On $\partial\Omega_A$, $\nabla\mathscr{L}_A \cdot F \leq 0$ by assumption. Hence Ω_A is forward invariant, and $\mathbf{x}(t) \in \Omega_A$ for all $t \geq t_0$.

Step 3 (LUB identification via A). Within Ω_A, all three terms of $\mathscr{L}_A$ are bounded. By LaSalle's invariance principle, the trajectory converges to the largest invariant set $\mathscr{M} \subset \Omega_A$ where $d\mathscr{L}_A/dt = 0$. On $\mathscr{M}$, the minimization of $\mathscr{L}_A$ implies: (i) each $V_i(x_i) \leq \theta$ (individual stability); (ii) $S(\mathbf{x}) = 0$ (shared consistency); (iii) $A(\mathbf{x}) = 0$ (abstraction potential vanishes).

Step 4 (*Conclusion*). By condition (A1), $A(\mathbf{x}) = 0$ if and only if $\varphi(\mathbf{x}) = \top = \mathrm{LUB}(W_1,\ldots,W_N)$. Therefore $\mathbf{x}^*(t) \to \mathrm{LUB}(W_1,\ldots,W_N)$. ■

Remark. The key additional structure over Theorems 1–2 is condition (A1): the abstraction potential A must be a faithful indicator of distance from LUB in the lattice. This is a design condition on the control policy, not an automatic consequence of the Lyapunov structure. In practice, A can be constructed as a lattice-distance function, e.g., $A(\mathbf{x}) = d_L(\varphi(\mathbf{x}), \top)$, where d_L denotes a suitable metric on the lattice L.

Interpretation. Convergence is lifted from intersection-based stability (Theorem 2) to abstraction-based integration at the LUB, provided the abstraction potential faithfully encodes the lattice structure.

E. Unified Interpretation

The three theorems form a hierarchy (individual stability → shared stability → abstract stability). All are instances of a single structure:

System dynamics are governed by decreasing instability and increasing abstraction.

F. LaSalle-Type Refinement

Define $\mathscr{M}_\Phi = \{\, z \mid \nabla\Phi(z) \cdot F(z) = 0 \,\}$. By LaSalle's invariance principle: $z(t) \to \mathscr{M}_\Phi \cap \Omega_\theta$. Thus convergence occurs to invariant structures within the stability region.

G. Cross-Domain Interpretation

This structure is universal:

- Physics: energy decreases
- Cognition: evaluative function decreases
- FEP: free energy decreases
- Predictive processing: prediction error decreases

All share the same convergence principle. The present framework extends this by showing that convergence is accompanied by abstraction, yielding a unified Lyapunov structure across physiical, cognitive, and social systems.

Stability arises through descent in potential; generality and inclusion arise through ascent in abstraction.

Theorem 1 (Individual Level): By modifying $V(x,t)$—increasing perceived threat, instability, or dissonance associated with certain states—the shape of the potential landscape is altered. As a result, the optimal trajectory

$x^*(t)$ converges toward a different TCZ.

Theorem 2 (**Social Level**)**:** Constraining shared stability structure narrows the system toward low-level Shared-TCZ, amplifying fragmentation and instability.

Theorem 3 (**Abstract Level**)**:** Increasing abstraction enables convergence toward high-level Shared-TCZ, producing integration, coherence, and stability.

Key Insight: Cognitive warfare operates not at the level of actions, but at the level of structures that generate actions.

This expresses boundary-level control with minimal intervention. The cost term $|V-\theta|^2$ is symmetric around the boundary: it attracts trajectories toward $V = \theta$ from both inside and outside. Cognitive warfare is most efficient when it operates where the system is most sensitive.

Accordingly, effective generation should construct a set of messages $M_t = \{ m_1, \ldots, m_K \}$ that are individually acceptable yet collectively induce movement toward a higher-level shared structure:

Theorems 1–3 are not independent results, but successive extensions of a single Lyapunov-based stability structure.

Consider a closed-loop system $\dot{z} = F(z)$. Let $\Phi(z)$ be continuously differentiable and define $\Omega_\theta := \{ z \mid \Phi(z) \leq \theta \}$. Assume $\nabla\Phi(z) \cdot F(z) \leq -\alpha(\Phi(z) - \theta)$, $\alpha > 0$, for all $z \notin \Omega_\theta$. Then $\mathrm{dist}(z(t), \Omega_\theta) \to 0$.

Proof. Define $y(t) = V(x(t)) - \theta$. Then $\underline{l}(t) \leq -\alpha y(t)$. By comparison principle:

Thus $x(t)$ approaches TCZ_θ. ■

Define composite Lyapunov function $\mathscr{L}(\mathbf{x}) = \Sigma_i V_i(x_i) + \lambda S(\mathbf{x})$. By LaSalle's principle, trajectories converge to the invariant set where $\mathscr{L} = \theta$, which implies $S(\mathbf{x}) = 0$, i.e., $\mathbf{x} \in \mathrm{TCZ}_{\mathrm{shared}}$. ■

Define $\mathscr{L}_A(\mathbf{x}) = \Sigma_i V_i(x_i) + \lambda S(\mathbf{x}) + \eta A(\mathbf{x})$. By LaSalle's invariance