Active Finite Distinction Systems

A Formal Core for Boundary Maintenance under Finite Capacity

Capacity Deficit, Dissipation, Externalization, Collapse, and Invariant Persistence

Formal Core Edition

v1.0 | May 2026

Yining Wu
Independent Researcher
yining.wu@alumni.upenn.edu

A formal core for what active finite systems must pay in order to maintain a boundary under limited capacity.

Companion formal core to *Distinction Theory: A General Theory of Finite Systems*, Zenodo, doi:10.5281/zenodo.20130174.

EPISTEMIC NOTICE

This manuscript does not present Distinction Theory as an undifferentiated theory of everything. It presents a restricted mathematical theory of *active finite distinction systems*: systems that maintain a boundary through state-dependent updates under finite representational capacity and finite resource budgets.

The core chain is not asserted as rhetoric. It is decomposed into formal definitions, conditional theorems, physical bridge assumptions, normal-form dynamics, and quarantined applications:

$$\begin{aligned}&\text{distinction} \to \text{boundary} \to \text{finite capacity} \to \text{capacity deficit}\\&\to \text{approximation} \to \text{irreversible update} \to \text{dissipation}\\&\to \text{prune/externalize/collapse} \to \text{invariant persistence.}\end{aligned}$$

The mathematical core can fail only by mathematical counterexample or by failure of stated hypotheses. The physical bridge can fail by violation of its thermodynamic assumptions. Domain applications can fail without propagating failure back to the core.

A failed claim is not reinterpreted into survival. Each claim carries an explicit proof type, dependency class, and failure consequence.

Abstract

This document defines active finite distinction systems as finite-capacity systems that maintain boundaries through state-dependent updates under resource constraints. It develops the formal core of Distinction Theory, including representational capacity, rate-distortion capacity deficit, conditional approximation proliferation, Landauer dissipation floor under physical bridge assumptions, prune–externalize–collapse trichotomy, and invariant-supported persistence. Domain applications are quarantined behind explicit bridge assumptions and operational tests.

How to cite this document:

Wu, Y. *Active Finite Distinction Systems: A Formal Core for Boundary Maintenance under Finite Capacity.* Zenodo, 2026.

Version relationship:

This document is the formal FDS Core companion to the full claim-space edition: *Distinction Theory: A General Theory of Finite Systems*, Zenodo, doi:10.5281/zenodo.20130174.

Reader Orientation

This document extracts the formal FDS core from the broader Distinction Theory claim-space archive. It is organized as a theory with three layers.

Layer 1: Formal FDS Core.
Definitions and conditional mathematical theorems about finite distinction systems. This layer uses partitions, information capacity, rate-distortion demand, conditional entropy, update maps, perturbation families, and invariant quotients.

Layer 2: Physical Bridge.
Conditions under which formal updates are realized by physical substrates coupled to thermal environments. This layer introduces Landauer costs, physical dissipation, free-energy budgets, and maintenance failure.

Layer 3: Quarantined Applications.
Biology, cognition, artificial intelligence, civilization, geometry, black holes, and other domains are treated only through explicit domain bridge assumptions and operational tests.

The purpose of this reconstruction is not to make the theory immune to criticism. It is to make every criticism land in the correct place.

Contents

Part I.

Formal Kernel: Active Finite Distinction Systems

1. The Primitive and the Scope of the Theory

1.1. The Distinction Primitive

Definition 1.1 (Distinction)**.** A *distinction* is an operation or relation that separates at least two alternatives within a possibility space. Formally, at the most elementary level, a distinction may be represented by a nontrivial partition

$$\mathcal{D} = \{A_i\}_{i=1}^{n}, \qquad A_i \cap A_j = \varnothing \ (i \neq j), \qquad \bigcup_i A_i = X, \qquad n \geq 2.$$

The primitive is not the claim that every distinction is physical, costly, or thermodynamic. A purely formal distinction is a logical separation. A realized distinction in an active finite system is an operational state maintained under constraints.

Proposition 1.2 (Denial performs distinction)**.** *Any coherent denial of distinction distinguishes the denial from the proposition denied. Thus the act of denial presupposes the operation it rejects.*

Proof. Let P be the proposition "distinction is required for coherent identity". A denial asserts $\neg P$ rather than P. The assertion is meaningful only if $\neg P$ is distinguished from P. Hence the denial performs a distinction. This is a logical observation, not a physical theorem. □

1.2. What the Primitive Does Not Prove

Not claimed. The distinction primitive alone does not derive thermodynamics, biology, cognition, artificial agency, black holes, spacetime geometry, or civilization. Those claims require finite-system restrictions, active-boundary hypotheses, physical bridge assumptions, and domain-specific mappings.

1.3. Finite Systems as the Domain of Application

The formal theory begins only when the primitive is restricted to realized bounded systems with finite representational and energetic resources. The central object is not an abstract distinction, but an active finite distinction system.

Definition 1.3 (Proof status labels)**.** Each claim in this document carries one of four proof statuses.

Formal. A mathematical statement derivable from definitions and stated hypotheses.

Conditional. A mathematical or physical statement derivable only under explicitly stated bridge assumptions.

Normal-form. A dynamical model describing a generic regime, not a universal law.

Domain-bridge. An applied claim requiring a mapping from the FDS formalism to a specific empirical domain.

1.3.1. From the Distinction Primitive to the FDS Object

The distinction primitive by itself is not a complete physical or computational theory. A bare distinction only asserts a separation: some difference is drawn between what is counted as one side and what is counted as another. To obtain a theory of active finite systems, the primitive must be realized under additional operational constraints: the distinction must be finite, maintained, updated, exposed to perturbation, and evaluated over time.

The twelve-component FDS object is therefore not introduced as an arbitrary list of variables. It is an explicit operational decomposition of what is otherwise often smuggled into informal uses of the word "distinction". Each component records one functional role required for a finite distinction to be actively maintained rather than merely declared.

First, a maintained distinction requires a partitioned domain: an internal state space X, an environment E, and a boundary B across which the self–environment distinction is operationally defined. Without this partition, there is no object whose boundary can be maintained.

Second, a finite system cannot preserve all distinctions available in its environment. It therefore requires an internal memory or model space M, a coarse-graining or projection map π, and a finite resource budget Φ. These terms make explicit the finiteness that is otherwise hidden in informal talk of an observer, agent, organism, or record-bearing system.

Third, an active distinction is not static. The system must receive observations through a channel Y, update its internal state through an update rule U, and possess an action or intervention space A through which its updates can matter for future boundary-maintenance loss. Without update and intervention, the system is at most a passive record or static partition.

Fourth, maintenance requires vulnerability. There must be a loss function ℓ specifying what counts as boundary-maintenance failure, a perturbation or pruning family $\mathcal{P}$ specifying the admissible transformations under which persistence is assessed, and an update timescale τ converting static information quantities into rates, costs, and dynamical constraints.

Thus the FDS tuple

$$S = (X, E, B, M, Y, A, U, \pi, \ell, \Phi, \mathcal{P}, \tau)$$

should be read as a minimal operational anatomy of an active finite distinction. Individual components may be compressed or represented differently in particular formalisms, but the functional roles they mark cannot be removed without either changing the object of study or reintroducing hidden assumptions.

2. The Finite Distinction System

2.1. The Twelve-Component Object

Definition 2.1 (Finite Distinction System)**.** A *finite distinction system* is a tuple

$$\mathcal{S} = (X, E, B, M, Y, A, U, \pi, \ell, \Phi, \mathcal{P}, \tau),$$

where:

X	is the internal state space of the system.
E	is the environmental state space.
B	is the boundary or interface variable separating system from environment.
M	is the internal memory, model, or representational state space.
Y	is the observation channel available at the boundary.
A	is the action or boundary-update space.
U	is the internal update map, typically $U : M \times Y \to M$ or a stochastic kernel $\mathbf{P}(M_{t+1} \mid M_t, Y_t)$.
π	is the finite distinction projection, coarse-graining, or partition map.
ℓ	is the boundary-maintenance loss function.
Φ	is the finite resource or free-energy budget functional.
$\mathcal{P}$	is the admissible family of pruning, coarse-graining, perturbation, or degradation operators.
τ	is the update timescale converting per-step information loss into rates or powers.

Remark 2.2. The twelfth component τ is not cosmetic. Without a timescale, information-theoretic quantities remain dimensionless counts. With τ, they become rates and can be connected to thermodynamic power bounds.

2.2. Active and Passive Boundaries

Definition 2.3 (Active boundary)**.** An FDS has an *active boundary* if its update rule is nontrivial and participates in future boundary maintenance:

$$\mathbb{P}\big(U(M_t, Y_t) \neq M_t\big) > 0$$

and there exists $k > 0$ such that

$$I(M_{t+1}; \ell_{t+k}) > 0.$$

The second condition excludes irrelevant state noise. A state change counts as active-boundary-relevant only if it is statistically related to future boundary-maintenance loss; empirical applications require the stronger causal qualification in Remark 2.8.

Remark 2.4 (Intervention criterion — cross-reference)**.** The mutual-information condition $I(M_{t+1}; \ell_{t+k}) > 0$ is a relevance screen; for the empirical causal strengthening, see Remark 2.8.

Definition 2.5 (Passive boundary)**.** A system has a *passive boundary* when its persistence is supported by static physical constraints or stable configurations without task-directed internal updating. Passive persistence may be real and stable, but it does not enter the deficit-dissipation cascade unless an active update channel is introduced.

Domain Restriction 0

The deficit-dissipation-pruning cascade applies only to active-boundary finite distinction systems. Passive crystals, inert rocks, stable nuclei, and purely formal mathematical objects are not governed by the FDS cascade merely because they possess boundaries.

Criterion 2.6 (Active Boundary Qualification)**.** An FDS enters the core Distinction Theory cascade only if its boundary-maintenance loss depends on nontrivial internal updates. Formally, the necessary qualification is

$$\mathbb{P}(U(M_t, Y_t) \neq M_t) > 0 \quad \text{and} \quad I(M_{t+1}; \ell_{t+k}) > 0$$

for some $k > 0$.

Remark 2.7 (Justification)**.** If $U(M_t, Y_t) = M_t$ almost surely, no internal update is performed and no internal correction channel exists. If U changes M_t but $I(M_{t+1}; \ell_{t+k}) = 0$ for all $k > 0$, the update is not statistically relevant to boundary maintenance under the chosen variables. In either case the system lacks active-boundary dynamics in the sense required by the cascade. Conversely, if both conditions hold, the system possesses a nontrivial internal update relevant to future boundary-maintenance loss. This is precisely the domain condition.

Remark 2.8 (Causal strengthening for empirical systems)**.** The mutual-information condition $I(M_{t+1}; \ell_{t+k}) > 0$ is a minimal formal relevance condition, not by itself a causal intervention test. In empirical applications, active boundary maintenance should be operationalized by an intervention or ablation criterion: there must exist an admissible null update channel $U_\emptyset$ (such as freezing, randomizing, or identity-updating the internal update) and some $k > 0$ such that

$$\mathbb{E}[\ell_{t+k} \mid do(U)] \neq \mathbb{E}[\ell_{t+k} \mid do(U_\emptyset)].$$

This excludes passive observers whose internal states merely correlate with future boundary loss through a common environmental cause.

2.3. Non-Domain Cases

The following are not core FDS cascade cases unless extra active-boundary structure is specified:

- inert passive solids;
- equilibrium crystals without active repair;
- closed unitary quantum evolution without coarse-grained irreversible update;
- static mathematical structures;
- read-only storage without overwrite, erasure, or maintenance update;
- systems whose internal changes do not affect boundary-maintenance loss.

Part II.

Information Theorems: Capacity, Approximation, and Complexity

3. Finite Distinction Representation

3.1. Representational Capacity

Definition 3.1 (Internal representational capacity)**.** For a finite memory/model space M, define the internal representational capacity

$$C_S = \log_2 |M|.$$

When M is continuous or effectively continuous, C_S is replaced by the appropriate operational capacity under finite resolution, noise, or coding constraints.

Definition 3.2 (Task-relevant environmental variable)**.** Let

$$Z_t = \psi(E_t, B_t)$$

be a task-relevant environmental statistic relevant to maintaining the boundary up to a specified distortion or loss tolerance. The theory does not require full representation of E_t; it requires only the representation needed for the boundary-maintenance task.

Theorem 3.3 (Finite Distinction Representation)**.** *For any FDS with finite internal model capacity C_S, the mutual information between the environment and the internal model satisfies*

$$I(E_t; M_t) \leq C_S.$$

Proof. By the data-processing inequality and the entropy bound on a finite memory state,

$$I(E_t; M_t) \leq H(M_t) \leq \log_2 |M| = C_S.$$

□

3.2. Operational Meaning

This theorem does not say that the environment is finite. It says that the internally maintained distinction structure is finite. The system may face an environment whose task-relevant complexity exceeds what the system can encode internally.

4. Capacity Deficit

4.1. Rate-Distortion Demand

Definition 4.1 (Admissible task statistics)**.** Let Ψ be a family of admissible statistics $\psi : (E, B) \to Z$ that are causally accessible through the observation channel Y. A statistic $Z_t = \psi(E_t, B_t)$ is task-relevant if it captures information needed to maintain the boundary within tolerated distortion. Admissible statistics must preserve sufficient information for the stated boundary-maintenance criterion; trivial constant maps are excluded unless the task itself is trivial. The family Ψ is fixed by the task specification and sensor–action architecture, not chosen post hoc after observing failure.

Definition 4.2 (Pre-registered admissible task families)**.** To prevent post-hoc selection of task statistics, the admissible family Ψ must be fixed before evaluating capacity deficit. Typical admissible families include:

(i) **Control or homeostasis tasks.** Ψ_{ctrl} contains statistics of (E_t, B_t) sufficient to select actions that keep the boundary variable B_t inside a specified viable set $\mathcal{V}$ up to tolerance ε.

(ii) **Predictive tracking tasks.** Ψ_{pred} contains statistics sufficient to predict a pre-specified future observable Y_{t+k} or target variable Z_{t+k} under a fixed loss function.

(iii) **Information-bottleneck tasks.** Ψ_{IB} contains compressed statistics that preserve task-relevant information about a pre-specified target Z under a stated rate-distortion or information-bottleneck objective.

The class Ψ is part of the task specification, not a variable chosen after observing system failure.

Definition 4.3 (Rate-distortion demand)**.** For an admissible statistic ψ, update window τ, and tolerated distortion ε, let

$$R^{(\tau)}_{\psi(E,B)}(\varepsilon)$$

be the minimum number of bits per update window τ required to encode $\psi(E, B)$ so that the expected boundary-maintenance distortion remains below ε. The minimal rate-distortion demand over the admissible class is

$$R^{(\tau)}_{\min}(\varepsilon) = \inf_{\psi \in \Psi} R^{(\tau)}_{\psi(E,B)}(\varepsilon).$$

Definition 4.4 (Capacity deficit)**.** The *capacity deficit* of an FDS at tolerance ε

over window τ is

$$\Delta_\varepsilon(\tau) = R^{(\tau)}_{\min}(\varepsilon) - C_S.$$

If $\Delta_\varepsilon(\tau) > 0$, the system is in predictive capacity deficit at tolerance ε.

Theorem 4.5 (Capacity Deficit Theorem)**.** *If*

$$R^{(\tau)}_{\min}(\varepsilon) > C_S,$$

then no purely internal model M_t of capacity C_S can encode even the least demanding admissible task statistic to distortion ε over window τ. The system must approximate, externalize, relax the task, or suffer boundary-maintenance failure.

Proof. For any admissible ψ, any representation of $\psi(E, B)$ achieving distortion at most ε over window τ requires at least $R^{(\tau)}_{\psi(E,B)}(\varepsilon)$ bits by the rate-distortion theorem. Taking the infimum, $R^{(\tau)}_{\min}(\varepsilon)$ is the minimal encoding demand across the admissible class. By Theorem 3.3, the system can internally maintain at most C_S bits. If $R^{(\tau)}_{\min}(\varepsilon) > C_S$, no admissible statistic can be encoded internally to the required distortion tolerance ε. Therefore at least one of the following must occur: representation is lossy beyond ε, representation is externalized outside M, the task tolerance is relaxed, or boundary-maintenance performance fails. □

Corollary 4.6 (Approximation Necessity)**.** *For an active-boundary FDS with $\Delta_\varepsilon(\tau) > 0$, approximation is not optional. It is forced by the information-theoretic mismatch between admissible task demand and internal capacity.*

Remark 4.7. The capacity deficit is not a vague cognitive gap. It is a computable gap between a task-relevant rate-distortion demand and an internal representational capacity.

5. Approximation Proliferation (Conditional)

5.1. Minimal Sufficient Complexity

Definition 5.1 (Minimal sufficient model complexity)**.** Let $\mathfrak{C}$ be a fixed coding family (representation language, model architecture, or hypothesis class) that does not change during the considered interval. Let $C(M)$ be a model-complexity functional, such as description length, state count, algorithmic cost, parameter count under a scheme in $\mathfrak{C}$, or thermodynamic maintenance load. Define

$$C_t^* = \min\{C(M) : \mathbb{E}[\ell_t(M)] \leq \ell_c\},$$

where ℓ_c is an admissible boundary-maintenance threshold, and the minimum is taken only over models inside $\mathfrak{C}$, not over future inventions of new representation languages.

Definition 5.2 (Persistent task novelty)**.** Fix a coding family $\mathfrak{C}$, a loss threshold ℓ_c, and the filtration $\mathcal{F}_t$ generated by the system's observations and internal states up to time t. The task process has persistent novelty at rate $\eta > 0$ if, for infinitely many update windows,

$$\mathbb{E}\left[\inf_{M\in\mathfrak{C}:\, C(M)\leq C_t^*} \mathbb{E}[\ell_{t+1}(M) \mid \mathcal{F}_t] - \ell_c \;\middle|\; \mathcal{F}_t\right]_+ \geq \eta.$$

Equivalently, the current minimum sufficient model fails to absorb future task-relevant variation without refinement, pruning, externalization, task relaxation, or compression improvement. The novelty rate η is relative to the fixed coding family $\mathfrak{C}$: even if the environment has positive entropy rate, a sufficiently expressive $\mathfrak{C}$ may absorb the variation without complexity growth.

Remark 5.3 (Entropy-rate sufficient condition)**.** A positive entropy rate $h(Z) > 0$ of the task-relevant process Z_t does not by itself imply model-complexity proliferation. It does so only when the induced innovations are not compressible within the fixed coding family $\mathfrak{C}$ at distortion ε, i.e. when the conditional rate-distortion residual remains bounded away from zero:

$$R^{(\tau)}_{Z_{t+1}\mid\mathcal{F}_t,\mathfrak{C}}(\varepsilon) \geq \eta > 0$$

over a positive-density set of update windows.

Theorem 5.4 (Conditional Approximation Proliferation under Fixed Coding Family). *Assume* $\Delta_\varepsilon(\tau) > 0$ *over a relevant window, persistent task novelty, no active pruning, no externalization, no task relaxation, and no improvement in compression scheme. Then the minimal sufficient model complexity is a submartingale:*

$$\mathbb{E}[C^*_{t+1} \mid \mathcal{F}_t] \geq C^*_t.$$

Remark 5.5 (Proof sketch)**.** Under capacity deficit, the system cannot internally encode all task-relevant distinctions to tolerance ε over the update window. Persistent novelty generates residual cases not captured by the current sufficient model. If the system must keep expected loss below ℓ_c and cannot prune, externalize, relax the task, or improve compression, the only remaining operation is to add or refine internal distinctions. Thus the minimum sufficient complexity cannot decrease in conditional expectation, giving the submartingale inequality. A fully formal proof would require a specified filtration $\mathcal{F}_t$, a measurable novelty process with a positive-rate definition, and a submartingale convergence argument. The present form is a conditional formal theorem.

Remark 5.6. This theorem does not say that observed representation length always increases. Better compression, abstraction, external memory, task relaxation, or pruning may reduce observed complexity. The theorem concerns minimum sufficient complexity under explicitly restricted conditions.

Remark 5.7. This result does not exclude discontinuous decreases in maintained complexity when a superior abstraction, compression code, pruning operation, or externalization channel becomes available. Historical examples include the transition from geocentric epicycles to Keplerian orbits, or the replacement of hand-crafted features by learned representations in machine learning. The theorem states only that in the absence of such improvements, persistent novelty forces the minimum sufficient complexity to be a submartingale.

Part III.

Physical Bridge: Irreversibility, Dissipation, and Free-Energy Budgets

6. Logical Irreversibility and Landauer Cost

6.1. Preimage Conditional Entropy

Definition 6.1 (Logical erasure per update)**.** For an update $M_{t+1} = U(M_t, Y_t)$, define the logical erasure in bits by

$$b_t = H(M_t \mid M_{t+1}, Y_t).$$

This quantity measures how much uncertainty remains about the prior memory state after the new memory state and current input are known. It is the preimage information lost by the update.

Remark 6.2. A forward divergence such as $D_{\mathrm{KL}}(\mathbb{P}(M_{t+1}) \,\|\, \mathbb{P}(U(M_t, Y_t)))$ does not measure erasure when M_{t+1} is defined by $U(M_t, Y_t)$; in that case the two distributions coincide. Logical erasure is a property of non-invertible preimages, not of forward equality.

Assumption 6.3 (Physical Landauer bridge)**.** A logically irreversible update implemented by a physical substrate coupled to a thermal environment at temperature $T > 0$ dissipates at least $k_{\mathrm{B}} T \ln 2$ per erased bit, under the standard physical conditions of Landauer's principle.

Theorem 6.4 (Landauer Dissipation Floor)**.** *For an active-boundary FDS physically implementing irreversible memory updates at timescale τ, the informational heat dissipation rate obeys*

$$\dot{Q}_{\mathrm{info}} \geq \frac{k_{\mathrm{B}} T \ln 2}{\tau} H(M_t \mid M_{t+1}, Y_t).$$

The total maintenance cost decomposes as

$$\dot{Q}_{\mathrm{total}} = \dot{Q}_{\mathrm{phys}} + \dot{Q}_{\mathrm{info}},$$

where $\dot{Q}_{\mathrm{phys}}$ includes non-informational physical losses such as control, coupling, leakage, friction, error correction, clocking, isolation, and transport.

Proof. By definition, $H(M_t \mid M_{t+1}, Y_t)$ is the number of erased bits per update. Under the physical Landauer bridge, each erased bit costs at least $k_{\mathrm{B}} T \ln 2$ of dissipated heat. Dividing by update timescale τ gives the rate bound. Physical losses independent of logical erasure add to the informational term. □

Remark 6.5. If $H(M_t \mid M_{t+1}, Y_t) = 0$, the update is logically reversible relative to Y_t and no Landauer erasure cost follows from this term. It does not follow that the system is passive or cost-free. Reversible active computation may still require nonzero $\dot{Q}_{\mathrm{phys}}$.

6.2. Failure Condition for the Physical Bridge

Failure Condition 6.6 (Landauer bridge failure)**.** If a physical system satisfying the stated conditions of the Landauer bridge performs logically irreversible many-to-one overwrite or erasure at a reliable average cost below $k_{\mathrm{B}}T \ln 2$ per erased bit, the physical bridge fails. Formal FDS theorems remain intact; thermodynamic conclusions depending on the bridge are demoted or abandoned.

7. Free-Energy Budgets and Maintenance Failure

Definition 7.1 (Available free-energy input)**.** Let $\dot{F}_{\text{in}}(t)$ denote the free-energy input rate available to support boundary maintenance, internal updating, pruning, externalization, repair, and control.

Definition 7.2 (Maintenance failure region)**.** An FDS enters the maintenance failure region when, over a relevant interval $[t_0, t_1]$,

$$\int_{t_0}^{t_1} \dot{Q}_{\text{total}}(t)\, dt > \int_{t_0}^{t_1} \dot{F}_{\text{in}}(t)\, dt$$

without compensating resource reserves, external subsidies, task relaxation, pruning, or externalization.

Assumption 7.3 (Loss–resource coupling)**.** In a physically realized FDS, boundary-maintenance loss may feed back into the available resource budget through a nonnegative damage functional $D_\ell(\ell_t) \geq 0$. Over one update window τ,

$$\Phi_{t+1} \leq \Phi_t + \int_t^{t+\tau} \dot{F}_{\text{in}}(s)\, ds - \int_t^{t+\tau} \dot{Q}_{\text{total}}(s)\, ds - D_\ell(\ell_t).$$

The damage functional may take various forms: linear ($\kappa\, \ell_t$), threshold-based (0 for $\ell \leq \ell_c$, positive above), or catastrophic (∞ for $\ell > \ell_{\text{death}}$). The formal FDS core does not require this coupling; it becomes active only in physical or domain-bridge models where boundary failure consumes repair capacity, destroys structure, or reduces future free-energy access.

Remark 7.4 (Model families for loss–resource coupling)**.** The damage functional D_ℓ is domain-specific. Common choices include linear damage $D_\ell(\ell) = \kappa\ell$, threshold damage $D_\ell(\ell) = \kappa(\ell - \ell_c)_+$, accelerated damage $D_\ell(\ell) = \kappa(e^{\alpha(\ell-\ell_c)} - 1)_+$, and catastrophic damage $D_\ell(\ell) = \infty$ for $\ell > \ell_{\text{death}}$. The FDS core does not select among these forms; empirical domain bridges must specify which coupling applies.

Theorem 7.5 (Prune–Externalize–Collapse Trichotomy)**.** *Let an active-boundary FDS satisfy $\Delta_\varepsilon(\tau) > 0$ over a relevant window and enter the maintenance failure region, with bounded $\dot{F}_{\text{in}}$ and no compensating external subsidy. Then the system cannot maintain the same active-boundary regime indefinitely; its trajectory must enter at least one of three outcome classes:*

(i) pruning*: reducing internal complexity or maintenance load;*

(ii) externalization*: shifting representational, energetic, or control load outside the internal memory budget M;*

(iii) collapse*: loss of boundary-maintenance capacity, task failure, or transition to a lower-complexity attractor.*

Proof. Under $\Delta_\varepsilon(\tau) > 0$ over the relevant window, the system faces a representational shortfall within its update budget. If continued internal updating and maintenance exceed the available free-energy input, the system cannot indefinitely maintain the same internal complexity and task burden. The admissible exits are: reduce the maintained internal load, move part of the load outside the internal budget, or fail to maintain the boundary/task. These are pruning, externalization, and collapse respectively. □

Remark 7.6. The trichotomy replaces the older binary phrase "prune or collapse". Many systems persist by externalizing distinctions into bodies, tools, texts, institutions, niches, databases, or environments. Externalization is not a loophole; it is a relocation of maintenance cost and boundary responsibility.

Part IV.

Persistence: Quotients, Perturbations, and Invariant-Supported Identity

8. Invariant-Supported Persistence

8.1. Identity Predicates and Quotient Maps

Definition 8.1 (Structural identity predicate)**.** Let $R_A : X \to \{0, 1\}$ be an identity predicate for a structure A in state space X. The statement $R_A(x) = 1$ means that state x realizes structure A under the chosen resolution.

Definition 8.2 (Invariant quotient)**.** Let $q : X \to T$ be a quotient or feature map into a reduced space T. The quotient is invariant under perturbation family $\mathcal{P}$ if

$$q \circ P_i = q$$

for all $P_i \in \mathcal{P}$.

Definition 8.3 (Invariant-supported identity)**.** A structural identity predicate R_A is invariant-supported under $\mathcal{P}$ if there exist $q : X \to T$ and $\bar{R}_A : T \to \{0, 1\}$ such that

$$R_A = \bar{R}_A \circ q$$

and

$$q \circ P_i = q \qquad \forall P_i \in \mathcal{P}.$$

Theorem 8.4 (Invariant Persistence)**.** *If R_A is invariant-supported under $\mathcal{P}$, then identity is preserved under every admissible perturbation:*

$$R_A \circ P_i = R_A \qquad \forall P_i \in \mathcal{P}.$$

Proof. Since $R_A = \bar{R}_A \circ q$ and $q \circ P_i = q$,

$$R_A \circ P_i = (\bar{R}_A \circ q) \circ P_i = \bar{R}_A \circ (q \circ P_i) = \bar{R}_A \circ q = R_A.$$

Thus the identity predicate is unchanged by all admissible perturbations. □

Corollary 8.5 (Unprotected identity)**.** *If no such invariant quotient exists, then the identity predicate is not algebraically protected by the perturbation family. The corollary does not imply that non-invariant identities must fail: active maintenance, repair, isolation, stochastic stability, or energy-consuming feedback may preserve them. It states only that algebraic protection through an invariant quotient is absent, so survival depends on ongoing compensatory mechanisms rather than structural persistence.*

Remark 8.6. The theory does not claim that every non-invariant identity immediately dies. It claims that non-invariant identities lack algebraic protection against the allowed perturbation family and therefore require active maintenance, isolation, repair, or external support.

8.2. Illustrative Examples

Example 8.7 (Error-correcting codes)**.** Let X be the space of physical code states and admissible corrupted states. Let P_i be an admissible error from a correctable error set. Let $q : X \to T$ map a physical state to its recovered logical equivalence class, i.e. the logical state obtained after syndrome extraction and recovery. Then for all correctable errors P_i,

$$q \circ P_i = q.$$

The physical microstate and syndrome may change, but the recovered logical identity is invariant under the admissible error family.

Example 8.8 (Software API contracts)**.** Let X be the full internal memory and implementation state of a service. Let P_i include garbage collection, memory reallocation, thread scheduling, or implementation refactoring. For a fixed admissible request class, let q map internal implementation states to contract-level input-output behavior. If the same valid request receives the same contract-level response before and after P_i, then the service identity is supported by the invariant quotient rather than by its microstate.

Example 8.9 (Topological winding numbers)**.** For a gapped system with a topological winding number, let q map a microscopic configuration to its winding class. Perturbations that preserve the gap and remain within the admissible locality and symmetry class satisfy $q \circ P_i = q$. The protected identity is therefore invariant under that restricted perturbation family, not under arbitrary destructive perturbations.

9. Coarse-Grained Anti-Recurrence

Definition 9.1 (Effective truncation map)**.** An effective truncation map is a non-injective coarse-graining or deletion operator

$$T : X \to X'$$

that identifies at least two distinct states: $x \neq x'$ but $T(x) = T(x')$.

Lemma 9.2 (Coarse-Grained Anti-Recurrence)**.** *In an effective representation containing a non-injective truncation map, recurrence to an erased micro-distinction is impossible within that representation without external records or additional inverse information.*

Proof. If $T(x) = T(x')$ for $x \neq x'$, then the image state does not contain enough information to determine whether the preimage was x or x'. Any inverse selection requires information not present in the truncated representation. Therefore recurrence to the specific erased micro-distinction is impossible inside the effective state description alone. □

Remark 9.3. This lemma is not a denial of microscopic reversibility, Hamiltonian recurrence, or unitary quantum evolution under their own hypotheses. It applies to coarse-grained effective representations containing non-injective truncation.

Proposition 9.4 (Coarse-graining and lumpability)**.** *Let (X_t) be a Markov process on a microscopic state space X, and let $T : X \to X'$ be a non-injective coarse-graining map. The projected process $X'_t = T(X_t)$ is Markovian only under the usual lumpability condition [14]: transition probabilities out of any two microstates in the same fiber $T^{-1}(x')$ agree after projection. If lumpability fails, the projected dynamics carries hidden-state memory; recurrence to a specific erased micro-distinction requires information not contained in X'_t alone.*

Remark 9.5. Thus Lemma 9.2 is not a claim about microscopic irreversibility. It is a claim about loss of inverse information in the effective representation. Lumpability identifies the exceptional cases in which coarse-grained dynamics remains Markovian; non-lumpable projections generically introduce memory, apparent irreversibility, or entropy production at the effective level.

Part V.

Normal-Form Dynamics: Phase A, Phase B, and Failure Modes

10. The Minimal Dynamical Schema

10.1. Regime Variables

Let C_t denote maintained internal complexity, Φ_t available free-energy reserve, S_t pruning or simplification effort, and E_t^{ext} externalization load. A generic continuous-time normal form may be written as

$$\frac{dC}{dt} = G(\Delta_\varepsilon(\tau), C) - \text{Prune}(S, C) - X(E^{\text{ext}}, C),$$

$$\frac{d\Phi}{dt} = \dot{F}_{\text{in}} - \dot{Q}_{\text{phys}}(C, B) - \dot{Q}_{\text{info}}(U, M, Y) - \Gamma(S) - \Xi(E^{\text{ext}}),$$

where G is deficit-driven complexity growth, Prune pruning, X externalization relief, Γ pruning cost, and Ξ externalization cost.

Normal-Form Status

This chapter does not assert a universal oscillator theorem. It defines a normal-form regime structure. Depending on parameters and reset maps, an FDS may exhibit oscillation, monotone collapse, metastable maintenance, one-time failure, or externalized persistence.

10.2. Phase A and Phase B

Definition 10.1 (Phase A)**.** Phase A is the growth-dominant regime in which capacity deficit and novelty drive model proliferation, exploration, expansion, or complexity accumulation:

$$G(\Delta_\varepsilon(\tau), C) > \text{Prune}(S, C) + X(E^{\text{ext}}, C).$$

Definition 10.2 (Phase B)**.** Phase B is the maintenance, pruning, externalization, or collapse-dominant regime in which accumulated complexity cannot continue unchecked:

$$\text{Prune}(S, C) + X(E^{\text{ext}}, C) \geq G(\Delta_\varepsilon(\tau), C)$$

or boundary maintenance fails.

Proposition 10.3 (Normal-Form Regime)**.** *Under finite capacity, finite free-energy input, positive deficit-driven complexity pressure, and nonzero maintenance cost, an active-boundary FDS cannot remain indefinitely in unbounded Phase-A complexity growth without pruning, externalization, task relaxation, or collapse.*

Proof. If C grows without bound while maintenance cost remains positive and nondecreasing in maintained complexity, then $\dot{Q}_{\text{total}}$ eventually exceeds bounded $\dot{F}_{\text{in}}$ unless pruning, externalization, task relaxation, or compression improvement offsets the growth. By Theorem 7.5, the system must exit unbounded Phase-A growth through one of the admissible channels. □

10.3. Oscillation Is Conditional

Proposition 10.4 (Conditional oscillation)**.** *If a system has replenishing free-energy input, threshold-triggered pruning or reset, and a return map whose fixed point is stable, then Phase-A/Phase-B oscillations may occur. Such oscillations are model-dependent and not universal consequences of the FDS core.*

Proof. Threshold-triggered reset maps can define a Poincare return map on a reduced state surface. If that map has a stable fixed point, oscillatory cycling follows. The result depends on the reset rule and stability hypotheses, and therefore belongs to normal-form dynamics rather than the formal core. □

Part VI.

Quarantined Applications: Domain Bridges and Operational Tests

11. Application Protocol

11.1. The Three-Part Application Form

Protocol 11.1 (Domain Bridge Protocol)**.** Every application of Distinction Theory outside the formal FDS core must be written in the following form:

$$\text{FDS theorem} + \text{domain bridge assumptions} + \text{operational test} \Rightarrow \text{domain conclusion.}$$

Failure of a domain conclusion may invalidate the bridge assumptions or the operational mapping without invalidating the formal FDS theorem.

11.2. Application Claim Template

Each application must specify:

(1) the system boundary B;

(2) the internal memory/model space M;

(3) the update operator U;

(4) whether the boundary is active under Criterion 2.6;

(5) the task-relevant variable Z_t;

(6) the rate-distortion demand $R^{(\tau)}_{\min}(\varepsilon)$ over the admissible class;

(7) the capacity estimate C_S;

(8) the erasure term $H(M_t \mid M_{t+1}, Y_t)$, if any;

(9) the free-energy or resource budget Φ;

(10) the perturbation family $\mathcal{P}$;

(11) the proposed invariant quotient q, if persistence is claimed;

(12) the falsification condition.

12. Example: Artificial Agents

12.1. Non-Claim

Not claimed. The FDS core does not prove that any named model class can or cannot realize artificial general intelligence. It only states conditional constraints on active finite systems.

12.2. Conditional AI Bridge

Let an artificial system have state

$$M_t = (\theta, c_t, m_t^{\text{ext}}),$$

where θ are parameters, c_t is transient context or working state, and m_t^{ext} is durable external memory or environmental state.

Proposition 12.1 (Passive-mapper limitation, conditional)**.** *If an artificial system has fixed parameters, no durable self-updating memory, no active pruning, no causal boundary control, and no externalization channel capable of reducing task-relevant capacity deficit, then it remains outside the strong FDS-agency class. Its apparent agency must be supplied by prompt structure, external operators, tools, environment, or wrapper systems.*

Proof. Under the stated conditions, the system lacks durable updates participating in future boundary-maintenance loss and lacks internal mechanisms for pruning or externalizing accumulated task-relevant distinctions. By Criterion 2.6, it does not qualify as a strong active-boundary agent. Any persistent agency must therefore be attributed to the larger coupled system that includes tools, memory, environment, or human operators. □

Remark 12.2. A tool-using agent with writable memory, causal interventions, active pruning, and durable boundary control may satisfy the FDS conditions even if its base model has fixed parameters. The relevant system is the whole coupled architecture, not the neural network weights alone.

13. Other Domain Examples (Minimal)

Further domain mappings follow the same protocol: a biological organism maps membrane, regulatory memory, metabolism, repair, and ecological niche to B, M, Φ, pruning, and externalization; a disease process can be studied as a competing active-boundary subsystem if it exhibits boundary maintenance, resource acquisition, and adaptive updating; a civilization maps institutions, infrastructure, records, energy throughput, and governance to the FDS components. All such claims are domain-bridge theses, not theorems of the formal core, and require the full Application Protocol for operational testing.

Part VII.

Appendices

A. Notation Summary

Symbol	Meaning
X	Internal state space
E	Environmental state space
B	Boundary/interface variable
M	Internal memory/model space
Y	Observation channel
A	Action or boundary-update space
U	Internal update operator
π	Finite distinction projection / coarse-graining
ℓ	Boundary-maintenance loss
Φ	Resource or free-energy budget functional
$\mathcal{P}$	Perturbation/pruning/coarse-graining family
τ	Update timescale
C_S	Internal capacity, typically $\log_2 \|M\|$
Ψ	Admissible class of statistics $\psi : (E, B) \to Z$
$R^{(\tau)}_{\min}(\varepsilon)$	Minimal rate-distortion demand, $\inf_{\psi\in\Psi} R^{(\tau)}_{\psi(E,B)}(\varepsilon)$
$\Delta_\varepsilon(\tau)$	Capacity deficit over window τ, $R^{(\tau)}_{\min}(\varepsilon) - C_S$
b_t	Logical erasure, $H(M_t \mid M_{t+1}, Y_t)$
q	Quotient or invariant feature map
R_A	Structural identity predicate

B. Theorem Dependency Table

Claim	Proof status	Dependencies and failure consequence
Active Boundary Qualification	Definitional / Domain criterion	Fails only if active-boundary definition is replaced or shown incoherent.
Finite Distinction Representation	Formal	Depends on finite internal capacity and information inequalities.
Capacity Deficit Theorem	Formal / information-theoretic	Depends on rate-distortion formalization and capacity estimate.
Approximation Proliferation	Conditional formal	Depends on persistent novelty and exclusion of pruning, externalization, task relaxation, and compression improvement.
Dissipation Floor	Physical bridge	Depends on Landauer conditions for physical irreversible erasure.
Trichotomy	Conditional physical	Depends on finite resource budget and maintenance cost exceeding input.
Invariant Persistence	Formal	Depends on existence of invariant quotient and perturbation family.
Normal-Form Regime	Normal-form	Describes regimes; does not claim universal oscillation.
Domain applications	Domain-bridge	Fail locally if mappings or operational tests fail.

B.1. Layered Defense Matrix

The three-layer architecture ensures that failure never propagates upward:

Layer	Content	Failure Consequence
L0 (Formal Core)	The FDS definitions and core theorems	Can only be killed by a mathematical counterexample or logical inconsistency.
L1 (Physical Bridge)	Landauer bridge, dissipation floor, free-energy budgets	Failure suspends only the thermodynamic corollaries; L0 survives.
L2 (Quarantined Applications)	Life, intelligence, AI, civilization	Failure archives the specific domain bridge; no upstream damage.

C. Failure Registry

C.1. Formal Core Failures

A formal core claim fails if its definitions are inconsistent or a counterexample satisfies the hypotheses while violating the conclusion.

C.2. Physical Bridge Failures

The physical bridge fails if physical irreversible erasure under stated thermodynamic conditions reliably violates the Landauer lower bound, or if the assumed thermal, locality, or coupling conditions do not apply.

C.3. Normal-Form Failures

A normal-form claim fails only as a universal claim if treated as universal. In this document it is not universal; it is a regime model. Specific normal forms may fail to fit specific systems.

C.4. Application Failures

An application fails if its domain bridge assumptions are false, its operational variables are misidentified, or its predicted test outcome fails. Such failure does not propagate to the formal FDS core unless the application was the only support for a formal premise, which is explicitly disallowed.

D. Bibliographic Orientation

The formal core is naturally adjacent to several existing bodies of mathematics and physics:

- Shannon information theory and rate-distortion theory [2, 3, 5];
- stochastic thermodynamics and Landauer's principle [7, 8, 10, 11];
- coarse-graining, renormalization group, and effective state descriptions;
- quotient spaces, invariants, fixed points, and algebraic topology;
- Markov blankets and active-boundary systems [12, 13];
- dynamical systems with resource constraints.

Key references for the formal framework include the core exposition [1] and a companion physical claim registry [16], to be archived separately. The rate-distortion bound follows the standard source-coding theorem; the Landauer bridge follows the established experimental and theoretical literature [9]; the collapse and persistence trichotomy connects to resource-constrained dynamical systems and stochastic thermodynamics.

These connections do not make Distinction Theory true by association. They specify where the theory must be formalized, tested, and, if necessary, broken.

E. Independent Criteria for Invariant-Supported Classification

A structure may be classified as invariant-supported under a perturbation family if it satisfies at least one of the following criteria:

(I) **Global vs. local character.** The identity condition R_A depends on quantized, global features (topological winding numbers, Chern numbers, symmetry charges) rather than continuous, local micro-coordinates.

(II) **Gap protection.** Destruction of the structure requires closing a spectral gap or crossing a non-perturbative phase boundary.

(III) **Coarse-graining stability.** Under iterative renormalization-group flow, the defining features of R_A flow to a non-trivial fixed point.

A structure satisfying none of these criteria is identifier-dependent: its persistence relies on active maintenance, repair, isolation, or stochastic stability rather than on algebraic protection through an invariant quotient.

Bibliography

[1] Y. Wu, *Distinction Theory: A General Theory of Finite Systems*, Zenodo (2026), doi:10.5281/zenodo.20130174.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal* **27**, 379–423 and 623–656 (1948).

[3] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE National Convention Record* **7**, 142–163 (1959).

[4] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression* (Prentice-Hall, Englewood Cliffs, 1971).

[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, Hoboken, 2006).

[6] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, "Thermodynamics of information," *Reviews of Modern Physics* **87**, 45–77 (2015), doi:10.1103/RevModPhys.87.45.

[7] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM Journal of Research and Development* **5**, 183–191 (1961), doi:10.1147/rd.53.0183.

[8] C. H. Bennett, "Logical reversibility of computation," *IBM Journal of Research and Development* **17**, 525–532 (1973).

[9] A. Bérut *et al.*, "Experimental verification of Landauer's principle linking information and thermodynamics," *Nature* **483**, 187–189 (2012), doi:10.1038/nature10872.

[10] U. Seifert, "Entropy production along a stochastic trajectory and an integral fluctuation theorem," *Physical Review Letters* **95**, 040602 (2005).

[11] J. M. Horowitz and T. Sagawa, "Equivalent information-theoretic and thermodynamic formulations of the second law in feedback-controlled systems," *Journal of Statistical Mechanics: Theory and Experiment* **2015**, P03022 (2015), doi:10.1088/1742-5468/2015/03/P03022.

[12] J. Pearl, "Bayesian networks: a model of self-activated memory for evidential reasoning," *Proceedings of the 7th Conference of the Cognitive Science Society*, 329–334 (1985).

[13] K. Friston, "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience* **11**, 127–138 (2010).

[14] J. G. Kemeny and J. L. Snell, *Finite Markov Chains* (Springer, New York, 1976).

[15] J. R. Norris, *Markov Chains* (Cambridge University Press, Cambridge, 1997).

[16] Y. Wu, *Physical Bridge Claim Registry for Active Finite Distinction Systems: Time, Observers, Horizons, Topological Persistence, and High-Risk Physical Extensions*, Zenodo, 2026. DOI to be added.