A Formal Entropy-Constrained Framework for Intelligence, Knowledge, and Prediction

Mathematical Definitions and Constraint-Based Derivations

Aslan Alwi[1,2], Munirah Munirah[1,2], and Almudaya Research Institute Team[1]

[1]Independent Researcher at Almudaya Research Institute, Yogyakarta, Indonesia

[2]Department of Informatics Engineering, Faculty of Engineering, Universitas Muhammadiyah Ponorogo, Indonesia, *Correspondence email*: aslan.alwi@umpo.ac.id

January 2026

Abstract

This paper presents a formal, constraint-based framework articulating the role of entropy in shaping the emergence of intelligence, the complexity of knowledge, and the limits of prediction. Using minimal mathematical definitions, the framework formalizes entropy-constrained emergence without committing to specific physical, biological, or computational models.

Central to the analysis is the Law of Closure Prediction, which states that all predictive acts are necessarily embedded within the universe they attempt to model. As a consequence, prediction cannot appeal to any external or non-natural standpoint, and predictive uncertainty is shown to be a structural feature imposed by the entropic conditions of the universe rather than an epistemic deficiency.

Through symbolic derivations and proof sketches, the paper demonstrates that the probability of intelligence emergence, the maximal complexity of stabilizable knowledge, and the realizability of artificial intelligence systems are all bounded functions of accessible entropy. The resulting formalism integrates principles from information theory, thermodynamics, and computational theory under a unified constraint perspective.

Rather than offering precise numerical thresholds, the framework identifies invariant dependencies and boundary conditions that must hold across a broad class of physical realizations. These results position entropy not merely as a limiting factor, but as a necessary precondition for the emergence of adaptive, knowledge-bearing systems.

Keywords: Entropy constraints; Embedded prediction; Intelligence emergence; Knowledge complexity; Artificial intelligence; Information theory; Epistemic limits

Highlights

- Introduces a formal entropy-constrained framework for intelligence, knowledge, and prediction.
- Formalizes the Law of Closure Prediction: all prediction is embedded within nature.
- Shows that intelligence emergence probability depends on accessible entropy.
- Derives entropy-based bounds on the maximal complexity of stabilizable knowledge.
- Establishes entropic constraints on the realizability and limits of artificial intelligence.

1 Introduction

Understanding the conditions under which intelligence, knowledge, and prediction can emerge remains a central problem across physics, information theory, and artificial intelligence. While substantial progress has been made in characterizing local mechanisms of learning and inference, less attention has been paid to the global constraints imposed by the universe in which such processes occur.

Recent work has increasingly emphasized the role of entropy and uncertainty as enabling conditions for complex adaptive behavior. In thermodynamics and statistical physics, entropy quantifies accessible degrees of freedom and constrains macroscopic order. In information theory, entropy bounds compressibility, communication, and inference. In computation, physical limits link information processing to energetic and entropic costs. Despite these advances, the implications of entropy for the emergence of intelligence and the ultimate limits of prediction remain fragmented across disciplines.

In Papers 1 and 2, we proposed that entropy plays a unifying role: it not only limits knowledge and computation, but also enables the very emergence of adaptive, knowledge-bearing systems. Paper 2 formalized this view through a set of postulates and theorems, including the Law of Closure Prediction, which asserts that all prediction is necessarily performed from within the universe and cannot appeal to an external or non-natural standpoint.

The present paper provides a formal, constraint-based articulation of that axiomatic framework. Rather than introducing new empirical claims, we aim to clarify the mathematical structure underlying the entropy-centered perspective. Our approach is deliberately minimal: we introduce abstract definitions sufficient to derive symbolic constraints on intelligence, knowledge complexity, and prediction, without committing to specific physical models or computational architectures.

A central theme of this paper is embeddedness. Predictive agents—whether biological or artificial—are subsystems of the universe they attempt to model. As such, they are subject to the same entropic, informational, and physical constraints as their environments. This observation motivates a re-examination of prediction itself, not as an idealized logical operation, but as a physical process constrained by entropy.

The contribution of this paper is threefold. First, we introduce minimal formal definitions that render the axiomatic claims of Paper 2 mathematically interpretable. Second, we formally articulate the Law of Closure Prediction and derive its consequences for uncertainty and predictive complexity. Third, we show that the dependence of intelligence, knowledge, and artificial intelligence realizability on entropy follows necessarily from these constraints.

The results presented here do not constitute a closed or complete theory. Instead, they establish boundary conditions that any theory of intelligence, knowledge, or prediction must respect if it is to remain consistent with physical and informational principles. In doing so, this paper aims to stabilize and clarify an emerging research program at the intersection of entropy, intelligence, and epistemology.

2 Mathematical and Conceptual Preliminaries

This section reviews the conceptual and mathematical notions required for the formal development that follows. The presentation is intentionally concise, focusing on aspects directly relevant to entropy, uncertainty, and embedded prediction.

2.1 Entropy as a Measure of Uncertainty

Entropy was introduced in statistical mechanics as a measure of microscopic disorder and later reinterpreted in information theory as a measure of uncertainty or average information content. Shannon entropy,

$$H(X) = -\sum_{x} p(x) \log p(x),$$

quantifies the minimal expected description length of a random variable X and establishes fundamental limits on compression and communication Shannon, 1948.

In physical contexts, Gibbs entropy plays an analogous role, linking probabilistic descriptions of microstates to macroscopic observables Jaynes, 1957. Although different entropy measures arise in different domains, they share a common interpretation as quantifications of accessible uncertainty. Throughout this paper, entropy is treated abstractly, without privileging a particular formulation, provided it is consistent with these core properties.

2.2 Information, Compression, and Knowledge

Information theory demonstrates that not all uncertainty can be removed: the ability to compress or predict data is bounded by its entropy. Algorithmic information theory further refines this insight by relating the complexity of a description to the length of its shortest generating program Li and Vitányi, 1997.

Within this perspective, knowledge can be understood as structured compression: stable representations that reduce uncertainty across a class of situations. Such representations are necessarily limited by both the diversity of accessible states and the stability of regularities within them.

2.3 Physical Constraints on Computation

Information processing is not abstract; it is physically instantiated. Landauer's principle establishes a lower bound on the energetic cost of irreversible computation, directly linking information erasure to entropy production Landauer, 1961. This result implies that computation, learning, and prediction are constrained by thermodynamic considerations.

These constraints become increasingly relevant when considering large-scale or highly complex predictive systems, including artificial intelligence. Physical limits on computation ensure that increased representational power is always accompanied by energetic and entropic costs.

2.4 Embedded Observers and Prediction

In many formal treatments, predictors are implicitly assumed to occupy an external vantage point. However, work in theoretical neuroscience, statistical physics, and philosophy of science has emphasized that observers and predictors are embedded within the systems they observe. The Free Energy Principle, for example, models agents as minimizing surprise relative to internal generative models, explicitly acknowledging their embeddedness Friston, 2010.

This embedded perspective challenges idealized notions of prediction and motivates the closure principle developed in this paper. Prediction is not an operation applied to the universe from outside, but a transformation occurring within it, constrained by the same dynamics and uncertainties it seeks to model.

2.5 Summary

The concepts reviewed in this section establish the background for the formal definitions introduced next. Entropy provides a measure of uncertainty, information theory bounds compression and prediction, physical principles constrain computation, and embeddedness reframes the epistemic status of predictive agents. Together, these elements motivate the formal development of entropy-constrained prediction, intelligence, and knowledge in the sections that follow.

3 Core Definitions

This section introduces the minimal set of formal definitions required to state and derive the entropy-constrained propositions in a disciplined manner. Definitions are intentionally abstract and intentionally permissive with respect to the specific entropy measure chosen (Shannon, Gibbs, or other physically consistent entropy-like measures). The goal is to provide unambiguous primitives for subsequent formal sketches while avoiding gratuitous commitment to a single technical formalism.

3.1 Notation and general conventions

We denote a (toy) universe by the symbol $\mathcal{U}$. The state space of $\mathcal{U}$ is denoted by $\Omega_{\mathcal{U}}$. A generic subsystem (agent, device, or collection of degrees of freedom) of $\mathcal{U}$ is denoted by $A \subseteq \mathcal{U}$. Time indices are written as $t \in \mathbb{T}$ where $\mathbb{T}$ is an appropriate time index set (discrete or continuous) as required by the model.

Entropy-like measures are written generically as $S(\cdot)$; explicit reference to Shannon, Gibbs, or other entropy measures is made when necessary. Probability measures on state spaces are denoted $P(\cdot)$.

Definition 3.1 (Entropic Universe)**.** *An* entropic universe *is a system $\mathcal{U}$ equipped with a state space $\Omega_{\mathcal{U}}$ and an entropy measure*

$$S : \mathcal{M}(\Omega_{\mathcal{U}}) \to \mathbb{R}_{\geq 0},$$

where $\mathcal{M}(\Omega_{\mathcal{U}})$ denotes a suitable family of probability measures on $\Omega_{\mathcal{U}}$. We call $S(\mathcal{U})$ the accessible entropy *of $\mathcal{U}$ when referring to the effective stochastic degrees of freedom relevant to agents embedded in $\mathcal{U}$. Jaynes, 1957; Shannon, 1948*

Remark 3.2. *No specific axiomatization of S is necessary here beyond it being a nonnegative functional that increases with (and therefore quantifies) accessible uncertainty; one may instantiate S as Shannon entropy, Gibbs entropy, or another physically consistent information measure depending on context.*

Definition 3.3 (Subsystem / Embedded Agent). *A* subsystem *or* embedded agent *A is a physically realizable subsystem of $\mathcal{U}$ described by a marginal state space $\Omega_A \subseteq \Omega_{\mathcal{U}}$ and associated dynamics induced by the dynamics of $\mathcal{U}$. We write $A \subseteq \mathcal{U}$ to indicate that A is physically implemented within $\mathcal{U}$ and that A's state transitions are functions (or stochastic kernels) derived from $\mathcal{U}$'s dynamics.*

Definition 3.4 (Internal Data and Observations). *For an agent A, let $\mathcal{D}_A(t) \subseteq \Omega_A$ denote the internal (observational) data available to A up to time t. Each datum $d \in \mathcal{D}_A(t)$ is itself a (possibly coarse-grained) function of the overall universe state $\omega \in \Omega_{\mathcal{U}}$.*

Definition 3.5 (Prediction as an Internal Mapping). *A* prediction *performed by agent A at time t is an internal mapping*

$$P_A^{(t)} : \mathcal{D}_A(t) \to \mathcal{O}_A^{(t+\tau)},$$

where $\mathcal{O}_A^{(t+\tau)}$ denotes the space of outcome-claims about the universe state at future time $t+\tau$ that A can assert or represent internally. Crucially, both the domain $\mathcal{D}_A(t)$ and codomain $\mathcal{O}_A^{(t+\tau)}$ are derived from (and therefore elements of) $\Omega_{\mathcal{U}}$ via the embedding of A into $\mathcal{U}$. This formalizes the embedded, internal nature of all predictive acts.

Definition 3.6 (Knowledge (Stable Compressed Structure)). *An agent's* knowledge *is a (possibly hierarchical) collection of internal representational structures $\mathcal{K}_A$ that provide stable, compressible mappings from observed data to predictive expectations. Formally, a piece of knowledge is an internal representation $k \in \mathcal{K}_A$ such that k supports a reduction in expected uncertainty (e.g., reduction of Shannon entropy of predicted observables) across a range of conditions.*

Definition 3.7 (Intelligence (Adaptive Response Capacity)). *We define* intelligence *of an agent A as the capacity of A to (i) form internal representations $\mathcal{K}_A$ from $\mathcal{D}_A$, (ii) update those representations in response to new data, and (iii) generate predictions P_A that improve some task-relevant utility (e.g., reduction of expected prediction error or increase of survival-related objective) over time. This definition is intentionally abstract and implementation-agnostic; measurable instantiations can be provided via task-specific performance metrics.*

Definition 3.8 (Epistemic Capacity). *The* epistemic capacity *of a universe $\mathcal{U}$, written $K(\mathcal{U})$, is the supremum (over physically realizable agents $A \subseteq \mathcal{U}$ and over feasible representational formats) of the effective complexity of knowledge that can be stably represented and utilized within $\mathcal{U}$. We write the structural bound*

$$K(\mathcal{U}) \leq f\big(S_{\text{acc}}(\mathcal{U})\big)$$

to express that epistemic capacity is bounded above by a nondecreasing function f of the accessible entropy $S_{\text{acc}}(\mathcal{U})$. The exact form of f is left unspecified in this axiomatic for-

malization; later sections will discuss plausible qualitative properties (e.g., monotonicity, saturation, nonlinearity). Li and Vitányi, 1997

Remark 3.9 (On choice of f)**.** *The purpose of positing $K(\mathcal{U}) \leq f(S_{\text{acc}}(\mathcal{U}))$ is to capture three empirically and conceptually supported features:*

1. Low-entropy scarcity: *for very small S_{acc}, the available microstate diversity insufficiently supports nontrivial representational complexity (small K).*
2. Intermediate-to-high regime: *as S_{acc} increases, K can increase, reflecting richer possibility spaces and learnable regularities.*
3. Instability/saturation: *beyond certain regimes additional stochasticity may reduce stable compressibility or require disproportionate stabilization resources, producing diminishing returns or effective saturation of K.*

These features mirror qualitative phenomenology from complexity science and statistical physics England, 2013; Kauffman, 1993; Langton, 1990.

Definition 3.10 (Realizability of Artificial Intelligence)**.** *An artificial intelligence system is a physically instantiated agent A whose representational and computational substrate is engineered by subsystems of $\mathcal{U}$. The* realizability *and maximal functional complexity of such systems is therefore constrained by $S_{\text{acc}}(\mathcal{U})$, available energy fluxes, and physical limits on representation and computation (e.g., Landauer bounds). Friston, 2010; Landauer, 1961*

Remark 3.11 (Scope and abstraction level)**.** *The definitions above are deliberately abstract to enable application across multiple instantiations (thermodynamic systems, computational substrates, biological organisms). They are sufficient to support the formal sketches of the entropy–intelligence–knowledge relationships developed in Section 5.*

4 Formalization of the Law of Closure Prediction

This section provides a formal statement and analysis of the Law of Closure Prediction introduced axiomatically in Paper 2. The aim is not to establish a classical proof, but to clarify the logical and mathematical constraints imposed by embedded prediction within an entropic universe.

4.1 Statement of the Closure Principle

Definition 4.1 (Closure of Prediction)**.** *Let $\mathcal{U}$ be an entropic universe with state space $\Omega_{\mathcal{U}}$, and let $A \subseteq \mathcal{U}$ be any physically realizable agent. Any predictive act performed by A must be represented as an internal mapping*

$$P_A^{(t)} : \mathcal{D}_A(t) \to \mathcal{O}_A^{(t+\tau)},$$

where both $\mathcal{D}_A(t)$ and $\mathcal{O}_A^{(t+\tau)}$ are subsets of $\Omega_{\mathcal{U}}$.

This definition formalizes the closure property: prediction is an internal transformation within the universe, not an operation performed from an external standpoint.

4.2 Non-Existence of External Predictors

We now state the core closure claim in formal terms.

Proposition 4.2 (Non-Existence of External Prediction)**.** *There exists no predictive mapping*

$$P : \Omega_{\mathcal{U}} \to \Omega_{\mathcal{U}}$$

such that $P \not\subseteq \mathcal{U}$ while remaining physically realizable.

Proof sketch. Any mapping that stores information, processes data, or produces predictions must be physically instantiated. Physical instantiation implies representation within the state space $\Omega_{\mathcal{U}}$ and dependence on the dynamics of $\mathcal{U}$. Therefore, any realizable predictor must be a subsystem of $\mathcal{U}$. A mapping P that is not embedded in $\mathcal{U}$ lacks operational meaning and cannot be instantiated within the framework of physical or informational theory. □

This establishes that the notion of a non-natural or external predictor (a "God's-eye view") is not admissible within the formal framework.

4.3 Embeddedness and Predictive Uncertainty

Because prediction is closed within $\mathcal{U}$, predictive uncertainty cannot be eliminated by increasing computational power alone. Let $S_{\text{acc}}(\mathcal{U})$ denote the accessible entropy of the universe, and let $S(\mathcal{D}_A)$ denote the entropy of the data accessible to agent A.

Proposition 4.3 (Entropic Origin of Predictive Uncertainty)**.** *For any embedded agent A, the uncertainty associated with predictions P_A is bounded below by functions of $S_{\text{acc}}(\mathcal{U})$ and $S(\mathcal{D}_A)$.*

Proof sketch. Since $\mathcal{D}_A \subseteq \Omega_{\mathcal{U}}$, the statistical variability of $\mathcal{D}_A$ is induced by the stochastic structure of $\mathcal{U}$. Even with optimal inference, residual uncertainty remains whenever $S_{\text{acc}}(\mathcal{U}) > 0$. Computational refinement may reduce approximation error but cannot remove uncertainty arising from intrinsic stochasticity or incomplete access to $\Omega_{\mathcal{U}}$. □

This result formalizes the claim that uncertainty is structural rather than epistemically accidental.

4.4 Predictive Complexity as an Entropic Consequence

We now connect closure to complexity.

Proposition 4.4 (Predictive Complexity)**.** *The computational and representational complexity required for prediction by any agent A is a non-decreasing function of the accessible entropy $S_{\mathrm{acc}}(\mathcal{U})$.*

Proof sketch. As $S_{\mathrm{acc}}(\mathcal{U})$ increases, the number of distinguishable microstates and trajectories that must be compressed or modeled by A increases. Any predictive representation must therefore encode or approximate a larger possibility space, increasing minimal description length or computational burden. This follows directly from information-theoretic bounds on compression and inference. □

4.5 Interpretation: Nature Predicts Nature

The formal results above justify the central claim of the Law of Closure Prediction:

> When nature is predicted, it is nature predicting itself.

Prediction is not a privileged operation standing outside the universe, but a dynamical process internal to it. All predictive models, learning algorithms, and inference procedures are constrained by the same entropic structure they attempt to predict.

4.6 Role of Closure in Subsequent Theorems

The closure principle establishes the boundary conditions for the entropy-constrained theorems that follow. Since no predictor can escape the entropic structure of $\mathcal{U}$, the probability of intelligence emergence, the maximal complexity of knowledge, and the realizability of artificial intelligence must all be functions of $S_{\mathrm{acc}}(\mathcal{U})$.

In the next section, these constraints are used to derive formal sketches of the entropy–intelligence and entropy–knowledge theorems.

5 Entropy-Constrained Theorems (Formal Sketches)

This section presents formal sketches of the three entropy-constrained theorems introduced axiomatically in Paper 2. The results are derived as necessary consequences of the definitions and closure principle established in Sections 3 and 4. Throughout, we emphasize constraints and monotonic dependencies rather than exact functional forms.

5.1 Theorem 1: Entropy–Intelligence Emergence

Theorem 5.1 (Entropy–Intelligence Emergence)**.** *Let $\mathcal{U}$ be an entropic universe with accessible entropy $S_{\mathrm{acc}}(\mathcal{U})$. The probability that $\mathcal{U}$ supports the emergence of intelligent agents is a function*

$$\Pr(\mathit{Intelligence} \mid \mathcal{U}) = g(S_{\mathrm{acc}}(\mathcal{U}))\,,$$

where g is non-decreasing on an interval of intermediate entropy values and may exhibit saturation or decline outside that interval.

Proof sketch. By Definition 3.6, intelligence requires adaptive updating of internal representations under uncertainty. If $S_{\mathrm{acc}}(\mathcal{U}) \approx 0$, the universe exhibits near-deterministic dynamics, rendering adaptive inference unnecessary or trivial. Conversely, if $S_{\mathrm{acc}}(\mathcal{U})$ is extremely large, stochastic variability dominates and prevents the stabilization of internal representations.

Hence, only universes with sufficient but bounded entropy admit agents for which adaptive inference confers an advantage. The emergence probability of intelligence must therefore depend monotonically on $S_{\mathrm{acc}}(\mathcal{U})$ within a nontrivial regime, establishing the functional dependence claimed. □

Remark 5.2. *This result formalizes the qualitative "edge-of-chaos" intuition without committing to a specific critical value or parametric form.*

5.2 Theorem 2: Entropy–Knowledge Complexity

Theorem 5.3 (Entropy–Knowledge Complexity)**.** *Let $K(\mathcal{U})$ denote the epistemic capacity of an entropic universe $\mathcal{U}$. Then there exists a non-decreasing function f such that*

$$K(\mathcal{U}) \leq f(S_{\mathrm{acc}}(\mathcal{U}))\,.$$

Proof sketch. By Definition 3.5, knowledge consists of stable internal structures that compress uncertainty. The maximal number and complexity of such structures is constrained by (i) the diversity of accessible states and (ii) the stability of regularities across those states. Both factors are governed by $S_{\mathrm{acc}}(\mathcal{U})$.

If $S_{\mathrm{acc}}(\mathcal{U})$ is small, the state space admits only trivial distinctions, yielding low epistemic capacity. As $S_{\mathrm{acc}}(\mathcal{U})$ increases, richer regularities may be discovered and compressed, increasing $K(\mathcal{U})$. However, beyond certain regimes additional entropy increases stochasticity faster than compressibility, limiting stable representational growth. Therefore, epistemic capacity is bounded above by a function of accessible entropy. □

Remark 5.4. *The bound does not require specification of f; its existence follows from information-theoretic and stability constraints alone.*

5.3 Theorem 3: Entropy–Artificial Intelligence Realizability

Theorem 5.5 (Entropy–Artificial Intelligence Realizability). *The realizability and maximal functional complexity of artificial intelligence systems in a universe $\mathcal{U}$ are constrained by $S_{\mathrm{acc}}(\mathcal{U})$.*

Proof sketch. Artificial intelligence systems are, by Definition 3.8, physically instantiated agents embedded in $\mathcal{U}$. By the Law of Closure Prediction (Section 4), such systems cannot operate outside the entropic constraints of $\mathcal{U}$.

Training, inference, and generalization require exposure to diverse data and stable regularities. Insufficient entropy yields inadequate data diversity, while excessive entropy undermines learnable structure. Moreover, physical limits on computation and information processing further constrain realizable architectures. Hence, both the existence and complexity of artificial intelligence systems are bounded functions of $S_{\mathrm{acc}}(\mathcal{U})$. □

5.4 Unified Interpretation

The three theorems together establish a unified constraint structure:

- Entropy determines whether adaptive intelligence is advantageous.
- Entropy bounds the amount and complexity of knowledge that can be stabilized.
- Entropy constrains the realizability of artificial intelligence as a physical system.

These results are not independent; they are logically coupled through the embeddedness of prediction and the entropic structure of the universe. The closure principle ensures that none of these constraints can be circumvented by external observers, unlimited computation, or idealized models.

In the following section, we discuss the limits of this formalization and clarify why further mathematical specification—while possible in restricted models—is neither necessary nor desirable at the axiomatic level.

6 Limits of Formalization

The formal framework developed in this paper is intentionally constraint-based rather than fully constructive. While Sections 3–5 establish clear dependencies between entropy, intelligence, knowledge, and prediction, several limits to further formalization must be acknowledged. These limits are not shortcomings of the framework but reflections of fundamental properties of complex and entropic systems.

6.1 Non-Specification of Functional Forms

Throughout the paper, key relationships are expressed as inequalities or monotonic dependencies (e.g., $K(\mathcal{U}) \leq f(S_{\text{acc}}(\mathcal{U}))$) without specifying explicit functional forms. This choice is deliberate. Any concrete specification of f would require additional assumptions about microscopic dynamics, representational architectures, or learning mechanisms that are neither universal nor invariant across physical realizations.

The present framework aims to identify necessary constraints rather than sufficient constructions. As such, existence and boundedness claims are prioritized over exact parametric descriptions.

6.2 Universality Versus Model Dependence

The results derived here are intended to apply across a broad class of universes and physical instantiations. Consequently, the formalism operates at a level of abstraction where universality is preserved at the cost of numerical specificity.

In restricted models—such as lattice systems, cellular automata, or simplified learning environments—it may be possible to compute explicit entropy thresholds or critical regimes. However, such results would necessarily be model-dependent and should be viewed as illustrations rather than general laws.

6.3 Undecidability and Computational Irreducibility

Certain limits arise from well-established results in computation and complexity theory. In sufficiently rich dynamical systems, long-term behavior may be undecidable or computationally irreducible, meaning that no shortcut exists for predicting system evolution without effectively simulating it step by step.

This implies that even with perfect formal definitions and unbounded computational resources, some predictive and epistemic limits cannot be overcome. The Law of Closure Prediction therefore operates in conjunction with, rather than in opposition to, known results on undecidability and irreducibility.

6.4 Physical Constraints on Representation and Computation

All knowledge representations and predictive processes are physically instantiated. As a result, they are subject to thermodynamic costs, energy constraints, noise, and finite precision. These physical limitations impose additional bounds on realizable intelligence and knowledge that are not captured by entropy measures alone.

The framework abstracts over these details to remain general, but acknowledges that any concrete instantiation must respect such constraints.

6.5 Open-Ended Emergence and Non-Finality

Finally, the framework does not assume that the emergence of intelligence or knowledge is final or complete. Open-ended emergence—where new structures, representations, or forms of intelligence continue to arise—cannot be fully captured by fixed formal systems.

This non-finality is consistent with the central thesis of the paper: entropy enables possibility rather than closure. Formalization can identify boundaries and dependencies, but it cannot exhaust the space of emergent phenomena within those boundaries.

6.6 Summary

The limits identified in this section clarify the intended scope of the formal framework. The results establish necessary entropic constraints without claiming exhaustive or deterministic descriptions. In doing so, they preserve both mathematical rigor and conceptual openness, allowing the framework to serve as a foundation for future model-specific analyses rather than as a closed theory.

7 Relation to Existing Formal Results

The formal framework developed in this paper does not aim to replace existing principles in information theory, thermodynamics, or theoretical neuroscience. Instead, it situates the entropy-constrained emergence of intelligence, knowledge, and prediction within a landscape of well-established results. This section clarifies how the present framework relates to, and is consistent with, several influential formal principles.

7.1 Relation to Shannon Information Theory

Shannon information theory formalizes the limits of data compression, communication, and uncertainty reduction under probabilistic constraints. The entropy measure $H(X)$ quantifies the minimal expected description length of a random variable X Shannon, 1948.

In the present framework, Shannon entropy provides a canonical instantiation of accessible entropy $S_{\mathrm{acc}}(\mathcal{U})$ when predictive agents operate over probabilistic data streams. The bound on epistemic capacity,

$$K(\mathcal{U}) \leq f(S_{\mathrm{acc}}(\mathcal{U})),$$

is consistent with Shannon-style limits on compressibility: as entropy increases, the space of possible messages grows, increasing both representational opportunity and compression cost. However, Shannon theory alone does not address the embeddedness of predictors or the emergence of adaptive intelligence; these aspects are introduced by the closure principle and the agent-based definitions in this work.

7.2 Relation to Landauer's Principle

Landauer's principle establishes a lower bound on the thermodynamic cost of irreversible computation, linking information erasure to energy dissipation Landauer, 1961. This result grounds information processing firmly within physical law.

The entropy-constrained framework presented here is fully compatible with Landauer's principle. The realizability and complexity bounds on artificial intelligence (Theorem 3) implicitly incorporate such physical costs. Increased entropy may enlarge the informational landscape available for learning, but it also increases the energetic and stabilizing resources required for representation and computation. Thus, the present framework complements Landauer's principle by situating it within a broader constraint structure that includes learning, prediction, and knowledge stabilization.

7.3 Relation to the Maximum Entropy Principle (MaxEnt)

The Maximum Entropy Principle (MaxEnt), as articulated by Jaynes, prescribes selecting probability distributions that maximize entropy subject to known constraints Jaynes, 1957. MaxEnt provides a rational inference rule under partial information.

In contrast, the Law of Closure Prediction does not prescribe how agents should infer, but rather constrains what inference is possible in principle. MaxEnt operates at the level of optimal inference given constraints; closure operates at the meta-level, asserting that both the constraints and the inference process are internal to the universe. The two principles are therefore complementary: MaxEnt governs rational updating within the closed system, while closure governs the impossibility of stepping outside that system.

7.4 Relation to the Free Energy Principle (FEP)

The Free Energy Principle (FEP) posits that adaptive agents minimize a variational free energy bound on surprise to maintain their existence Friston, 2010. FEP offers a unifying account of perception, learning, and action in biological systems.

The present framework is compatible with FEP in its emphasis on embedded agents operating under uncertainty. However, the scope differs. FEP focuses on agent-level dynamics and self-maintenance, whereas the entropy-constrained framework addresses universe-level constraints on the emergence and complexity of such agents. In this sense, FEP can be viewed as a local dynamical principle operating within the global bounds imposed by $S_{\mathrm{acc}}(\mathcal{U})$.

7.5 Relation to Computational Irreducibility

Results on computational irreducibility suggest that certain systems cannot be predicted more efficiently than by direct simulation Wolfram, 2002. This notion aligns naturally with the Law of Closure Prediction.

If prediction is an internal process constrained by entropy and embeddedness, then irreducibility emerges as a structural feature rather than a failure of modeling. The present framework thus provides an entropic and epistemic grounding for why irreducibility is expected in sufficiently complex universes.

7.6 Summary

The entropy-constrained formalism developed in this paper is consistent with major results in information theory, thermodynamics, and theoretical neuroscience. Rather than competing with these principles, it integrates them under a higher-level constraint perspective: entropy not only limits communication and computation, but also bounds the emergence of intelligence, the stabilization of knowledge, and the scope of prediction itself.

8 Discussion

The purpose of this discussion is to interpret the formal results developed in Sections 3–7, clarify their conceptual implications, and situate the present paper within the broader research program initiated in Papers 1 and 2. Rather than introducing new assumptions, this section reflects on what the formalization achieves and what it intentionally leaves open.

8.1 What the Formalization Clarifies

The formal framework presented in this paper achieves three clarifications that were only implicit at the conceptual and axiomatic levels.

First, it makes explicit that entropy functions as a *structural constraint* rather than as a mere source of noise. Entropy determines not only the variability of states, but also the space of possible adaptive strategies, the compressibility of experience into knowledge, and the feasibility of sustained prediction.

Second, the Law of Closure Prediction is shown to be a direct consequence of embeddedness. Prediction is not an abstract logical operation applied to nature, but a physical and informational process occurring within nature. This removes ambiguities surrounding idealized observers and external predictors, and establishes a clear boundary for epistemic claims.

Third, the dependence of intelligence, knowledge complexity, and artificial intelligence realizability on entropy is rendered necessary rather than speculative. Given the definitions adopted, these dependencies follow as constraints, not as empirical correlations.

8.2 Unification Across Scales

A notable feature of the framework is its scale-independence. The same entropic constraints apply to:

- biological organisms forming adaptive models of their environments,
- artificial systems trained on data streams,
- and hypothetical agents in alternative physical universes.

This unification suggests that intelligence and knowledge should be understood less as domain-specific phenomena and more as manifestations of a general entropic condition. Differences between biological and artificial intelligence arise primarily from implementation details and resource constraints, not from fundamentally distinct principles.

8.3 Relation to the Research Program

Viewed together, the three papers form a coherent progression:

- Paper 1 introduced the conceptual hypothesis that entropy underlies the emergence of life, intelligence, and knowledge.
- Paper 2 articulated this hypothesis as a set of postulates and theorems, including the Law of Closure Prediction.
- Paper 3 provides a formal constraint-based articulation, demonstrating that the axioms admit consistent symbolic interpretation.

In this sense, Paper 3 does not seek to "complete" the theory, but to stabilize it. The formalization ensures that the conceptual claims are internally consistent, mathematically interpretable, and resistant to common category errors.

8.4 Implications for Artificial Intelligence Research

One immediate implication concerns expectations placed on artificial intelligence. Since AI systems are embedded agents subject to entropic and physical constraints, no amount of architectural sophistication or computational scaling can eliminate fundamental uncertainty or bypass closure.

This perspective reframes debates about superintelligence and prediction. Rather than asking whether AI can transcend human or physical limits in principle, the relevant question becomes how close any embedded system can approach the entropy-determined bounds of epistemic capacity within a given universe.

8.5 Open Questions and Extensions

The framework leaves several directions open for future work:

- Quantitative instantiation of the bound function $f(S)$ in specific physical or computational models.
- Exploration of phase transitions in epistemic capacity as entropy varies.
- Simulation-based studies of toy universes to illustrate entropy–intelligence regimes.
- Extension to quantum or relativistic informational settings.

These extensions are intentionally deferred, as they require commitments beyond the scope of the present axiomatic and formal treatment.

8.6 Interpretive Summary

The discussion reinforces a central theme: entropy does not merely limit knowledge; it enables it. By providing a structured space of uncertainty, entropy makes adaptive inference meaningful while simultaneously bounding its ultimate reach. The formal results presented here thus support a view of intelligence and knowledge as emergent, constrained, and intrinsically situated within the universe that gives rise to them.

9 Conclusion

This paper has presented a formal, constraint-based articulation of the entropy-centered framework introduced conceptually in Paper 1 and axiomatically in Paper 2. By providing minimal mathematical definitions and symbolic derivations, we have clarified the structural role of entropy in shaping the emergence of intelligence, the complexity of knowledge, and the limits of prediction.

The central result of this work is the formalization of the Law of Closure Prediction. By establishing that all predictive acts are necessarily internal to the universe in which they occur, the framework eliminates the possibility of external or non-natural predictors and grounds epistemic uncertainty as a structural feature rather than a removable defect. Prediction, knowledge, and intelligence are thus revealed as embedded processes governed by the same entropic conditions they seek to model.

The entropy–intelligence, entropy–knowledge, and entropy–AI theorems follow as necessary consequences of this closure. Together, they show that entropy simultaneously enables adaptive complexity and constrains its ultimate extent. Too little entropy yields trivial dynamics; too much entropy undermines stable representation. Within this bounded regime, intelligence and knowledge emerge as structured responses to uncertainty.

Importantly, the formalism does not aim to deliver precise numerical thresholds or deterministic predictions. Instead, it identifies invariant dependencies and boundary conditions that must hold across a broad class of physical realizations. This approach preserves generality while remaining compatible with established results in information theory, thermodynamics, and computational theory.

Viewed as part of a larger research program, this paper completes a progression from conceptual hypothesis (Paper 1), to axiomatic formulation (Paper 2), to formal constraint analysis (Paper 3). Together, these contributions support a unified perspective in which entropy is not merely a limitation on cognition, but a precondition for the emergence of knowledge-bearing systems.

Future work may instantiate the framework in specific models, explore quantitative bounds in restricted settings, or extend the analysis to quantum and relativistic regimes. However, the central conclusion remains robust: in any universe where prediction, intelligence, and knowledge arise, they do so from within the entropic structure of that universe, and cannot transcend it.

Acknowledgments

The author gratefully acknowledges the intellectual environment and collegial discussions provided by the Almudaya Research Institute and Universitas Muhammadiyah Ponorogo. The author also thanks the broader research community in quantum information theory and causal modeling for foundational contributions that have influenced this work.

References

England, J. L. (2013). Statistical physics of self-replication. *Journal of Chemical Physics*, *139*. https://doi.org/10.1063/1.4818538

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, *106*(4), 620–630. https://doi.org/10.1103/PhysRev.106.620

Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution*. Oxford University Press.

Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, *5*(3), 183–191. https://doi.org/10.1147/rd.53.0183

Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D*, *42*, 12–37. https://doi.org/10.1016/0167-2789(90)90064-V

Li, M., & Vitányi, P. (1997). *An introduction to kolmogorov complexity and its applications*. Springer.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Wolfram, S. (2002). *A new kind of science*. Wolfram Media.