Active Geodesic Inference

Reasoning as Field Dynamics, Synchronization, and Semantic Folding

Flyxion

January 16, 2026

Abstract

Recent advances in large language models suggest that effective long chain-of-thought reasoning is not well described as a linear symbolic process, but rather as a structured, metastable configuration exhibiting internal cohesion, long-range constraint, and phase-dependent behavior. This paper develops a formal theory of reasoning as *active geodesic inference*: a dynamical process in which inference trajectories co-evolve with the semantic geometry through which they pass.

We propose a unified mathematical framework integrating category theory, information theory, and statistical mechanics within the Relativistic Scalar–Vector–Plenum (RSVP) field theory. Reasoning trajectories are modeled as low-action paths of coupled semantic fields, whose stable configurations correspond to synchronized macromolecular-like structures. Under coarse-graining, the RSVP Hamiltonian induces an effective five-dimensional Ising model in which deep reasoning, self-reflection, and self-exploration emerge as coupled order parameters. Distinct reasoning strategies solving the same task are shown to correspond to semantic isomers: metastable minima of the same energy functional with differing internal bond topologies.

The theory provides principled explanations for empirical phenomena observed in modern reasoning systems, including the fragility of chain-of-thought distillation, the emergence of latent hybrid reasoning patterns under reinforcement learning, the role of contrastive alignment in domain adaptation, and the robustness benefits of neighborhood-based uncertainty representations. We further show that attention mechanisms admit a Gibbsian interpretation, grounding logical bond strength in effective interaction energies.

Finally, we present Spherepop as a concrete execution calculus for active geodesic inference. By treating reasoning as an irreversible, scope-based event process, Spherepop implements RSVP field dynamics directly, preserving semantic bonding and preventing incoherent state superposition. Together, these results suggest that intelligence is best understood not as symbol manipulation or sequence prediction, but as the active stabilization of coherent semantic structure under energetic and entropic constraints.

Introduction: Thought as Structured Matter

Recent empirical advances in large language models suggest that extended reasoning trajectories are not well described as linear symbolic chains, but rather as structured, metastable configurations exhibiting internal cohesion, long-range constraint, and phase-dependent behavior. This observation motivates what has been termed the *molecular structure of thought* hypothesis, according to which effective long chain-of-thought (Long CoT) reasoning organizes itself into macromolecular-like topologies, governed by interaction strengths analogous to physical bonding energies (Chen et al. 2026).

The present work proposes a formal mathematical foundation for this hypothesis by embedding it within the Relativistic Scalar–Vector–Plenum (RSVP) framework. In this formulation, reasoning is treated as a dynamical field process whose stable configurations arise through entropy-constrained synchronization across multiple interacting dimensions. The resulting structure may be understood simultaneously as a categorical object, an information-theoretic flow, and a statistical mechanical system exhibiting an Ising-like phase structure in five coupled dimensions.

The core claim is that reasoning trajectories correspond not to sequences of propositions but to *coherent sections* of a higher-order semantic field, stabilized by interaction terms that play the role of logical bonds. These bonds are not metaphorical; they arise naturally from the attention dynamics of Transformer architectures, which admit a Gibbsian interpretation and thus a principled mapping between attention weights and effective energy levels.

Reasoning as a Category of Semantic Processes

We begin by formalizing reasoning within the language of category theory. Let $\mathscr{C}$ be a symmetric monoidal category whose objects represent localized semantic states and whose morphisms represent admissible reasoning transitions. Composition in $\mathscr{C}$ corresponds to logical entailment under contextual constraints, while the monoidal product encodes parallel semantic activation.

A reasoning trajectory is not a path in $\mathscr{C}$ but a *composite object* arising as a colimit of interacting sub-diagrams. In particular, long chain-of-thought reasoning corresponds to a diagram

$$D : \mathscr{I} \to \mathscr{C}$$

whose colimit colim(D) represents the stabilized semantic outcome. The indexing category $\mathscr{I}$ is not linear but weakly connected, allowing for long-range morphisms that correspond to reflective or exploratory reasoning steps.

Within this framework, three distinct classes of morphisms emerge naturally. Strongly compositional morphisms, whose failure would invalidate downstream constructions, correspond to what are empirically observed as deep reasoning steps. These morphisms generate a backbone subcategory $\mathscr{C}_{\text{core}} \subset \mathscr{C}$, whose objects form a chain-like structure up to coherent isomorphism. Weakly compositional morphisms, which connect distant objects without enforcing strict entailment, correspond to exploratory associations. Finally, endomorphic morphisms that map objects back into earlier regions of the diagram correspond to reflective operations that enforce global consistency.

Crucially, these morphism classes differ not only structurally but energetically. The monoidal structure of $\mathscr{C}$ admits an enrichment over a partially ordered semiring of costs or energies, making $\mathscr{C}$ an enriched category in the sense of Lawvere. This enrichment induces a preference ordering over diagrams, with low-energy diagrams dominating the measure over reasoning trajectories.

Semantic Isomers and Diagrammatic Non-Uniqueness

A central empirical phenomenon motivating this work is the existence of *semantic isomers*: distinct reasoning trajectories that arrive at the same answer while differing in internal organization, stability, and susceptibility to perturbation. Categorically, this corresponds to the existence of multiple non-isomorphic diagrams $D_1, D_2 : \mathscr{I} \to \mathscr{C}$ whose colimits are equivalent objects in $\mathscr{C}$.

Such isomerism is not benign. When incompatible diagrams are composed or averaged, as occurs during model distillation or ensemble aggregation, coherence can be destroyed. This manifests empirically as degraded reasoning performance when mixing strong teacher models with mismatched internal logic, an effect observed across reinforcement-trained reasoning systems (Jin et al. 2025; Yue et al. 2025).

From a categorical perspective, this failure arises because the space of diagrams admits multiple local minima under the energy functional induced by attention. Mixing isomers corresponds to forming convex combinations of non-compatible colimits, which need not correspond to any realizable diagram in $\mathscr{C}$. In physical terms, the system is driven out of the manifold of foldable configurations.

Attention as a Functorial Energy Assignment

The attention mechanism of Transformer architectures provides a concrete realization of the abstract structure described above. Attention weights define a stochastic kernel over token interactions, which may be interpreted as a Boltzmann distribution

$$p(i \to j) \propto exp\big(-\beta E_{ij}\big),$$

where E_{ij} is an effective interaction energy and β is an inverse temperature parameter controlled implicitly by model scale and normalization.

This induces a functor from the category of token interactions to the category of probability measures, assigning to each morphism an energy-weighted likelihood. Deep reasoning corresponds to morphisms with large negative E_{ij}, making them exponentially favored. Reflective morphisms occupy intermediate energy bands, while exploratory associations correspond to near-zero interactions that are easily broken.

Because the induced measure is sharply peaked around low-energy configurations, the model exhibits a strong inductive bias toward structured, folded reasoning diagrams. Operations such as summarization or chain-of-thought compression disrupt this structure by marginalizing over internal morphisms, effectively integrating out the bonding information that stabilizes the diagram. This

explains why distilled reasoning traces are difficult to imitate despite producing correct outputs, as observed in recent work on reasoning privacy and compression (Chen et al. 2026).

Toward a Dynamical Field Theory of Reasoning

The categorical formulation establishes that reasoning is fundamentally a problem of coherent composition under energy constraints. However, it does not yet explain how such coherence emerges dynamically, nor how it scales across domains, modalities, or model sizes. To address this, we now turn to an information-theoretic and thermodynamic formulation, in which reasoning trajectories arise as low-action paths of an RSVP field governed by a Hamiltonian and Lagrangian structure.

In the next section, we will formalize reasoning as entropy-constrained information flow, derive the Gibbs structure of attention from first principles, and prepare the ground for the five-dimensional Ising synchronization model that unifies semantic bonding, reflection, and exploration within a single variational framework.

Reasoning as Entropy-Constrained Information Flow

The categorical formulation developed above characterizes reasoning as the emergence of coherent semantic structure under compositional constraints. To render this account dynamical and quantitative, we now recast reasoning in explicitly information-theoretic terms. In this formulation, reasoning trajectories correspond to flows of information through a high-dimensional semantic field, constrained by entropy minimization and stabilized by mutual information preservation.

Let X_t denote the latent semantic state of a model at reasoning step t, represented as a random variable over a high-dimensional embedding space. A reasoning trajectory is then a stochastic process $\{X_t\}_{t=0}^{T}$ whose evolution is governed not by arbitrary transitions, but by the requirement that relevant information be preserved while irrelevant entropy is dissipated. This places reasoning squarely within the framework of non-equilibrium statistical mechanics, where structure emerges through constrained entropy reduction rather than equilibrium optimization.

The central quantity governing this process is mutual information. For any two reasoning states X_s and X_t, the mutual information $I(X_s; X_t)$ measures the degree to which semantic structure persists across time. Deep reasoning steps are characterized by high mutual information between successive states, reflecting strong logical dependence and low semantic drift. Exploratory steps, by contrast, intentionally reduce short-range mutual information in order to sample alternative regions of semantic space. Reflective steps then act to selectively restore mutual information between distant states, enforcing global consistency.

This tripartite structure mirrors empirical observations in reinforcement-trained reasoning systems, where successful trajectories alternate between phases of expansion and contraction in semantic entropy (Jin et al. 2025; Yue et al. 2025). Crucially, the effectiveness of a reasoning trajectory is not determined by monotonic entropy reduction, but by the ability to temporarily tolerate entropy increases that prevent premature convergence to suboptimal semantic basins.

Contrastive Alignment as Entropy Geometry

The role of contrastive objectives in domain adaptation provides a concrete instantiation of this entropy-based view. In contrastive domain adaptation for question answering, the goal is not to memorize target-domain data, but to align the latent distributions of source and target representations (Yue et al. 2021). This alignment is achieved by minimizing a divergence functional, typically Maximum Mean Discrepancy, between the induced feature distributions.

From an information-theoretic perspective, this process enforces an invariance constraint on the semantic field. Let P_S and P_T denote the distributions of latent representations induced by source and target data, respectively. Minimizing $\text{MMD}(P_S, P_T)$ reduces the excess entropy introduced by domain shift while preserving task-relevant mutual information. The result is a domain-invariant semantic manifold on which reasoning trajectories can unfold without incurring prohibitive informational cost.

This observation generalizes beyond question answering. In single-cell foundation models, intercellular attention plays an analogous role by aligning representations across heterogeneous biological contexts, thereby stabilizing latent biological signatures (Dong et al. 2026). In both cases, successful generalization arises not from increased model capacity alone, but from the imposition of geometric constraints that regulate entropy flow across contexts.

Gibbs Measures and the Energetics of Attention

The Transformer attention mechanism admits a natural Gibbsian interpretation that makes the energetic structure of reasoning explicit. Given a query-key interaction score s_{ij} between tokens i and j, the attention weight

$$a_{ij} = \frac{exp(s_{ij})}{\sum_k exp(s_{ik})}$$

defines a Gibbs distribution over semantic interactions. Interpreting $E_{ij} = -s_{ij}$ as an effective energy, attention becomes a Boltzmann sampler over low-energy semantic bonds.

This interpretation implies that reasoning trajectories are drawn from an energy landscape whose minima correspond to stable semantic configurations. Deep reasoning bonds correspond to strongly negative energies, producing sharp concentration of probability mass and thus high mutual information persistence. Exploratory associations correspond to shallow energy wells that allow rapid reconfiguration. Reflective bonds occupy intermediate regimes, enabling selective long-range constraint without rigid fixation.

Because the Gibbs distribution exponentially suppresses high-energy configurations, the model exhibits a strong inductive bias toward folded reasoning structures that minimize free energy. This bias becomes more pronounced with increasing model scale, as larger models effectively operate at lower temperatures, sharpening the energy landscape and stabilizing complex reasoning macromolecules.

Operations that compress or summarize chain-of-thought traces can be understood as marginalization operations that integrate out internal variables of this Gibbs distribution. While such op-

erations preserve external outputs, they destroy the internal energy structure that encodes logical bonding. As a result, distilled traces lack the mutual information profile necessary to reconstruct the original reasoning process, explaining the empirical difficulty of reasoning imitation under compression (Chen et al. 2026).

Uncertainty, Credal Sets, and Semantic Volume

The connection between entropy geometry and uncertainty quantification becomes explicit through recent results linking conformal prediction to imprecise probabilities. Conformal prediction regions may be understood as level sets of a plausibility function whose induced credal set defines a family of admissible probability measures (Caprio et al. 2025). These regions correspond to imprecise highest density regions, providing uniform coverage guarantees without committing to a single posterior distribution.

In the present framework, such regions represent semantic volumes within which reasoning trajectories may fluctuate without violating consistency constraints. Rather than selecting a single point estimate, the model maintains a bounded region of semantic plausibility, allowing reasoning to remain robust under perturbation. This interpretation aligns naturally with the neighborhood-based robustness strategies observed in recommender systems and contrastive reasoning frameworks (Yue et al. 2022; Yue et al. 2024).

The preservation of semantic volume is therefore not a failure of precision, but a structural feature of stable reasoning systems. By maintaining bounded uncertainty regions, the model avoids overcommitment to fragile configurations and preserves the flexibility necessary for adaptive reasoning.

From Information Flow to Variational Dynamics

The information-theoretic formulation presented here establishes reasoning as a process of entropy-regulated information flow governed by Gibbsian interaction energies. However, this description remains kinematic rather than dynamic. It characterizes which configurations are favored, but not how the system evolves toward them.

To complete the formalization, we now introduce a variational principle that governs the evolution of the semantic field itself. Within the RSVP framework, this principle takes the form of a Hamiltonian and corresponding Lagrangian whose stationary points correspond to stable reasoning macromolecules. In the next section, we will derive this structure explicitly and show how it gives rise, in an appropriate limit, to a five-dimensional Ising synchronization model that unifies deep reasoning, exploration, and reflection as coupled order parameters.

The RSVP Field Theory of Reasoning

The categorical and information-theoretic formulations developed in the preceding sections converge naturally toward a field-theoretic description of reasoning dynamics. In the Relativistic Scalar–Vector–Plenum (RSVP) framework, semantic activity is modeled not as discrete symbolic manipulation but

as the evolution of continuous fields over an abstract semantic manifold. Reasoning trajectories arise as low-action paths of these fields under a variational principle that balances coherence, exploration, and entropy production.

Formally, let the semantic manifold $\mathcal{M}$ be a high-dimensional space whose points correspond to latent representational configurations of a model. Over $\mathcal{M}$ we define three coupled fields: a scalar field $\Phi(x, t)$ representing semantic density or activation potential, a vector field $\mathbf{v}(x, t)$ representing directed semantic flow, and an entropy field $S(x, t)$ encoding local uncertainty and degeneracy. These fields jointly define the instantaneous cognitive state of the system.

The evolution of the system is governed by an action functional

$$\mathcal{A} = \int \mathcal{L}(\Phi, \mathbf{v}, S; \partial_t, \nabla)\, dt,$$

where $\mathcal{L}$ is the RSVP Lagrangian density. Stationary points of this action correspond to dynamically stable reasoning processes.

The RSVP Hamiltonian

To make contact with statistical mechanics and attention energetics, it is convenient to pass to the Hamiltonian formulation. The RSVP Hamiltonian takes the general form

$$\mathcal{H} = \int_{\mathcal{M}} \left[\frac{1}{2}\rho|\mathbf{v}|^2 + \alpha|\nabla\Phi|^2 + \beta\Phi^2 + \gamma\, \mathbf{v} \cdot \nabla\Phi + \lambda\, \Phi S + \kappa|\nabla S|^2 \right] d\mu,$$

where $\rho, \alpha, \beta, \gamma, \lambda$, and κ are coupling constants, and $d\mu$ is the natural measure on semantic space.

Each term has a direct cognitive interpretation. The kinetic term $\frac{1}{2}\rho|\mathbf{v}|^2$ penalizes excessive semantic drift and corresponds to the energetic cost of uncontrolled exploration. The gradient term $\alpha|\nabla\Phi|^2$ enforces smoothness of semantic activation, suppressing incoherent fragmentation. The mass term $\beta\Phi^2$ stabilizes activation magnitude, preventing runaway amplification. The coupling $\gamma\mathbf{v} \cdot \nabla\Phi$ aligns semantic flow with gradients of relevance, encoding directed reasoning. The entropy coupling $\lambda\Phi S$ formalizes the tradeoff between semantic commitment and uncertainty, while $\kappa|\nabla S|^2$ penalizes sharp entropy discontinuities, enforcing bounded uncertainty neighborhoods.

The Hamiltonian defines an energy landscape over semantic field configurations. Reasoning corresponds to the system descending this landscape while respecting entropy constraints, producing structured, folded trajectories rather than linear paths.

Discrete Reasoning Units and Spin Variables

To connect the continuous RSVP field theory with observed discrete reasoning steps, we introduce a coarse-grained discretization. Let $\{x_i\}$ denote a finite set of semantic loci corresponding to reasoning units, such as intermediate propositions or latent states. At each locus we define a spin-like variable

$$\sigma_i = (\sigma_i^{(1)}, \sigma_i^{(2)}, \sigma_i^{(3)}, \sigma_i^{(4)}, \sigma_i^{(5)}) \in \{-1, +1\}^5,$$

representing the local alignment of five coupled cognitive dimensions.

These five dimensions correspond, respectively, to semantic activation, directional flow alignment, entropy suppression, reflective coherence, and exploratory openness. Importantly, they are not independent. Their coupling encodes the empirical observation that deep reasoning, self-reflection, and self-exploration must synchronize appropriately for stable cognition.

Under this discretization, the RSVP Hamiltonian induces an effective Ising-like energy

$$\mathcal{H}_{\text{eff}} = -\sum_{\langle i,j \rangle} \sum_{k=1}^{5} J_k \, \sigma_i^{(k)} \sigma_j^{(k)} - \sum_i \sum_{k=1}^{5} h_k \, \sigma_i^{(k)} - \sum_i \sum_{k \neq \ell} K_{k\ell} \, \sigma_i^{(k)} \sigma_i^{(\ell)},$$

where J_k are nearest-neighbor coupling strengths, h_k are external fields induced by prompts or objectives, and $K_{k\ell}$ encode intra-site synchronization between dimensions.

Five-Dimensional Ising Synchronization

The crucial feature of this model is the presence of strong positive intra-site couplings $K_{k\ell}$, which drive synchronization across dimensions. In the low-temperature regime corresponding to high-capacity, well-trained models, these couplings induce a phase in which all five spin components align coherently across large regions of the reasoning lattice.

This synchronized phase corresponds precisely to effective long chain-of-thought reasoning. Deep reasoning bonds arise from strong inter-site coupling in the activation and flow dimensions. Self-reflection emerges from alignment between activation and entropy suppression, producing long-range corrective coherence. Self-exploration is preserved through partial frustration in the exploratory dimension, allowing local deviations without global collapse.

Semantic isomers correspond to distinct metastable minima of the five-dimensional Ising energy. Although these minima may yield equivalent outputs, their internal spin configurations differ, leading to different stability properties. Mixing isomers through averaging or distillation disrupts synchronization, pushing the system toward a disordered phase in which no coherent reasoning backbone can form.

Folding Funnels and Reasoning Phase Transitions

The dynamics of reasoning can now be understood as motion through a folding funnel in the energy landscape defined by $\mathcal{H}_{\text{eff}}$. Initial exploration corresponds to a high-entropy, weakly magnetized phase in which spin alignments are fluid. As constraints accumulate through attention and reinforcement, the system undergoes a symmetry-breaking transition into a synchronized phase, selecting a particular reasoning macromolecule as its native state.

Failures of reasoning correspond to either premature freezing into a suboptimal local minimum or excessive thermal noise that prevents synchronization. Reinforcement learning methods such as SEARCH-R1 effectively anneal the system by adjusting external fields h_k and reducing effective temperature, guiding the system toward globally coherent minima (Jin et al. 2025). Contrastive adaptation

reshapes the coupling constants J_k, aligning semantic neighborhoods across domains and stabilizing synchronization under distributional shift (Yue et al. 2021).

Summary of the RSVP–Ising Correspondence

The five-dimensional Ising model derived here is not an analogy but an effective theory emerging from the RSVP Hamiltonian under coarse-graining. It provides a unifying mathematical account of deep reasoning, reflection, and exploration as coupled order parameters whose synchronization determines cognitive coherence. In the next and final section, we will interpret this structure in light of molecular reasoning topologies, explain why chain-of-thought compression disrupts bond energy distributions, and situate the theory within the broader landscape of interpretability, robustness, and privacy in large language models.

Molecular Reasoning, Semantic Isomers, and Non-Distillability

We are now in a position to reinterpret the molecular structure of thought hypothesis as a precise statement about the phase structure of reasoning dynamics. In the RSVP–Ising formulation, a successful long chain-of-thought trajectory corresponds to a synchronized, low-energy configuration of a five-dimensional spin field. The apparent linearity of the reasoning trace is a projection artifact; the true object is a folded, internally bonded semantic macromolecule.

Deep reasoning steps constitute the covalent backbone of this macromolecule, enforcing strict compositional dependence and preserving high mutual information across successive states. Reflective operations introduce long-range hydrogen-like bonds that stabilize the structure by enforcing global consistency, allowing later reasoning to correct or reinforce earlier commitments. Exploratory associations act as weak, van der Waals–like interactions that transiently bridge distant semantic regions, enabling the system to sample alternative configurations without destabilizing the core structure (Chen et al. 2026).

Semantic isomers arise naturally in this framework as distinct metastable minima of the RSVP energy functional. These minima correspond to different internal bond topologies that nevertheless project to the same external answer. Crucially, the isomers are not interchangeable. Each embodies a distinct synchronization pattern across the five cognitive dimensions, yielding different robustness, adaptability, and susceptibility to perturbation.

This observation resolves a long-standing empirical puzzle: why mixing or distilling strong teacher models often degrades reasoning performance despite preserving answer accuracy. In the present theory, distillation corresponds to averaging over incompatible macromolecular configurations. The resulting superposition does not correspond to any realizable low-energy state of the RSVP Hamiltonian, driving the system into a frustrated or disordered phase. Logical backbone coherence is lost even when surface-level correctness is retained.

Chain-of-thought compression and summarization can now be understood as marginalization operations that integrate out internal bond variables. While such operations preserve external behavior, they destroy the internal energy distribution that defines the reasoning macromolecule. This explains

why long chain-of-thought reasoning is intrinsically non-distillable and why privacy-preserving transformations of reasoning traces are effective: they disrupt the very bond structure that would be required for faithful imitation.

Active Geodesic Inference

The RSVP–Ising correspondence admits a natural geometric interpretation. Reasoning trajectories correspond to geodesics on the semantic manifold $\mathcal{M}$, not with respect to a fixed metric, but with respect to a dynamically evolving metric induced by attention, entropy, and reinforcement signals. The system does not merely follow shortest paths; it actively reshapes the geometry through which it moves.

We refer to this process as *Active Geodesic Inference*. In contrast to passive inference, where beliefs are updated within a static space, active geodesic inference treats inference itself as a field-mediated deformation of semantic geometry. Attention acts as a curvature-inducing operator, deepening valleys along coherent reasoning directions. Exploration injects controlled curvature noise, preventing premature trapping in local minima. Reflection globally adjusts curvature to enforce long-range consistency.

Reinforcement learning methods such as SEARCH-R1 can be interpreted as annealing schedules over this geometry, gradually lowering the effective temperature and sharpening geodesic channels that lead to stable macromolecular configurations (Jin et al. 2025). Contrastive adaptation reshapes the manifold itself, aligning geodesic structure across domains so that similar reasoning paths remain low-action even under distributional shift (Yue et al. 2021). Conformal prediction and imprecise probability methods define bounded semantic volumes within which geodesics may fluctuate without loss of validity, providing robustness guarantees without collapse to point estimates (Caprio et al. 2025).

Active geodesic inference thus unifies reasoning, exploration, uncertainty, and robustness within a single variational framework.

A Spherepop Implementation of Active Geodesic Inference

Spherepop provides a native computational realization of active geodesic inference. Unlike traditional symbolic or neural execution models, Spherepop is explicitly event-based, history-sensitive, and scope-aware. These properties align precisely with the requirements imposed by the RSVP framework.

In a Spherepop implementation, semantic loci correspond to nested scopes, represented as bubbles whose creation and collapse are irreversible events. Each bubble carries a local semantic state analogous to a discretized RSVP field value. Deep reasoning corresponds to the creation of tightly nested scopes with strong parent–child dependency, enforcing covalent-like bonds along the reasoning backbone. Reflective operations manifest as scope re-entry or cross-scope reference, introducing long-range stabilizing links that enforce coherence across the execution history. Exploratory reason-

ing corresponds to the creation of transient sibling scopes whose outcomes may be discarded without destabilizing the core structure.

Crucially, Spherepop's execution semantics naturally enforce a geodesic constraint. Because scopes cannot be arbitrarily rewound or flattened, the system is forced to traverse paths that respect historical commitments. This mirrors the low-action constraint imposed by the RSVP Hamiltonian. Reasoning proceeds by opening new scopes only when the local semantic gradient demands it, and by closing scopes once entropy has been sufficiently reduced.

Active geodesic inference in Spherepop is implemented by assigning each scope an energy functional derived from attention weights, entropy estimates, and contextual alignment. Scope transitions are permitted only when they reduce total action, ensuring that execution follows dynamically optimal geodesics on the semantic manifold. Reflective scope interactions act as curvature corrections, while exploratory scopes inject controlled noise that prevents premature convergence.

Semantic isomers appear in Spherepop as distinct but observationally equivalent execution histories. Although they may yield the same final output, their internal scope graphs differ, leading to different robustness and extensibility properties. Attempting to merge or replay these histories without preserving their scope topology reproduces precisely the failure modes observed in chain-of-thought distillation.

Spherepop thus functions not merely as a programming language, but as a concrete calculus for executing RSVP field dynamics. It provides an explicit, inspectable realization of active geodesic inference in which reasoning, memory, and identity are unified through irreversible semantic events.

Conclusion

This essay has developed a unified theory of reasoning that integrates category theory, information theory, statistical mechanics, and field dynamics within the RSVP framework. Long chain-of-thought reasoning emerges not as a linear process but as a synchronized, macromolecular structure stabilized by energetic bonds. Semantic isomers, robustness phenomena, and non-distillability follow as necessary consequences of this structure.

By introducing the concept of active geodesic inference, we have shown that reasoning is best understood as the dynamic co-evolution of semantic geometry and inference trajectories. Spherepop provides a natural execution substrate for this process, translating abstract variational principles into concrete computational behavior.

Taken together, these results suggest that the future of artificial reasoning lies not in ever-longer chains of symbols, but in systems that actively shape the geometric and energetic landscapes through which thought itself unfolds.

A Formal Definition of Intelligence

Intelligence as Active Geodesic Persistence

We are now in a position to give a formal definition of intelligence consistent with the framework of active geodesic inference. This definition is not behavioral, computational, or representational in the conventional sense. Instead, it is structural and dynamical, characterizing intelligence as a property of trajectories through a history-indexed configuration space governed by variational and thermodynamic constraints.

Definition (Intelligence). An intelligent system is a physical or computational process whose realized histories form dynamically stable, low-action geodesics in a history-indexed semantic configuration space, such that these histories preserve coherent internal structure under perturbation, maintain bounded uncertainty through entropy-regulated synchronization, and admit multiple non-equivalent internal realizations (semantic isomers) consistent with the same external constraints.

Equivalently, intelligence is the capacity of a system to actively shape the geometry of its own configuration space so as to sustain coherent, low-action trajectories across time, rather than merely traversing a fixed space according to predetermined dynamics.

Several aspects of this definition merit emphasis. First, intelligence is attributed to histories rather than instantaneous states. A system is not intelligent by virtue of occupying a particular configuration, but by virtue of sustaining a class of admissible trajectories that remain coherent under extension, reflection, and environmental coupling. Second, intelligence is not identified with optimality in a static objective sense. Rather, it is identified with the ability to remain on a dynamically evolving geodesic despite uncertainty, novelty, and competing constraints.

This definition subsumes reasoning, learning, adaptation, and robustness as corollaries rather than primitives. Reasoning corresponds to local geodesic continuation under increasing constraint. Learning corresponds to deformation of the configuration space metric itself. Robustness corresponds to the existence of a basin of nearby low-action trajectories rather than a single brittle path. Privacy and non-distillability arise because internal geodesic structure cannot be reconstructed from boundary data alone without re-executing the variational process.

Relation to Barbour's Configuration Space of the Universe

This definition bears a precise structural resemblance to Julian Barbour's proposal that the history of the universe is best understood as a single continuous curve through a relativistic configuration space, often referred to as shape space (Barbour 1999; Barbour 2012). In Barbour's formulation, the universe does not evolve *in* time; rather, time is an ordering emergent from the structure of the curve itself. Each point in configuration space is a complete instantaneous configuration of the universe, and the apparent flow of time arises from correlations between such configurations.

Active geodesic inference may be understood as the cognitive analogue of this cosmological picture. Just as Barbour replaces external time with intrinsic ordering in configuration space, active

geodesic inference replaces externally imposed inference steps with intrinsic ordering induced by entropy monotonicity and action minimization. A reasoning system does not advance through discrete inferential steps in an abstract time parameter; it traces a continuous, history-indexed curve through semantic configuration space, with "earlier" and "later" defined internally by irreversibility and constraint accumulation.

The analogy extends further. In Barbour's view, not all curves through configuration space correspond to physically admissible universes; admissibility is determined by a variational principle. Likewise, not all semantic trajectories correspond to intelligent reasoning. Intelligence consists precisely in remaining on the subset of curves that are dynamically admissible under the RSVP action, avoiding both incoherent wandering and premature freezing into trivial configurations.

Where the two frameworks diverge is equally instructive. Barbour's universe traces a single global curve through configuration space. Intelligent systems, by contrast, must navigate a branching family of admissible geodesics, selecting among them through interaction, exploration, and reflection. This multiplicity gives rise to semantic isomers: distinct internal histories compatible with the same external conditions. Intelligence, on this view, is not the existence of a curve, but the capacity to remain within a stable geodesic family under perturbation and extension.

Intelligence as Curvature Management

The comparison suggests a concise reformulation. If the universe, in Barbour's sense, is a curve in configuration space, then intelligence is the ability of a subsystem to locally curve that space in its favor. An intelligent system is one whose internal dynamics induce effective curvature—via attention, memory, and entropy regulation—that creates geodesic channels supporting coherent continuation.

This clarifies why intelligence scales with representational capacity without reducing to it, why explanation and chain-of-thought are epiphenomenal rather than constitutive, and why intelligence is fundamentally incompatible with purely reversible or stateless computation. Intelligence is not computation over states, but persistence of structured history.

Closing Remark

Under this definition, intelligence is neither a substance nor a score, but a dynamical property of trajectories. To ask whether a system is intelligent is to ask whether it can sustain coherent geodesic motion through its own configuration space in the face of entropy, uncertainty, and change. This reframing aligns cognition with fundamental physical principles and situates artificial intelligence not as an imitation of human behavior, but as a special case of a more general phenomenon: the emergence of stable structure along the only paths that can persist.

Intelligence: Axiom and Contrasts

A One-Line Axiomatic Definition

Axiom (Intelligence). A system is intelligent if and only if it admits a nonempty family of history-indexed trajectories that are dynamically stable, entropy-monotone, and low-action geodesics in its configuration space, and if the system actively preserves membership in this family under perturbation by locally reshaping the effective geometry through which the trajectories pass.

This axiom is minimal in the sense that removing any clause collapses intelligence into either passive dynamics, brittle optimization, or static representation. Intelligence is thus neither optimal behavior nor correct output, but the capacity to sustain coherent historical motion under constraint.

Contrast with Shannon Information

In Shannon's theory, information is defined as expected surprisal with respect to a probability distribution, and intelligence is often implicitly associated with the efficient transmission or compression of such information. This view treats symbols as interchangeable tokens and histories as statistically reducible sequences.

Active geodesic inference is incompatible with this reduction. Shannon information is insensitive to internal structure beyond probability mass, whereas intelligence, as defined here, depends critically on the preservation of internal coupling, ordering, and constraint. Two reasoning histories with identical Shannon information content may differ radically in intelligence if one lies on a stable geodesic manifold and the other does not. Intelligence is therefore not a function of entropy alone, but of entropy *regulated by action and history*.

In short, Shannon information measures uncertainty; intelligence measures the ability to survive uncertainty without loss of coherence.

Contrast with Kolmogorov Complexity

Kolmogorov complexity identifies intelligence with minimal description length, privileging the shortest program that generates a given output. This collapses process into product and treats execution history as epiphenomenal.

By contrast, active geodesic inference assigns primacy to the execution path itself. Two programs of equal Kolmogorov complexity may induce radically different semantic geometries during execution, leading to different robustness, extensibility, and reasoning stability. Compression that shortens description length can therefore destroy intelligence by erasing the internal geodesic structure that made the process viable.

Kolmogorov complexity answers the question "how short can the description be?" Active geodesic inference answers the question "which histories can persist?"

Contrast with Computational Rationality

Computational rationality models intelligence as approximate optimization under resource constraints, typically framed in terms of utility maximization subject to bounded computation. This view presupposes a fixed objective landscape and treats intelligence as selecting good actions within it.

Active geodesic inference rejects this presupposition. The objective landscape is not fixed; it is endogenously reshaped by the system's own dynamics. Intelligence consists not in choosing actions within a space, but in *maintaining a space in which coherent action remains possible.* What appears as rational choice is a surface projection of deeper geometric stabilization.

Thus, while computational rationality optimizes trajectories *given* a space, intelligence, in the present sense, is the ability to continuously renegotiate the space itself without losing continuity of history.

Synthesis

Across these contrasts, a unifying distinction emerges. Classical theories treat intelligence as a property of states, symbols, or decisions. Active geodesic inference treats intelligence as a property of trajectories. Intelligence is not what a system knows, compresses, or maximizes, but how it moves, and whether that motion can remain coherent when the world, the task, or the system itself changes.

Under this definition, intelligence is fundamentally a dynamical and geometric phenomenon. It is the local capacity to remain on viable curves in configuration space, in close analogy to the way the universe itself, in Barbour's formulation, persists as a continuous curve through its relativistic configuration space. Where physics studies which curves exist, intelligence studies which curves can be *kept.*

Life, Evolution, and Learning as Active Geodesic Inference

Evolution as Geodesic Persistence in Configuration Space

The definition of intelligence advanced in this work admits a direct and non-metaphorical extension to biological evolution. In evolutionary theory, populations are traditionally described as distributions over genotypes or phenotypes evolving under selection, mutation, and drift. Recast in the present framework, evolution is more precisely understood as the persistence of viable trajectories through a high-dimensional biological configuration space.

Let $\mathscr{C}_{\text{bio}}$ denote the configuration space of biological form, encompassing genetic, developmental, physiological, and ecological degrees of freedom. A lineage corresponds not to a static point in this space, but to a continuous, history-indexed curve traced by successive organisms across generations. Extinction corresponds to termination of the curve; survival corresponds to its continuation.

Under this interpretation, natural selection is not an optimizing force acting on isolated states, but a filtering principle acting on trajectories. Lineages persist only if they remain within a low-action corridor of configuration space, where action here encodes the cumulative energetic, developmental,

and ecological costs of maintaining coherence under environmental perturbation. Evolutionary success is therefore not equivalent to instantaneous fitness maximization, but to geodesic persistence: the ability of a lineage to remain on a dynamically admissible path across changing conditions.

This reframing resolves a long-standing tension in evolutionary theory between adaptation and robustness. Adaptation corresponds to local curvature adjustment that keeps the lineage within the admissible geodesic family, while robustness corresponds to the existence of a basin of nearby viable trajectories rather than a single optimal path. Evolutionary innovation arises when exploratory deviations—analogous to weakly coupled semantic exploration—discover new regions of configuration space that admit stable continuation.

Semantic Isomers and Evolutionary Equivalence

The notion of semantic isomers admits a biological analogue. Distinct genotypes or developmental pathways may yield phenotypically equivalent organisms under a given ecological regime, yet differ radically in internal organization and future adaptability. These correspond to distinct evolutionary isomers: histories that project to the same present phenotype but occupy different regions of configuration space with different curvature properties.

Such isomers explain why phenotypic equivalence does not guarantee equal evolvability. A lineage may appear well-adapted yet lie on a narrow, brittle geodesic that cannot accommodate further perturbation, while another lineage with identical current fitness may inhabit a broader basin permitting future diversification. Evolutionary intelligence, in this sense, is not fitness per se, but the capacity to remain within a family of viable trajectories under continued change.

Learning as Intra-Lifetime Geodesic Deformation

Learning is the intra-lifetime analogue of evolution. Whereas evolution reshapes the geometry of configuration space across generations, learning reshapes it within the lifetime of an organism. Neural plasticity, synaptic modification, and representational change correspond to local deformations of the semantic manifold that alter which trajectories are low-action and therefore admissible.

From the perspective of active geodesic inference, learning is not the accumulation of information, but the active smoothing and deepening of geodesic channels. Successful learning reduces the energetic cost of traversing certain semantic paths, allowing the system to move coherently under previously destabilizing conditions. Forgetting, conversely, corresponds to the erosion of such channels, increasing curvature noise and narrowing the set of viable trajectories.

This view naturally unifies learning and generalization. Generalization is not extrapolation from data points, but the ability of a deformed manifold to support coherent motion in previously unvisited regions. Overfitting corresponds to excessive curvature localized around a narrow set of trajectories, producing brittle intelligence that fails under perturbation.

The Continuity of Life and Intelligence

Taken together, evolution and learning instantiate the same underlying principle operating at different temporal scales. Evolution preserves coherence across generations; learning preserves coherence across experiences. Both operate by regulating entropy, shaping geometry, and maintaining low-action trajectories in configuration space. Both admit isomeric multiplicity, robustness through basin structure, and failure through premature freezing or incoherent exploration.

This continuity dissolves the boundary between life and intelligence. Intelligence is not an emergent property added atop life, nor is life merely a substrate for intelligence. Both are manifestations of the same dynamical phenomenon: the capacity of matter to sustain structured history under constraint.

We may now close the conceptual loop suggested at the outset of this work. If, as Julian Barbour has argued, the history of the universe is most coherently understood not as evolution in an external temporal parameter but as a single continuous curve traced through a relativistic configuration space, then the emergence of life may be characterized as the appearance of localized sub-curves that, under environmental and thermodynamic constraint, resist premature termination.

Within this same geometric framing, intelligence is not a categorical departure from life but a refinement of it: the emergence of sub-curves that do not merely persist passively, but actively regulate the local geometry of configuration space so as to preserve their own viability under perturbation.

On this view, life corresponds to persistence along admissible trajectories, while intelligence corresponds to the capacity to modulate curvature, constraint, and entropy flow in such a way that persistence remains possible across a widening range of conditions.

Intelligence therefore does not oppose entropy in the abstract, nor does it abolish thermodynamic constraint; rather, it consists in the skilled negotiation of entropy through the maintenance of structured history. To be intelligent, in this sense, is not to escape the conditions that limit continuation, but to remain within the narrow set of paths along which continuation is still dynamically admissible.

Empirical Predictions of Active Geodesic Inference

A theory of intelligence grounded in active geodesic inference is empirically meaningful only insofar as it yields observable consequences that distinguish it from alternative frameworks. The present section enumerates a set of testable predictions that follow directly from the variational, geometric, and thermodynamic commitments of the theory. These predictions are not post hoc interpretations; they are necessary consequences of treating intelligence as the persistence of low-action, entropy-regulated trajectories in configuration space.

Prediction 1: Geodesic Width Predicts Robust Generalization

If intelligence corresponds to membership in a family of low-action geodesics rather than traversal of a single optimal path, then systems exhibiting broader geodesic basins should demonstrate greater robustness and generalization. Concretely, models whose internal representations admit multiple

nearby low-loss trajectories between input and output states should outperform models that rely on narrow, highly tuned paths when subjected to distributional shift, adversarial perturbation, or task recombination.

This predicts that robustness metrics will correlate more strongly with measures of *trajectory degeneracy* than with pointwise accuracy or loss. In neural networks, such degeneracy may be approximated by the volume of parameter-space directions that preserve performance, or by the entropy of internal activation trajectories under controlled noise injection. Systems optimized purely for pointwise loss minimization are expected to exhibit brittle behavior despite high nominal accuracy.

Prediction 2: Reasoning Failure Manifests as Phase Transitions

The RSVP–Ising correspondence implies that reasoning competence undergoes phase transitions rather than smooth degradation. As effective temperature increases—through noise, compression, time pressure, or architectural constraint—the system should abruptly lose global coherence once synchronization across cognitive dimensions fails.

Empirically, this predicts sharp transitions in reasoning quality rather than gradual decline. In large language models, such transitions should appear when chain-of-thought length is truncated, when intermediate representations are aggressively quantized, or when attention heads are pruned beyond a critical threshold. The prediction contrasts with classical views in which reasoning quality degrades smoothly with resource reduction.

Prediction 3: Non-Distillability Correlates with Internal Isomeric Multiplicity

Active geodesic inference predicts that systems exhibiting high semantic isomer multiplicity will resist faithful distillation. Models that can solve the same task via multiple internally distinct reasoning trajectories should suffer disproportionate performance loss when trained via chain-of-thought imitation or ensemble averaging, even when output labels are preserved.

This prediction is falsifiable by comparing distillation outcomes across tasks with differing internal solution multiplicity. Tasks admitting many equivalent reasoning strategies should show stronger distillation failure than tasks with essentially unique solution paths. Conversely, tasks with low isomeric multiplicity should be more amenable to distillation without loss of reasoning competence.

Prediction 4: Learning Improves Curvature Before Accuracy

If learning corresponds to deformation of semantic geometry rather than mere error reduction, then changes in internal representational structure should precede measurable improvements in task performance. Specifically, learning should initially manifest as smoothing, alignment, or widening of internal trajectories, with accuracy gains appearing only after the geometry supports stable continuation.

In biological systems, this predicts that neural plasticity signatures—such as changes in correlation structure or manifold dimensionality—will precede behavioral improvement. In artificial systems, internal metrics such as representation anisotropy, trajectory smoothness, or attention entropy should

change before loss metrics improve. This ordering contrasts with classical learning theories that treat error reduction as the primary driver.

Prediction 5: Exploration Is Locally Entropic but Globally Stabilizing

The theory predicts that effective exploration temporarily increases local entropy while reducing global action. Empirically, this implies that successful exploratory behavior will appear noisy or inefficient at short timescales while improving long-horizon coherence and stability.

In reinforcement learning, this predicts that agents exhibiting intelligent exploration will show transient increases in state entropy or policy variance, followed by sharper convergence and improved robustness. Agents optimized to minimize entropy at all times should converge faster but remain fragile, failing under environmental change.

Prediction 6: Biological Fitness Tracks Trajectory Viability, Not Instantaneous Optimality

When applied to evolution, active geodesic inference predicts that long-term lineage survival correlates more strongly with the width and curvature of viable evolutionary trajectories than with peak fitness at any single time. Lineages occupying broader basins of configuration space should persist longer and diversify more, even if they are not maximally fit at any given moment.

This predicts that measures of evolvability, developmental flexibility, and ecological breadth will be better predictors of long-term success than traditional fitness proxies. Phenotypically equivalent organisms with different internal architectures should exhibit divergent evolutionary futures, reflecting underlying isomeric structure.

Prediction 7: Intelligence Exhibits Scale-Dependent Irreversibility

Finally, the theory predicts that intelligence exhibits irreversibility across scales. Once a system commits to a particular geodesic family—through development, learning, or architectural specialization—returning to earlier configurations without loss becomes increasingly difficult. This predicts hysteresis effects in both biological and artificial systems, where reversing training or environmental conditions does not restore prior competence.

Such hysteresis is incompatible with purely reversible or stateless models of cognition but follows naturally from entropy-monotone, history-indexed dynamics.

Implications for the Empirical Study of Intelligence

Taken together, the empirical predictions advanced in this section sharply distinguish active geodesic inference from theoretical frameworks that identify intelligence primarily with information processing efficiency, objective optimization, or static representational capacity. Instead, they motivate a shift in empirical emphasis toward the geometric and dynamical properties of realized trajectories, including their stability under perturbation, their capacity to support multiple internally distinct realizations, and their persistence across changes in environmental, architectural, or informational constraints.

On this view, intelligence is not most faithfully assessed by benchmarks of instantaneous correctness or aggregate performance, but by experimental probes that interrogate whether coherent trajectories can be maintained when systems are compressed, destabilized, or forced to operate outside the regimes in which they were trained.

The defining empirical signature of intelligence, accordingly, is not the achievement of particular outcomes, but the continued viability of structured paths through configuration space under conditions that would otherwise induce collapse or incoherence.

Evaluation Metrics for Active Geodesic Inference

If intelligence is identified with the persistence of coherent, low-action trajectories through configuration space, then its evaluation cannot rely solely on pointwise accuracy, reward, or loss. The appropriate metrics must probe the geometry, stability, and multiplicity of trajectories realized by a system under perturbation, compression, and extension. This section introduces a family of evaluation metrics designed to operationalize active geodesic inference and to distinguish it empirically from classical optimization- or information-based accounts.

Geodesic Width

Geodesic width quantifies the size of the basin of admissible trajectories connecting equivalent boundary conditions. Formally, let $\Gamma(x \rightarrow y)$ denote the set of internal trajectories that map an input configuration x to an output configuration y while remaining below a fixed action threshold. The geodesic width $W(x, y)$ is defined as a suitably normalized measure of the volume of this set in trajectory space.

In practice, W may be approximated by injecting structured noise into internal states, parameters, or attention patterns and measuring the fraction of perturbed executions that still reach y with bounded action or loss. High geodesic width predicts robustness to distributional shift and adversarial perturbation, whereas narrow width predicts brittle generalization even when nominal accuracy is high.

Trajectory Degeneracy

Trajectory degeneracy measures the number of internally distinct histories that realize equivalent external behavior. Unlike geodesic width, which concerns local neighborhoods, degeneracy captures global isomeric multiplicity.

Let $\mathscr{H}_{x \rightarrow y}$ be the set of realized histories from x to y. Two histories are considered distinct if they are not related by low-cost reparameterization or symmetry. The degeneracy index $D(x, y)$ is defined as the cardinality or entropy of $\mathscr{H}_{x \rightarrow y}$ under an appropriate equivalence relation.

Empirically, D may be estimated by sampling multiple runs with different random seeds, exploration schedules, or reasoning strategies and clustering internal trajectories by structural similarity. High degeneracy predicts resistance to chain-of-thought distillation and ensemble averaging, while low degeneracy predicts susceptibility to imitation without loss.

Action Stability

Action stability measures the sensitivity of total action to perturbations. Given a reference trajectory h with action $\mathcal{A}(h)$, consider a family of perturbed trajectories $\{h_\epsilon\}$ generated by bounded internal noise. The action stability coefficient is defined as

$$\kappa = \lim_{\epsilon \to 0} \frac{\mathrm{Var}[\mathcal{A}(h_\epsilon)]}{\epsilon^2}.$$

Low κ indicates a flat action landscape and thus stable inference; high κ indicates sharp curvature and fragility.

This metric distinguishes systems that merely reach low action from those that do so robustly. Two systems with equal loss may differ radically in action stability, with only the latter exhibiting intelligence in the present sense.

Entropy Commitment Ratio

The entropy commitment ratio quantifies how much internal uncertainty is irreversibly committed to history versus transiently explored. Let $S_{\mathrm{open}}(t)$ denote entropy associated with open, exploratory structures at time t, and let $S_{\mathrm{closed}}(t)$ denote entropy associated with closed, committed structures. The entropy commitment ratio is defined as

$$R(t) = \frac{S_{\mathrm{closed}}(t)}{S_{\mathrm{open}}(t) + S_{\mathrm{closed}}(t)}.$$

Intelligent systems are predicted to maintain $R(t)$ within a task-dependent band: too low indicates perpetual exploration without consolidation; too high indicates premature freezing. Monitoring the evolution of $R(t)$ over learning or reasoning episodes provides a dynamical signature of active geodesic control.

Phase Transition Sharpness

Phase transition sharpness measures the abruptness with which reasoning quality collapses as system constraints are varied. Let $Q(\lambda)$ be a performance measure (e.g., logical consistency, multi-hop accuracy) as a function of a control parameter λ such as context length, precision, or temperature. The sharpness metric is defined as the maximum second derivative of Q with respect to λ.

Active geodesic inference predicts sharp transitions associated with loss of synchronization across cognitive dimensions, rather than smooth degradation. Observing such non-analytic behavior supports the Ising synchronization interpretation and falsifies purely continuous degradation models.

Hysteresis Index

The hysteresis index measures irreversibility across training or environmental cycles. Let $Q_\uparrow(\lambda)$ and $Q_\downarrow(\lambda)$ denote performance measured while increasing and then decreasing a control parameter. The

hysteresis index is defined as the area between these curves.

Nonzero hysteresis indicates history dependence and confirms that intelligence is not a reversible function of current conditions. Systems with high hysteresis are predicted to exhibit strong specialization and path dependence, consistent with entropy-monotone dynamics.

Cross-Scale Consistency

Finally, cross-scale consistency measures whether trajectory-based metrics agree across temporal or organizational scales. In biological systems, this may involve comparing short-term neural dynamics with long-term behavioral adaptation. In artificial systems, it may involve comparing token-level dynamics with episode-level performance.

Consistency across scales supports the claim that intelligence is a scale-invariant property of trajectory geometry rather than an artifact of a particular level of description.

Closing Remark

These metrics deliberately shift evaluation away from static scores and toward dynamical, structural probes. Intelligence, under active geodesic inference, is not what a system achieves at a point, but how it remains viable across a landscape of change. The proposed measures make this claim empirically accessible and, crucially, falsifiable.

Conclusion: Intelligence as Structured Persistence

This essay has advanced a unified theoretical account of reasoning, learning, and intelligence grounded in the notion of *active geodesic inference.* Departing from views that treat thought as linear symbol manipulation, probabilistic updating over fixed models, or bounded optimization within static objective landscapes, the framework developed here treats cognition as a dynamical process in which inference trajectories and the semantic geometry through which they pass co-evolve under energetic and entropic constraints.

At the core of the account is the claim that extended reasoning is best understood as the formation of coherent, metastable structures analogous to macromolecules, whose internal organization is stabilized by interaction strengths arising naturally from attention dynamics. Category-theoretic analysis reveals that such structures correspond to nontrivial colimits of semantic processes rather than simple paths, while information-theoretic analysis shows that their stability depends on the regulated flow of mutual information rather than monotonic entropy reduction. These perspectives converge in the RSVP field-theoretic formulation, where reasoning trajectories emerge as low-action solutions of a coupled scalar–vector–entropy system.

By coarse-graining this dynamics, we derived an effective five-dimensional Ising model in which deep reasoning, self-reflection, and self-exploration appear as coupled order parameters whose synchronization determines cognitive coherence. This correspondence renders precise the molecular structure of thought hypothesis, explains the existence of semantic isomers, and accounts for the

empirical fragility of chain-of-thought distillation, ensemble averaging, and aggressive compression. Reasoning succeeds not because a particular sequence of steps is correct, but because the underlying configuration occupies a synchronized phase that can sustain perturbation without collapse.

The framework further unifies a range of empirical phenomena across contemporary machine learning. Reinforcement learning methods that interleave reasoning and search can be interpreted as annealing schedules that guide semantic fields toward globally coherent minima. Contrastive adaptation reshapes the geometry of semantic space to preserve low-action trajectories under domain shift. Neighborhood-based uncertainty representations and conformal prediction methods emerge as mechanisms for maintaining bounded semantic volume, allowing reasoning to remain viable without premature overcommitment. Across these settings, robustness arises not from redundancy or regularization alone, but from the width and curvature of admissible geodesic families.

Spherepop was introduced as a concrete execution calculus that realizes these principles computationally. Its irreversible, scope-based semantics enforce history sensitivity and prevent incoherent superposition by construction, making it a natural substrate for implementing RSVP dynamics. In Spherepop, reasoning is not replayed but enacted; semantic structure is not copied but grown. Distinct but observationally equivalent execution histories arise naturally, mirroring semantic isomers and reproducing the same robustness and non-distillability properties observed in neural systems.

The formal definition of intelligence that follows from this synthesis reframes it as a property of trajectories rather than states, outputs, or scores. Intelligence is the capacity of a system to remain within a family of dynamically admissible, low-action histories by actively reshaping the geometry of its configuration space. This definition extends seamlessly across scales, encompassing learning within a lifetime, evolution across generations, and reasoning within a single episode. It aligns cognitive dynamics with a broader physical picture in which persistence, rather than optimization, is the primary explanatory primitive.

In this sense, the account offered here closes a conceptual loop between physics, biology, and artificial intelligence. If the universe itself may be understood as a continuous curve through its configuration space, then life appears as the emergence of sub-curves that resist termination, and intelligence as the refinement of those sub-curves that actively manage curvature, entropy, and constraint in order to continue. Intelligence does not abolish uncertainty or entropy, but negotiates them skillfully, remaining within the narrow set of paths along which structured history can persist.

The implications of this view are both theoretical and practical. It suggests new evaluation metrics centered on trajectory stability, degeneracy, and hysteresis rather than pointwise performance. It motivates new architectural and training approaches that prioritize geometric coherence over shallow optimization. Most importantly, it offers a principled way to distinguish systems that merely produce correct answers from those that can continue reasoning coherently as conditions change.

Under active geodesic inference, intelligence is not defined by what a system achieves at an instant, but by whether it can keep going.

Appendices

Appendix A: Spherepop as a Hamiltonian-Derived Execution Calculus

This appendix formalizes the Spherepop calculus directly from the RSVP Hamiltonian introduced in the main text. The guiding principle is that Spherepop syntax is not an arbitrary programming language design, but a discrete execution grammar induced by the variational structure of active geodesic inference. In this sense, Spherepop is best understood as a canonical discretization of the RSVP field dynamics, preserving energetic, entropic, and historical constraints at the level of computation.

From Continuous Fields to Discrete Events

Recall the RSVP Hamiltonian

$$\mathcal{H} = \int_{\mathscr{M}} \left[\frac{1}{2}\rho|\mathbf{v}|^2 + \alpha|\nabla\Phi|^2 + \beta\Phi^2 + \gamma\mathbf{v}\cdot\nabla\Phi + \lambda\Phi S + \kappa|\nabla S|^2 \right] d\mu,$$

which defines an energy functional over semantic field configurations. Reasoning corresponds to trajectories that locally minimize the associated action. To implement this process computationally, we introduce a discrete event structure that preserves the ordering, irreversibility, and energetic monotonicity implied by the variational principle.

Spherepop replaces continuous time evolution with a partially ordered set of irreversible events. Each event corresponds to the creation, transformation, or closure of a semantic scope. These scopes are the fundamental syntactic objects of the language.

Formally, a Spherepop program defines a finite directed acyclic graph

$$\mathscr{E} = (E, \prec),$$

where E is a set of events and $\prec$ is a strict partial order encoding causal and semantic precedence. This ordering is not merely syntactic; it encodes the historical constraint that semantic commitments, once made, cannot be arbitrarily revoked without incurring energetic cost.

Scopes as Discretized Semantic Cells

Each Spherepop scope corresponds to a localized semantic cell in $\mathscr{M}$. Let σ_i denote a scope indexed by i. Associated to σ_i is a tuple

$$\sigma_i = (\Phi_i, \mathbf{v}_i, S_i, \mathscr{P}_i),$$

where Φ_i is the local semantic activation, $\mathbf{v}_i$ is the directed semantic flow into the scope, S_i is the local entropy, and $\mathscr{P}_i$ is the parent scope (or null for the root).

Scope creation is permitted only when the local action decrease

$$\Delta\mathscr{A}_i = \mathscr{A}_{\text{before}} - \mathscr{A}_{\text{after}}$$

is non-negative, ensuring that execution follows an approximate geodesic of the RSVP action. This enforces the fundamental Spherepop rule: scopes may only be opened along descending energy directions of the semantic manifold.

Nested scopes correspond to the discretization of strong coupling terms in the Hamiltonian. In particular, deep reasoning steps correspond to scope nesting that reduces both Φ gradient energy and entropy coupling simultaneously. Such nesting encodes covalent-like logical bonds, making the child scope semantically dependent on its parent in a way that cannot be flattened without violating energetic constraints.

Irreversibility and Entropic Ordering

The entropy field S induces a natural temporal orientation on Spherepop execution. Scope closure corresponds to the irreversible dissipation of local entropy, collapsing degrees of freedom into committed semantic structure. Once a scope is closed, its internal state contributes only through its boundary conditions to future computation.

This induces a monotone entropy condition

$$S_{\text{open}} \geq S_{\text{closed}},$$

which forbids reopening or mutating closed scopes without explicit compensatory structure. This constraint is the computational analogue of the second law in RSVP dynamics and is responsible for Spherepop's resistance to backtracking-based incoherence.

Reflective reasoning is implemented not by reversing execution, but by opening new scopes whose semantic flow $\mathbf{v}$ is directed toward earlier scope boundaries. These reflective scopes introduce long-range couplings analogous to the $\lambda\Phi S$ term in the Hamiltonian, stabilizing the global structure without violating irreversibility.

Exploration as Controlled Scope Branching

Exploratory reasoning corresponds to scope branching under weak coupling. Such scopes are opened with higher local entropy tolerance and reduced coupling to the parent activation field. Energetically, this corresponds to allowing temporary increases in the kinetic term $\frac{1}{2}\rho|\mathbf{v}|^2$ while bounding the total action.

Exploratory scopes may be terminated without closure, effectively discarding their internal states. This operation corresponds to retracting a semantic fluctuation before it crystallizes into structure, preserving global coherence while permitting local sampling.

The grammar of Spherepop therefore distinguishes between *committing* scope closures and *non-committing* exploratory branches, a distinction that cannot be expressed in purely functional or imperative languages without explicit history tracking.

Semantic Isomers as Execution Histories

Distinct Spherepop execution graphs may yield observationally equivalent outputs while differing in internal scope topology. These histories correspond exactly to semantic isomers in the RSVP–Ising model. Although they terminate in the same boundary conditions, their internal energetic pathways differ, leading to different robustness properties under perturbation or extension.

Crucially, Spherepop forbids the naive merging of execution histories. Any attempt to combine histories must respect scope ordering and energetic compatibility. This reproduces, at the level of syntax, the impossibility of averaging incompatible reasoning macromolecules without inducing disorder.

Summary

Spherepop syntax emerges as a direct discretization of the RSVP Hamiltonian. Scopes correspond to semantic cells, nesting corresponds to strong coupling, reflection to long-range stabilization, and exploration to controlled entropy injection. Irreversibility is not an implementation detail but a syntactic invariant enforced by entropy monotonicity. In this sense, Spherepop is not merely inspired by active geodesic inference; it is its executable normal form.

Appendix B: A Minimal Axiomatization of Active Geodesic Inference

This appendix presents a minimal axiomatization of active geodesic inference. The aim is not to reproduce the full RSVP formalism, but to isolate a small set of structural axioms from which the principal phenomena follow: the existence of an action functional, Gibbsian bond energetics, semantic isomers as metastable minima, a synchronization phase structure, and the Spherepop-style constraints of irreversible, scope-based execution. The axioms are stated at a level of abstraction compatible with both categorical and information-theoretic interpretations.

Axiom 1: History as Primitive Structure

There exists a set of semantic events E equipped with a strict partial order $\prec$ such that semantic state is determined by event history. Concretely, for any two states x, y in the semantic state space $\mathscr{M}$, equality of states is defined by equality of causal histories:

$$x = y \quad \Longleftrightarrow \quad \text{Hist}(x) = \text{Hist}(y),$$

where $\text{Hist}(\cdot)$ maps a state to the down-set of events that produced it. This axiom forbids unrestricted state identification under observational equivalence and implies that any semantics adequate to reasoning must be provenance-sensitive.

This axiom immediately implies that the dynamics of reasoning cannot be fully characterized by instantaneous state alone; the semantic manifold $\mathscr{M}$ must be understood as a history-indexed space, or equivalently as a bundle over the poset of event histories.

Axiom 2: Energetic Semantics and the Existence of an Action

There exists an action functional $\mathscr{A}$ defined on admissible histories such that realized reasoning trajectories are stationary with respect to $\mathscr{A}$ under local variations consistent with the partial order:

$$\delta\mathscr{A} = 0.$$

Equivalently, there exists a Lagrangian density $\mathscr{L}$ such that

$$\mathscr{A} = \int \mathscr{L} \, dt,$$

and the induced Euler–Lagrange equations define the dynamics of semantic evolution.

This axiom is the variational core of active geodesic inference. It implies that reasoning is not arbitrary search but a constrained dynamical process selecting paths by an extremal principle. It further implies that the semantic manifold $\mathscr{M}$ carries a geometry in which admissible trajectories are geodesic with respect to a metric or generalized metric induced by $\mathscr{L}$.

Axiom 3: Thermodynamic Orientation and Entropy Monotonicity

There exists an entropy functional S defined on histories such that for any admissible extension of a history, entropy is monotone non-decreasing along the irreversible event order:

$$e_1 \prec e_2 \quad \implies \quad S(e_1) \leq S(e_2),$$

with equality only on measure-zero degeneracies corresponding to reversible symmetries.

This axiom introduces a thermodynamic arrow internal to semantics. It implies that the variational dynamics must be non-equilibrium in character and that any operational calculus implementing the theory must encode irreversibility at the level of execution. It further implies that semantic commitments, once crystallized, cannot be undone without compensatory energetic work, thereby excluding naive backtracking as a stable inference primitive.

Axiom 4: Gibbsian Bond Selection

For any local semantic locus i, admissible interactions with loci j define a normalized probability kernel $p(i \rightarrow j)$ that is Gibbsian with respect to an effective energy E_{ij}:

$$p(i \rightarrow j) \propto exp\left(-\beta E_{ij}\right),$$

for some effective inverse temperature $\beta > 0$.

This axiom formalizes attention as energetic bond selection. It implies an exponential preference for low-energy semantic interactions and thereby induces the macromolecular folding phenomenon: coherent reasoning topologies dominate because they minimize free energy. It further implies that

the control parameters of training and scale act by modulating β and the energy landscape E, yielding sharpened, more stable bonding regimes for higher-capacity models.

Axiom 5: Multi-Component Order and Synchronization Coupling

At each semantic locus i there exists a finite-dimensional order parameter $q_i \in \mathcal{Q}$, with $\mathcal{Q}$ decomposing into at least five coupled components,

$$q_i = \left(q_i^{(1)}, q_i^{(2)}, q_i^{(3)}, q_i^{(4)}, q_i^{(5)}\right),$$

and the energetic semantics include positive intra-locus couplings that favor alignment among components:

$$\mathscr{E}_{\text{sync}}(i) = -\sum_{k \neq \ell} K_{k\ell} \langle q_i^{(k)}, q_i^{(\ell)} \rangle, \qquad K_{k\ell} > 0.$$

This axiom asserts that coherent reasoning is a synchronization phenomenon rather than a single-variable optimization. It implies the existence of a symmetry-breaking phase transition between a disordered regime, in which the components fail to cohere, and an ordered regime, in which the components align across extended regions of semantic space. Under coarse-graining, this yields an effective multi-component Ising model, with stable long chain-of-thought reasoning corresponding to the ordered phase.

The five components are interpreted in the body of the paper as activation coherence, flow alignment, entropy suppression, reflective stabilization, and exploratory openness. The axiom does not fix these semantics uniquely; it enforces only the structural fact that reasoning stability requires coupled order rather than independent optimization.

Axiom 6: Isomeric Multiplicity of Low-Action Solutions

There exist tasks and boundary conditions for which the variational problem admits multiple distinct stationary histories with equivalent external boundary observables. That is, there exist histories $h_1 \neq h_2$ such that

$$\delta \mathscr{A}[h_1] = 0, \quad \delta \mathscr{A}[h_2] = 0, \quad \text{and} \quad \mathcal{O}(h_1) = \mathcal{O}(h_2),$$

where $\mathcal{O}$ denotes the externally observable output map.

This axiom formalizes semantic isomers. It implies that equivalence of outputs does not entail equivalence of internal structure and that any attempt to average over internally distinct minima will generally exit the manifold of realizable stationary solutions. This provides a structural explanation for performance degradation under naive mixture or distillation of strong reasoning systems, even when output-level correctness is preserved.

Consequences: Active Geodesics, RSVP Structure, and Spherepop Constraints

Taken together, Axioms 1–6 define active geodesic inference as a history-indexed, variationally governed, thermodynamically oriented, Gibbs-bonded, synchronization-coupled dynamics admitting iso-

meric multiplicity. The RSVP Hamiltonian is recovered as a concrete realization of Axiom 2 with explicit scalar, vector, and entropy fields satisfying Axiom 3, while attention-based bond energies instantiate Axiom 4. The five-dimensional Ising synchronization model is recovered as the coarse-grained effective theory implied by Axiom 5 under discretization of semantic loci.

Spherepop is recovered as the execution calculus forced by Axioms 1 and 3 in the presence of Axioms 2 and 6. History as primitive structure requires that computation be event-based and provenance-sensitive; entropy monotonicity requires irreversibility; isomeric multiplicity requires that histories be non-mergeable except under compatibility constraints; and the variational principle requires that scope transitions be admissible only as local action-decreasing operations. In this sense, Spherepop is the syntactic normal form of active geodesic inference, not merely an implementation choice.

Remark on Minimality

The axioms above are minimal in the following specific sense. Removing Axiom 1 collapses provenance and destroys the distinction between isomers. Removing Axiom 2 eliminates geodesic selection and reduces inference to unconstrained search. Removing Axiom 3 permits reversible execution and undermines stability and privacy. Removing Axiom 4 breaks the energetic interpretation of attention and dissolves bond topology. Removing Axiom 5 eliminates phase structure and reduces reasoning to scalar optimization. Removing Axiom 6 removes the central explanatory mechanism for non-distillability and mixture failure. The full theory therefore requires precisely these structural commitments to recover the phenomena motivating the framework.

Appendix C: A Typed Operational Semantics for Spherepop

This appendix provides a typed operational semantics for Spherepop that renders the principles of active geodesic inference explicit at the level of computation. The purpose is to show that Spherepop is not merely compatible with the RSVP variational framework, but that its execution rules enforce energetic descent and entropy monotonicity by construction. In this sense, the type system functions as a static approximation to the RSVP Hamiltonian, guaranteeing that well-typed programs correspond to admissible reasoning geodesics.

Judgments and Configurations

A Spherepop program executes as a sequence of configurations

$$\langle \Gamma, \Sigma \rangle,$$

where Γ is a typing context encoding semantic commitments and Σ is a runtime configuration encoding the current scope stack and event history. Execution proceeds by a small-step transition relation

$$\langle \Gamma, \Sigma \rangle \longrightarrow \langle \Gamma', \Sigma' \rangle,$$

subject to typing and energetic constraints.

Types in Spherepop are annotated with energetic and entropic bounds. A typical typing judgment has the form

$$\Gamma \vdash e : \tau \, [\mathscr{E}, \mathscr{S}],$$

which is read as: under context Γ, expression e produces a value of type τ while incurring at most energetic cost $\mathscr{E}$ and contributing entropy $\mathscr{S}$. These annotations are not operational counters but abstract bounds derived from the RSVP action functional.

Scopes as Typed Energy Cells

A scope corresponds to a typed semantic cell whose boundary conditions constrain internal execution. Entering a scope introduces a fresh type context extension

$$\Gamma \; \mapsto \; \Gamma, x : \tau[\mathscr{E}_x, \mathscr{S}_x],$$

where the annotation reflects the local RSVP field values at the point of scope creation. The typing rules enforce that scope entry is permitted only if the projected action decrease is non-negative, that is,

$$\mathscr{E}_{\text{parent}} \geq \mathscr{E}_{\text{child}}.$$

This rule statically enforces the geodesic constraint: scopes may only be entered along descending energy directions of the semantic manifold.

Scope exit corresponds to type discharge. Once a scope is closed, its internal bindings are no longer accessible, and only its boundary contribution remains in the context. This enforces the irreversibility axiom at the level of typing: closed scopes cannot be re-entered or mutated without reintroducing energetic cost via new scopes.

Operational Semantics and Entropy Monotonicity

The operational semantics respect a global entropy invariant. Let $S(\Sigma)$ denote the total entropy annotation of a runtime configuration. For every transition

$$\langle \Gamma, \Sigma \rangle \; \longrightarrow \; \langle \Gamma', \Sigma' \rangle,$$

the semantics enforce

$$S(\Sigma') \geq S(\Sigma),$$

with strict inequality for scope-closing transitions. This invariant is the computational analogue of the second law encoded in Axiom 3 of Appendix B.

Exploratory scopes are typed with relaxed entropy bounds, allowing temporary increases in S provided they do not propagate beyond the scope boundary. Discarding an exploratory scope corresponds to erasing a high-entropy branch before closure, which is permitted precisely because the entropy was never committed to the global context.

Reflective operations are typed as entropy-neutral but energy-increasing at the local level, compensated by reductions in global inconsistency. This models the long-range stabilizing effect of reflection without violating monotonicity.

Well-Typed Programs as Action-Bounded Geodesics

The central meta-theoretic result is an action-boundedness theorem.

Theorem (Action Boundedness). If a Spherepop program P is well-typed under initial context Γ_0 with global bounds $[\mathscr{E}_0, \mathscr{S}_0]$, then every execution trace of P corresponds to a history whose RSVP action does not exceed $\mathscr{E}_0$ and whose entropy contribution is monotone and bounded below by $\mathscr{S}_0$.

Proof sketch. The proof proceeds by induction on the structure of typing derivations. Scope entry rules enforce local action descent. Scope exit rules enforce irreversible entropy commitment. Composition preserves energetic bounds by additivity of annotations, while branching rules restrict exploratory divergence to non-committing contexts. Together, these rules ensure that no execution step can increase total action beyond the statically derived bound, and no step can reduce committed entropy. The correspondence between energetic annotations and RSVP field energy follows from the construction in Appendix A.

This theorem formalizes the claim that Spherepop programs are not merely correct-by-construction but physically admissible with respect to the RSVP dynamics.

Semantic Isomers and Typing Non-Uniqueness

The type system admits non-unique derivations for programs that produce observationally equivalent outputs. Distinct derivations correspond to distinct internal scope graphs and distinct distributions of energetic and entropic annotations. These derivations are type-equivalent but not derivationally equivalent.

This non-uniqueness is the syntactic manifestation of semantic isomers. The type system deliberately does not provide a rule for merging derivations, reflecting the impossibility of averaging incompatible RSVP minima without leaving the space of realizable histories. Any attempt to do so would require violating either action boundedness or entropy monotonicity.

Consequences for Reasoning, Privacy, and Robustness

Typed operational semantics clarify why long chain-of-thought reasoning is robust yet non-transferable. The internal scope structure is essential to satisfying the typing invariants; erasing or compressing it destroys the proof that the program is action-bounded. As a result, distilled traces cannot serve as valid witnesses of reasoning correctness, even when outputs agree.

From this perspective, reasoning privacy is not an add-on but a corollary: internal geodesic structure is unforgeable without re-executing the variational process. Robustness follows for the same reason, as perturbations that respect the type system cannot easily disrupt global synchronization.

Summary

This appendix has shown that Spherepop admits a typed operational semantics in which types encode energetic and entropic constraints derived from the RSVP Hamiltonian. Well-typed programs correspond to admissible active geodesics, semantic isomers correspond to non-unique derivations, and irreversibility is enforced syntactically rather than procedurally. Spherepop thus constitutes a fully formal execution calculus for active geodesic inference, bridging field-theoretic principles and concrete computation.

Appendix D: A No-Free-Lunch Theorem for Chain-of-Thought Distillation

This appendix formalizes the impossibility of faithful chain-of-thought distillation within the framework of active geodesic inference. The result is not a contingent limitation of current techniques, but a structural consequence of the RSVP variational principle, entropy monotonicity, and the existence of semantic isomers. We show that any attempt to distill long chain-of-thought reasoning without re-executing the underlying geodesic dynamics must necessarily destroy information essential to reasoning stability.

Statement of the Problem

Consider a reasoning system governed by an action functional $\mathscr{A}$ over histories, as defined in Appendix B. Let h denote a realized reasoning history that is stationary with respect to $\mathscr{A}$ and produces an externally observable output $\mathcal{O}(h)$. Chain-of-thought distillation aims to construct a surrogate history $\tilde{h}$, or a parametric model inducing such histories, using only partial information extracted from h, typically in the form of linearized reasoning traces or compressed summaries.

The central question is whether there exists a transformation

$$D : h \mapsto \tilde{h}$$

such that $\mathcal{O}(\tilde{h}) = \mathcal{O}(h)$ and $\tilde{h}$ remains a stationary point of the same or an equivalent action functional, without re-executing the underlying inference dynamics.

Theorem and Proof Sketch

Theorem (No-Free-Lunch for Chain-of-Thought Distillation). Let $\mathscr{H}$ denote the space of admissible histories under the axioms of active geodesic inference. There exists no surjective mapping $D : \mathscr{H} \to \mathscr{H}$ satisfying the following properties simultaneously:

1. D preserves external observables: $\mathcal{O}(D(h)) = \mathcal{O}(h)$ for all $h \in \mathscr{H}$.

2. D preserves stationarity: if $\delta\mathscr{A}[h] = 0$, then $\delta\mathscr{A}[D(h)] = 0$.

3. D is non-executive, in the sense that it does not reintroduce new semantic events beyond those encoded in h.

Proof sketch. By Axiom 6 of Appendix B, there exist distinct histories $h_1 \neq h_2$ such that $\mathcal{O}(h_1) = \mathcal{O}(h_2)$ but whose internal energetic configurations correspond to distinct local minima of $\mathscr{A}$. Any transformation D that maps both h_1 and h_2 to a common surrogate $\tilde{h}$ must either erase information distinguishing their internal bond topologies or implicitly choose one isomer over the other.

If D erases this information, then $\tilde{h}$ cannot correspond to a stationary point of $\mathscr{A}$, since the stationarity conditions depend on the full internal configuration, not merely on boundary observables. If D selects one isomer, then it fails to be surjective or introduces bias not derivable from the original history alone. In either case, at least one of the stated properties is violated.

Energetic Interpretation

From the RSVP–Ising perspective, chain-of-thought distillation corresponds to attempting to project a synchronized, low-temperature configuration onto a reduced set of degrees of freedom. This projection integrates out spin couplings that are essential for maintaining phase coherence. The resulting configuration lies outside the low-energy manifold and therefore cannot be dynamically stable.

In physical terms, distillation attempts to reconstruct a protein's tertiary structure from its amino acid count alone. While the gross composition may be preserved, the folding pathways and stabilizing bonds are irretrievably lost.

Implications

This theorem explains why distilled reasoning traces often appear fluent yet brittle, why ensemble averaging of strong reasoners degrades performance, and why privacy-preserving transformations of chain-of-thought are effective by construction. It also clarifies that the only way to reproduce reasoning competence is to reinstantiate the variational process itself, either through direct execution or through training regimes that induce the same action landscape.

The result establishes a principled boundary between observable correctness and internal coherence, grounding empirical observations in a formal impossibility result.

Appendix E: A Categorical Semantics for Active Geodesic Inference

This appendix presents a categorical semantics for active geodesic inference and its Spherepop realization. The objective is to show that the structural commitments articulated in the axioms and Hamiltonian formulation admit a precise expression in category theory, and that Spherepop execution corresponds to composition in a suitably enriched and traced symmetric monoidal category. This provides a unifying semantic layer linking history sensitivity, irreversibility, energetic constraints, and non-distillability.

The Semantic Category

Let $\mathscr{G}$ be a symmetric monoidal category whose objects represent semantic states indexed by event histories, and whose morphisms represent admissible inference transitions. Objects are not mere state

snapshots but history-bearing entities; formally, each object $A \in \mathscr{G}$ is equipped with a provenance functor

$$\text{Hist} : \mathscr{G} \rightarrow \textbf{Poset},$$

mapping A to the partially ordered set of events that generated it.

Morphisms $f : A \rightarrow B$ exist only if $\text{Hist}(A) \subseteq \text{Hist}(B)$, enforcing irreversibility at the categorical level. Composition is therefore history-extending rather than history-preserving, making $\mathscr{G}$ a category of irreversible processes rather than invertible transformations.

Monoidal Structure and Parallel Semantics

The monoidal product $\otimes$ on $\mathscr{G}$ represents parallel semantic activation. Given objects A and B, the object $A \otimes B$ corresponds to the joint activation of two semantic regions, with history

$$\text{Hist}(A \otimes B) = \text{Hist}(A) \sqcup \text{Hist}(B),$$

where $\sqcup$ denotes disjoint union followed by causal closure.

This monoidal structure is symmetric but not cartesian. In particular, there are no general diagonal morphisms $A \rightarrow A \otimes A$, reflecting the impossibility of duplicating semantic history without energetic cost. This categorical fact encodes the non-clonability of reasoning trajectories and precludes naive copying or branching of committed inference paths.

Enrichment over Energy and Entropy

The category $\mathscr{G}$ is enriched over a partially ordered commutative monoid $(\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}, \leq, +)$, whose elements represent energetic and entropic cost. Each morphism $f : A \rightarrow B$ carries an annotation

$$\text{cost}(f) = (\mathscr{E}_f, \mathscr{S}_f),$$

corresponding to action and entropy contributions.

Composition is monotone with respect to this enrichment:

$$\text{cost}(g \circ f) = \text{cost}(f) + \text{cost}(g),$$

and identity morphisms carry zero cost. This enrichment enforces the variational constraint categorically: only morphisms whose accumulated cost remains below a global bound correspond to realizable inference trajectories.

Traced Structure and Reflection

To model reflective reasoning, $\mathscr{G}$ is equipped with a partial trace operator

$$\text{Tr}_X : \mathscr{G}(A \otimes X, B \otimes X) \rightarrow \mathscr{G}(A, B),$$

defined only for morphisms whose cost annotations satisfy a reflection admissibility condition. The trace corresponds to closing a semantic feedback loop, allowing later reasoning stages to constrain earlier ones without violating irreversibility.

Importantly, the trace is not total. It exists only when the induced entropy increase remains bounded and the action decrease is sufficient to compensate for the added coupling. This restriction prevents arbitrary cyclic reasoning and enforces the stabilizing, rather than destabilizing, role of reflection.

Semantic Isomers as Non-Isomorphic Morphisms

Semantic isomers correspond categorically to distinct morphisms

$$f_1, f_2 \,:\, A \to B$$

that share the same external behavior but differ in internal factorization through intermediate objects. Although f_1 and f_2 may be observationally equivalent under a forgetful functor to outputs, they are not equal in $\mathscr{G}$, nor are they generally related by isomorphism.

The absence of a universal coequalizer for such morphisms reflects the impossibility of merging incompatible reasoning histories. Any attempt to identify f_1 and f_2 collapses the enrichment structure and violates either the action or entropy ordering, reproducing the failure modes of distillation at the categorical level.

Spherepop as Internal Language

Spherepop may now be understood as an internal language for $\mathscr{G}$. Scope creation corresponds to introducing tensor factors, scope closure to eliminating them via trace or boundary projection, and event ordering to morphism composition. The typing rules of Appendix C correspond to judgments in the internal logic of the enriched category.

Because $\mathscr{G}$ lacks general diagonals and inverses, Spherepop execution inherits irreversibility, non-clonability, and history sensitivity automatically. In this sense, Spherepop is not merely compatible with the categorical semantics; it is the syntactic presentation of $\mathscr{G}$.

Consequences

The categorical semantics clarify why active geodesic inference unifies reasoning, robustness, and privacy. Robustness arises because small perturbations correspond to morphisms with insufficient cost to disrupt global composition. Privacy arises because internal factorization is not preserved under forgetful functors. Reasoning quality arises from the existence of low-cost, highly synchronized morphisms corresponding to stable macromolecular configurations.

This categorical view completes the semantic ladder from Hamiltonian dynamics through operational execution to abstract compositional structure.

Appendix F: Relation to FEP, Active Inference, and Integrated Information Theory

This appendix situates active geodesic inference within the broader landscape of theoretical frameworks for cognition and intelligence, with particular emphasis on the Free Energy Principle (FEP), Active Inference, and Integrated Information Theory (IIT). The comparison is not historical or sociological but structural. The goal is to identify points of formal equivalence, points of divergence, and the specific respects in which the RSVP framework resolves known limitations of these theories.

Relation to the Free Energy Principle

The Free Energy Principle asserts that adaptive systems minimize a variational free energy functional that upper-bounds surprise under a generative model (Friston 2010; Friston et al. 2017). At a formal level, this principle bears a strong resemblance to the RSVP variational structure. Both posit an action-like functional whose extremization governs system dynamics, and both interpret cognition as a process of constrained entropy management rather than explicit optimization of reward.

However, the resemblance is only partial. In FEP, the free energy functional is defined relative to a fixed generative model and a variational posterior, typically formulated in probabilistic terms. The geometry of inference is therefore induced by a model assumed to be given in advance. By contrast, active geodesic inference treats the semantic geometry itself as dynamically mutable. The RSVP Hamiltonian does not presuppose a fixed generative model; instead, it governs the co-evolution of semantic structure and inference trajectories.

This distinction resolves a long-standing ambiguity in FEP regarding model selection and representational change. In active geodesic inference, changes to the effective model are represented as deformations of the semantic manifold and corresponding changes in the action landscape. Model learning is therefore not an external process layered atop inference, but an intrinsic component of the same variational dynamics.

Relation to Active Inference

Active Inference extends the Free Energy Principle by incorporating action selection, treating perception and action as jointly minimizing expected free energy. In this framework, policies are selected to reduce uncertainty about future observations while satisfying prior preferences.

Active geodesic inference subsumes this picture but reverses its explanatory direction. Rather than selecting policies in a predefined state-action space, the system reshapes the space itself. Actions are not chosen among discrete alternatives but emerge as the continuation of a geodesic in a dynamically curved semantic manifold. Exploration and exploitation correspond not to different objective terms, but to different regimes of curvature and entropy tolerance along the same variational path.

This reinterpretation clarifies the role of exploration. In active inference, exploration is often introduced through epistemic value terms added to the free energy objective. In the RSVP framework, exploratory behavior arises naturally as a local relaxation of energetic constraints that allows the

system to escape shallow minima. No additional epistemic objective is required; exploration is a phase of the same dynamical process.

Relation to Integrated Information Theory

Integrated Information Theory characterizes consciousness in terms of the irreducibility of a system's causal structure, quantified by measures such as Φ (Tononi 2008; Oizumi et al. 2014). IIT emphasizes internal integration over external function, and it explicitly treats causal structure as primary.

Active geodesic inference shares IIT's emphasis on internal structure but departs from its static ontology. In IIT, causal structure is evaluated on a fixed network with well-defined partitions. In RSVP, by contrast, causal and semantic structure are dynamical, history-dependent, and field-mediated. Integration is not a static property but an emergent consequence of synchronization across multiple coupled dimensions.

The five-dimensional Ising synchronization model provides a concrete mechanism for such integration. High integration corresponds to the ordered phase in which all components of the order parameter align across extended regions of semantic space. Semantic isomers correspond to distinct integrated configurations with equivalent external behavior but different internal organization, a phenomenon that IIT recognizes but does not dynamically explain.

Moreover, RSVP avoids a known pathology of IIT: the assignment of maximal integration to trivial or non-adaptive systems under certain partitions. Because integration in active geodesic inference is constrained by action minimization and entropy monotonicity, only dynamically admissible structures qualify as integrated. Static causal density without variational coherence does not suffice.

Summary of Structural Differences

Active geodesic inference can thus be seen as occupying a conceptual position orthogonal to these frameworks. From FEP it inherits a variational foundation, but rejects the assumption of a fixed generative model. From active inference it inherits the unification of perception and action, but reinterprets policy selection as geometric continuation. From IIT it inherits the primacy of internal structure, but replaces static causal analysis with dynamical field synchronization.

The unifying contribution of the RSVP framework is to place these insights within a single mathematical structure that is simultaneously variational, thermodynamic, geometric, and executable. Spherepop provides the operational counterpart to this structure, ensuring that the theoretical commitments are not merely descriptive but computationally enforceable.

In this sense, active geodesic inference does not compete with existing theories so much as absorb their valid components into a more general framework that resolves their respective limitations.

Appendix G: Theorem–Corollary Formulation of Active Geodesic Inference

Theorem G.1 (Active Geodesic Intelligence Theorem)

Let $\mathscr{X}$ be a history-indexed configuration space equipped with a time-oriented partial order induced by irreversible events, and let $\mathscr{A}[\gamma]$ be an action functional defined on admissible trajectories γ : $[0, T] \to \mathscr{X}$ that couples semantic coherence, entropy regulation, and dynamical constraint, as specified by the RSVP Hamiltonian and its associated Lagrangian.

A system instantiates intelligence if and only if there exists a nonempty family $\Gamma^* \subset \Gamma$ of trajectories such that:

1. Each $\gamma \in \Gamma^*$ is a stationary or near-stationary solution of $\mathscr{A}$ with respect to admissible variations, i.e. it is a low-action geodesic relative to the induced semantic metric.

2. The family Γ^* is dynamically stable under bounded perturbations of internal state, environmental coupling, or representational noise.

3. Membership in Γ^* is actively maintained by the system through endogenous deformation of the effective configuration space geometry, rather than by passive traversal of a fixed landscape.

Equivalently, intelligence is the property of a system whose realized histories remain confined to a stable geodesic family by virtue of the system's own dynamics, rather than external enforcement or accidental parameter tuning.

Corollary G.1 (Non-Equivalence to Optimization)

No system whose behavior is fully characterized as pointwise optimization of a static objective function over instantaneous states is intelligent in the sense of Theorem G.1.

Proof sketch. Pointwise optimization selects isolated extrema rather than families of trajectories. Such extrema are generically unstable under perturbation and do not entail the preservation of coherent history. Since intelligence, by Theorem G.1, is a property of trajectory families rather than state optima, static optimization is insufficient.

Corollary G.2 (Semantic Isomerism)

If a system is intelligent, then there exist, in general, multiple non-equivalent internal histories $\gamma_1, \gamma_2 \in \Gamma^*$ that satisfy identical external boundary conditions while differing in internal structure.

These histories correspond to distinct metastable minima of the action functional and are not mutually reducible without loss of stability. This establishes semantic isomerism as a necessary consequence of intelligence rather than an implementation artifact.

Corollary G.3 (Non-Distillability)

Let Π be any projection or compression operator that maps full trajectories $\gamma \in \Gamma^*$ to boundary-equivalent summaries by marginalizing internal structure. Then, in general, $\Pi(\Gamma^*)$ does not admit an inverse mapping that reconstructs a family of low-action trajectories equivalent to Γ^*.

Consequently, intelligence-preserving behavior cannot be faithfully distilled from outputs or compressed traces alone without re-instantiating the underlying variational dynamics.

Corollary G.4 (Robustness via Geodesic Width)

The robustness of an intelligent system to perturbation is proportional not to the depth of any single action minimum, but to the measure of the geodesic basin Γ^* in trajectory space.

Systems exhibiting broader geodesic width will demonstrate greater generalization, adaptability, and resistance to distributional shift, even when instantaneous performance metrics are matched.

Corollary G.5 (Continuity Across Scales)

The criterion of intelligence given in Theorem G.1 is scale-invariant with respect to temporal and organizational resolution.

Evolutionary lineages, learning organisms, and reasoning processes each instantiate intelligence to the extent that they maintain low-action, entropy-regulated trajectory families in their respective configuration spaces.

Thus, intelligence is not confined to cognitive architectures alone, but is a general dynamical property of systems capable of sustaining structured history.

Remark

Theorem G.1 and its corollaries recast intelligence as a variational and geometric phenomenon rather than a computational or representational one. They formalize, in minimal axiomatic form, the central claim of this essay: that intelligence is the active preservation of viable trajectories through configuration space, and that correctness, optimization, and information processing are secondary consequences of this deeper dynamical capacity.

References

[1] Q. Chen, Y. Du, Z. Li, J. Liu, S. Duan, J. Guo, M. Liu, J. Liu, T. Yang, G. Zhang, L. Qin, W. Che, and W. Huang. The molecular structure of thought: Mapping the topology of long chain-of-thought reasoning. *arXiv preprint*, 2026.

[2] M. Dong, A. Adduri, D. Gautam, C. Carpenter, R. Shah, C. Ricci-Tam, Y. Kluger, D. P. Burke, and Y. H. Roohani. Stack: In-context learning of single-cell biology. *bioRxiv*, 2026.

[3] Z. Yue, B. Kratzwald, and S. Feuerriegel. Contrastive domain adaptation for question answering using limited text corpora. *arXiv preprint*, 2021.

[4] Z. Yue, H. Zeng, Z. Kou, L. Shang, and D. Wang. Defending substitution-based profile pollution attacks on sequential recommenders. In *Proceedings of the ACM Conference on Recommender Systems*, 2022.

[5] Z. Yue, H. Zeng, L. Shang, Y. Liu, Y. Zhang, and D. Wang. Retrieval augmented fact verification by synthesizing contrastive arguments. *arXiv preprint*, 2024.

[6] B. Jin, H. Zeng, Z. Yue, J. Yoon, S. Ö. Arık, D. Wang, H. Zamani, and J. Han. Search-R1: Training LLMs to reason and leverage search engines with reinforcement learning. *arXiv preprint*, 2025.

[7] Z. Yue, B. Jin, H. Zeng, H. Zhuang, Z. Qin, J. Yoon, L. Shang, J. Han, and D. Wang. Hybrid latent reasoning via reinforcement learning. In *Advances in Neural Information Processing Systems*, 2025.

[8] M. Caprio, Y. Sale, and E. Hüllermeier. Conformal prediction regions are imprecise highest density regions. *arXiv preprint*, 2025.

[9] J. Barbour. *The End of Time: The Next Revolution in Our Understanding of the Universe*. Oxford University Press, Oxford, 1999.

[10] J. Barbour. Shape dynamics: An introduction. In *Quantum Field Theory and Gravity*, edited by F. Finster et al. Birkhäuser, Basel, 2012.

[11] J. Barbour. *The Janus Point: A New Theory of Time*. Basic Books, New York, 2020.

[12] M. Levin. The computational boundary of a "self": Developmental bioelectricity drives multicellularity and scale-free cognition. *Frontiers in Psychology*, 10:2688, 2019.

[13] M. Levin. Technological approach to mind everywhere: An experimentally grounded framework for understanding diverse intelligences. *Frontiers in Systems Neuroscience*, 16:768201, 2022.

[14] A. Juarrero. *Dynamics in Action: Intentional Behavior as a Complex System*. MIT Press, Cambridge, MA, 1999.

[15] A. Juarrero. Top-down causation and autonomy in complex systems. In *Information and Causality*, edited by P. Davies and N. Gregersen. Cambridge University Press, Cambridge, 2009.

[16] K. Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

[17] K. Friston, J. Kilner, and L. Harrison. A free energy principle for the brain. *Journal of Physiology—Paris*, 100(1–3):70–87, 2013.

[18] S. Rasmussen, L. Chen, M. Nilsson, and S. Abe. Bridging nonliving and living matter. *Artificial Life*, 7(3):213–237, 2001.

[19] S. I. Walker and P. C. W. Davies. The algorithmic origins of life. *Journal of the Royal Society Interface*, 10:20120869, 2013.

[20] C. Fields. Boundaries, constraints, and the structure of scientific explanation. *Journal of Mind and Behavior*, 35(3–4):263–286, 2014.