On the Emergence of Cognition From Complexity

Cole B. Johnson

Abstract

Complexity drives the emergence of life, yet frameworks remain insufficient to capture how abstract systems such as genetics, cognition, and culture arise from increasingly intricate structures. In this paper, we introduce two theories: Phase Embeddings and Extended Assembly Theory. Phase Embeddings explains how each major transition between paradigms of complexity – stellar to mineral, mineral to biological, biological to cognitive, etc. – builds upon the contents of its predecessor, forming hierarchical layers of complexity. Extended Assembly Theory augments Assembly Theory by incorporating compression, temporal constraints, and the basis-specific interactions between form and function, thus accounting for phase transitions where abstract units (e.g., ideas, memes) operate with environmental reference under conditions of scarcity. We demonstrate how cognition qualifies as a significant phase transition, driven by universal selection principles that promote stability, dynamic persistence, and novelty generation. We then validate our theories by examining the cultural phase transition, where rapid growth in complexity arises from meme-based evolution. Finally, we argue under our proposed frameworks that consciousness, unlike previous phases, actively assimilates new transitions, except when emerging forms violate embodiment, a boundary that draws an intrinsic horizon to conscious evolution.

Keywords Assembly Theory · Phase Transitions · Self-Replication · Origins of Life · Cognition

Introduction

Life is nearly undefinable; like consciousness, our belonging to it blinds us to its true nature. Regardless, many have characterized aspects of life under the different paradigms of reason throughout time. In one of the first scientific attempts, Schrödinger, argued that life defies entropy by consuming energy to maintain order, a fragile balancing act against the universe's tendency toward chaos [1]. Emerging from this disorder, life constructs itself through complexity, a process recently beginning to be explained by Assembly Theory which posits the emergence of complexity through selection mechanisms that facilitate assemblies resulting in configurations that would be combinatorically improbable through random events and, eventually, without self-replication [2].

Despite objections to this theory of life, the notion of complexity does arise from the unification of these frameworks. Wong et al. posit a hierarchical selection process that begins with principal forces, such as kinetic energy, driving the formation of stable energetic dynamical systems where configurations which maintain structural integrity are favored [3]. The channeling of outside energy into a stable system induces a negentropic state – imagine shaking two glass panels containing scattered metal balls. This will result in a lattice-like structure that becomes increasingly resistant to disruption, much like the process of heat-treating metal.

These stable systems act as scaffolds, supporting secondary functions that enhance their persistence, including novelty production. This hierarchy explains how complexity builds over time, with lower-level stability enabling higher-order processes that generate adaptive and innovative features essential for evolution.

When this kind of order is combined with self-replication, it becomes the origin of complexity. The antifragility of such ordered systems is exceeded only by the dynamic meta-stability, where cycles of self-reproducing processes, such as autocatalytic sets, reinforce and extend their stability, as described by Filisetti et al [4]. By the logic of Assembly Theory, this dominance over time makes self-replicating assemblies the sole viable components of highly complex systems, as their repeated assemblies avoid the combinatorial explosions that would otherwise render complexity impossible.

This type of complexity is found in every mature, stable system, with extremely rapid growth often be attributed to the role of phase transitions. Kauffman abstractly describes phase transitions as events that fundamentally alter selection pressures [5]. By opening new energy pathways, they enable systems to leap into higher-order complexity.

These transitions are not gradual but occur explosively, creating novel systems that exploit negentropic (negative entropy) channels to amplify global complexity.

Interestingly, phase transitions drive complexity growth even outside the chains of self-replicating systems we typically associate with life, such as in the abiogenesis of life itself. Reality, as some argue, may be rooted in computational frameworks, while others posit it is fundamentally random, with emergent structures trickling up from this basis. In the computational view, phase transitions lead to events like Fontana's "turning gas," where a leap in computation produces a sharp rise in Kolmogorov Complexity upon the advent of self-replication, as seen in Blaise Agüera y Arcas work [6]. Conversely, in a random basis of reality, Kauffman highlights the emergence of autocatalytic networks from prebiotic "soup," demonstrating how explosive transitions can create the foundation for increasingly intricate systems. Thus, whether rooted in computational frameworks or arising from randomness, complexity can emerge through phase transitions, even in the absence of self-replicating chains, and can ultimately give rise to self-replication itself.

Despite serving as a promising framework for understanding life, Assembly Theory and similar approaches fail to address properties of phase transitions that result from it – most notably, cognition. In the pursuit of understanding what life is, we inevitably rely on our cognition to explore and define its nature. Given that cognition is both the tool and the subject of inquiry, it is inherently inconsistent to construct a theory of life that excludes the emergence of cognition. A theory that cannot account for cognition's evolution from complexity cannot justifiably claim to explain the emergence of life as a whole.

Thus, in this paper, we aim to formalize the nature of phase transitions in evolving systems and extend Assembly Theory to account for cognition. In Section 1, we will argue for cognition as a significant phase transition and, by examining the differences between transitions, develop the concept of phase embeddings. In Section 2, we will address the insufficiency of Assembly Theory in modeling phase transitions, particularly cognition, and propose a formal extension to rectify these limitations. Section 3 will demonstrate the explanatory power of this extended framework by applying it to other phase transitions, with a focus on cultural evolution. Finally, in Section 4, we will define consciousness within the framework of phase embeddings and speculate about the nature of the next phase transition.

I Cognition as a Phase Transition & Phase Embeddings

Cognition is often overlooked in discussions of phase transitions, even in arguments regarding the nature of life. By examining evolving systems over time, we demonstrate that cognition mirrors universal principles of assembly and selection, continuing natural trends of the hierarchy of complexity, as well as introducing novel paradigms itself. To illustrate this, we build off the work proposed by Wong et al. We then develop the notion of cognition's evolution as a phase transition through arguments made by Kauffman, having satisfied the precondition of positioning it inside the hierarchy of complexity.

Selection and Complexity in Cognition

Wong et al. develop a framework in which complexity arises through hierarchical selection processes. These processes begin with basic forces like kinetic energy, enabling the formation of stable configurations that serve as scaffolds for more advanced systems. To illustrate this framework, Wong et al. explore a sequence of related phenomena: stellar evolution, mineral evolution, and biological evolution. Each system showcases increasing configurational diversity, functional persistence, and shared mechanisms of selection.

Stellar evolution in our current understanding of the universe is the earliest stage of complexity. Stars form as hydrogen-helium masses undergoing nuclear fusion, producing heavier elements through successive fusion stages. Explosive events like supernovae and neutron star collisions generate the diversity of atomic nuclei observed today. The building blocks are atoms and nuclei, with stability as the primary selection criterion – from the identity relation, we know that configurations that resist decay persist, illustrating static persistence at the foundational level.

Mineral evolution represents the next step in complexity, where planetary differentiation drives the formation of increasingly diverse minerals. Processes like condensation, crystallization, and chemical interactions generate over 5,900 documented mineral species, each stage building on prior configurations. Here, chemical elements and compounds serve as the building blocks, and stability under local environmental conditions determines persistence.

Biological evolution introduces adaptability as a key driver of selection. Genetic recombination and mutation create vast combinatorial spaces, while natural selection retains traits that enhance survival and reproduction. Genes are the fundamental units, with configurations evolving based on their ability to persist in changing environments. This dynamic persistence represents a shift from mere stability to systems actively maintaining and adapting themselves.

Extending this framework to cognitive evolution, we propose ideas as its fundamental building blocks. Like genes, ideas interact within vast combinatorial spaces, producing novel configurations that undergo selection. Cognitive systems exhibit antifragility, thriving under disruption and refining ideas to meet challenges. Configurations that foster adaptability, innovation, and persistence are favored, mirroring the selection dynamics of earlier systems.

Despite their differences, these systems share key features. Each operates within massive combinatorial spaces, driven by diverse, interacting components – atoms, elements, genes, or ideas. Processes generate diverse configurations, from nuclear fusion to cognitive synthesis, while selection favors stability and functionality. Wong et al.'s hierarchical selection framework formalizes this progression:

- *First-order selection (Static Persistence)*: Configurations resisting decay, such as stable nuclear particles, durable minerals, and conserved genetic traits.
- *Second-order selection (Dynamic Persistence)*: Processes sustaining systems, including energy dissipation, autocatalysis, and homeostasis.
- *Third-order selection (Novelty Generation)*: Systems generating new configurations and functions, such as novel isotopes, mineral phases, or adaptive traits.

Cognitive evolution aligns with these principles. Its foundational ideas exhibit static persistence, adaptive frameworks demonstrate dynamic persistence, and the creation of innovative paradigms reflects novelty generation.

Cognition as a Phase Transition

Phase transitions in complexity theory, as posited by Kauffman, occur at the 'edge of chaos', a critical point where systems balance stability and adaptability. Cognitive systems achieve optimal operation at the edge of chaos by balancing stability – preserving established frameworks – and plasticity – allowing reorganization in response to new challenges and generation of novelty. This balance supports innovation without descending into disarray.

This critical balance is not accidental but a result of selection. Evolutionary cytoarchitectonic theory suggests that natural selection tunes neural and cognitive architectures to this critical point. As Kauffman asserts, "If it proves true that selection tunes genomic systems to the edge of chaos, then evolution is persistently exploring networks constrained to this fascinating ensemble of dynamical systems" (1993, p. 522). This persistent exploration enables cognitive systems to remain dynamic yet robust, ensuring they are equipped to navigate the complexities of their environments.

Phase transitions in cognition parallel those in physical and biological systems, where thresholds are crossed, producing dramatic increases in complexity. Just as abiogenesis transitioned molecular systems to life, cognition emerged from a pre-cognitive soup, characterized by loosely structured interactions.

Theories on the guiding factors of these transitions vary. Social collaboration through language, as highlighted in The Enigma of Reason, provided selective pressures favoring complex behaviors. Shared knowledge systems and collaborative problem-solving enhanced adaptability and survival. For instance, myths and collective reasoning frameworks unified communities, fostering cognitive advancement.

Alternatively, the bicameral hypothesis, proposed by Julian Jaynes, posits that early cognition involved structured but non-introspective mechanisms, such as perceived divine commands. This pre-conscious state transitioned to introspective thought, driven by both biological changes and environmental pressures, enabling more sophisticated reasoning and adaptability.

Thus, cognition is driven by function and selection, as articulated in Wong et al.'s framework, establishing cognition as a complex, evolving system. Additionally, this progression of complexity aligns with Kauffman's criteria for phase transition events.

Phase Embeddings

Phase transitions in complex systems are not isolated phenomena but are deeply interconnected through what can be described as negentropic embeddings. These embeddings reveal how the units of each system are derived from the generation of the preceding phase transition, creating a chain of dependencies that define the evolution of complexity. This interconnectedness leads us to the concept of embedded phases, which positions each phase transition as nested within the structures formed by earlier transitions, building a hierarchy of systems upwardly reliant on one another for their existence and resources.

The relationship between embedded phases is evident across several domains. Mineral evolution, for instance, is fundamentally dependent on the outputs of stellar nucleosynthesis. Atoms like carbon, oxygen, and silicon – essential to mineral formation – are forged during the nuclear fusion processes of stars and released during supernovae and neutron star collisions. Similarly, the genetic backbone of life relies on the chemical elements produced and cycled through mineral evolution. DNA, RNA, and proteins are constructed from elements like carbon, nitrogen, and phosphorus, illustrating the reliance of biological systems on the chemical complexity provided by minerals.

This embedding implies similarities with Schrödinger's concept of life as a defiance of entropy – similar forms of entropy are found at every level of these interconnected systems. At the atomic level, the second law of thermodynamics introduces entropy by dictating that the number of particles exiting an interaction exceeds the number entering. At the chemical level, bond reconfigurations are governed by probabilistic models of entropy, which influence molecular transformations. At the genetic level, genetic entropy drives variability, enabling adaptation and evolution. These parallels between the entropic behavior of embedded phases are why there is continuity in the characteristics of the evolution of complexity across the disparate domains, and perhaps hint at why it is so difficult to define life itself.

Tracing this chain backward, we find that the atomic units utilized in stellar evolution are themselves products of the primordial singularity, where the universe's first elements – hydrogen and helium – were formed. At deeper levels, we encounter even more fundamental forms of entropy. Nuclear entropy governs the interactions of atomic nuclei during stellar processes, while quantum entropy, particularly the entropy of entanglement, reveals the quantum correlations between particles that distribute information across systems. These fundamental layers further embed each phase within an intricate web of dependencies.

These phase embeddings operate as negentropic microcosms, with each phase both reliant on and limited in its access to the systems preceding it. For instance, molecular bonds are embedded within atoms, yet their behavior is constrained by atomic properties. Similarly, life is embedded within molecular chemistry, relying on stable bonds for its persistence but incapable of directly engaging with the subatomic behaviors underlying those bonds. Cognition, too, is embedded within life, dependent on biological stability and self-reproduction for its existence; just as life would cease if molecular bonds were to decay, cognition would collapse without the scaffolding of living systems. Here, 'access' takes on different meanings depending on the phase: molecules access through spatial proximity, genes through informational awareness, and cognition through intent. In all cases, phases derive their output through interactions with units within their own phase or those of preceding ones.

This notion of phase embeddings invites reflection on how humans conceive the limits of observability. Traditional perspectives frame observability in two ways: spatially, as the bounds of the observable universe, or temporally, as the limit imposed by the present and the primordial singularity. However, the framing of complexity evolution through negentropic embeddings challenges these spatiotemporal premises. Much like geocentrism once constrained our understanding of the cosmos, these dimensional perspectives limit our grasp of the systems in which we are embedded. Instead, it is more compelling to think of the observable as the chain of negentropic embeddings within which we are positioned. Each phase transition restricts access to its predecessor while providing the foundation for the next, shaping what is discoverable to us. This notion reframes observability not as a static boundary but as a dynamic range within which we operate and expand through technology.

II Extending Assembly Theory

Assembly Theory (AT), introduced by Sharma et al., explains how complexity emerges from simple building blocks under selection pressures. Thus, it is a useful framework to examine the dynamics of evolving systems within phase transitions. However, while demonstrated successfully in chemical processes, AT requires extension to apply across all phase embeddings.

Assembly Theory

AT redefines objects not as static entities but as products of their formation histories. At the core of AT is the *assembly index* (AI), a measure of the minimal number of steps required to construct an object from its elementary components. This index quantifies the memory embedded in an object's construction, distinguishing it from entropy by its explicit dependence on historical causality. Complementing the AI is the concept of *copy number*, which reflects the abundance of identical objects within a system. Together, these metrics reveal the fingerprints of selection: objects with high assembly indices that appear in large quantities are evidence of non-random processes driving their formation. AT has been validated experimentally in chemical systems where molecular assembly can be traced using mass spectrometry.

The mathematical formulation of AT further refines this framework. Assembly (A) combines the AI and copy number to capture the cumulative selection required to produce an ensemble of objects:

$$A = \sum_{i=1}^{N} \frac{n_i \cdot e^{a_i}}{N_T} - 1$$

Here, a_i is the assembly index of object i, n_i its copy number, N the number of unique objects, and N_T the total number of objects in the ensemble. This equation encapsulates the balance between the difficulty of discovering new configurations and the ease of replicating selected ones.

However, the potential of assembly spaces to grow super-exponentially reveals a critical role for selection. Without constraints, combinatorial possibilities expand uncontrollably, rendering the exploration of new objects infeasible. AT resolves this by introducing historical contingency, where assembly processes are channeled along specific paths dictated by prior constructions. Selection emerges naturally in this context, as pathways that optimize stability and replication are favored.

A necessary condition for sustaining such growth is the equilibrium between exploration and exploitation. This is formalized through the balance of discovery time (T_d) – the time required to find novel configurations – and production time (T_p) – the time to replicate existing ones. When $T_p = T_d$, the system avoids both stagnation and runaway combinatorial explosions, enabling stable complexity growth. Imbalances, such as $T_d \gg T_p$, as seen in prebiotic chemistry, lead to inefficiencies, while systems achieving near-equilibrium, like biological evolution, sustain complexity through selective retention.

Form vs. Function & Temporal Cost

In chemical systems, AT functions effectively because there is no distinction between form and function. The final assembly – such as a molecule – fully encapsulates the information it represents and the function it serves. This alignment allows AT to capture the recursive assembly of atoms into molecules, where the process and outcome are intrinsically linked.

However, genetic systems highlight a fundamental departure. A DNA sequence, while an assembly product, is functionally inert without the mechanisms to decode and act upon it. John von Neumann anticipated this, describing a Turing-complete system with three essential components: a tape to hold information (DNA), a machine to read the tape (ribosomes), and a machine to copy the tape for replication (DNA polymerase) [7]. Schrödinger's earlier insights into life's negentropic nature similarly recognized the role of information, though without mechanistic detail. In this framework, the genetic system operates on a basis analogous to computation where the conserved chemical machinery serves as a stable foundation for genetic encoding. Thus, the act of assembling within genetic systems is nonsensible without specifying the basis upon which assembly occurs.

Cognitive systems further challenge AT by introducing significant temporal constraints. In chemical systems, reactions typically occur on millisecond-to-second timescales and rarely degrade, making the cost of assembly negligible. In contrast, the assembly of ideas in cognitive systems occurs within an inherently temporally scarce environment, where life itself is subject to evolving selection pressures and limited energy inputs. In other words, biological life ends, and when it does, the basis upon which ideas are stored changes. Thus, there is a temporal cost associated with assemblies not present in other preceding phases.

The time required to assemble cognitive units grows exponentially with the complexity of the assembly, making it increasingly unlikely for such assemblies to occur within a feasible timeframe. This logic parallels that of Sharma et al., but while combinatorial explosion in traditional AT results in an assembly space too vast to explore, in cognitive systems, the temporal constraints of biological life impose an additional limitation. As a result, the shortest assembly pathway in cognition does not align with the assumptions of traditional AT, which posits that only the combinatorically shortest pathway could naturally occur. For example, the development of Newtonian mechanics was likely facilitated by religion as a cognitive mechanism, providing generations of social stability necessary for the discovery of natural laws. Without temporal constraints, however, the shortest assembly pathway would likely resemble the one predicted by traditional AT, as the limitations imposed by time would no longer dictate the assembly process.

These challenges reveal AT's broader limitations. While the theory works well for many chemical systems, it struggles in edge cases where the relationship between form and function is less direct. Genetic systems expose AT's neglect of the computational basis, where decoding mechanisms are as essential as the information encoded. Cognitive systems, operating under severe temporal constraints, pose an even greater challenge.

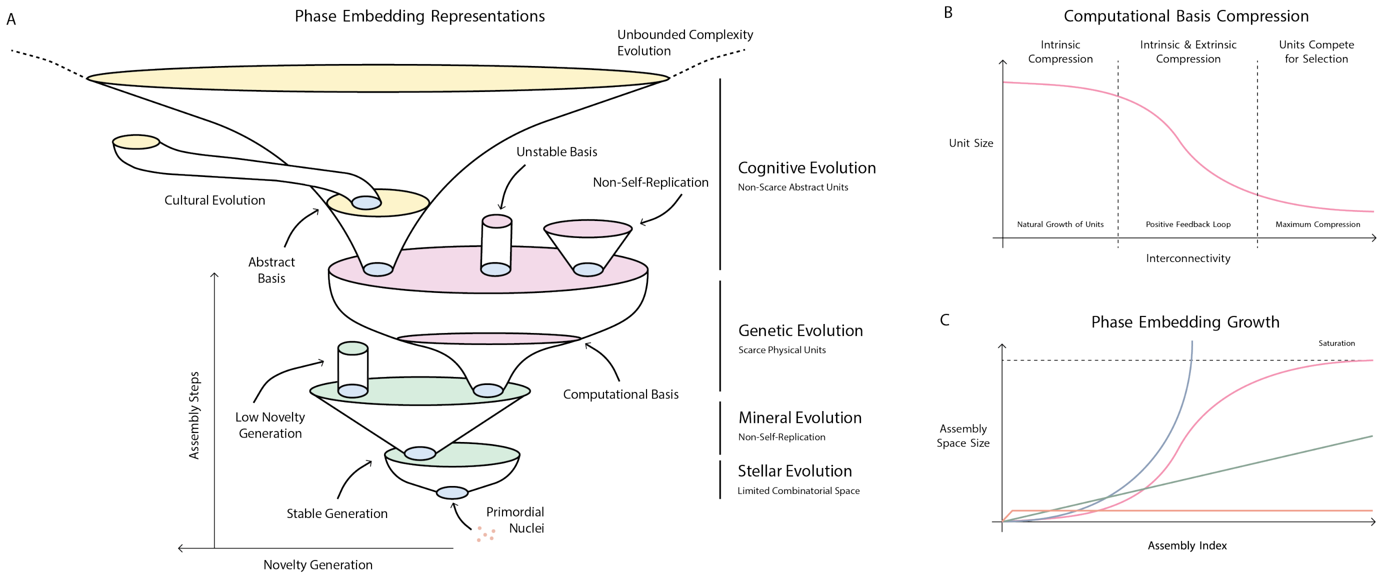

Fig. 1. Phase Embeddings, Compression, and Growth Dynamics. Figure A: Phase transitions (blue ellipses) begin with the primordial nuclei produced by the Big Bang. Stellar evolution introduce complexity from supernovae and other events which create heavier elements, eventually resulting in a stable state (green ellipses). The novelty generation in this phase is very limited from a small combinatorial space. The stability enables a the evolution of minerals, which are not dynamically meta-stable as they cannot self replicate. The stability resulting from mineral evolution then enables systems to generate high levels of novelty, and those which do not are selected against. Genetic evolution introduces a computational basis (purple ellipses) which are stable physical systems that can self-reproduce. Phase transitions which do not self-replicate or are unstable after genetic evolution are selected against. Cognitive evolution abstracts the computational basis and enables unbounded complexity evolution as there is no scarcity in the units that it assembles. At thresholds of cognitive evolution, culture, consciousness, and other phases transition, and are assembled upon as well. Figure B: The computational basis introduced by genetic evolution enables the compression of units. In disconnected systems, unit sizes are large and can only be intrinsically compressed. As the interconnectivity of the basis increases, a positive feedback loop between unit size decreasing and the increase in connectivity is introduced through extrinsic compression. As maximal compression is reached in highly interconnected systems, units themselves begin to compete for selection as there is temporal cost associated with the self-replication of units. Figure C: The relation between the size of the assembly space and the assembly index is a framework to understand the rate of complexity evolution in systems. In systems which have a limited combinatorial space (orange line), such as stellar evolution, the assembly space size is quickly reached after a small number of assemblies. In non-self-replicating systems (green line), a linear growth in complexity occurs, such as seen in mineral evolution. In systems with scarce physical units (red line), a logistic growth occurs, as seen in population dynamics in genetic systems. Finally, in systems with non-scarce abstract units (blue line), an unbounded exponential growth is observed, enabling phase transitions at many different assembly indexes to occur.

Compression of Units

Compression is a fundamental aspect of cognition, transforming raw sensory data into abstract ideas that are cognitively useful. Unlike physical assemblies, these ideas are not inherently interpretable without the neural circuits and broader cognitive frameworks in which they are embedded. This transformation involves two distinct forms of compression: intrinsic and extrinsic.

Intrinsic compression occurs within the structure of the idea itself, minimizing redundancy and optimizing efficiency. For instance, sparse coding in the visual system uses the minimal number of neurons to represent visual input, ensuring a compact and efficient encoding [8]. In contrast, extrinsic compression depends on the brain's existing knowledge base and expectations, prioritizing novelty over redundancy. The Bayesian Brain Hypothesis exemplifies this process, suggesting that the brain predicts expected sensory input and focuses on discrepancies – much like delta encoding in machine learning [9].

This process of compression inherently involves loss. Unlike traditional Assembly Theory (AT), which assumes that unselected components remain available in the assembly pool, compression dissipates unused elements, making them unavailable for future assemblies. This challenge extends beyond cognition to genetic systems, where compression also plays a critical role. Genetic sequences, while compact, rely on selective expression mechanisms such as transcription and translation. In these processes, non-coding regions (introns) often remain dormant, while coding regions (exons) represent the compressed, functional components.

AT's foundational assumption of alignment between form and function breaks down in systems which compress their units. In genetic systems, compressed sequences must be decoded and expressed through additional mechanisms, introducing a discrepancy between the form of the assembly and its function. Similarly, in cognition, intrinsic and extrinsic compression transform raw input into abstract ideas that cannot be understood independently of the systems in which they operate. AT does not model this translation, nor does it account for the selective loss inherent in the process.

Thus, to extend AT to these domains, it must account for the transformation of form into function, the dissipation of unselected components, and the constraints imposed by compression and temporal cost on assembly dynamics.

Extended Assembly Theory

Intrinsic compression functionally reduces the complexity of a unit (idea, in the case of cognition) while preserving its essential information. This is achieved through the computational basis (e.g., neural circuits), which maps high-dimensional inputs into lower-dimensional representations. The efficiency of this mapping can be quantified as:

$$\gamma_{\text{intrinsic}}(a) = \frac{\text{Dim}(I_{\text{comp}})}{\text{Dim}(I)} = \eta,$$

where η represents the efficiency of the computational basis, $\text{Dim}(I)$ is the dimensionality of the original idea, and $\text{Dim}(I_{\text{comp}})$ is the dimensionality after compression.

Extrinsic compression leverages the relationships between an idea and its neighboring ideas to reduce complexity further. This is mathematically expressed as:

$$\gamma_{\text{extrinsic}}(a, N) = e^{-\lambda N},$$

where λ reflects the effectiveness of additional neighboring ideas in reducing complexity, and N is the number of neighboring ideas.

Together, intrinsic and extrinsic compression combine to produce the total compression factor:

$$\gamma(a, N) = \gamma_{\text{intrinsic}}(a) \cdot \gamma_{\text{extrinsic}}(a, N) = \eta e^{-\lambda N}.$$

This compression reduces the effective assembly index of an idea:

$$a_{\text{eff}} = \gamma(a, N) \cdot a = \eta e^{-\lambda N} a.$$

In cognitive systems, the temporal cost of assembling ideas grows exponentially with complexity. Compression mitigates these costs by reducing the effective assembly index. The discovery time (τ_d) and production time (τ_p) for an idea are given by:

$$\tau_d(a) = \tau_{d0} e^{\beta_d a_{\text{eff}}} = \tau_{d0} e^{\beta_d \eta e^{-\lambda N} a} \text{ and } \tau_p(a) = \tau_{p0} e^{\beta_p a_{\text{eff}}} = \tau_{p0} e^{\beta_p \eta e^{-\lambda N} a}.$$

Here, τ_{d0} and τ_{p0} are base timescales, and β_d and β_p reflect how costs scale with complexity. As N increases, $e^{-\lambda N}$ decreases, reducing τ_d and τ_p. Without sufficient neighboring ideas, compression is inadequate, and temporal costs become prohibitive.

The growth of cognitive assemblies depends directly on these reduced temporal costs. The rates of discovery and production are:

$$\frac{dN_{a+1}}{dt} = \frac{k_d}{\tau_d(a)} \cdot (N_a)^\alpha = \frac{k_d}{\tau_{d0} e^{\beta_d \eta e^{-\lambda N} a}} \cdot (N_a)^\alpha,$$

$$\frac{dN_a}{dt} = \frac{k_p}{\tau_p(a)} \cdot N_a = \frac{k_p}{\tau_{p0} e^{\beta_p \eta e^{-\lambda N} a}} \cdot N_a,$$

where k_d and k_p are rates of discovery and production, and α reflects selection pressure.

We can formalize the relationship between neighboring ideas and the feasibility of idea assembly. We refer to this as the Contextual Density Threshold, which states that there exists a critical number of neighboring ideas N_c such that for all $N \geq N_c$, the discovery time $\tau_d(a)$ remains below a maximum allowable time $\tau_{\max}$. From the inequality for discovery time,

$$\tau_d(a) = \tau_{d0} \exp\left(\beta_d \eta e^{-\lambda N} a\right) \leq \tau_{\max}.$$

We define N_c as the critical threshold,

$$N_c = \frac{1}{\lambda} \ln\left(\frac{\beta_d \eta a}{\ln\left(\frac{\tau_{\max}}{\tau_{d0}}\right)}\right),$$

which establishes that once the cognitive network reaches N_c, the discovery time $\tau_d(a)$ becomes feasible for assembling complex ideas.

The Contextual Density Threshold mirrors the result found by Sharma et al., which they use to explain the impossibility of randomly assembling chemicals with high complexity. Just as chemical compounds of sufficient complexity require selection mechanisms, ideas must self-replicate and evolve within cognitive systems under selection mechanisms. Compression and contextual density enable these processes, fostering the emergence of cognitive complexity.

III Assembling Culture

To validate the explanatory power of Extended Assembly Theory, we test it against a phase transition which traditional Assembly Theory fails to explain: culture. Culture represents a distinct evolutionary system characterized by its ability to transmit and refine ideas, behaviors, and practices across generations.

Culture as a Phase Transition

Culture exhibits the defining characteristics of a phase transition, its evolution resulting in rapid complexity emergence, novel selection pressures, and the creation of negentropic pockets. Moreover, as Boyd and Richerson argue, cultural evolution operates as an evolutionary system distinct from but intertwined with biological evolution [10]. This system relies not on genetic inheritance but on the transmission of learned behaviors and shared practices, allowing for far greater adaptability and speed. Similarly, Dawkins introduces the concept of memes as replicators akin to genes, suggesting that cultural transmission mirrors the mechanisms of biological inheritance [11]. Memes, like genes, are subject to variation, selection, and replication, driving the evolution of cultural systems, much as the units in other phase transitions.

The emergence of culture represents an explosion of complexity, catalyzed by the development of symbolic thought and language. Language, in particular, enabled unprecedented levels of cooperation and social organization, forming the foundation for exponential population growth and global civilizations. As Tomasello observes, language serves as a medium for transmitting cultural knowledge, bridging individual cognition to collective problem-solving and technological innovation [12]. This cultural complexity manifests through cumulative evolution, where innovations build upon prior discoveries. Boyd and Richerson (1985) highlight how the ability to inherit cultural knowledge across generations allows humans to innovate collectively, surpassing the limits of individual cognition. Moreover, the novelty generation resulting from the sustainment of innovation creates a feedback loop analogous to ones found in other phase transitions which cause dynamic meta-stability (e.g. autocatalytic sets in genes).

Cultural evolution introduces novel selection pressures distinct from those of biological evolution. Henrich and Boyd argue that these traits such as cooperation, reputation, and technological proficiency enhance group survival, creating societies that are resilient and adaptive in the face of complex challenges [13]. For example, moral systems promote trust and cooperation, enabling large-scale societal structures, while technological advancement fosters positive feedback loops to generate novelty. Unlike the slower process of genetic evolution, cultural evolution is rapid, allowing societies to adapt quickly to changing environments.

Cultural systems create negentropic pockets that resist change and sustain the transmission of knowledge across generations. These pockets stabilize collective knowledge, enabling the efficient organization and preservation of information essential for cultural and technological progress. Henrich emphasizes the importance of such systems in fostering cumulative cultural evolution, where each generation builds on the achievements of its predecessors [14]. This creates a feedback loop that drives exponential growth in cultural complexity. As with other phase transitions, the stability provided by these negentropic pockets fosters complexity explosions, ultimately leading to the next phase transition.

The Evolution of Cultural Complexity

Cultural evolution depends on abstract units whose meanings are contingent upon the basis of cognition and social systems interpreting them. Memes, unlike molecules, do not exist as static entities; their value is contextual, evolving through interactions with the systems that host them. The inability of AT to incorporate these dependencies renders it inadequate for explaining cultural evolution. Without modeling the interplay between memes and the selective pressures and interpretations of cultural systems, it cannot capture the mechanisms driving cultural change or the behaviors resulting from these pressures.

The evolution of culture hinges on the efficient formation, storage, and transmission of ideas, achieved through intrinsic and extrinsic compression. These mechanisms reconcile the abstract nature of memes with their roles in cognitive and social contexts.

- *Intrinsic Compression*: Ideas are simplified into compact, self-contained forms, such as language, gestures, and symbols, minimizing redundancy and enhancing efficiency.

- *Extrinsic Compression*: Ideas are integrated into existing frameworks, such as myths, stories, and traditions, enabling coherence with broader cultural narratives. This form of compression allows memes to reference and depend on other ideas, embedding them within a larger cognitive and social basis.

Cultural evolution is inherently bound by time. Memes and the systems that support them exist within finite temporal spans, creating selective pressures that favor innovations reducing temporal costs. Extrinsically compressed memes, for example, benefit from the longevity of their associated frameworks. Cultural innovations that decrease temporal costs, such as the development of language or the printing press, are naturally favored, as they enable faster propagation and broader adoption. The varying rates at which cultural ideas emerge and spread are thus explained by their impact on temporal efficiency. Ideas that streamline communication and action are selected for and propagate more quickly, shaping the trajectory of cultural evolution.

Cultural evolution depends not only on the efficiency of individual memes but also on the interconnectedness of the systems in which they exist. Each generation's inevitable end necessitates the efficient transfer of cultural ideas, prioritizing compression to minimize loss. This continuous compression indirectly selects for increased social connectivity on the community level and thus evolutionarily on the individual level. By referencing a broader network of external memes in highly interconnected societies, compressed memes reduce the cognitive and social load required to sustain them. This explains the historical and contemporary preference for highly connected social networks, especially in the digital age. Connectivity enhances the efficiency of meme propagation, enabling cultural systems to grow increasingly complex and resilient.

As lossless compression approaches its asymptotic limits, selection shifts focus from refining the compression of the memes to acting directly on the memes themselves. The Contextual Density Threshold describes this transition, where the fitness of cultural units determines their persistence and propagation. Memes that fail to meet these selective criteria are eliminated, while advantageous ideas proliferate. This dynamic is evident in the historical transition from religious to scientific dominance. For millennia, religious paradigms compressed philosophical and metaphysical ideas, preserving them and fostering cultural stability. This stability enabled an explosion of complexity resulting in the discovery of natural laws, culminating in breakthroughs which produced advantageous scientific ideas. Once these memes surpassed the contextual density threshold, new selection pressures favored their dissemination. As scientific memes grew in volume and prominence, religious ideas were selected against due to the asymptotic limit of lossless compression being reached, shifting selection to the underlying memes themselves and reducing their generational transmission under temporal constraints.

IV Consciousness & Amorphous Forms

To illustrate the concept of phase embeddings introduced in Section 2, we will apply it to both rhetorical analysis and predictive arguments, mirroring the validation of physical theories by testing against existing frameworks (rhetoric) and forecasting future outcomes.

Characteristics of Consciousness

The notion of consciousness, despite myriad definitions across numerous fields, generally shares three central characteristics: self-awareness, intentionality, and embodiment. Self-awareness provides a subjective lens through which an entity perceives its own existence; intentionality extends this perception outward; and embodiment assigns both self-awareness and intentionality to a physical substrate. Together, these elements create the unified experience we associate with being conscious.

Self-awareness is the ability to perceive oneself as a distinct entity with internal states. This recognition is not merely an abstract concept but is underpinned by biological mechanisms, particularly the default mode network (DMN). The DMN, a neural system active during introspective thought, integrates self-referential information, memories, and emotions [15]. This integration enables an entity to reflect on its own existence, forming the basis of the self. Philosophically, René Descartes encapsulated the essence of self-awareness with his assertion, "Cogito, ergo sum" (I think, therefore I am). This statement positions thought itself as undeniable proof of existence. Thomas Nagel builds on this idea with his concept of qualia, the "what it's like" aspect of experience. In his famous thought experiment, Nagel argues that the subjective experience of being – a bat, for instance – is fundamentally inaccessible to anyone outside that perspective [16]. Qualia, then, are central to the uniqueness of consciousness, resisting reduction to purely objective or third-person descriptions. Psychologically, the development of self-awareness is observable in children. Milestones such as recognizing oneself in a mirror or understanding that others have distinct thoughts and emotions (Lewis & Brooks-Gunn, 1979) show the gradual emergence of self-awareness.

Intentionality is the property of mental states being directed toward something – whether an object, idea, or goal. This capacity transforms self-awareness into purpose, imbuing actions and thoughts with direction and meaning. Neuroscientifically, intentionality is closely tied to the prefrontal cortex, the brain region responsible for planning, decision-making, and goal-directed behavior (Miller & Cohen, 2001). Damage to this area disrupts an individual's ability to form intentions, emphasizing its critical role in consciousness. Philosophically, Franz Brentano identified intentionality as the hallmark of mental acts, distinguishing conscious processes from unconscious ones [17]. Intentionality provides the structure for beliefs, desires, and goals, which collectively guide behavior. Cognitive models, such as those proposed by Baars, highlight how intentionality operates within a feedback loop, where perception shapes intention and intention reshapes perception [18]. Theological perspectives, such as those of Aquinas, extend intentionality to metaphysical realms, proposing that human intentionality reflects a higher order. While modern interpretations often secularize this view, the notion that intentionality arises from and contributes to a greater system remains central to understanding its adaptive value.

Embodiment integrates sensory input, emotional states, and motor outputs into a cohesive experience. Antonio Damasio's somatic marker hypothesis illustrates how bodily states influence decision-making and awareness [19]. Emotions, as somatic markers, guide choices by linking abstract concepts to physical sensations. For instance, the nervous tension before a risky decision reflects the body's role in evaluating outcomes, anchoring abstract thought in tangible experience. Philosophically, Maurice Merleau-Ponty emphasized the lived experience of the body, arguing that consciousness cannot be separated from its physical form [20]. The body is not an object to be observed but the medium through which the world is experienced. Psychological conditions such as anosognosia, where individuals lack awareness of bodily deficits, further underscore the interdependence of consciousness and embodiment.

Interpretation of Consciousness

Within the framework of embedded phase transitions, the traits that compose consciousness – self-awareness, intentionality, and embodiment – give rise to the notion:

Consciousness wishes itself upon its creations and its creations upon itself,
with little care for which direction yields illusion.

Phase transitions, by their nature, are additive processes. They do not subtract from the diversity or volume of units within the preceding phase; instead, they build upon them, creating novel configurations and expanding complexity. For example, the emergence of biological systems did not reduce the variety of chemical compounds but instead leveraged them to form intricate biochemical networks. The complexity of life arose from the chemical foundations established in earlier phases. This additive characteristic implies that phase transitions do not induce selection pressures that necessitate resistance from entities within a phase. In other words, the natural progression of transitions is one of expansion, not conflict. Entities within a phase have no intrinsic reason to resist transitions; instead, they adapt and incorporate the changes into their existence.

While previous phases transition without resistance from selection or assimilation, consciousness is unique in its relationship with phase transitions. Unlike earlier systems, consciousness inherently selects for transitions and seeks to assimilate them into itself. This distinction arises from the interplay of self-awareness, intentionality, and embodiment. Conscious entities possess an acute awareness of their own existence and the phases they inhabit and cause. Beyond awareness, consciousness directs itself toward specific outcomes. Intentionality transforms passive recognition of transitions into deliberate efforts to influence them, thus inducing a selection for transitions aligned with the selection for self-aware systems. Conscious beings, thus, must view transitions as extensions of their embodied experience. Because phases are embedded within the units of their predecessors, transitions are not external disruptions but natural evolutions of the systems consciousness inhabits. For instance, just as a biological entity sees molecular changes as intrinsic to its form, consciousness perceives phase transitions as integral to its identity. Together, these traits position consciousness as an active force in driving and assimilating phase transitions. Rather than passively undergoing change, consciousness seeks to integrate transitions into its framework, making them a part of its continued existence.

This unique characteristic of consciousness is evident in human behavior and societal trends, where numerous examples illustrate how consciousness seeks to embody and integrate emerging phases. One prominent example is the growing investment in brain-computer interfaces, driven by the anticipation of an impending "AI revolution." Industries are dedicating substantial resources to developing these technologies, reflecting a deeper instinct within human consciousness: the desire to integrate the next phase of technological advancement into our own experience. By creating interfaces that directly connect the mind to machines, humans are not merely adapting to new capabilities but actively seeking to assimilate this technological evolution into their embodied reality – to belong to and own the next phase transition.

Amorphous Forms

Because consciousness tends to assimilate transitions rather than allow distinct phases to emerge, generally only random, unintentional events might result in non-assimilated phases. Assimilation prevents the emergence of a distinct system capable of independent complexity evolution, effectively nullifying the transition as a true phase shift. However, there are exceptions to this rule.

Exceptions arise when a transition fundamentally violates one of the characteristics essential to the phase which generated it. Of the three – self-awareness, intentionality, and embodiment – characteristics defining consciousness, only embodiment can be disrupted in a way that prevents assimilation. Self-awareness and intentionality, by their nature, can be violated and assimilated into a conscious being just as contemporary technology is currently being integrated into humans [21], [22]. Embodiment, however, assigns consciousness a distinct physical substrate. Thus, violations to embodiment yield phases impossible to assimilate.

Non-constant morphologies, or amorphous forms, represent a class of embodiment violations characterized by fluid, transient structures that defy stability. These forms challenge consciousness by their very nature; they lack the consistent boundaries required for stable interaction or integration. A flawed yet illustrative analogy is that of air: while humans can sense and manipulate air, it resists direct assimilation into the body's form. Consciousness, similarly, could perceive amorphous forms but would lack the capacity to integrate them into its embodied framework. Such entities remain external, inaccessible as extensions of consciousness.

Non-material morphologies, or morphless forms, extend this violation further. These forms exist without any physical substrate, operating purely as informational or energetic entities. An example is digital information networks which, debatably are a phase transition produced by consciousness, yet unable to be assimilated into it. Consciousness, reliant on its embodied medium, would have no means of interacting with or assimilating such forms.

Successive phases can access and build upon their predecessors, integrating their structures and systems, but not vice-versa. For instance, genes evolve combinations of chemicals, and cognition reasons about genes, but the reverse does not hold: chemicals do not act on genes, and biology does not act on cognition to generate novelty. When a future phase transition results in forms that consciousness cannot assimilate – whether amorphous or morphless – consciousness would likely remain unaware of their existence. Even if conscious systems were involved in triggering such transitions, they would lack the capacity to engage with the resulting entities to generate novelty.

Conclusion

The emergence of cognition as a phase transition challenges existing frameworks like Assembly Theory, which, while effective in modeling chemical systems, cannot account for the dynamics of genetics, cognition, and culture – abstract units like ideas, mechanisms of compression, and temporal constraints. To address this, we introduced two theories: Phase Embeddings, which explain how successive phases build upon and interact with their predecessors, and Extended Assembly Theory, which leads to the Contextual Density Threshold Theorem to model cognition's evolution. Validating this framework through cultural evolution, we demonstrated how Extended Assembly Theory explains phenomena Assembly Theory cannot, such as the rapid complexity growth driven by memes, the role of interconnectivity, and the historical transition between paradigms of reason. Extending this, we defined consciousness through traits common to almost all interpretations and argued that phases with conscious entities tend to assimilate transitions rather than generating distinct phases, except when transitions violate embodiment, as seen in amorphous or morphless forms. Consciousness may thus face a horizon beyond which it cannot evolve or perceive. Situating ourselves within this chain of phase embeddings, we recognize our consciousness as assembled from complexity's origins, and we are compelled to question whether intent is indistinguishable from awareness and awareness indistinguishable from being. The lines between selection and telos blur not only within our conscious selves but also in future phases, which we may only ever be able to imagine.

References

[1] Erwin Schrödinger. *What is Life? The Physical Aspect of the Living Cell.* Cambridge University Press, 1944.

[2] Aakash Sharma, Daniel Czégel, Michael Lachmann, and Leroy Cronin. Assembly theory explains and quantifies selection and evolution. *Nature*, 622:321–328, 2023.

[3] Michael L. Wong, Carol E. Cleland, Daniel Jr. Arend, et al. On the roles of function and selection in evolving systems. *Proceedings of the National Academy of Sciences*, 120(43):e2310223120, 2023.

[4] Alessandro Filisetti, Alex Graudenzi, Roberto Serra, et al. A stochastic model of the emergence of autocatalytic cycles. *Journal of Systems Chemistry*, 2:2, 2011.

[5] Stuart A Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution.* Oxford University Press, 1993.

[6] Blaise Agüera y Arcas, Jyrki Alakuijala, James Evans, Ben Laurie, Alexander Mordvintsev, Eyvind Niklasson, Ettore Randazzo, and Luca Versari. Computational life: How well-formed, self-replicating programs emerge from simple interaction, 2024.

[7] John Von Neumann and Arthur W Burks. *Theory of Self-Reproducing Automata.* University of Illinois Press, 1966.

[8] Bruno A Olshausen and David J Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609, 1996.

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

[10] Robert Boyd and Peter J Richerson. *Culture and the Evolutionary Process.* University of Chicago Press, 1985.

[11] Richard Dawkins. *The Selfish Gene.* Oxford University Press, 1976.

[12] Michael Tomasello. *The Cultural Origins of Human Cognition.* Harvard University Press, 1999.

[13] Joseph Henrich and Robert Boyd. Culture and the evolutionary origins of human cooperation. *Nature*, 410:409–412, 2001.

[14] Joseph Henrich. *The Secret of Our Success: How Culture is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter.* Princeton University Press, 2016.

[15] Scott C. Johnson, Leslie C. Baxter, L. Sherman Wilder, James G. Pipe, Joseph E. Heiserman, and George P. Prigatano. Neural correlates of self-reflection. *Brain*, 125(Pt 8):1808–1814, August 2002.

[16] Thomas Nagel. What is it like to be a bat? *The Philosophical Review*, 83(4):435–450, 1974.

[17] Franz Brentano. *Psychology from an Empirical Standpoint.* Duncker & Humblot, 1874.

[18] Bernard J Baars. *A Cognitive Theory of Consciousness.* Cambridge University Press, 1988.

[19] Antonio Damasio. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness.* Harcourt Brace, 1999.

[20] Maurice Merleau-Ponty. *Phenomenology of Perception.* Gallimard, 1945.

[21] Amber Case. We are all cyborgs now. TED Talk, 2010.

[22] Sherry Turkle. *Alone Together: Why We Expect More from Technology and Less from Each Other.* Basic Books, New York, 2011.