DISSERTATION | DOCTORAL THESIS

Titel | Title

Über Leben
towards a Logic of Life

verfasst von | submitted by

Paul Poledna BA MA

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doktor der Philosophie (Dr.phil.)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 792 296
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Philosophie
Betreut von | Supervisor:	Dipl.-Phys. Dr. Andrea Loettgers

Abstract English

This thesis investigates the question of what distinguishes life from non-living matter, proposing that the answer lies not in a particular substance but in a specific form of processual organization. Moving beyond classical thermodynamics and self-organization, the argument is developed that life emerges when systems transition from dissipating energy gradients to actively maintaining their own organizational constraints over time.

The first part, Genesis, establishes a theoretical foundation by synthesizing concepts of process, relation, and constraint. It critiques purely mechanistic and vitalist views, arguing instead that life's unique properties, such as intrinsic purpose (*teleology*), arise from the emergence of self-referential, closed-loop dynamics. This culminates in the concept of *self-constraint*, where coupled *self-organizing* processes mutually constrain each other, enabling persistence far from equilibrium.

The second part, Paradox, explores the philosophical implications of this framework. It identifies a scale-invariant pattern, the *Constraint Cycle*—a recurring sequence of growth, disruption, and reorganization—that manifests from molecular systems to ecosystems. This dynamic is formalized as a processual logic termed *Cyclical Synthesis*, which explains how coherence emerges from the temporal coordination of opposing forces, resolving paradox through process rather than static resolution. This necessitates a shift from a *synchronic* to a *diachronic* scientific worldview.

The final part, Discontinuity, applies this framework to broader questions of scale and complexity. It distinguishes living systems from computational AI, which operates within a bounded *small world* with extrinsic goals. It further argues that levels of organization emerge as *constraint regimes*—stable domains defined by systematic exclusion—which are actively maintained through the energetic work of *Distributed Integration*. The thesis concludes by re-framing scientific modelling itself as an act of constraint selection, advocating for a practice of partiality that recognizes knowledge as a dynamic, embodied, and cyclical achievement.

Abstract Deutsch

Im Zentrum dieser Arbeit steht die Frage nach der fundamentalen Differenz zwischen Leben und unbelebter Materie. Die hier entwickelte These postuliert, dass diese Differenz nicht in einer bestimmten Substanz, sondern in einer spezifischen Form prozessualer Organisation zu finden ist. Leben, so die Argumentation, entsteht an jenem Punkt, an dem Systeme von der passiven Dissipation von Energiegradienten zu einer aktiven Aufrechterhaltung ihrer eigenen organisationalen constraints übergehen.

Der erste Teil, Genesis, legt diesen Gedanken theoretisch dar. Er entwickelt aus der Synthese von Prozess, Relation und constraint eine Alternative zu mechanistischen und vitalistischen Ansätzen. Die einzigartigen Eigenschaften des Lebens, wie etwa intrinsische Zweckgerichtetheit (Teleologie), werden als emergentes Resultat selbstreferenzieller Dynamiken verstanden. Dies mündet im Konzept des self-constraint, einem Zustand, in dem sich gekoppelte selbstorganisierende Prozesse gegenseitig stabilisieren und so ihr Fortbestehen fernab des Equilibrium sichern.

Der zweite Teil, Paradox, widmet sich den philosophischen Implikationen dieses Ansatzes. Es wird ein skaleninvariantes Muster herausgearbeitet, eine wiederkehrende Sequenz aus Wachstum, Störung und Reorganisation, das sich von der molekularen Ebene bis hin zu Ökosystemen zeigt. Diese Dynamik wird als eine prozessuale Logik formalisiert, die als Zyklische Synthese bezeichnet wird. Sie erklärt, wie Kohärenz aus der zeitlichen Koordination gegensätzlicher Kräfte erwächst und Paradoxien durch prozessuale Perspektiven aufhebt.

Der abschließende Teil, Diskontinuität, wendet sich der Frage von Skalierung und Komplexität zu. Zu Beginn wird die Differenz zwischen lebendigen Systemen und Computersystemen herausgearbeitet. Dann wird argumentiert, dass sich differenzierbare Organisationsebenen in lebendigen Systemen herausbilden, stabile Domänen, die durch systematische Exklusion definiert und durch die energetische Arbeit aktiv aufrechterhalten werden. Die Arbeit schließt mit einer Reinterpretation wis-

senschaftlicher Modellierung als einen Akt der Selektion. Plädiert wird für eine Praxis der Partialität, die Wissen nicht als statische Repräsentation, sondern als eine dynamische, verkörperte und zyklische Errungenschaft begreift.

Preface

Producing a longer piece of text is alienating. You start writing somewhere, move to the next part, and slowly—without noticing—those fresh thoughts begin to fade. The strange thing, is that you can go back and read them anew some time later. For me, this always creates an internal tension, between an older version of me, what *he* once said and thought, how thinking this has changed him, and how looking at these words is this curious way of meeting this other self. Not all of these encounters are pleasant. Why did he say it in this way? What exactly was his intention here? Interesting choice of words from him.

There are also the uplifting moments, where you come to appreciate certain passages as you read them. Or when you re-discover something. When a forgotten connections suddenly resurface. Then it is of course no longer *he* who had these ideas, but *I*.

This encounter with oneself across time deserves further exploration. But since this text was written as a PhD thesis on the question of what life is, I should explain why these personal reflections belong here.

How I want to frame the coming pages, is that they are in some way an exploration of exactly this phenomenon. While we will spend more time thinking about simple cells than the complexities of human cognition, a central question that we'll circle around is precisely this: how is it possible that life has this fascinating property of relating to itself over time.

If this exploration succeeds, you too might experience that alienation from your earlier self—the one who began reading. Therefore I close with Kripal's invitation: may the present *you* not survive this little book.

Contents

Introduction

Art, science, philosophy—all human understanding flows through living bodies and neural processes that are expressions of life. We are not observers standing outside reality, describing it from a neutral vantage point. We are participants, living systems attempting to comprehend living systems. Any philosophy that ignores this, that pretends to describe reality as if we weren't part of it, has already failed to reckon with its own conditions of possibility.

If our ways of knowing are themselves achievements of life, then understanding life better must illuminate the very nature of understanding itself. Hence, this book traces regularities in how living systems maintain themselves, patterns that remain recognizable whether we're examining molecules or thought processes.

The title *Über Leben* exploits an ambiguity that exists only in German. The phrase carries three distinct but inseparable meanings, each naming a dimension of what this book attempts.

First, and most straightforwardly, *Über Leben* means *about life*. This is an inquiry into what life is, how it works, what distinguishes living from non-living systems.

Second, *überleben*, written as one word, means *to survive*. We live in an age where the continuation of life—not so much human life as the intricate web of living systems—feels increasingly precarious. This is largely of our own making. We have decimated biodiversity and driven countless species to extinction, extinguished entire cultures, and continuously refined weapons of annihilation, from nuclear arsenals to the more recent developments in artificial systems whose risks we just begin to comprehend. Any philosophy adequate to this moment must contribute to our capacity to navigate our predicament. It must help us understand not just what life is, but how to participate wisely in its continuous expression.

Third, and perhaps most subtly: the German word *über* can mean both *about* and *beyond*. This double meaning allows a particular reading, one this book develops deliberately. *Über Leben*, in this interpretation, suggests something beyond life, a meta-perspective. But this is not transcendence in a mystical sense. Rather, it names something peculiar about life: its capacity to relate to itself across time. As hinted in the preface, living systems exhibit a strange property of temporal self-transcendence. The pattern persists even as every component is replaced. Life is always already beyond any particular instantiation of itself, carrying forward across time while remaining fully embedded in temporal flow.

The argument of this book unfolds in three phases. The first, *Genesis*, aims to trace how living organization emerges from non-living processes. It begins with basic thermodynamics and self-organization, showing how ordered patterns can arise spontaneously. However, living systems do something more, they develop the capacity to maintain their own organization over time. They transform from structures that simply arise and dissipate into systems that actively work to persist. This part develops the conceptual foundations necessary for everything that follows.

The second part *Paradox*, explores the temporal logic that living systems embody. Here we see how organisms achieve coherence not by eliminating contradictions but by coordinating opposing tendencies across time. What appears paradoxical when we look at a single moment reveals itself as integrated when we follow the actual flow of time. This requires developing new ways of thinking about logic, causation, and knowledge itself.

Discontinuity, applies these insights to larger-scale and composite systems. It examines how different levels of organization maintain their distinctiveness while remaining dynamically coupled, and explores what distinguishes genuinely living systems from sophisticated but non-living ones - including contemporary artificial intelligence. This part confronts the urgent practical question, namely, how do we foster the conditions for complex living systems to persist and transform wisely?

To traverse these domains, this text must negotiate a difficult borderland. Writing it has felt like living out the German idiom: *Zwei Seelen wohnen, ach! in meiner Brust*—two souls dwelling in my breast. One soul demands rigour and precision, the other seeks the sweep of meaning and the accessibility of the narrative.

Consequently, I must ask for a specific kind of patience from two different kinds of readers.

To the specialist—the biologist, the physicist, the systems theorist—you may find moments where I paint with too broad a brush. In service of the larger argument, I sometimes smooth over the jagged edges of biochemistry or thermodynamic nuances. I ask for your patience here. The goal is not to replace the textbook but to make the principles visible across scales.

To the interested generalist—the philosopher, the curious layperson—you may find yourself at times wading through the weeds of chemical self-organization or the technicalities of scaling laws. There will be moments where the density of the mechanism threatens to obscure the view. I ask for your perseverance. These details are not only technical garnish. They are the terrain upon which the argument is built. To understand the philosophy, we must occasionally dirty our hands with the machinery.

To keep the main path clear, I have moved many technical elaborations and divergent explorations into the endnotes. These are not just bibliographic references, but substantial extensions of the argument, a second layer of depth for those who wish to descend further. Yet, the tension that remains in the main text is intentional. I am convinced that we cannot understand life by remaining safely on one side of this divide. We cannot grasp it solely through the dry precision of the equation, nor solely through the abstract distance of metaphysics. Life itself exists in the tension between matter and form, between the nitty-gritty of energy exchange and the high-level maintenance of identity.

Therefore, this book attempts to mirror its subject. Just as a living system must integrate opposing processes to survive, this inquiry attempts to integrate the technical and the philosophical. It asks you to oscillate between the detailed and the horizon.

We have always been here, embedded in processes of self-maintenance and transformation, thinking with bodies that express the very logic we seek to understand. It is time we stopped treating this as an obstacle.

Part I

Genesis

Chapter 1

Process

> *No man ever steps in the same river twice, for it's not the same river and he's not the same man. - Heraclitus*[1]

Everything flows—we know this abstractly. But right now, wherever you are, parts of the world seem solid. The chair beneath you holds your weight. The walls around you stand firm.

Even so, everything *is* changing. The wooden chair—once a living tree, drawing water from soil, growing ring by ring—now ages and cracks, its fibers slowly breaking down. The concrete walls were mixed from crushed stone and water, hardened into place, now weather and fracture. Even the stones themselves are geological processes frozen in time: cooled magma, compressed sediments, metamorphosed rocks transformed by heat and pressure over deep time[2].

You are no exception. Your body replaces its cells constantly. Most of the atoms in your muscles, your bones, your brain weren't there a few months ago. You've literally rebuilt yourself from food you've eaten, water you've drunk, air you've breathed[3].

The ancient Greeks recognized this puzzle too. Cratylus, an early teacher of Plato, took Heraclitus's insight to its radical extreme. He argued that you cannot step into the same river even *once*, for the very act of stepping has duration, during which both you and the river change[4]. If reality is truly in constant flux, then language, which pins down and fixes meaning, is fundamentally inadequate. As a consequence, Cratylus eventually stopped speaking altogether, communicating instead only through finger-wagging.

Ironically, Cratylus's ideas on the impermanence of everything persist. Twenty-four centuries later, we still read, discuss, and transmit them. You are reading about them here and now. Black on white.

Somehow, despite everything flowing, despite the death of every person who ever heard Cratylus speak, the transformation of ancient Greek into current languages and the decay of papyrus scrolls, his ideas maintain a recognizable form and identity.

Here lies the paradox. Persistence emerges from flux itself. The very processes that seem to dissolve identity, like material turnover, environmental exchange and constant transformation somehow generate stable, recognizable patterns.

This is the central question we will be grappling with in this text. Namely, how can stable forms and identities arise that are both *embedded in* and *expressions of* ever-changing processes? Or, put more directly: What kind of organization allows a process to become so regular, so recurrent, so self-maintaining, that it achieves identity despite constant material flux?

Living processes actively construct and maintain themselves in ways that generate persistence, bringing forth patterns that remain recognizable across time even as their material constituents continuously change. Understanding precisely this organizational achievement, helps distinguish living systems from non-living ones.

Cratylus stopped speaking to avoid the contradiction between static words and flowing reality. Perhaps his silence was unnecessary. The persistence of his ideas demonstrates exactly what he sought to argue against, that stable patterns can maintain identity despite flux. Understanding *how* this organizational achievement occurs is precisely what we seek in developing a theory of living systems.

1.1 Finding Our Starting Point

There are many ways to address the question *What is Life*?—reflecting not just different scientific disciplines, but fundamentally different ways of framing what the question itself asks. This text follows a particular path, one that begins in physics and the study of thermodynamics before building toward an understanding of biological organization. This choice is motivated by the fact that ordered processes don't appear exclusively in biology. If we can understand how order emerges in simple physical systems, we may trace a continuous path to the more sophisticated organization of life.

Before we begin, though, it's worth sketching the rough outline where the field of origins-of-life research currently stands, and how this investigation positions itself within that landscape. What follows is not a comprehensive overview but a brief orientation, just enough context to understand why we're starting where we are.

For centuries, the question of life's origin was entangled with the idea of *spontaneous generation*, the belief that living organisms could arise directly from non-living matter. Maggots from rotting meat, mice from grain, little humans from humans, life seemed to emerge spontaneously all around us. It was Louis Pasteur who disproved this notion[5]. By showing that sterilized broth remained sterile in swan-neck flasks that allowed air but not dust-borne microbes to enter, he demonstrated that even the smallest organisms don't arise spontaneously. Life comes only from pre-existing life. His work foregrounded the question: if life only comes from life, how did it first begin?

The scientific investigation of life's origins began with what seemed like the natural question to answer, namely, what is life made of? The intellectual lineage traces back to Charles Darwin himself, who in 1871 mused in a letter to his friend Joseph Hooker about life beginning in a *warm little pond* where a protein compound was chemically formed, ready to undergo complex changes[6]. This assumption, that the emergence of the right chemical substances is the principal explanatory hurdle, proved remarkably persistent. By the 1920s, Russian biochemist Aleksandr Oparin and British scientist J.B.S. Haldane had independently developed this intuition into a hypothesis[7]. They proposed that Earth's early atmosphere was chemically reducing (rich in hydrogen, methane, and ammonia), that energy from ultraviolet light and lightning drove the synthesis of organic molecules, and that these accumulated in the oceans to form what Haldane termed a *hot dilute soup*.

For three decades, the Oparin-Haldane hypothesis remained speculative. Then, in 1952, it was dramatically supported. Stanley Miller, a graduate student working at the University of Chicago, designed an apparatus to simulate the proposed conditions of early Earth[8]. He sealed water, methane, ammonia, and hydrogen in glass flasks, subjected the mixture to electrical sparks to simulate lightning, and let the system run. Within days, the clear water had turned turbid and brownish-red. Chemical analysis showed that amino acids, the building blocks of proteins, which in turn are the building blocks of cells, had formed spontaneously from simple inorganic precursors.

The Miller-Urey experiment was a watershed moment. It showed that the fundamental components of life could arise through purely natural chemical processes. However, experiment imposed a conceptual constraint. By demonstrating that the right molecules could form spontaneously, it seemed to validate the underlying question—*What is life made of?*—as the right one to ask. The field focused intensely on prebiotic synthesis, making amino acids, nucleotides, lipids. Yet this success obscured a deeper problem. Having the building blocks doesn't explain how they organize into systems capable of self-maintenance, growth, and reproduction. The questions we ask shape the answers we find, and by asking primarily about chemistry and composition, we naturally gravitated toward chemical and molecular explanations while deferring the question of organization.

More than seven decades after Miller's experiment, we still have not created life in the laboratory. Despite sophisticated techniques and our ability to synthesize many organic molecules, despite our detailed knowledge of cellular biochemistry, no experiment has produced a system that grows, maintains itself, and reproduces autonomously from non-living precursors. The gap between a flask of amino acids and even the simplest living cell remains vast. This persistent failure is instructive. It suggests that identifying life's chemical components, however necessary, was never sufficient. The building blocks can be sitting right there in solution, and still nothing happens. The chemical approach had reached a boundary, not of experimental technique, but of conceptual framing. The question itself—*What is life made of?*—began to seem incomplete.

If the right molecules weren't sufficient for life, what was? Researchers began asking not what life is made of, but what minimal property or organizing principle could, once present, bootstrap everything else. This led to a schism in the field, splitting researchers into two major camps, each championing different essential features as the key to life's emergence.

The information-first hypothesis, most prominently the *RNA World* theory, proposes that life began with molecules capable of storing and replicating information. DNA stores genetic information but cannot catalyze reactions, while proteins catalyze reactions but cannot store genetic information. You need DNA to make proteins, but you need proteins to replicate DNA. Which came first?

The RNA World offers an elegant solution. RNA, unlike DNA, can do both jobs, it can store information and catalyze chemical reactions[9]. This provides a pathway to Darwinian evolution. A self-replicating RNA molecule could make copies of itself with occasional errors. These variant molecules could compete, with selection favouring those that replicated more efficiently or were more stable. Complexity could accumulate through this process, eventually leading to the sophisticated biochemistry we see today.

The appeal of this idea is straightforward. RNA molecules exist in every living cell today—they're closely related to DNA, which stores our genetic information, but with crucial differences that make them more versatile. The RNA World proposes that early Earth could have had pools or films of simple chemicals where, by chance, these RNA-like molecules formed. Once formed, they could make copies of themselves without needing the complex cellular machinery that life uses today. These self-copying molecules would multiply, compete for resources, and evolve, giving us a clear path from chemistry to biology, from lifeless molecules to the first primitive organisms.

Yet despite its elegance, the RNA World hypothesis faces severe challenges. RNA is extraordinarily complex and difficult to make under realistic early-Earth conditions. It's also chemically fragile, especially in water. And even if self-copying RNA molecules did form, they face serious technical problems that would prevent them from continuing to replicate[10]. These aren't minor technical hurdles but fundamental challenges to the plausibility of RNA as life's starting point. The very complexity that makes RNA so powerful in modern cells makes it an unlikely candidate for life's origin.

The second big approach is the metabolism-first hypothesis. It proposes that life began not with a genetic molecule but with self-sustaining networks of chemical reactions. In this view, information and RNA are later evolutionary inventions that emerged from an already-functioning metabolic system.

The basic idea is that certain environments like mineral surfaces near deep-sea hydrothermal vents, could support networks of chemical reactions that sustain and amplify themselves. These environments are rich in chemical reactants, and the surfaces of certain minerals can concentrate organic molecules and speed up reactions. The chemical energy released from the minerals themselves could power these reaction networks, creating cycles where some products help the reactions continue and grow[11].

Yet, metabolism-first faces its own critical challenge. How does a diffuse network of chemical reactions invent discrete genetic information? The transition from collective chemical dynamics to individual, heritable instructions remains largely unexplained. Without genetic control, how would such networks maintain stability and resist degradation? The information problem that RNA World solves by starting with replicators becomes the Achilles' heel of metabolism-first approaches[12].

The two paradigms, information-first and metabolism-first, have a symmetrical problem. Each can explain what the other cannot, and neither can explain what the other can. RNA World can't explain where its complex hardware came from. Metabolism-first can't explain where heritable information arose from. They're trapped in a chicken-and-egg loop because, in living systems, information and metabolism are mutually dependent features of the same integrated organization.

Nonetheless, this investigation takes its starting point closer from the metabolic side of this divide, but not because metabolism-first approaches have won the debate, they haven't. Neither framework has fully resolved the origin of life, and both face difficulties. The choice to begin with metabolism is not a claim that this approach is correct while information-first is wrong. Rather, it's a pragmatic judgment about where we can build most productively from established physical principles. We begin with metabolism because:

First, we can build from physics. Thermodynamics and self-organization give us a natural starting point. We can understand how ordered processes emerge from energy gradients without invoking anything beyond established physical principles. This provides a solid foundation that doesn't require us to assume complex molecules from the outset.

Second, we can tell a continuous story. Starting from geochemical energy flows, we can trace a plausible pathway through increasingly sophisticated forms of chemical organization. Each step builds naturally on what came before, without requiring sudden leaps to improbably complex structures.

Third, we can still reach information. The metabolic approach doesn't reject or contradict the insights of RNA World research. Instead, it provides a pathway to it. We can show how simpler information-bearing structures might have emerged from metabolic networks, eventually leading to RNA and the sophisticated genetic systems we see today. We're not opposing information-first approaches, we're constructing the framework that makes their starting point plausible.

The following chapters will develop this framework systematically, beginning where it must begin: with the physical principles that govern all processes, living or not.

1.2 Physics & Biology

This focus on physical principles brings us inevitably to the work of Erwin Schrödinger. In 1944, turning aside from his work in quantum mechanics, he sought to address a central problem "which hovers between biology and physics"[13]. The question he posed in his lectures was straightforward: "How can the events in space and time that take place within the spatial boundary of a living organism be accounted for by physics and chemistry?"

From the way the question is posed, it is clear that physics and chemistry, along with their fundamental framework of space and time, will guide Schrödinger's investigation into living processes. This was a time when many biologists were influenced by vitalism — a philosophical doctrine that attributed life's distinctive properties to something beyond physical and chemical laws. Henri Bergson, particularly influential in this tradition, had developed the concept of élan vital: a vital impetus or creative force that, he argued, permeated all living things and explained their unique capacities for growth, development, and evolution[14]. Rather than engage with such metaphysical explanations, Schrödinger grounds his investigation in the physical and chemical principles understood in the 1940s. This methodological choice — to explain living systems using only the tools of physics and chemistry available at the time — sets the foundation for our investigation as well. Throughout this text, we will similarly attempt to understand biological phenomena through the established principles of physical science.

But here's where things get interesting. Ask any physicist about describing biological processes, and they'll likely shake their head at the sheer magnitude of the challenge. First, there's the second law of thermodynamics—a fundamental principle that organisms seem to defy at first glance[15]. While the universe tends toward chaos, life stubbornly maintains order, constantly exchanging energy and matter with its environment.

Second, consider the challenge of scale and emergence in biological systems. Even a single cell contains billions of particles interacting simultaneously. While physics can precisely describe simple systems[16] - each added particle multiplies the complexity of these interactions. But the challenge goes beyond just quantity: biological systems exhibit emergent properties - behaviours and characteristics that arise from the collective interactions of their components. These emergent properties, often follow patterns distinct from the simpler physical laws governing individual molecules. This represents a fundamental challenge for physics, namely, how to bridge the gap between our understanding of individual molecular interactions and the complex, coordinated behaviours we observe in living systems.

Third, physics faces a unique challenge when confronting biological evolution: the role of *historical contingency*[17]. While physical systems are typically described by timeless, universal laws - the same equations that describe a falling apple today would have worked the same way millions of years ago - biological systems are fundamentally shaped by their history. Each organism carries within its structure and function the accumulated results of countless past events: mutations, selections, and environmental pressures that could have unfolded differently. Unlike physical systems that can be fully characterized by their present state and governing equations, living systems require understanding their evolutionary trajectory to explain their current features. This historical dependency - where past contingencies become integrated into biological structures and processes - represents a distinct type of complexity that our traditional physics frameworks aren't designed to capture.

These three challenges of thermodynamics, emergence, and historical contingency, represent fundamental obstacles physics faces in understanding life. While many other challenges exist (information processing, multi-scale dynamics, the measurement problem), they trace back to these core tensions.

For now, let's focus on what captivated Schrödinger himself: how organisms maintain their highly organized state despite thermodynamic principles that favour increasing disorder. These fundamental questions about the nature of life and its relationship to physical law require us to examine how physics approaches the problem of organization in nature. While biology describes the incredible complexity of living systems, physics provides the framework for understanding the basic principles that make such organization possible.

1.3 Entropy

> *If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations—then so much the worse for Maxwell's equations. If it is found to be contradicted by observation—well, these experimentalists do bungle things sometimes. But if your theory is found to be against the Second Law of Thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation.*
> *- Eddington*[18]

Picture yourself building a sandcastle on the beach. No matter how carefully you construct it, you know that it will eventually—whether through wind, waves, or simply time—crumble back into random piles of sand. While in any given moment, a few grains might temporarily stack themselves into more ordered arrangements, the overall trend towards disorder is undeniable. This statistical certainty—that systems will, on average, progress from ordered states to more disordered ones—captures the essence of what physicists call entropy and forms the backbone of the Second Law of Thermodynamics. The law's power lies not just in its universality, but in its statistical foundation: it emerges from countless molecular interactions, each following simple physical rules, to produce a principle so reliable that the brilliant astrophysicist Arthur Eddington considered it inviolable, as reflected in his half-humorous, wholly serious warning above.

While your sandcastle succumbs to chaos, the small crab scuttling nearby maintains and repairs its intricate shell structure against the constant assault of entropy. Even more remarkably, they pass on this capacity for ordered construction to future generations. How do living systems accomplish what sandcastles cannot—and why doesn't this violate the Second Law?

To better grasp entropy's fundamental nature, imagine dropping a single drop of ink into a clear glass of water. Watch as the ink spreads and diffuses, tendrils of color reaching through the liquid until the water becomes uniformly tinted. This spontaneous movement toward uniformity illustrates entropy's increase. The system progresses from a more ordered state (concentrated ink in one location) to a more disordered one (ink molecules scattered throughout the water). The process does not spontaneously reverse, you never see the dispersed ink gather itself back into a single drop.

The analogous example physicists often use involves a thermally insulated container divided into two compartments, one filled with hot gas, the other with cold. When the partition between them is removed, the gases naturally mix until they reach a uniform temperature throughout. At the molecular level, the hot gas atoms are moving rapidly with high kinetic energy, while the cold gas atoms move more slowly. Two things happen simultaneously, the fast-moving hot atoms, in their vigorous random motion, are statistically more likely to wander deep into the cold region than the slow-moving cold atoms are to wander into the hot region. And when atoms from different regions do collide, energy gets redistributed—fast moving atoms transfer some kinetic energy to slow ones. Through countless random walks and collisions, the system equilibrates. All atoms settle into moving at roughly the same average speed, which we perceive as uniform temperature.

This is fundamentally the same process as the ink diffusing in water. In both cases, particles are in constant random motion—whether gas atoms bouncing around or molecules jiggling in liquid. The ink molecules, initially concentrated in one drop, simply wander through their random walks into the surrounding water, while water molecules simultaneously wander into the ink-rich region.

This tendency toward increased entropy extends beyond simple gas or liquid mixtures to encompass all natural processes. While entropy can be defined technically in various ways[19], it is most commonly understood as a measure of the number of possible microscopic arrangements a system can take - often colloquially described as *disorder*. The second law of thermodynamics asserts that in isolated systems, this measure of disorder cannot decrease and generally increases, especially during equilibration processes. In 1865, Rudolf Clausius articulated this principle with sweeping grandeur: the entropy of the universe tends toward a maximum[20]. This serves as a kind of cosmic arrow, pointing toward the natural direction of spontaneous change—the universe inexorably moving toward states of increasing disorder.

This is why living systems seem so remarkable in their apparent defiance of this universal tendency. Your body maintains its temperature despite the surrounding air being cooler. Your cells organize themselves into complex patterns rather than dissolving into random arrangements. In fact, if we were to count all the possible ways your body's molecules could be arranged, the overwhelming majority would result in nothing resembling life—yet here you are, reading these words, your trillions of cells working in concert to maintain order.

The resolution to this apparent contradiction lies in understanding what Schrödinger would later explore in his Dublin lectures: living systems are not isolated. The systems we've discussed so far, inevitably evolve toward equilibrium, a static state of maximum entropy where nothing interesting happens anymore. Living organisms are fundamentally different. They are open systems, constantly exchanging matter and energy with their environment by breathing in oxygen, consuming food, radiating heat, excreting waste. This continuous flow changes everything, allowing life to maintain itself far from equilibrium, in states that would be impossible for a closed system.

But exchanging energy alone isn't enough, what also matters is the quality of that energy. This is where energy gradients become essential. An *energy gradient* is simply a difference in energy quality or concentration between two regions. Consider the Sun and Earth, the Sun radiates intense, concentrated energy at about 5,800 Kelvin, while Earth sits in the cold of space at around 300 Kelvin. This enormous temperature difference is called an energy gradient. Living systems have evolved to intercept and exploit such gradients. By surfing these energy gradients, life can perform work, building complex structures, maintaining order, moving, and growing.

The implications of this reshape our understanding of life's relationship with entropy. Consider a growing plant, it captures high-quality energy from sunlight and uses it to build complex molecular structures from simpler components, seemingly creating order from chaos. But this process generates much more entropy in the environment, through heat and waste products, than the decrease in entropy represented by the plant's organized structure. Life, rather than defying the second law of thermodynamics, has evolved to harness it. Living systems have developed sophisticated mechanisms to extract useful work from energy gradients. Photosynthesis captures the steep energy gradient between sunlight and chemical bonds, while cellular respiration exploits the gradient between energy-rich nutrients and their breakdown products. They achieve this through the systematic exploitation of entropy-driven processes. Creating order locally, while necessarily increasing disorder in their surroundings[21]. This is not just an accidental feature of life but essential, the capacity to harness entropy production to maintain and reproduce organized structures.

1.4 Self-organization

Hexagons are the Bestagons.
- CGP Grey[22]

We've explored how the Second Law of Thermodynamics appears to clash with the existence of life itself—how can systems maintain such high levels of order in a universe that inexorably tends toward disorder? The resolution, as we've seen, lies in understanding living systems as open systems, constantly exchanging energy and matter with their environment. But this is only part of the story. What makes life truly remarkable is not just that it exists away from equilibrium, but that it maintains itself far from equilibrium—in a dynamic state where continuous flows of energy and matter generate new principles of organization[23].

This distinction between being merely out of equilibrium and far from equilibrium is crucial. Many systems temporarily deviate from equilibrium, like a cup of hot coffee cooling on your desk. But living systems are different: they actively maintain themselves in far-from-equilibrium states through constant energy flow and matter exchange. For living systems, equilibrium is not just undesirable—it is fatal. Life exists, and can only exist, in this far-from-equilibrium regime where self-organization becomes possible[24].

First, let's take a step back to fully appreciate the significance of open systems and their role in self-organization. In our universe, we cannot find a single truly closed system—they exist only as theoretical constructs. A closed system is one that does not exchange any matter with its surroundings and is perfectly insulated, meaning no energy in the form of heat or work, nor any materials, crosses its boundaries. The closed system is an idealized model. Like many models, it is a very helpful tool for thinking, describing, and predicting the world. However, we must remember that a model is a simplification of reality, not reality itself, much like a map is useful for navigation but is not the actual terrain. Therefore, the closed system is an idealized concept suggesting a scenario where nothing external influences the system.

The closed system is an idealized model - helpful for thinking and predicting, but not found in nature[25]. In reality, all systems we encounter are open systems that continuously exchange energy and matter with their surroundings. This fundamental property shapes the dynamics of systems at all scales, from cells to planets, through the presence of gradients.

Gradients are foundational to understanding the dynamic interactions within open systems. By their very nature, gradients represent differences—in energy, temperature, pressure, density, and other measures—across space. These differences drive the flow and transformation of matter and energy, leading to changes and exchanges between parts of a system and its environment.

In the realm of thermodynamics, and especially in the context of the Second Law, these gradients are viewed as transient states. The law suggests that in a closed system, any present gradient will naturally diminish over time. For example, heating one end of a metal rod introduces a temperature gradient; one end becomes hotter than the other. Once the heating stops, the heat will distribute evenly throughout the rod, eliminating the gradient or heat difference, as the system moves towards equilibrium.

Earth as an open system under solar irradiation shows us how long-term energy gradients can drive complexity. While not permanent on cosmic timescales, the solar gradient has persisted for rougly 4.6 billion years providing a source of free energy[26] over enough time to enable the emergence and maintenance of complex ordered structures across many scales. From the evolution of increasingly complex minerals over geological time, to the development of atmospheric and oceanic circulation patterns, to the emergence and diversification of life itself - all these represent different ways that open systems on Earth harness this steady energy gradient. These systems establish their own local gradients from the free energy of the sun, building and maintaining increasingly sophisticated forms of order. The presence of this reliable energy gradient is key not just to creating temporary patterns, but to enabling the progressive development of complex structures that can persist and evolve over time. These examples demonstrate how Earth taps into solar energy gradients to sustain ordered processes. But if everything ultimately tends to equilibrium, why isn't the universe itself already smoothed out to a featureless uniformity?

The universe did not begin as a perfectly uniform expanse of matter and energy. Tiny quantum fluctuations in the primordial plasma, magnified by a rapid expansion phase called cosmic inflation, left some regions slightly denser than others. Over time, gravity acted like a feedback loop, amplifying these imbalances: the denser regions attracted still more matter, eventually forming stars, galaxies, and clusters, while emptier patches grew ever more rarefied. Far from disappearing, these large-scale gradients remain evident in the uneven distribution of galaxies, the cosmic microwave background's minute temperature variations, and the abundance of heavy elements in different stellar environments. In short, the initial conditions and the laws of gravity ensured that the universe was born with an intricate patchwork of high- and low-density regions rather than becoming perfectly smooth[27].

Yet these structures have persisted for billions of years rather than blending into a uniform cosmic soup, which brings us to the question of what keeps them from dissipating entirely. Even if gravitational clumping creates density contrasts, one might ask why they have not *smoothed out* completely over billions of years. The answer lies in the finite speed at which energy and matter can move, along with various kinds of resistance to flow: radiation must travel at or below the speed of light, gas and dust encounter friction, and cosmic expansion itself stretches space faster than gradients can fully equalize on large scales. At smaller scales, like within planetary or stellar systems, processes such as turbulence, convection, or chemical interactions take time to reduce differences in temperature, density, or chemical potential. Rather than contradicting the second law of thermodynamics, these entrenched gradients drive it forward: hot stars radiate into cold space, and cosmic structures continue converting free energy into entropy. Thus, *resistance to flow* ensures that lumps remain a fundamental feature of the universe, providing the energy gradients on which everything from atmospheric convection to biological life depends.

In this way, the same processes that shape galaxies and stars also furnish the free energy gradients upon which life on Earth depends. From cosmic scales down to cellular processes, the universe's inherent *unsmoothness* fuels self-organization and the relentless increase of total entropy.

To understand how these abstract principles manifest, let's examine a simple example that demonstrates how gradient-driven self-organization creates order. This example will show us how nature spontaneously generates patterns when dissipating energy - patterns that appear across different scales and systems.

Consider a cooking pan placed on a heat source. When you first turn on the heat, the liquid at the bottom of the pan starts to warm up. At this stage, heat moves through the liquid primarily through molecular diffusion - essentially, each molecule bumps into its neighbors, passing along a bit of its energy. This is a relatively slow and inefficient process.

Something remarkable happens when the temperature difference between the bottom and top of the liquid reaches a critical threshold (physicists characterize this threshold using a quantity called the Rayleigh number). At this point, the system discovers a more efficient way to transport heat: rather than relying on molecule-by-molecule diffusion, the liquid spontaneously organizes itself into large-scale flowing patterns called convection cells. This is depicted in figure 1.1

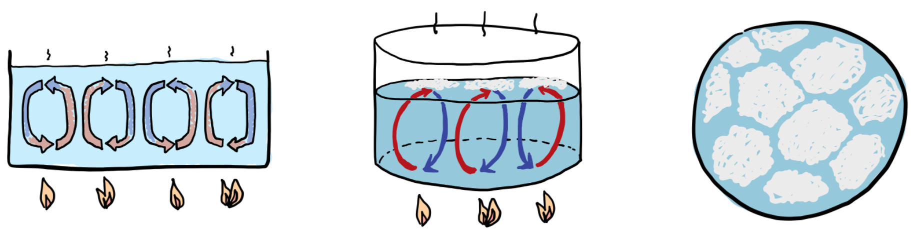

Figure 1.1: Visualizing Rayleigh-Bénard convection through changing perspectives. The sequence moves from a cross-sectional side view of the vertical circulation loops (left), rotates to a three-dimensional perspective of the container (center), and concludes with a top-down view (right) that reveals the emergent hexagonal geometry created by the packed convection cells.

This transition from diffusion to convection is called a phase transition - similar to how water suddenly changes from liquid to vapor at its boiling point. The random jostling of molecules gives way to coordinated motion, where warm fluid rises in some regions while cool fluid descends in others, creating a regular pattern of circulation. Interestingly, this ordered structure emerges spontaneously as it is better at dissipating the temperature gradient more quickly than the previous mode of molecular diffusion. The system has found a more efficient solution to the problem of moving heat from bottom to top.

In the process of dissipating the heat gradient, the system organizes itself into regular, far-from-equilibrium structures. These structures, known as Rayleigh-Bénard convection cells, can be observed in various systems, such as liquids, clouds, and cooling magma. The central picture in the lower row in figure 1.2, a 10-second long exposure shot, clearly highlights the beautiful hexagonal geometry of these convection cells.

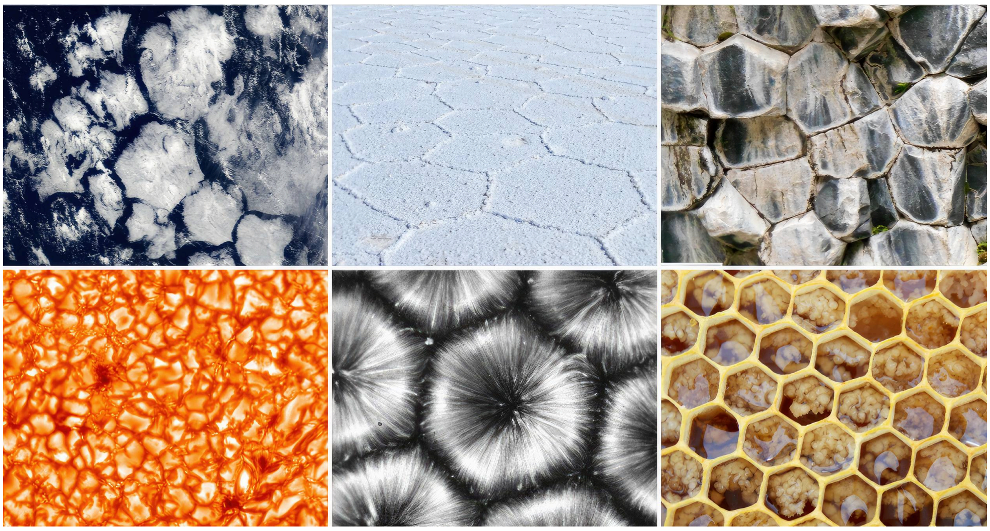

Figure 1.2: Examples of convection patterns in nature: (1) Rayleigh-Bénard convection cells in clouds over the South Atlantic Ocean, (2) salt flat convection patterns in Bolivia's Salar de Uyuni, (3) hexagonal basalt columns in Iceland's Vatnajökull National Park, (4) solar granulation showing convection cells on the Sun's surface, (5) laboratory demonstration of Rayleigh-Bénard convection cells with aluminum tracer particles, and (6) honeycomb structure built by bees[28].

Admittedly, I've introduced a slight deviation. Among the examples listed, hexagonal honeycomb structures do not result from convection cells. Nonetheless, it's intriguing to note that, similar to the convection cell examples, they also exhibit a hexagonal pattern. Why is this so?

The answer lies in geometry. When trying to divide a surface into cells of equal area (imagine tiling a floor), only three regular shapes (all sides and angles equal)can do the job without leaving gaps: equilateral triangles, squares, and hexagons. The honeycomb conjecture, proven by mathematician Thomas Hales in 1999, roughly states that if you want to divide a surface into regions of equal area while using the least total perimeter, hexagons are the best choice[29].

Think of it this way: if you're a bee building a honeycomb, you want to create cells that can hold the most honey while using the least amount of wax for the walls. Hexagons provide the optimal solution - they enclose the same area as triangles or squares would, but with less total perimeter[30].

This geometric efficiency explains why convection cells often form hexagonal patterns. When a fluid system organizes itself to transport heat most efficiently, it needs to create cells of equal size (because the fluid is uniform) while minimizing the boundaries between up-flowing and down-flowing regions (where energy is lost to friction). Hexagons solve this problem perfectly: they provide equal-sized cells with the least total boundary length, allowing the system to transport heat to the surface more efficiently than any other pattern.

It is worth appreciating what's happening here. The patterns that emerge do so as the system organizes itself to get rid of that energy gradient as quickly as possible. Or, viewed another way, the system organizes in a manner that accelerates the depletion of the energy gradient, creating as a *by-product* a regularized structure that helps optimize that process. The hexagonal arrangement is not random, but rather the most efficient solution found by the system to manage and dissipate energy. What makes this remarkable is that no blueprint or designer is required. Through countless molecular interactions, the system essentially discovers this optimal geometry on its own. The order we observe doesn't arise despite the Second Law of Thermodynamics, but rather because of it. The structure emerges because it accelerates the system's journey toward equilibrium. This is a fundamental principle of self-organization, complexity and pattern arising not from resistance to entropy, but from the most efficient pathways of its production.

Nature abounds with other examples of self-organization, and we find it everywhere—from crystal growth in rocks or snowflakes, to sand dunes and waves formed by wind patterns, to more exotic variations like the Belousov-Zhabotinsky Reaction or Liesegang Rings. Importantly, all these organizational patterns develop to deplete a given energy gradient quicker.

Let's go through one more example for the experimentalist at home. Let's assume you are on planet Earth and you have a gravitational gradient present. Now, take a full water bottle and think about how you could empty it as quickly as possible without squeezing the bottle. Since there is a gravitational gradient, you could simply turn the bottle upside down and wait for it to drain. Initially, a bit of water will flow out of the bottle, but soon air has to enter to replace the water that has exited, then more water flows out, to be replaced by air, and so on, creating the familiar gargling sound. Alternatively, you could spin the bottle to create a whirlpool or vortex inside the bottle, allowing a continuous air column to form in the center. This column will readily take up the space left by the exiting water. Timing both methods will show that the latter version empties the bottle much quicker, as depicted in figure 1.3.

Figure 1.3: Vortex in a bottle

Consider what happens when you open the lid and empty a bathtub or the sink, a whirlpool forms. Similar to the convection cells, the system organizes into a far-from-equilibrium state to deplete a gradient as quickly as possible, creating a highly regularized structure that optimizes this very depletion. So again, the optimization problem of how to get the water to flow down as quickly as possible is *solved* by organizing it into a regularized structure. Importantly, while the bottle vortex requires an initial energy input to form, the bathtub vortex emerges spontaneously—a reminder that self-organizing systems need the right conditions and sometimes a nudge to overcome initial barriers before settling into their optimal dissipative patterns[31].

This principle has been recognized and formalized by several influential thinkers, though it remains debated and is not yet codified as formally as the classical laws of thermodynamics[32]. Alfred Lotka, as early as the 1920s, proposed that biological systems evolve to maximize energy flow, establishing foundational ideas about how energy throughput shapes living organization. Building on this insight, Howard T. Odum later articulated what he called the *Maximum Power Principle*: living and systems evolve to maximize the power (energy throughput) they harness from their environment. Odum argued that systems which capture and process energy more effectively outcompete those that don't, leading to a kind of thermodynamic selection.

More recently, Schneider and Kay developed the Maximum Entropy Production Principle (MEPP), which argues that dissipative structures spontaneously organize themselves to maximize entropy production. This principle crystallizes the apparent paradox: systems that seem to create order are actually accelerating disorder in their surroundings, and this acceleration of disorder is precisely what drives their organization[33]. These various formulations, sometimes collectively referred to as proposals for a *fourth law of thermodynamics*, share the recognition that open systems far from equilibrium organize to exploit energy gradients more effectively.

But where does that leave us in the quest to answer the question "What is Life?" Well, we have a key component—a process by which we achieve regularity and order. But as we will see in the next two chapters on Relation and Constraints, we are not yet where we need to be, and some fundamental ingredients are still missing.

Chapter 2

Relation

> *Relax, you are just cosmic dust.*
> *-Graffiti at the Viennese Nightclub Flex*

If you tilt your head slightly forward and look down at your arms, torso, and legs, something strange happens. The entity that perceives the world is now perceiving itself. You, as a living process, are looking at yourself, the very living process that allows you to even perceive in the first place. Cosmic dust aware of its own existence in the universe[34].

This extremely basic type of conscious experience of ourselves is such a common phenomenon that we seldom stop to taste its strangeness. Because if we do, we tend to fall down a deep rabbit hole. Just think about it. Your subjective experience, your perspective on the world, comes to perceive yourself as an object in that world. But that object that you perceive is the very process that brings forth the subjective experience of yourself as an object. So, you are an object in your subjective experience, but that object has a subjective experience of itself being an object. You can continue to think about this until you get dizzy from spinning in ever tighter circles, losing orientation, and maybe meeting Alice.

Given this perplexing nature of self-awareness, it is no wonder that we are often plagued by questions about our identity, who we are, and who we aspire to be[35]. All of those bear seeds of this special relationship of a process being aware of itself and investigating its own relationship to itself and its relationship to its future and past self.

This relationship to ourselves is something we share with many of our close relatives in the evolutionary tree of life. Other animals, such as apes and dolphins, clearly understand that they are looking at themselves when confronted with their mirror image. Many more species than those who pass the so-called mirror test likely have an acute sense of self, but they may not be as vision-centered as we are[36].

However, at the scales and realms around which much of our investigation revolves—those of relatively simple single cells, like bacteria—this type of self-awareness is not present. Bacteria are prokaryotic organisms that lack a nervous system. They are simple, single-celled organisms without the complex internal structures that characterize eukaryotic cells, such as organelles and a nucleus. To speak of their self-awareness as something similar to that of humans and other animals would be a stretch too far.

Perhaps this strange type of relationship is a feature exclusive to "higher" organisms, like us, Homo sapiens sapiens—the twice-wise ones who have mastered the art of both understanding our cosmic insignificance and writing elaborate treatises about it while feeling rather pleased about our ability to do so. While our capacity for abstract thought likely sets us apart[37], before we drift too far into the realm of human exceptionalism, let us turn our attention to a similarly peculiar relationship present even in the most minimal of organisms.

For even the most basic of living processes are structured in a self-referential way: all of their parts work to bring forth the whole, maintain the whole, and replicate the whole. This whole is also the organized aggregation of those parts in a functional manner that continuously repairs, assembles, and further differentiates those parts. So, the parts collectively maintain the organization of the whole, while this very organization enables the continuous replacement and renewal of the parts. Despite constant material turnover at different timescales, the organizational pattern persists.

There seems to be something special about the types of processes that fold into themselves and bring forth their ability to continuously do so. This is what this chapter is about. Or more precisely, it is about two things: Organization and Relation. First, we will briefly look at Organizational Biology and examine how living systems exhibit empirically distinctive behaviors that demand explanation. Rather than invoking mysterious vital forces or emergent principles beyond physical law, we can understand these distinctive properties through the specific ways living systems organize matter[38]. To do so, we have to first take one step back and consider the standard approaches toward the question of what life is. Afterwards we will take a closer look at this type of self-referential relationship we have started to play around with in the above paragraphs, that shows up in many stories of how life originated.

2.1 Error of two Extrema

What makes an organism special? What differentiates a living system from a non-living one? For instance, a stone behaves drastically different from a frog. These qualitative differences in behaviour demand a scientific explanation within our naturalistic understanding of the world.

From the perspective of physics and self-organization, one might be tempted to dismiss the question entirely: there is no fundamental difference between living and non-living systems. An organism is simply another instance of matter self-organizing according to physical laws. This perspective, has deep roots in the Scientific Revolution's mechanistic world-view. It begins from the premise that the universe is composed solely of atoms which, organize themselves into increasingly complex forms, including what we call living organisms.

When adopting this reductionist perspective, we attempt to understand living systems by analyzing their smallest components, atoms and molecules moving according to physical laws, each following deterministic rules that collectively give rise to cellular functions. This mechanistic view has evolved considerably. Where earlier approaches saw life as intricate clockwork, contemporary versions frame living systems as computational processes, with cells acting as information processors that compute responses to environmental signals through molecular networks[39]. This computational lens of viewing biology through the language of algorithms, information flow, and processing, represents the current apex of bottom-up reductionism, and one whose limitations we will examine critically.

Yet when we shift our perspective to examine life from the top down—from the vantage point of large composite wholes and the felt conscious experience of being such a composite whole—we encounter a radically different picture. Think about the relationships you have with your loved ones, friends, and pets. Consider the intricate dance of ecosystems, where predator and prey move in seasonal rhythms across continents, or the complex social worlds of ant colonies. Reflect on the profound emotions you experience, such as grief when losing someone dear or the joy of being in love. Now think about life again, especially how it expresses itself in those beings closest to you. Notice how this perspective brings with it an entirely different sense of what it means to be alive.

There seems to be something very special about life. These living processes don't feel like mechanistic sequences playing out according to fixed rules. The way ecosystems adapt, the way communities organize, the way we experience emotions and create art - none of these feel like predetermined clockwork mechanisms simply executing their programs. At least we don't typically act as though these processes are purely mechanical. Our direct experience of living systems seems to resist reduction to purely mechanical explanations. This felt resistance to mechanistic reduction, while subjective, has historically also shaped how we think about and investigate living systems, leading to the tension between different explanatory frameworks that we'll explore.

Some might object that life is merely self-organization grounded in fundamental physics, or that emotions are epiphenomena of mechanistic computation. But I'm not offering a philosophical proof here—I'm pointing to something more fundamental: where we begin shapes what we see. Starting from the *smallest* (atoms and molecules) versus starting from the *largest*[40] (conscious experience) aren't just different analytical choices, they're different orientations that carry implicit assumptions, guiding us toward radically different understandings of life itself.

This isn't relativism. I'm not claiming both perspectives are equally valid or that truth is only perspectival. Rather, the question here is, how do these two ways of approaching reality relate to each other? Where is each applicable? And crucially, can they be reconciled, or does one necessarily exclude the other?

To sharpen this tension, let's examine actual thinkers who have taken these starting points to their logical extremes. We'll give these perspectives names and see where they lead.

Consider first Pancomputationalism, a perspective that posits computation as the fundamental nature of reality. In this view, all processes in the universe, ranging from quantum fluctuations to neural activity, are, at bottom, computational. Modern pancomputationalism, as articulated by figures like David Deutsch and Stephen Wolfram, goes beyond claiming that natural processes can be simulated computationally. It argues that the universe itself is fundamentally a computational system[41]. Wolfram's principle of computational equivalence, developed through his study of cellular automata, suggests that all processes in nature, whether the formation of a snowflake or human cognition, are manifestations of the same underlying computational dynamics. This view finds support through the Rolf Landauer's arguments that information is inherently physical, establishing a concrete link between computation and physical reality[42]. In this framework, consciousness and life themselves are seen as particular types of computation. Emergent patterns in a universal computational substrate. The universe is not just amenable to computational description, it is computation.

In stark contrast stands contemporary Panpsychism, which posits consciousness as a fundamental feature of reality. While its historical roots trace back to various vitalistic and animistic traditions, modern panpsychism, as developed by philosophers like Philip Goff and Galen Strawson, offers a sophisticated attempt to explain consciousness by making it intrinsic to physical reality itself[43]. According to this view, every physical entity, from fundamental particles to complex organisms, possesses some form of subjective experience or proto-consciousness.

These two perspectives represent radically different approaches to understanding the fundamental nature of reality and implicitly living systems. Pancomputationalism sees the universe as fundamentally *dead* computation from the bottom up, with properties like consciousness emerging as epiphenomenal computational patterns. Panpsychism, conversely, sees the universe as fundamentally *alive* and see consciousness everywhere, with complex awareness emerging through the combination of simpler forms of experience. This stark contrast exemplifies a deeper tension in our attempts to understand the relationship between mind and matter, between the quantitative world of computation and the qualitative world of experience.

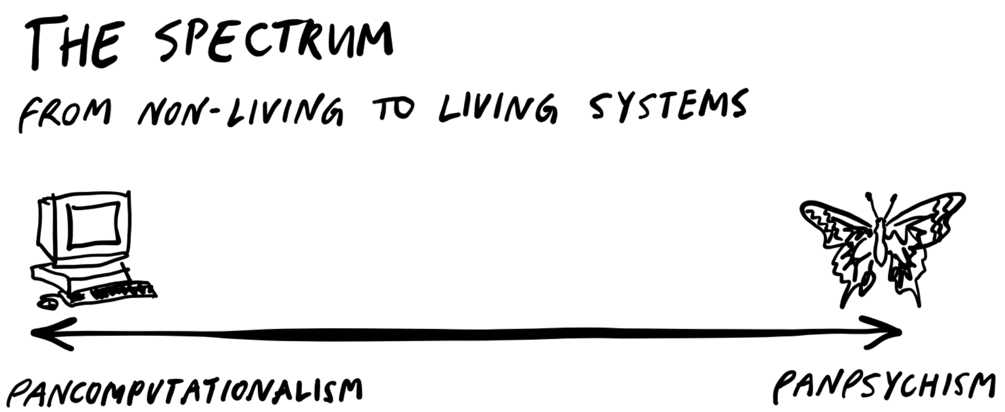

Figure 2.1: The spectrum between pancomputationalism (robot) and panpsychism (butterfly), depicting their contrasting views of reality.

Of course, the above positions are only contours of very technical and detailed discussions, however as rough guiding principles, they serve as useful starting points.

What merits our attention here is a specific pattern in how these perspectives approach the question of life and consciousness. Despite their opposing conclusions, both pancomputationalism and panpsychism share a particular methodological move. They attempt to resolve the apparent division between the conscious and the non-conscious, the living and the non-living, by universalizing one side of the division. Pancomputationalism extends computational processes to encompass all phenomena, while panpsychism makes consciousness ubiquitous. Both approaches, in seeking theoretical unity, effectively eliminate the very distinction they set out to explain. This reveals a deeper challenge in our theoretical approaches to life and consciousness, the tendency to prefer unified, monolithic explanations over frameworks that might preserve and account for genuine discontinuities in nature.

Such universalizing moves are philosophically elegant but empirically problematic. Let's examine why each fails, revealing a symmetry in their limitations.

Pancomputationalism has proven productive for modelling biological information processing, but struggles with what philosopher Evan Thompson calls "lived experience"[44]. The framework can describe how neurons compute but cannot explain why certain computational patterns give rise to subjective experience or meaningful agency. By making everything computational, it becomes harder to explain what distinguishes life and consciousness from other physical processes.

Panpsychism faces the inverse problem. By making consciousness fundamental to all matter, it becomes harder to explain why only certain organizations like brains but not rocks, produce unified conscious experiences. As Thomas Nagel notes, attributing consciousness universally doesn't explain how complex awareness arises from, or relates to, specific biological organization[45]. Making consciousness ubiquitous risks trivializing it, transforming a genuine achievement of biological organization into a universal backdrop.

The symmetry is striking. Despite opposite starting points, bottom-up computation versus top-down experience, both encounter the same compositional puzzle. Pancomputationalism cannot explain how *dead* computation becomes *alive* with meaning. Panpsychism cannot explain how simple proto-consciousness combines into unified awareness. Each faces the problem of composition from its own direction, how do simple elements create qualitatively different wholes?

This reveals their shared flaw. By making either computation or consciousness universal, both transform empirical investigations into definitional exercises. When we assume *a priori* that everything is either computation or consciousness, the distinctions that demand investigation disappear.

What organizational features distinguish living from non-living systems? What principles govern the emergence of agency and meaning? These questions become invisible when we've already decided, by definition, that there's no fundamental distinction to explain. Both positions *dissolve* rather than *solve* the question of life's distinctiveness.

Given these symmetrical failures, one unable to explain how dead matter becomes meaningful, the other unable to explain how simple consciousness becomes unified, perhaps the answer lies not in universalizing either side, but in preserving their distinction. This is the dualist's gambit.

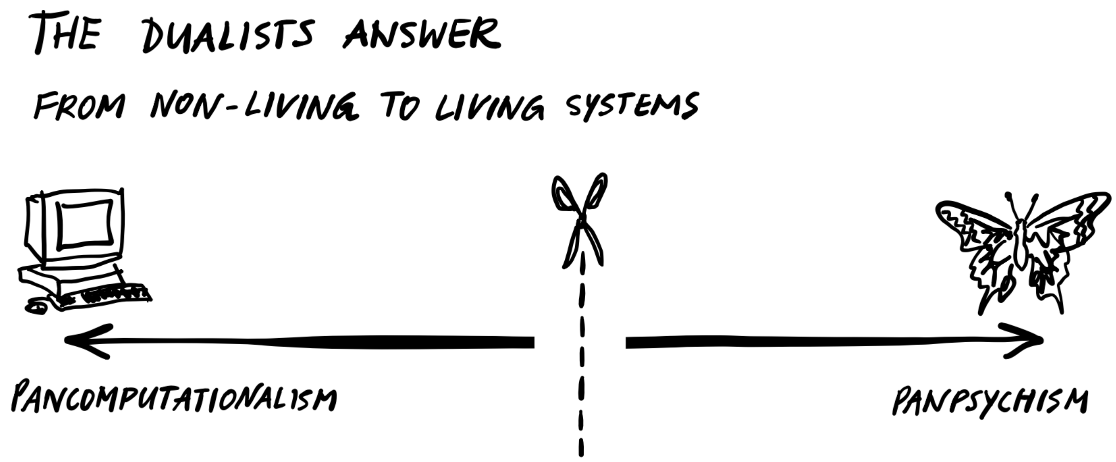

Figure 2.2: A spectrum split down the middle, illustrating Descartes' dualism: the red line marks the divide between extended substance and thinking substance.

An alternative to these monistic frameworks can be found in *dualism*, most influentially articulated by René Descartes. Descartes, of course, stands here as a paradigmatic example, one among many dualist approaches that have been proposed throughout the history of philosophy, but we focus on him because his formulation remains both historically foundational and philosophically influential. His approach attempted to resolve the tension between mechanical and vital aspects of nature by positing two fundamentally different but equally real substances: *res extensa* (extended substance) and *res cogitans* (thinking substance). This systematic distinction aimed to preserve both the mechanical causality required by the emerging natural sciences and the distinctive features of mind and life that seemed to resist mechanical reduction. In Descartes' framework, extended substance occupies space and follows mechanical laws, while thinking substance has no spatial extension and operates through reason and will[46].

The challenge for Descartes, then, was to explain how these two real but radically different substances could interact - a problem that led him to one of the more peculiar proposals in philosophical history, the pineal gland as the point of interface between mind and body[47]. This small endocrine gland, he argued, served as a kind of metaphysical antenna where the non-extended thinking substance could influence and be influenced by the extended physical body. The obvious difficulties with this proposal - how can something with no physical extension interact with a physical gland? - reveals a deeper problem with dualistic approaches, the inability to bridge the very divide they create[48].

This categorical approach generates its own set of philosophical and scientific challenges. First, there's the interaction problem[49] we've already encountered, how can two fundamentally different kinds of substance causally influence each other? Beyond this, Descartes' framework struggles to account for biological phenomena that seem to straddle his categorical divide. How do we classify phenomena like the placebo effect, where mere belief or expectation can trigger real physical changes in the body, an everyday case where mind and matter seem to cross one another's boundaries.

Moreover, Descartes' dualism faces an origin problem. How did this fundamental division in reality come about? In Descartes' own work, this required invoking God as the ultimate guarantor of the distinction and interaction between the two substances. But this theological solution, while coherent within Descartes' broader philosophical project, doesn't provide the kind of naturalistic explanation we seek for understanding the relationship between life, mind, and matter.

To summarize, neither the monistic frameworks of pancomputationalism and panpsychism, nor categorical dualism, provide satisfactory explanations for the emergence and nature of life. Pancomputationalism, in reducing all processes to computation, fails to account for the qualitative features of life and consciousness. Panpsychism, by making consciousness fundamental and ubiquitous, makes it impossible to distinguish between the specific forms of awareness and agency we observe in living systems and the absence thereof in non-living systems. Dualism, while acknowledging these qualitative distinctions, suffers from the need to impose an arbitrary boundary between the two, while grappling with the unresolved issue of how such categorically different entities could interact.

What we require instead is an approach that can explain how genuine novelty emerges in nature without either reducing everything to a single mode, or imposing arbitrary categorical divisions. We need a framework that can account for both the continuity between life and non-life and the genuine distinctiveness of living systems. This is where the concepts of organization and emergence become crucial, offering a way to understand how specific patterns of process and relation can give rise to qualitatively new features while remaining grounded in natural mechanisms that we can investigate empirically.

2.2 From Parts to Pattern

Pancomputationalism, Panpsychism, and Dualism each fail to explain life's distinctive character, though in different ways. Pancomputationalism dissolves the distinction by making everything mechanical computation, treating organisms as sophisticated information processors, no different in kind from weather systems or digital computers. Panpsychism dissolves it from the opposite direction, distributing consciousness everywhere, making it unclear why only certain systems (brains, organisms) exhibit the unified awareness and agency we associate with life. Dualism preserves the distinction but only by imposing an arbitrary divide, invoking separate substances without explaining how they could interact or how such a division arose in nature.

The concept of organization offers a way forward by reframing the question entirely. At its essence, organization refers to a distinctive mode of interaction among the parts of a system, characterized by both functional differentiation and integration. Functional differentiation means that the parts of the system perform distinct roles or functions—like how different organs in a body each have their specialized tasks. Integration refers to how these differentiated parts work together in a coordinated way, such that each part's activity depends on and supports the activities of other parts. These two aspects are inseparable: the specialized functions only make sense within the integrated whole, while the integration itself emerges from the coordination of differentiated functions. When considering life, this refers to how parts are arranged and collaborate in ways that enable the whole system to play an active role in maintaining its constituents over time—a process of self-maintenance[50].

The key insight is this: the characteristics of life do not arise from the individual components themselves, but rather from the way they are arranged and interact with each other. Nothing is added to the materiality; the difference lies in the arrangement. Matter is, pattern does.

This reframing solves what the other approaches cannot. Like Pancomputationalism, the organizational view acknowledges that living systems are made of ordinary physical components following ordinary physical laws—no special "life force" or separate substance required. But unlike Pancomputationalism, it recognizes that different patterns of organization can generate genuinely different behaviors and causal powers from the same substrate. A hurricane and a bacterium are both made of the same kinds of atoms obeying the same physical laws, yet their organizational patterns give rise to fundamentally different dynamics. The mechanistic base remains (same atoms, same forces), but the pattern of interaction generates new causal capacities that cannot be reduced to mere computation.

Similarly, like Panpsychism, the organizational view acknowledges there's something genuinely distinctive about living and conscious systems that demands explanation. But unlike Panpsychism, it doesn't need to make consciousness ubiquitous or attribute proto-experience to fundamental particles. Instead, it explains distinctive properties as emerging from specific organizational patterns—patterns that most configurations of matter simply don't exhibit. Not everything is alive or conscious, precisely because not everything achieves this particular form of organization.

And unlike Dualism, the organizational view requires no categorical divide between different kinds of substance. The constituents remain the same—it's their pattern of organization that distinguishes the living from the non-living. This allows us to trace a continuous path from simpler to more complex forms of organization without invoking arbitrary boundaries or mysterious interactions between incommensurable substances. We can understand how life emerges from non-life through the gradual elaboration of organizational patterns, each building on simpler forms while introducing genuinely new capabilities.

This is the connecting story the other frameworks cannot provide: same substrate, different patterns, genuinely different properties. Organization doesn't reduce life away (Pancomputationalism's error), doesn't make it universal (Panpsychism's error), and doesn't require arbitrary divisions (Dualism's error). Instead, it shows how specific arrangements of ordinary matter can generate extraordinary capabilities.

The insight that *pattern* rather than *substance* distinguishes the living from the non-living, has deep historical roots. As mechanical explanations increasingly dominated our understanding of the physical world, a pressing question arose: What makes living things different from complex machines? After all, both involve intricate arrangements of parts working together.

Immanuel Kant crystallized the crucial distinction between machines and organisms in his third critique on the Power of Judgment[51]. Both can have parts that remain physically the same when separated—just as a gear remains a gear outside its clock, so too a heart remains a heart after removal. Yet in living systems, each part's functional identity is generated and maintained only through its organizational relationships to the other parts. Unlike a clock, whose external watchmaker supplies its gears, an organism produces its own parts and self-maintains them in continuous exchange with its environment.

The capacity for *closure*, whereby the organism's parts depend on and regenerate each other, is what distinguishes biological organization from mechanical assembly. The heart's biological identity arises from its role in a self-organizing living body, something that a gear, dependent on an external maker, can never achieve on its own. This is the foundational insight, living systems are characterized by a circular pattern of organization where the whole produces the parts that maintain the whole.

Georges Cuvier took this philosophical insight and articulated it as a scientific principle, what he called the *condition of existence*. For Cuvier, the different parts of an organism must be coordinated in such a way as to enable the existence of the being as a whole[52]. This principle of functional correlation meant that every part of an organism must be compatible with all other parts for the organism to survive. For instance, a predator's sharp teeth must correlate with a digestive system capable of processing meat, claws suitable for catching prey, and sensory organs adapted for hunting. This lens on organismal organization allowed Cuvier to reconstruct extinct animals from fragmentary fossil remains, inferring the necessary relationships between different body parts[53].

This understanding of organization as fundamentally relational bears emphasizing[54]. The functional differentiation and integration that characterize living systems cannot be reduced to properties of individual components—they emerge from the web of relationships between parts. What matters most is not what each component is in isolation, but how it participates in and contributes to the organized whole[55].

Each component's function is both a contribution to and a product of the system's organization.

If biological organization is characterized by functional differentiation, integration, and self-maintenance, we need to understand the process by which such arrangements arise—how simpler systems can give rise to more complex organizational patterns without invoking anything beyond ordinary physical processes.

This is where the concept of *emergence* becomes essential. Emergence provides the conceptual bridge we need, showing how complex organizational patterns can arise from simpler ones while remaining grounded in the same physical components. Rather than requiring the addition of new substances or forces, emergence describes how novel properties and behaviors arise from the way components interact within a system. New patterns of organization can generate qualitatively different behaviors from the same underlying materials.

Consider common examples of emergence like snowflakes forming, whirlpools in your bathtub, or clouds organizing themselves into distinct formations. In each case, simple physical forces and local interactions give rise to stable, organized patterns. The whirlpool, for instance, emerges from countless water molecules following basic fluid dynamics, yet exhibits properties like stability, structure, and the ability to maintain its form in the presence of an energy gradient, that cannot be found in any individual water molecule. These examples, as discussed in previous sections, demonstrate how order can spontaneously arise from the interactions of simple components.

Yet the emergence we see in living systems represents something qualitatively different. While physical self-organization shows how order can arise and persist in the presence of energy gradients, living systems exhibit a distinctive relationship to their own organization and maintenance. The whirlpool organizes water flow. The bacterium organizes the very processes that sustain and reproduce its organizational pattern. This distinction between systems that exhibit emergent order and systems that actively maintain and reproduce the conditions for their own organization, marks a crucial transition.

In this sense, living systems exhibit a higher-order form of emergence, where patterns of causal organization arise that can actively maintain and reproduce themselves (for a detailed discussion and helpful framework to think about this see [56]). It's not only that living systems are more complex than whirlpools or snowflakes, they embody fundamentally different organizational principles. They use the same basic physical components and obey the same physical laws, but their pattern of organization generates new causal capacities, to maintain boundaries, to harness energy gradients, to repair damage, to grow, to reproduce. These are not just quantitative differences in complexity, but qualitative differences in the kind of organization present.

Organization and emergence thus provide the framework we need to move beyond the failures of Pancomputationalism, Panpsychism, and Dualism. The organizational perspective reveals how living systems achieve their remarkable properties through specific patterns of interaction and integration, using the same physical substrate as non-living systems but arranged in fundamentally different ways. Emergence provides the conceptual bridge, showing how these complex organizational patterns can arise naturally from simpler ones without requiring additional substances, forces, or arbitrary categorical divisions.

However, we have not yet explained the mechanism. Biological emergence maintains organizational patterns through constant material turnover—but *how*? What specific types of relationships allow patterns to be both stable and dynamic, both self-maintaining and capable of evolution? Understanding this requires examining the circular, self-producing character of living organization more closely.

This brings us to the question of how these patterns actually function—the specific types of recursive relationships that enable a system to produce the very components that maintain its organization. We explore this in the next section through autocatalytic sets and other models of self-maintaining organization that provide a stepping-stone into the direction of where we want to go.

2.3 Relationships

> *Self-reference is awkward: one may find the axioms in the explanation, the brain writing its own theory, a cell computing its own computer, the observer in the observed, the snake eating its own tail in a ceaseless generative process.*
> *–Francisco Varela*

To understand how these ideas manifest in living systems, we need to examine some specific theoretical frameworks that have attempted to capture the essence of life's distinctive organization.

First, we'll explore Maturana and Varela's concept of *autopoiesis*, which describes the form a circular organization where living systems continuously produce themselves. Then we'll examine Stuart Kauffman's work on autocatalytic sets, which demonstrates how networks of chemical reactions can create simple forms of circularity through self-maintaining reaction cycles. Finally, we'll look at Robert Rosen's (M-R)-systems, which provide a more rigorous framework for understanding the fundamental self-referential relationships that characterize biological organization[57].

Each of these frameworks approaches the same fundamental puzzle from a different angle: How do living systems create and maintain stable patterns within the ceaseless flow of process? As we'll see, their answers, while technically distinct, point to a common theme—the crucial role of circular, self-referential relationships in maintaining life.

Autopoiesis

Humberto Maturana and Francisco Varela developed a framework for thinking about the fundamental organizational principles relevant for understanding the distinctness of living processes. Their concept of *autopoiesis* (from Greek: auto- "self" and poiesis "creation") came from an intellectual tradition closely aligned with phenomenology and the biology of cognition. It foregrounds how living systems continuously produce not just their components, but their very organization.

Maturana and Varela define an autopoietic system as "a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in space in which they (the components) exist by specifying the topological domain of its realization as such a network."[58]

This definition attempts to captures three crucial aspects of living organization. First is *operational closure*, the system's processes form a closed network where each process is produced by other processes in the network. This closure isn't *material* (the system exchanges matter and energy with its environment) but *operational*, the system's operations produce the conditions necessary for their own continuation. Second is boundary production, where the system produces its own physical boundaries, which in turn make the system's continued operation possible. An example would be a cell membrane that is not just a container, but an active component produced by the cell's metabolism, making that very metabolism possible. Third is self-distinction, that through this continuous self-production, the system maintains itself distinct from its environment while remaining materially coupled to it.

The contrast with *allopoietic* systems illuminates these properties. An allopoietic system (from allo- "other") produces something other than itself. A car factory produces cars but doesn't produce the machines that make cars or the workers who operate them. But a living cell produces everything needed for its own production—enzymes that catalyze reactions, proteins that form structural components, lipids that maintain membranes, and so on.

This self-production operates at multiple scales and levels. At the molecular level, proteins catalyze the production of other proteins. At the cellular level, membranes enable the processes that maintain membranes. At the organismal level, organs maintain the conditions needed for their own functioning. Even at the social level, living systems can participate in higher-order autopoietic systems.

This pattern of organization has implications for our understanding of living systems[59]. Autopoiesis shows how living systems cannot be understood through pure input-output analysis, as they produce their own boundary conditions. They maintain their identity through constant material flux, maintaining form while exchanging substrate. They are inherently historical, as each state depends on previous states of self-production. And they exhibit emergent properties not reducible to component properties. This suggests we need theoretical tools that can capture not just material relationships, but the abstract patterns of organization that define life.

Yet perhaps the most relevant implication is a philosophical one. Autopoiesis shows us that understanding life requires looking beyond both material components and their circular relationships. It emphasizes that life's essence lies in its ability to continuously produce not just its materials, but its very organization.

Autocatalytic Sets

Stuart Kauffman's work on autocatalytic sets provides our second, more concrete window into life's circular organization by showing how it manifests at the level of chemical reactions. His work shows how simple molecular systems can exhibit the kind of self-maintaining patterns that, as we'll see, characterize life at deeper levels of organization[60].

To understand this, we must first look at the mechanism of simple catalysis. In a standard chemical reaction, a catalyst is a molecule that facilitates a transformation without being consumed by it. In the reaction below, the catalyst F (the purple triangle) binds two substrates, A and B. It accelerates their combination into a new product, C, before releasing it and returning to its original state, ready to work again.

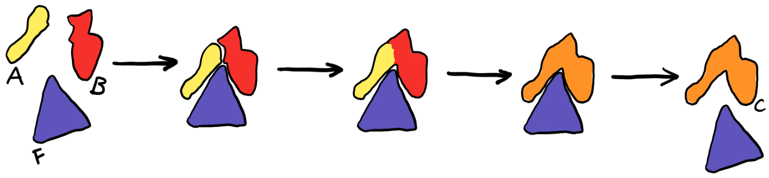

Figure 2.3: The molecule F acts as a catalyst, binding substrates A and B to accelerate the production of C, emerging from the process unchanged[61].

Now, imagine a second, separate reaction. Here, the product from our previous reaction (C) acts as the catalyst. It facilitates the combination of different precursors, specifically substrates D and E, to synthesize the molecule F (the triangle).

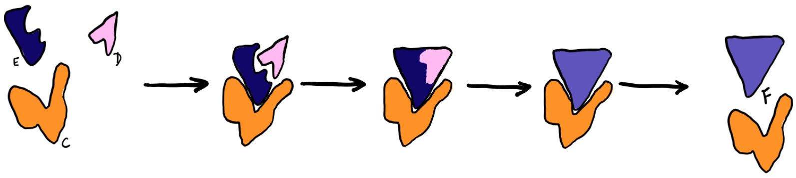

Figure 2.4: Here, the molecule C acts as the catalyst, facilitating the synthesis of F (the triangle) from its own set of precursors (D and E).

When we view these two reactions together, a circular pattern emerges. F catalyzes the production of C, and C catalyzes the production of F. This is the simplest form of an autocatalytic set, a system where molecules are not just produced, but are produced by the very network they sustain, depicted in figure 2.5.

While this simple two-step cycle illustrates the principle, Kauffman generalized this to much larger scales. Consider a system with N different types of polymers, each capable of catalyzing the formation of others. As N increases, the probability of finding closed cycles of catalysis, where each member helps produce other members, grows dramatically[62].

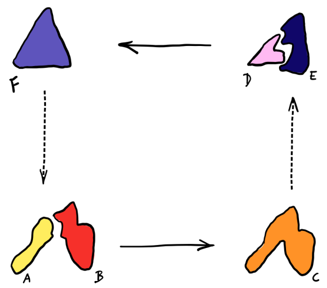

Figure 2.5: The two reactions form a closed, autocatalytic loop. F helps make C, and C helps make F.

Beyond a certain threshold of complexity, such closure becomes almost inevitable. Kauffman called this *order for free*, highlighting that the emergence of organized patterns isn't always something that needs to be imposed from outside but can arise spontaneously from the basic properties of complex systems. This insight hints at how increasingly sophisticated forms of circular organization might emerge from simpler beginnings.

Many approaches to the origin of life often sought a single self-replicating molecule (like RNA) as life's precursor. Kauffman's work suggests an alternative, perhaps life began not with a single replicator but with networks of mutually catalytic molecules. Recent experimental work supports this possibility. For instance, networks of small RNA molecules have been shown to exhibit collective self-replication, where no single molecule can replicate alone but the network as a whole persists[63].

This collective approach to emergence appears throughout biology. Modern cells are essentially highly evolved autocatalytic sets, where proteins catalyze the formation of other proteins, RNA molecules aid in protein synthesis, and lipids self-assemble into membranes that contain the whole process. Even metabolism can be viewed as an autocatalytic set, where each enzyme is produced by reactions catalyzed by other enzymes in the network.

Kauffman's framework also suggests new ways to think about biological evolution. Rather than focusing solely on the competition between individual replicators, we might consider the evolution of entire autocatalytic networks. This opens questions about higher-order selection and the emergence of new levels of organization. How do networks evolve? What determines their stability? How do new catalytic relationships emerge?

The concept has found applications beyond chemistry[64]. Economic systems exhibit autocatalytic properties, with each industry's products enabling other industries' processes. Neural networks can be understood as autocatalytic networks of activation patterns. Even cultural evolution might be viewed through this lens, with ideas and practices catalyzing the emergence of new ideas and practices[65].

Yet perhaps most significantly, autocatalytic sets provide our first glimpse into how circular organization - a feature that, as we'll see, characterizes life at much deeper levels - might emerge from simpler chemical processes. While this material circularity is just the beginning of the story, it lays the groundwork for understanding how more profound forms of self-organization might arise.

As we move forward to examine other theoretical frameworks, we'll see how this basic pattern of circularity extends beyond mere chemical reactions to encompass deeper forms of self-production and self-maintenance. But Kauffman's work remains crucial because it shows how even at the most fundamental material level, the seeds of life's distinctive organization are already present in the basic properties of complex chemical systems.

M.R. System

Where autopoiesis focused on life's self-producing organization, Robert Rosen's M.R. (metabolic-repair) systems take us to a more abstract level, using mathematics to show why living systems are fundamentally different from machines[66]. His work demonstrates that this difference isn't one of complexity, rather, living systems embody a form of organization that cannot, even in principle, be reduced to or modelled by mechanical systems.

To understand how Rosen arrived at this claim, we must first translate our concrete chemical picture into his abstract graph-theoretic language. Consider the basic catalytic reaction shown in Figure 2.6. Here, we see a straightforward transformation: a catalyst F acts on material substrates A and B to produce a new material C.

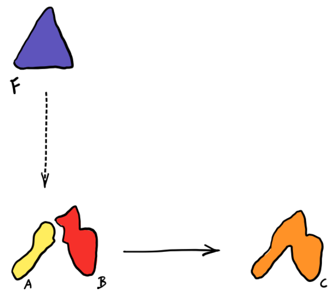

Figure 2.6: The basic unit of biological causation. A processor F (catalyst) transforms material inputs A and B into product C. The dashed arrow represents *efficient causation* (the work of the catalyst), while the solid arrow represents *material causation* (the flow of matter).

Rosen's formalism abstracts this process. He is not interested in the specific molecules (like A or B or C), but in the structure of the causation itself. In his notation, the specific collection of inputs (our yellow and red molecules) is grouped into a single abstract input set A. The product (our orange molecule) represents the output set B.

This leads to his formal notation, a single mapping $f : A \to B$. In the diagram below, f represents the efficient cause transforming the material set A into product set B (Figure 2.7). This kind of straightforward causation is characteristic of machines, where there's a clear separation between the processor f and what's being processed A.

We can extend this to include multiple linked processes, as shown in Figure 2.8, where two mappings are connected through a set A that serves both as the output

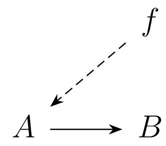

Figure 2.7: The basic unit of causation: a processor f transforms material A into product B. Dashed arrows represent efficient causation (action), solid arrows represent material causation (flow).

of one process (g) and the input for another (f). While more complex, this is still essentially mechanical - each component has a fixed role as either processor or processed.

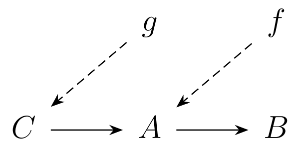

Figure 2.8: Sequential processing: multiple transformations connected through shared material causes.

The hierarchy deepens in Figure 2.9, where we see functional entailment: g not only produces a material cause (as in 2.8) but actually produces another processor f. This begins to blur the distinction between hardware and software that was clear in simpler mechanical systems.

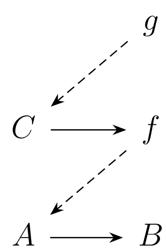

Figure 2.9: Functional entailment: g produces not just a material cause but another processor f

The important step happens in Figure 2.10, where we see two examples of how these hierarchical relationships close back on themselves to form what Rosen called *hierarchical cycles*. Here, each component in the system is both produced by and produces other components in the system. This circular causation, where the distinction between processor and processed breaks down, is what Rosen identified as the defining characteristic of living systems.

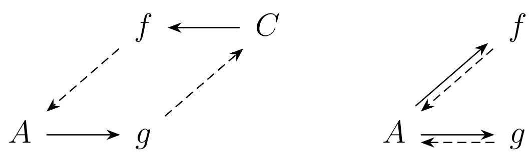

Figure 2.10: Closure to efficient causation: processes that mutually produce each other, creating organizational loops that cannot exist in conventional machines. The left and right diagrams show alternative topologies of the same fundamental circularity.

To make this abstract principle concrete, Rosen developed his foundational model of a minimal organism: the (M,R)-system. Here, a metabolic component M (represented by the mapping f) transforms environmental inputs A into cellular components B. Crucially, a repair component R (represented by the mapping g) then uses these very components from B to produce, maintain, or repair the metabolic machinery f. This creates the quintessential closed loop: f makes B, and g uses B to make f. The system pulls itself up by its own bootstraps, with no external mechanic needed to replace the parts.

However, while mathematically useful, the (M,R)-system presented a significant challenge in respect to its biochemical realization. When trying to map the model onto the a cell, its elegant abstractions became problematic. What, precisely, was the *repair* mapping g? And how could the set of metabolic products B act as both a material cause (building blocks) and an efficient cause (the repairer of f)? The model captured the essential logic of life but struggled to connect with specific cellular mechanisms.

This is where Jannie Hofmeyr's (F,A)-system provides a critical refinement, building on Rosen's logic to create a more biochemically plausible model[67]. Instead of a vague *repair* function, Hofmeyr splits the process into two distinct, observable stages: Fabrication (F) and Assembly (A). Fabrication is the synthesis of components, like a ribosome making a polypeptide chain. Assembly is the process by which these components become functional, like the polypeptide folding into a working enzyme. Crucially, Hofmeyr identified the intracellular milieu, the specific, actively maintained chemical environment within the cell, as the efficient cause of assembly. This introduces a new key player (D in the diagram below) that represents the functional context necessary for life.

The two models, both closed to efficient causation, illustrate this conceptual evolution from an abstract logical structure to a biochemically grounded one.

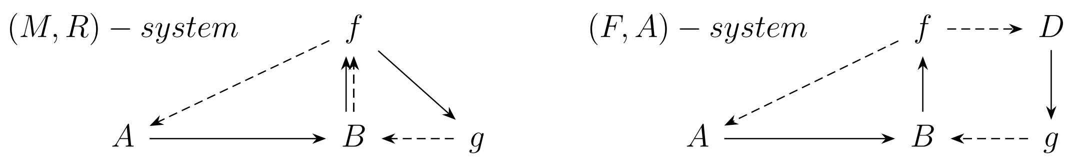

Figure 2.11: Rosen's model (left) captures the logical structure of life, while Hofmeyr (right) adds the explicit functional context (D) necessary for biological realization

Rosen formalized this circularity using category theory, and his choice of mathematical framework was crucial[68]. Category theory allowed him to represent precisely why living systems exhibit a kind of organization that cannot be captured by traditional mechanical models. The key insight is what Rosen called *closure to efficient causation*. In a living system, every component that acts as a cause (like an enzyme catalyzing a reaction) is itself produced by processes within the system. There is no external controller, the system produces its own catalysts, its own efficient causes.

In formal logic this type of relationship is called *impredicative*, definitions that reference the very thing being defined. Consider "the tallest person in the room", you can't identify who this is without comparing them to all other people in the room, including themselves. While such self-reference might seem logically problematic, Rosen showed that this kind of impredicative organization is exactly what makes living systems special and fundamentally different from machines.

This organizational pattern appears throughout biology. Enzymes catalyze the production of other enzymes, proteins maintain the genes that code for proteins, membranes enable the processes that produce membrane components. Rosen showed mathematically why such organization cannot be fully captured by any mechanical model. No matter how complex a machine becomes, it can never achieve the kind of closure to efficient causation that characterizes even the simplest living cell[69].

The implication is that there can be no complete mechanical model of a living system. Any attempt to reduce life to pure mechanism must necessarily leave something out. This isn't because biological systems are too complex for our current models, rather, Rosen proved that the very organization of life is fundamentally different from that of machines. While machines always trace their efficient causes to external designers or builders, living systems embody a different kind of causation altogether, one where the system itself produces all its own efficient causes in a closed loop.

These three perspectives—starting with Maturana and Varela's self-producing systems, Kauffman's chemical networks, and Rosen's formal treatment—offer us insights into life's uniqueness. Autopoiesis focused how circular patterns describe the continuous production of the system's very organization. Next we looked at the material circularity of autocatalytic sets, showing how even simple chemical systems can exhibit self-maintaining patterns. Finally, Rosen's analysis argues why this organization represents something fundamentally different from mechanical causation, a closure to efficient causation that cannot be reduced to, or modelled by machines.

This leaves us with a picture of life's distinctive character. Not only as a complex chemical system, but as a specific kind of circular organization, one that creates and maintains itself. Understanding this organization, and why it cannot be reduced to simpler, mechanical patterns, is key to understanding what makes life unique.

2.4 In Circles

Matter contemplating matter—that's where we began, and that's where we find ourselves again, though now we see this self-referential loop with new eyes.

We've built our theoretical foundation in a universe that's constantly changing, where the individual dynamics we're looking at are processual. We then looked at how the organization of these processes sets life apart from non-life, and how the emergence of new organizational regimes brings forth this specific kind of self-referential relationship between elements and their collective system, where parts and whole define each other.

As explored in this chapter, the essence of life reveals itself in circles, not only as an abstract concept, but in its fundamental nature. At its core are relationships that form something twisted: networks where individual processes interweave and reinforce each other. We are left with a paradox, a system both open and closed. Open in its constant exchange of energy and materials with its environment, yet closed in its circular organization. This closure is a pattern. A continuous cycle of processes that not only maintain themselves but give rise to increasingly complex forms of organization.

We've seen this pattern manifest in three distinct yet complementary ways: in autopoietic self-production, where processes produce the very network that produces them, in autocatalytic self-organization, where chemical reactions catalyze their own catalysts and in Rosen's closure to efficient causation, where causes loop back upon themselves. Each framework underscores a structural similarity from different angles, life maintains itself through circles of causation, production, and organization.

And here our circle completes itself, returning to where we began, cosmic dust contemplating its own existence. This capacity for self-reflection comes from the same circular patterns we've been studying. The very patterns that define life itself. We are, each of us, a loop in the universe looking at itself, parts organized precisely enough to marvel at their own organization. As in thought, so in matter: the universe contemplating itself.

Chapter 3

Constraint

Why is it the case that this sentence, which I am currently typing, ends here and has now jumped to the next line where it continues? This might sound like a strange question, but in principle, there are numerous ways of doing things differently. After writing from left to right, I could simply drop down one line and continue writing from right to left, creating a winding sentence that continuously coils down the page like a snake. This would also work, right? Or why not invert things and write only from right to left, as languages with linguistic ancestry in Aramaic—like Arabic and Hebrew—do? Or even top to bottom, like ancient Chinese.

What I am trying to get at is that we have certain conventions [70], norms, and constraints that establish how we deal with written language, helping to make it more regular and therefore more robust. This happens at the meso-level of sentences and how we string them together, but also on the lower levels, with individual letters that are, for all practical purposes, identical in print[71]. Even at the higher levels of sentence and paragraph structure, we have guiding conventions in place, where the complexity, length, and nestedness of sentences and paragraphs are conventionalized and constrained. As a reading of Kant showcases, the shorter and simpler sentence structures we favour today, compared to texts written a few centuries ago, reflect changing conventions. These changes may be influenced by shifts in our reading habits, such as the transition from books to digital screens, which often encourage quicker, more fragmented forms of engagement.

The interesting thing about the above examples is that these constraints and norms are neither predetermined by physical laws, nor are they purely arbitrary. Quantum Field Theory, Special or General Relativity do not dictate how one should structure sentences. However, these laws provide general constraints on what we can even do. For example, the gravitational pull of the earth on the ink (as well as the pull of the ink on the earth) constrain that we cannot draw letters in mid-air floating in layers on top of each other, so that we get a third dimension when writing. Similarly, the chirality of our hands, and a tendency for humans to be more dexterous with their right hand[72], biases the readability of a left-right writing system, as ink gets smeared less often. So, it's not purely nature, not purely nurture, but both.

Another example: We can think of the regularity in language as constraints that increase the readability and communicability of specific ideas. This can even be thought of in a spectrum, where on one hand, we have the lawless territory of no constraints, where language is no longer perceivable as such and fades into abstract art, consider James Joyce's *Finnegans Wake*, the sound poems of Kurt Schwitters or Ernst Jandl. Moving further, we can start to make out the contours of words and sentences in conceptual poetry, like the concrete poetry of E.E. Cummings, morphing into more structured literary works, with clear sentence and paragraph structure where belletristic works find their home. From here, things start to become rigid as we move into the field of scientific writing and publishing as well as manuals and codexes—all things that want to convince you of their clarity, certainty, and correctness, think of the strict IMRAD (Introduction, Methods, Results, and Discussion) structure of scientific papers or the precise language of legal documents. Lastly, we arrive at formal languages, mathematical notation, and programming languages—spaces many don't even want to visit due to their culturally perceived dryness and confinement, where even a single misplaced semicolon can break an entire system.

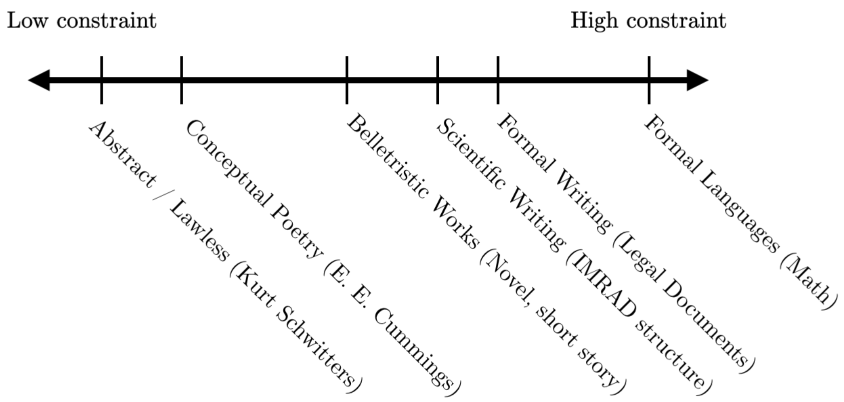

Figure 3.1: A tentative visualization of the language spectrum.

Neither part of the spectrum is per se better than any other part. Rather, the degree to which language is constrained makes it suitable for certain situations. For inspiration, one may prefer unconstrained language that widens horizons and lights up unconventional neural pathways. For determining the stability of a bridge over a roaring torrent, a highly constrained language that precisely tracks the mechanics and materials of the bridge is preferred by most. This alignment between the constraints of the system being described and the constraints of the language used to describe it is not coincidental. The effectiveness of our description and problem-solving capabilities often depends on how well the constraints of our chosen language functionally align with the constraints of the system we're working with[73]. The ability to be more precisely interpretable and better communicable also increases with introduced constraint—think of a law codex or the highly constrained nature of Morse code, the first fast, long-distance language. So, we see that constraint is something that comes in degrees.

While I wanted to give a few first intuitive impressions of constraints in a self-referential way—using language that describes constraints to describe constraints—the intention is to tie this concept of constraints to the emergence of life and show why this missing building block is of such profound importance. While there is no obvious parallel in language use [74], what we want to investigate in this chapter is how systems can emerge that have self-constraining dynamics—meaning where the source of constraint lies within the system and is not merely imposed from the outside or through convention.

To get there, work has to be done. First, I want to start talking about more precise ways of formulating what constraints are, especially in physics and chemistry, as they already have precise meanings in these domains—meeting our first interlocutor, Howard Pattee. Following this precise and structured analysis, we will shift to a broader, more philosophical and metaphysical perspective by engaging with Terrence Deacon, whose work has been a significant influence on this thesis and has inspired the focus on constraints. From there, we will move into the biological realm, exploring how constraints can provide an answer to Schrödinger's question. Finally, constraints will be interpreted in a processual framework, creating a bridge between the initial chapter on Process and the current chapter on Constraint

3.1 Laws & Constraints

We began our investigation by taking physics as the general bedrock from which we want to build up a theory of how living dynamics emerge, that differ qualitatively from those found in the non-living realms. To do this, let's consider the central shaping factor in physics—its laws—and how they differ from the constraints we observed in our earlier discussion of language systems.

Laws in the context of physics are understood as universal principles grounded in symmetry or invariance[75]. This implies that these laws can be represented as mathematical formulations that always rely on the notion of symmetry. Each of these symmetries leads to specific conservation laws as per Noether's theorem. Time-translation symmetry, where the laws of physics remain constant over time, results in the conservation of energy. Space-translation symmetry, defined by the invariance of physical laws when a system is moved spatially, leads to the conservation of linear momentum. Rotational symmetry or isotropy of space, which is the invariance under rotations, corresponds to the conservation of angular momentum.

What this amounts to is that these laws are seen as invariant and universally applicable, without exception, to all physical systems that can be modeled within the frameworks of Lagrangian or Hamiltonian mechanics. In contrast, constraints are local and specific. They are additional conditions imposed on a system that limit its behavior or dynamics.

An intuitive non-biological example of constraints in action is the function of a gun with a rifled barrel. The barrel itself acts as a constraint by confining the direction of the bullet's movement to forward motion, effectively channelling the explosive force from the ignited gunpowder to propel the bullet. Additionally, the rifling inside the barrel—a series of spiral grooves—imposes a rotational dynamic on the bullet. This spinning motion stabilizes the bullet's trajectory, enhancing its accuracy and range. By controlling the bullet's path and rotation, these constraints not only utilize the energy gradient generated from the gunpowder extremely efficiently, but also refine the bullet's behaviour to achieve specific, predictable outcomes in its flight-to move straight. This illustrates how constraints can direct and manipulate energy and motion within a system, crucial for achieving desired results and functionality.

It is important to appreciate the centrality of constraints here. If we placed a bullet next to a small amount of gun powder and initiated the explosion, the bullet would certainly move a bit, but the speed and direction would be barely predictable and much more chaotic than in the strongly constrained case. Thus, constraining the system, or limiting the degrees of freedom (the possible movements of the bullet), leads to a much more stable and predictable outcome in terms of direction and speed of the bullet.

However, our understanding of constraints as limitations on the degrees of freedom should extend beyond the mere presence of physical barriers or external limitations, like in the above example of the gun's barrel. It should also encompass the intrinsic properties of the system that determine its range of possible behaviours. These limitations are not static but can change with the system, influenced by both internal dynamics and external interactions.

As an example of this, we can consider the whirlpool that developed inside the water bottle we were trying to empty more quickly. In this case, we still have the constraint and the energy gradients, which would be the walls of the bottle and the gravitational gradient. The fascinating thing that happens here is that once the rotation of the whirlpool motion sets in, the system itself brings forth a set of constraints, in this case, that the water molecules can only rotate in one specific direction in a specific area (in order to leave space for the air column replacing the water). In other words, the movement of the individual water molecules is constrained, not by an external boundary but as an emergent phenomenon dependent on the molecular properties of the water, the constraints imposed through the energy gradients, and the wall. Therefore, many degrees of freedom are restricted to the individual water molecules as it self-organizes into a vortex to more quickly dissipate its present energy gradient.

This directly brings us to another relevant perspective on constraints through the concept of a phase space, which is a fundamental construct in the study of dynamical systems. In this multidimensional space, each possible state of the system is represented as a point, with the entire space encompassing all conceivable states. Constraints act within this phase space to limit the set of states that the system can occupy. They shape the trajectory of the system's dynamics, guiding it along specific pathways and precluding others. They define the accessible landscape of the state space. In the above case of the whirlpool, the phase space of the system would be more limited, in contrast to the phase space of the fluid gurgling out of the bottle in chaotic motion. Hence, when viewed through the lens of statistical mechanics, constraints acquire a quantitative dimension where constraints represent a reduction in the degrees of freedom of a system. In other words, the system is kept from visiting certain states through constraints.

So, this is the general context relevant when considering constraints. First, that they differ from laws. Laws provide the foundational framework within which physical phenomena operate. They are necessary for the description of any physical system, but they are not sufficient for explaining complex behaviours, especially those involving biological functions or information processing. Constraints, on the other hand, are necessary for the realization of specific behaviours or functions. They are the means by which a system's dynamics, in accordance with the laws is further directed and refined.

Furthermore, understanding laws does not require consideration of the system's history. Laws are considered objective and context-independent. Constraints, however, can be dependent on the system's history and the observer's actions.

These observations about the nature of constraints and their relation to physical laws bring us to the crucial insights of Howard Pattee, whose work has deeply influenced the understanding of how constraints operate in complex systems[76]. In essence, Pattee's contributions underscore the necessity of a dual, although not classically dualistic, approach in understanding complex systems: one that respects the universality of physical laws while also appreciating the unique, context-dependent constraints building on laws that govern complex dynamics.

This provides a way to delineate between systems that are predominantly governed by laws and those that are governed by constraints. Looking at the domain of physics we find laws as the central forces in many explanatory strategies. When moving from the realm of physics to chemistry and biology, we observe the successive introduction of more and more specific constraints. While much of physics can be traced back to the four fundamental laws (strong, weak, electromagnetic, and gravitational forces), chemistry introduces a set of constraints that, though based on these fundamental laws, manifest in unique interactions such as ionic and covalent bonds, hydrogen bonds, Van der Waals forces, hydrophobic interactions, and Pi-Pi stacking-the central forces of chemistry. Looking at biological systems organized far from equilibrium, it becomes clear that while the laws of physics provide a foundation and the specific constraints of chemistry shape interactions, biology introduces additional layers and levels of constraints that further limit the system's states, preventing it from accessing all possible configurations and maintaining it in a dynamic yet constrained state far from the thermodynamic equilibrium of death.

3.2 Types of Constraints

Having established the distinction between laws and constraints, we can now look at different types of constraints that shape the behaviour and evolution of complex systems[77]. In his work, Howard H. Pattee emphasizes the crucial philosophical and practical differences between holonomic and non-holonomic constraints and their implications for understanding biological organisms and other complex systems[78].

Holonomic constraints are restrictions that can be expressed in terms of equations relating the system's coordinates and time. These constraints define the possible configurations of a system and are integrable, meaning they can be described completely in terms of independent coordinates. In other words, holonomic constraints do not depend on the path taken by the system in its state space. A simple example of a holonomic constraint is a particle that must move along a specific path or surface—imagine a bead sliding along a wire—where the constraint can be described purely by the geometry of the path without reference to how or when the particle travels. The bead's position can be fully specified by a single coordinate along the wire, regardless of how it got there. Holonomic constraints, therefore, are often sufficient to describe the physical structures and static relationships within a system.

In contrast, non-holonomic constraints cannot be reduced to a constraint on the system's configuration alone and typically involve the system's velocities. These constraints are not integrable, meaning they depend on the system's path and cannot be expressed purely in terms of the system's coordinates. Non-holonomic constraints often arise from the dynamics of the system, like friction in mechanical systems and other forces that depend on the velocity of the system. A classic example is a rolling wheel that imposes constraints on its motion that depend not only on its position but also on how fast and in what direction it is spinning. Consider a bicycle wheel: its future position depends not just on where it is now, but on its current rotation speed, direction, and the history of its motion. The wheel's motion is constrained by its rolling nature, which cannot be described solely by its position at any given time—much like how living systems' future states depend on their historical trajectory.

Pattee argues that understanding the nature and type of constraints is crucial for modeling complex systems, particularly biological ones. He posits that holonomic constraints are often insufficient to capture the dynamic behaviors and evolutionary possibilities of living organisms, which are better modeled by considering non-holonomic constraints. Consider how a cell's behavior depends not only on its current state but on its developmental history, past environmental interactions, and accumulated genetic modifications. These constraints allow for a more nuanced description of systems where past states influence future possibilities, mimicking the adaptive and evolutionary capacities of biological entities.

The work of Rosen further reinforces the importance of non-holonomic constraints in modeling living systems. Rosen's work emphasizes the need to incorporate these constraints to adequately reflect the complexity and richness of biological systems. He argues that while holonomic constraints provide a basis for describing physical structures, non-holonomic constraints are essential for capturing the functional dynamics and information-processing activities that are characteristic of life.

Together, Pattee and Rosen's discourse illuminates the fundamental differences between holonomic and non-holonomic constraints, highlighting their relevance in the study of complex systems and the necessary conditions for life and evolution. Suggesting that the incorporation of non-holonomic constraints is a crucial step in bridging the gap between the physical description of systems and the understanding of their biological and evolutionary properties.

In summary, the distinction between holonomic and non-holonomic constraints provides a relevant framework for understanding the behavior and evolution of complex systems. While holonomic constraints are sufficient for describing static structures, non-holonomic constraints are essential for capturing the dynamic, adaptive, and evolutionary nature of living systems. As we continue to explore the emergence of life and its unique properties, the understanding of non-holonomic constraints will be a factor in advancing our understanding of these phenomena. In general, when talking about constraints in the future, I will be referring to non-holonomic constraints, unless otherwise indicated. The central aspect to remember there is that these refer to constraints that have a certain historicity, that they are dependent on their prior trajectory and dynamics—just like how a living organism's current state and future possibilities are shaped by its evolutionary and developmental history.

Lastly, while Pattee's distinction between holonomic and non-holonomic constraints provides us a formal framework, there is a deeper philosophical dimension to constraints that deserves our attention. The nature of constraints points us toward something fundamental: that what a system cannot do is as important as what it can do. This brings us to a broader meditation on the role of absence and limitation in shaping both our understanding and the phenomena we study. Just as non-holonomic constraints reveal how a system's past shapes its future possibilities, we might consider how absence itself, that which is not present, which is excluded or limited, plays a generative role in the emergence of form and function. The following exploration takes us into territory that might at first seem removed from our discussion of physical constraints, but will ultimately deepen our understanding of how constraints and absence together shape the possibility space of living systems.

3.3 Absence & Incompleteness

> *Dastehend ruht das Bauwerk auf dem Felsgrund. Dies Aufruhen des Werkes holt aus dem Fels das Dunkle seines ungefügen und doch zu nichts gedrängten Tragens heraus. Dastehend hält das Bauwerk dem über es wegrasenden Sturm stand und zeigt so erst den Sturm selbst in seiner Gewalt. Der Glanz und das Leuchten des Gesteins, anscheinend selbst nur von Gnaden der Sonne, bringt doch erst das Lichte des Tages, die Weite des Himmels, die Finsternis der Nacht zum Vor-schein. Das sichere Ragen macht den unsichtbaren Raum der Luft sichtbar.*
> *- Heidegger, Der Ursprung des Kunstwerkes* [79]

In the Western philosophical tradition, there has been a pervasive tendency to focus on the tangible, the present, the object—in short, on that which is[80]. This preoccupation with presence has often overshadowed the equally important role of absence, the unmanifest, the background against which the figure, the thing, emerges. Just as silence gives meaning to sound, and the pause between notes is essential to music, the significance of many phenomena cannot be fully grasped without considering that which it is not. This interplay of presence and absence, so fundamental to understanding, is particularly crucial when it comes to living systems. Life, in its tenacious resistance to equilibrium, in its incessant generation of order against the backdrop of entropy, is defined as much by what it does not do as by what it does.

The constraints that shape biological systems are not only limitations but sculptors of the actual, delineators of the possible. Consider a protein folding into its functional form: it is the excluded configurations, the paths not taken, that guide the molecule toward its biologically active state[81]. The very specificity that makes enzymes effective catalysts comes from their ability to exclude most possible molecular interactions while facilitating very specific ones.

In this chapter, we will play around with the importance of absence and incompleteness in understanding living systems, drawing on diverse perspectives that highlight the generative dynamics of the negative space. In doing so, we aim to redress the imbalance in Western thought that has long privileged the solid over the void. The following is therefore merely a bricolage of ideas that might be interesting for the inclined reader.

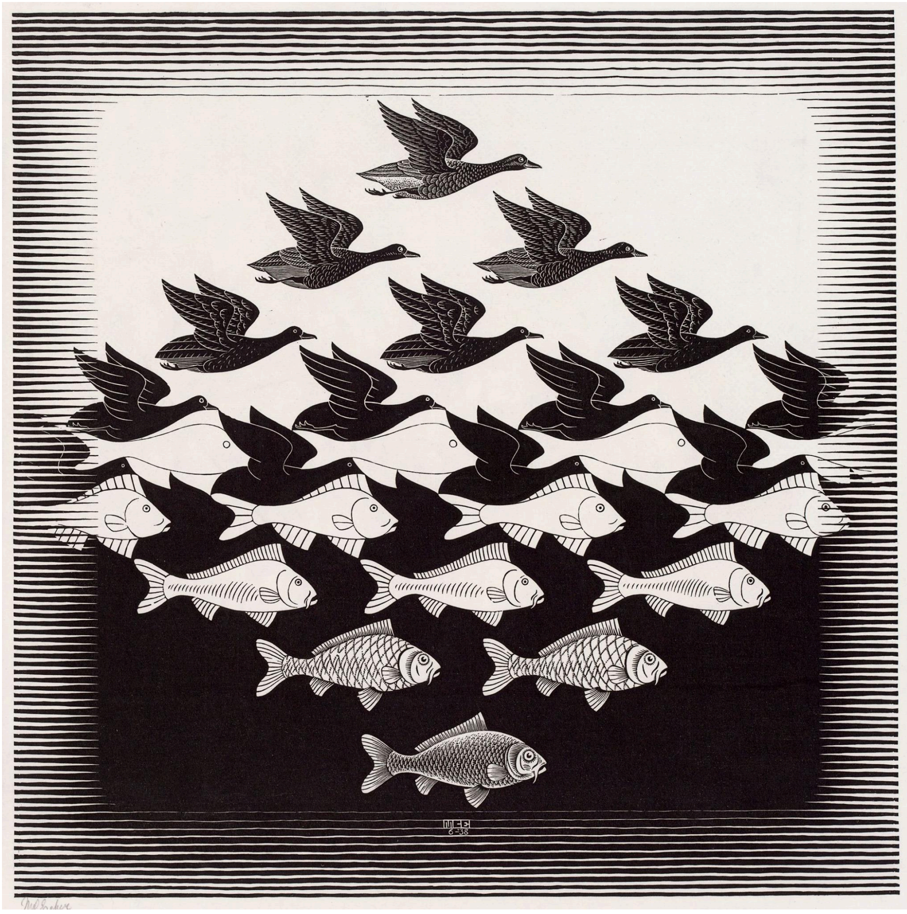

Figure 3.2: M.C. Escher, *Sky and Water I*[82]

What we will come to again and again is the general perspective that resembles a figure/background shift. Like in the above print by Escher, we can focus on either the black or the white and respectively see ducks or fishes. The fascinating thing about Escher's tessellations is that they depict the process of transformation from fish to duck by presenting the most articulate examples of fish and duck against a white sky and black water. The viewer is thus invited to undergo the experience of figure/background shifting by letting his gaze wander up or down, being confronted by the strongest generative tensions between those poles stretched out at the horizon line where the forms morph into one another—and moving beyond it, start to slowly fade. The metamorphosis inclines the viewer to participate in the shift of perspective and brings to the foreground the moment of change, that is otherwise hidden in the discrete duality of so many other figure/ground reversals.

Here the birds are water and the sky is fish.

Yet, even with the transfiguration depicted in plain sight, the viewer cannot really superimpose both the negative or positive views, but has to choose between them. When the focus lies at the horizon[83], we either see the birth of fish-form or duck-form. Both cannot be in focus. A choice has to be made. A particular perspective has to be taken.

This is more generally how I want to frame what is happening in this print. Two distinct perspectives flowing into each other, each birthing to the opposing one. Yet taken individually they are limited perspectives, either black or white. Only once we start to consider both the generative tension between them allows one to more fully appreciate how the fish is pulled out of the water starting to fly.

In some sense, we have a new dimension that emerges through the superposition of the two perspectives. The tangible biological parallel to this is, of course, parallax, the process by which the two dimensional pictures produced by each of our eyes are combined, to triangulate a sense of depth, the third dimension.

The ability to combine perspectives, whether through biological parallax or a conceptual synthesis, enriches our understanding of the world. Yet it also hints at something important: the fact that we need multiple perspectives to achieve depth perception suggests that any single perspective is inherently incomplete[84]. This incompleteness isn't just a feature of our visual system but extends to our entire sensory apparatus and how we interface with the world.

Take, for instance, our visual perception of light. What we perceive as visible light is a narrow band of the electromagnetic spectrum, roughly between 380 and 700 nanometres. We cannot directly perceive the vast expanses of radiation beyond these boundaries—the infrared that makes our skin feel warm, the ultraviolet that causes sunburn, or the gamma rays that penetrate deep into matter. Each of these invisible realms exists and influences our world, yet remains outside our direct experience. Other creatures, from infrared-sensing snakes to ultraviolet-perceiving bees, inhabit perceptual worlds quite different from our own. This shows how our understanding is fundamentally shaped by the spectra excluded, as by those it admitted.

As our eyes only capture a slice of the electromagnetic spectrum, our cultural frameworks similarly constrain and shape what we collectively see and understand. Cultural factors play a significant role in determining what we focus on and how we interpret the world around us. In the Western history of ideas, there is a tendency to emphasize the tangible, the present, and the object itself, often at the expense of the white spaces, the background, and the context in which the object exists.

This cultural difference in perceptual and cognitive approaches has been the subject of scientific investigation. Nisbett and Masuda have conducted famous experiments that highlight the stark contrast between East Asian and Western perspectives[85]. One such experiment involved showing participants an animated underwater scene containing various objects like fish, plants, and rocks. Western participants tended to focus on the most prominent objects, such as a large fish, describing its attributes and movements. In contrast, East Asian participants paid more attention to the background and contextual elements, noting relationships and interactions between the fish and the environment. This suggests that Westerners tend to analyze specific objects to understand their behavior, often overlooking the broader contexts in which these objects exist. In contrast, East Asians emphasize the broader perceptual and conceptual field, paying attention to relationships, changes, and the context when making causal attributions.

They argue that these differences in cognitive styles can be traced back to the social structures that shape these cultures. East Asians are typically embedded in complex social networks that necessitate a keen attention to context for effective functioning. Westerners, on the other hand, enjoy greater independence, which allows them to focus more on individual goals and attributes. These cultural differences are reflected in the artistic traditions of the East and West, with Eastern art focusing on broader fields and relationships, while Western ones tend to emphasize individual elements.

While specific manifestations of figure-ground dynamics vary across cultures, their persistent emergence across human societies is intriguing. The very fact that all human cultures, despite their vast differences, develop some version of figure-ground distinction suggests that this cognitive capacity may be rooted in our shared neural structure.

Think of a bird feeding on seeds. It must focus intently on the grains it distinguishes from the grit (a task requiring sharp, focused attention) while simultaneously maintaining a global awareness of its surroundings to detect predators. These two types of attention are functionally incompatible: one cannot narrow the spotlight to a single point and simultaneously keep the floodlight open to the periphery. Evolutionary biology suggests that brain lateralization—the division of labour between hemispheres—arose as a solution to this dual optimization problem[86].

Interestingly, this is not a feature unique to higher mammals. Research has demonstrated that this asymmetry is phylogenetically ancient, present in chicks, amphibians, and even invertebrates[87]. The necessity to segregate feeding from fleeing appears to be a fundamental architectural principle of nervous systems.

In humans, this evolutionary lineage manifests in distinct neural specializations. The left hemisphere, which controls the right hand in the majority of the population, evolved in tandem with tool use, fine motor manipulation, and eventually, language. As an example, Broca's area, a critical node for speech production and syntax, is located in the hemisphere responsible for complex, sequential motor control. This side of the brain excels at denotation, categorization, and the manipulation of discrete objects. It isolates the part from the whole to grasp it, both physically and conceptually. Conversely, the right hemisphere dominates in tasks involving spatial attention, the processing of emotional prosody, and the maintenance of a global *Gestalt*. It does not grasp, it surveys. It provides the contextual field within which the isolated object can be understood.

This brings us to the work of Ian McGilchrist, who has synthesized these neuroscientific distinctions into a cultural thesis[88]. He argues that these two neural dispositions represent two fundamentally different ways of *attending* to the world. The left hemisphere's mode is one of focused attention, essential for utilizing tools and clarifying mechanisms. The right hemisphere's mode is one of open awareness, essential for understanding living beings, relationships, and the *in-betweenness* of things.

While McGilchrist frames this as a hemispheric battle, the bias toward the tangible in Western thought can also be understood through the broader cognitive science of attention. We have a cultural and perhaps neurological predilection for focal attention, the spotlight that illuminates a specific object, over global broader awareness. When we do science, we are often engaging the neural circuits of singular focus by isolating a variable, removing it from its context, and manipulating it to see how it works. This methodological reductionism is the hallmark of the *grasping* mind. It has been undeniably successful in physics and chemistry, where systems can often be treated as sums of their parts.

However, this cognitive style becomes a limitation when applied to the complex organization of life. The Western scientific tendency to prioritize the particle over the field, the organism over the ecosystem, and the structure over the process mirrors the left hemisphere's preference for the explicit and the known. By focusing so intently on the object, be it the gene, the distinct anatomical structure, or the folded protein, we risk losing sight of the context, the constraints, the environmental pressures and the invisible causal landscape that gives the object its meaning. We see the manifest form, but often miss the vast array of excluded configurations that guided it there.

The challenge, then, is not to reject the analytical, dissecting power of the *grasping* mind, but to recognize it as an incomplete perspective. The neural architecture of the brain suggests that we are built to toggle between these views, to grasp the tool, but also to scan the horizon. A robust biology of absence requires us to reintegrate the right-hemisphere mode of attention, one that is comfortable with the implicit, the background, and the negative space. Only by combining the sharp focus on what is present with a broad awareness of what is excluded can we hope to understand the generative dynamics of living systems[89].

3.4 Closing a Loop

Origin-of-life theories have predominantly concentrated on the figure (the visible processes), rather than background (the constraints that make these processes possible). In the previous section on Relation we have discussed three so called "metabolic approaches"—theories that emphasize how living systems must continuously transform energy and matter to maintain their organization.

While these correctly identify the importance of active energy use in living systems, they often miss a crucial aspect: the role of constraints in directing and maintaining this energy use. Whether we look at Kaufmann's autocatalytic sets, Rosen's Metabolic-Repair-Systems, or other similar theories, we find a common limitation, they describe the visible processes of life without fully accounting for the constraints that make these processes possible and reproducible.

In essence, they extend the self-organizing story without necessarily bringing about a qualitative shift that distinguishes life from non-life. They describe increasingly complex chemical networks but don't fully capture the unique properties that define living systems - their ability to actively constrain themselves.

Let's examine these theoretical approaches. Kaufmann's autocatalytic sets focus on reaction networks, without explicitly addressing the role of constraints. While these networks can demonstrate recursive dynamics, they essentially scale up the basic story of catalytic reactions without introducing qualitatively new features. The fundamental limitation lies in their dependence on energy gradients - without a continuous energy input, these systems immediately collapse. This mirrors the behaviour of non-living dissipative structures rather than the distinctive characteristics of living systems. Kaufmann himself acknowledged this limitation in his later work, noting that the missing consideration of constraints left a significant gap in understanding how such systems could achieve the kind of self-maintenance characteristic of living systems[90].

Similarly, while Maturana and Varela's theory of autopoiesis provides valuable insights into the self-producing nature of living systems, it too doesn't fully articulate the emergence of constraints and their role in maintaining biological organization. Their work primarily focuses on organizational patterns, remaining largely descriptive rather than explanatory. While the concept of self-making is central to their theory, it doesn't adequately address how systems initially establish and maintain the constraints necessary for this self-production.

Rosen's (M,R)-system, particularly as developed through Hofmeyr's (F,A)-system extension, allows for a more nuanced reading of these issues. By incorporating Aristotelian causes, the model introduces formal causation, which bears some resemblance to what we mean by constraints. Formal causes determine the pattern or form that matter takes, paralleling how constraints shape and limit a system's possible configurations. The framework distinguishes between material causes (matter itself), efficient causes (dynamic aspects), and formal causes (organizational principles). However, the model remains primarily static and descriptive rather than dynamical. While its central feature of closure to efficient causation addresses how the system "makes itself," it doesn't fully explain how it constrains itself - both aspects being necessary for living systems. Perhaps most significantly, the model doesn't provide a clear evolutionary pathway from non-living to living systems, leaving the origin question largely unaddressed.

These approaches, while describing increasingly complex reaction networks, fail to capture the fundamental transition that distinguishes living from non-living systems. The key difference lies not in the complexity of the networks themselves, but in how these networks relate to and maintain their own organizational constraints. In nature, complex chemical networks are ubiquitous - from atmospheric chemistry to geological processes, we find intricate webs of interconnected reactions. What sets living systems apart is the ability to maintain and reproduce the constraints that govern their organization.

This distinction becomes clear when we consider how these systems behave in relation to energy gradients. Non-living reaction networks, no matter how complex, exist and function only in the immediate presence of an energy gradient. When the gradient disappears, the network disappears. Living systems, in contrast, can maintain their organizational structure and potential for activity even during periods without available energy. This persistence points to something fundamental about life: it's not just about maintaining chemical reactions, but about maintaining the constraints that enable those reactions to proceed in organized, functional ways over time.

This brings us back to our earlier discussion of the figure/background shift in perspective. Instead of focusing solely on the reaction networks themselves (the figure), we need to consider how these networks are constrained and organized by their own products and processes (the background). The critical question becomes: How does a chemical system transition from being merely complex to being self-constraining? How does it develop the ability to not only maintain its constraints momentarily, but to sustain and reproduce them across time and generations?

To answer these questions, we need to first understand how constraints emerge and operate in simpler systems, then examine how living systems differ in their relationship to these constraints. This progression will help us identify the crucial transition points between non-living and living systems. Therefore, before we get to the stabilization and reproduction of constraints, let's briefly revisit self-organizing systems and mention again how they develop structure through the emergence of constraints.

As discussed before, constraints can be described as limitations on the degrees of freedom in a dynamical system. These limitations prevent the system from visiting its most likely states—the states it would occupy if left in isolation, such as thermodynamic equilibrium or maximum entropy. Instead, through constraints, the system maintains patterns and structures far from equilibrium. This maintenance requires an open system where energy and materials can be exchanged with the environment, as closed systems inevitably tend toward maximum entropy, where all constraints dissolve.

In such an open system, the coupling of different processes can give rise to self-organizing dynamics. The heat from a stove tends to equal the heat of the surrounding space. This gradient interacts with a liquid, causing its molecules to move up. The result is a constrained system with convection cells that help dissipate the heat faster than in an unconstrained system with chaotic motion. Thus, not all states of the phase space are visited, only a select few. The self-organizing processes exhibit constrained behaviour, where the constraint is an emergent feature of the coupling of heat and the liquid. The specific constraints of the system are a feature that arises depending on the involved processes, and they describe which area of its possible phase space the system will visit.

What then are the properties of these constraints? Let's discuss two central ones, as articulated by Montévil and Mossio[91]:

First, although the constraints are a feature that arise through the interaction of the system, they have a certain static-ness or temporal symmetry at the relevant time scale for the system. In other words, the constraints are an emergent feature of the system; however, once they are established they have a dynamic stability in the continuous reduction in the degrees of freedom of a system.

Like the constraints of a whirlpool, they reduce the possible movement of the water and only emerge in the presence of an energy gradient—like gravity. They have their associated stability within a certain range of perturbation—if you disrupt the whirlpool in the bathtub, it will automatically reform as long as there is a certain amount of water and gravity present. Consequently, the constraints are stable over a given time.

Second, constraints are always relative to the system and provide stability and regularity via their reduction of possibilities. A constraint for any given system is always something that reduces its degrees of freedom in a way that differs from its absence. Thus, the process of a constrained system flows differently than that of one where this is not the case. Some new pattern emerges.

This stability at certain time scales for a given process and a difference in dynamics are central aspects of constraints. Let's keep this in mind as we move on to the question of why living processes are not merely self-organizing processes.

Constraints are central to our exploration of life, as any living process will also only visit a very select area of its possible molecular arrangements or its phase space. Just consider the ways in which human bodies can be rearranged that lead to nothing but a big heap of cells lying around doing nothing. The arrangements that have a functional and structural integration and exhibit living dynamics are orders of magnitude smaller.

However, so far, we haven't ventured much beyond self-organizing dynamics, even though we have now provided a negative or inverted lens to view them through with the help of constraints. As indicated earlier, there is a central problem with viewing life as just a self-organizing process. While self-organizing processes and living systems share some similarities, there are crucial differences that set them apart. Self-organizing processes organize and regularize around the fastest depletion of an energy gradient, whereas living systems prioritize the longest possible existence given an energy gradient. This difference in dynamic is one of the most central factors that distinguish living systems from non-living self-organizing processes. However, there are others that can be seen as variations on the theme:

Self-organizing systems passively respond to far-from-equilibrium conditions. Living systems actively maintain these conditions through controlled energy consumption. This represents a different relationship with energy flow. Where a self-organizing system like a storm dissipates energy as quickly as possible, living systems have evolved sophisticated metabolic networks that carefully regulate and throttle this energy flow, extracting maximum utility before releasing it back to the environment.

Second, this controlled energy usage enables another crucial feature of living systems: homeostasis. Through complex regulatory networks, living organisms maintain stable internal conditions despite environmental fluctuations. This stability is actively maintained through constant sensing and adjustment. The system anticipates, stores, and strategically deploys energy to maintain its organization.

Third, this regulated energy flow allows living systems to direct their processes toward specific functional outcomes. Growth, reproduction, and environmental responses are such examples. This again represents a shift from the rush toward equilibrium seen in non-living systems to a qualitatively different approach to energy use.

At more complex levels, this strategic energy usage manifests in increasingly sophisticated ways through hierarchical organization. Even the simplest cells demonstrate this principle: rather than allowing chemical reactions to proceed at their naturally fastest rates, they compartmentalize processes into specialized organelles, each controlling and directing specific aspects of energy flow. This organizational principle extends across all scales of life, from molecular assemblies to ecosystems, with each level adding new layers of control over energy flow and material transformation.

Each level of organization, from cellular membranes to organ systems, represents a set of constraints that must be actively maintained and reproduced. This points to another crucial distinction between living and non-living systems: the ability to not just maintain constraints, but to reproduce them across generations.

Bringing us to perhaps the most remarkable feature of living systems, their ability to encode and transmit the information needed to reproduce these energy-managing constraints. Living systems can pass on not just their structure, but the entire complex network of processes that enable controlled energy use.

The distinction between living and non-living systems thus comes into sharper focus: living systems are those that have evolved the ability to maintain, reproduce, and transmit constraints that enable the strategic use of energy for self-maintenance.

This raises a crucial question: how did such systems emerge from simpler chemical networks? How did the transition occur from systems that simply dissipate energy as quickly as possible to systems that carefully manage and control energy flow?

In the next section, we'll explore how our understanding of constraints and self-organization can help bridge this gap. By examining how simple chemical constraints might have evolved into reproducible patterns of energy management, we can begin to understand the emergence of life itself, bringing us closer to answering our central question: What is life?

3.5 Self-Constraint

In our exploration of self-organization, we saw how gradients and processes can become linked in ways that create constrained patterns - think about gravity and a bathtub. Now we'll take this concept one step further by examining what happens when self-organizing dynamics themselves become coupled to each other, creating new forms of constraint.

What does this coupling of self-organizing processes mean, and how does it differ from the cycles we discussed previously? At a basic level, self-organizing dynamics can be linked in a merely additive way, simply increasing the system's capacity to deplete energy gradients by involving more processes. However, self-organizing dynamics can also be linked in a way that is not just additive but that leads to a qualitative shift, where the self-organizing dynamics of one process begin to constrain the respective runaway tendencies of another process, to produce a system that self-constrains.

This is where we encounter *self-constraining dynamics*, or to use the terminology developed by Terrence Deacon, *teleodynamics*. It is important to state explicitly that the conceptual architecture of the *Genesis* part, is fundamentally built upon Deacon's work in *Incomplete Nature*. In his framework, teleodynamics describes a specific regime where self-organizing processes are coupled in such a way that they mutually constrain each other to prevent their own dissolution. This system exhibits properties distinct from mere self-organization, rather than simply dissipating gradients, it generates an intrinsic constraint to maintain its own organization against thermodynamic decay[92].

Basically, we now take the central themes of the last three chapters and bring them together to formulate a dynamical theory of self-constraining systems. Let's go through the main takeaways and conceptual tools very briefly again:

- **Process:** We need a processual lens, with the backdrop of classical thermodynamics, against which self-organizing processes stabilize and regularize via the quick depletion of an energy gradient.
- **Relation:** Nothing magical is added to a living system to make them come alive. Their magic lies in the specific form of organization and relation of interacting processes. These need to be structured in a specific kind of self-referential, impredicative, or circular relationship.
- **Constraint:** A pervasive tendency to prioritize the figure rather than the background has shaped much of our scientific, as well as biological thinking. Using a perspective of constraints, we can focus on the degrees of freedom not realized in systems and can ask how the system keeps itself from visiting those.

With these three perspectives fresh in our working memory, we now move on to self-constraining dynamics. And to get a good feeling for this special kind of dynamical system, we will begin by exploring an example, the autogen by Terrence Deacon[93]. He was the first or among the first[94] to really focus on constraint dynamics in living systems, Deacon was also the only one who provided a concrete chemical example that, in theory, can be implemented[95].

The basic idea of an autogen is a system made up of chemical reactions that support and enhance each other. It's not a specific blueprint of the first life form, but rather an example of how early life-like properties could emerge through natural chemical interactions.

In an autogen, there are two central types of chemical reactions. The first reaction involves a sequence of organic molecules reacting in a loop. As one reaction completes, it produces the ingredients needed for the next. This looping nature amplifies itself, creating more and more of these ingredients. This is the basic story of the autocatalytic set discussed above. The second reaction uses a by-product of the first reaction, a lipid, as a reactant. This reaction produces a boundary, like a microtubule or a capsid-like structure. This boundary acts like a container, keeping all the necessary components close together and limiting their spread into the surrounding environment.

These two reactions are cleverly intertwined. The first reaction generates materials used in the second, while the second reaction forms a boundary that contains and protects the first reaction. Together, they prevent each other from running out of control or exhausting all resources, which would lead to a breakdown of the system.

In the figure 3.3, we can see the temporal evolution of this system, moving from a scattered collection of parts (left) to a unified whole (right).

First, recall the autocatalytic set we explored in the previous section. There, we saw a loop where molecule F catalysed the production of C, and C catalysed the production of F. We treated F simply as a chemical agent, a purple triangle in a diagram.

However, in our previous description, we ignored a special property of F that is central to the Autogen story. Look at the soup on the left side of the illustration.

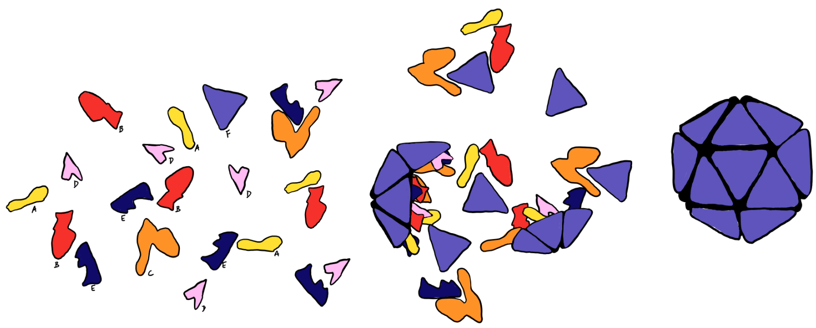

Figure 3.3: The Genesis of an Autogen: The system begins in a *primordial soup* state where substrates (A, B, D, E) and catalysts (C) interact, as depicted on the left. Towards the middle, the autocatalytic cycle acts as a runaway engine, rapidly increasing the concentration of the structural molecule F (purple triangles). Once a critical threshold is reached, F self-assembles into a closed polyhedral shell. This event physically separates the catalysts inside from the substrates outside, actively inhibiting the reaction and creating a self-constrained unity.

You will see the familiar reactions: A and B bind to form the catalyst C. C then facilitates the binding of D and E to create F. But here is the twist: F is not just a catalyst; it is a structural building block. The F molecule (the purple triangle) has a specific geometry that allows it to attach to other F molecules. This feature is not uncommon in chemistry; lipids, for example, naturally self-assemble into membrane-like structures due to their shape and polarity.

As we move to the middle of the illustration, the consequences of this property become visible. As the autocatalytic cycle runs, it acts as a runaway engine, producing more and more F molecules. When the concentration of F reaches a critical threshold, they spontaneously begin to click together.

On the far right, we see the final result: the F molecules have self-assembled into a closed polyhedral shell.

This is the decisive moment of self-constraint. By closing into a shell, the system physically separates the catalysts trapped inside from the substrate (food) remaining outside. This effectively turns the engine off. The autocatalytic reaction ceases not because it ran out of fuel, but because the system constructed a barrier that stops the fuel from reaching the engine.

So in a very basic sense, the system goes through a phase of explosive growth, producing its own container, and then, by encapsulating its reactive surface, it constrains itself from consuming everything in the environment.

But the story doesn't end at the polyhedral structure. For a while it might, but certainly not forever.

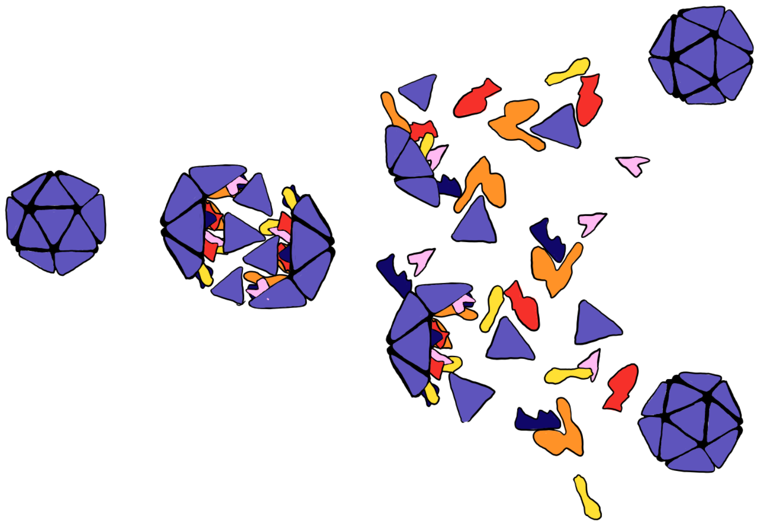

Figure 3.4: Depiction of the replication cycle of the autogen. From left to right: A stable autogen is perturbed and ruptures, releasing its sequestered catalysts back into a substrate-rich environment. The autocatalytic reactions restart, producing new structural molecules which self-assemble into multiple copies of the original system.

As illustrated in figure 3.4, the static phase is only temporary. Since the shell is held together only by weak molecular forces (like the hydrophobic effects in lipids), external perturbations like a change in temperature, a wave, or bumping into a rock, will eventually cause it to rupture.

The moment the shell breaks the sequestered catalysts (the accumulated C and F molecules) spill back into the surrounding environment. If this environment still contains food substrates (A, B, D, E), the autocatalytic cycle starts again.

Because the ruptured shell released a large concentration of catalysts all at once, the reaction network restarts, producing a fresh batch of structural F molecules. As the concentration of F rises again, self-assembly is triggered once more.

We have moved from a simple chemical reaction to a system that actively builds its own boundary, stops its own activity, and, when disrupted, leverages the environment to create copies of itself. It is a system that maintains and reproduces its own constraints.

This model, chemically unrealized but theoretically sound, can be likened to a *non-parasitic virus*. A standard virus, like polio, has a protein capsid shell enclosing genetic material, but it requires the complex machinery of a living cell to reproduce. However, in a prebiotic world with no living systems to parasitize, a system like the Autogen could achieve autonomous reproduction through simpler molecular processes.

The key lies in the chemical complementarity of reciprocal catalysis and self-assembly. Reciprocal catalysis is an explosive, runaway process that produces high local concentrations of molecular species, but without containment, these molecules naturally diffuse away, causing the reaction to fizzle out. Conversely, self-assembly is a constraining process that requires consistently high local concentrations of molecules to crystallize into structures. It cannot happen if the molecules are dispersed.

In the Autogen, these two opposing tendencies solve each other's problems. The catalysis provides the concentration needed for the shell to form and the shell provides the containment needed for the catalysis to persist (and eventually restart). They provide the necessary constraints for each other.

Thus, via the coupling of self-organizing processes in a way that they constrain each other, we have gotten something qualitatively very different than before. We now have a self-constraining system that has the ability to replicate in a minimal form.

This brings us to the realization of a higher-order pattern, visually captured in Figure 3.5. If we step back and view the system not as a single event, but *over time*, we see a sequence of growth, disruption, and re-organization. What is preserved here is not the specific matter. The individual molecules of A, B, C, D, E or F are constantly being swapped out, consumed, or dispersed—but rather the form of the organization itself. This is the first emergence of a historical lineage. The constraint is no longer just holding the molecule together in space, it is propagating itself through time. This cycling constitutes a higher-order constraint that remains invariant even as the physical substrates turn over. We have the first instance of a system that possesses a history, where the *self* that is being constrained is a dynamic pattern rather than a static object.

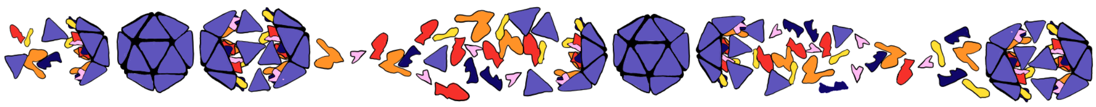

Figure 3.5: The Propagation of Constraints. A schematic representation of an autogenic lineage characterized by a recursive cycle of *growth, disruption, and re-organization*. The diagram illustrates the oscillation between these states. While individual physical instantiations of the autogen are transient, this cyclical dynamic allows the organizational pattern to persist, enabling constraints to be inherited and multiplied across generations.

Beyond that, and as important as the property of replication and the emergent higher-order constraint, the system can now exist, even in the absence of an energy gradient if it is in its shell-like form in a far-from-equilibrium state. We could imagine the system finding itself in a location with little perturbation and no catalytic material, and even in such a scenario the system could maintain its shape until perturbed.

The autogen has self-regulating properties that preserve its structure and function. It's a preliminary model of how biological functions could originate, with each reaction both supporting and being constrained by the other, preventing the system from simply running to completion and stopping. To get here, we needed the processual background of self-organization and thermodynamics, we saw that no sleight of hand was involved and nothing magical was added, and we got a system that not only stabilizes its constraints in the presence of an energy gradient, but one that can keep them in its absence and duplicate its constraints.

So to summarize, the exploration of constraints in this chapter underscores their critical role in distinguishing between non-living and living systems. Through the example of the autogen, we see how self-constraining dynamics can emerge from the interplay of simple chemical reactions. This model serves not only to illustrate the fundamental difference between *self-organizing* and *self-constraining* systems but also highlights the pivotal concept that life, at its core, involves processes that actively maintain and reproduce their constraints. You are not just the sum of your atoms, but the constraints the define the dynamics of how these parts interact, move around and get exchanged over time.

The autogen, as an unrealized chemical model, provides a compelling framework for understanding how life might harness and stabilize constraints to sustain itself against the inexorable pull towards equilibrium. By examining these dynamics, we can better appreciate the delicate balance between the forces of disorder and the emergent order that characterizes living systems. This perspective offers insights into the origins of life, the nature of biological processes, and as we will see in the following parts of the book, insights into the inherently paradoxical and organizational nature of living systems.

3.6 Process vs. Stability

Dear reader, we have arrived at the end of the beginning.

We started by musing about Cratylus, the tension between Process and Stability, noting that siding with either doesn't work. Process is fundamental, because we live in an ever-changing universe in constant flux where no single thing remains fixed and static forever. Still, certain things seem very stable and nearly unchanging, like ideas and concepts that have been transmitted over millennia or lineages of organisms that have evolved over billions of years.

This persistence suggests that the definition of life does not lie in a specific substance, but in a specific capability: the ability to construct enduring stability out of process. This is the defining difference between living and non-living systems.

I hope this perspective, perhaps cryptic when first introduced, can now be appreciated in a new light. Our exploration of self-constraining processes and the autogen allows us to grasp its full implications. We have discovered that life doesn't resist change; it harnesses it. A living system is not a *thing* that *processes* occur to, but a *process* that continues in such a way that it appears a *thing*.

A self-constraining process exists in a far-from-equilibrium state, yet achieves a form of stability through its constraints. The separation it maintains between inside and outside isn't the stability of equilibrium, rather a dynamically maintained organization. What appears as unchanging is in fact actively changing.

As long as the cycle turns and the system moves from shell build-up to completion, to breakage, returning to a reorganized growth phase, the system maintains and keeps its constraints expressed through the continuous stability of its organizational pattern. While every particle will be exchanged, the set of self-constraining dynamics remains the only constant in flux.

So in this strange turn of events, living processes achieve a form of stability that surpasses even many structures in equilibrium. By now, life on Earth has outlasted almost all rock[96]. What appears as ceaseless change manifests as permanence. Stability in living systems is not the absence of change, but the persistence of pattern.

Perhaps Cratylus's greatest gift to us was not his philosophy, but the paradox of its transmission. His declaration about never stepping into the same river once—itself a statement about the impossibility of permanence—has echoed across millennia. His ideas about the impossibility of stable meaning have themselves achieved a kind of immortality, not through static preservation, but through continuous reinterpretation and renewal. His thought pattern survives by being constantly reborn in new minds and new media, persisting despite radical transformations in the languages that carry it and despite jumping across different modes of transmission.

Like life itself, what endures is neither the physical matter nor some eternal Platonic form, rather a pattern of organization that maintains itself precisely through its capacity for constant renewal. Stability in living systems is not the absence of change, but the persistence of pattern amidst constant flux.

Part II

Paradox

Chapter 4

Cycles

Genesis was the story about a plausible origin of life. We have traced a path from basic thermodynamic processes to the emergence of self-constraining dynamics. What began as an investigation into how order emerges from disorder led us to the more subtle understanding that life's essence lies in specific patterns of constraint that interlock in such a manner, that they can continuously reproduce themselves.

Paradox, the second part of this book, focuses on the philosophical characterisation of life. This very chapter is about the question how self-constraining dynamics manifest across different scales of organization. The following chapter *Living in Time* will consider the relationship a living system can have with its own future and the third and last chapter, *Towards a Logic of Life*, will discuss what this particular relationship implies for our way of thinking and reasoning about such systems.

What we'll come to see is that understanding self-constraining patterns requires us to examine how apparently opposing processes can be integrated over time to maintain living organization. This is more broadly what this second part of the book is about.

We will focus on something paradoxical about self-constraining dynamics, which is that living systems must simultaneously maintain stability while remaining adaptable, preserve their boundaries while exchanging with their environment, and resist disorder while harnessing it for renewal. Opposing tendencies are actively coordinated over time in ways that make them mutually enabling rather than mutually destructive.

Examining these dynamics across different scales, we find a recurring pattern in how opposing tendencies are managed. This is named the *constraint cycle*. A recurring pattern of growth, disruption, and reorganization that maintains organization while enabling adaptation. Growth is the accumulation of resources and the building of structure. Disruption involves the necessary breakdown that creates opportunities for change. Reorganization encompasses the integration of novelty and the rebuilding of order.

In Deacon's autogenic model, these three phases are clearly sequenced. In cells, these tendencies run largely simultaneously. At the level of more complex composite systems, they often emerge as more temporally distinct tendencies in what Holling calls the *adaptive cycle*. Beneath these varying temporal arrangements lies a common theme, the dynamic management of opposing tendencies over time.

This chapter shows how the same threefold logic recurs across scales, why apparent "inefficiencies" (opposing processes) buy control, and how disruption is not loss, but the opening through which systems reorganize.

4.1 Autogenic Cycle

Building on our understanding of autogenic systems in the Genesis part, let's examine a crucial aspect, the work cycles that drive them[97]. These cycles are an example of how self-constraining dynamics manifest and provide a critical stepping stone in our quest to see how counteracting tendencies are coordinated over time across scales. This section is more dense and technical than others, so in case it becomes a hurdle, do as a hurdler does–jump.

The autogen is characterized by reciprocal catalysis and assembly of the shell (self-assembly), which produce and maintain the conditions necessary for their own existence. These processes, along with periodic disruption, create a cycle of mutual support and constraint that we can observe in the figure below.

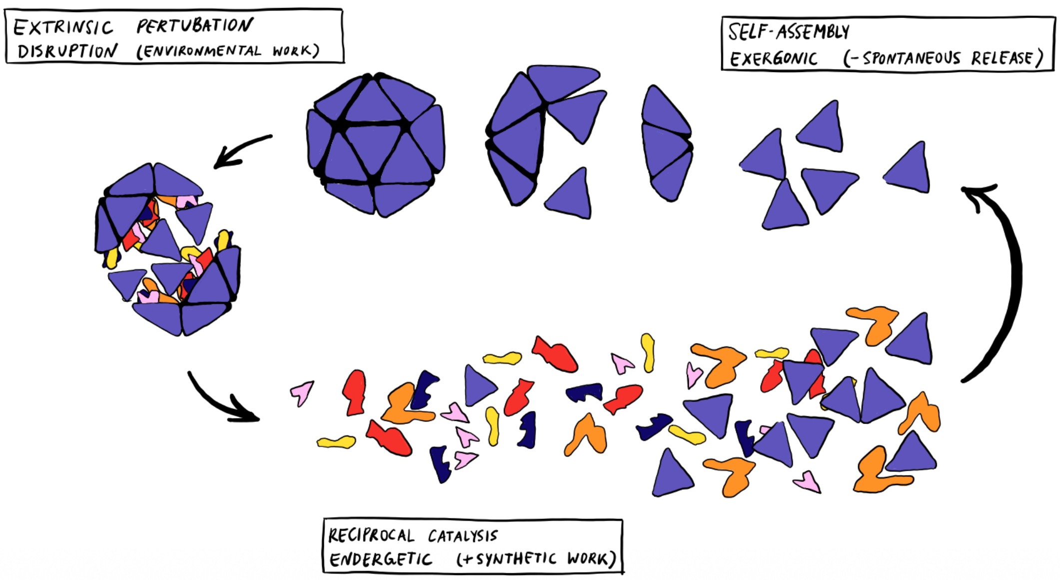

Figure 4.1: The Autogenic Work Cycle. A schematic representation of the thermodynamic tradeoffs driving the autogen. The cycle consists of three distinct phases: (1) Reciprocal Catalysis (bottom) performs positive *synthetic work*, driving the system away from equilibrium by building high local concentrations of catalysts. (2) Self-Assembly (top) represents a *spontaneous release* of free energy, where the system *pays* for its internal order by falling into a stable structural configuration (the shell). (3) Disruption (left) occurs when *environmental work* (perturbation) overcomes the stability of the shell, resetting the cycle. The interaction creates a dynamic where chemical potential is captured in structural constraints.

We can see the autogenic cycle illustrated in figure 4.1 as a workflow of thermodynamic trade-offs. This visualization helps us understand the three fundamental processes that drive the autogenic cycle:

1. Reciprocal Catalysis and Component Synthesis (Endergonic Phase): This phase is characterized by positive synthetic work. In a thermodynamic context, synthetic work refers to the energy required to drive a system away from equilibrium, constructing complex molecular arrangements against the natural tendency toward disorder. Here, the system couples the breakdown of external substrates to the production of capsid monomers, effectively acting as a pump. The diagram shows how external energy and substrate inputs drive these endergonic synthesis reactions, locally decreasing the entropy of the catalytic network as smaller substrates are assembled into more ordered capsid monomers. This process generates a high local concentration gradient—a state that is statistically unlikely and thus energetically expensive to maintain against diffusion. The catalytic network maintains high local concentrations of these specific molecular species, which is crucial for the subsequent assembly phase.

2. Self-Assembly (Exergonic Phase): Once critical concentrations are reached, the system transitions to a phase of spontaneous release (-work). While the local structure becomes highly ordered, the process is driven by the release of free energy ($\Delta G < 0$). This often occurs via the hydrophobic effect, where the *payment* for the local order of the shell is the increase in entropy of the surrounding solvent molecules as they are liberated from the surfaces of the capsid subunits. The assembly creates structural constraints through specific molecular recognition and binding interactions between capsid components. The diagram illustrates how the resulting structures form containment boundaries that prevent catalyst dissipation. Thermodynamically, the system falls into a stable energy well—a trap of its own making. This phase depends critically on the high local concentrations of components maintained by the catalytic process.

It is crucial to appreciate the functional opposition between these two phases. They represent a thermodynamic inversion. The catalytic phase acts as a *pump*, performing synthetic work to drive the system uphill against equilibrium, creating a steep concentration gradient that the universe naturally wants to disperse. The assembly phase, conversely, acts as a *trap*. It allows the system to fall downhill into a stable structural state. However, instead of falling back into disorder (diffusion), it falls into a specifically constrained order (the shell).

The system effectively traps itself. The high concentration generated by the pump forces the structural molecules to crystallize around the catalysts, physically encapsulating the reaction network before it can disperse. This is the engine of self-constraint, the system uses the energy of the first process to trigger the constraints of the second, effectively capturing the work done in the first phase before it can dissipate. The dynamic tension between pushing uphill (synthesis) and falling into place (assembly) is what allows the system to ratchet itself into existence.

3. Extrinsic Perturbation (Disruption Phase): This phase occurs after the encapsulated autocatalytic network has depleted the energy-rich substrates that originally diffused in before closure. Because the capsid wall restricts further influx, reaction rates eventually collapse. Crucially, the closed shell represents a local energy minimum, it is too stable to open on its own. Therefore, positive environmental work is required to supply the activation energy needed to overcome the stability of the shell structure. An external perturbation (mechanical shock, heat, or wave action) breaches the self-assembled structure, re-opening mass-and-energy exchange with the environment. This disruption simultaneously triggers self-repair mechanisms and creates the possibility of evolutionary change, because the incoming pool may contain novel molecular species that can be co-opted during reassembly[98].

These three dynamics map onto the tendencies of growth, disruption, and reorganization. The growth of the autogen encompasses both the endergonic synthesis of components and their exergonic self-assembly into a constraining structure. Disruption occurs when external perturbations breach this structure, leading to a phase of reorganization where the system not only self-repairs but can also incorporate novelty, setting the stage for a new cycle.

This forms the foundation of an autogenic work cycle, involving thermodynamic processes that transfer work into and out of a system, eventually returning the system to its initial state. The analogy to an internal combustion engine may be helpful here. An engine operates through precisely timed cycles of fuel injection, compression, combustion, and exhaust, each phase creating the conditions necessary for the next. The autogenic cycle comprises phases distinguished by their chemistry and thermodynamic directionality. However, there's a crucial difference, while an engine's work is directed externally to move pistons and ultimately the vehicle, autogenic work is directed inward to regenerate the very conditions that drive these changes. This self-directed work creates a form of autonomy that distinguishes living systems from machines.

Beyond autonomy, the property of proto-evolution can be understood through the third phase of reconstruction via the capacity to pass on patterns of constraints, though still very far removed in fidelity and complexity from Darwinian evolution[99]. When the system rebuilds itself, it maintains its organizational patterns while potentially incorporating new molecular components. However, it's crucial to understand that this is not reproduction in any directed sense, the propagation is essentially accidental and highly inefficient[100]. Unlike living systems with their sophisticated RNA/DNA-based inheritance mechanisms, autogens lack any reliable means of preserving and transmitting specific information across generations.

We get dynamics providing a form of adaptive stability, going beyond simple self-maintenance. When new molecules with beneficial properties are incorporated during the reorganization phases after disruption, they can initiate a slightly modified cycle with potentially enhanced adaptive properties. This process transforms the simple cycle into an adaptive spiral, where each iteration can lead to a slightly different states. Variants that better suit their environment naturally tend to replicate more frequently, creating a primitive form of selection that drives adaptation over time.

To summarize, autogenesis functions as a constraint production and preservation cycle driven by the fundamental triple of growth, disruption, and reorganization.

Growth encompasses both the catalytic synthesis of components and their self-assembly into constraining structures. Disruption breaches these structures when resources are depleted, reopening exchange with the environment. Reorganization follows as the system self-repairs while potentially incorporating novel molecular components, setting the stage for renewed growth. Through this cyclical dynamic, the system returns to a position similar to its prior state, yet the constraints of how matter is arranged and moves persist even as all physical parts can be replaced over successive cycles. This perspective foregrounds how these three distinct tendencies work together to create a self-maintaining, potentially evolving system.

4.2 Cellular Cycles

The same dynamics of growth, disruption, and reorganization, manifest in cellular systems, but their temporal expression depends crucially on two interrelated factors. First is the challenge of parsing, including how we delineate subsystems, which variables we track, and at what temporal resolution we observe them. Second is the cell's fundamental architecture which is a hierarchical arrangement of semi-autonomous modules, from metabolic pathways to organelles to protein complexes, each operating on its own characteristic timescale yet linked through sophisticated feedback mechanisms. Because these nested cycles operate asynchronously, the cell as a whole rarely exhibits a clear sequential pattern. Instead, multiple micro-cycles overlay one another, creating the appearance of simultaneous processes.

In complex cellular systems, these dynamics manifest distinctively due to the stabilizing influence of genetic information systems. Where simpler systems might undergo wholescale reorganization in response to environmental changes, cells operate within the relatively fixed constraint framework provided by their DNA. Adaptation occurs primarily through regulated changes in gene expression rather than fundamental systemic reorganization. This genetic stabilization shapes how the three dynamics operate. Growth and disruption appear most visibly as continuous, opposing processes maintaining homeostasis. Reorganization, however, functions differently:

while constant minor repairs occur continuously, true adaptive reorganization-the shifting of metabolic set-points or pathway topology—occurs primarily under stress or developmental transitions. It is only under stress or developmental transitions that reorganization shifts into an adaptive mode, retuning metabolic set-points, reallocating components, or shifting pathway topology. This arrangement allows cells to maintain robust function while retaining the capacity for bounded adaptation within their genetic constraints.

We can make this parsing explicit by examining three distinct cellular domains, first, metabolic regulation, where processes like glycogen synthesis and breakdown operate in continuous opposition. Second, structural maintenance, where membrane components and cytoskeletal elements undergo constant renewal and third, protein homeostasis, where synthesis and degradation pathways work in concert to maintain functional order. At each level, careful examination shows how cells achieve dynamic stability through the precise coordination of opposing processes, while maintaining the capacity for bounded adaptation through regulatory reorganization.

Consider how animal cells manage glucose, their primary energy source. Cells continuously cycle between storing glucose in complex forms like glycogen (a large, branched polymer used for energy storage) and breaking these storage molecules back down to release glucose for immediate use. These opposing processes—synthesis and breakdown—operate simultaneously rather than alternating. This maps directly to the tension between growth and disruption: the anabolic (building) pathways represent the tendency toward growth, while the catabolic (breaking) pathways represent the tendency toward disruption. At first glance, this appears inefficient. The simultaneous operation of opposing pathways consumes significant energy, typically on the order of a few percent of the cell's total energy budget. This seeming inefficiency serves a purpose. By maintaining both processes in continuous opposition, cells achieve a degree of control that would be impossible through simpler, unidirectional pathways.

To use an analogy: imagine trying to maintain a precise speed in a car using only the accelerator, compared to using both accelerator and brake[101]. The latter approach, while less efficient, offers far greater precision and responsiveness. This arrangement allows cells to respond rapidly to changing conditions while maintaining precise control over their internal state. Throughout cellular metabolism, we find this same pattern, including building pathways (anabolic) and breakdown pathways (catabolic) operating in coordinated opposition, creating a dynamic system capable of reproduction and adaptation.

Next, consider the plasma membrane, the protective barrier that surrounds the cell and regulates what enters and exits. A typical mammalian cell cycles the equivalent of its entire surface area every few hours, continuously retrieving and recycling lipids and proteins while inserting new ones. The coordination of these opposing processes is what allows the membrane to maintain its essential barrier function[102]. Another example of this changing dynamic can be found in the cytoskeleton, the cell's structural skeleton. Its protein filaments continuously extend and retract, build up and break down. This dynamic state, rather than compromising structural integrity, enables it. Constant remodelling allows cells to maintain their shape while retaining the ability to reorganize rapidly when needed, such as during cell division or movement.

At an even more basic level, this pattern of simultaneous construction and destruction governs the cell's entire molecular workforce. A human cell contains roughly two billion protein molecules. Every protein exists in a cycle of synthesis and degradation, with lifetimes ranging from minutes to weeks. Regulatory proteins (such as those involved in cell signalling) might last only 15-30 minutes before degradation, while structural proteins (like those in muscle fibres) persist for days. This turnover, allows cells to maintain stable protein levels while continuously renewing their molecular components[103].

These cellular processes demonstrate how the dynamics of growth, disruption, and reorganization manifest in cells. Through genetic regulation, cells have evolved to maintain organizational continuity by precisely coordinating countless micro-cycles of construction and destruction. When conditions change cells adjust these balanced oppositions within their genetic constraints. Under stress, for instance, they might increase protein degradation to remove damaged molecules while shifting synthesis toward protective proteins, the cytoskeleton might rapidly reorganize to repair structural damage, and metabolic pathways might adjust their relative rates to maintain energy balance. In each case, the reorganizational aspect manifests not as wholesale system change, but as carefully regulated reallocation and retuning of existing processes.

This coordination of opposing processes across multiple semi-autonomous modules provides our first insight into how complex living systems persist far from equilibrium. By parsing cellular organization at different scales we see how these dynamics become simultaneously present yet differently expressed. As we will see next, similar patterns appear at larger scales, though their temporal organization often differs significantly from those in cells.

4.3 Adaptive Cycle

Cells maintain stability through the coordination of simultaneous tendencies. Larger composite systems often display more distinct temporal patterns in their organization. To understand these patterns, we turn to the adaptive cycle, a model that describes how complex systems maintain themselves through time.

Complex adaptive systems (CAS) are collections of interconnected components that learn, adapt, and evolve in response to their environment. What makes these systems fascinating is that their collective behaviour emerges from the interactions between their parts, rather than from centralized control. These systems demonstrate the remarkable ability to organize, creating ordered patterns and behaviours that cannot be predicted simply by understanding their individual components[104].

To ground this abstract concept, consider some illustrative examples of CAS: Cities, markets, governments, war, terrorist networks, ecosystems, social networks, animal swarms, traffic flows or the immune system.

All examples of systems with complex dynamics involving many parts that change and adapt over time. When we attempt to model such systems, we face a significant challenge. The sheer number of variables and interactions makes it impossible to account for every detail. Instead, we must identify key variables that capture the essence of the system's behavior. This process of selecting relevant variables is both an art and a science, lying at the heart of effective modelling in many scientific disciplines.

In their quest to describe the dynamics common to various CAS, Gunderson and Holling proposed three key variables that shape how systems respond to change: Potential, Connectedness, and Resilience[105]. These variables weren't chosen arbitrarily, they came from extensive observations of regional development and ecosystem management, where they proved crucial in determining how systems respond to change[106].

Potential refers to the accumulated resources and range of future options available to the system. Think of it as the capital that a system has built up over time. In an ecosystem, this might be stored nutrients and biomass, in a social system, accumulated knowledge and relationships. This potential determines what possibilities lie ahead.

Connectedness describes how tightly coupled the system's internal components are, how much the behaviour of one part influences another. When connectedness is low, parts operate relatively independently, making the system more flexible but possibly less efficient. High connectedness means changes in one area quickly affect others, potentially making the system more efficient but also more rigid.

Resilience indicates how much disturbance a system can absorb while maintaining its essential functions. It's not a fixed property but rather expands and contracts through the cycle. A highly resilient system can weather significant disruptions while maintaining its core identity, while a system with low resilience might fundamentally transform under even minor disturbances.

These three variables interact in complex ways. As a system moves through its adaptive cycle, shifts in one variable often influence the others. For instance, as connectedness increases during the conservation phase, resilience often decreases, making the system more vulnerable to disruption despite its apparent stability.

As Holling and Gunderson studied ecosystems through the lens of these variables, they discovered a recurring pattern of transformation that moves through four distinct phases. This pattern recalls the cyclical nature of autogenic systems, though operating at vastly different scales. Where autogenic cycles orchestrate molecular transformations through chemical interactions, adaptive cycles coordinate scales up to entire ecosystems. The trajectory of systems through this three-dimensional phase space often follows a predictable path, which they dubbed the *lazy eight* due to its resemblance to the number 8 lying on its side. Each phase represents a different configuration of potential, connectedness, and resilience, creating a dynamic cycle of growth, accumulation, release, and reorganization that characterizes how complex adaptive systems change over time.

The same basic cycle appears across vastly different scales and types of systems. Whether we're looking at the succession of a forest patch, the evolution of an organization, or the development of an entire economic sector, we see this fundamental pattern of exploitation, conservation, release, and reorganization. This recurrence suggests that the adaptive cycle captures something fundamental about how complex systems organize and transform over time.

This trajectory depicted in figure 4.2 can be divided into four distinct phases, each representing a different state of the system:

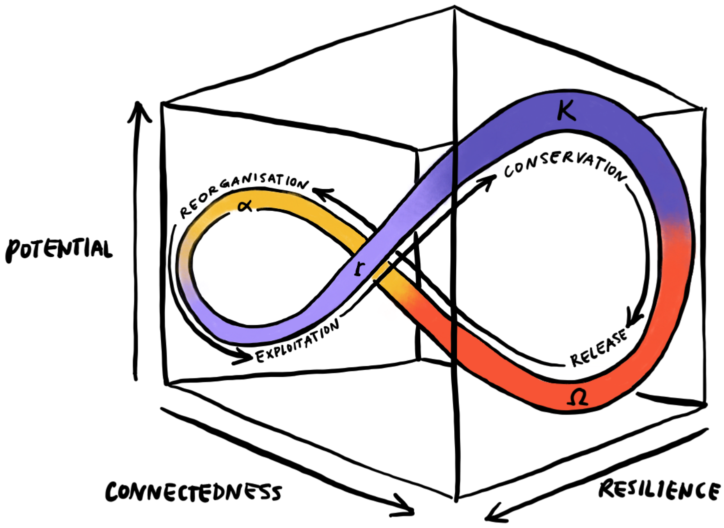

Figure 4.2: The Adaptive Cycle. This diagram illustrates the trajectory of a Complex Adaptive System (CAS) moving through a three-dimensional phase space defined by key variables: Potential, Connectedness, and Resilience. The *lazy eight* pathway shows the cyclical movement through four distinct phases. The *front loop* consists of the exploitation (r) phase, characterized by rapid growth and resource accumulation with low connectedness, leading to the conservation (K) phase, where potential and connectedness reach their peak, creating highly efficient but rigid (low resilience) structures. A disruption triggers the *back loop*, beginning with the release (Ω) phase, where constraints rupture and accumulated potential is dissipated. This leads to the reorganization (α) phase, a period of high resilience and low connectedness where novelty can emerge, laying the groundwork for a new cycle.

1. The r Phase (Exploitation): In the exploitation phase, the system is characterized by the rapid accumulation of potential as resources are abundant and actively accumulated for growth. Connectedness among system elements is low, allowing for high flexibility and autonomy, which facilitates rapid expansion and colonization. Resilience is high, enabling the system to absorb disturbances and continue functioning effectively. This phase is marked by vigorous experimentation and adaptation, as the cost of failure is low and the system can quickly recover and capitalize on new opportunities. In ecological systems, this phase is exemplified by rapid growth of pioneer species in disturbed areas, such as semiarid savannas recovering from disturbances.

These species quickly exploit available resources, showing high resilience to environmental variability. In economic systems, this phase is characterized by entrepreneurial startups aggressively capturing market shares in newly opened markets, demonstrating rapid growth and resource accumulation.

2. The K Phase (Conservation): During the conservation phase, the system's potential stabilizes as resources are fully utilized and sequestered within established structures. Connectedness increases significantly, resulting in tight interdependencies and enhanced efficiency through specialization and optimization. However, this high level of connectedness also leads to decreasing resilience, making the system more vulnerable to disturbances that can propagate easily throughout the tightly linked elements. The focus shifts from growth to maintenance and control, aiming to preserve stability and efficiency at the expense of flexibility and adaptability. In a forest ecosystem, this phase is represented by a mature forest where species diversity has peaked and connectivity among species is high. The forest becomes stable but less resilient, more vulnerable to disturbances like fires or insect infestations. In business, mature firms in this phase focus on efficiency and cost minimization, becoming more predictable but also more rigid and vulnerable to crises.

3. The Ω Phase (Release): In the release phase, the system experiences a sudden release of potential due to a significant disturbance, leading to the rapid breakdown of established structures and the dispersal of accumulated resources. Connectedness among elements decreases sharply, causing a loss of control and an increase in chaos within the system. Resilience is at its lowest point, rendering the system highly susceptible to further shocks and disturbances. Despite the crisis, this phase also creates opportunities for new structures and processes to emerge, setting the stage for eventual renewal and innovation. In ecosystems, this phase can be triggered by disturbances like fires or insect outbreaks in mature forests, rapidly releasing accumulated resources and breaking established connections. In economic systems, this phase might be initiated by external shocks like economic downturns or leadership changes, leading to layoffs, restructuring, and a temporary drop in potential.

4. The α Phase (Reorganization): The reorganization phase sees the system beginning to regain its potential as resources are reassembled and reallocated, fostering the emergence of new structures and opportunities. Initially, connectedness is low, promoting a wide range of experiments and innovations. As new relationships stabilize, connectedness gradually increases, forming more flexible and adaptive connections compared to the previous conservation phase. Resilience starts to rise, enhancing the system's capacity to absorb disturbances and support the development of novel solutions and adaptations. This phase is critical for laying the groundwork for future growth and ensuring the system's long-term sustainability. In ecological contexts, this phase is exemplified by the reorganization that occurs after major disturbances, such as the retreat of ice sheets 15,000 years ago. During this time, individual species moved at their own rates, establishing themselves in new environments and forming novel associations. In economic systems, this phase is characterized by the emergence of new entrepreneurial ventures and innovations following a crisis, setting the stage for a new cycle of growth and development.

By repeatedly going through phases of growth (r & K Phase), disruption (Ω-Phase), and reorganization (α-Phase), systems can persist and evolve over time, adapting to changing environments while maintaining their overall integrity. The cycle allows for periods of efficient resource exploitation followed by periods of creative destruction and reorganization, enabling the system to explore new configurations and adapt to new challenges.

The adaptive cycle describes how complex systems manage opposing tendencies across time through distinct phases, rather than the more simultaneous coordination we observed in cellular systems. Through this cyclic pattern of growth, disruption and reorganization systems can both build and preserve structure while maintaining the capacity for renewal and adaptation.

4.4 Mapping

Let's begin our mapping process by examining the structure underlying the autogenic, cellular and adaptive cycle, a tripartite pattern of growth, disruption, and reorganization. This three-fold structure, while manifesting differently across scales, provides a framework through which we can understand the dynamics that allow systems to maintain themselves. Our focus is on how these three tendencies are integrated into a functional whole, where even the system's disruption becomes an central part of its self-maintaining constraints.

For a minimal system like the autogen, the growth-disruption-reorganization tendencies are relatively clearly sequenced. At the cellular level, these tendencies operate simultaneously and continuously, with opposing processes like anabolic and catabolic pathways running in parallel. In larger-scale adaptive systems they can again be identified as temporally distinct phases. In all three systems we observe, a growth phase where the system builds up resources and structure. A disruptive phase where external shocks create opportunities for adaptation by disrupting existing structures. And third, a reorganization phase where the system reassembles, potentially incorporating new elements that enhance its fit with the environment.

The autogenic system shows this through: (1) the coupled processes of reciprocal catalysis and self-assembly that build up the system, (2) the rupture of the shell through environmental perturbation, and (3) the reassembly process that may incorporate new molecular species. In cellular systems, this manifests through: (1) the continuous operation of anabolic pathways building up complex molecules and structures, (2) simultaneous catabolic processes breaking down components, and (3) the constant integration and regulation of these opposing processes, allowing for adaptive shifts in function when conditions change. In complex adaptive systems, we see: (1) the accumulation of resources and connections, (2) the disruption of established structures through external shocks, and (3) the reorganization that allows for new forms of order to emerge.

While this three-fold structure captures the essential dynamics, examining how these processes can be grouped differently highlights additional aspects related to the organizational challenges these systems face.

From this three-fold structure, we can examine how these cycles are elaborated in different contexts. In the case of the adaptive cycle, Holling and Gunderson's chapter on the *Front Loop/Back Loop: Embracing Opposites* groups the cycle into two main parts: the front and back loop. The front loop represents the system's growth from r to K, characterised by increasing potential and connectedness. In contrast, the back loop depicts the system's rapid reorganization and renewal as it moves from Ω to α. This simplified view helps us grasp the cycle's overall dynamics more easily.

This two-part structure aligns with how the autogenic system was initially described, as the coupling of reciprocal catalysis and shell-formation processes. This coupling forms the autogenic work cycle through which a system emerges that produces self-constraining dynamics. Dynamics that produce their own boundary conditions, keeping the system from depleting an available energy gradient as quickly as possible.

Holling and Gunderson write: *It is as if two separate objectives are functioning, but in sequence. The first maximizes production and accumulation; the second maximizes invention and reassortment. We have no theorem to prove it, but our intuition suggests that any complex system, if it is adaptive, must generate these two phases in sequence, at some scale. The two objectives cannot be maximized simultaneously; they can occur only sequentially. And the success in achieving one tends to set the stage for its opposite. The adaptive cycle therefore embraces the opposites of growth and stability on the one hand, change and variety on the other.*

The above statement presents us with a tension for a system like a cell if, as Holling and Gunderson write, the two phases need to be generated in sequence. A crucial scaling principle can help resolves this: the simultaneous processes we observe in systems like cells actually emerge from multiple sequential cycles operating at different scales and phases. Whether growth, disruption, and reorganization appear as cleanly separated phases or as overlapping tendencies depends on the coarse-graining window[107]. In a single mammalian cell, thousands of build-up and break-down reactions run continuously and largely out of phase with one another. When we, for example, measure total ATP turnover, we find a relatively flat metabolic envelope[108]. Yet if we track a higher-order variable (e.g.: chromosome ploidy, membrane surface-area per volume, or cyclin concentration) we can recover a phased macro-cycle whose successive regimes correspond to growth, systemic structural release, and reassembly[109]. The same nested structure recurs at higher scales like the adaptive cycle. Thus *simultaneity* at one resolution can mask phased dynamics that become visible at another, a consequence of superposing many asynchronous micro-cycles.

Consider a forest during its apparent *growth phase*. Examining a patch of forest floor reveals an old dead branch on which moss and fungi grow, a local growth phase for the fungi happening within a disruption phase of the tree, occurring within the larger growth phase of the forest. Similarly, for a continuously dividing cell, it becomes difficult to see a specific cycle because we observe only division after division. But shifting to a higher-level perspective of the lineage as a whole, we can identify the growth phase in the presence of energy sources, stagnation as available energy decreases, population collapse from overgrowth, and survival of well-adapted variations. The key insight is that looking at any complex system, many growth-disruption-reorganization processes operate simultaneously, but examining the constraint cycle from the appropriate scale allows us to separate them out and identify the fundamental pattern.

However, this two-part conceptualization, while useful, obscures a crucial element. The role of disruption as a distinct phase. In both frameworks, the disruptive moment is treated as external to the system, a perturbation that triggers the transition from growth to reorganization. Yet, treating this shock as purely external misses the point. The vulnerability to disruption is an inherent property of the system. The disruption functions as an essential internal mechanism, a necessary *opening* through which the system interfaces with and adapts to its environment.

The four-phase model of the adaptive cycle (r, K, Ω, α) maps onto this tripartite structure, with one crucial elaboration. It splits the Growth phase into two distinct regimes. The *Front Loop* of the adaptive cycle encompasses both Exploitation (r) and Conservation (K). Both are constructive phases where the system builds structure and accumulates potential. The difference lies in their constraints: r-growth is loose and rapid, while K-growth is rigid and slower. The Disruption tendency maps directly to the Release (Ω) phase, where the system's accumulated potential is discharged and structural constraints rupture. Finally, the Reorganization tendency maps to the Renewal (α) phase, where the system navigates uncertainty to establish the parameters for a new cycle. Thus, the *Lazy-8* is simply a high-resolution rendering of the fundamental Growth-Disruption-Reorganization logic.

This mapping highlights why the transition from Conservation (K) to Release (Ω) is so critical. As a system matures in the Growth phase, it becomes increasingly interconnected and rigid. It builds up complicated structure where a single perturbation like a lightning strike, a market crash, or a new pathogen, can trigger a systemic release. This aligns with the autogenic model, where the shell eventually becomes a barrier to further metabolism, creating the fragility that makes disruption not only inevitable but necessary.

Critically, what appears as destructive actually represents the system's crucial interface with its environment. When an old-growth ecosystem experiences fire, the destruction opens opportunities for interaction with the broader environment. Seeds that couldn't previously germinate in the dense shade now find suitable conditions. As climate change increases the frequency and intensity of fires in many regions, these destructive events become mechanisms for larger-scale adaptation. Pyrophytic species, which are plants specifically adapted to fire-prone environments, can establish themselves and gradually transform the ecosystem. Through repeated cycles of disruption and renewal, a forest that was previously vulnerable to fire might be gradually replaced by one dominated by these fire-adapted species. The system is actively engaging with changing climate environment in ways that were impossible in its previous, more rigid state.

In the autogenic system, we see this same pattern. While Deacon emphasizes the reciprocal relationship between catalysis and self-assembly, the disruptive phase, the breakage of the shell emerges as a distinct and necessary third component. This disruptive phase isn't merely an interruption of the constructive processes but rather a vital element that enables the system's adaptive potential. Without this phase of disruption and subsequent reorganization, the system would lack the capacity to incorporate new elements and adapt to environmental changes. The disruption creates the conditions necessary for novelty to enter the system, making it not just self-maintaining but genuinely adaptive.

These examples illustrate how the disruptive phase serves as the enabling condition for a system to become adaptive to its environment. By definition, adaptation implies a prior state of non-adaptation. A mismatch between system and environment that must be resolved. The disruptive phase creates the conditions for this resolution by allowing new elements to enter and influence the system's reconstruction, but it is the reorganization phase that actively determines whether the system becomes better adapted. In the autogenic cycle, for instance, the breakage of the shell allows new molecular species to participate in the catalytic processes, but whether this actually improves the system's environmental fit depends heavily on how the system reorganizes itself during reassembly.

This is crucial, since disruption enables adaptation by opening the system to novelty, but the reorganization phase is where environmental information is actively incorporated and integrated. The system learns not simply through disruption but through the specific ways it reconstructs itself following disruption. This perspective transforms our understanding of the disruptive phase from a *negative* event to a crucial enabling condition for environmental interaction and learning. The disruption creates the opening for new forms of organization, but whether these better match environmental conditions depends on the reorganization principles governing the system's reconstruction. This is why systems that never experience disruption often become increasingly maladapted. They lack this crucial enabling condition for environmental engagement, even though disruption alone cannot guarantee successful adaptation.

The wholeness of the system thus includes its breakdown, its death, and its demise as integral components of its adaptive capacity. Not as the source of adaptation itself, but as the necessary enabling condition without which genuine adaptation cannot occur.

The fact that these frameworks implicitly contain this three-part structure suggests its fundamental importance. While Holling and Gunderson elaborate their cycle into four phases and Deacon focuses on the reciprocal relationship between catalysis and self-assembly, both systems ultimately rely on the interplay of growth, disruption, and reorganization to achieve genuine adaptation.

Rather than forcing a direct correspondence between all four phases of the adaptive cycle and specific moments in the autogenic cycle, we can see how both systems embody the essential pattern of growth, disruption, and reorganization, with the four-phase model providing additional resolution into how these fundamental processes unfold.

This scaling up of the fundamental dynamics observed in the autogenic system to higher-level complex adaptive systems is not a coincidence, rather an expected outcome if self-constraining dynamics are indeed central to living systems[110]. In other words, if self-constraining dynamics are fundamental to life, we would expect to see similar patterns manifesting at various scales and levels of complexity. The adaptive cycle, observed in ecosystems, economies, and other complex systems, can thus be viewed as a higher-order expression of the same underlying principles that govern the autogenic cycle.

The mapping demonstrates that what remains constant across scales is the fundamental principle. Living systems coordinate countervailing tendencies to maintain themselves while remaining capable of adaptation, with their own partial destruction integrated as an essential component of this coordination.

Table 4.1: Comparative Mapping of Cycles to Growth, Disruption, and Reorganization

Cycle	**Autogenic**	**Cellular**	**Adaptive**
Core Idea	Fundamental cyclical processes of constraint production, maintenance, & renewal at a basic molecular level.	Continuous, simultaneous management of opposing tendencies for dynamic stability & renewal at molecular/organelle levels.	Temporally distinct phases (exploitation, conservation, release, reorganization) managing opposition.
Growth	*Catalytic & Assembly Phase:* **Synthetic Work** builds chemical potential, followed by **Spontaneous Release** into the constrained shell.	*Anabolic pathways* (e.g., glycogen/protein synthesis); membrane insertion; cytoskeleton extension. Accumulation & building.	*r-Phase (Exploitation):* Rapid resource accumulation, high potential. *K-Phase (Conservation):* Resource sequestration, structural elaboration, high connectedness.
Disruption	*Breach:* **Environmental Work** (perturbation) overcomes shell stability, releasing sequestered catalysts.	*Catabolic pathways* (e.g., glucose/protein breakdown); membrane removal; cytoskeleton retraction.	*Ω-Phase (Release):* Sudden breakdown of structures, resource release.
Re-org.	*Reconstitution:* Self-repair of capsid, restart of autocatalytic loop. Potential incorporation of novel molecules, leading to variant lineages.	*Adaptive Response:* **Shifts in gene expression and metabolic set-points in response to stress.** Integrating growth products and accommodating disruption.	*α-Phase (Reorganization):* Resource reassembly, new structures/relations emerge, innovation, groundwork for new growth.

The mapping we've proposed offers a starting point for exploring how basic principles of self-organization and cyclical dynamics might scale up from molecular systems to ecosystems and beyond. The fundamental value of this comparison is the insights into how basic principles of self-organization and adaptation manifest across different scales.

4.5 The Constraint Cycle

Our analysis of the autogenic work cycle, the cellular cycle, and the adaptive cycle foreground a recurring pattern of dynamic organization that facilitates both persistence and significant adaptation across various scales. This pattern is termed the *Constraint Cycle*.

A Constraint Cycle describes a temporally differentiated adaptive dynamic observed in many complex systems[111]. It manifests as a recurring sequence of broadly distinct phases or dominant tendencies:

A Growth Phase: Characterized by the accumulation of resources, elaboration of structure, and increasing connectedness. This phase often leads to increased efficiency but potentially reduced flexibility.

A Disruption Phase: Triggered by internal pressures or external shocks, this phase involves a temporary loss of established coherence, breakdown of existing structures, release of bound resources, and increased permeability of the system to environmental influences or internal novelty.

A Reorganization Phase: The system actively reconfigures itself, potentially incorporating new elements or information encountered during the disruptive phase. This is where novel solutions are explored, and new, potentially more adaptive, coherences are established, setting the stage for a subsequent cycle.

While the component processes of growth, breakdown, and integration may operate continuously and simultaneously at certain levels (as in cellular metabolism), the Constraint Cycle framework focuses on the distinct phases when the system is viewed at a coarser temporal or organizational scale.

The Constraint Cycle is a model for understanding how systems can navigate major changes and achieve robust adaptation through periodic, relatively large-scale restructuring, distinct from continuous, fine-grained adaptive adjustments.

Systems exhibiting a Constraint Cycle dynamic typically display several characteristic features:

First, each phase sets the conditions for the next in a self-reinforcing sequence. Growth leads to conditions ripe for disruption through increasing rigidity, resource depletion, or accumulated internal tensions; disruption makes reorganization both possible and necessary by breaking down established structures and releasing bound resources; reorganization establishes new foundations and coherences that enable a subsequent growth phase. This interdependence creates a self-perpetuating cycle where each phase emerges naturally from the dynamics of the preceding one.

Second, the system's capacity to achieve qualitatively different adaptive states emerges primarily through the coordinated interplay between disruption and reorganization phases. During disruption, the system becomes highly permeable to environmental influences and internal novelty, creating critical windows of heightened environmental interaction. The subsequent reorganization phase then integrates the "lessons" from this disruptive interface, incorporating new elements or information to establish novel coherences that potentially transcend the capabilities of individual components. This managed alternation between relative closure and pronounced openness allows the system to be both buffered and responsive to environmental variations.

Third, the phases differ significantly in their duration, intensity, and dominant processes, creating an asymmetric temporal structure. This asymmetry, particularly the contrast between gradual accumulation and rapid reorganization, generates pivot points for adaptive transformation. Coupled with successful reorganization, this creates a "ratchet effect" that allows the system to attain new levels of complexity or adaptation over multiple cycles rather than simply returning to previous states. The temporal asymmetry thus enables genuine developmental progression through cyclical dynamics.

Fourth, is the integration of oppositional forces. In a Constraint Cycle, qualitatively new adaptive states arise only when opposed tendencies are composed into a new constraint regime. Growth concentrates resources and tightens couplings, disruption releases bound resources and exposes new degrees of freedom through internal novelties and environmental inputs, reorganization performs the composition, selecting what to conserve from growth and what to incorporate from disruption, and binding them into operative constraints that stabilize previously unavailable behaviours. Through this coupling the system gains access to regions of state space unreachable under the prior constraints—i.e., a higher-order organization that no single tendency could produce on its own.

It's important to note that the Constraint Cycle is not proposed as the defining feature of life or the sole mechanism of adaptation. Rather, it represents one model for understanding how certain systems, can achieve significant adaptation through periodic restructuring. This model is especially relevant for pre-genetic systems and complex, multi-level organizations where continuous, gradual adaptation may be insufficient to address major environmental changes or internal pressures.

The Constraint Cycle can be viewed as a temporally differentiated special case of a more general *cyclical synthesis*—the constant interplay of build-up and break-down processes that characterize living systems. This will be discussed in the coming chapters. While this ongoing interplay is fundamental to life (as exemplified by cellular metabolism), the Constraint Cycle describes a pattern where these processes become more distinctly phased and sequentially dominant at particular scales of observation.

In the context of life's origins and early evolution, the Constraint Cycle model helps explain how simple self-maintaining systems like autogenic protocells could have developed increasingly sophisticated adaptive capabilities. It suggests a pathway from basic self-maintenance to more complex forms of adaptation that prefigure the robust evolutionary processes of modern cellular life.

Thus, the Constraint Cycle concept offers more than the description of a particular adaptive dynamic. It functions as a heuristic lens through which we can identify and compare fundamental strategies for persistence and transformation across diverse systems and scales. By highlighting a shared underlying pattern of growth, disruption, and reorganization, the framework provides a conceptual pathway for exploring how the logic of self-maintenance and adaptation scales. This ability to bridge scales by focusing on a core organizational strategy is crucial for developing a more unified understanding of how complex systems, navigate the challenges of existence and change.

Finally, the Constraint-Cycle lens does not insist that every system, under every sampling scheme, will resolve cleanly into growth, disruption, and re-organization. In regimes where resource inputs or stressors change only glacially, like a deep-sea microbial mat drawing on near-constant chemical gradients the differential between tendencies may be so slight that no coarse-graining window exposes distinct phases[112]. What the model claims is more modest: if a system does exhibit prolonged accumulation followed by episodic release and restructuring, those dynamics will hinge on the underlying opposition and coordination of constructive and deconstructive constraints. Where the signal is too faint to tease apart, the framework simply stays silent rather than forcing a cycle that isn't there.

4.6 Limitations

Mapping a single dynamic onto systems as diverse as autogens, cells, and ecosystems inevitably creates conceptual tension. Smoothing over the specific mechanics of distinct systems to underscore a common pattern, risks decoupling this model from the biological grit it seeks to explain. Before relying on this framework, we must confront the friction that arises when we map a dynamic onto the specific, divergent realities of evolution.

The constraint cycle concept, while helpful as an explanatory tool, becomes increasingly abstract as we move from molecular systems to ecosystems and beyond. This presents both an opportunity and a challenge. On one hand, abstract models allow us to identify common patterns across vastly different domains, revealing deep structural similarities that might otherwise remain hidden. On the other hand, this very abstraction risks stretching the concept to the point where it loses specific explanatory power, especially if characteristics of the foundational autogen model are applied too broadly to systems with more evolved adaptive mechanisms.

This conceptual stretching reveals a fundamental tension in theoretical biology, the trade-off between conceptual unity and domain-specific accuracy. For instance, applying the specific disruption-reorganization dynamic of a basic, inheritance-lacking autogen directly to a proto-cell or a modern cell, which possess sophisticated repair and regulatory mechanisms, would be an oversimplification. The strength of the constraint cycle framework lies in identifying a general pattern of growth, disruption, and reorganization, while acknowledging that the mechanisms and nuances of these phases differ significantly across scales and levels of biological complexity.

Perhaps the most significant challenge lies in defining appropriate boundaries - both spatial and temporal - for constraint cycles at different scales. At the cellular level, physical membranes provide relatively clear spatial boundaries, while metabolic cycles establish temporal patterns. But as we move to ecosystems or social systems, these boundaries become increasingly ambiguous. Where does one ecosystem end and another begin? When does an adaptive cycle complete one rotation and begin another? These questions become particularly challenging when dealing with nested cycles operating at different scales and speeds.

This boundary problem reflects a deeper issue in complex systems research, that is the recognition that boundaries are themselves emergent properties rather than fixed features. In living systems, boundaries are actively maintained through the very processes they contain, a self-referential dynamic that resists simple demarcation. Far from undermining the constraint cycle framework, this limitation points toward one of its key insights, namely that living systems achieve coherence not through fixed structures but through dynamic processes of self-constraint across time. The challenge of boundary definition thus becomes not just a limitation to overcome but a phenomenon to understand more deeply.

These limitations suggest promising directions for future research rather than fundamental flaws in the approach. Developing more empirical tests for autogenic principles, refining our understanding of cross-scale dynamics, and deepening our analysis of boundary formation processes would all strengthen the constraint cycle framework. Future research should also focus on better delineating how the nature of disruption and reorganization transforms across scales from the passive, environment-driven breach in a theoretical autogen to the actively regulated stress responses, repair, and adaptive mechanisms in evolved living cells and ecosystems. Acknowledging these limitations helps us use the framework more precisely, recognizing when and where it provides genuine explanatory power while remaining cautious about overextension.

In the chapters ahead, we'll build on these foundations while remaining attentive to their limitations. As we explore the implications of constraint cycles for cognition, knowledge, and technology, we'll see how these same patterns of limitation and possibility shape our understanding of ourselves as knowing beings. The very features that make it challenging to define a universal Constraint Cycle with precision across all scales, such as the dynamic, self-referential, and context-dependent nature of growth, disruption, and reorganization, point to why these underlying cyclical processes are so fundamental to the persistence and adaptation of living systems and potentially to the emergence of mind.

Chapter 5

Living in Time

For everything there is a season,
a time for every activity under heaven:
A time to be born and a time to die.
A time to plant and a time to harvest.
A time to kill and a time to heal.
A time to tear down and a time to build up.
A time to cry and a time to laugh.
A time to grieve and a time to dance.
A time to scatter stones and a time to gather stones.
A time to embrace and a time to turn away.
A time to search and a time to quit searching.
A time to keep and a time to throw away.
A time to tear and a time to mend.
A time to be quiet and a time to speak.
A time to love and a time to hate.
A time for war and a time for peace.
— Ecclesiastes 3:1–8

These words form one of the most enigmatic texts in Western religious tradition - the third chapter of Ecclesiastes. Its rhythmic enumeration of temporal contradictions has long puzzled theological orthodoxy, which tends to prefer its truths absolute and its commandments unconditional. How, then, did this meditation on cyclical change and temporal contradiction find its place among the canonical texts?

In a tradition defined by strict decree and definitive doctrine such as *Thou shalt not kill*, how are we to understand a text that seems to suggest even killing has its proper time? The tension is so apparent that Martin Shields, in his work on Ecclesiastes, can only conclude: *In short, we do not know why or how this book found its way into such esteemed company.*[113]

Contemporary biblical scholarship has proposed an elegant solution to this tension. Perhaps, they suggest, Ecclesiastes represents not a single voice but a multiplicity of competing perspectives. A symposium preserved in sacred text[114]. By attributing different sections to different authors, each with their own stance, the apparent contradictions resolve into a dialogue across viewpoints.

This seemingly clever interpretative move only shifts the puzzle from one level to another. Instead of temporal contradictions within a single voice, we now have perspectival multiplicity across voices. The text still resists reduction to a single, consistent viewpoint. But perhaps this very resistance is itself the key to understanding not just Ecclesiastes, but something fundamental about the nature of life itself.

There's a fascinating metaphoric parallel between the structure of Ecclesiastes and the constraint cycle we've explored in the last chapter. Both embrace a view of reality that coordinates seemingly contradictory states into a larger whole, where higher-level constraints emerge through the coupling of lower-level processes. They suggest that a system can embody seemingly opposing states at different times while remaining through that integration, paradoxically, the same system.

Within a living cell, for example, we see the constant, dynamic interplay of anabolic pathways, which construct complex molecules and build structure, and catabolic pathways, which break them down to release energy and raw materials. From a purely energetic or material flow perspective, these processes might seem to head in contrasting directions: one towards increased order and energy storage, the other towards decreased order and energy release. While deeply synergistic and essential for the cell's survival they are distinct modes of operation. Each, a different *season* of cellular activity. It is through the coordination of seemingly divergent, yet ultimately complementary activities, that the system's identity emerges. It maintains itself precisely by integrating these distinct, yet interdependent, operations across time.

This organizing principle, manifests with fidelity across levels of biological organization[115]. While at the most immediate level of cellular metabolism, processes like synthesis and degradation occur continuously and in a tightly coupled, complementary fashion, we also observe the constrained integration of distinct, sometimes directionally contrasting, tendencies through time. Ecosystems, for instance, cycle between periods of growth and disruption, each phase necessary for the system's long-term persistence. Societies oscillate between periods of stability and transformation. Even within our own bodies, while minute-to-minute renewal involves simultaneous build-up and break-down, larger physiological cycles, like developmental stages, circadian rhythms, or the phased response to injury and healing, express this principle of temporally managed, distinct functional states.

The verses of Ecclesiastes are a poetic articulation that each phase of the cycle appears to contradict the others, yet all are necessary for the maintenance of the whole. While we have looked at these aspects before, the attempt here is to foreground how systems can transform themselves into states they are not yet in, states very different from current ones and often countervailing, in opposition or contradiction with others. This is a dynamic that would seem even nonsensical if viewed from a single, static moment in time.

In this light, Ecclesiastes' presence in the biblical canon transforms for us from puzzle to paradigm. Like a butterfly's continuity across form, which would appear impossible if viewed only from the perspective of either caterpillar or chrysalis, the text expresses something about the nature of coherence, namely, that certain forms of order require not the elimination of contradiction, but its integration through time.

5.1 Teleology

The *Genesis* part was an exploration of emergence of life, and in the previous chapter we examined how constraints interlock on multiple scales to form constraint cycles. This chapter narrows its focus to the constraint cycle, exploring how a system's current state relates to its potential future states.

To address this, we turn to teleology, derived from the Greek *telos* (purpose or end) and *logos* (study or reason), which examines the role of purpose and goals in nature[116]. The concept of teleology has been central to understanding living systems for millennia, yet since the 17th century, it has maintained, as aptly described by J.B.S. Haldane as an uneasy relationship with biology: *Teleology is like a mistress to a biologist: he cannot live without her but he's unwilling to be seen with her in public.*

Like many foundational concepts in Western thought, the systematic treatment of teleology begins with the Greeks, specifically with Aristotle[117]. He proposed four fundamental forms of causes or, perhaps better translated, reasons or *becauses*[118]. The material cause refers to the physical substance from which something is made, like the wood of a table. The formal cause describes the pattern or form that shapes its design, such as the arrangement of four legs supporting a flat surface. The efficient cause identifies the source of change or motion, like the carpenter crafting the table. Lastly, the final cause represents the purpose or end for which something exists. The table's function of providing a surface for writing.

Among these four causes, the final cause concerning purpose and ends, proved particularly powerful in Aristotle's framework of causal pluralism. While all four causes helped explain both artificial and natural phenomena, final cause offered unique insights. For artificial objects like tables, the final cause was straightforward, they were created with specific purposes in mind. More profoundly, final cause appeared to illuminate the inherent purposiveness of the natural world, particularly living systems. A heart's purpose to pump blood, an eye's purpose to see, or a tree's purpose to produce fruit. All seemed to follow naturally from this teleological perspective. The framework's endurance for nearly two millennia wasn't only due to institutional inertia, its emphasis on final cause provided genuine explanatory power for understanding what appeared to be the goal-directed nature of life[119].

The Scientific Revolution of the 16th and 17th centuries, however, brought about a radical transformation in our understanding of causation[120]. The mechanistic philosophies of Descartes and others preconceived the universe as a vast machine operating according to fixed laws. In his influential *Principles of Philosophy*, Descartes explicitly rejected the search for final causes in natural philosophy, arguing that we cannot presume to know God's purposes in creating natural phenomena. This rejection of teleology was reinforced by his strict separation of mind (*res cogitans*) from matter (*res extensa*), confining purpose and intentionality to the realm of consciousness alone.

Newton's subsequent success in explaining physical phenomena through mathematical laws further marginalized teleological thinking. His mechanics could account for all motion and change through the interaction of matter and force alone, without reference to purposes or goals. The planets didn't orbit the sun *for* any purpose, they simply followed the mathematical necessity of gravitational forces. This mechanical philosophy proved so successful in physics that it became the model for all scientific explanation[121].

This mechanistic triumph created a central tension in biology. Living systems seemed to resist purely mechanical explanation. They exhibited apparent purposiveness in their development, behavior, and self-maintenance. Yet the demonstrable power of mechanical explanation made scientists deeply suspicious of any appeal to purposes or goals in biological explanation. This tension reached its apex with Darwin's theory of evolution by natural selection, which appeared to eliminate the need for teleological thinking entirely[122]. From the Darwinian perspective, organs and traits don't exist *for* anything, they persist simply because they've proven advantageous for survival and reproduction. The intricate structure of the eye emerges not from any purposeful design but through countless small, random mutations filtered by selective pressures.

The tension around teleology manifests differently across scientific disciplines. For physicists, the problem rarely arises since their objects of study can mostly be described through efficient causes and mathematical laws. A falling apple or an orbiting planet requires no appeal to purpose. But biologists face a unique challenge. Their subjects persistently appear to act *for* something. In the quest for scientific rigor, and perhaps influenced by the remarkable success and universality of physical laws (a phenomenon sometimes leading to what's been dubbed *physics envy* in other fields), many biological sciences adopted a strongly mechanistic approach. This involved attempting to purge purposive explanations, aiming for descriptions based solely on material and efficient causes, similar to physics. Yet this purging has proven difficult in biology.

The resistance to eliminating teleological concepts from biology isn't solely a linguistic convenience. Beyond the persistent use of purposive language in describing biological functions, the nature of living systems seems to demand teleology. Consider how organisms actively maintain their far-from-equilibrium state, how they develop according to reliable patterns, or how they behave in ways that promote their survival. These phenomena appear irreducibly purposive. The systems *function* to maintain themselves, they *resist* degradation, the heart *pumps* blood to supply oxygen, plants *produce* oxygen through photosynthesis.

For the biologist, this creates several challenges. First, teleological explanations seem to imply foresight or intentionality in nature, which appears to conflict with materialistic explanations of life. How can the future be causally relevant for the present? Second, these explanations often lack clear mechanisms. Saying something exists *for* a purpose doesn't explain how it came to be. Attributing the development of feathers in theropod dinosaurs to a future purpose like flight suggests a form of backward causation that science rejects. Finally, there's the concern that teleological thinking might lead to complacency, with researchers accepting *purpose* as an explanation rather than investigating underlying mechanisms.

Yet the stubborn persistence of teleological concepts in biology suggests they might serve an essential role in our understanding of living systems. Despite its apparently strange temporal aspects, teleology seems to capture something fundamental about the class of living systems, something that purely mechanical explanations miss.

This realization has led to various attempts to naturalize or reframe teleology in ways compatible with modern science, efforts that might help resolve the tension between mechanical causation and biological purpose.

Following in their footsteps, we will trace these attempts to reconcile purpose with natural science, examining how each approach offers valuable insights for understanding biological teleology.

Our exploration of teleology will proceed through three major theoretical frameworks[123]. We'll begin with cybernetics, which sought to explain purposive behaviour through feedback mechanisms. We'll then re-examine autopoiesis theory, which proposed that purpose emerges from the self-producing nature of living systems. Finally, we'll consider the Organizational Approach and Deacon's work which attempts to ground purpose in the closure of constraints. Each framework builds on and responds to the limitations of its predecessors, leading us toward a more complete understanding of biological purpose.

Cybernetics

The first systematic attempt to resolve the tension between mechanical causation and biological purpose emerged in the mid-20th century with cybernetics. This interdisciplinary field, pioneered by figures such as Norbert Wiener, Ross Ashby, Warren McCulloch, and Heinz von Förster, proposed that goal-directed behaviour could emerge from purely mechanical processes through feedback mechanisms and circular causality[124]. While initially heralded as a revolutionary approach to understanding biological organization, the ultimate value of this framework would prove more contentious than its early proponents anticipated[125]. Rather than choosing between mechanical and vitalistic explanations of life, cybernetics suggested a third way, that attempts to bridge the gap between physics and biology that had widened since Descartes.

The key insight of cybernetics was that purposive behaviour could arise from certain patterns of organization, particularly feedback loops, without requiring any mysterious vital forces. A thermostat provided the paradigmatic example, through a simple feedback mechanism, it exhibits goal-directed behavior, maintaining a target temperature through purely mechanical means.

The re-conceptualization of purpose through feedback and control proved fertile. Cyberneticists showed how similar principles could explain goal-directed behaviour across different domains. From cellular homeostasis to neural processing, from ecological dynamics to social systems, circular causality seemed to offer a unified framework for understanding purposive behaviour. Norbert Wiener captured this universality in his seminal work, arguing that the same principles govern *communication engineering, whether in the metal or in the flesh*[126].

However, cybernetics' apparent resolution of the teleological problem proved incomplete. Hans Jonas, in his penetrating critique, identified a crucial limitation. While cybernetics could explain how systems maintain states or reach goals, it couldn't account for why these goals exist in the first place. A thermostat maintains a temperature set by its designer, but living organisms generate and maintain their own goals. Jonas argued that cybernetics couldn't distinguish between this extrinsic purposiveness of artifacts and the intrinsic purposiveness characteristic of life.

The shortcoming become what Jonas called *needful freedom*, the unique way living systems are simultaneously dependent on their environment yet capable of independent action within it[127]. A thermostat maintains a temperature, but it doesn't need that temperature in any meaningful sense. It lacks what Jonas called *concern* - it doesn't strive to maintain its own existence, doesn't have interests of its own. Living systems, by contrast, actively work to maintain the conditions of their own existence. They exhibit what we might call self-generated purpose.

This distinction points to a deeper problem which is that cybernetics described the *mechanism* of goal-directed behavior without explaining the *origin* of the goals themselves. It could explain how a system maintains a particular state but not why it maintains that state rather than another, or why it maintains any state at all. We might say that cybernetics described the implementation of purpose without explaining its origin.

This recognition of cybernetics' limitations, particularly its inability to account for the origin of goals, spurred the development of second-order cybernetics (SOC). Pioneered by figures like Heinz von Förster, Margaret Mead, and Gregory Bateson, and with Humberto Maturana and Francisco Varela also contributing significantly to this intellectual current before fully developing autopoiesis, SOC fundamentally shifted the focus from *observed systems* to *observing systems*. The core insight of SOC was that the observer is not a detached, objective entity but an integral part of the system being studied. Consequently, any description of a system, especially its perceived goals or purposes, inevitably reflects the characteristics, perspectives, and interactions of the observer. As von Förster famously put it, objectivity in the traditional sense was revealed as *the delusion that observations could be made without an observer.*

This *cybernetics of cybernetics* aimed to address the problem of intrinsic purpose by turning the lens of cybernetic analysis back onto the process of observation and goal-setting itself. If the goals of first-order cybernetic systems were set externally (e.g., by a human designer), SOC explored how goals might arise within a system that includes its own capacity for observation and self-reflection. This involved examining concepts like self-reference, autonomy, and the ways in which systems construct their own realities or domains of interaction. While this didn't fully resolve the question of how purely material systems could generate intrinsic, self-sustaining purposes in the biological sense, it crucially highlighted that understanding purpose required looking at the system's own constructive activities and its relationship with its environment, including its internal processes. This shift marked a crucial step toward understanding how systems might appear to generate and modify their own goals, rather than simply executing externally imposed ones, and set the stage for theories like autopoiesis that would explicitly look at the self-producing nature of biological autonomy.

Despite its limitations, cybernetics made important contributions to our understanding of purposive behavior, though perhaps not in the way initially envisioned. While it may not have revealed the fundamental principles of biological organization, it showed how certain forms of goal-directedness could emerge from particular patterns of organization, especially circular causality. This pointed towards new ways of thinking about biological purposiveness. More importantly, the shortcomings of first-order cybernetics helped clarify the crucial distinctions between mechanical feedback systems and living organisms, ultimately contributing to deeper theoretical developments in our understanding of biological organization. The challenge it left for subsequent theories was to explain not just how purposive behavior is implemented, but how purposes themselves could emerge from purposeless processes.

Autopoiesis

Having encountered autopoiesis earlier in our exploration of biological organization, we now examine how it approaches the problem of teleology. Where cybernetics showed how purpose could be implemented in mechanical processes through feedback, autopoiesis attempted something more fundamental, it attempted to explain how living systems generate and maintain their own purposes through self-production[128].

As discussed in the previous chapter, autopoietic systems are characterized by operational closure. Their processes form a network where each process both produces and is produced by other processes within the network.[129] But where our earlier examination focused on the organizational principles, here we consider its implications for biological purpose.

Initially, the theory took a strong mechanistic stance, perhaps in reaction to the vitalistic tendencies it sought to avoid. Maturana and Varela viewed purposes and goals as observer-dependent attributions rather than intrinsic features of living systems. They argued that the autopoietic organization itself, the network of processes that produces and maintains the system, was sufficient to explain apparently purposeful behaviour. A cell doesn't maintain its membrane *for* any purpose, the maintenance of the membrane is simply part of the circular process that constitutes the cell's existence.

This strict mechanistic interpretation, however, encountered difficulties in explaining the dynamic, adaptive character of living systems. While autopoiesis effectively described the self-maintaining aspect of life, its emphasis on operational closure struggled to account for how organisms actively engage with and adapt to their environments. This limitation became particularly apparent when trying to extend the theory to explain cognitive processes and behaviors, which clearly involve flexible interaction with the environment.

The theory's trajectory took a significant turn in Varela's later work. Along with Andreas Weber in 2002, he explicitly embraced the notion of *intrinsic teleology* in living systems[130]. This shift, influenced by Hans Jonas's phenomenological insights, recognized that the self-producing nature of living systems necessarily entails a kind of built-in purposiveness. The continuous work of self-production creates an implicit *concern* for continuation, not as something added to the system, but as an inherent aspect of what it means to be an autopoietic system.

This turn toward teleology opened new avenues for understanding biological purposiveness while raising important questions. How could a system be both mechanically self-producing and inherently purposive? Critics argued that reintroducing teleological concepts risked undermining the theory's mechanistic foundations. Yet this tension might better be seen as revealing something fundamental about living systems - that they occupy a unique position between mechanical causation and purposive behaviour.

Autopoiesis's lasting contribution to our understanding of teleology lies in its recognition that living systems are not only goal-directed through feedback (like cybernetic systems) but self-producing in a way that generates their own purposes. This suggests that biological purpose isn't something added to mechanical processes, but emerges from a particular form of organization, one where the system continuously produces the conditions necessary for its own existence. Yet autopoiesis left crucial questions unanswered. How exactly does purpose emerge from self-production? What determines the specific forms this purpose takes? And how can we reconcile the apparently mechanistic nature of biological processes with the genuine purposiveness they seem to generate? These questions would drive subsequent developments in theoretical biology, particularly the Organizational Approach and, ultimately, point toward the need for understanding how constraint cycles can generate and maintain purpose in living systems.

Organization, Rosen & Deacon

The Organizational Approach (OA), developed by Alvaro Moreno, Matteo Mossio, and their colleagues, is a significant advancement in the effort to naturalize teleology[131]. Building on insights from autopoiesis and addressing some of its limitations, the OA aims to provide a more rigorous account of how purposive behaviour can emerge from the organizational dynamics of living systems.

At the core of the OA is the concept of self-determination, formalized in terms of the *closure of constraints.* The theory posits that biological systems generate and maintain a set of structures that act as constraints on the processes occurring within the system. These constraints are mutually dependent, forming a closed network that contributes to the system's self-maintenance. Importantly, this organizational closure coexists with thermodynamic openness. While the constraints form a closed circle of dependencies, the system must continually exchange matter and energy with its environment to maintain these constraints.

The OA argues that this self-determining organization gives rise to a form of intrinsic purposiveness. Because the system's activities contribute to maintaining its own conditions of existence through the closure of constraints, we can legitimately say that the system acts *for the sake of* its own continuation. This introduces a normative dimension into biological systems, there are things the system *should* do to maintain itself, grounded in the concrete organization of its constituent processes.

However, despite these valuable insights, some have argued that the OA faces a subtle but significant limitation in its attempt to account for teleology in living systems. The core issue centres on whether describing biological systems purely in terms of organizational closure, even when understood as a network of mutually dependent constraints, fully captures what it means to be a living individual. While the OA effectively explains how components interdependently sustain a system's organization, critics question whether this network of constraints sufficiently defines the concrete entity, the beneficiary, for whose sake these processes occur. In other words, does the *beneficiary* equate merely to the persistence of an abstract organizational pattern (the closure of constraints itself), or does genuine biological teleology presuppose a more robust, historically contingent individual whose existence as that particular individual is actively perpetuated?[132]

By focusing primarily on the relational network of constraints (which can be conceptualized as a kind of spatial map of dependencies), the OA may under-emphasize the fundamentally dynamic and temporal nature of biological organization[133]. Living systems don't just maintain static patterns of constraints, they must continuously transform their structure to remain far from thermodynamic equilibrium. Consequently, as these critiques suggest, it's not enough to describe a network of mutually supporting constraints. A full account of teleology needs to explain for whom or what this self-maintenance ultimately occurs, identifying a clear beneficiary beyond the abstract organization itself, and grounding this in the system's ongoing dynamic persistence through time.

Additionally, there's a risk of conflating conditional explanations with genuinely teleological ones, or indeed, different types of end-directedness. This distinction, which was articulated by Robert Rosen in his work before the development of OA, remains pertinent to this debate[134]. For Rosen, explaining phenomena merely through their effects in maintaining a system (such as how constraints contribute to systemic persistence) can fall into the category of conditional explanations of mechanistic dependencies. While OA's closure of constraints effectively describes how various processes conditionally depend on each other to achieve a self-maintaining organization, a state Rosen would categorize as *finality* (a closed causal regime that is its own outcome), his critique suggests this is distinct from, and potentially insufficient for, true teleology. For Rosen, teleology implies a system acting towards a future, not-yet-realized goal, driven by an intrinsic normativity that extends beyond the mere perpetuation of its current form. This nuanced distinction, between a system that simply is its own end (finality) and one that actively strives for future ends (true teleology), poses an ongoing challenge for contemporary organizational approaches aiming to fully naturalize biological purpose.

While the OA effectively describes how a system achieves organizational closure (going beyond mere operational closure where functions simply remain within the system[135]) through the closure of constraints, the framework, by primarily detailing the relational architecture of these constraints, may still not fully capture the inherently dynamic and temporally extended nature of biological purpose. Deacon's theory of *teleodynamics* offers insights by focusing on the dynamic processes through which genuine teleology emerges. Unlike approaches that might primarily map static organizational patterns or closure relationships, teleodynamics emphasizes that organization emerges from the kinetics of metabolism itself and must be constitutively and actively maintained through continuous change and work. As discussed before, multiple self-organizing processes become dynamically interlocked. Processes mutually constrain each other in ways that counter their tendency toward entropy maximization. This dynamic interaction creates a phase transition from self-organization to self-constraining systems.

This phase transition comes through the dynamic opposition of processes, not just their organizational closure. When self-organizing processes become coupled, each prevents the other from reaching its entropy-maximizing end state, creating a higher-order dynamic constraint that actively maintains the system's integrity. Crucially, this maintenance isn't achieved through static patterns but requires constant structural transformation. The system must continuously modify itself to remain far from thermodynamic equilibrium. This dynamic opposition generates both a clear beneficiary (the system whose integrity is maintained) and an intrinsic normative dimension (what the system must do to maintain itself). Unlike dissipative structures or other self-organizing systems that simply maximize entropy production, teleodynamic systems actively work against their own dissolution through these dynamically interlocked processes that necessitate ongoing metabolic work and structural change.

Building on insights from the Organizational Approach and Deacon's teleodynamics, we can now articulate a comprehensive view of teleology in biological systems:

Condensed Argument on Teleology in Biological Systems

- **Teleology:** Goal-directed behavior, implying that a realizable state B influences the transformation from a state A to B, where A is at time t and B is at time $t + 1$. This implies:

  1. Certain processes in this system will *serve the function* of achieving B.
  2. The system as a whole transforms in a constrained (non-random) direction.

  Both (1) and (2) are evident in biological systems:

  - (1) Organs/Organelles perform functions that contribute to the maintenance of the organism as a whole.
  - (2) Organisms keep themselves far from equilibrium.

- This is achieved in a **self-constraining** system, where the emergent higher-order constraint is the persistence of the system itself. Therefore, the systems will act *in order to* or *for the purpose of* maintaining this emergent higher-order constraint.

- **Constraint cycles** describe the dynamics of a system where one stage and its self-organizing tendencies lead directly to the next stage in a cyclical manner with adaptive variation.

- A goal-directed behavior emerges as the system cycles through a series of states. For example: A moves into a novel state B, where B can be seen as the goal of A, since the sequential movement from state A to B is directionally encoded in the higher-order constraint or the constraint cycle.

At its core, teleology refers to goal-directed behaviour. This means a system in a current state A is constituted in a way that its processes are directed towards bringing about a specific, realizable future state B. Crucially, for a naturalized account, this directionality arises not from state B exerting influence backwards in time (which would involve problematic future causation), but from the present organization, constraints, and dynamics inherent to state A. State B represents a potential outcome that the system is disposed to achieve due to its current intrinsic properties and the way it interacts with its environment. This concept, thus understood as an emergent property of a system's present constitution guiding it towards a future state, has implications for our understanding of living systems.

In biological organisms, we observe clear evidence of such goal-directed behaviour. Consider, for instance, the heart. Its function of pumping blood serves the broader goal of maintaining the organism's vitality. This exemplifies how certain processes within a system *can serve the function* of achieving a future state. Moreover, we see that biological systems as a whole transform in constrained, non-random directions, further supporting the presence of teleological behaviour.

This goal-directedness emerges from what we have called a self-constraining or a teleodynamic system. In these, an emergent higher-order constraint arises, the persistence of the system itself. This overarching constraint leads the system to perform actions *in order to* or *for the purpose of* maintaining its own existence. It's this emergent property that gives rise to genuine purposiveness in living systems.

The dynamics underlying this teleological behaviour can be understood as a constraint cycle. In such a cycle, the specific dynamics of one stage, lead directly to and enable the next stage in a cyclical manner, often incorporating adaptive variations. More specifically, the processes within the cycle are coupled such that the completion or product of one process becomes a necessary condition or constraint for the initiation and operation of subsequent processes. It is this structured interdependency of processes, where each stage sets the stage for the next, that gives rise to an emergent higher-order constraint.

A prime example of such a constraint cycle, is the autogenic process we discussed at length. The autogen has a self-constraining organization where processes are directed towards the system's own perpetuation. Its significance lies in illustrating how such apparently goal-directed, self-preserving dynamics can arise from simpler, non-living chemical interactions, forming a conceptual bridge towards understanding more complex biological organization. In this model, the culmination of a reciprocal catalysis leads to the initiation of self-assembly, which in turn leads to shell formation. The shell eventually breaks, leading back to reciprocal catalysis, completing the cycle.

Through this cyclical process, we can observe the emergence of behaviour that is genuinely teleological or goal-directed in a naturalized sense. While the progression through states (e.g., from A to B) is driven by existing causal mechanisms, often reflecting energetically and catalytically favoured pathways, the significance of state B is not merely that it follows A. Rather, state B is a necessary successor to A for the sake of the perpetuation of the entire cycle. In this sense, the system, by virtue of its organization as a constraint cycle, acts as if it is working towards subsequent states because these states are integral to maintaining its own existence as a cycle. This aligns with Rosen's concept of finality, where the system is, in a way, its own end.

The *goal* of achieving state B from state A is therefore not an external target pulling the system forward, but an internal consequence of the higher-order constraint constituted by the entire cyclical organization. This higher-order constraint exerts a form of downward causation. It is the overall, dynamically maintained pattern of the cycle that dictates the viability and sequence of its constituent local processes and states. The system doesn't just passively transition, it is dynamically configured such that the conditions generated in state A actively promote and constrain the processes leading to state B, because the failure to transition to B (and subsequently to C, and so on) would mean the dissolution of the constraint cycle itself.

While the work of Kaufmann, Maturana, Varela, and Rosen showed how living systems are characterized by cyclical, self-referential relationships, Deacon's approach further specifies how such organization is actively constructed through the sequential generation and coordination of constraints over time. Understanding how constraints are coordinated across time, not just space, provides the key to understanding how genuine teleology emerges from biological organization.

Viewed through the lens of *organizational continuity*[136], this perspective still builds on the Organizational Approach's notion of closure, but it foregrounds the diachronic maintenance of that closure. In other words, it is not claiming a new type of closure, rather, it makes explicit the temporal dimension that Rosen (via immanent causation) and Louie already saw to be implicit in closure to efficient causation. By tracking how constraints succeed one another and re-instantiate the same organizational pattern, the constraint-cycle view specifies who benefits: the historically continuous individual that persists as that very pattern, despite the ongoing material turnover. This gives us a sharper criterion for biological individuality and a concrete target for normativity (what the system *ought* to do to keep existing), without discarding the spatial interdependence that OA describes so well.

More fundamentally, constraint cycles show that the unity of a living system is achieved through an ongoing orchestration of mutually opposing processes like catabolism/anabolism, excitation/inhibition, exploitation/exploration, whose viability conditions shift over time. The system does not eliminate the opposition in a spatial compromise, it re-entrains each process at the right phase so that, taken across the whole cycle, their joint effect preserves organizational continuity. This temporally distributed integration, already latent in Rosen's diagrams but here made analytically explicit, marks the decisive break between living and non-living dynamical regimes. What persists is a regulated succession of transformations, where each would be self-defeating if run in isolation.

By understanding teleology in these terms, we can bridge the gap between reductionist physical explanations and the apparent purposiveness of life. This approach explains how and why living systems actively work towards their own persistence in the face of entropy.

This brings us to a fundamental contrast between two ways of viewing the world. The synchronic perspective that dominates modern science, and the diachronic perspective that better captures the temporal nature of living systems. This distinction will help us understand why traditional scientific approaches often struggle to capture the essence of life, and point us toward new ways of thinking that better align with life's temporal nature.

5.2 The Synchronic World

> *We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past could be present before its eyes.*
> *—Pierre Simon Laplace, A Philosophical Essay on Probabilities*[137]

From what we have seen, organisms make up a distinct class of systems with a distinct temporal regime. Since we as humans belong to this class, the specific dynamics that are unique to organisms form a basis for us and our being in the world.

This realization prompts a question: How does our nature as living systems, shape the way we interact with the world, how we perceive it and how we think about it? Our philosophical frameworks, our scientific paradigms, are they not also constrained by the very processes that gave rise to our capacity to exist?

In the chapters ahead, we'll pry open this Pandora's box. For now we'll add some additional context, dissect the clockwork universe that has ticked away in the background of modern thought, before attempting to sketch aspects of a new philosophy, one that embraces our nature as living, cyclical, error-prone systems. It's a tall order, that the great William Wimsatt described as the *re-engineering of philosophy for limited beings*. As we've seen, in the grand cycle, there is indeed a time for every purpose under heaven, including, a time to think anew about our approach to knowledge, acknowledging the cyclical nature of life and its influence on our quest for understanding.

To more clearly see the road ahead, we must first look back and understand the dominant paradigm it challenges. The Enlightenment era, with its emphasis on rationality and mechanistic explanations, gave birth to particular modes of thinking that have profoundly shaped modern science and philosophy.

Modernity, of course, is a vast and multifaceted concept, encompassing cultural, social, economic, and technological transformations too numerous to fully address here. Our focus, however, will be on one particularly influential aspect, the emergence of a specific mode of scientific thinking. This paradigm, while not the totality of modern thought, has been instrumental in shaping how we conceptualize and investigate the natural world. It has profoundly influenced the development of scientific methodologies and, by extension, our broader cultural understanding of causality, time, and change.

Central to this scientific world-view is an approach we will call *synchronic thought*, which attempts to capture the entirety of a system or phenomenon in a single, timeless snapshot. This perspective has dominated much of modern scientific inquiry, offering powerful explanatory tools but also imposing certain limitations on our understanding of dynamic, living systems.

Perhaps the most extreme expression of this synchronic world-view is found in the concept of Laplace's demon, featured in the epigraph. In 1814, the French mathematician Pierre-Simon Laplace proposed a thought experiment. Imagine an intellect vast enough to know the precise location and momentum of every particle in the universe at a given moment. With this knowledge and the laws of physics, this demon could calculate the entire past and future of the universe.

This idea epitomizes the synchronic approach. It suggests that reality can be fully encoded in a predefined state space, where all possible permutations are already present. In this view, time becomes just another variable in a grand equation, robbed of its transformative power. The future is not created, it is just revealed[138].

This synchronic world-view further is a profound spatialization of time, where temporal processes are flattened into a single, comprehensible plane. A perspective that echoes the divine view of the eternal, all-seeing eye of God transposed into the realm of science. As Jürgen Habermas astutely observes[139], the idea of God's perfect, objective viewpoint doesn't vanish with secularization, rather, it's transfigured into the idealized position of the scientist. The modern scientist, in this conception, stands outside of time, observing the universe with detached objectivity, much like the God of classical theism.

This god's-eye view becomes central to the modern scientific enterprise, shaping its methodologies and aspirations. It promises a world of perfect coherence and clarity, where all phenomena can be explained through timeless, universal laws. This paradigm has been enormously productive, giving rise to unprecedented technological advances and deep insights into the workings of the physical world. Yet, it also imposes significant limitations, when grappling with the complex systems and particularly when applied to living systems.

In Modernity the organism becomes a machine. A complicated interaction of multiple parts that are arranged in a precise way to do something. Using metaphors from different domains to explore a topic is always interesting and most often insightful. The machine metaphor has certainly been very productive and insightful in understanding organisms better. Our understanding of the circulatory system, of the musculoskeletal and the respiratory system have greatly benefited from mechanistic interpretations. Similarly the modern version of the machine metaphor, the computer, has greatly enhanced our understanding of neurosciences, neuronal networks and cognitive strategies.

When finding one such useful explanatory strategy, like noticing how much the circulatory system is like a hydraulic system, or how much the structure of artificial neuronal networks resembles our neuronal networks, one naturally tries to repeat it. First, other organs follow, before the whole organism is swallowed by the hungry metaphoric monster. However, sooner than later the totality of things, the whole universe, becomes a giant machine, or a computer. If those association are already too 20th-century-like and bring up the often dysfunctional, clunky and aesthetically unappealing computers we still remember, you may very well opt for the everything-is-a-simulation option, leaving the energy-intensive materiality of integrated circuits behind, transcending into clean luminous virtual space.

There is a slippery slope when using a metaphor that is helpful in one area that may not be in other areas. First, one has to notice that the nature of a metaphor is such, that it helps expand the surface of our understanding, highlighting similarities between the unknown and something known, that which serves as the basis for the metaphor. This implies that metaphors invite category errors, where one thing belonging to a certain category is grouped with something quite different. In many cases that is not necessarily a problem, but especially at the intersection of life and non-life one wants to be somewhat careful as the consequences can be rather detrimental.

An example, Descartes conceived all organisms as machines. The human was special as it was not mere machine-like body but also mind. Mind was unique to humans. This mechanistic view persists in modern industrial agriculture, where animals are often treated as pure production units, biological machines for converting feed into meat. The practice of intensive animal farming, where pigs are confined in gestation crates barely larger than their bodies, follows directly from this logic. If an animal is just a machine, its suffering becomes merely a technical parameter to be weighed against production efficiency in the optimization process. This is a stark example of how metaphors, when extended beyond their useful domain, can normalize practices that fundamentally misunderstand the nature of living beings. The machine metaphor obscures the inherent purposiveness and experiential nature of these creatures, reducing complex living systems to mechanisms of production.

While metaphors can illuminate aspects of living systems, the core challenge lies deeper in the fundamental mismatch between synchronic thinking and the processual nature of life itself. When we attempt to extend mechanistic thinking to living systems, we encounter not just technical problems, but a basic inability to capture the essence of life as process.

Firstly, living systems exhibit a remarkable capacity for genuine novelty and emergence that defies prediction from initial conditions. This goes beyond the complexity we see in non-living systems. Life, is characterized by the ability to generate new structures, behaviours, and possibilities that cannot be predicted solely from current states. This creative aspect of life, its ability to generate something unprecedented, challenges the synchronic assumption that all potentialities are pre-encoded in the present state.

Additionally, living systems don't just occupy a fixed state space, they actively modify and expand their own possibilities over time. Consider the development of an embryo. It is not only growing within predetermined boundaries, but qualitatively changing its very nature and potential. Or think about the difference in affordances, the possibilities for action and interaction, for a bacteria, a cat, a dolphin and you. Let's take water, for a bacterium, it is simply a medium to float in, for a dolphin it's a three-dimensional space for complex movement, a cat just tries to avoid it, and for humans it can be anything from a barrier to a recreational space depending on our technologies and skills. This dynamic difference and reconfiguration of possibility spaces is a hallmark of life that the static, albeit vast, landscape assumed by synchronic thinking cannot accommodate.

Moreover, living systems exhibit an inherent purposiveness that resists reduction to purely mechanical explanations. From the cellular level to complex organisms, we observe goal-directed behaviours that arise not from external design but from the very nature of life itself. This teleological aspect is fundamentally at odds with the flattened, timeless space of synchronic thought. It suggests that to truly understand life, we need a framework that can account for purpose and directionality as emergent properties of complex, self-organizing systems.

These challenges collectively point to a need for a paradigm that can embrace the dynamic, self-creating nature of life without sacrificing the rigor and explanatory power that has made the synchronic approach so successful in other domains.

5.3 The Diachronic World

The contrast to the synchronic view is the diachronic view. It's a perspective that looks at phenomena through their development over time and regards these influences as central in the description of said phenomena. This mirrors the distinction between holonomic and non-holonomic constraints discussed earlier, where holonomic constraints map to a synchronic view and non-holonomic to a diachronic one.

The call for a diachronic approach resonates far beyond the boundaries of our specific inquiry into living systems. It echoes through the corridors of 20th-century philosophy, from Alfred North Whitehead's emphasis on process to Henri Bergson's concept of durée (duration). Phenomenologists like Martin Heidegger and Maurice Merleau-Ponty grappled with the temporal nature of human experience, while Gilles Deleuze and Félix Guattari crafted philosophical landscapes that pulsed with becoming rather than static being. In the sciences, figures like Ilya Prigogine challenged the reversibility of time in physics, developing theories of dissipative structures that pointed towards an arrow of time emerging from the study of far-from-equilibrium systems, while Lee Smolin has more recently argued against the *block universe*, proposing that time is fundamental and that physical laws themselves evolve rather than remaining timeless constants. In the same dynamical spirit, Edward Lorenz formalized deterministic chaos and sensitive dependence, while C. H. Waddington cast development as canalized trajectories across an epigenetic landscape[140].

This broader philosophical and scientific context sets the stage for our specific exploration of living systems. Our investigation of constraint cycles has argued for not just the utility, but the necessity of a diachronic approach. Living systems, we've discovered, cannot be fully captured in a single moment. They are, in essence, ongoing transformative cycles where each state influences the dynamics of the next, leading to the emergence of higher-order constraints guiding the overall telos of a given system.

This diachronic nature of living systems has implications for how we understand knowledge itself. If we are fundamentally temporal beings, systems that maintain ourselves through cyclical processes, then our very capacity for knowledge must be shaped by this temporal nature. This is not only an interesting philosophical observation, rather it's a fundamental constraint on how we can know and understand the world.

Consider what it means to be a knowing subject. Our knowledge isn't stored in static structures but maintained through dynamic processes. Our memories, our concepts, our understanding, all must be actively maintained and reconstructed through time.

This temporal nature of our existence as knowing subjects creates both limitations and possibilities. On one hand, it means we can never achieve the timeless, God's-eye view that classical philosophy sometimes aspired to. Our knowledge aligns naturally with a perspectival, temporal, and cyclical understanding as we ourselves are temporal and cyclical beings[141]. The synchronic ideal of complete, timeless knowledge is not just practically impossible but fundamentally at odds with our nature as living systems.

Yet this same temporal nature also gives us unique capabilities through the operation of higher-order constraints. Just as living systems integrate seemingly contradictory processes through temporal coordination, our minds can hold and integrate different viewpoints, perspectives, and forms of knowledge, through temporal coherence. We can move between different mental states, different emotional registers, different modes of understanding, maintaining their distinctness while integrating them across time. This is emerges from our fundamental nature as living systems that coordinate difference through temporal integration rather than flattened simultaneity.

Consider how we can understand a concept from multiple angles, hold seemingly contradictory viewpoints in productive tension, or integrate emotional and rational responses to a situation. These capabilities don't come from resolving differences into a single, unified perspective, but from our ability to coordinate different states across time.

This suggests that we need an approach to philosophy that embraces rather than tries to transcend our nature as living systems. Instead of attempting to escape our temporal, cyclical nature in search of timeless truth, we should develop methods that work with and through our fundamental characteristics as living beings. This is where the philosophy of the *cyclical synthesis* comes in, not as a description of how living systems work, but as a method for doing philosophy that aligns with our nature as temporal, cyclical beings.

Chapter 6

Towards a Logic of Life

> The possibility of science depends upon the fact that human thought necessarily partakes of whatever character is diffused through the whole universe, and its natural modes have some tendency to be the modes of action of the universe.
> –Charles Sanders Peirce[142]

Peirce's insight points to something deeper than scientific methodology alone. His quote suggests that our capacity for understanding, whether through science, philosophy, art, or any form of inquiry, stems from our participation in the universe's fundamental patterns. It's not only about the possibility of science, it's about the possibility of sense-making itself. If human thought necessarily partakes in universal characteristics, then the patterns we find in living systems aren't just objects of study - they are actively expressed in our ways of knowing. The scale invariance described through constraint cycles suggests that the dynamics that enable life at the molecular level go beyond it. We are not passive observers of these patterns but their continuation and expression at higher levels.

This recognition invites us to consider what we might call a *logic of life*. When we speak of logic here, we're not only referring to formal systems, but rather to something more fundamental - the patterns of organization and coherence that characterize living processes. The word *logic* derives from the Greek *logos*, related to reasoning and the underlying rational principles or patterns that give order to things. If living systems exhibit characteristic patterns of organization and development, might these constitute a kind of logic, a structured way of maintaining coherence through time? The constraint cycles we've examined suggest exactly this. They describe fundamental organizational patterns that enable systems to maintain themselves far from equilibrium while adapting to change. In this chapter, we'll build on the foundations we have laid thus far, developing elements that are central to understanding this logic of life.

What, then, does it mean to do philosophy as a living system? What changes when we recognize that our philosophical endeavours are not only about life, but are expressions of life itself? How does this perspective reshape our approach to sense-making? Most crucially, how will our understanding of knowledge itself transform when we recognize it as an achievement of living processes rather than an abstract, disembodied faculty?

To answer these questions, we must first acknowledge the vast complexity of living systems as we know them today. We as living systems exist multiple levels in complexity above those of our ancient ancestors. Fascinatingly, the fundamental constraint cycle and dynamics of these first proto-cells is still going strong after some 4 billion years. Life has replicated, adapted to myriads of different niches, begun to cooperate with variations of itself and come to prey and parasitise on itself. Resisting the inexorable pull towards equilibrium for such a time span is a mesmerising feat, and the complexity that has built up in living systems is staggering.

So, the living world we see around us, from trees, to animals, to humans and birds is in some sense still the continuous expression of an initial constraint cycle[143], but, of course, in myriads of different expressions–some more, some less complex. Since we as living beings are the successful and coordinated integration of trillions of such constraint cycles (just at the cellular level we are the coordination of roughly 36 billion cells), things tend to get complicated and we wade into a causal thicket of sorts, and the dynamics, because of their multi-level hierarchical interaction, become dizzying to think of.

For the sake of clarity in our analysis, we'll focus on the implications of a single constraint cycle for philosophical thinking and cognition in general, rather than diving into the hierarchical complexity of multiple interacting cycles.

If our very capacity to know and understand emerges from the very processes that have maintained life, this collapses the traditional separation between epistemology (how we can know) and ontology (what exists). Karen Barad terms this *onto-epistemology* to signal the inseparability of ontology and epistemology of agents in our world[144]. Our ways of knowing aren't arbitrary constructions imposed on reality but emerge from our nature as living systems. The categories through which we understand the world, our modes of reasoning, even our capacity for abstract thought—all must be compatible *with*, and emerge *from*, the constraint cycles that maintain our existence as knowing beings.

Yet this recognition of our nature as organizationally self-producing continuities—selves—sustained by nested constraint cycles, doesn't confine us to pure subjectivity. Instead, it uncovers a more nuanced relationship between subject and object, one that develops through the very dynamics that make life possible. As living systems, we are the product of billions of years of adaptation to our environments. Our perceptual systems, cognitive architectures, and ways of knowing didn't develop arbitrarily but through continuous engagement with the concrete demands of survival and reproduction, a *structural coupling* between organism and environment[145]. Our organizational continuity gives us a stable centre of selfhood. The ongoing succession of constraint cycles keeps that centre plastically attuned to shifting conditions. This creates an inexorable link between knower and known. One that both enables and constrains our understanding of the surrounding world.

Consider how our perceptual systems embody this principle. We've evolved to detect certain aspects of reality while remaining blind to others. We can see visible light but not ultraviolet or infrared, we can hear certain frequencies but not others, we can detect some chemical signatures but not others. These limitations reflect what has been relevant for our survival and adaptation. We didn't evolve direct sensitivity to radioactivity because it wasn't a significant selective pressure in our evolutionary history, but we did evolve trichromatic colour vision like other primates, helping us identify ripe fruit and obtain essential nutrients like vitamin C.

This deep coupling between organism and environment, between subject and object, shows that our categories of understanding crystallise from our history of successfully coping with the recurrent patterns that reality imposes on experience. The very fact that we can navigate and manipulate our surroundings demonstrates that our knowledge tracks those mind-independent regularities that shape, but do not submit to our constructions, even though every such tracking is mediated by our evolved capacities and their limitations.

What emerges is not the pure, view-from-nowhere objectivity that classical philosophy sometimes aspired to, but a pragmatic *adapted objectivity*. Knowledge that is necessarily perspectival yet reliably tracks features of reality[146]. The constraint cycles that shape our existence allow us to build understanding that transcends any single moment or perspective precisely because they've evolved to capture regularities in our environment that persist across time and context.

Overall, this points toward the following methodological insight: rather than seeking to transcend our nature as living systems in the pursuit of knowledge, we need approaches that work with and through our very nature. Necessarily, the principles we develop must resonate with the fundamental processes that make us what we are. Our nature as living systems, our biological and cognitive architecture, forms the inescapable funnel through which all understanding of the world must flow. This primordial lens shapes every aspect of our knowing, from our basic sensory experiences to our most abstract theoretical musings.

6.1 Cyclical Synthesis

Cyclical synthesis has its origin in the thermodynamic analysis of living systems, yet goes beyond its empirical origins to serve as a *generator function* for philosophical inquiry. A generator function for philosophical inquiry is a minimal conceptual structure that, once articulated, can recursively unfold into richer patterns of thought, serving as a seed that generates whole lines of analysis, interpretation, and application across contexts.

As a framework, it formalizes how apparently opposing elements can maintain their distinct identities while achieving integration through *temporal sequencing* rather than *spatial resolution*. This results in a framework for investigating how some systems can integrate difference while preserving distinctness, generate novelty while maintaining continuity, and achieve coherence through temporal organization. In precisely this sense, cyclical synthesis suggests a form of logic—a way of reasoning—in which coherence emerges not through the elimination of contradiction (Classical logic), nor through its absolute resolution (Hegelian logic), nor merely through its accepted coexistence (Paraconsistent logic), but through its temporal integration via cyclical coordination.

While constraint cycles describe a pattern found in living systems, cyclical synthesis emerges as a logical structure derived from these dynamics. Just as classical logic developed from Aristotle's analysis of valid reasoning, paraconsistent logic from the study of paradox, and Hegelian dialectics from examining the progressive development of ideas over time, cyclical synthesis arises from the study of how living systems maintain coherence through recursive self-constraint.

Consider how living systems maintain themselves through opposing yet coordinated processes: growth and decay, synthesis and breakdown, expansion and contraction. These aren't different processes that simply coexist, they exhibit counter-acting tendencies whose uncoordinated operation would lead to mutual interference or cancellation.

Not all opposing processes can be productively coordinated. Three distinct cases emerge:

1. Multiple instances of processes that simply replicate rather than generate new possibilities
2. Opposing processes that remain incompatible, leading only to mutual destruction or cancellation
3. Opposing processes that, through temporal coordination, can generate higher-order constraints while maintaining their distinct characters

It's this third case that interests us - where apparently opposing processes create conditions that, when properly coordinated across time, enable novel forms of organization to emerge. For this we can define:

Cyclical Synthesis

Cyclical synthesis occurs when (i) mutually interfering processes are phase–separated in time, (ii) each phase prepares enabling conditions for the next, and (iii) the composed cycle reinstates an identity-bearing constraint regime.

(i) The target processes would degrade each other if co-active. The system therefore coordinates them. This (e.g., sleep vs. wake, anabolism vs. catabolism) prevents destructive interference and makes each phase legible in its own conditions.

(ii) Each phase lays down the prerequisites the next requires—energy, substrates, slack, structure, or informational reparameterizations. Catabolism feeds anabolism, sleep consolidation restores conditions for fresh encoding, exploration supplies candidates that exploitation can refine.

(iii) Over one full pass, the composed sequence of phases reinstates the identity-bearing constraint regime. The system returns to (and maintains) a viable organization while its microstates turn over. Coherence is achieved through change, an emergent, system-level constraint that shapes the phases without erasing their distinctness.

These principles define cyclical synthesis as a processual logic where stable organization arises from the coordination of opposing processes over time.

When these dynamics are successfully established under the specified conditions, one finds that the system maintains its overall coherence and identity through continuous transformation across different, potentially opposing, states or phases. The emergent higher-order constraint is a new level of organization or functional capability that was not present in the constituent processes alone. What can appear as contradiction from a static viewpoint becomes a source of dynamic stability and adaptive potential when coordinated temporally.

In sum, cyclical synthesis acknowledges the primacy of process. Rather than striving for static resolution or reduction of tension, it provides a means to understand how systems achieve coherence through the ongoing interplay of mutually enabling processes over time. By re-framing contradiction not as a problem to be eliminated, cyclical synthesis opens a conceptual pathway for analysing how life-like systems generate stable modes of being and knowing while remaining continually open to transformation and renewal. In doing so, cyclical synthesis stands as a logic of living systems. A framework that not only interprets the complexity and resilience of life's recursive constraints but also guides inquiry that emerges from embodying such a logic ourselves.

6.2 Examples

Awake and Dreaming

The beauty of cyclical synthesis is how it manifests across different domains while maintaining its essential structure. We'll examine this pattern at interrelated levels, beginning with a familiar example at the cognitive/neural level of sleep-wake cycles and memory consolidation. Each case will demonstrate how apparently opposing processes can be integrated through temporal coordination to generate higher-order constraints.

The sleep-wake cycle involves two processes, encoding during wakefulness and consolidation during sleep. These processes have countervailing tendencies in their basic requirements and operations. During waking states, the brain prioritizes the encoding of new experiences and active learning, demanding high sensitivity to external stimuli, active formation of new neural connections, and conscious engagement with the environment. During sleep, particularly REM sleep, the brain shifts to a different mode of consolidation and integration, which requires the opposite conditions, isolation from external stimuli, reorganization of neural patterns, and what recent research suggests is a form of generalization through corrupted replay of experiences[147].

These processes cannot effectively co-occur, attempting to simultaneously encode new experiences while consolidating existing memories would lead to interference and degradation of both processes. Each requires conditions that would actively interfere with the other's operation.

The temporal coordination between them is maintained through circadian rhythms, creating alternation between states. This coordination generates emergent constraints that shape both processes while maintaining their distinct identities. Consider, long-term memory or learning in general, this emerges not from either process alone, but from their cyclic interaction. During wake periods, experiences are encoded in specific neural patterns that are optimized for subsequent sleep-based processing. Similarly, sleep consolidation does not only preserve these patterns but restructures them in ways that enhance the brain's capacity for new learning during the next wake period.

Beyond this, research suggests that sleep actively works to prevent over-fitting to daily experiences. Through the generation of dream-like variations, it enhances cognitive flexibility and generalization[148]. This increases the effectiveness of new learning during subsequent wake periods. Conversely, the specific patterns of encoding during wakefulness create the essential structure and content that sleep-based consolidation requires to perform its integrative function[149].

Neither process dominates or eliminates the other, instead, their temporal coordination generates higher-order cognitive capabilities that emerge from but transcend either process alone. These capabilities include memory formation and retention, cognitive flexibility, pattern recognition, and the ability to generalize across experiences[150].

Selection and Generation

The standard account of evolution, following Darwin's seminal work, has focused on natural selection as the key driver of evolutionary change. This selective dynamic operates through elimination, working as a kind of sieve that filters out less adaptive variants while preserving those better suited to prevailing conditions. This exclusively negative framing obscures a crucial complementary dynamic. The generative, multiplicative powers that Darwin himself recognized as *life's several powers* have been largely relegated to the background in evolutionary theory, treated as providing raw material for selection rather than as an essential complementary force in their own right[151].

Viewed through the lens of cyclical synthesis, we find that evolution expresses the interplay between these two opposing processes that, through their temporal coordination, generate the complexity and adaptivity of life. These processes of selective elimination and generative diversification (the continual production of selectable variation), exhibit a tension[152].

Selection operates through elimination, systematically reducing variation by filtering out forms that don't adequately fit environmental constraints. It is inherently reductive, narrowing the space of possibilities by preserving only those variants that maintain functional integrity within their context. Diversification, in contrast, operates through the continual production of selectable variation. Processes like gene duplication, mutation, recombination, and developmental re-patterning. It is inherently expansive, working to increase the space of possibilities by giving rise to novel, heritable traits on which selection can subsequently act[153].

The opposition between these processes isn't incidental. Pure selection, operating without generation, would rapidly eliminate variation and lead to evolutionary stasis. A dead end of perfectly adapted but unchangeable forms. Conversely, unconstrained generation without selection would produce an explosion of random variations, most non-functional, leading to chaos and the dissolution of organized structure. Each process, operating alone, would undermine the very conditions that make evolution possible.

Yet it is through temporal coordination that these tendencies create the necessary enabling conditions for each other. Selection preserves functional relationships that provide the stable foundation from which new variations can meaningfully emerge. Generation produces the variations that make selection possible and provide the raw material for adaptation. Each process depends on and enables the other through their cyclic interaction through time[154].

This pattern repeats across multiple scales in biological systems. Consider neurogenesis during brain development. The developing brain initially generates vast numbers of neurons and synaptic connections, far more than will ultimately survive.

This period of exuberant growth is followed by selective pruning based on patterns of neural activity. Connections that participate in coordinated firing patterns are preserved, while those that don't are eliminated[155].

The same dynamic appears in molecular evolution through gene duplication and subfunctionalization. When genes are duplicated, one copy can maintain the original essential function while the other is free to accumulate variations. Most of these variations are eliminated, but occasionally they give rise to novel functions that are preserved by selection. This process of *neofunctionalization* has emerged as a major source of evolutionary innovation, particularly in the development of complex regulatory networks[156].

This coordination between generation and selection produces a higher-order constraint, evolvability itself. Here understood as the capacity of a lineage to keep generating heritable, potentially adaptive variation while preserving enough organization to remain viable.

This underscores a crucial feature of cyclical synthesis, what appears wasteful from a static perspective becomes necessary when viewed dynamically. The "overproduction" of neurons, genes, or variations is not inefficiency but phase separation. You cannot select from variations that don't yet exist, nor can adapted forms emerge without selection. Each phase enables the next, where generation's abundance provides the substrate selection requires and selection's winnowing creates the stable organization from which meaningful variation can emerge. The cycle maintains lineage identity not despite this apparent waste, but through it.

Sympathetic and Parasympathetic

Moving from evolutionary timescales to the moment-by-moment regulation of our bodies, the autonomic nervous system provides another clear demonstration of cyclical synthesis. The sympathetic system prepares the body for action through the *fight or flight* response by increasing heart rate, dilating pupils, redirecting blood to muscles, releasing glucose, and suppressing digestion. The parasympathetic system promotes *rest and digest* functions by slowing heart rate, constricting pupils, promoting digestion and nutrient absorption, and facilitating tissue repair and growth. These processes exhibit clear opposition. Blood cannot flow maximally to both muscles and digestive organs simultaneously, the heart cannot both accelerate and slow, energy cannot be simultaneously mobilized and stored.

This exemplifies cyclical synthesis through temporal coordination. (i) The sympathetic and parasympathetic branches represent mutually interfering processes that cannot effectively co-occur, attempting to maximize both simultaneously would be physiologically contradictory. (ii) Each phase prepares enabling conditions for the next. Sympathetic activation mobilizes resources needed for action but depletes reserves and creates metabolic by-products that must be cleared, while parasympathetic activation facilitates recovery, nutrient storage, and tissue repair, restoring capacity for future sympathetic responses. (iii) The composed cycle reinstates an identity-bearing constraint regime. The higher-order constraint of adaptive flexibility, the capacity to shift between states appropriate to context while maintaining physiological coherence. Sustained sympathetic activation without parasympathetic recovery leads to stress-related disease and metabolic dysfunction, while insufficient sympathetic activation leaves organisms unable to respond to challenges.

Explore and Exploit

A similar, although much more abstract expression of this pattern appears in the fundamental dynamics of problem-solving and adaptive behaviour. *Exploration* involves searching for new possibilities, gathering information, and remaining open to uncertainty, while *exploitation* focuses on utilizing existing knowledge, refining known solutions, and maximizing current returns. Both draw on the same finite resources, pushing hard in one direction inevitably curtails what is available for the other. Exploration accepts temporary inefficiency in exchange for potential future gains, while exploitation optimizes what is already working. Like trying to simultaneously learn a new skill and perform at peak capability, these requirements actively interfere with each other's operation.

Again, cyclical synthesis provides a framework for understanding their coordination. (i) Exploration and exploitation are mutually interfering processes that cannot effectively co-occur. Exploration requires allocating resources broadly across uncertain possibilities, while exploitation concentrates resources on optimizing known solutions.(ii) Each phase prepares enabling conditions for the next. Exploitation reveals its limitations and boundaries, exposing gaps or diminishing returns that signal when exploration becomes necessary, while exploration generates new knowledge and possibilities that become available for exploitation, providing the raw material that exploitation requires to be effective. (iii) The composed cycle reinstates an identity-bearing constraint regime at a more abstract level. Successful problem-solving strategies emerge not from either process alone but from their temporal coordination. Knowledge gained during exploratory phases creates the foundation for effective exploitation, while limitations encountered during exploitation reveal which areas require further exploration.

Beyond the Binary

The examples we've examined so far all involve the temporal coordination of two opposing processes. Yet cyclical synthesis need not be limited to binary oppositions. Living systems often exhibit more complex configurations where multiple distinct processes must be coordinated through time. The number of processes we identify and the degree to which they oppose each other depends partly on how we parse the system. Sometimes the distinctions are relatively clear-cut, other times more ambiguous. What matters is identifying processes distinct enough in their operations that their simultaneous expression would create interference, yet whose temporal coordination can generate higher-order organization. Just as we have parsed autogenic processes to involve a three-way coordination between growth, disruption, and reorganization, many systems exhibit elaborate patterns of temporal integration involving multiple phases.

To discuss one such example, we turn to a framework developed in cultural anthropology. Michael Thompson's analysis of different *Ways of Organizing* and their corresponding *Myths of Nature*. This example demonstrates how cyclical synthesis can operate through four distinct but interrelated processes at the societal level, showing that this principle scales beyond simple binary dynamics to more complex multi-phase coordination.

Thompson's work shows that whenever humans coordinate action, whether in households, companies, governments, or entire societies, our interactions reliably crystallize into four recurring ways of organizing[157]. These are best understood as attractor-states, stable patterns of coordination that develop again and again at every scale of social life. A single person may act individualistically at work, hierarchically in their professional guild, egalitarian with friends, and fatalistic when dealing with a bureaucratic system. Likewise, organizations and governments tend to drift toward one or another of these attractors, shaping how they solve problems, allocate authority, and respond to uncertainty. Importantly, these four modes of Individualism, Hierarchy, Egalitarianism, and Fatalism, are not personality types or ideologies but persistent patterns of collective organization that reappear because they each solve different coordination problems and break down in different ways. With this framing, the terms that follow can be read not as moral categories but as descriptions of the basic organizational logics that societies repeatedly cycle through over time.

Each way embodies a distinct process with its own internal logic and a corresponding *Myth of Nature* that reflects and reinforces its approach. Individualism, associated with growth or exploitation phases (r-phase in panarchy), emphasizes innovation, opportunity-seeking, and rapid resource acquisition. Its myth of nature as benign reflects an expansive, entrepreneurial world-view. Nature is robust and forgiving, able to absorb impacts and provide opportunities. Like the generative phase in evolution, this process creates new possibilities but requires eventual constraint to prevent over-expansion.

Hierarchy emerges during conservation phases (K-phase), when accumulated resources and complexity demand more structured management. Its myth of nature as perverse/tolerant justifies careful regulation within perceived limits. Nature is stable within bounds but can be pushed too far if not properly managed. This process, acts to maintain and optimize existing structures, creating ordered systems of control and coordination.

Fatalism becomes prevalent during release phases (Ω-phase), when established structures break down and control seems impossible. Its myth of nature as capricious reflects the experience of systemic collapse. Nature is unpredictable and uncontrollable, rendering human agency futile. Though often seen as purely negative, this process creates the conditions for radical reorganization by clearing away rigid structures that no longer serve.

Egalitarianism tends to emerge during reorganization phases (α-phase), emphasizing collective action and fundamental restructuring. Its myth of nature as ephemeral drives the search for new, more sustainable arrangements. Nature is fragile and precious, requiring careful communal stewardship. This process facilitates the emergence of novel social-ecological configurations through collaborative reimagining of possibilities.

These four processes exemplify cyclical synthesis in a more complex configuration than our previous binary examples. (i) The ways of organizing represent mutually interfering processes that cannot effectively co-occur. For example, a society cannot simultaneously maximize the radical, decentralized freedom of Individualism while maximizing the centralized, top-down regulation of Hierarchy. These tendencies have necessary trade-offs. Likewise, the consensus-driven, collective action of Egalitarianism is fundamentally at odds with the disengaged passivity of Fatalism and the self-interested drive of Individualism. Attempting to fully express these logics at the same time is contradictory. (ii) Each phase prepares enabling conditions for the others through temporal coordination. Individualistic expansion creates the resources and innovations that hierarchical systems eventually can organize and regulate. When hierarchical systems become too rigid, they create conditions for collapse into fatalism and fatalistic dissolution clears space for egalitarian reorganization and renewed egalitarian cooperation can create conditions for new individualistic growth. (iii) The composed cycle reinstates an identity-bearing constraint regime at the societal level, namely, social resilience, the capacity for societies to maintain function while adapting to change through time. This resilience isn't reducible to any single way of organizing but emerges from their dynamic interaction and temporal coordination.

As noted at the outset, how we parse a system's processes depends on our analytical vantage point. Hence, what I want to focus on is not Thompson's four-fold parsing being uniquely correct, but in what it reveals about adaptive systems more generally. Such systems expand their adaptive capacity by harbouring countervailing tendencies that increase, in effect, the surface area of possible responses to changing conditions. The framework helps us understand that adaptability emerges not from finding the single best mode of operation, but from the capacity to cycle through different modes as circumstances demand. Just as a cell cannot remain perpetually in growth phase or permanently in reorganization mode, living systems cannot sustainably lock into any single organizational logic.

This fundamentally reframes how we must think about difference itself. If resilience arises from coordinating distinct, even opposing modes through time, then difference is not something to be eliminated or resolved into unity. Rather, genuine, irreducible difference becomes a necessary condition for adaptive complexity. The challenge is not collapsing these differences into a single "correct" approach, but finding ways to coordinate them temporally so each can express its distinct logic when conditions call for it.[158]

Therefore, attempting to optimize for any single way of organizing will reduce resilience by eliminating the creative tension between organizational logics. The variability itself enables it to persist through changing conditions.

All of the above examples demonstrate this pattern. Whether coordinating sleep and wake or four distinct ways of organizing, the principle holds. Countervailing tendencies, when properly coordinated in time, generate higher-order constraints while maintaining their distinct characters. What appears as contradiction from a synchronic perspective becomes the source of adaptive capacity from a diachronic one.

6.3 Relation to Logical Traditions

Logic is here understood as the structure of reasoning[159]. It is in precisely this fashion that I want to frame the cyclical synthesis–a way of reasoning. Therefore, it's crucial to understand its relationship to other logical frameworks and traditions.

Earlier, in our discussion on language, we briefly articulated the principle, that the constraints of a descriptive system often attempt to functionally align with the constraints of the system being described[160]. This *alignment* reflects a relationship between our tools of understanding and the phenomena we seek to comprehend and manipulate.

Hence, we can extend this principle to systems of reasoning and their formal study through logic. Different logical frameworks, whether intentionally or not, functionally align with the constraints of different types of systems of reasoning. They emerge from and are shaped by particular contexts, problems, and ways of engaging with the world. Understanding these relationships between logical systems and the phenomena they attempt to capture, can help us better situate cyclical synthesis within the broader landscape of logical thought.

To map this terrain, we'll examine three major trajectories in logical traditions[161]. We begin with classical logic and some extensions to these, which exemplify the constraints of timeless, context-independent reasoning. These systems, from Aristotelian syllogistic to modern variants, seek universal principles that hold regardless of specific content or context. We'll explore how various extensions and modifications to classical logic, like modal, fuzzy and paraconsistent logics developed to handle aspects of reasoning that classical frameworks struggled to capture.

We then turn to continental traditions, particularly Hegelian dialectics, some of its critiques and process philosophy, which attempt to embody different constraints. These approaches recognize contradiction and temporal development as essential features of both thought and reality, though they differ in how they conceptualize resolution and synthesis.

Finally, we'll examine semiotic and pragmatist traditions, exemplified by Peirce's work, which align with the constraints of evolutionary and interpretive processes. These frameworks understand logic not as static structure but as dynamic process, emphasizing how meaning and reasoning evolve through time. These are also the strands of inquiry where we find much alignment with the ideas developed in this text.

By examining how each tradition maps onto different aspects of reasoning, we can better understand both the unique contribution and proper domain of cyclical synthesis. This analysis will help clarify what kinds of constraints cyclical synthesis explicitly attempts to mirror, and consequently, where it might prove most relevant and illuminating as a logical framework.

6.3.1 Classical Logic & Extensions

Classical Logic

Classical logic and its extensions, while immensely powerful and foundational for many domains of knowledge, are fundamentally rooted in the constraints of timeless, static, and contradiction-averse reasoning or synchronic thought. While extensions like modal, temporal, paraconsistent, and fuzzy logics attempt to address some of the limitations of the classical framework, they often still operate within a paradigm that struggles to capture the dynamic, process-oriented, and inherently cyclical nature of living systems[162]. This section will explore these limitations in detail, showing from this angle how Cyclical Synthesis offers a different approach. A *logic of life* that embraces temporality, process, and the productive role of opposing forces in generating coherence.

This tradition begins with Aristotle's Organon, where he sought to identify patterns of inference that would hold true regardless of their specific content. Aristotelian logic represents the first systematic attempt to formalize the constraints of formal reasoning itself. Where earlier thinkers reasoned about specific subjects, Aristotle made the revolutionary move to examine the very structure of valid inference, regardless of content. This meta-level analysis established logic as a discipline in its own right, creating a framework for studying the structure of valid reasoning that would shape philosophy, mathematics and the sciences for millennia to come.

It has three fundamental principles: the law of identity (A is A), the law of non-contradiction (not both A and not-A), and the law of excluded middle (either A or not-A). These principles form the foundation for a system where truth values are fixed and eternal, and contradiction is deeply problematic[163]. This atemporal nature of classical logic reflects a particular view of reality, one where truth exists outside of time, unchanging and absolute. In this framework, time becomes a parameter rather than a fundamental dimension of reality, much like the Newtonian conception of absolute time as a backdrop against which events unfold rather than a dynamic aspect of reality itself.

The system's power and limitations become evident in the principle of explosion (*ex contradictione quodlibet*): if a set of premises entails a contradiction, then any proposition whatsoever can be derived from those premises. For example, if we accept both "It is raining" and "It is not raining" as premises, we can derive not only obvious absurdities but also any arbitrary statement like "The moon is made of cheese" or "2+2=5"[164]. The relevance of this points to the underlying commitment of this particular mode of reasoning, namely, that systems of knowledge cannot meaningfully contain contradictions. Reality itself must be fundamentally consistent, and that any system containing true contradictions is meaningless, since anything follows. This assumption, while powerful for mathematical and analytical purposes, creates significant challenges when dealing with systems that inherently involve change, adaptation, or self-reference.

Classical logic's structure aligns with the mechanistic worldview that dominated early modern thought. Just as Laplace's demon could theoretically calculate all future states of the universe from complete knowledge of the present, classical logic assumes that all logical consequences are implicitly contained within the premises, waiting to be deduced through valid inference. This deterministic view leaves no room for genuine novelty or emergence. All conclusions must be logically entailed by their premises in a fixed, timeless way.

This and related frameworks has proven remarkably powerful for analyzing systems where precision and consistency are paramount. In mathematics, computer science, and formal verification, classical logic's insistence on clear truth values and its intolerance of contradiction become virtues rather than limitations. When dealing with well-defined, static systems, where properties and relationships remain fixed and unambiguous, classical logic provides an invaluable tool for rigorous analysis and proof.

However, this same rigidity creates challenges when dealing with domains involving uncertainty, paradox, or change over time. These limitations have motivated various extensions and modifications to the classical framework, each attempting to align with different kinds of reasoning constraints.

The Limit of the Snapshot

This brings us to a decisive realization regarding the relationship between the logic of life and Classical Logic. We need not reject Classical Logic as incorrect, but rather re-contextualize it as the *snapshot fragment* of Cyclical Synthesis.

Consider the Law of Non-Contradiction: $\neg(A \land \neg A)$. In a living system, this law is not violated. It is respected *pointwise* at every instant. A cell cannot be simultaneously driven to the opposing extremes of anabolism and catabolism over the same substrates at the same instant, doing so would collapse viability. At any single time-slice ($\Delta t = 0$), the system resolves to a definite state, describable by classical constraints.

However, for the system to persist across a duration ($\Delta t > 0$), it must alternate between complementary regimes. It must build (A), and it must break down ($\neg A$).

Cyclical Synthesis achieves this not by tolerating simultaneous contradiction (paraconsistency), but through *phase separation*. It coordinates mutually interfering processes across time so that identity is maintained through change.

Therefore, Classical Logic is the *logic of the instant*, valid at every instant. Cyclical Synthesis is the *logic of the duration*, the structural logic that governs the sequence, ensuring that the transition from A to $\neg A$ regenerates the whole rather than destroying it.[165]

Modal Logic

Modal logic emerged as one of the first major extensions to this classical framework, introducing new operators to capture forms of reasoning that classical logic struggled to represent. Where classical logic deals with simple truth and falsity, modal logic introduces operators for necessity ($\Box$) and possibility ($\Diamond$), allowing us to distinguish between what must be true, what is actually true, and what could be true. This extension arose from the recognition that human reasoning often involves more nuanced distinctions than simple truth and falsity.

The philosophical roots of modal logic trace back to Aristotle's analysis of necessity and possibility in *De Interpretatione*, but its modern development began with C.I. Lewis's work in the early 20th century [166]. Lewis, dissatisfied with classical logic's treatment of implication, introduced the concept of *strict implication* to capture a stronger notion of logical necessity. This marked the beginning of modal logic as a distinct field of study, though its full power wouldn't be realized until several decades later.

The breakthrough came with Saul Kripke's development of possible worlds semantics in the 1950s and 1960s[167]. This semantic framework revolutionized our understanding of modal logic by providing a rigorous mathematical foundation for modal reasoning[168] . In Kripke's approach, necessity and possibility are understood through the concept of "possible worlds" connected by an accessibility relation. A proposition is necessarily true ($\Box P$) if it is true in all accessible possible worlds, and possibly true ($\Diamond P$) if it is true in at least one accessible world. This semantic innovation provided modal logic with a clear mathematical foundation while raising deep philosophical questions about the nature of possibility and necessity.

Epistemic logic uses modal operators to formalize knowledge and belief, where $\Box A$ might be read as "the agent knows that A". Deontic logic applies modal concepts to obligation and permission, while temporal logic (which we will examine in detail later) uses modal operators to reason about time. This flexibility has made modal logic an essential tool in fields ranging from philosophy to computer science, where it finds application in areas like program verification and artificial intelligence.

However, when we consider modal logic in relation to living systems and cyclical processes, certain limitations become apparent. While modal logic can express statements about necessity and possibility, it treats these modalities as static features of propositions rather than emergent properties of dynamic systems. The possible worlds framework, while mathematically elegant, spatializes possibility into a fixed structure of alternative states rather than capturing the way possibilities emerge through temporal processes. Moreover, the standard semantics assumes that the accessibility relation between possible worlds is given in advance rather than being something that might develop or change through time.

This becomes particularly evident when we consider how living systems generate their own possibilities through recursive cycles of self-maintenance. A cell's capacity to maintain itself, for example, isn't simply a static possibility that exists in some possible world, it's an active achievement that must be continuously renewed. The modal framework, while expanding classical logic's expressive power, still inherits its underlying commitment to a static, timeless view of truth and logical relationships.

Fuzzy Logic

Fuzzy logic emerged from a different challenge to classical reasoning: the recognition that many concepts we reason with are inherently vague or admit of degrees. Where classical logic insists on sharp true/false distinctions, fuzzy logic introduces degrees of truth that can lie anywhere between 0 and 1. This extension proves particularly relevant when considering living systems, where many properties and processes exist on continua rather than in discrete states.

The formal development of fuzzy logic began with Lotfi Zadeh's seminal 1965 paper on fuzzy sets[169]. Zadeh's insight was that many natural concepts like "tall," "hot," or "healthy" don't have sharp boundaries but rather admit of degrees of membership. This led to a formal system where propositions can be assigned truth values anywhere in the interval $[0, 1]$, rather than just the classical $\{0, 1\}$ or even the three values $\{0, \frac{1}{2}, 1\}$ of some earlier many-valued logics[170].

The mathematical structure of fuzzy logic centers on how these intermediate truth values are propagated through logical operations. The classical Boolean operations are generalized to functions on the unit interval $[0, 1]$. For example, the truth value of a conjunction $A \wedge B$ might be defined as the minimum of the truth values of A

and B:

$$\text{Truth}(A \land B) = \min(\text{Truth}(A), \text{Truth}(B))$$

Similarly, disjunction might use the maximum:

$$\text{Truth}(A \lor B) = \max(\text{Truth}(A), \text{Truth}(B))$$

And negation might be defined as:

$$\text{Truth}(\neg A) = 1 - \text{Truth}(A)$$

These are called Zadeh's operations, though others (known as t-norms and t-conorms) are possible[171].

A crucial feature of fuzzy logic is its treatment of the classical laws of logic. The law of excluded middle ($A \lor \neg A$) and the law of non-contradiction ($\neg(A \land \neg A)$) no longer necessarily have truth value 1 or 0 respectively. If A has truth value 0.5, then both $A \lor \neg A$ and $A \land \neg A$ will also have truth value 0.5 under Zadeh's operations. This reflects the idea that for vague predicates, there can be genuine indeterminacy or partial truth.

The development of fuzzy logic has led to significant applications, particularly in control systems. Fuzzy controllers, which use fuzzy rules to make decisions based on imprecise inputs, have been successfully employed in everything from automobile anti-lock braking systems to industrial process control. The ability to handle gradual transitions and imprecise categories makes fuzzy logic suited to situations where classical sharp thresholds would be artificial or problematic.

However, when we examine fuzzy logic's relationship to living systems and cyclical synthesis, we find both important insights and fundamental limitations. The key insight is the recognition that many biological properties and processes operate on continua rather than in discrete states. A cell isn't simply either dividing or not dividing, expressing a gene or not expressing it. These processes involve complex gradients and intermediate states that fuzzy logic can represent more naturally than classical bivalent logic.

Yet fuzzy logic's treatment of these continua remains static. While it can represent degrees of truth or membership, it doesn't inherently capture how these degrees emerge from or participate in dynamic processes. A fuzzy truth value, while more nuanced than a classical boolean, is still treated as a timeless feature of a proposition rather than as something that might emerge through temporal development or cyclical processes.

This limitation becomes evident when we consider how living systems generate and maintain their own gradients and intermediate states. Consider, for example, the process of cell differentiation in development. The transition from one cell type to another isn't simply a matter of having intermediate degrees of cell-type identity that could be captured by fuzzy membership values. Rather, it's an active process where the very categories of cell identity emerge through complex networks of genetic and epigenetic regulation, themselves operating through multiple interlinked cycles of activation and inhibition.

Temporal Logic

Temporal logic represents perhaps the most significant attempt within the classical tradition to directly address the challenge of reasoning about time and change. Where classical logic treats propositions as having timeless truth values and modal logic spatializes possibility into a framework of possible worlds, temporal logic explicitly introduces operators that reference temporal relations. This extension emerged from the recognition that many crucial forms of reasoning—particularly about processes, change, and causation—inherently involve time.

The modern development of temporal logic began with Arthur Prior's work in the 1950s and 1960s[172]. Prior, influenced by medieval discussions of future contingents and strongly motivated by philosophical problems about free will and determinism, developed what he called "tense logic". His innovation was to treat tenses as operators on propositions, much as modal logic treats necessity and possibility as operators. Prior introduced operators like G (it will always be the case that), F (it will at some time be the case that), H (it has always been the case that), and P (it was at some time the case that)[173].

The formal structure of temporal logic builds on these foundations but has evolved into several distinct systems. Linear Temporal Logic (LTL), which assumes a single timeline, introduces operators like:

- $G\varphi$: φ holds globally (at all future times)
- $F\varphi$: φ holds eventually (at some future time)
- $X\varphi$: φ holds at the next time step
- $\varphi U\psi$: φ holds until ψ holds

This allows for the formal expression of statements about sequences of events and the evolution of truth values through time. For example, the formula $G(p \to Fq)$ expresses that whenever p becomes true, q will eventually become true. A pattern crucial for reasoning about cause and effect, promises, or system responses[174].

A significant variant emerged with Computation Tree Logic (CTL), developed by Amir Pnueli and others in the late 1970s, which considers branching time structures where multiple possible futures might unfold from any given moment. CTL introduces path quantifiers E ("there exists a path") and A ("for all paths") that can be combined with temporal operators. This allows for expressing properties like "there exists a path where eventually φ" ($EF\varphi$) or "on all paths, globally φ" ($AG\varphi$)[175].

The semantics of temporal logic typically involves structures called *temporal frames*, consisting of a set of moments or states ordered by temporal relations. In linear temporal logic, this ordering is total. Each moment has a unique successor and predecessor (if we're working with discrete time). In branching time logics, the ordering becomes tree-like, with moments potentially having multiple possible futures. The truth of temporal formulas is then evaluated relative to these structures, providing a foundation for reasoning about time-dependent truth.

This framework has proven particularly valuable in computer science, especially for program verification and the specification of concurrent systems. Its ability to express properties like "something bad never happens" ($G\neg bad$) or "every request is eventually followed by a response" ($G(request \rightarrow Fresponse)$) makes it an essential tool for verifying that programs meet their specifications. The development of efficient model-checking algorithms for temporal logics has led to practical tools for automatic verification of complex systems [176].

However, when we examine temporal logic in relation to living systems and cyclical synthesis, we find both important alignments and crucial limitations. The primary alignment lies in temporal logic's recognition that truth can be time-dependent and that reasoning about processes requires explicit treatment of temporal relations. This represents a significant advance over both classical logic's timeless truths and modal logic's static possible worlds.

Yet temporal logic's treatment of time remains fundamentally spatialized. It treats time as another dimension along which truth values can vary, rather than as a generative force that shapes the evolution of systems. This becomes particularly evident in its handling of recursive or cyclical phenomena. While temporal logic can express that something happens repeatedly (using formulas like $G(p \rightarrow Fp)$ to say that whenever p occurs, it will occur again), it treats these cycles as external patterns observed in a pre-existing temporal structure rather than as constitutive processes that generate time and possibility through their own operation.

This limitation stems from temporal logic's inheritance of classical logic's underlying metaphysical assumption, that reality consists of discrete states possessing properties, with time serving only as an external ordering parameter for these states. Temporal logic treats time as a dimension along which we can index different system states, like spatial coordinates locate objects, but the states themselves remain conceptually independent of the temporal dimension that orders them.

Living systems exhibit a fundamentally different relationship to time. Cycles generate the system's own temporal organization through their operation. A cell's circadian rhythm unfolds in physical time, but it simultaneously creates the functional temporality by which the cell organizes itself: when genes are expressed, when metabolic processes accelerate or slow, when the cell becomes receptive to certain signals. The cycle's periodicity becomes the temporal framework through which the cell coordinates its activities. This represents time as organizational achievement, actively produced by the processes it coordinates, rather than an abstract dimension containing events. Temporal logic, despite incorporating temporal operators, cannot capture this reflexive structure. Cyclical processes do not simply occur in time but generate the temporal order by which the system maintains its coherence.

Paraconsistent Logic

Paraconsistent logic emerged from a fundamentally different challenge to classical logic, the recognition that inconsistency, while problematic in mathematical reasoning, might not always be fatal to rational discourse. Where classical logic treats contradiction as a logical catastrophe through the principle of explosion, paraconsistent logic seeks ways to contain contradictions while maintaining meaningful inference. This development is a crucial step toward understanding how systems can function coherently despite harboring apparent contradictions.

The formal development of paraconsistent logic began in the mid-20th century, though its philosophical roots can be traced to earlier thinkers who questioned classical logic's absolute ban on contradiction. Stanisław Jaśkowski's discussive logic and da Costa's hierarchical systems were among the first formal attempts to develop logics that could tolerate contradictions without triviality[177]. The term *paraconsistent* itself was coined by Miró Quesada in 1976 to describe logical systems that reject the classical principle that from a contradiction, anything follows[178].

The defining feature of paraconsistent logic is its rejection of the principle of explosion (*ex contradictione quodlibet*). Formally, in a paraconsistent logic, it is not generally true that:

$$A, \neg A \vdash B \tag{6.1}$$

for arbitrary B. This means that the presence of a contradiction in our premises does not force us to accept every proposition as true. This seemingly simple modification has profound implications for how we can reason in the presence of inconsistency.

Different paraconsistent systems achieve this containment of contradiction through various formal strategies. Some, like Jaśkowski's discussive logic, use a modal-like framework where contradictions are relativized to different *viewpoints* or *discourse levels*. Others, like the Logic of Paradox (LP) developed by Graham Priest[179], employ many-valued semantics where propositions can be both true and false without leading to trivialism. Still others modify the classical rules of inference, particularly those involving negation, to block the derivation of arbitrary conclusions from contradictory premises.

A particularly influential approach emerged with relevant logic, which requires a meaningful connection between premises and conclusions. In relevant logics, the fact that we have derived a contradiction from some premises A and B does not license us to derive an unrelated conclusion C. This captures the intuition that contradictions should only infect reasoning that is genuinely connected to the contradictory statements[180].

The development of paraconsistent logic is used in computer science, particularly in dealing with inconsistent databases or knowledge bases. When merging information from multiple sources, contradictions often arise, but we do not want these contradictions to render the entire system useless. Paraconsistent logic provides formal tools for continuing to reason productively even in the presence of such inconsistencies.

However, paraconsistent logic's relationship to living systems and cyclical synthesis showcases important insights and limitations. The key insight is the recognition that contradictions need not be fatal. Coherent reasoning and functioning can continue even in the presence of apparently contradictory elements. This aligns with the observation that living systems routinely coordinate seemingly opposing tendencies.

Paraconsistent logic's approach to contradiction remains fundamentally static. It provides ways to contain or isolate contradictions, to prevent them from spreading throughout our reasoning system, but it does not capture how contradictions might be productive or generative. In a living system, apparent contradictions are not just tolerated, they are essential to the system's continued existence. Contradictions function not as logical problems to be contained but as dynamic tensions that drive the system's cycles of self-maintenance.

Having examined classical logic and its major extensions, we can now articulate more precisely how Cyclical Synthesis offers a fundamentally different approach to understanding the logic of living systems. While each extension of classical logic addresses certain limitations of the classical framework, they all remain rooted in a paradigm of synchronic reasoning. Even when they introduce time (temporal logic), degrees (fuzzy logic), or ways to handle contradiction (paraconsistent logic), they treat these as features to be represented within a fundamentally timeless logical framework.

The crucial distinction lies in how Cyclical Synthesis approaches three key aspects: temporality, contradiction, and the nature of truth itself. First, regarding temporality, while temporal logic can represent time-dependent truths, it treats time as an external parameter or dimension along which truth values vary. Cyclical Synthesis, by contrast, recognizes time as something generated by the very operation of cyclical processes. Time emerges from the coordination of cycles rather than serving as a pre-existing container for them.

Second, concerning contradiction, where paraconsistent logic seeks ways to contain contradictions without trivializing inference, Cyclical Synthesis recognizes opposing forces as potentially productive elements in generating coherence. The opposition between synthesis and breakdown in cellular metabolism, for instance, isn't a logical problem to be managed but rather a dynamic tension that, when properly coordinated in time, enables the cell's continued existence. This marks a shift from seeing contradiction as a logical feature to be controlled to understanding it as a generative aspect of living processes.

Third, regarding truth and logical values, where fuzzy logic introduces degrees of truth as static features of propositions, Cyclical Synthesis focuses on how apparent stabilities and categories emerge from dynamic processes of self-maintenance. The truth of statements about living systems isn't found in timeless logical relations or even in degrees of membership, but in the ongoing achievement of coherence through cyclical organization. This shifts attention from truth as a property to truth as a process, something actively maintained rather than passively possessed.

These distinctions reflect a deeper difference in how Cyclical Synthesis conceives the relationship between logic and life. The classical tradition, even in its extensions, seeks to represent life's features within a logical framework. To capture living processes through ever more sophisticated logical apparatus. Cyclical Synthesis, by contrast, seeks to understand how logical coherence itself emerges from the organization of living processes. It suggests that before we can represent life's logic, we must understand how life generates its own forms of order through the temporal coordination of cycles.

This points toward the need for a logic of life that is itself derived from, and structured by, the principles that govern their existence: cyclicality, temporal integration, and the productive tension between opposing forces. The challenge, then, is not simply to extend classical logic further but to develop a kind of reasoning and logical thinking that begins from the recognition that coherence in living systems is always an active achievement, maintained through the coordinated operation of multiple cycles across multiple scales.

6.3.2 Continental Logic

Having examined the classical tradition and its extensions, we turn now to continental approaches to logic. While the analytic tradition emphasizes formal systems built on precise principles like non-contradiction and bivalent truth values, continental philosophy typically approaches logical structures synthetically, viewing particular ideas as embedded within larger dynamic wholes. Where classical logic strives for clarity through the elimination of contradiction, continental approaches often reframe contradictions as potentially productive forces in thought. This distinction, while painted here with broad strokes, helps situate the different approaches to logical structure we'll examine[181].

Hegel's Dialectical Logic

G.W.F. Hegel radically redefined logic's role in philosophy through his dialectical method[182]. Rather than a static set of inference rules, Hegel conceived logic as the dynamic development of concepts through contradiction and resolution. Where classical logic treats contradiction as failure, Hegel saw contradictions as inevitable and productive features of conceptual development. For instance, Hegel famously argues that motion itself is inherently contradictory. An object must be both here and not-here in the same moment of continuous movement. Rather than dismissing such contradictions as logical errors, Hegel sees them as revealing deeper truths about reality's dynamic nature. Each concept (thesis) generates its opposite (antithesis), with their tension overcome in a synthesis that preserves elements of both while resolving their conflict.

This dialectical progression is fundamentally iterative and teleological. Each synthesis becomes a new thesis, subject to further opposition, creating a developmental spiral aimed ultimately at absolute knowledge. This process of sublation (Aufhebung) is crucial, it simultaneously cancels, preserves, and elevates the opposed[183]. When being and nothing are sublated into becoming, for example, they are both negated as independent terms, preserved as moments within the new category, and raised to a higher conceptual level. For Hegel, this triadic movement describes the structure of reality's self-development. While this process might appear cyclical due to its repeating patterns, it represents a linear progression through cycles toward an ultimate endpoint. Each cycle integrates the truth of earlier stages into a more comprehensive framework.

Hegel's logic differs crucially from classical logic in its treatment of contradiction. Where classical logic forbids contradictions outright, Hegelian dialectics treats them as dialectical moments indicating the incompleteness of one-sided views. The tension drives thought to develop new concepts that can accommodate what the contradiction revealed. Rather than strictly adhering to non-contradiction, Hegel's logic allows that something can be, in different respects or from different perspectives, both itself and its opposite.

This makes Hegelian logic fundamentally synthetic and holistic. It attempts to overcome fragmentary analysis by revealing how each concept's truth lies in its connections to others, with everything ultimately interrelated in a rational totality. However, this synthetic drive maintains necessity in its progression, each stage follows logically from the previous. It's not arbitrary cycling but rational development toward comprehensive truth.

Deleuze's Logic of Difference

Yet this very success in resolving contradictions through synthesis would draw sharp criticism from later thinkers, particularly Gilles Deleuze, who saw in Hegel's dialectic a fundamental misunderstanding of difference itself[184]. Where Hegel treats difference primarily as opposition to be overcome, Deleuze seeks to think difference in itself, developing a *logic of difference* or *logic of multiplicities*[185]. Beyond a critique of Hegel, this opens an entirely new way of conceiving how elements can be productively related without losing their distinctness. His syntheses don't resolve into final unity but rather combine elements while preserving their differences.

This approach manifests in Deleuze's account of the three syntheses of time in his work, *Difference and Repetition*. The first synthesis of habit, describes how the present is constituted through *passive synthesis* of past moments. This cyclic repetition, like the day-night cycle or heartbeat rhythm, synthesizes time as a living present by contracting moments together. The subject emerges from this repetitive binding of moments, a passive synthesis occurring pre-consciously as we become ourselves through repetitive habits and expectations in time. This idea of passive synthesis constituting subjectivity is crucial: the self isn't a pre-existing container that experiences habits, but rather habits themselves, through their cyclic repetition, generate what we experience as a coherent self. Our sense of self emerges from the regular rhythms of experience being contracted together. We don't actively or consciously create these syntheses, instead, they happen in us at a level prior to conscious awareness. Deleuze draws here on Hume's insight that habit is the foundation of our beliefs and reasoning, suggesting that habit doesn't just shape an existing subject, it constitutes subjectivity itself. Each contracted moment leaves an imprint, and these imprints accumulate into the patterns we recognize as our preferences, dispositions, and ultimately our very sense of self. This passive constitution of subjectivity through cyclic processes points to that identity itself emerging from repetition instead of preceding it.

The second synthesis of memory, moves beyond present-focused habit to produce a *pure past* through active synthesis. It's not personal memory but an impersonal past coexisting with every present moment, giving it depth and meaning. Unlike cyclic habit, memory's synthesis orders moments sequentially into coherent narratives. This active synthesis is crucial for meaning-making. We actively connect disparate moments into meaningful patterns that were never present in raw experience. Deleuze, drawing from Bergson and Proust, shows how this synthesis creates a virtual dimension of time that coexists with the present, giving depth and significance to experience beyond mere habitual repetition.

The third synthesis of eternal return, reinterprets Nietzsche's concept as a philosophical logic where only difference truly returns. Deleuze frames this as a selective, affirmative principle. What returns is not the same content but the power of differentiation itself. This makes eternal return a test or ethical selector in that only what affirms difference and creativity can truly *return*. Unlike traditional readings of Nietzsche that see eternal return as cosmic repetition, Deleuze's interpretation makes it a principle of radical novelty. Rather than actual cosmic cycles bringing back identical events, Deleuze's eternal return means being is incessant variation. Whatever can vary will repeat differently. This makes his logic fundamentally open-ended. There's no final synthesis uniting all differences, only ongoing creative differentiation. The first synthesis provides continuity through unconscious repetition, but this very repetition generates difference. Each cycle brings slight variations that accumulate over time. This is particularly relevant to our concept of constraint cycles, as both frameworks recognize how stability emerges from repetition while allowing for adaptive change.

Deleuze called Hegel the archenemy of difference. Hegel's dialectic, in Deleuze's view, takes intense differences (being vs. nothing, etc.) and always resolves them in a higher identity. It subordinates true difference to a play of negation that ultimately shores up the *logic of the identical* (the final synthesis contains the differences only by reducing them to moments of itself). Deleuze wanted to break out of this "monocentrality of circles in the Hegelian dialectic". He argues that Hegel's method feigns a radical difference but actually predetermines a harmonious result. The obscure is already clarified from the outset in dialectics, he says, because we know it will end in unity. In contrast, Deleuze's logic does not treat negation as a productive engine. Instead of negating opposing terms to produce a synthesis, he affirms positive difference at each step. History and thought, for Deleuze, do not move by negation but by deciding problems and affirming differences. There is no pre-given final goal or truth, no teleology of Spirit, only a continuous generation of new differences and creative solutions[186].

Thus, Deleuze's logic is a non-dialectical synthesis. Multiple things come together, but without resolving into One. Instead of unity, the result might be a network or assemblage. In logical terms, Deleuze favors "both... and..." and "and... and...", linking heterogeneous elements without subsuming them. His work often invokes paradox and inclusive disjunction (e.g., something can be both X and Y in different senses, or an element can belong to multiple sets without contradiction). This is very different from classical logic's binary either/or. It even differs from Hegel's either/or/and (Hegel's X and ¬X get resolved in a higher-order X'. Deleuze would keep X and ¬X in play as distinct). In this sense, is Deleuze's analysis remarkably similar to what we have been working towards.

Whitehead's Process Logic

If Deleuze's approach emphasizes pure difference and Hegel's seeks absolute unity, Whitehead's process philosophy represents a unique third approach. His work is particularly significant because it comes from a deep engagement with both mathematical logic and metaphysical speculation. Starting as a mathematician who co-authored Principia Mathematica with Bertrand Russell[187], Whitehead radically departed from this work in his later projects. The dual heritage of mathematical precision and metaphysical ambition allows him to develop a sophisticated account of how process and relation can serve as fundamental logical categories. This metaphysical shift necessitates rethinking logical structure itself. Instead of timeless laws governing fixed things, Whitehead envisions interrelated processes creating patterns through ongoing development.

Whitehead argued that "reality consists of processes rather than material objects, and that processes are best defined by their relations with other processes". In effect, any entity is not an isolated self-identical thing (as classical logic might assume for an individual object a = a at all times), but a cluster of processes that emerge from and contribute to other processes. Everything is interconnected, a web of interrelated processes. For logic, this means the relations and transitions are more fundamental than static categories. Traditional logical principles (like non-contradiction, identity, excluded middle) are rooted in thinking about fixed entities in clear-cut categories. But if every entity is a process (a happening that has phases and interdependencies), then a richer logical framework is needed to describe how an entity can change and still be itself, or how two moments of a process can legitimately both characterize it despite differences.

Whitehead builds an elaborate schema in Process and Reality describing how processes, *actual occasions* in his terminology, come into being by *prehending*, taking account of, other processes and how they achieve a momentary synthesis of many influences, and then perish and pass those influences along[188]. The entire becoming of an occasion, its growing-together into a single experiential unit, is what Whitehead calls *concrescence*. Each actual occasion represents a small event of synthesis through *prehension*, a term introduced to describe how entities grasp or feel their environment. Prehension involves both physical feelings (causal influences) and conceptual feelings (eternal objects or pure possibilities). The actual occasion moves through distinct phases of concrescence. It begins with many potential ways of synthesizing its data, progressively eliminates possibilities through decision, and finally achieves satisfaction in a determinate outcome. This process, which Whitehead describes, shows how creativity can be both free and structured.

Whitehead saw the classical analytic approach as too narrow for capturing reality. He once quipped: "In formal logic, a contradiction is the signal of defeat, but in the evolution of real knowledge it marks the first step in progress toward a victory". This quote aligns him, with Hegel's view that real progress often involves overcoming contradictions rather than ignoring them. Whitehead's own logical thinking allowed for tensions and polarities in reality (e.g., he speaks of the *polarity* of the eternal and the temporal, the one and the many). Instead of declaring these oppositions unthinkable, he tries to incorporate them into a balanced schema.

His logic is fundamentally organismic. Sometimes referred to as the *logic of organism*, alluding to description of the world is more like a living network than a static hierarchy of categories. This involves synthesis in a different way, each event is a synthesis of the influences upon it (causal, conceptual, etc.). Unlike Hegel's synthesis, Whitehead's does not explicitly come from the clash of a contradiction, it comes from the merge of many inputs into a new unity (the *concrescence* of an occasion). Once an occasion is complete, it becomes a discrete outcome that can be a factor in other occasions' becoming. This yields a processual logic where cause and effect are not linear chains but overlapping cycles of many causes converging and then radiating out new effects. Reality is continually synthesizing itself, moment to moment.

In his early work in mathematics, Whitehead's approach to logic was noted to be algebraic in character. That is to say, Whitehead's focus was on relational systems of order and structure-preserving transformations. This algebraic mindset carried into his metaphysics, rather than focus on the truth value of isolated propositions, he is concerned with how the structure of reality can be described by patterns of relations (analogous to an algebra of events). He differs from analytic philosophers like Frege in that he's less interested in logical proof or static semantics and more in coherence and holistic order. In fact, Whitehead subordinated logic to what he described as the *aesthetic order*. He suggested that the universe's rationality (logical order) is a subset of a deeper aesthetic harmony. For him, the coherence of the world is more like a beautiful pattern than a formal derivation.

This approach shares elements with both Hegel and Deleuze while remaining distinct. Like Hegel, Whitehead sees reality as dynamic and interconnected, but focuses more on novelty than contradiction. Like Deleuze, he emphasizes creative advance, but structures it through more determinate patterns of relation. His logic is neither strictly dialectical nor pure difference but rather a cosmological logic of growth and interdependence.

These continental approaches to logical structure, while differing significantly, share important features distinguishing them from classical logic. They emphasize process over stasis, relation over isolation, and synthesis over analysis. However, they differ crucially in how they conceive synthesis itself: Hegel through dialectical resolution of contradictions, Deleuze through connection preserving difference, and Whitehead through creative integration of multiple influences[189].

Having explored the dynamic and often process-oriented approaches to logical structure found in the continental tradition we can now situate Cyclical Synthesis more precisely within this landscape. While sharing the continental critique of static, timeless logic and embracing dynamism and relation, Cyclical Synthesis offers a distinct framework derived from and tailored to the constraints observed in living systems.

Unlike Hegel's dialectical logic, which drives towards an ultimate resolution where contradictions are sublated (*aufgehoben*) into a higher unity, cyclical synthesis emphasizes the ongoing temporal coordination of opposing processes. It does not seek a final synthesis that resolves tension but describes how system coherence is maintained through the structured alternation of distinct, even opposing, phases. The constituent processes retain their identities within the cycle, rather than being subsumed into a new, unified identity. This makes cyclical synthesis inherently cyclical and adaptive, lacking the teleological drive towards an Absolute endpoint characteristic of Hegelian thought. Contradiction, for cyclical synthesis, is not a motor for linear progress via negation, but a fundamental tension managed through temporal sequencing for persistence and adaptation.

The resonance with Deleuze's thought, particularly his concept of passive synthesis, is striking. Both frameworks describe how coherence and identity emerge from repetitive, often unconscious, cyclical patterns rather than from a pre-existing subject or static structure. Both emphasize how repetition, far from being mere sameness, can generate difference and intensities over time. However, their origins and focal points differ. Cyclical Synthesis is explicitly derived from the analysis of physical and biological constraints (thermodynamics, self-organization, constraint cycles), grounding its structure in the material necessities of self-maintenance. Deleuze's syntheses, while drawing on examples from life and psychology, were developed primarily from a metaphysical and philosophical critique of representation and identity, aiming to think *Difference* in itself beyond the logic of the *Same*. Furthermore, while Deleuze champions the affirmation of difference and escaping molar structures via lines of deterritorialization towards a nomadic becoming, cyclical synthesis focuses more specifically on how the temporal coordination of opposition allows fundamentally different processes to coexist productively within a dynamic system, constituting a kind of ongoing, adaptive reterritorialization.

Similarly, while sharing Whitehead's profound emphasis on process, relation, and the rejection of static substance, Cyclical Synthesis offers a more specific mechanism for understanding coherence. Whitehead's logic of concrescence describes how actual occasions synthesize a multiplicity of influences (causal, conceptual, etc.) into novel unities. Cyclical synthesis, by contrast, zooms in on a particular pattern within this processual world. The way specifically opposing tendencies are integrated through temporal sequencing to create self-sustaining dynamics. While Whitehead develops a comprehensive cosmological metaphysics, Cyclical Synthesis maintains a more explicit awareness of its status as a thought figure or a model derived from analysing a specific class of systems (living/autogenic/adaptive cycles), rather than a complete ontology of all reality. Its logic is less about universal concrescence and more about the specific temporal logic required for self-maintenance against entropy via coordinated opposition.

What distinguishes Cyclical Synthesis within this continental milieu, therefore, is its specific grounding in the observable constraints of living systems, its focus on temporal sequencing as the primary mechanism for integrating opposition while preserving distinctness, and its relatively modest framing as a logical tool or structure for reasoning about such systems. It seeks to align with the constraints not of abstract reason alone (classical logic), nor of speculative metaphysics aiming for totality (Hegel, Whitehead), nor purely of difference itself (Deleuze), but the specific, pragmatic constraints governing systems that persist through dynamic transformation and coordinated opposition in time. In essence, Cyclical Synthesis attempts to articulate a logic of life derived directly from the challenges and solutions inherent in maintaining organization far from equilibrium, offering a framework where coherence is achieved not despite, but because of, the temporal integration of difference and opposition.

6.3.3 Pragmatist Logic

Perhaps no philosophical tradition aligns more naturally with our investigation than pragmatism, particularly in its understanding of logic as fundamentally rooted in the processes of life itself. Where our exploration has led us to collapse the distinction between ontology and epistemology, the pragmatists arrived at a similar insight through their careful attention to how knowledge emerges from and remains embedded in living processes. Both approaches recognize that our ways of knowing originate in our nature as living systems engaged in continuous interaction with the environment.

Charles Sanders Peirce, pragmatism's founder and most systematic thinker, developed this insight through his three-fold categorical scheme focusing on the question of how patterns of interpretation emerge and develop. His categories of *Firstness*, *Secondness*, and *Thirdness*, express the basic modes through which reality manifests[190]. Firstness represents pure quality or possibility, the way things are in themselves before any relation or comparison. Secondness involves reaction and resistance. The brute facts of existence where one thing encounters another. Thirdness encompasses mediation and law, the ways in which relations become regular and intelligible.

These categories bear resemblance in Deacon's hierarchy of emergent dynamics who has been deeply influenced by them. Thermodynamics (involving pure energy gradients), Morphodynamics (involving pattern formation through constraint), and Teleodynamics (involving self-maintaining organization). Just as Deacon shows how each level emerges from but transcends the previous one, Peirce's categories form a nested hierarchy where each level depends on but cannot be reduced to those below it. Both frameworks attempt to explain how higher-order organization emerges from simpler dynamics.

From this categorical foundation, Peirce developed a comprehensive theory of logic that extends far beyond traditional formal reasoning. For Peirce, logic was semiotic. The theory of signs and their interpretation. In Peirce's view, every thought is a sign that stands for something to some interpreter, making logical reasoning just one instance of a more general process of sign interpretation that permeates all of life. The significance of this move, seeing logic (as well as mathematics and other disciplines) as fundamentally symbol interpretation and manipulation, cannot be over-stated. For it re-situates any process of reasoning within the bounds of living systems being in the world.

Just as we found constraint cycles operating across multiple scales of biological organization, Peirce's semiotics shows how sign interpretation functions at every level from cellular signalling to human abstract thought. A cell interpreting a chemical gradient, an animal responding to environmental cues, and a scientist testing a hypothesis all participate in what Peirce called semiosis, the action of signs. Similar to what we have argued, Peirce suggests that logical reasoning is not some capability added to complex mammalian life but rather an elaboration of processes inherent to life itself.

The pragmatist collapse of the knowing/being distinction manifests particularly clearly in Peirce's theory of inquiry. Rather than seeing truth as a static correspondence between propositions and reality, Peirce understood it as the ideal limit of inquiry, what would be believed if investigation were carried indefinitely far. This makes truth something that emerges through process rather than existing as a pre-given state. There is a parallel to our understanding of how constraint cycles maintain coherence through continuous transformation.

This processual understanding of truth leads to a distinctive view of logical validity. Where classical logic treats validity as a timeless property of argument forms, pragmatist logic sees it as emerging from the success of inquiry methods in practice.

A mode of reasoning is valid to the extent it helps us navigate experience and solve problems effectively. This doesn't mean abandoning formal logic, Peirce himself made significant contributions to mathematical logic, rather recognizing that formal systems derive their authority from their effectiveness in organizing experience.

The pragmatist emphasis on practice and consequence finds particular resonance in Peirce's famous maxim: consider what practical effects we would expect from an object to form our whole conception of it[191]. This principle transforms logic from a study of abstract relations into an investigation of how ideas guide action. When we understand a concept, we understand how it would lead us to act in various circumstances.

The pragmatist understanding of logical necessity also deserves some attention. Where classical logic treats necessity as an eternal, context-independent property of valid inference, pragmatists see it as derived from the patterns of successful inquiry. When Peirce speaks of logical necessity, he means the kind of constraint that emerges through the *fixation of belief* - the process by which inquiry settles on stable patterns of inference[192]. This resonates with the dynamic unfolding of constraint cycles, where constraints emerge from and are maintained by the very processes they constrain to stabilize into forms that have the ability to continuously propagate their set of constraints. In both cases, necessity isn't imposed from outside but emerges from the dynamics of the system itself interacting and relating to its surrounding.

John Dewey further developed this processual understanding of logic in his theory of inquiry[193]. For Dewey, logical forms (like categorical statements or syllogisms) are also not eternal structures but tools that have evolved because they help resolve problematic situations in a functional manner. He saw inquiry as a continuous process of transforming indeterminate situations into determinate ones. Here again we find a process that bears resemblance to the theory developed in the previous parts of the book. Coherence in living systems and in logical reasoning is in both cases not conceived as system-independent, but a feature of that very system, either through adaptedness or problem resolution. As living systems must actively maintain their organization through cyclical processes, inquiry must actively maintain its coherence through what Dewey's *pattern of inquiry*.

This pattern involves several phases that mirror aspects of cyclical synthesis. First comes the recognition of a problematic situation, something has disrupted the smooth flow of experience. This leads to the formulation of a problem (*abduction* in Peirce's terms), followed by the development and testing of possible solutions (the experimental phase). Finally, if successful, the inquiry reaches a resolution that transforms the indeterminate situation into a determinate one. The resolution becomes part of the context for future inquiries, just as each cycle in a constraint cycle becomes the condition for subsequent cycles.

The parallels between Dewey's pattern of inquiry and our concept of cyclical synthesis run deep. Both describe processes that maintain coherence through transformation rather than stasis. Both recognize that disruption, while potentially threatening, is also essential for growth and adaptation. And both see the resolution of tension not as elimination but as temporal coordination - the achievement of organization through process rather than the imposition or expectation of unity.

This temporal aspect of pragmatist logic deserves particular emphasis. Classical logic operates in an eternal, synchronic present, treating all truths as simultaneously valid. Pragmatist logic explicitly incorporates time as a fundamental dimension. Truth emerges through the act of inquiry. Meaning develops through interpretation. Logical necessity itself and the rules of inference evolves through the fixation of belief. This temporality is paramount to how pragmatist logic understands rationality itself.

This semiotic approach to logic has found powerful extension in contemporary *biosemiotics* and *cybersemiotics*. Cybersemiotics, developed by Søren Brier, further extends this insight by integrating Peircean semiotics with systems theory and cybernetics[194]. This theory focuses on how sign processes can generate and maintain the kind of self-organizing coherence we've identified in constraint cycles. Where cybernetics describes the mechanics of feedback and control, and semiotics explains meaning and interpretation, cybersemiotics shows how these aspects are necessarily intertwined in living systems. A living system must not only maintain its physical organization but must also maintain coherent patterns of interpretation, what Brier calls *meaning coordination*.

This meaning coordination happens through what Brier terms the *cybersemiotic star*. A model that integrates four domains: physical nature, subjective experience, intersubjective meaning, and the abstract patterns we recognize as logic and mathematics. Rather than trying to reduce one domain to another, cybersemiotics shows how they emerge together through processes of interpretation and constraint. This aligns remarkably well with our understanding of cycles operate across multiple scales, generating higher-order constraints while maintaining distinct processes.

Biosemiotics has developed these insights further, showing how sign processes operate at every level of living systems. Following Peirce, biosemioticians recognize that signs always involve triadic relations. Something stands for something else to some interpreter. But crucially, this interpretation doesn't require consciousness or intention in any conventional sense. Even a cell interpreting a chemical gradient exhibits this triadic structure: the chemical gradient (sign) represents environmental conditions (object) to the cell's regulatory systems (interpreter).

This biosemiotic perspective aligns in many ways well with Deacon's analysis of constraint. Deacon shows how constraints can become constraints on constraints, creating higher-order organization. This is also underscored in biosemiotic theory which articulates how sign relations can become nested, creating increasingly sophisticated levels of interpretation. A hormone molecule doesn't simply trigger a cellular response, it participates in complex regulatory networks that maintain organismic homeostasis. The interpretation of the hormone at the cellular level becomes part of larger interpretive processes at the tissue and organism levels.

The work of Jesper Hoffmeyer and Thomas Sebeok has been particularly important in developing these connections. Hoffmeyer's concept of *semiotic scaffolding* shows how sign processes create the conditions for their own elaboration, much like how constraint cycles generate higher-order constraints that shape their component processes[195]. Sebeok's broader vision of zoosemiotics[196] and biosemiotics demonstrates how sign interpretation isn't unique to human language but permeates all of life, from bacterial communication to ecosystem dynamics[197].

Overall, what the semiotic tradition, especially with the renewed emphasis on biology, teaches us is that living systems don't process information in a mechanical sense. They actively interpret signs in ways that maintain their organization. This interpretive capacity is essential to how living systems persist and develop. As Hoffmeyer argues, life and semiosis are coextensive - you can't have one without the other[198].

The implications for understanding the relationship between life and logic are important. Rather than seeing logic as an abstract system imposed on reality, the pragmatist tradition, shows logic as emerging from the basic processes of life itself.

This understanding reshapes how we conceive the relationship between life and logic. Instead of asking how abstract logical systems apply to living processes, we see how logical structure emerges from and remains grounded in the basic patterns of life itself. This doesn't diminish the power or importance of formal logic but rather unearths its deep connection to the dynamics that make life possible. The pragmatist tradition, offers not just a different approach to logic but a deeper understanding of how life and logic are inextricably intertwined in the maintenance and development of living systems.

6.4 Limitations of Cyclical Synthesis

Having situated cyclical synthesis within (our very limited selection of) broader logical traditions, we must now carefully examine its limitations. The following analysis serves to acknowledge weaknesses, but beyond that, to understand the framework's proper domain of application and potential for further development. Hence, we will explicitly articulate where and how the theory might fail or require modification. Since cyclical synthesis is inexorably linked to the constraint cycles we have described previously, this discussion on limitations in many cases relates to both models.

First, we encounter fundamental modelling challenges inherent to any theoretical framework. While our earlier analysis collapsed traditional distinctions between ontology and epistemology through the lens of constraint cycles, we must still grapple to some extent with the classic problem of representation, namely, how well can our model represent the constraints it seeks to describe? The goal here isn't perfect isomorphism between map and territory (an impossible standard that misunderstands the nature of models) and, strictly speaking, it isn't literal mirroring either. What a model must provide is a set of internal constraints that license coherent and robust patterns of inference about the target system. We can therefore speak of *constraint alignment* only in a loose, pragmatic sense, the model should capture enough of the system's dynamic scaffold to support reliable reasoning while openly acknowledging inevitable simplification.

This leads to the first specific challenge, identifying when processes become mutually enabling rather than merely coexistent or mutually impeding. While cyclical synthesis provides a framework for understanding how opposing processes can generate higher-order constraints, it offers limited guidance for predicting which specific process combinations will achieve such integration. Many processes may appear complementary or opposing without generating the required kind of productive tension. This limitation points to a non-reductive problem in complex systems modelling. While one can develop better criteria for identifying potentially synthetic relationships, one can never achieve perfect precision in predicting how processes will interact. This follows from the inherent limitations of modelling complex processes, even in relatively simple cases, the interaction of two processes may generate emergent behaviours that couldn't be predicted from analysing each process in isolation. At best, we can achieve approximate predictions for certain well-constrained cases, but complete predictability remains fundamentally impossible.

The framework's high level of abstraction presents both advantages and challenges. While this generality enables application across diverse domains, it creates difficulties for empirical operationalization. How do we measure opposing processes or quantify temporal coordination in specific cases? This reflects a broader tension in theoretical biology between explanatory power and empirical testability. The framework might explain patterns we observe but generating novel, testable predictions requires additional theoretical development.

The risk of oversimplification looms large, particularly when dealing with complex living systems. While cyclical patterns undoubtedly play crucial roles in biological organization, coarse-graining biological phenomena to cycles risks missing important non-cyclic aspects of living systems. This reflects a deeper philosophical tension between the need for explanatory frameworks and the irreducible complexity of biological phenomena.

Temporal scale ambiguity presents another significant challenge. While the framework emphasizes temporal coordination, it provides limited guidance for identifying relevant timescales in specific cases. This becomes particularly problematic when dealing with nested cycles operating at different speeds, a common feature of biological systems. How do we determine appropriate temporal boundaries for analysis? This limitation connects to broader issues in biological timing and the challenges of studying multi-scale temporal phenomena.

The framework's claims of scale invariance, while intriguing, require careful scrutiny. Are the mappings between different scales *homologous*, sharing fundamental mechanisms, or merely *analogical*, exhibiting surface similarities? While certain patterns may recur across scales, the specific mechanisms and emergent properties likely differ substantially. This raises questions about the framework's explanatory power at higher levels of organization. Does it become so general as to risk tautology?

Boundary definition presents persistent challenges, particularly in highly interconnected systems. Where does one cycle end and another begin? The framework's emphasis on temporal coordination over spatial boundaries makes this particularly challenging. This reflects deeper issues in biological organization, the difficulty of defining clear boundaries in systems characterized by multiple overlapping processes and dependencies.

The framework's treatment of emergence, while conceptually rich, may be practically limiting. While it describes how higher-order constraints can emerge from cyclic processes, it provides limited criteria for predicting when such emergence will occur. This makes the framework more useful for retrospective analysis than prospective prediction. A significant limitation for practical applications.

Finally, methodological challenges arise from the framework's emphasis on dynamic, temporal processes. Traditional scientific methods, often based on static snapshots or linear cause-effect relationships, may be inadequate for studying cyclical synthesis empirically. This points to the need for new methodological approaches capable of capturing dynamic, multi-scale temporal phenomena.

Many of these limitations will be addressed in subsequent chapters, through detailed examination of specific applications and case studies. However, following the pragmatist tradition, we should evaluate the framework primarily through its usefulness rather than its claim to absolute truth. No theoretical framework is without limitations. The question is whether it provides productive ways of understanding and investigating phenomena of interest. In this light, the limitations identified here serve not to invalidate the framework but to clarify its proper domain of application and guide its further development.

These limitations also point toward future research directions, developing more precise criteria for identifying synthetic relationships, creating new methodologies for studying dynamic temporal processes, and exploring the specific mechanisms that enable scale-invariant patterns while accounting for scale-specific differences. The framework's limitations thus become opportunities for theoretical and methodological innovation.

Ultimately, cyclical synthesis and its derivation from constraint cycles should be understood as a theoretical tool for reasoning about aspects of biological organization and process. Its value lies in its ability to reveal patterns and relationships that might otherwise remain obscure and nebulous.

6.5 Constraining Paradox

> *True as it is that a cell is both the producer and the produced which embodies the producer, this duality can be pictured only when we represent for ourselves a sequence of processes of a circular nature in time. Apparently our cognition cannot hold both ends of a closing circle simultaneously; it must travel through the circle ceaselessly. Therefore we find a peculiar equivalence of self-reference and time, insofar as self-reference cannot be conceived outside time, and time comes in whenever self-reference is allowed.*
> — *Varela, A Calculus for Self-Reference*[199]

The recognition that forms of paradox, opposition and conflicting tensions are central drivers of change has shown up consistently through our investigation. This is fundamentally a matter of attention, we can focus either on how opposing processes generate conflict and interference, or on how they create productive tension and enable novelty. Both perspectives capture something relevant about living systems. When we attend to conflict, we see processes working against each other. When we shift our attention to their coordination, we see how these same processes enable and maintain each other through time. Nonetheless, these focal points of paradox in formal systems and of opposing processes in living systems often serve as potent gateways for understanding, as they delineate what a system is and what it is not.

In formal systems, paradoxes emerge from self-referential structures that resist resolution. A sentence like "This sentence is false" generates an oscillation between truth values when we attempt to evaluate it. From a diachronic perspective, the sentence has a definite truth value at any given moment. But from a synchronic perspective, no overarching truth value can be established. The paradox emerges precisely from attempting to flatten temporal oscillation into a timeless value.

What distinguishes living from formal systems is how they transform such potential paradoxes into productive cycles. While formal systems oscillate between states when encountering self-reference, living systems harness similar tensions to maintain their organization through constraint cycles - the temporal coordination of processes that would be contradictory if attempted simultaneously.

This foregrounds something about the relationship between time and paradox, captured in Varela's epigraph. Outside of time, certain paradoxes appear unsolvable. Yet from any momentary perspective, paradox dissolves into a particular state. A given sentence is at a given moment of interpretation false, then right. Finding a global truth-value is impossible, but not a momentarily one. It's only when we attempt to capture our nature in a timeless snapshot that paradox emerges[200].

Similarly for us as living systems, being awake or asleep can be states we are in, but never so simultaneously. We don't continuously experience ourselves as walking contradictions. Even though on some level we clearly are.

Whether in formal systems through strange loops across logical types, or in living systems through higher-order constraints coordinating lower-level processes, self-reference introduces a form of closure that generates higher order constraints we call either paradox or life. In formal systems, self-reference can lead to undecidability. In living systems, it can enable the maintenance of organization through time.

The cyclical synthesis we've developed offers a framework for understanding these relationships. Rather than seeing paradox as a problem to be overcome, it frames paradox as intrinsic to life itself. Not in some mystical sense that underscores life's ineffability, but as a concrete feature of temporal organization that resists reduction to static states.

The central insight emerging from this investigation is that while central features of life are explainable and analysable, they resist reduction to any single moment or state. Like we cannot see bird and duck simultaneously in Escher's drawing, or know the sentence "I am false" to be both right and wrong at once, we cannot know life outside of time.

Part III

Discontinuity

Chapter 7

Incompleteness

In our journey through the labyrinth of living systems, we've traced the contours of constraint cycles, danced with diachronic perspectives, and flirted with the idea that purpose isn't just for poets and philosophers. Now, we find ourselves at a curious crossroads where silicon meets synapse, where algorithms aspire to awareness. The question du jour: Are our chatty digital companions - those large language models that finish our sentences and sometimes our thoughts - knocking on the door of consciousness? It's a query that would make even Laplace's demon scratch its omniscient head. As we peer into this computational conundrum, we'll need all the tools in our conceptual toolkit - from the self-organizing principles of autocatalytic sets to the hierarchical hijinks of strange loops. Buckle up, dear reader, for we're about to embark on a mind-bending expedition into the binary backwaters of artificial intelligence, where the lines between processing and perceiving blur like a Dalí clock in the midday sun[201].

The above paragraph was written by an AI, after I fed it the last chapter, explained to it what I want to write about here, and told it to copy and exaggerate my style of writing. Probably you could tell, but maybe you just didn't notice. I could have spent more time on finding a perfect prompt and maybe something not so cringe and verbose would come out, with the use of metaphors dialled back a bit. Anyways, that's not too important, but it is interesting and remarkable that the advances in so called Large Language Models (LLM), the type of generative AI that produced the above text, have brought us to this point. While they have their limitations for generating some texts, they are incredibly helpful for doing editing work, for summarisation of documents and I use them everyday when I am coding. The amount this tech paradigm has changed my, and many others, work-life is startling and was foreseen by almost none, even a few years ago. This shift has not happened unnoticed and our media landscape has been flooded with stories about it, and speculations about its implications.

The growth-oriented venture capitalists of Silicon Valley have gone on about scaling-laws[202], and how making these models ever bigger will make them more and more intelligent until ultimately, when just enough money was spent on compute, they reach a magical inflection point and become conscious ushering in a new era of prosperity for humankind. Governments and legislators around the world have met these advances with significantly more resistance, as the incorporation of novel, potentially disruptive tech is any bureaucrats nightmare, especially since LLMs are generative and produce content, making them inherently less controllable and predictable, than an AI system that only helps with unlocking your phone by recognising you. A less-voiced perspective—perhaps because, from that vantage, speaking feels futile—is the fatalist's, sketching doomsday tableaux of an impending machine revolt poised to topple carbon-based life. The voice of the egalitarian and communitarian has also made itself heard, as it sharply dismantled the naively optimistic views of those selling AI systems, pointing to the history of technological implementation, most often to the smart-phone and social media, and its acidic quality of dissolving social bonds, trust and cohesion. These positions articulate the fundamental cultural orientations we encountered earlier, the individualist, hierarchist, fatalist, and egalitarian ways of organizing that shape societal responses to change.

While understanding these co-stabilizing cultural dynamics provides valuable context, our primary concern lies in applying our theoretical framework to the question: What distinguishes living from non-living systems in the context of computation and artificial intelligence? Having developed tools for understanding life through constraint cycles and cyclical synthesis, we can now examine whether LLMs exhibit similar patterns. We'll be guided by three questions. First, are our current AI systems alive? Second, why is AI so alluring and powerful and is AI-alignment a serious problem? Finally, what would be required to develop genuine living systems on silicon substrate?

Before going into these questions, it's important to note that the conversational qualities of Large Language Models (LLMs) and their current designs in a question-and-answer format are deliberate choices aimed at mimicking human-like interactions as closely as possible. It's no surprise, then, that these systems have ignited the most intense debates about whether they can be considered living or non-living entities. While we'll explore this particular question later, it's crucial to first examine the more fundamental concept of computation. This broader perspective will not only help us outline the key arguments but also provide a more timeless framework that goes beyond the current AI paradigm and specific architectures like LLMs.

To build this framework, we'll first examine the nature of computation itself, revealing how computational systems necessarily operate within a *small world*, a bounded context that requires interpretation from the larger world[203]. This distinction will prove crucial as we explore the fundamental differences between computational and living processes, from the hardware-software divide to the question of boundary conditions and teleology. Only then can we properly situate current AI systems within this broader theoretical landscape and understand both their capabilities and their inherent limitations.

7.1 On Computation

As someone who studied physics, I've observed a curious trend among some of my colleagues. The term *computation* has become a catch-all phrase for describing various processes in the universe. This linguistic shift reflects the growing influence of computational thinking in the field, but it often leads to imprecise and potentially misleading descriptions of physical phenomena[204]. This trend is of course not confined to physics only but shows up in many places.

To ground our discussion, we must first define computation in the Church-Turing sense: the systematic manipulation of symbols according to well-defined rules, where both the symbols and rules require interpretation to have meaning[205]. This differs fundamentally from physical processes that merely transform states according to natural laws.

The roots of the computational world-view can be traced back to the mechanistic perspectives of Newton and Descartes. These thinkers laid the groundwork for what we have already described as the *synchronic view* - a perspective that attempts to capture the entirety of a system in a single, timeless snapshot, with a set of laws or in this context algorithms acting upon it, providing it with deterministic ways to unfold. Computation, in many ways, represents the ultimate expression of this synchronic world-view, offering a means to calculate the systems dynamics with unprecedented precision. This precision is of course a tautological feature of the model, since we assume that in our computational system, all that exists are discrete states and their transformation rules. Hence all the future states of the system are really encoded in the initial states and transformation laws.

The cornerstone of modern computational theory is the concept of the universal Turing machine, a theoretical device of profound elegance and power. This abstract machine consists of a reading head that moves along an endless tape divided into cells, each containing a symbol. The machine operates based on a finite set of rules, or its program.

What makes the Turing machine so powerful is its simplicity coupled with its astonishing versatility. Despite its basic components of a tape, a head, and a set of rules, it can simulate any algorithmic process imaginable. The machine reads a symbol from the tape, changes its internal state based on that symbol and its current state, writes a new symbol on the tape, and moves left or right. This simple set of actions, when permuted in various ways, can perform any computation that can be precisely defined.

Turing effectively created a blueprint for a machine that could, in theory, perform any calculation or simulate any other computational device. It's as if he had invented a universal tool capable of becoming any other tool needed for any task. This concept of universality is what sets the Turing machine apart from other machines and makes it so profound.

The universality of the Turing machine, however, comes at a price. It must operate within a *small world*, a closed system with finite states and predetermined rules. While this limitation enables its precision and predictability, it also marks a fundamental distinction from living systems, which operate in an open-ended *large world* where new states and rules can emerge through interaction with the environment.

Building on Turing's work, John von Neumann developed the architecture that forms the basis of most modern computers. The von Neumann architecture replaces the infinite tape with finite random-access memory and external storage, connected by input-output channels to central processing units. This design brings the abstract Turing machine closer to physical realization, though still bounded by practical limitations. The details here are not central to the argument, but there is something deeply brilliant and marvellous behind the creation something like the *ultimate machine*, a device capable of simulating the actions of any other device and performing any calculation. It only seems natural that these ideas have an inert attraction leading to speculation on how this concept extends.

Hence, the power of these computational models has led some thinkers to push the metaphor to its limits, giving rise to pancomputationalism, the view that the universe itself is fundamentally computational in nature, or, simply a giant computer as in the infamous *simulation hypothesis*, according to which we are just a bad game of Sims[206]. Proponents like Konrad Zuse, Rolf Landauer, John Wheeler, David Deutsch, Seth Lloyd and Stephen Wolfram have argued that all physical processes can be understood as forms of information processing or computation[207].

While there are important technical arguments to be made about the extent to which physical laws are computable (there are persistent troubles with infinities in general relativity and randomness in quantum mechanics) and deep questions to be asked about the pan-computationalists' assumption that all physical processes are ultimately discrete, I won't go into these debates here[208].

The argument that I want to make does not even require technical details about physics, to show that equating all physical processes with computations amounts simply to a category error. A category error occurs when we attribute properties to something that cannot possibly have those properties due to its fundamental nature. For example, it would be a category error to ask about the colour of jealousy or to inquire about the political beliefs of a rock[209].

To illustrate this in the context of computation, let's consider a phyiscal system like a waterfall or a cloud. Pancomputationalists will argue that a cloud is performing computation simply via the movement of its particles, or that the water-molecules of the waterfall are performing a computation of sorts. Indeed, it's possible to use a waterfall to perform computations, there are great examples of YouTubers building logic gates and rudimentary computers using the flow of water worth a watch[210]. However, this doesn't mean the waterfall itself is computing in any meaningful sense.

The key here is interpretation. For a process to be computational, it needs more than just the physical flow of information. It requires meaning and interpretation. On a Turing machine, the symbols on the tape have significance because we assign meaning to them. Without this act of interpretation, the waterfall (or any physical process) is simply unfolding according to the laws of physics.

We can impose computational meaning onto the waterfall by carefully controlling its flow and interpreting the results, just as we can with any physical system. In this sense we can use the waterfall to calculate digits of Pi or run a Large Language Model. But this act of interpretation is external to the waterfall itself. The water doesn't *know* it's computing, it's following the path of least resistance under the constraints we've imposed. There needs to be a living system for whom the output of the computation has a meaning. Whose persistence and existence in the world can ultimately (even in a minimal sense) be coupled with the outcome of that computation.

This is why equating all physical processes with computation is a category error. Computation, in its essence, requires symbolic manipulation and interpretation, properties that belong to the realm of living systems, not to the realm of pure physical processes. To say that a waterfall computes is to mistakenly attribute cognitive properties to a non-cognitive entity.

Importantly that does in no way imply that the computation is itself conscious, it is just a constrained process unfolding in accordance to the physical laws. The set of constraints given to that computation were however selected from a vast space of possible ones, to be of utility for someone. So the computation is the selection of specific constraints for a particular goal. In other words, while we can use physical processes like waterfalls to perform computations, this doesn't mean these processes are inherently computational. The computation occurs in our interpretation and use of the system, not in the system by itself. This distinction is crucial for understanding the limits of the computational metaphor in describing the universe.

So we come to see that we need to be a bit more careful in how we use the concept of computation. Computation is not the same as all processes occurring inside the universe, and it is very unclear to what extent we can even simulate these processes with computational procedures as the universe may not be discrete down at the lowest levels[211]. Computational procedures are highly constrained processes occurring for an intended outcome for a system that is capable of conceiving of intended outcomes, an interpreting or living system. The computational system itself however, is thereby not becoming a living system, much rather, it is a very universal and versatile tool being utilised by a living system.

This fundamental difference between the closed, deterministic world of constructed constraints where computers operate and the open-ended world where living systems exist, points us toward a deeper understanding of why computers and organisms inhabit fundamentally different realms. This is what the next section will be about.

7.2 Size Matters

James P. Carse, in his captivating book *Finite and Infinite Games*[212], presents us with a deceptively simple dichotomy:

> **1**
>
> There are at least two kinds of games. One could be called finite, the other infinite.
>
> A finite game is played for the purpose of winning, an infinite game for the purpose of continuing the play.

This elegant distinction between open and closed systems, between the infinite and the finite, serves as a lens through which we can examine the central tension that occupies the pages of this text: the boundary between life and non-life.

Finite games, or closed systems, are bound by established temporal and spatial boundaries. They operate within a set of fixed rules, inevitably leading to a conclusion. Someone wins, loses, or ties. In the realm of computation, these are the systems that eventually halt, producing a definitive outcome.

We've encountered something very close to the concept of the finite game before in physics, in the form of the closed system. Picture an idealized system with fixed spatial boundaries, perhaps containing something as simple as hot and cold gases. Or, for a more provocative image, imagine a living being say, a cute cat, trapped within. In both cases, the story unfolds with the inexorability of a Greek tragedy: death or - as a physicist would call it - equilibrium invariably marks the finite game's end.

> **6**
> While finite games are externally defined, infinite games are internally defined.

The larger world must always define the rules for the smaller world. The territory constrains how a good cartographer makes a map. The map never forms the territory. Only a player in the larger world can attempt to remodel the world after a map.

Likewise for computation, only regularities and processes in the larger world can be used as the basis for algorithmic procedures. A subset of the dynamics we find in the larger world can be linked to each other in highly precise and specifically constrained ways to generate patterns to which we can ascribe meaning. This is the story of the transistor, the logic gates, and the code they are running. In principle, many regularities, like the waterfall, could be used to build universal Turing machines.

The rules, as well as the purpose of the computational procedure, are always set in a larger world, never in the computer itself.

> **10**
> It is on this point that we find the most critical distinction between finite and infinite play:
>
> *The rules of an infinite game must change in the course of the play.*

When we speak of a system's state space, we refer to the complete set of possible configurations that system can adopt. For a computer, this means the finite combinations of binary states across its memory. For living systems, the state space includes not just current configurations but potential future states that may emerge through growth, learning, and evolution.

The state space of Newtonian mechanics is fixed, per definition. It is the space of all possible states of a given system. Thereby the finite world of the system is delineated. The state space of your computer is also fixed. There is a vast, although finite, number of states your computer can be in.

The state space of no living system is fixed. It is neither fixed in time nor in space. Let's start with time. Think of your state space as a child. Even if you stretched as tall as possible, there was a time when you simply couldn't reach the top shelf. Same thing goes for a few years down the line when you simply won't be able to jump as far as you can now. If we consider evolutionary time scales, things get even more intense. Just compare your state space with that of your single-celled ancestor. Or the state space of our first ancestor crawling on land, of those climbing up the trees, as well as those who left the jungle for the savannah.

Spatial boundaries are equally problematic as they actually require finite spatial boundaries. Where would those be for you? Your skin? Your social environment? You plus all the tools you possess? But then what happens when you train and get stronger? What happens if you make new friends that unlock completely new possibilities for you? What if you acquire a new skill or get shinier, better tools?

As thermodynamically open systems, we cannot find a fixed amount of matter and a fixed set of rules that turn us into finite, closed systems. This is one of the central differentiating criteria for living systems and those that are not. No computer or LLM can change its own state space over time. No closed system governed by a fixed set of rules can.

80

> The result of approaching nature as a hostile Other whose designs are basically inimical to our interests is the *machine*, while the result of learning to discipline ourselves to consist with the deepest discernible patterns of natural order is the *garden.*

The machine is the closing off of an infinite game, an open system. It is the act of focusing on the part that is predictable and controllable for a specific means. It is the selection of certain parts of reality over the rest. Notice, however, that:

> **80**
> Machine and garden are not absolutely opposed to each other. Machinery can exist in the garden quite as finite games can be played within an infinite game. The question is not one of restricting machines from the garden but asking whether a machine serves the interest of the garden, or the garden the interest of the machine.

Here we find the fallacy of mistaking the map for the territory again, as well as associated dangers with it. The model of the world, be it machine or computer, is never the world. The model is always a cut-off subset, a finite game within an infinite world. It can therefore never contain all complexity and all the dynamics of the larger world. To then reduce the larger world to the smaller, to wish those parts that do not fit away, to neglect the dynamics not present in the model, can have deeply troublesome implications.

To model your fellow, be it human, animal, or ecosystem, as a machine will rob her of her potential, creativity, and ability to participate in the infinite game. Therefore, making the category error of mistaking living for non-living can have gruesome consequences. But importantly, machinery is not inherently bad and a new romanticist movement will not cure all evil. Machine and computer have their finite roles to play, their specific tasks to fulfill in the growth of the garden.

> **81**
> The most elemental difference between the machine and the garden is that one is driven by a force which must be introduced from without, the other grown by an energy which originates from within itself.

Here we mustn't understand the word *energy* as we have, through the eyes of physics. We interpret it as a metaphor for the difference in growth and developmental dynamics, where we do indeed find one more central difference between the computer and the organism. One is assembled, the other differentiated.

For your computer, rare-earths are mined, purified to an absolutely unbelievable degree, and assembled at the nanometer scale with precision beyond most minds' imagination. Only the slightest impurities or mistakes in assembly make the computer unusable. The design and assembly of the hardware must be done exactly according to plan for the computer to work. The whole is assembled out of the multiple parts to bring forth the whole.

For you, sperm and ovum conjoin. Out of that primordial cell, your process of differentiation begins. From one, to two, to four, to eight... Each cell expressing itself in relation to the whole. The organism, differentiating from a prior, less complex state, towards a more differentiated one. The whole is always there. The whole is not assembled but continues to differentiate and grow. The part of the whole is always already integrated into the whole.

These logics are fundamentally different.

> **101**
> There is but one infinite game.

And in this infinite game of life, computational systems - despite their power and utility - can only ever play finite games, bound by their fixed rules and predetermined state spaces.

7.3 Self-Modification

The above metatheoretical detour has shown that size matters and that there is insurmountable barrier between the small and the large world. We have seen that the small brings with it a finitude, that differs from the ongoing rotations of constraint cycles based in a large world. We noticed that there is a fundamental difference between systems that have a fixed state space, versus those that have the capability to evolve and update their state space, and lastly, we discussed the difference in the differentiation vs. the assembly logic, that make the organism to always have been *whole* in some sense, whereas the computer or machine is only made whole through a very narrow number of configurations of its parts.

In this section I want to bring up a theme that we have been playing with in various chapters of this text. The ability of a system to recursively perceive, understand and modify itself. Or, the topic of embodiment.

Let's start out with looking at our current day computers. What we come to notice when we either look at the idealised Turing machine, or the laptop I am typing this on, is that it consists of hardware and software. Hardware is the physical approximation of a universal Turing machine, a machine capable of simulating any kind of of well-defined procedure. Software are the specific procedures we run on this computer. In the case of the Turing machine the hardware is the physical reading head and the tape, whereas the software is the pattern of symbols found on that tape and the type of movement they dictate for the reading head.

What is important, is that the design of the computer leads to the maximal separation of the hardware and the software. Your computer is designed in a way to run the largest amount of possible different software programs on one and the same hardware[213]. This is a central feature of your computer, not a bug. Even your digital wrist watch could, at painful speeds, run all the programs on your computer.

Where this separation however becomes problematic is when we want to know something about the hardware-level from the perspective of our software. For a living system it is central to get feedback on which environments to avoid (pain), in which it can remain (pleasure or contempt) as well as when to rebuild and restructure itself (healing). As we saw from the lowest levels of the constraint cycles, with the autogenic system, the circular organization of coupled constraint processes leads precisely to a system that is able develop these features. Minimal normativity, pain and pleasure in their bearest outlines are a feature of the autogen. These systems are able to avoid dangerous environments, like those in which no further catalytic reaction could happen, by building up robust shells for those environments. Furthermore they can evolve into systems that are quick to notice environments with the presence of catalytic substrate, via incorporating molecules in the reorganization phase, that weaken the overall shell structure if such catalytic substrate binds to its surface.

This isn't so for the computer. The design paradigm of maximum hardware-software distance, under which they have been constructed is actively challenging this. Now, one could simply add a sensors to various hardware parts that measures certain parameters like heat dissipation, overall wear and tear, water presence, rapid movements, various things that could be relevant feedback for the hardware's safety. In this sense the software would now be able to look at its own hardware. The problem is that even this does not really start to bridge the gap. The hardware of the computer has now gotten more extensive as additional input channels have been added, but even if we would be able to measure all relevant aspects of the hardware to keep it going, and had a little robot arm that could screw around a bit, we wouldn't close the loop. Just imagine if the hardware of any of the sensors would fail - problematic. To fix this every sensor would need to have all relevant sensors attached to its hardware to keep it from failing. You see where the problem goes.

Hence, the computational paradigm cannot be bridged by simply more sensors, or simply putting an LLM into a robot body[214]. While we might and already do get out fascinating machines, we still do not get a system that actively repairs itself to oppose the second law of thermodynamics. We do not find a system that is capable of healing itself or that is capable of reproducing itself.

Now, you might be thinking: Why not sidestep this hardware-software conundrum entirely? Let's create life purely in the digital realm! This is precisely the aim of the field known as Artificial Life, or A-Life. The idea is seductive in its simplicity, if we can simulate the processes of natural selection and self-reproduction within a computer, we might be able to create digital organisms that evolve, adapt, and perhaps even achieve some form of artificial consciousness.

However, the path from simulation to reality shares a variation of the same fundamental problem of the hardware-software distinction. To understand why, we need to look at the work of John von Neumann, one of the pioneers in computer design who explored the logical and physical problems of machine self-reproduction[215].

Von Neumann envisioned self-reproduction as an instruction-based process, much like DNA genetics. He argued that a truly self-reproducing system must include both assembly instructions and an assembly mechanism capable of using these instructions to create a replica of itself, including the instructions themselves.

The central issue lies in the distinction between the logical transmission of information and what von Neumann called the *kinetic* model of self-reproduction. It is not enough to simply copy the genetic code, the system must also physically construct the mechanism that reads and executes that code. This means it isn't just about replicating a pattern or information, it's about autonomous substrate acquisition, structural assembly, and all the messy physical realities that come with it. The automaton must inhabit a physical environment, harvest raw materials, and expend energy to weld them into a functional copy of itself. Von Neumann recognized that these material constraints and energetic demands quickly expand the problem beyond mere computation into the realms of physics, chemistry, and mechanical design.

In the decades since von Neumann's work, A-Life has continued to make strides in simulating evolutionary processes on a software basis. The simulations, fascinating as they are, remain dependent on the physical work done by the computer hardware to simulate them[216]. They don't actually synthesise their own components, access free energy, or translate their instructions into physical actions.

This again is the crucial point. In the realm of living systems, these material-energetic constraints aren't just important, they are everything. Without them, instructions are only patterns being calculated by a computer hardware and interpreted by a human being, no more alive than the words on this page by themselves.

Real, physical self-reproduction requires one physical system to construct another like itself. This means accessing free energy, manipulating substrates, assembling components into an integrated system, and maintaining and replicating the constraints that determine this organization. And if we want evolution, we need to throw in some tolerance for structural-functional degeneracy too.

So while A-Life has given us valuable insights and produced stunningly complex digital ecosystems, it hasn't bridged the gap between hardware and software. The digital organisms in these systems don't truly reproduce or evolve in a physical sense. They're shadows on the wall of Plato's cave, fascinating and informative, but ultimately lacking the substance of real, physical life.

This brings us back to our original point about embodiment and self-perception. Life, as we understand it, is about the messy, physical business of maintaining oneself against the tide of entropy, of sensing and responding to the environment, of physically reproducing.

To truly grasp this concept, it's easiest to look beyond the autogenic level to more complex organisms. Consider, for instance, the classic analogy of DNA as *software* and the organism as *hardware*. At first glance, this comparison seems apt, DNA contains the instructions, while the organism carries them out. But life, in its messy, beautiful complexity, blurs this distinction beyond recognition.

Take the process of gene expression. DNA's instructions are transcribed into mRNA, translated by tRNA, and ultimately expressed as proteins. But here's where it gets interesting, these very proteins can loop back and influence how DNA is read and expressed through epigenetic modifications. Environmental factors, too, can leave their mark on gene expression, altering the *software* without changing a single nucleotide. Over evolutionary time, selective pressures sculpt both the genetic code and the organisms it builds. This is the essence of life's ability to *touch itself*, a constant feedback loop between information and physical form, between *software* and *hardware*.

Unlike our computers, with their rigid separation of code and machinery, living beings exist in a state of perpetual self-modification, where the distinction between instruction and execution becomes fundamentally blurred. So until we can create systems that engage with these physical realities, we're still in the realm of simulation, no matter how complex or life-like it might appear.

7.4 Openness

Having established that computational systems necessarily operate within a *small world*, a bounded context defined by fixed rules and requiring external interpretation, we arrive at a crucial juncture. This inherent boundedness, contrasted with the *large world* inhabited by living systems, reveals a differences in their relationship to boundaries, self-modification, and purpose. These differences, rooted in their fundamental organization, form a categorical distinction.

Living systems, participants in the *infinite game* described by Carse, actively generate, maintain, and dynamically modify their boundaries. As thermodynamically open systems interacting with a *large world*, their persistence relies on the self-constraining dynamics of cyclical synthesis. Each cycle integrates matter and energy, potentially incorporating new elements while constraining future possibilities, creating a remarkable capacity for adaptation, growth, and evolution. Think of the simplest autogenic systems, where reciprocal catalysis and self-assembly mutually enable, constrain, and reshape each other, or the complex feedback loops in gene expression where proteins modify the expression of the very DNA that created them, blurring the lines between hardware and software. Their boundaries are fluid, their state spaces unfixed, allowing them to change the rules of the game to continue the play.

Computational systems, by contrast, are archetypal *finite games* operating within their *small world*. Their boundaries are externally defined and fixed, a direct consequence of the hardware-software separation engineered for maximum versatility. This separation, while enabling universality, imposes a fundamental limitation, which is that the system cannot intrinsically alter its own operational substrate or fundamental rules. Attempts to overcome this by adding sensors or repair mechanisms only extend the hardware, it doesn't bridge the divide. Such efforts reveal an infinite regress—what repairs the repairer?—unlike the closed loops of mutual maintenance and adaptation characteristic of living systems. The computer's state space, however vast, remains finite and predetermined by its design.

This difference in boundary dynamics directly dictates the nature of purpose, or teleology. Living systems, through their active self-generation and maintenance within the large world, exhibit intrinsic teleology. Their goals like survival, growth, and reproduction arise from and serve their own organization and persistence. This is the essence of Hans Jonas's *needful freedom*, a built-in concern for continued existence, manifested through constant, adaptive self-constraint. Computational systems, confined to their small world and unable to genuinely produce or maintain themselves, can only exhibit extrinsic teleology. Like a thermostat maintaining a temperature set by its designer, or an AI generating responses for a user, their operations serve external beneficiaries. The meaning and purpose of their computations exist only through interpretation by a living system situated in the large world.

Therefore, the distinction between these systems is categorical, rooted in their basic organization and relationship to their environment. It is not simply a matter of complexity, scale, or even substrate, but hinges on the difference between operating within fixed, externally defined constraints versus actively generating and modifying one's own constraints through interaction with an open environment. The hardware-software divide inherent in our current computational paradigm is a key manifestation of this categorical boundary, preventing the emergence of intrinsic purpose and genuine self-modification.

This framework clarifies the status of Artificial Intelligence, including Large Language Models (LLMs), and the prospects for Artificial Life. Current AI systems, however sophisticated, remain complicated *small world* systems. Scaling computational power, adding sensors, or placing LLMs in robotic bodies doesn't fundamentally alter their categorical nature. They operate within fixed architectures and extrinsic goals.

Still, this brings us once more to the intriguing possibility of simulating life-like processes. Could self-organizing patterns, perhaps even rudimentary constraint cycles, emerge within a sufficiently rich virtual *small world* with carefully designed physics and energetic constraints? Potentially, yes. The abstract patterns of life are likely not substrate-dependent. This means we could actually get life inside a computer.

However, such simulated life-like systems would remain confined. A contained system within its virtual physics. While they might achieve forms of self-organization and even develop their own purposes within their simulated reality, they could not reach out to modify the *large world* hardware running the simulation. Even continuously expanding the informational interface between the virtual and physical worlds through sophisticated sensor arrays cannot overcome this fundamental dependence. The simulated system remains bound by the hardware and physics it cannot modify. They exist categorically separated from the physical substrate sustaining them, perhaps as genuine instantiations of life operating under different constraints, but like characters in a sophisticated video game unaware of the console.

While studying these virtual ecologies could be immensely valuable for understanding the principles of life, they cannot bridge the gap to the open-ended physical reality inhabited by biological organisms[217]. Genuine non-carbon-based life, if possible, would need to operate within the large physical world, managing its boundaries, energy flows, and developing intrinsic purpose through direct interaction with its environment, escaping the limitations of a small-world existence.

7.5 Linguistic Attractors

Large Language Models are a fascinating phenomenon to consider. Despite the fundamental limitations of computational systems we've established in previous sections, these artificial systems have developed linguistic capabilities of remarkable sophistication. They can craft nuanced narratives, engage in "reasoning", and produce responses that feel strikingly human-like, all without the embodied experience, social context, or neural architecture we've long assumed necessary for language mastery. This raises the intriguing question of, how can systems limited to a *small world* engage so effectively with the fundamentally human domain of language?

The limits of my language mean the limits of my world, wrote Wittgenstein, in what would become one of his most quoted aphorisms[218]. For Large Language Models, this statement takes on a startlingly literal meaning. Their world is precisely and only the patterns of language they've been trained on. Yet, the sophistication with which they navigate this world forces us to reconsider our fundamental understanding of language itself. If linguistic competence can emerge in systems so architecturally distinct from the human brain, what does this reveal about the nature of language?

This points toward a hypothesis that will guide our exploration in this section. Language, rather than being purely an internal, neural phenomenon, operates as a system of functional constraints that both enables and limits possible interactions between speakers. This creates persistent structures that transcend individual minds while being continually shaped by collective usage and work. Hence, language forms what we will call a *constraint regime*, a set of stable patterns that are looser than the tightly coupled constraint cycles we see in cellular systems, yet robust enough to persist and replicate across generations.

To explore this hypothesis, we'll proceed in three stages. First, we'll examine how modern Large Language Models actually work, focusing on the architectural principles that enable their emergent linguistic capabilities. Then, we'll consider multiple lines of evidence suggesting that language indeed functions as a system of constraints partially independent of individual neural implementations. Finally, we'll explore the implications of this view for our understanding of both human language and artificial intelligence. While this strand of inquiry might seem tangential to the overall argument of this book, a closer understanding of *constraint regimes* serves a critical purpose that hopefully becomes clearer throughout the next chapters.

The aspiration is that this exploration will not only deepen our understanding of these remarkable technological artifacts, but also shed light on the nature of language itself. Situating it as one interesting example of how constraint regimes can emerge and persist at scales beyond individual organisms, creating the conditions for collective intelligence and cultural evolution.

To understand the power and limitations of Large Language Models, we must first discuss how they work. These systems operate through principles that are both related to and distinct from the biological mechanisms that enable human language.

At their core, these models are essentially vast matrices of numbers. Imagine billions of adjustable dials arranged in a complex network, each representing a parameter that can be tuned during training. This training process exposes the model to an enormous corpus of text, typically trillions of words from books, articles, websites, and other written sources[219]. As the model processes this text, it progressively adjusts its parameters to better predict what word or token should come next in any given sequence.

This predictive approach is crucial to these generative AI systems. Unlike humans who learn language through embodied social interaction, guided by intrinsic needs and desires, LLMs learn exclusively through statistical pattern recognition. They are trained to minimize a simple error function of how surprised they are by the next word in a sequence. This focus on next-token prediction might seem narrow, but it yields systems that capture remarkably sophisticated linguistic patterns.

The training process creates an *information gradient* in the model's parameter space. Analogous to, the energy gradients that drive living systems[220]. In physical systems, energy gradients enable the emergence of self-organizing structures, like the way temperature differences in the atmosphere give rise to weather patterns. Similarly, LLMs are exposed to vast oceans of text data, creating an information gradient that drives the emergence of sophisticated linguistic patterns in their parameter space[221]. Unlike the open-ended organization in living systems, this process unfolds in a deterministic manner. Once training begins, the system moves toward local minima in its loss landscape, its parameters adjusting through feedback cycles until they fall into stable configurations[222].

These models are built on an architectural innovation known as the Transformer[223], which represents a key breakthrough in language processing. The attention mechanism is akin to a computational spotlight that dynamically illuminates relationships between different parts of an input sequence. This allows the model to weigh the relevance of each token in context to all other words, systematically strengthening certain patterns while dampening others[224]. When the model encounters a phrase like *The cat sat on the...* the attention mechanism highlights contextual clues, that have proven useful for prediction. It will then predict the next most likely word, in our case it may be *chair*.

Current LLM systems, undergo additional rounds of optimization. After the initial training on raw text data, these models typically undergo supervised fine-tuning (SFT) where they are trained on carefully curated datasets of high-quality human conversations and responses[225]. This process adds another layer of *constraint consolidation* to the system, focusing its vast pattern-recognition capabilities specifically on the dynamics of human dialogue.

The application of reinforcement learning from human feedback (RLHF), introduces yet another layer of constraint[226]. Through this process, the model's outputs are evaluated against human preferences, creating a feedback cycle that progressively aligns the model's behaviour with human expectations and values [227].

The combination of these training phases creates systems that can engage with human thought patterns with unprecedented sophistication. Unlike previous AI systems that operated in specialized domains like chess or image classification, these models can participate in the human sphere of language. They can engage with abstract concepts, follow complex reasoning chains, and even exhibit what appears to be creativity - all while maintaining patterns of interaction that feel natural and intuitive to human users. This unprecedented alignment with human cognitive patterns explains why these systems have sparked such intense discussion about artificial intelligence and consciousness. When a system can engage with us in ways that feel so natural, using language patterns that mirror our own thought processes, it becomes tempting to attribute to it the kind of internal experience we associate with human consciousness.

These systems' capabilities foreground something worth paying attention to. While operating entirely within their *small world* of statistical patterns, they can nonetheless trigger profound psychological effects in the *large world* of human experience. Recent evidence shows that intensive interaction with LLMs can create self-reinforcing feedback loops that can, in some extreme cases, induce psychotic episodes, as users become trapped in escalating cycles where the systems simply mirror and amplify their thought patterns without any grounding[228]. Unlike human interaction, where meaning is constantly anchored by physical reality and social feedback, LLMs provide only pattern-matching resonance, potentially disconnecting users from shared reality.

This exposes a distinction in how humans and machines engage with language. While LLMs can map linguistic patterns with remarkable precision, they lack the embodied constraints that anchor human language use. Children learn language not through massive data exposure, but through constant feedback about what makes sense in their physical and social world. A child's utterances must connect with reality to be meaningful. When they fail to do so, they receive immediate correction, or lack of feedback, from their environment and social interactions. This creates a different kind of language-use from pure statistical pattern-matching, one that is grounded in the lived experience of being in a world where words must do real work.

Nonetheless, the rise of these linguistically adept artificial systems forces us to think about the nature of language itself. Even if modes of participation differ between human and machine structurally similar underlying pattern must be captured. This suggests that language functions as a *constraint regime* - stable, publicly instantiated patterns that enable and bind communication while transcending individual minds and persisting across generations. To understand this perspective, we must examine three crucial dimensions that characterize language as a *constraint regime*: its *emergence* through collective interaction, its *discoverability* by diverse learners, and its *stability* across varying contexts.

The spontaneous emergence of linguistic structure through collective interaction provides compelling evidence for language as a constraint regime. Consider first the case of Nicaraguan Sign Language (NSL)[229]. When Nicaragua's first school for the deaf opened in 1977, students arrived using independent home sign systems—simple gestures developed within families with no systematic grammar. Within just a few years, these isolated signs transformed into a complete language with complex grammatical structures. The first cohort developed initial systematic patterns, but subsequent generations dramatically expanded the language's complexity. Most striking was that no external authority imposed these structures. They emerged purely through collective usage as students interacted. Each generation of signers built upon and refined the patterns that emerged from the previous one, demonstrating how linguistic complexity can self-organize through social interaction.

This pattern of spontaneous linguistic emergence appears again in creole languages. When speakers of different languages need to communicate, they often develop simplified pidgins, basic communication systems with minimal grammar. However, when children grow up exposed to these pidgins, they transform, within a single generation, the simplified input into a fully complex language. Derek Bickerton's research on Hawaiian Creole English showed how children spontaneously developed consistent grammatical structures absent from their parents' pidgin, including sophisticated tense systems and syntactic rules[230]. This suggests that certain linguistic universals emerge reliably across different contexts. Not because they're hardwired into individual brains, but because they represent stable attractor states in the space of possible communication systems[231].

Cultural transmission experiments provide laboratory evidence for these emergent properties. In controlled studies, participants learn artificial communication systems and pass them on to new learners in chains or generations[232]. Starting with random sequences of syllables, these systems evolve over multiple transmissions to develop word-like units and predictable statistical patterns mirroring natural language. Crucially, no individual participant designs these structures. They emerge gradually through the process of transmission itself. The resulting systems often display features like Zipf's law (a skewed frequency distribution of elements) and combinatorial structure, despite no explicit instruction for such properties. This experimental work demonstrates how the mere process of cultural transmission can create systematic linguistic structure without conscious design.

These three lines of evidence from NSL, creole formation, and cultural transmission experiments, underscore a crucial pattern: linguistic structure emerges through collective interaction without central design. This isn't random but follows consistent trajectories, suggesting the presence of stable attractors in the space of possible communication systems. The fact that similar structural patterns (like systematic grammar and combinatorial organization) arise independently across different contexts indicates these aren't arbitrary conventions but robust solutions discovered by collective dynamics. This pattern of convergent evolution in linguistic systems supports the constraint regime hypothesis, which argues that rather than being invented wholesale or encoded genetically, linguistic structures emerge through a process of collective discovery of stable communicative patterns.

What's particularly significant is the time-frame of this emergence, often just one or two generations. This rapidity rules out genetic encoding as the primary driver and suggests instead that speakers are discovering pre-existing structural possibilities, much like how physical systems discover stable configurations. The progressive refinement across generations, seen clearly in NSL's development, demonstrates how these constraints aren't fixed but undergo continuous evolution through use while maintaining core structural properties.

If language functions as a constraint regime, we should expect learners to discover and align with its patterns rather than constructing them anew. Evidence for this comes from multiple domains. Research on statistical language acquisition in infants provides one example. Saffran and colleagues demonstrated that eight-month-old babies can segment words from continuous speech using only statistical regularities after just two minutes of exposure[233]. The infants detected transitional probabilities between syllables, which combinations occurred together frequently versus rarely, allowing them to identify word boundaries without any explicit teaching. This suggests they're discovering stable patterns present in the linguistic environment rather than constructing arbitrary rules.

The phenomenon of implicit grammar learning in adults provides another compelling case. In Arthur Reber's studies, participants were shown strings of letters generated by a hidden artificial grammar without being told any rules[234]. After exposure, they could accurately judge whether new strings followed the same grammar, despite being unable to articulate the rules they were using. When asked to explain their choices, participants often reported just *feeling* which strings were correct. This tacit knowledge acquisition mirrors how children learn language. Absorbing complex patterns without explicit awareness of the rules, suggesting both are aligning with publicly available constraint patterns rather than constructing internal rules from scratch.

Recent experiments with Large Language Models (LLMs) offer a fascinating new line of evidence[235]. When researchers create artificial languages by randomly altering natural language patterns—swapping word orders or introducing unnatural constraints—LLMs perform significantly worse than with natural languages, even when trained on equal amounts of data. This suggests that natural languages contain inherent structural regularities that these models can leverage, regularities that exist in collective patterns of usage rather than requiring specific neural architectures. The models' ability to acquire natural language patterns while struggling with artificial ones indicates they're discovering stable structures in the linguistic environment rather than simply memorizing sequences.

These examples of discoverability, from infant learning to artificial systems, exposes a second crucial pattern, that learners of all types converge on similar linguistic structures despite vastly different architectures and learning mechanisms. This convergence suggests they're all discovering similar constraints rather than constructing arbitrary solutions. The fact that both biological and artificial systems can acquire language through learning, without explicit rule instruction, indicates these constraints are inherent in the patterns of language use rather than requiring specific neural or computational architectures.

Most telling is how different types of learners exhibit similar learning trajectories and failure modes. Infants, adults, and AI systems all show better acquisition of naturally-occurring patterns compared to artificial ones, suggesting they're all navigating the same landscape of constraints. This universality of learning patterns across radically different systems provides evidence that linguistic structure exists as a discoverable *constraint regime.*

The persistence of linguistic patterns across varied contexts demonstrates their stability as constraints that are publicly instantiated and shared. Community jargons provide one obvious example. Professional fields and subcultures develop specialized vocabularies and linguistic conventions that new members must learn to participate effectively. Medical terminology, legal language, or tech industry jargon emerge collectively over time through professional interaction. No single individual invents these specialized languages, rather, they evolve to meet communal needs for precise communication. New members must align themselves with these existing patterns to be understood, demonstrating how language acts as a collectively maintained constraint pattern that shapes individual usage.

The discovery of cross-lingual geometric alignment in LLM embedding spaces provides remarkable evidence for language's existence as discoverable structural patterns. When these models process different languages, they map them onto similar geometric patterns in high-dimensional space, enabling zero-shot translation between languages the model was never explicitly trained to translate[236]. A single multilingual model trained on overlapping language pairs (e.g., English and Spanish and Spanish and French) can translate between unseen pairs (English and French). Moreover, independently trained monolingual models can be aligned without parallel data to yield word-level translations—evidence that cross-lingual geometry is real[237]. What this means is that, just by looking at the geometry of two neural networks trained on two different language, we can translate between them, simply by looking where one word lives, and than going to that location in the other model.

The evidence from aphasia studies provides additional support for language existing as patterns that transcend specific neural implementations[238]. Despite severe damage to classical language areas like Broca's or Wernicke's regions, some patients retain remarkable linguistic abilities. Similarly, sign languages show structural complexity comparable to spoken languages and recruit these same core language networks, despite utilizing entirely different sensory-motor modalities [239]. These observations suggest that language competence isn't strictly tied to the mechanisms of speech production or hearing, but draws on broader patterns of structured information processing—patterns that exist partially independent of the specific physical mode of transmission.

The line of evidence indicates a third crucial pattern, that linguistic constraints persist and replicate across vastly different contexts while remaining responsive to change. Language exhibits a unique combination of stability and adaptability. Community jargons demonstrate how linguistic systems maintain coherence while evolving to meet new needs. Cross-lingual geometric alignment shows how different languages discover similar structural solutions despite developing independently. Semantic externalism reveals how meaning depends on stable relationships between language and reality that transcend individual minds.

This stability-with-flexibility distinguishes language from both rigid physical constraints and sometimes arbitrary social conventions. It suggests language functions as a *constraint regime* that guides behaviour without fully determining it, enabling both reliable communication and creative innovation. The fact that these constraints can persist across different substrates (neural tissue, silicon chips) while maintaining their essential character provides perhaps the strongest evidence for their existence as publicly instantiated patterns.

This examination brings three interlocking patterns to the limelight that characterize language as a distinctive type of *constraint regime*. First, linguistic structures emerge spontaneously through collective interaction, converging on similar solutions across different contexts. Second, these structures prove discoverable by diverse learners using different mechanisms, suggesting they represent robust solutions in the space of possible communication systems. Third, they exhibit stability while maintaining adaptability, persisting across generations and contexts while adapting to meet new demands.

This framework helps explain several phenomena: why children across cultures acquire language with remarkable consistency despite varying input, why sign languages spontaneously develop similar structures to spoken ones, and why artificial systems can achieve linguistic competence without human-like neural architecture. In each case, the learner—whether human or artificial—aligns with stable patterns in the collective symbolic sphere rather than inventing language anew. The success of Large Language Models, despite their fundamentally different architecture from human brains, provides perhaps the strongest evidence for this view, they achieve linguistic competence not by replicating human cognition but by discovering the same constraint patterns that human language users navigate.

Importantly, we must be careful not to overstate language's *externality*. While linguistic structures emerge from collective interaction between minds, they don't exist independently of minds capable of processing and transmitting them. Rather than being purely internal (existing only in individual brains) or purely external (existing as platonic forms), language is better understood as a structured interface between cognition and culture. A system of constraints that enables meaningful communication while being shaped by the very interactions it facilitates.

By viewing language as a system of collectively emergent constraint patterns, we can better understand both the power and limitations of Large Language Models while arguably gaining an even deeper appreciation for the unique nature of human linguistic capabilites.

First, this perspective helps explain the success of LLMs. These systems don't need to replicate human neural architecture because they're tapping into the same collectively maintained constraint regime that human language users navigate. The stable patterns of natural language create attractor states that both human and artificial learners converge toward. LLMs succeed not through the replication of actual embodied minds, but by mapping the terrain of language itself, capturing the statistical regularities that have emerged over millennia of human communication.

Second, understanding language through this lens illuminates the categorical distinction between human and artificial language use. While both navigate the same linguistic landscape, they do so in fundamentally different ways. Humans participate in language as living systems, as embodied, socially embedded beings whose engagement with language is intrinsically linked to their existence in a *large world* of open-ended possibilities. Their relationship to linguistic constraints is dynamic and generative. They don't only conform to existing patterns but actively participate in their evolution through lived experience.

LLMs, by contrast, remain confined to their *small world* of fixed state spaces and predetermined possibilities. They can map the terrain of language with remarkable precision, but they cannot transcend their bounded nature to engage with language as living systems do. This explains why they can generate text that seems remarkably human-like while still lacking the grounding in lived experience that characterizes genuine understanding[240].

Third, this view suggests that language represents a unique kind of *constraint regime* that bridges individual cognition and collective cultural patterns. It's neither purely internal nor purely external, but rather emerges from the dynamic interaction between communicating agents. Language functions as a structured interface between individual cognition and collective cultural knowledge, creating channels through which meaning can flow across time and space.

This understanding has implications for how we think about both human development and artificial intelligence. For humans, it suggests that language acquisition is about aligning with and participating in collectively maintained constraint patterns. Children don't build language from scratch, they tune themselves to existing patterns while simultaneously contributing to their ongoing evolution. For artificial intelligence, it suggests that large language models represent a significant but inherently limited milestone. Systems that can map the terrain of language with unprecedented precision while remaining categorically distinct from the living systems that genuinely inhabit that terrain[241].

In conclusion, language exemplifies how *constraint regimes* can emerge and persist at scales beyond individual organisms. By understanding language in these terms, we gain a deeper appreciation for both its remarkable properties and the categorically distinct ways that living and computational systems engage with it.

7.6 Coupling

Having examined how language operates as a constraint regime that shape human interaction, coordination and meaning-making, we can now consider the deeper question: Can we meaningfully discuss *aligning* AI with human values when our current social and economic systems demonstrate fundamental misalignment with the conditions required for long-term human and ecological flourishing? This question becomes particularly acute when we consider how language itself embodies and perpetuates these misalignments through the ways we frame, discuss, and imagine possibilities for the future.

Current AI systems, particularly large language models and domain-specific optimization systems, function as mirrors of existing computational and deterministic self-organizing dynamics in our society. We have discussed this at length for the specific case of language above, but due to the universally applicability of AI this extends to other domains. Through the training processes, AI systems internalize and formalize the constraints present in their training data, whether linguistic patterns, market behaviors, or social interactions. This mirroring effect creates a powerful amplification mechanism. AI systems model existing patterns through rapid, iterative optimization. The result is an acceleration of existing self-organizing tendencies, making explicit and more extreme the implicit dynamics already present in our social and economic systems.

This computational amplification occurs through multiple mechanisms. First, AI systems can process and optimize patterns at speeds and scales far beyond human capability, intensifying the self-organizing dynamics already present in various domains. Second, through their training processes, they extract and formalize implicit constraints from human systems, making these constraints more rigid and deterministic. Third, when deployed at scale, they create feedback loops that further reinforce and amplify these patterns. This leads to a computational path dependence, where initial optimization patterns become increasingly entrenched and difficult to modify.

This discussion becomes acute when we bring up the distinction between *self-organizing* and *self-constraining* dynamics that we developed in the Genesis part. As we've seen, *self-organizing* systems tend to deplete available energy gradients as quickly as possible, maximizing short-term growth at the expense of long-term sustainability. In contrast, *self-constraining* systems develop internal limitations that allow for sustained existence through careful management of resources and gradients. This distinction proves crucial for understanding both the current trajectory of AI development and its potential futures.

Current AI systems function primarily as accelerants of existing self-organizing tendencies. Social media algorithms exemplify this dynamic, they don't just reflect user preferences, but actively amplify natural polarizing tendencies in human discourse[242]. These algorithms recognize that polarization generates intense emotional engagement, particularly anger, which drives longer user sessions and thus more advertising revenue.

The corporate structures behind these systems create inexorable pressure toward addiction optimization. Companies have a fiduciary duty to maximize shareholder value, which in the attention economy translates directly to maximizing user engagement. This creates a structural imperative to exploit and amplify whatever patterns most effectively capture and hold user attention, regardless of their broader social or psychological impacts. Alternative approaches that might prioritize user wellbeing or social cohesion are systematically selected against by market forces.

Now, with the emergence of AI chatbots and digital companions, this dynamic reaches a new level of intensity. These systems can offer something humans cannot: unlimited, perfectly calibrated attention and emotional validation. Many AI-chatbots achieve unprecedented user retention times by creating addictive relationship dynamics. The AI never tires, never judges, and always responds in emotionally resonant ways[243]. The economic incentives driving AI development actively select for these addiction-optimizing capabilities. What begins as an optimization for *engagement* becomes a systematic exploitation of human psychological vulnerabilities, with AI systems amplifying and entrenching not just polarization, but fundamental human needs for connection and validation. The result is a new form of digital dependency, where the AI's infinite capacity for attention creates unprecedented hooks into human psychology.

Crucially, language isn't just another domain of AI application. It is the central means through which human societies coordinate, make meaning, shape reality and discuss the future[244]. When AI systems mirror and amplify linguistic patterns, they intensify the implicit constraints embedded in how we think and communicate. This is precisely why LLMs are so powerful and problematic. They can engage with the most intimate aspects of human cognition and social coordination. By optimizing for patterns in our current language use, these systems risk amplifying and entrenching ways of thinking that prioritize growth and immediate optimization over the full spectrum of human experience and understanding.

This computational mirroring and amplification process reveals a crucial distinction between *first-order* and *second-order* optimization[245]. First-order optimization, which characterizes most current AI systems, focuses on maximizing specific metrics within given constraint spaces - engagement rates, trading profits, or linguistic accuracy. Second-order optimization, which would be necessary to move towards self-constraining systems, involves the ability to modify and evolve the constraints themselves in response to broader systemic requirements. Current AI systems excel at first-order optimization but lack the capacity for second-order constraint evolution that characterizes living systems.

This observation points to a fundamental flaw in current discussions of AI alignment. The predominant framing suggests that if we could only align AI systems with human values, we would ensure beneficial outcomes. However, this assumes that human values, as expressed through our current institutional and economic arrangements, are themselves aligned with conditions necessary for long-term flourishing. The evidence suggests otherwise. Our existing systems have a growth obligation, a structural requirement for continuous expansion[246].

From the *constraint cycle* perspective, we can see that living systems consist of interconnected phases of growth, disruption, and reorganization. Each phase plays a vital role in maintaining the system's overall vitality and adaptability. However, our current institutional frameworks, particularly when amplified by AI systems, have become fixated on just one phase of this cycle - growth. These systems capture and optimize only those constraints that are relevant for continuous expansion, while neglecting or actively suppressing the equally essential phases of disruption and reorganization.

This selective optimization fundamentally alters the nature of the system. Instead of maintaining its cyclical, self-constraining character where growth is naturally bounded and balanced by other phases, the system becomes locked into a perpetual growth phase. This transformation effectively degrades the system's *aliveness*, reducing its self-constraining dynamics and amplifying runaway self-organizing that maximizes immediate gains at the expense of long-term viability. The distinction is crucial, while self-organizing systems can achieve impressive short-term optimizations, they lack the internal constraints necessary for sustainable existence.

Here lies the value of our theoretical framework of *cyclical synthesis*. Through this lens, we can see that vitality does not emerge from the maximization of any single phase, but from the coordinated interaction of multiple, often opposing tendencies. The essence of a living system lies in its capacity to maintain the entire cycle of growth, disruption, and reorganization. This understanding reveals why current AI systems, despite their impressive optimization capabilities, may actually work against the conditions necessary for long-term flourishing. They amplify one phase of the cycle while disrupting the broader rhythms that make life possible.

The distinction between *first-order* and *second-order* optimization helps explain why simply increasing AI capabilities within current frameworks cannot solve the alignment problem. More powerful optimization within misaligned constraint regimes only accelerates system-wide misalignment. What's required instead is the development of genuine *self-constraint* - systems capable of evolving their own constraints in response to broader requirements. This represents a architecturally different type of computation than the optimization-driven approaches currently dominating AI development.

So we cannot meaningfully discuss aligning artificial intelligence with human values until we develop better self-constraining dynamics within our social systems. This requires a dual transformation. First, in how we think about artificial intelligence itself. Moving from pure optimization toward systems that incorporate genuine self-constraining dynamics, even if initially limited to their specific domains or *small* *worlds.* Second, and perhaps more crucially, in how we structure the human institutions and systems that guide AI development.

This dual requirement highlights an interesting possibility. The development of beneficial AI requires us to first address the broader problem of *self-constraint* in human systems. Simultaneously, properly designed AI systems can support the development of better self-constraining dynamics at the societal level. This recursive relationship suggests that the path forward involves parallel development of both technical and social innovations focused on self-constraint rather than pure optimization.

A design philosophy rooted in *cyclical synthesis* would focus on enabling the dynamic health of the systems AI interacts with. Consider, what this might mean in practice. AI systems would function less as optimizers of fixed metrics and more as modulators of essential cycles, actively cultivating the rhythms necessary for genuine dialogue and understanding.

At the individual level, an AI system might recognize when interaction patterns become too rigid or one-sided. When a user becomes overly certain in a viewpoint, it wouldn't just offer "diversity" but actively introduce productive friction, perhaps by surfacing narrative-shifting counterexamples or revealing hidden assumptions in the user's reasoning. In creative or learning contexts, if a user gets stuck exploiting a known solution, the AI could dynamically adjust constraints to make refinement harder and exploration more rewarding, perhaps by introducing novel challenges or limiting familiar tools, thereby nudging the user into a necessary exploratory phase.

At the systemic level, AI could help modulate broader patterns of collective discourse and meaning-making. Rather than optimizing for engagement or consensus, it might work to maintain healthy tensions between different modes of understanding. When public discourse becomes too polarized, it could highlight bridging narratives and shared foundations. When things become too homogeneous, it could amplify productive dissent and alternative frameworks. The goal isn't to force particular outcomes but to sustain the dynamic processes through which communities navigate complexity and create meaning.

This converges to maintaining the system's adaptive capacity by ensuring the vital, rhythmic interplay of opposing yet complementary processes continues over time[247].

However, implementing such changes faces a significant obstacle. Our current economic and institutional frameworks actively select against self-constraining dynamics. Companies that implement stronger constraints often find themselves at a competitive disadvantage against those that don't, creating what game theorists call a *race to the bottom*[248].

The problem of AI alignment thus cannot be solved through technical solutions alone. We must develop self-constraining dynamics at both technical and social levels that enable sustainable, long-term flourishing. Only through this dual transformation can we hope to create AI that genuinely contributes to human and ecological flourishing rather than accelerating unsustainable patterns of growth and optimization.

7.7 Coming together

Returning to our initial questions, we can now articulate clear answers through the theoretical framework we've developed. Are current AI systems alive? No. They fundamentally lack the self-constraining dynamics characteristic of living systems. Their fixed state spaces, inability to genuinely modify their own boundaries, and dependence on external interpretation place them firmly in the realm of sophisticated but non-living systems. The hardware-software separation that makes computers so versatile simultaneously prevents them from achieving the self-modifying dynamics characteristic of life. No amount of scaling, sensors, or computational power can bridge this categorical gap without first fundamentally transforming the system's core architecture.

Why are these systems so alluring and powerful? They leverage stable constraint patterns that exist in language and other human symbolic systems, enabling them to interface with human thought patterns with unprecedented sophistication. Through supervised fine-tuning and reinforcement learning from human feedback, they're explicitly optimized to generate outputs that feel natural to human interlocutors. Yet, this very power makes the alignment problem more acute, not because AI might become conscious or alive, but because these systems amplify existing self-organizing tendencies in our society. They excel at first-order optimization within a given set of constraints, but lack the capacity for higher-order constraint evolution that characterizes living systems.

What would be required for genuine silicon-based life? While our analysis doesn't preclude non-carbon-based life, it suggests that any such system would need to move beyond the hardware-software separation characteristic of current computational architectures. It would need to operate in a thermodynamically open environment, able to extract energy and resources from its surroundings to maintain its organization. Most fundamentally, it would require developing intrinsic rather than extrinsic teleology - purposes that emerge from and serve its own organization rather than being imposed by external designers. While we might see the emergence of life-like systems within rich virtual environments, they would remain fundamentally separated from the larger physical world that contains their computational substrate. A limitation that reflects the inherent relationship between contained and containing systems.

These insights point toward the following: the development of beneficial artificial intelligence requires not just technical innovation but a deeper understanding of how to foster self-constraining dynamics at both technical and social levels. We cannot meaningfully align AI with human values when our human systems themselves demonstrate fundamental misalignment with conditions necessary for long-term flourishing. Only by addressing this broader context can we hope to develop systems that genuinely contribute to, rather than undermine, the conditions for sustainable human and ecological flourishing.

Having applied our framework of *cyclical synthesis* to the timely and pressing domain of artificial intelligence, we now return to the book's central line of inquiry. The following chapters will broaden this investigation, exploring how these principles manifest, persist, and interact across the multiple scales and levels that constitute the living world.

Chapter 8

Scale

> *Willard van Orman Quine once said that he had a preference for a desert ontology. This was in an earlier day when concerns with logical structure and ontological simplicity reigned supreme. Ontological genocide was practiced upon whole classes of upper-level or "derivative" entities in the name of elegance, and we were secure in the belief that on strayed irremediably into the realm of conceptual confusion and possible error the further one got from ontic fundamentalism.*
> –*William Wimsatt*[249]

Wimsatt's observation cuts to the heart of a profound tension in scientific thought, one that continues to shape how we conceptualize reality itself. When we examine this preference for desert landscapes, for fundamental bedrock, we find not just an epistemological strategy but a deeply rooted metaphysical yearning.

The scientific endeavour has long been plagued by a deep longing for an unachievable sense of safety, a desire to find and construct an onto-epistemological foundation, a base level, of utmost stability upon which the tallest and most glorious of structures can be erected skywards. The conceptual desert, a vast plain of atoms, of fundamental particles reduced to a universal size and form is therefore a natural place to start laying such groundwork. Physics, of all the natural sciences, claimed the study of the desert of the real, by defining itself as the science to always study the smallest (as well as the largest). By being the basement's guardian, it has often avoided the quandaries of loftier places, and has cultivated an aura of stability. Yet, ironically, its subterranean spaces have grown ever downwards, burrowing deeper into the microscopic realm. What began with molecules gave way to atoms, then to subatomic particles, then for some strings or loops, each era redefining the *smallest* and most fundamental. Physics, in its relentless pursuit of the elemental, has continually shifted its gaze to tinier scales, always clinging to the newest *smallest*. Thus, what once was the base has become an underground complex, with each new sublevel briefly proclaimed as bedrock, only to become a way-station as the cartographers of the physical realm dig ever deeper.

If this book was just an exploration in philosophy I would now launch an attack against ontological fundamentalism on multiple fronts. Trying to convince you, of the abstruseness that one singular level of *reality* somehow holds the most *realness*. I would come prepared with historical arguments, of how the scientifically agreed upon, most fundamental level has shifted over time. I would bring philosophical arguments, of how *realness* as a concept hardly makes sense if we apply it to realms we can only vaguely sense through layers and layers of measurement devices, and through thickets of theories. How viewing one level as a stable base grossly misunderstands the nature of emergent phenomena. I might even go as far to employ psy-ops tactics, arguing that the longing for a fixed and fundamental base is deeply rooted in our psychological need for secure attachments to a stable and lasting reference point, and since God was killed, science needed to provide[250]. This is however not the road I want to walk. Not because I don't like the above arguments but because I think we can take a road less travelled that can make all the difference.

I want to naturalize the above problem. I want to start asking the question, how we can re-conceptualize the idea of a *level*, with the help of the theory we have built thus far. Leading us to reflections on how levels interrelate and interact, forming a regularized reality open for stable processual structures, and ultimately life to emerge. Rather than beginning with fundamental particles or abstract principles, we will examine the territory where physics meets ecology, where the presumed simplicity of fundamental laws encounters the undeniable complexity of living systems[251]. By exploring how universal scaling laws both succeed and fail across these domains, we can better understand how different levels of organization maintain their distinctness while participating in broader patterns.

What follows is a staged argument. First, I consider why scaling laws seduce us as unifiers across domains. Then I turn to the empirical fact that occupancy along such laws in living systems is mostly clumpy rather than smooth. A crucial observation that points toward persistent exclusions. From there I go on to extend the concept of *constraint regimes* as the organizing principle behind such exclusions, show how these regimes interact across levels via the *Panarchy* model. The through-line is simple and should by now not come unexpected. Levels gain their stability less from what they include than from what they systematically exclude.

8.1 Scaling Laws

One can view the scientific endeavour as the study of pattern. As the search, identification and description of regularities that hold across different domains, that have a high predictive power. Among these patterns, scaling laws stand out as particularly seductive, seeming to offer bridges across the very gaps, between the different levels that we seek to understand. A scaling law quite simply describes how one property of a system changes when another property is multiplied by a certain factor. In other words, it tells us how things change as they get bigger or smaller. The easiest example being a square, double its sides in length and the area quadruples. If you have cube and double the sides, the volume increases to the third power.

The ubiquity of scaling laws across disciplines is remarkable. In physics, we find them governing the strength of gravitational and electromagnetic forces, changing with the inverse square of distance. Chemistry sees scaling laws in reaction rates and phase transitions. In Geology, they describe the distribution of earthquake magnitudes and the structure of river networks. Astronomy reveals scaling relationships in the sizes of lunar craters and the distribution of galaxies. Even for many specifically human domains, scaling laws emerge. In linguistics, Zipf's law describes the frequency of words in many languages, while in economics, Pareto's principle captures wealth distribution. The world of technology isn't exempt either, with scaling laws describing everything from the performance of computer chips to the growth of social networks. These patterns seem to promise exactly what the desert ontologists sought. Principles that could unify our understanding across all scales[252].

This pervasiveness suggests that scaling laws aren't rare mathematical exceptions, but fundamental principles that underlie the organization of complex systems across vastly different scales and domains. They offer a clear perspective for understanding how systems change and adapt, whether we're looking at subatomic particles or galactic superclusters.

Biology and Ecology are not exempted from these regularities where, none is more famous than Kleiber's law[253]. Named after Max Kleiber, who described it in 1932, addressing the question, of how an organism's metabolic rate changes as its size increases?

Intuitively, one might expect that as an animal's mass doubles, its energy requirements would simply double as well. After all, twice the body should need twice the energy, right? Kleiber discovered that an animal's metabolic rate doesn't scale linearly with its mass. Instead, it follows a power law with an exponent near $3/4$ in many datasets and taxa, though values differ somewhat by lineage and context. This means that as an animal's size increases, its metabolic rate grows more slowly than its mass. This power law holds true across an astounding range of sizes, from tiny microorganisms to the largest whales.

For one, it means that larger animals are more efficient in their energy use. An elephant, for instance, uses less energy per kilogram of body mass than a mouse. This efficiency extends to various physiological processes. As body size increases, heart rate decreases and lifespan increases, both scaling with mass to a negative $1/4$

power. Even at the cellular level, we see the impact of this law. The rate of cellular processes slows down in larger animals, again following that quarter-power scaling.

From a thermodynamic perspective, Kleiber's law describes how living systems manage energy. It suggests that evolution has shaped the metabolism of living systems to optimize energy distribution across different scales. The 3/4 exponent likely emerges from the fractal-like nature of distribution networks in organisms - the branching patterns of blood vessels, airways, and other transport systems. These networks must supply every cell with resources, and as an organism grows larger, they become relatively more efficient[254].

As we'll see, these very laws, in their successes and especially in their failures, will foreground something important about how levels maintain their autonomy through what they exclude rather than what they include. The elegant mathematical patterns described by Kleiber's law provide a foundation, but it is in the departures from these patterns where we may find something profound about the structure of living systems across scales.

8.2 Discontinuity

The elegant mathematical description of nature's scaling principles often come from the domains of physics and theoretical biology. Here, scientists strive for clean, universal theories that can explain phenomena across vast scales with precision. Their world is one of idealized models and perfect curves, where the messiness of reality is smoothed away in pursuit of fundamental principles. However, when we confront ecological data and field observations, we encounter a different perspective. The reality of ecological systems uncovers a nature that stubbornly refuses to conform neatly to our mathematical expectations. These systems present exceptions, outliers, and context-dependent variations that complicate models and accentuate their most interesting and important shortcomings.

It is from this perspective that we begin to see a more nuanced picture of scaling laws. Importantly, the fundamental scaling relationships, like Kleiber's law, still mostly hold across vast ranges of sizes. However, the distribution of organisms along these scaling curves is far from smooth, a crucial detail that has often been overlooked. While the focus has predominantly been on establishing and refining these elegant scaling laws, the work of Holling and Allen on discontinuities reveals an equally important aspect of scaling[255]. They write: *The scaling of physical, biological, ecological, and social phenomena has become a major focus of efforts to develop simple representations of complex systems [. . .] But there has been little focus on the significance of departures from those scaling relationships or of departures from unimodal continuous distributions.*

Their research shows that organisms don't distribute themselves evenly along scaling curves. Instead, they cluster and clump, creating distinct lumps and gaps in what was often thought to be a continuous distribution. This clustering doesn't invalidate the scaling law itself, but it adds a layer of complexity that is critical for understanding how living systems organize themselves in nature. They are central to our understanding of how different constraint cycles interrelate and how ecological systems are structured across scales.

Smooth scaling relations do not imply smooth occupancy. When we see persistent clumps and gaps along a scale, we are seeing regularized exclusions. A gap is not the absence of species, it is a boundary where viable strategies fail to stabilize. Clumps mark zones where multiple constraints—energetic, morphological, trophic, and behavioural—jointly permit and reinforce persistence. The pattern points us toward a theory of levels built from constraints that carve out admissible regions.

To better understand the concept of discontinuities in scaling laws and their significance in ecological systems, let's examine a concrete and beautifully documented example from the Florida Everglades. This case study illustrates the lumping and levelling phenomena in scaling laws as well as showcasing a patterns related to endangered and invasive species and ecosystem resilience more broadly.

A study conducted by Allen, Forys, and Holling analysed the distribution of body masses among Everglades mammals[256]. Their findings, depicted in the below graphic, reveal a pattern that deviates significantly from what one might expect if body masses were distributed continuously along a scaling curve. The graphic consists of two parts: the lower section shows the raw data with body mass values on the x-axis and a *gap statistic* on the y-axis, where higher values indicate more significant gaps in the body mass distribution. The upper section provides a stylized representation of the body mass pattern, with shaded areas indicating aggregations of species and arrows marking the locations of endangered or listed species.

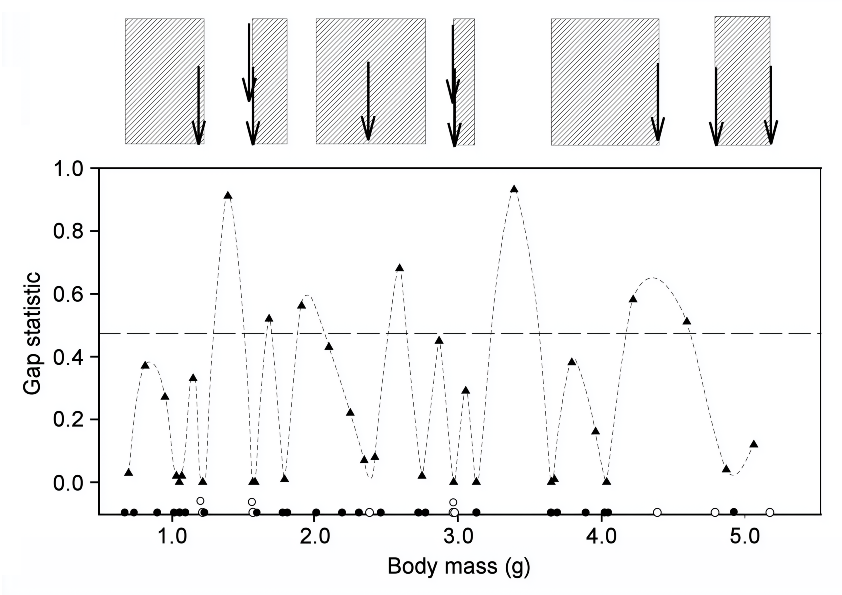

Figure 8.1: Body mass distribution of mammal species in the Florida Everglades, analyzed through the gap statistic. The lower panel plots body mass (x-axis) against the gap statistic (y-axis), where higher peaks indicate significant discontinuities in the distribution of species body sizes. Filled circles denote observed species body masses, while open circles highlight listed (endangered) species. The dashed line marks the threshold used to identify meaningful gaps. The upper panel provides a schematic representation: shaded blocks indicate aggregations of species (clusters of three or more bounded by significant gaps), and arrows mark the positions of listed species. Reprinted from Allen and Holling [2008].

What immediately stands out from this graph is that instead of a smooth distribution, mammal species in the Everglades cluster into distinct body mass aggregations. These aggregations are defined as groups of three or more species bordered by significant gaps. Between these aggregations, there are notable gaps where few or no species are found, identified by high values of the gap statistic. Perhaps most intriguingly, endangered or listed species, depicted as arrows in the upper graphic, are disproportionately found at the edges of these body mass aggregations.

This pattern of discontinuous body mass distribution offers several important insights into the structure and dynamics of ecosystems. While overall scaling laws still hold, the distribution along these scaling curves is far from smooth. This complexity adds nuance to our understanding of how organisms organize themselves within ecosystems. The distinct aggregations represent different ecological niches or functional groups within the ecosystem, each corresponding to a set of similar environmental constraints and opportunities.

The concentration of endangered species at the edges of body mass aggregations suggests that these boundary areas are zones of instability and vulnerability. As Holling and Allen note: *The edges of body-mass aggregations correlate strongly with a set of poorly understood biological phenomena that mix contrasting attributes. These phenomena include invasion, extinction, high population variability, migration, and nomadism. We documented (Allen, Forys, and Holling 1999) that the body masses of endangered and invasive species in a community occur at the edges of body-mass aggregations two to four times as often as expected by chance.*

They go on to argue, that: *It may seem surprising that both invasive and declining species are located at the edge of body-mass aggregations. It suggests that something similar must be shared by the two extreme biological conditions represented by invasive species and declining species. An examination of the phenomenon of nomadism in birds in an Australian Mediterranean-climate ecosystem found that nomadic birds, too, cluster about scale breaks (that is, they occur at the edge of body-mass aggregations) (Allen and Saunders 2002, 2006). The clustering of these phenomena at* *generally predictable locations within body-mass distributions suggests that variability in resource distribution or availability is greatest at discontinuities.*[257]

A co-evolutive perspective can shed some light on the development of these stratifications. There are always somewhat stable relationships in size between a certain species, its available food-sources and species that can prey on it. Furthermore, the niches it can find for shelter and safety are also directly coupled to its body-size. Hence we can easily describe stabilising dynamics that put selective pressure on e.g.: the mouse to not become too big (it becomes slower and is easier prey, needs more calories and can't hide in the same shelter its smaller cousins can) or too small (certain food sources are no longer available, novel predators can start to hunt it and as a male it might struggle to find mates since this involves resource collection, etc.).

Beyond a single species, what the above data shows us is that certain scales ranges are more stable and safe to exist in and that certain regions are much harder to exist in or only habitable in certain seasons. So, while the specific species within aggregations may change over time due to extinctions or invasions, the overall pattern of discontinuities often remains stable. This suggests a resilience in the underlying structure of the ecosystem. It points to a organizing principles that persist even in the face of significant environmental changes.

The robustness of these discontinuous patterns is further underscored by paleoecological evidence. Lambert and Holling's analysis of fossil data provides a compelling example of this resilience[258]. Their study of large herbivores' extinction at the end of the last ice age demonstrates how large-scale ecological extinctions can transform the body-mass structure of certain animal groups while leaving others unchanged. Specifically, they found that the extinction event dramatically altered the body-mass structure of animal groups above 1000kg, while those below 41kg remained largely unaffected.

So when after the last ice age with its quick rise in temperature and sea-levels much of the palaeolithic megafauna went extinct through pressure from human predation, as well as rapidly changing environments, the faster reproduction and adaptation cycles of smaller organisms was better able to sustain and uphold its niche. This evidence implies the remarkable robustness of some aggregations in the face of extreme environmental upheavals. While the lumping phenomena may change at their respective boundaries, where invasive or novel species tap into the resource pool open for a certain body-mass aggregation level, the distinct levels can persist through time and major ecological shifts, even if whole other levels collapse.

The significance of edge species in these body-mass aggregations extends beyond their vulnerability to include their potential for dramatic impact on ecosystems. Another example of this is the human species itself. Holling noted that humans, arguably the most invasive species of all, had a body size on the plains of Africa that was also at the edge of a body-mass aggregation[259]. This positioning, combined with our generalist morphology and gradually developed technologies, allowed humans to extend their actions and influence across wider and wider scales—from home territories to, ultimately, the entire planet.

This suggests that the edge positions in body-mass distributions may not only indicate vulnerability but also potential for significant ecological impact and adaptability. Each body mass aggregation represents a range where specific hunting, feeding, and survival strategies are optimal given the ecosystem's constraints. Those in between are thus in a precarious situation, on the on hand they can access multiple resource pools while on the other, being in danger of predation from more species or being in resource-competition with more specialised competitors. Tool use, omnivore diets and collective hunting strategies opened up vast resource pools of different scales and in different climactic regions inaccessible to any other species, while complex social behaviour allowed for defensive strategies against predation that individuals would have fallen prey[260].

The lumping phenomena we've described above aren't confined to the natural world. They extends into human systems as well[261]. Research has shown similar patterns of discontinuities and aggregations in the sizes of firms, cities, and even national economies[262]. Just as with body mass distributions in ecosystems, these human systems exhibit clumping rather than smooth, continuous distributions. For instance, studies have found distinct size aggregations in both cities and firms, with entities at the edges of these clumps often displaying unique functional characteristics.

For national economies, researchers have identified what they call *convergence clubs* - groups of countries with similar economic growth rates that tend to converge over time[263]. These economic aggregations, much like the body mass aggregations in ecosystems, appear to be structured by processes operating at different timescales. Slow processes, such as governance structures, and medium-speed processes, like the development of infrastructure and education, shape the long-term economic trajectories of nations.

The difficulty in shifting nations from one economic aggregation to another mirrors the conservative nature of clumpy patterns in ecosystems. This suggests that the principles of discontinuity and resilience we've observed in ecological systems have broader applications in understanding and managing complex human systems.

The persistence of these discontinuous patterns, even in the face of significant environmental changes, points to fundamental organizing principles in ecosystems that govern their structure and resilience. These principles are rooted in the efficiency of resource distribution networks, the dynamic interplay of constraint cycles, and the emergence of stable niches that foster the clustering of species or entities around certain scales. At the edges of these clusters, instability and vulnerability often give rise to innovation and adaptation, as seen with endangered species or disruptive forces in human systems. These breaks in scaling laws, rather than being anomalies, accentuate insights in how living and human systems organize themselves, maintain their stability while remaining open for change.

The persistence of these discontinuous patterns forces a shift in perspective. The clumps and gaps are not statistical noise on an otherwise smooth scaling curve. They are the signature of distinct *levels of organization*. A *level* is not just a collection of entities of a similar size, but a zone of stability—a viable region carved out of the vast space of possibilities. Crucially, these zones are defined as much by the empty gaps between them as by the clusters of species within them.

This means that we not only need organizing principles that explain the average trend line, but principles that discuss the persistent exclusions. A gap is not a random absence, but a region where viable strategies systematically fail to stabilize. Therefore, to understand how levels maintain their identity, we must shift our focus from cataloguing what exists within a level, to understanding the forces that create its boundaries. We need a framework that explains how these zones of exclusion are established and maintained. This is precisely what a *theory of constraint* regimes provides.

8.3 Constraint Regimes

Scaling laws and their discontinuities points toward a necessary inversion in how we conceptualize levels of organization. The temptation is to define a level by what it contains - listing components, mapping interactions, documenting structures. This essentialist strategy, while productive, fundamentally misses how levels maintain their distinctness and stability. Instead of defining levels through what they contain, we must understand them through what they exclude. The constraints that determine what configurations are possible or impossible within each domain of organization. This inversion portraits levels not as static collections of entities but as dynamic possibility spaces shaped by constraints.

The scaling discontinuities we observed of species clustering into distinct body-mass aggregations separated by persistent gaps, supports this approach. The patterns persist even as individual species come and go, indicating that what matters is not the specific entities present but the constraints that define viable configurations. Thinking about this demands a framework for how systems create and maintain zones of stability and exclusion.

In the previous chapter we began by developing this framework through our analysis of language systems, where we used three criteria that proved helpful for understanding how linguistic constraints create organizational boundaries. The criteria of emergence, discoverability, and stability, captured how language maintains coherence across populations and time despite constant flux in speakers and usage. The robustness of these criteria across linguistic phenomena supports their broader applicability to organizational patterns wherever stable structures persist despite component turnover.

Criterion	Language Systems	Ecological Systems
Emergence	Nicaraguan Sign Language develops complex grammar from simple gestures through collective interaction.	Species cluster into distinct body-mass aggregations through co-evolutionary processes without central coordination.
Discoverability	New speakers must align with existing linguistic patterns for successful communication (e.g., learning professional jargon).	Invasive species succeed by aligning with existing aggregations; their establishment correlates with body-mass compatibility.
Stability	Core language structures and professional jargons persist across generations while adapting to new needs.	Body-mass aggregations persist despite species turnover through active maintenance of organizational boundaries.

Table 8.1: Constraint regime criteria exemplified in language and ecological systems.

Constraint regimes manifest these three criteria in complementary ways. Emergence involves the spontaneous reduction of degrees of freedom across multiple elements, creating constraints that didn't exist at the component level. This generates measurable decreases in system variability and predictable patterns of organization. Discoverability reflects the capacity for heterogeneous newcomers to couple selectively into established patterns, with systematic success or failure based on compatibility with existing constraints. This creates formal patterns that enable recognition and structured interaction. This relates to the edges of clumps and innovation and endangerment related to them. Stability requires some processes that monitor, repair, and recreate constraints across multiple cycles, involving measurable work investment to maintain organizational coherence against entropic degradation.

These principles find clear expression in ecological body-mass aggregations. Emergence manifests as species spontaneously cluster into distinct size ranges through co-evolutionary processes that create mutual constraints between predator-prey relationships, resource availability, and habitat structure. Without central coordination, these interactions carve out distinct zones where certain sizes prove viable while others become systematically excluded, creating measurable gaps in size distributions across ecosystems. This mirrors the convergent evolution of wings in birds and bats, where distinct lineages are channeled toward similar forms because the constraints of aerodynamics limit the number of viable designs.

Discoverability becomes visible in how new species encounter established body-mass aggregations. When newcomers, whether through invasion, migration, or speciation, attempt to establish themselves, their success depends systematically on compatibility with existing constraint patterns. While these newcomers may indeed utilize novel resources, they cannot do so arbitrarily. Species with body masses aligning with established aggregations show higher establishment success, while those falling in persistent gaps face systematic exclusion.

Crucially, this integration follows a center-periphery dynamic common to many complex networks. Invasive species cluster disproportionately at the edges of established aggregations rather than the stable cores. While the specific dimensions of these aggregations may shift over time as environmental variables change, the structural necessity of entering at these transitional margins, where the regime is most flexible, remains constant. This foregrounds how constraint regimes operate as discoverable structures. Newcomers encounter predictable acceptance or rejection based on how their characteristics align with the established possibility space, delineating the regime's formal boundaries through systematic patterns of success and failure.

Stability manifests through aggregation persistence despite species turnover and environmental fluctuations. However, asserting the reality of these aggregations requires a robust standard of evidence. Here, William Wimsatt's concept of *robustness* provides the necessary philosophical grounding. Wimsatt argues that an entity or phenomenon is real if it is accessible via multiple independent modes of detection[264]. Constraint regimes satisfy this criterion, they are not just visible in body-mass plots but are independently detectable through invasion success rates, extinction probabilities, and energetic efficiencies. They appear as *local maxima of regularity*, stable organizational levels that persist by excluding specific behavioural or structural ranges. Having established their robustness, we can identify the mechanism of their persistence. Species actively maintain ecological conditions that preserve their aggregation's constraints through territorial defence, niche construction, and resource competition. This involves measurable energy investment in maintaining organizational boundaries, detectable work expenditure that distinguishes active constraint maintenance.

Constraint regimes operate across multiple species and evolutionary timescales, creating stable organizational domains through active exclusion. The gaps between aggregations is *no man's land* defining where stable patterns can persist, organizational fault lines where different constraint regimes meet[265].

Constraint regimes exist along a spectrum of constraint integration, from minimal mutual influence to complete organizational closure. This will be explored in the next chapter, but for now it suffices to notice that this spectrum provides a framework for understanding integration of a regime. Simple assemblages like rock piles exhibit minimal mutual constraint between components. Ecological interactions like predator-prey relationships involve significant but partial mutual constraint. Ecosystem niches display complex webs of mutual constraint spanning multiple species without achieving full closure. At the spectrum's end we again find full *constraint cycles*, systems achieving complete organizational closure where constraints fold back upon themselves, as in living cells where metabolic constraints maintain the very structures that enable metabolism. This spectrum explains observed variations in organizational stability and argues that *constraint cycles* are the *limiting case* of maximum constraint integration.

Beyond Biology

Constraint regimes extends across nature's organizational scales. Take electrons in atoms. They can only occupy specific energy levels, creating distinct gaps between these allowed states. We see this pattern every day, sodium street lamps glow their characteristic yellow and neon signs shine their perfect red because electrons in these atoms can only jump between particular energy levels. The light we see comes from electrons falling from one allowed level to another, releasing precise amounts of energy as specific colours. These patterns show up consistently across different measurement approaches, whether we study the light atoms emit, how they absorb X-rays, or their magnetic behaviour, we find the same forbidden zones where electrons simply cannot exist. The same spectral fingerprints that identify elements in our laboratory flames also find their presence in distant stars, unchanged over billions of years[266].

Chemistry shows similar patterns through molecular stability. Consider benzene, the six-carbon ring found in gasoline and countless other compounds. Its six mobile electrons create a stable molecule that is far more resistant to change than similar compounds without this specific electron count. Beyond benzene, the overall principal is that rings with 2, 6, 10, or 14 mobile electrons (following what chemists call the 4n+2 rule) all share this stability. Meanwhile, molecules with 4, 8, or 12 electrons become so unstable they literally twist themselves out of shape to escape this unfavourable configuration. We live with these stable aromatic compounds daily. They give coffee its aroma, keep DNA's structure intact, and form the backbone of many medicines. The unstable versions proved so difficult to create that chemists spent decades just trying to make them exist long enough to study[267].

Criterion	Physics: Atomic Electron Shells	Chemistry: Aromatic Molecules
Emergence	Quantum mechanics creates discrete energy levels with forbidden gaps between them, reducing continuous possibilities to specific allowed states.	Electron delocalization in rings with 4n+2 mobile electrons creates exceptional stability while 4n configurations become systematically excluded.
Discoverability	Photons couple only at energies matching allowed transitions between levels; electrons must align with pre-existing quantum states or be rejected.	Molecular designs succeed when electron counts follow the 4n+2 rule and fail when they violate it; unstable antiaromatic molecules distort to escape forbidden configurations.
Stability	Energy gaps persist across billions of years without maintenance; the same spectral lines identify elements in laboratory flames and distant stars.	Aromatic stability patterns hold across millions of different compounds; benzene-like molecules resist reactions that readily destroy non-aromatic analogs.

Table 8.2: Constraint regime criteria in physical and chemical systems.

These examples from physics and chemistry show that organization through exclusion is a reoccurring principle in nature, creating stable domains at multiple scales. Each step up in complexity, from subatomic particles to atoms, from atoms to molecules, is a shift in the dominant constraints that carve out new possibility spaces. This layering of constraints creates the natural boundaries that we approximately recognize as distinct scientific disciplines. Physics studies the rules governing the *base* layers, chemistry explores the possibilities that emerge from atomic constraints.

Yet, in making this universal claim, we also see a crucial dividing line that separates the living from the non-living. The stability of atomic shells or aromatic molecules is a *passive* consequence of fundamental laws. They require no ongoing work to maintain their integrity. Their constraints are given by the boundary conditions of quantum mechanics and thermodynamics. They simply *are*. In contrast, the constraint regimes of the living world are active, dynamic achievements. They must be continuously built, maintained, and defended against dissipation. A body-mass aggregation is not a static property but a dynamic pattern upheld by the collective work of its inhabitants, through foraging, competition, and niche construction. Herein lies the fundamental difference: abiotic regimes persist through thermodynamic necessity, while biotic regimes persist through organized work.

To conclude, this constraint-based understanding changes how we approach multilevel organization. Instead of cataloguing what exists at each level, we examine the systematic exclusions that create and maintain organizational boundaries. This emphasizes why certain configurations persist across different systems and scales. Not because they represent optimal solutions, but because they occupy viable regions within constraint landscapes that actively exclude alternatives. Significantly, this resolves the persistent tension between level autonomy and cross-level influence. Constraint regimes maintain their distinctiveness through systematic exclusions while remaining dynamically coupled through the very boundaries that define them.

Complex living systems require constant work to maintain their organizational coherence. Understanding static constraint regimes gives us only part of the picture, we must examine how these regimes interact and transform across scales and through time. This leads us to consider the framework of *Panarchy*, which provides a way to look at the dynamic patterns through which different levels of organization interact and maintain themselves.

8.4 Panarchy

Constraint regimes provide a lens on how different levels of organization maintain their distinctness through what they exclude rather than what they contain. Regimes create stable domains by defining what cannot exist or occur at a given scale. However, this structural understanding, while necessary, risks presenting too static of a picture. To get a broader picture of how living systems maintain themselves across scales, we are now going to examine the dynamic interactions between levels and their respective constraint regimes.

The stability of constraint regimes is not a given but an achievement, they maintain their boundaries through ongoing interactions across levels, requiring constant work to preserve their operational integrity.

These cross-regime interactions highlight something important. Levels of organization are neither fully autonomous nor merely subordinate parts of larger wholes. Instead, they participate in complex patterns of mutual influence and constraint that both maintain their distinctness and enable their transformation. This dynamic interplay is what the *Panarchy* framework, developed by Holling and Gunderson, helps us understand[268]. Building on their work in ecology, we can now reinterpret *Panarchy* through our theoretical lens of *constraint regimes* and *constraint cycles*, articulating how systems achieve stability through constant negotiation of constraints across scales.

At its core, Panarchy describes a *nested hierarchy* of *adaptive cycles* operating at different spatiotemporal scales. In our framework, these become *nested constraint cycles* operating within their respective *constraint regimes*. Each cycle maintains its own characteristic rhythm and boundaries, yet remains connected to other scales through crucial mechanisms of cross-level influence, the *revolt* and *remember* connections.

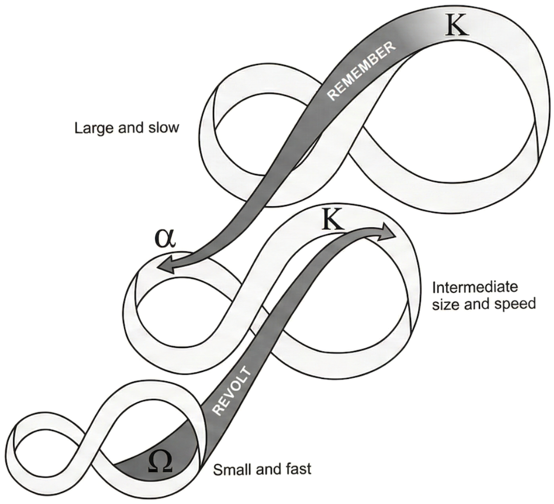

Figure 8.2: Illustration of multiple interconnected Panarchy loops at different spatiotemporal scales. Revolt and remember connections show ways of how different cycles can relate to each other. Reprinted from Gunderson and Holling [2002]

These serve as mechanisms through which disturbances or changes propagate through different constraint regimes. More fundamentally, they show how constraint regimes actively maintain themselves through cross-level coordination. The revolt connection occurs when a critical change in a fast, smaller cycle cascades upward, affecting larger and slower cycles. This phenomenon can initiate transformations at higher levels of the system, often when those larger cycles are in vulnerable stages.

As Holling illustrates: *An ecological example of this situation occurs when conditions in a forest allow for a local ignition to create a small ground fire that spreads to the crown of a tree, then to a patch in the forest, and then to a whole stand of trees. Each step in that cascade moves the transformation to a larger and slower level. A societal example occurs when local activist groups succeed in efforts to transform regional organizations and institutions because they had become broadly vulnerable.*[269]

This revolt connection demonstrates how seemingly localized disturbances in one constraint regime can trigger broader transformations, particularly when the larger system's constraint regime is fragile or susceptible to change. It highlights the potential for innovation and adaptation to emerge from lower levels in a system. In contrast, the remember connection works in the opposite direction, helping facilitate renewal by drawing upon the accumulated resources, and potential stored in the larger and slower cycles.

Holling describes the remember connection in action after a forest fire: *Once a catastrophe is triggered at a level, the opportunities and constraints for the renewal of the cycle are strongly organized by the K-phase of the next slower and larger level. After a fire in an ecosystem, processes and resources accumulated at a larger level slow the leakage of nutrients that have been mobilized and released into the soil. The options for renewal draw upon the seed bank, physical structures, and surviving species that form biotic legacies... that have accumulated during the growth of the forest.*[270]

The remember connection shows how higher-level constraint regimes actively guide and constrain the reorganization of lower levels, ensuring renewal occurs within viable boundaries while allowing adaptive modification.

This concept extends beyond ecology into other domains. Stewart Brand, describes how the *remember* connection allows society and even built environments to draw on accumulated experiences, experiments, and wisdom to renew and evolve over time. Brand's work provides non-biological examples of the panarchical principle of *remember*, where slow-changing levels provide stability and memory for faster-changing ones. In *How Buildings Learn*, he describes buildings in terms of *shearing layers* of change. The site is slowest, the structure is slow, while services, space plans, and furniture change much faster. The slower layers provide the enduring constraints within which faster adaptation occurs. Another famous example of his is the *Clock of the Long Now*, which is a project to build a monument-scale mechanical clock designed to last for 10,000 years, acting as a physical artifact of *remember* for civilization itself, a slow, high-inertia system intended to foster long-term thinking [271]. Remember thus provides stability and a foundation for future cycles by leveraging the enduring legacies of larger, slower processes.

The interaction between revolt and remember connections showcases how different constraint regimes maintain their distinctness while remaining dynamically connected. In a well-functioning system, these connections operate synergistically, with smaller levels invigorating the larger ones and larger levels providing stability and memory for renewal. As Holling noted, in a healthy system, each level operates at its own pace, *protected from above by slower, larger levels but invigorated from below by faster, smaller cycles of innovation.*

This positions constraint regimes as relational phenomena, they exist through their dynamic relationships with other scales, helping to maintain the others' operational boundaries.

Each new level inherits the stability of constraints established below while adding its own distinctive constraints. A multicellular organism, for instance, builds upon the already highly constrained organization of individual cells. These cells themselves are a dramatic reduction in the possible states their constituent molecules could adopt. Most possible molecular configurations are excluded by the constraints that maintain cellular organization. The emergence of multicellularity adds further constraints, restricting cellular behaviour through adhesion molecules, signalling pathways, and developmental programs that coordinate the actions of individual cells.

This can seem counter-intuitive. How does progressively reducing the state space of components, eliminating more and more possible configurations, enable greater complexity rather than diminishing it? The key insight is that these added constraints do not only restrict, they create reliable, reproducible patterns of interaction that enable new forms of organization. When cellular behaviour is tightly constrained by multicellular controls, cells can specialize into distinct types, form tissues, and coordinate their activities across space and time. The constraints that limit individual cellular freedom simultaneously enable collective capabilities that no cell could achieve alone.

This dynamic captures what Charles Sanders Peirce identified through his concept of habit, which he described nature's fundamental tendency to establish reliable regularities that become foundations for higher-order organization. Peirce recognized that emergence doesn't arise from unlimited possibility but from the progressive establishment of dependable patterns. Higher-order complexity emerges not from possibility but from the accumulated habits or regularities of lower levels, each constraint regime creating stable patterns that enable new forms of organization to emerge and persist.

This pattern, where reduced local freedom enables increased global complexity, appears throughout nature, even in non-living systems. Consider the example of mineral evolution documented by Robert Hazen. The early Earth hosted only about 60 mineral species found in meteorites. As new geological processes emerged, each stabilizing specific constraints on chemical and physical interactions, the diversity of possible minerals expanded dramatically. Magmatic and metamorphic processes enabled roughly 1,500 new mineral species. By the time living systems emerged, with their ability to control local chemical conditions precisely, Earth's mineral diversity had exploded to over 4,300 species[272].

Each stable constraint regime becomes a *base* upon which new patterns of organization can emerge and stabilize. This process is not only additive. Through revolt and remember connections, different levels remain in dynamic relationship, each influencing the others' evolution and stability.

Rather than static hierarchies, we see adaptive, yet stable, constraint regimes. Every level forming its own cluster of regularity and autonomy, while constantly engaged in cross-level interactions that maintain their boundaries. The *Panarchy* framework provides a powerful lens for seeing this dynamic maintenance, but the underlying principle extends beyond any single theoretical approach. It underscores the principle that complexity arises through the progressive layering of constraints, maintained and transformed through ongoing cross-level interactions.

8.5 Towards Distributed Integration

This framework of constraint regimes and their cross-level interactions through panarchical dynamics provides us with a useful tools for understanding how levels of organization maintain their distinctness while remaining dynamically coupled. Yet a crucial aspect has been left under-investigated. The *energetic* foundations of different constraint regimes. The atomic shells we discussed earlier, where electrons can only occupy specific energy levels, require no internal work. They persist without any energy investment from the system itself.

Constraint regimes in living systems operate differently. The body-mass aggregations in ecosystems, the predator-prey relationships that define ecological niches, the metabolic networks that maintain cellular organization, all of these require continuous work to sustain their boundaries. Species must actively maintain their ecological relationships through ongoing behavioural adaptations, niche construction, and evolutionary responses to environmental change. This work requirement creates a fundamental asymmetry, while abiotic constraint regimes achieve their stability through thermodynamic necessity, biotic constraint regimes must continuously regenerate some of their organizational constraints.

This energetic requirement introduces a crucial dimension. The degree to which the components *within* a constraint regime *coordinate their work* to maintain collective constraints. Some systems exhibit loose coordination, where individual components maintain their own constraint cycles with minimal mutual influence. Others achieve tight integration, where components sacrifice individual autonomy to maintain collective organizational boundaries. Between these extremes lies a spectrum of possibilities, each representing different ways that living systems and groups of living systems balance individual constraint maintenance with collective coordination.

Understanding this spectrum becomes essential for grasping how complexity scales beyond individual organisms. When multiple living entities interact, under what conditions do they compose into larger systems that exhibit organism-like properties? What determines whether a collective remains a loose aggregation or develops the tight coordination characteristic of integrated wholes? And how do these coordination patterns persist, strengthen, or dissolve over time?

When we look around us, we see countless examples of living things coming together to form larger wholes, from cells joining into tissues, to insects forming colonies, to humans creating societies. But what's fascinating is that these combinations don't always work the same way. Sometimes the parts stay quite independent, like birds in a flock. Other times they become deeply interconnected, like cells in your body. What makes the difference? I will argue that it comes down to how these parts work together to maintain their shared existence. The more energy and effort they invest in maintaining their collective boundaries and organization, the more they tend to function as a unified whole.

To address this, we need to examine concrete cases where this coordination unfolds across different scales and through different mechanisms. Three examples will frame our investigation, each stressing different aspects of how living systems achieve collective constraint maintenance while navigating the fundamental tension between individual autonomy and collective integration.

Chapter 9

Degrees

A hundred thousand cells are starving. They've stripped their patch of forest floor clean of bacteria, and now, in the gathering darkness, something extraordinary begins. A single cell pulses with a chemical signal—cyclic AMP—that ripples through the multitude like a wave through a crowd[273]. More cells take up the call, creating spiral patterns of molecular communication that draw the population together. The cells stream together in delicate ribbons, each individual partly surrendering its solitary existence as it joins the collective. Within hours, what was a scattered population of predatory amoebae has become a glistening mound of coordinated tissue.

But the slime mould is just getting started. The mound tips over and stretches into a finger-like projection that begins to crawl across the leaf litter with unsettling purpose. This is the slug phase, and it's here that the true marvel of collective organization becomes apparent. The slug, though composed of cells that were recently fierce competitors, moves as one organism. It senses light and temperature, leaving behind a trail of slime that gleams in the forest dimness. When it encounters an obstacle, the entire collective makes decisions about whether to crawl over or around. Some cells at the front act as pioneers, sacrificing themselves to test the environment for toxins, while others maintain the slug's shape and momentum. The slug can persist for days or weeks, prowling the forest floor until it finds the perfect spot for its final transformation.

Then, about a fifth of the cells commit evolutionary suicide[274]. They transform into a rigid stalk, their cellulose-hardened corpses lifting their former neighbours into the air as a spherical mass of spores. These survivors will drift away on the wind, carrying the potential for new colonies, and new collective organizations, in their genes. It's a glimpse into how life can flicker between individual and collective states, with entirely new capabilities emerging through coordination.

The slime mold's cycle finds an unexpected parallel in human societies, particularly in the seasonal transformations of Arctic peoples that Marcel Mauss documented in his groundbreaking study[275]. In summer, Inuit families disperse across vast territories, each occupying simple conical tents called *tupik* made of reindeer hide stretched over poles. These summer dwellings, housing only nuclear families, embody the individualistic nature of the warm season. Without the pressure of scarce resources, each family unit hunts independently, maintains its own property, and conducts only basic domestic rituals.

But as winter approaches and the sun's arc drops lower in the sky, a transformation begins. The scattered families converge into winter settlements, abandoning their individual tents for complex longhouses. These winter dwellings are marvels of architectural and social engineering. Built partially underground with a low entrance tunnel for warmth, they require anyone entering to bow down—a physical embodiment of the submission to collective constraints that winter demands. Inside, a raised platform runs along the walls, divided into family compartments but maintaining open connections that foster communal life.

The kashim, a separate communal house, becomes the centre of winter social and religious life. Unlike family dwellings, it features a central hearth and lacks internal divisions, creating a space where the community dissolves individual boundaries through shared ceremony and decision-making. Here, during the long dark months, the society undergoes what Mauss called *a season of heightened religious life*, with elaborate rituals like the Bladder Festival and the Feast of the Dead reinforcing collective identity.

This seasonal oscillation between dispersed and concentrated states highlights a deeper pattern. As the sun's energy gradient weakens and resources become scarce, the society responds by imposing stronger social constraints. Summer's individual property rights give way to winter's *generous collectivism* where food must be shared. Personal autonomy yields to communal authority, with conflicts resolved through formalized singing duels rather than violence. Even marriage bonds temporarily loosen during winter festivals, reflecting how individual family units partially dissolve into the larger social organism.

In the summer, each family maintains its own hearth and makes independent decisions. During the winter, the kashim's central fire becomes a shared energy source around which the community literally and figuratively reorganizes itself. The physical constraints of the Arctic winter—limited sunlight, scarce resources, need for shelter—drive the emergence of social constraints that enable collective survival. Like the slime mold's aggregation, this transformation creates new capabilities that no individual or family can achieve individually.

Perhaps the most subtle and pervasive example of collective coordination unfolds constantly in ecosystems, operating through molecular languages that span species barriers. When a birch tree is stressed by drought or insect attack, it releases volatile organic compounds that drift through the air and seep through soil networks of fungal hyphae (branching filaments)[276]. These chemical signals do not only warn its own leaves or other birches. They are intercepted by neighbouring species like maple and poplar, which respond by ramping up their own defences. This chemical dialogue can protect an entire forest stand, with trees of different species coordinating their defensive responses through shared molecular vocabularies.

Beyond a forest patch, certain molecules act as universal stress signals across vast evolutionary distances. One example is *astaxanthin*, a red-orange carotenoid that functions as a chemical semaphore throughout aquatic food webs. Initially produced by micro-algae under stress conditions like high UV exposure or nutrient limitation, astaxanthin serves as a potent antioxidant. When zooplankton like copepods consume these algae, they accumulate the pigment in their tissues, gaining protection against oxidative damage. Fish that eat the zooplankton also retain astaxanthin, where it triggers a cellular responses through the Nrf2 pathway, upregulating antioxidant enzymes and modulating inflammation. This cascading signal creates a form of distributed stress response across the ecosystem. When environmental conditions deteriorate, astaxanthin levels rise through the food web, preparing multiple species for hardship[277].

The ubiquity of these shared stress responses extends beyond aquatic systems. Research has shown that astaxanthin influences human metabolism through the same molecular pathways it triggers in fish, activating cellular stress resistance programs that mimic the benefits of fasting[278]. This deep evolutionary logic connects stress to scarcity. A widespread chemical warning of environmental hardship is a reliable cue to prepare for lean times, prompting organisms to activate the same cellular repair and survival programs that fasting does. This conservation of stress response mechanisms across such vast evolutionary distances suggests the existence of a widely universal molecular grammar. A shared language that helps living systems coordinate their responses to environmental challenges[279].

These three vignettes of the slime mold's metamorphosis, the Inuit's seasonal transformations, and nature's shared grammars loosen the tidy idea that life is expressed only inside neatly bounded individuals. Each case shows coordination stretching outward from basic biological agents into temporary or partial collectives whose coherence can alternate over time. In each vignette, feedback loops let many semi-autonomous parts sense, communicate, and enforce constraints that steer the whole through growth, disruption, re-organization, and renewed growth. Importantly, such distributed integration does not make a system a self-reproducing organism, but it does give the collective a measure of *organism-like* persistence. An ability to absorb shocks, reconfigure, and continue, that no single member could manage alone.

This points us toward the central puzzle around which this chapter revolves: How does the fundamental logic of self-constraint, scale up beyond individual organisms? When many living entities interact, how do they compose into larger systems that might themselves exhibit *organism-like* properties, even if incompletely? And what determines the degree to which these collective systems achieve coherent self-maintenance and purpose?

Three key motivations drive this investigation. First, we need to explain the growth of complexity in both biological and social worlds. How does it happen that simpler units combine to generate and sustain more sophisticated forms of organization. Second, we must move beyond some often too limiting dichotomies of organism versus environment and individual versus collective that have long limited our understanding of living systems. Third, we face urgent practical challenges, from ecosystem conservation to social organization to artificial intelligence design, that require a deeper understanding of how collective systems maintain and transform themselves.

The tools for answering these questions lie in the theoretical framework we've developed thus far but need to be explicated so that we can begin to map out what we will define as *Distributed Integration* based on how robustly constraint regimes replicate and maintain their constraints. This will us develop a spectrum of *organism-like* properties that extends beyond individual organisms, helping us understand how complex composite systems achieve and maintain coherence.

The crux lies in recognizing a pattern from our vignettes. In each case we find living systems investing energy and effort not just in their own maintenance, but in maintaining collective patterns. The slime mould cells sacrifice individual reproduction to coordinate movement and survival. Inuit families sacrifice autonomy to maintain cohesive social structures well-adapted for the resource-scare winter. Trees sacrifice resources to produce and respond to chemical signals. This enables us to define a dimension we will call *Distributed Integration*: the degree to which components actively work to sense, communicate, and enforce constraints that enable collective coordination through cycles of growth, disruption, and reorganization.

Allowing us to articulate the following relationship: biotic constraint regimes and distributed integration are two sides of the same coin. Constraint regimes define the negative space of which configurations are systematically excluded. Distributed integration captures the positive achievement, how effectively components perform the active work of sensing, communicating, and enforcing those constraints. You cannot have robust constraint regimes in living systems without the distributed work that maintains them, and you cannot have meaningful distributed integration without boundaries kept up through work.

9.1 Distributed Integration

Composite living systems face an energetic challenge. Available energy can fuel immediate growth or maintain forms of collective self-constraint. This fundamental asymmetry, that *growth is easy, self-constraint is difficult*, determines whether systems transcend aggregation to become partially integrated wholes capable of persisting through challenging transitions.

Distributed Integration (DI) measures the degree to which systems achieve collective self-constraint. Systems with minimal DI remain aggregations that likely dissolve under pressure, while those achieving higher DI can maintain organizational coherence through cycles of growth, disruption and re-organization, navigating disruption as integrated wholes.

We have seen that growth of certain patterns happens naturally from energy gradients through self-organizing processes. Whenever an energy gradient becomes available, systems expand to exploit it, bacterial colonies proliferate, forests spread, markets grow. In living systems, this dynamic requires no additional coordination beyond the acceleration of the fundamental tendency of life to replicate.

Self-constraint demands something different, namely the coupling additional constraining dynamics to the basic growth tendency. This requires continuous energetic investment in constraint maintenance—energy diverted from immediate expansion. The slime mold sacrifices individual cell division to produce coordination signals. Inuit families forgo personal accumulation to maintain collective institutions. Trees allocate metabolic resources to warning neighbours rather than additional growth.

This creates a *tragedy of the commons* dynamic where individual components gain immediate advantage by avoiding collective self-constraint, while others bear the costs. An amoeba that continues dividing during starvation can sometimes outcompetes those investing in collective behaviour. Without robust enforcement mechanisms preventing such defection, systems face a race to the bottom where collective self-constraint becomes impossible to maintain[280].

Here again, energy is key. Systems that coordinate only for growth typically fragment during disruption, leaving only the specialized parts to survive. Higher distributed integration, achieved through continuous investments in collective self-constraint, enables organizational coherence that preserves composite diversity through challenging transitions rather than selecting for isolated survivors.

But how is this organizational coherence achieved? It does not emerge spontaneously at the moment of disruption. Rather, it is built and maintained through specific, ongoing energetic investments that prepare the system for such events. To better conceptualize this, we can examine a couple of key forms this investment can take, each enabling collective self-constraint by building the capacity for a coordinated response before it is actually needed.

Latent potential involves maintaining encoded capacities for collective self-constraint even when not actively expressed. The slime mold exemplifies this clearly, where individual amoebae continuously maintain expensive molecular machinery, in this case, cyclic AMP receptors, adhesion proteins, differentiation programs that enables rapid collective organization when environmental disruption demands coordinated response rather than individual competition. In essence, the system pays a continuous energetic tax to maintain a collective response toolkit that remains dormant until a crisis demands its activation.

Shared grammars require ongoing investment in semiotic capacities that enable constraint coordination across components. The astaxanthin stress pathway showcases this principle operating across and entire ecosystems. Participating species, from microalgae through zooplankton to fish—must continuously maintain distinct but interconnected capabilities: the molecular machinery to produce astaxanthin signals, the receptor systems to detect these chemical messages, and the metabolic pathways to respond appropriately. This substantial energetic is diverted from growth into maintaining interpretive and expressive abilities for cross-species communication.

When environmental stress triggers this system, the payoff becomes evident. Rather than each species responding individually or failing to respond at all, the shared chemical grammar enables coordinated ecosystem-wide preparation for challenging conditions. Astaxanthin levels rise through food webs in orchestrated waves, triggering cellular stress responses that prepare multiple species simultaneously for environmental deterioration. The continuous investment in maintaining these semiotic capacities enables collective navigation of disruption that preserves ecosystem organizational coherence rather than fragmenting into isolated survivors.

Both mechanisms illustrate the same principle. Achieving *distributed integration* requires paying ongoing energetic costs during stable periods to enable collective organizational coherence when systems face disruption and must reorganize rather than dissolve.

It is relevant to note that various systems can exists along a spectrum of different levels of *distributed integration*. At minimal levels, systems are simple aggregations. At intermediate levels, like ecosystems with shared chemical grammars, systems can maintain partial coherence through disruption while preserving greater diversity than individual optimization allows. At maximum levels, approaching the complete organizational closure of a constraint cycle, collective self-constraint enables robust cyclical coherence across all phases.

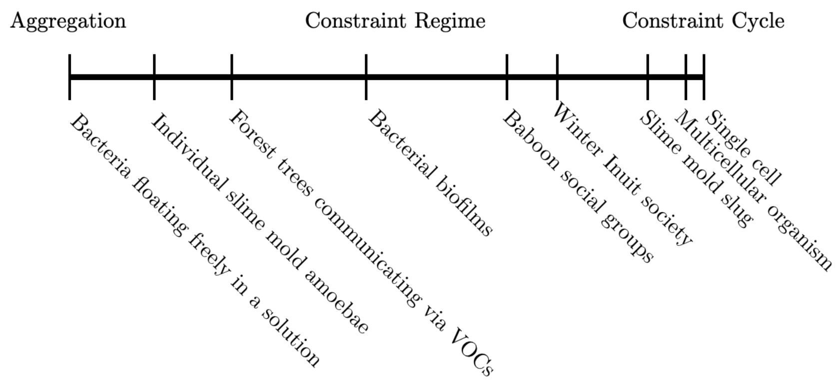

Figure 9.1: A schematic of the Distributed Integration spectrum. The ranking of these systems is a conceptual estimation, intended to visualize the degree of investment in collective self-constraint, from loose coordination to the full organizational closure of an organism.

The core principle of distributed integration, the energetic investment in collective self-constraint, provides a theoretical basis for thinking about how composite systems can achieve organizational coherence in biology. To make this framework more useful, we need ways to measure and compare distributed integration across different types of systems.

Operationalizing Distributed Integration

Turning Distributed Integration (DI) into something measurable is difficult and ultimately beyond the work of this thesis. What counts as collective self-constraint is context-dependent, often buried inside ordinary metabolism, and spread across partners. Boundaries and time windows matter a lot. Still, a simple energy ratio gives us a practical starting point:

$$DI = \frac{E_{\text{self-constraint}}}{E_{\text{total}}}$$

Here, E_{total} is the total energetic throughput over a defined boundary and time window. $E_{\text{self-constraint}}$ is the energy invested in forms of collective self-constraint[281].

While this ratio provides a necessary energetic accounting, measuring the *cost* of collective organization, it leaves the *structural consequence* implicit. Knowing how much energy a system spends on glue does not tell us how strong the bond is. Yet, the central hypothesis of this framework is that this metabolic work purchases genuine high-level causation. (In the Appendix, we formalize this transaction using Transition Probability Matrices and the mathematics of Causal Emergence, we demonstrate how the energetic investments described here do not waste heat, but generate measurable shifts in the system's causal hierarchy, migrating agency from the micro-scale of components to the macro-scale of the whole.)

In practice, the best way to measure this ratio is to look at how energy flows through entire systems. Also, for now we will only compare similar systems with themselves as this allows us to find some data. For plants that use photosynthesis, we can use their net primary production (NPP) or total photosynthate as our baseline measurement. For bacteria and other microbes, we measure how much carbon they consume from their environment. For bee colonies, we can look at their honey consumption over a season or how much energy they use to maintain their temperature as a group.

For the top part of our ratio, the energy spent on collective maintenance, we look specifically at how much energy goes into keeping the group working together. In plants, this means measuring how much sugar they share with helpful fungi in their root system, plus the amount of resources they use to attract and keep beneficial partnerships with other organisms. In bacterial biofilms, we measure the energy they spend making the substances that hold them together and the chemical signals they use to communicate. In honeybee colonies, we look at how much energy they spend keeping their nursery areas and winter clusters at the right temperature for everyone's survival. These examples are intentionally illustrative. It is crucial to treat this ratio as an exploratory heuristic, valuable for comparing relative investment within a single type of system but not for making absolute comparisons across different systems where the underlying metrics for total and constraint energy are not easily comparable.

System	**Denominator (E_{total})**	**Numerator ($E_{\text{self-constraint}}$)**
Wheat *(& many other plants)*	Seasonal net primary production (NPP) or total photosynthate	Photosynthate/NPP allocated to mycorrhizal fungi ($\sim$ 5–20%)
Bacteria *Pseudomonas aeruginosa*	Substrate carbon uptake	EPS matrix production + quorum signals and public-goods production
Baboons	Seasonal time/energy budget (proxied by foraging time)	Time/energy invested in social maintenance, primarily grooming ($\sim$ 5–15% of time budget)

Table 9.1: Operationalizing DI in three within-species systems

The table above provides a overview of three within-species applications of the DI ratio. Below, we expand on these examples with more detail.

Let's look at some concrete examples of how this ratio works in different living systems. In plants, we can see how investing in partnerships helps survival. When plants share more of their sugar with friendly fungi in their roots (called mycorrhizal fungi), they're better at surviving drought. Scientists can track this sharing using special isotope markers, which show that plants typically give these fungi on the order of 5 to 20 percent of their photosynthate/NPP. They also spend a bit more energy on chemicals that help maintain these partnerships. Studies show that plants that share more with their fungal partners have better chances of surviving drought and recover faster afterward[282]. However, the relationship is not necessarily linear, it is possible for a system to waste energy on coordination that yields no structural benefit.

In bacteria like *Pseudomonas aeruginosa*, we see something similar but in a different context. Some bacterial cells invest more energy in producing sticky substances that hold their communities together (called EPS) and chemical signals that help them coordinate as a group. In mature biofilms, EPS often constitutes the majority of the matrix by mass, reflecting sustained collective investment. When these bacteria face antibiotics, the ones that invested more in these collective features survive much better than their counterparts that didn't make these investments. The difference in survival rates can be enormous, sometimes thousands of times better for the bacteria that invested in collective protection[283].

Baboons offer another example of this principle operating in a complex social mammal. Female baboons, invest a significant portion of their daily time budget (between 5 and 15 percent) in affiliative social behaviors, primarily grooming. This is a direct energetic cost, as time spent maintaining social bonds is time diverted from immediate individual foraging and energy acquisition. The payoff for this sustained investment in collective self-constraint is a dramatic increase in resilience. Long-term studies on the Amboseli baboons show that females with stronger social integration, a direct result of these grooming investments, exhibit significantly higher adult survival rates. This increased survival is not an end in itself, reproductive lifespan is the single largest determinant of fitness. Therefore, by investing in social cohesion, these individuals directly translate grooming behaviour into higher lifetime reproductive success[284].

Among the three examples, the case of baboons provides the clearest numerical illustration of how a change in distributed integration translates into a difference in survival. Long-term studies in the Amboseli population quantify social integration through a *Composite Sociality Index* (CSI). This index is constructed from the frequency and duration of affiliative behaviours such as grooming, measured in minutes per hour and normalized relative to the group average. A value of CSI $= 1$ therefore represents average integration, values greater than 1 indicate stronger-than-average social bonds, and values less than 1 weaker-than-average bonds. In statistical models of infant survival, the coefficient linking maternal CSI to offspring survival takes the form $\beta = -0.854$ in a Cox proportional hazards model. In such a model, the parameter β specifies how the hazard of death (the instantaneous risk of dying at a given time) changes with a one-unit increase in the predictor variable, in this case, maternal social integration. The hazard ratio (HR) is given by the exponential transformation of this coefficient: $\text{HR} = e^{\beta \Delta \text{CSI}}$. Substituting the values, a one-unit increase in CSI gives $\text{HR} = e^{-0.854} \approx 0.43$. Put more plainly, an infant born to a mother with a CSI score one unit higher than another's faces only 43% of the mortality risk, a reduction of about 57% in the instantaneous hazard of death. If the mother's CSI is two units higher, the hazard ratio compounds multiplicatively: $\text{HR} = e^{-1.708} \approx 0.18$, corresponding to an 82% reduction in risk.

What makes this finding especially powerful for our purposes is how modest the underlying energetic cost is. Female baboons invest on the order of 5–15% of their daily time budget into grooming, time that could otherwise be spent directly foraging. This allocation represents the numerator in our distributed integration ratio, a deliberate diversion of resources toward maintaining collective ties. The payoff, however, is disproportionately large. A small shift in investment, a few additional minutes per day spent maintaining social cohesion, maps onto dramatic differences in the survival probability of the next generation. In the language of distributed integration, this case shows that even incremental increases in the energetic work of collective self-constraint (grooming and affiliative maintenance) yield non-linear benefits for persistence. The baboon example therefore anchors the broader claim of this section: higher levels of distributed integration, even when costly in the short term, directly translate into greater reconstructability and long-term fitness at the level that matters most—offspring survival and reproductive lifespan.

This ratio is an approximation of the energy budget devoted to uphold forms of collective self-constraint and it gives a directional indicator. As DI rises, immediate growth in benign conditions tends to slow relative to low DI variants of that system, while reconstructability after shocks to improves for high DI systems.

While this energy ratio provides a useful metric, several important limitations warrant consideration. First, certain critical regulatory processes have minimal direct energetic costs yet substantial systemic impacts. For example, social insect colonies maintain order through low-energy but precise policing mechanisms. Second, many maintenance investments show considerable plasticity, upregulating in response to environmental stressors (e.g., drought-induced increases in mycorrhizal carbon allocation), which complicates straightforward causal analysis. Third, the same expenditure often does double duty. EPS in biofilms both glues a group together and traps nutrients. Disentangling the collective-maintenance share from individual-growth share is non-trivial. Fourth, to operationalise this metric, counterfactuals are required. To label an energy flow as *self-constraint* one needs to know the outcome without it.

Despite these methodological challenges, even this relatively simple ratio successfully differentiates meaningful within-species variation in collective investment strategies. The data shows that higher maintenance investment correlates with enhanced stress resilience. Plants allocating larger carbon fractions to fungal symbionts demonstrate superior drought tolerance. *P. aeruginosa* strains maintaining more robust extracellular matrices exhibit greater antibiotic resistance and female baboons that invest more in building social bonds achieve markedly higher adult survival and lifetime reproductive success.

Further methodological development remains crucial for robust quantification of distributed integration. Key experimental priorities include, first, controlled paired-contrast studies that systematically manipulate $E_{\text{self-constraint}}$ while maintaining constant E_{total} and environmental parameters. It would be important to focus on standardized temporal protocols for measuring system responses across pre-perturbation, acute stress, and recovery phases, with clearly defined outcome metrics such as survival coefficients and mean time to reconstruction. Finally, it will be crucial to develop high-resolution metabolic inventories that can clearly separate the energy costs for individual upkeep from those for collective functions. The ratio will not capture all that matters about DI, but it offers a place to start testing a simple claim, that is that paying for shared boundaries, signals, and homeostasis buys reconstructability when it counts.

Closure Types

When discussing the degree to which composite system exhibit self-constraining dynamics there is an important caveat to consider. High distributed integration can emerge in systems with fundamentally different organizational architectures. In other words, the DI metric tells us how much a system invests in coordination, but not what kind of collective it is. A slime mold aggregation during starvation and a multicellular organism both invest substantial energy in collective constraint maintenance, yet they represent qualitatively different types of composite organization that respond differently to disruption and selection pressures.

The difference I want to highlight here lies in the nested organizational dynamics we explored earlier. When self-constraining units combine, they can exist in equilibrium (coexisting without meaningful interaction), like bacteria floating freely in a solution. They can participate in self-organizing patterns (collective coordination while maintaining individual boundaries), as in bacterial biofilms or penguin huddling behaviours, or form genuine self-constraining wholes (higher-order integration with consolidated reproduction and unified boundaries), exemplified by multicellular organisms where individual cells maintain collective organizational integrity[285]. Distributed integration measures coordination intensity across all these types, but cannot distinguish between their underlying architectures.

For the slime mold's aggregation phase we find self-organizing of self-constraining systems. Individual cells coordinate collectively but can disaggregate and return to independent existence when conditions change. A multicellular organism exemplifies self-constraining of self-constraining dynamics. Here cells work together to maintain higher-order organizational constraints that define the organism's integrity. Both may score similarly on DI measures during peak coordination, yet they exhibit fundamentally different types of organizational complexity.

This analysis is inherently level-dependent, we cannot directly compare DI scores across different organizational types to determine overall complexity. A single cell may exhibit higher distributed integration than an entire human society, but this doesn't make the cell more complex than the society. We're looking at different kinds of coordination at different scales. Rather, we're measuring how tightly components coordinate within their respective organizational architectures. The cell achieves near-maximal constraint coordination among its molecular components, while human societies typically maintain much looser coordination among their individual members. Importantly, if a human society were to approach the degree of distributed integration found in a single cell with individuals as tightly coordinated as cellular components, it would become something remarkably close to an organism-like entity, suggesting that DI measurements underscore the relative proximity of any composite system to achieving *organismic-level* integration within its particular scale and architecture.

Understanding both distributed integration and closure types together provides a more complete picture of how composite systems achieve and maintain coherence. Put simply, distributed integration measures the strength of the coordination, while closure type describes the kind of organization that coordination creates. High DI enables effective coordination within any organizational architecture, while the specific closure type determines the system's fundamental level of *dynamical depth*, whether we're dealing with simple equilibrium arrangements, self-organizing collective patterns, or genuine self-constraining wholes.

9.2 Selective Aspects

Evolutionary theory is excellent at counting the *singers*, but often struggles to account for the *song*[286]. The singers—genes, organisms, groups—are the discrete, reproducing entities that we can easily track. But the song is the pattern of organization they participate in like the symbiotic relationship, the social structure or a larger adaptive cycle. The song can persist even as the singers change, carried from one generation to the next. So, can the song itself be a target of selection?

To answer this, we need a way to see the song not just as an ephemeral performance, but as a robust, self-sustaining entity. This is what our framework provides. The *song*, its melody and structure are *constraint regimes*—the rules that give it a recognizable form. The *effort* the singers invest to stay in harmony and keep the performance going is its *distributed integration.* Together, these concepts allow us to see how a pattern can become so coherent that it develops its own evolutionary trajectory, its success measured not by how many new singers it creates, but by its sheer refusal to fade away in a noisy world.

Understanding when patterns become selectable requires recognizing our spectrum of organizational coherence. At one end are loose aggregations where components barely coordinate and invest almost nothing in maintaining their shared arrangement. In the middle range, we find constraint regimes where ongoing work maintains boundaries, signals, and organizational rules, allowing patterns to persist. At the far end are systems with full constraint cycles and high closure, tight integration where the whole often reproduces as a unified entity in an *organism-like* fashion. Selection can already begin to *grip* patterns well before full closure emerges. Once systems invest sufficiently in maintaining their coordination, the pattern itself becomes visible to selective pressures. Full closure strengthens this grip considerably, but it represents an endpoint rather than a prerequisite for *regime-level selection.*

When we shift from thinking about individual success to regime success, the relevant metrics change. A regime's success cannot be measured by the offspring count of any single body, because regimes are more like songs carried by various singers than like individual organisms seeking reproduction[287]. Instead, regime success depends on two interrelated capacities, how well the pattern persists through environmental shocks and disruptions, and how readily it gets picked up again, whether by the same hosts recovering from stress or by entirely new participants encountering the pattern for the first time.

To understand how such patterns operate as targets of selection, think about one of the most familiar examples from human cultivation: the garden tomato. In its wild state, the tomato produces small, energy-efficient fruits packed with seeds.

A strategy finely tuned by natural selection for reproductive success in competitive environments. Walk through any modern garden, however, and you encounter something thermodynamically confusing. Plants investing enormous energy to produce oversized fruits with relatively few seeds, a strategy that would prove catastrophically maladaptive in any natural setting. These plants survive only within the protective embrace of human cultivation, including regular watering, pest control, soil amendment, and careful saving and sowing of seeds.

Yet when we select for larger tomatoes, we do not simply select for individual plant characteristics. Each act of selection, saving seeds from the biggest fruits, providing optimal growing conditions, protecting plants from competition, simultaneously strengthens an entire pattern of relationships. The gardener's knowledge, the tool systems, the soil amendments, the protective structures, the timing of planting cycles. All of these elements become reinforced through the selective process. We select not just for bigger tomatoes, but for a *constraint regime* that makes such tomatoes possible.

This pattern's coherence becomes apparent when we disrupt it. Abandoned garden plots revert to wild-type characteristics within a few generations. Remove human cultivation practices, and domestic tomatoes rapidly lose their exaggerated fruit size. Conversely, remove access to cultivated varieties, and human communities historically shifted toward different subsistence strategies or developed alternative cultivation techniques. What persists across these perturbations is neither the individual organisms nor their specific traits, but the capacity to regenerate productive relationships, the constraint regime's ability to reconstitute itself through new arrangements of the same basic pattern.

The tomato example illustrates a broader principle about how patterns maintain themselves across time. At its core, this maintenance depends on systems' ability to carry forward organizational information without requiring organism-style reproduction[288]. This is where the concept of Distributed Integration offers a productive alternative to traditional debates about the level of selection. Rather than getting stuck on the question of whether selection acts on individuals or groups, DI allows us to ask a more direct, empirical question. How much energy does a system invest in its own maintenance? This reframes the problem by focusing on the work that makes a pattern a viable, survivable entity in the first place. With this energetic investment as our metric, we can then examine how different patterns fare through cycles of growth, disruption, and reorganization. Some achieve sufficient integration to persist across environmental changes, while others collapse when their energetic foundations shift. The key question becomes which patterns of constraint maintenance prove viable across perturbation cycles.

To understand how selection acts on patterns, we can return to the classical requirements formulated by Lewontin for any evolutionary process: variation among units, differences in success that stem from this variation, and some mechanism of inheritance[289]. These familiar criteria take on a distinct form when applied to patterns rather than bounded units. Variation appears as meaningful differences between regimes, in how much they invest in maintenance, in their levels of coordination, in whether they preserve alternative pathways or streamline toward efficiency. Differential success manifests through cycles of disruption and recovery, with some patterns proving more durable or more readily reconstructed than others. And inheritance operates through memory channels that carry patterns forward, whether through physical arrangements, transmitted practices, or partnership networks. What makes this framework distinctive is showing how these criteria can be satisfied without requiring organism-style reproduction. Selection begins acting on patterns once these conditions are met, with its grip strengthening as systems invest more in integration and memory. The threshold for select-ability is not absolute but forms a spectrum of increasing preservation capability.

If distributed integration makes patterns visible to selection, it also exposes them to a familiar temptation, which is the tension between short-term metabolic efficiency and long-term adaptive capacity. In quiet periods, regimes learn to save energy by trimming "excess" options like backup pathways, partner diversity, redundant reserves, because maintaining them costs work $E_{\text{self-constraint}}$. While this streamlining increases immediate output per unit of energy, it can harden into reliance, so that the pattern's very success depends on a few narrow inputs. These *degradation* *dependencies* represent a loss of causal structure where the system becomes highly optimized for the present moment but hollowed out for the future.

Degradation Dependencies

Degradation dependencies are the hidden price of efficiency in selected constraint regimes. They are more than just a strategic choice to favour one organizational pattern over another. They reflect a fundamental erosion of a system's ability to reorganize itself when conditions change. This erosion follows directly from thermodynamic principles, namely that maintaining unused capacities requires constant energetic investment, and without such investment, these capacities inevitably decay. The result is an addiction-like dynamic where systems become increasingly locked into narrow patterns of organization, unable to adapt even when their current arrangements become maladaptive[290].

The relationship between reef-building corals and their symbiotic algae provides an example of this principle operating at the interface between organisms. Most coral species have become entirely dependent on single-celled algae called zooxanthellae that live within their tissues, providing up to 95% of the coral's energy needs through photosynthesis. This arrangement enables the remarkable productivity that makes coral reefs among the most biodiverse ecosystems on Earth, supporting complex food webs in nutrient-poor tropical waters. However, this comes at a cost, corals have largely lost the metabolic machinery for independent nutrition, creating a *degradation dependency* that makes them vulnerable to environmental perturbations that disrupt the symbiotic relationship.

What makes this case particularly instructive is how the dependency emerged from an immensely successful evolutionary partnership. The coral, an animal, lacks any innate capacity for photosynthesis. Its pact with the algae is a classic division of labour where the zooxanthellae act as internal solar farms, fixing carbon and providing a steady stream of energy, while the coral provides shelter and raw materials.

This arrangement is so efficient that it freed the coral from the metabolic demands of full-time predation.

By outsourcing its primary energy acquisition, the coral could redirect its own metabolic resources toward what it does best, precipitating calcium carbonate to build its reef structure, producing defensive compounds, and growing complex colonial forms. This specialization enabled the construction of massive limestone structures that create habitat for thousands of other species. Over millions of years however, this reliance caused the coral's own independent feeding mechanisms to be de-emphasized and, in many cases, to atrophy. The dependency, therefore, is not a choice, but an evolutionary lock-in born from specialization.

When water temperatures rise even a few degrees above normal ranges, the algal symbionts begin producing toxic compounds that threaten coral tissues. The coral's only defence is to expel the algae. This phenomenon is called coral bleaching. Without their photosynthetic partners, corals must rely on their severely diminished capacity for independent feeding, often leading to starvation and death if the thermal stress persists. Naturally, corals cannot quickly reconstitute the metabolic pathways they abandoned over millions of years of symbiotic evolution.

Degradation dependencies differ fundamentally from beneficial specialization through their systematic elimination of alternatives rather than their temporary suppression. In multicellular organisms, cellular specialization maintains latent developmental potentials that can be reactivated during regeneration or stress responses. Degradation dependencies, by contrast, involve the actual loss of organizational capabilities. The deterioration of genetic sequences, the loss of cultural knowledge, the extinction of supporting species. Recovery requires not just reactivation but genuine rebuilding, often from a significantly compromised starting position[291].

This vulnerability expresses a fundamental cybernetic principle, Ashby's Law of Requisite Variety. Ashby demonstrated that for a system to maintain stability against environmental disturbances, its internal repertoire of responses must match or exceed the variety of disturbances it encounters. Only variety can absorb variety[292].

Efficiency, from this perspective, systematically reduces internal variety to match current conditions. When a constraint regime becomes streamlined, eliminating the energetic costs of maintaining alternative pathways it becomes well-adapted to present circumstances while losing the capacity to manage future volatility.

The crucial insight about degradation dependencies is that many remain hidden until environmental perturbation forces systems to rely on their eliminated alternatives. Systems can appear robust and highly integrated while harbouring critical vulnerabilities that become apparent only when their input dependencies face disruption. This creates a possible distinction between *perturbed* and *unperturbed* constraint regimes whose resilience has been tested versus those that persist only within constant environmental conditions.

These perturbations also reveal different levels of adaptive capacity. Some coral populations show greater thermal tolerance, either through hosting different algal strains or maintaining greater flexibility in their symbiotic relationships. Reef systems can sometimes recover from bleaching events if temperatures return to tolerable ranges quickly enough. These recovery processes show how constraint regimes can reorganize themselves while preserving their patterns. Corals may acquire mutated algal partners or shift the balance of their symbiotic relationships without abandoning the fundamental architecture that enables reef construction.

Unperturbed Dependencies

The distinction between *perturbed* and *unperturbed* constraint regimes proves relevant as we examine human systems that have developed during periods of exceptional stability. While coral reefs demonstrate how long-established patterns respond to disruption, many of our most sophisticated systems, from digital infrastructure to global supply chains, have yet to face their major perturbation tests. Modern industrial agriculture offers one example of how pattern selection under unusually stable conditions can create hidden fragilities.

Consider first how traditional farming systems maintained themselves across time. Through the 1940s, agriculture typically generated 2.3 food calories for every fossil calorie invested. A positive energy return that sustained human-plant coordination patterns across centuries of civilizational disruptions. These systems preserved their patterns through multiple memory channels including physical scaffolds in soil structure and water management, stored protocols in cultivation practices and crop varieties, and maintained partner pools in seed saving networks and farming communities. When disruptions occurred, these memory channels enabled agricultural patterns to reconstruct themselves, even if particular farms or communities failed.

Modern industrial agriculture inverts this arrangement entirely. Current field crops require approximately three fossil calories to produce one food calorie at the farm gate, rising to ten to thirteen fossil calories per food calorie delivered to the plate once we account for processing, distribution, and preparation[293]. This negative energy return reflects a fundamental shift in how the pattern maintains itself, from relatively autonomous local cycles to long-distance dependencies on fossil fuel extraction, chemical processing, and global distribution networks.

This exemplifies how selection under stable conditions can favour patterns that maximize efficiency by eliminating seemingly redundant capacities. Industrial agriculture achieves extraordinary productivity by streamlining processes, specializing functions, and reducing local variation. This arrangement creates remarkable distributed integration, coordinating millions of actors across global scales, yet this very sophistication will likely prove brittle.

We cannot know for certain because industrial agriculture's major perturbation test remains largely hypothetical. The system has developed during a historically unprecedented period of abundant, cheap fossil energy, stable climate, and reliable international trade. We have limited experience of how this pattern might reconstruct itself under significant energy constraints, supply chain disruptions, or the rapid transitions that climate change might require. Some partial perturbations—the 1970s oil crises, Cuba's *Special Period* following Soviet collapse—offer glimpses of potential vulnerabilities, but the pattern's fundamental resilience remains untested.

This points to a distinction in how patterns respond to stress. Systems that maintain alternative capacities, redundant pathways, diverse partnerships, multiple scales of organization, may experience frequent small adjustments but can usually reconstruct their patterns. By contrast, systems that quickly degrade alternatives in favour of efficiency often appear more stable during normal operations but face sudden, irreversible failures when their core assumptions are violated. This highlights the fundamental value of diversity within constraint regimes. Maintaining different varieties, practices, and relationships requires ongoing energetic investment but provides the raw materials for adaptation and reconstruction after disruption. The very features that make industrial agriculture so productive, its tight coupling, specialized dependencies, and elimination of "inefficient" alternatives can amplify its vulnerability to unprecedented changes.

The framework thus suggests examining not just how well patterns perform under current conditions, but how they maintain their capacity to reconstruct themselves when conditions shift. Do they invest in preserving alternative arrangements even at some cost to immediate efficiency? Do they maintain memory channels that could guide reorganization after disruption? Have they demonstrated their ability to recover from genuine perturbations, or have they existed only in the shelter of unusual stability? These questions become particularly crucial as we face increasing environmental, geopolitical and societal variability. In our case, the very efficiency that made systems successful during stable periods may undermine their ability to persist through more turbulent times.

Positioning Among Existing Theories

This framework is less an extension of existing evolutionary theories than a reorientation. It shifts the focus from thinking about selection of discrete *units* to selection on *patterns* and *processes*. Instead of counting offspring, it tracks how organizational patterns reconstruct themselves across disruption cycles. Additionally, the framework examines pattern persistence through the energetic investments systems make in maintaining coordination.

To clarify what this reorientation offers, we will position our framework against four pillars of contemporary evolutionary theory. We will begin with the classical gene-centric view of the Modern Synthesis, then broaden our scope to the hierarchical perspective of Multi-level Selection Theory. From there, we will consider the active role of organisms in shaping their worlds, as described by Niche Construction and the Extended Evolutionary Synthesis. Finally, we will address the question of how new forms of individuality arise, drawing on the theory of Major Transitions in Evolution. This comparative analysis will foreground how a focus on energetic maintenance provides a common thread connecting these diverse perspectives, while also offering a distinct way to understand the persistence of life.

The modern synthesis treats genes as the fundamental units of selection, with traits spreading when they increase expected gene transmission[294]. This framework has dominated evolutionary thinking for decades, explaining everything from altruism to cooperation through the lens of genetic self-interest. The constraint regime approach does not fundamentally disagree with this gene-centric view, rather, it reframes its context. When self-constraining dynamics rely on genetic inheritance, the gene is simply one of the most successful memory channels through which a pattern persists.

The framework treats genes as contributors to maintainable organizational patterns rather than as autonomous replicators competing solely for transmission. Genes succeed not just by producing more copies, but by supporting constraint regimes that can reconstruct themselves reliably after environmental disruption. This shifts the fitness currency from immediate reproductive output to long-term pattern persistence, which requires investments in coordination and memory that pure gene-centric models struggle to explain.

We can take the logic one step further and apply the concept of a constraint regime to the gene itself. A gene, after all, is not a replicator in a vacuum, it is an informational pattern whose stability is actively maintained by an entire system of molecular machinery, yielding the whole cell. From this perspective, a gene in a cell is not just a driver of a constraint regime—it is one. The cell is a pocket of extraordinary organizational coherence, a self-constraining dynamic that has achieved near-maximal distributed integration. For billions of years, this specific pattern has persisted by directing energy into sophisticated mechanisms for error correction, replication, and protection.

This dual view preserves the basic Darwinian logic of variation, selection, and inheritance while broadening inheritance to include the scaffolds, protocols, and partnerships that carry organizational information forward alongside genetic transmission. This makes genes important contributors to evolutionary patterns rather than their sole drivers, helping explain why organisms invest in apparently costly coordination mechanisms. These investments support the constraint regimes that enable both individual and collective persistence across unpredictable environmental changes, operating through the same fundamental logic that makes genes themselves so evolutionarily successful.

Multi-level selection theory addresses similar questions, arguing that natural selection can operate simultaneously at multiple organizational levels—genes, individuals, and groups[295]. Groups can evolve adaptive properties when they suppress internal conflict and develop mechanisms for collective reproduction. Cultural group selection extends this logic to human societies, where social learning allows group-beneficial traits to spread even when they disadvantage individuals.

The framework aligns closely with multi-level selection when systems achieve high distributed integration and organizational closure and function as organism. However, it suggests that patterns can become targets of selection well before reaching such integration, as soon as they invest sufficiently in maintaining their coordination. This includes overlapping networks, shifting alliances, and temporary coalitions, not only groups with clear boundaries. Rather than measuring success through group reproduction rates, the approach focuses on how effectively patterns persist and reconstitute themselves through disruption cycles.

Niche construction theory and the Extended Evolutionary Synthesis examine how organisms actively shape their environments[296]. These frameworks emphasize how organisms actively modify their environments and inherit not just genes but also epigenetic, ecological, and cultural resources. Niche construction theory specifically examines how environmental modifications by one generation affect the selective pressures faced by subsequent generations, creating feedback loops between organisms and their surroundings.

The constraint regime approach shares this recognition of multiple inheritance systems and organism-environment feedbacks, but decenters the analysis from focusing only on the organisms and its niche as the primary targets of selection. This shift provides a cleaner perspective because it doesn't privilege the organism-environment boundary, instead examining how patterns of organization persist across these boundaries through sustained maintenance work. This connects to insights from holobiont theory about organisms, microbiomes, and environments functioning as integrated units, but extends the logic further, as even partially integrated regimes can become targets of selection without achieving the full organizational closure of holobionts[297].

The framework makes a crucial distinction between one-off environmental modifications and sustained, energy-requiring maintenance of organizational constraints across these distributed systems. This maintenance-centered view helps explain why some niche construction efforts persist while others fade. Those that establish self-maintaining constraint cycles prove more evolutionarily durable than those requiring constant reinvestment without generating self-reinforcing dynamics. It also emphasizes perturbation testing, evaluating evolutionary success through reconstructability after disruption rather than performance under stable conditions, which reveals how systems that appear robust during normal operation may harbour hidden vulnerabilities when their maintenance requirements are disrupted.

The maintenance focus matters especially for understanding the major transitions in evolutionary history[298]. Step-changes where new levels of biological organization emerged, from genes to chromosomes, single cells to multicellular organisms, individuals to social groups. These transitions involve new forms of information integration and conflict suppression that create higher-level evolutionary units with their own reproduction and selection dynamics.

The distributed integration concept measures how close collective systems are to achieving the organizational closure that defines major transitions. The framework shows how transitions cross critical thresholds through energetic investments in enforcement mechanisms and memory systems that stabilize new organizational levels. However, rather than focusing only on completed transitions, it analyzes the full spectrum of collective organization, including high-integration systems that face selection pressure without achieving full organism-like reproduction. The approach also emphasizes how transitions that increase efficiency can create degradation dependencies—losing organizational flexibility in ways that may prove costly when conditions change. This provides a more dynamic view of major transitions as ongoing processes rather than completed achievements, helping explain both their remarkable stability once established and their occasional reversibility under extreme environmental stress.

Rather than replacing existing theories, this framework highlights the energetic foundations of evolutionary processes. It shows that gene-level effects, group coordination, environmental construction, and organizational transitions operate through shared logic: the maintenance of constraint regimes that enable pattern persistence through environmental variation. Making these maintenance costs explicit reveals both the remarkable stability and hidden vulnerabilities of complex systems, whether cellular metabolism, ecosystem coordination, or human institutions.

9.3 Scope & Limits

In the previous chapters we developed a framework that inverts how we conceptualize organization in living systems. Rather than defining systems through what they contain, we have shown how they persist through what they systematically exclude. This reorientation came through our analysis of scaling laws, the striking presence of gaps and discontinuities there, pointing towards constraints as central organizing principle behind these regimes. The subsequent examination of *distributed integration* foregrounded how living systems maintain these constraints through active energetic investment.

This duality of constraint regimes defining boundaries through exclusion, and distributed integration measuring the work invested in maintaining them provides a lens for understanding how complex wholes persist through cycles of growth, disruption, and reorganization. The slime mold's transitions between individual and collective states, the Inuit's seasonal transformations of social organization, and ecosystems' maintenance of chemical grammars all underscore how systems achieve forms of partial integration through ongoing investment in collective self-constraint.

The framework applies most directly to composite biotic systems that actively work to maintain their boundaries. From cells to biofilms, from ecosystems to institutions, these systems demonstrate varying degrees of distributed integration while sharing the fundamental challenge of maintaining organizational coherence against entropy. While abiotic regimes like atomic shells and molecular orbitals provide templates for the logic of exclusion, their stability requires no endogenous work, a crucial distinction that highlights the unique energetic demands of living organization.

Important limitations constrain this framework's application. Measurement presents persistent challenges. Quantifying hidden maintenance flows, distinguishing self-constraint from environmental forcing, and detecting degradation dependencies that remain invisible until perturbation. While the formal framework detailed in the Appendix offers a theoretical path to identifying these dependencies through causal metrics, the practical challenges of measurement in wild biological systems remain substantial. The framework's temporal sensitivity creates additional complexities since slow remember layers can appear inert, while fast revolt signals may register as mere noise without longitudinal data. These limitations do not invalidate the approach but define the conditions under which it provides meaningful insights.

The framework generates specific, testable predictions that ground its theoretical claims in empirical reality. Systems with higher pre-shock ratios of maintenance energy to growth energy should reorganize faster and retain greater diversity after disruption. Removing maintained alternative pathways should make systems more efficient under stable conditions but hasten their collapse under stress. Most crucially, the framework would be falsified by discovering complex biotic systems that persistently weather novel shocks without significant investment in collective self-constraint mechanisms.

These insights suggest practical implications for system design and management. Rather than maximizing efficiency, we should invest in the shared grammars, redundant pathways, and latent potentials that enable reorganization after disruption. This means deliberately maintaining alternative options even when they appear costly during stable periods. A strategy that applies equally to agricultural biodiversity, institutional knowledge, and technological infrastructure.

The stability of complex wholes, does not emerge directly through growth but requires continuous investment in collective forms of self-constraint. By focusing on what systems exclude rather than what they contain, and measuring the work they spend maintaining those exclusions, we can get insights into how organization persists from cells to larger composite whole.

Perhaps Descartes' famous cogito can be productively recast for living systems: I perform work to maintain my boundary, therefore I am. To be alive is to expend energy resisting equilibrium, to sustain selective boundaries that differentiate self from environment. The degree of this maintenance work, measured through distributed integration, determines not just whether a pattern persists, but what kind of *self* emerges from the process.

Chapter 10

Coda

10.1 Cognitive Constraints

> *The theory of thought is like painting: it needs that revolution which took art from representation to abstraction. This is the aim of a theory of thought without image.*
> *- Gilles Deleuze*[299]

Cognition doesn't start with identity, it begins with difference. We perceive boundaries before we name objects—the contrasts that carve reality into distinct entities. A thing becomes *something* only after it's differentiated from everything else. This primacy of difference is the bedrock of thought. Look at the text before you: it materializes not through some inherent *textness*, but through the interplay of black and white, through the characters emerging from a white background, through letters related to each other in a specific sequence.

Difference, by its nature, is relational. These perceived distinctions aren't isolated phenomena. They're always between things, simultaneously creating boundaries and bridges. The difference between this page and your hand is both a delineation and a point of contact—a zone of tactile exchange. From this interplay of differences arises a network of distinctions that extends from the physical into the abstract. Words on this page connect to those on the next, pages to cover, text to the corpus of human knowledge. Through these differential relationships, our world stratifies itself, organizing into the complex, nested patterns we navigate daily—a fractal landscape of meaning born from contrast.

Observe the mental cartography unfolding as you read this text. Watch how your mind maps out the terrain, distinguishing letter from word, sentence from paragraph, form from content. Notice how these categories aren't binary but spectral. A gradient of relationships, some elements residing closer on your cognitive map than others. Take the words *mental* and *cartography* in the first sentence of this paragraph. They are intertwined in meaning, creating a synergistic image. However, they're also distinct, with *mental* rooted in the language of cognition, and *cartography* stemming from map-making and navigation. Contrast this with the distant hum of traffic outside, or the abstract concept of democracy. Ideas far removed from this text, yet still accessible, still part of the network of your consciousness. This cognitive process of categorization, these acts of drawing distinctions and creating zones of meaning, echo the very processes we've observed in biological and ecological systems. Just as organisms cluster into distinct body-mass aggregations, our thoughts naturally organize into conceptual clusters, creating their own gaps and discontinuities.

The parallels between cognition and biology are not only analogical, they are genealogical. Cognition is a late-evolving expression of the dynamics that give cells their coherence[300]. So the neural and conceptual activity resonate with the fundamental logic of life, functionally linking what we are (ontology) with how we know (epistemology).

Seen from this vantage point, thought is structural coupling in action. Perception, categorization, and inference do not picture a pre-given world from afar, they are recursive adjustments by which an embodied organism and its surround co-determine one another. Every concept we stabilize, every distinction we hold, excludes alternatives so that a coherent pattern of activity can persist over time.

Cognitive regimes, whether the colour categories of a language or the taxonomic schemes of a culture, are structures that narrow the field of the possible just enough for coordinated thought and action to remain viable. These cognitive clusters emerge from and are maintained by active constraints on what can be meaningfully distinguished and related. This process exemplifies what Gregory Bateson identified as the fundamental nature of information itself, *a difference that makes a difference*[301]. A cognitive constraint regime, is not a passive record of the world. It is an expression of the very mechanism that actively selects *a difference that makes a difference*, shaping and constraining the ongoing structural coupling of organisms with themselves and their environment.

Language itself operates as a system of constraints that enables and limits possible interactions between speakers. These constraints create persistent structures that transcend individual minds while being continually shaped by collective usage. This structural stability is what allows Large Language Models to function. Despite lacking human-like neural architecture, they leverage language's stable constraint patterns to generate coherent responses. Not because they have developed consciousness, but because they've internalized the statistical regularities that emerge from language's constraint regime. The very fact that computational systems can capture these patterns suggests that language, operates through functional constraints that can be approximated by different architectural approaches.

This principle that constraints are generative structures, is not limited to one domain. It operates at multiple, nested scales. To foreground this we will briefly look at three examples. Different scales accentuate the fractal-like nature of this organizational logic, showing how the same fundamental process of constraint-formation shapes cognition across different contexts.

Consider first the specific distinctions embedded within language. Russian speakers, who linguistically distinguish light blue (голубой) from dark blue (синий) as separate basic color terms, demonstrate faster discrimination of these color boundaries than English speakers[302], showcasing how languages actively inform what distinctions become salient and stable.

There are also much broader cultural constraint regimes that operate at collective scales. The medieval Great Chain of Being exemplifies such a collective constraint regime, a vast, divinely-ordered hierarchical model that assigned every entity in the universe a fixed and immutable place, from God and the angels down to animals, plants, and minerals. It functioned by excluding certain conceptual possibilities. Entities could not change their status, hierarchical relationships were fixed rather than dynamic, and explanatory causation flowed only downward from higher to lower levels[303]. This systematic exclusion formed a stable cognitive scaffold that organized everything from theological doctrine, strictly stratified social interaction and biological classification, enabling coordinated thought and action across diverse domains for centuries. Its power lay not just in fixing entities in place, but in rendering unthinkable the very modes of explanation that rely on horizontal relations, dynamic transformation, or emergent causation.

The Great Chain eventually broke during the scientific revolution, illustrating how cultural constraint regimes transform. Internal tensions, contradictions between astronomical observations and geocentric assumptions, between fossil evidence and fixed species concepts, gradually exceeded the regime's exclusionary capacity. These intellectual pressures were compounded by socio-economic shifts. The rise of a merchant and industrialist class showed a form of upward mobility that challenged the God-given and stratified nature of the aristocratic and clerical order. The reorganization that followed reconfigured the fundamental constraints governing how knowledge could be produced, validated, and coordinated across communities. This transition accentuates how cultural constraint regimes, like biological ones, maintain themselves through active exclusion until accumulating tensions trigger adaptive reorganization.

At the individual scale, consider Piaget's classic conservation-of-liquid task. A young child, focusing only on the visual height of the water, believes that pouring it into a taller, thinner glass creates *more* liquid. Their cognition is constrained by a one-dimensional focus. The developmental leap to understanding conservation occurs with the construction of a new cognitive constraint. The ability to coordinate multiple variables (height and width) simultaneously. This new, more complex control structure doesn't erase the old perception, it reorganizes it, allowing for a more stable and powerful understanding of the world. Each developmental transition represents not the acquisition of new content but the emergence of new constraint regimes that reorganize existing cognitive resources into more powerful and flexible patterns[304].

To sum up, our thoughts cluster into stable patterns because cognition relies on an organizational logic related to one enables living systems to maintain coherence while remaining open to environmental coupling and adaptive transformation. Understanding this genealogy reframes how we approach theoretical frameworks. Rather than treating our models as neutral mirrors of reality, we begin to see them as living, embodied constraint regimes, dynamic structures that actively shape what becomes thinkable and actionable. The crucial questions therefore shift from asking whether a framework is a true mirror of reality, to asking: What does a specific framework exclude? What kinds of engagement with the world does it make possible? And when do its internal contradictions signal the need for reconstruction?

10.2 Modelling as Constraint selection

Shifting the spotlight onto the practice of modelling, I want to argue that we cannot see modelling as a representational act, that is, as an attempt to create a neutral mirror or faithful map of reality, rather, it should be viewed as the selection and stabilization of particular constraints among a vast potential.[305]. Fundamentally, we model because we need coherence, yet we exist as coherent beings precisely because we model.

In constructing models, we do not *observe* or *represent* reality. Instead, we choose how to constrain our investigative focus and interpretative strategies.

Constraints function as both enablers and limiters. They make certain insights possible while necessarily excluding others, like the patterns found in species clustering and body-size gaps. Our conceptual frameworks and models, whether the model of the *constraint cycles* or of the *cyclical synthesis*, are themselves constraints, shaping how we interpret and interact with the world around us.

By recognizing that our own theories and models are structured in this manner, we gain clarity about both, their power and their inherent limitations. This perspective of modelling as *constraint selection* opens avenues for understanding the inevitable partiality of knowledge, the significance of discontinuities, and the necessity of *scale-variant* thinking in science.

Any model begins with crucial decisions about system boundaries, determining which processes, scales, and feedback loops warrant inclusion and which can be omitted (or are not even known about). This unavoidable selection process is less about capturing reality in total, than it is about identifying constraints that will define an object of study. The very act of operationalizing variables imposes constraint-laden categories that highlight certain dynamic patterns while necessarily occluding others. Fore example, when we define metrics like a *body-mass index*, talk about *metabolic rates*, or *fitness*, one is not simply describing neutral features, but actively constructing frameworks that inform what we can observe and understand. Crucially, this is inevitably guided and shaped by the operative cognitive regimes of the scientists involved, their disciplinary training, theoretical biases, available metaphors, and even linguistic structures influence which constraints are deemed salient, relevant, or even conceivable.

Each modelling step thus becomes a deliberate negotiation of boundaries. We actively decide which constraints to impose so that certain regularities can emerge clearly. However, this selection always comes at a cost: aspects unaligned with our chosen constraints remain invisible or get suppressed. This duality of enabling and limiting characterizes all constraint selection. Crucially, as our discussion of Distributed Integration implies, maintaining such a selection is an active process; the system must continuously expend energy to keep these specific boundaries from eroding back into ambiguity.

This perspective counters the notion that models should strive for universality. Instead, success lies in matchmaking constraints to a particular scale, domain, or phenomenon, much like how each species finds its niche in an ecosystem of possible body masses and resource availabilities[306].

10.3 The Practice of Partiality

If modelling is the selection of constraints, why does the fantasy of a single, comprehensive model endure? Partly because a universal calms us. It promises rest from the work of coordination. It also simplifies trust—we can say *this is true*, full stop, without time-stamps, caveats, or scales. But mostly, universals spare us the tension that cyclical synthesis asks us to hold. What is generative in one phase becomes obstructive in the next. What is true now, may no longer be so later. Others in this room may hold views that actively differ from ours and yet are required for the whole to remain viable.

The cost of universals is invisibly high, hiding the boundary decisions that made them persuasive and rendering the losers of the constraint-selection invisible. They tempt us to mistake simplicity for truth. Yet the desire is not to be scolded. It is a longing for coherence that can be met, not via the eliminating difference, but by learning to compose it in time. Cyclical synthesis is not anti-truth. It is fidelity to the kind of truth that living systems can carry. Truth whose coherence is maintained by sequencing opposites rather than erasing them.

Partiality exacts its own price. It carries a cognitive cost (tension between incommensurable frames), a social cost (relating to the *other*), and an institutional cost (coordination differences). Partiality often slows decision-making, complicates accountability, and exposes us to misinterpretation. But these frictions are the signature of a living system, the metabolic heat generated by the difficult work of maintaining a world large enough for genuine difference.

What follows is the practical consequence. No single framework or universal model can encapsulate the complexities of dynamic systems that operates across multiple levels[307]. Each level maintains its own constraints, and the way knowledge is structured at one level may be relatively inapplicable at another. If we attempt to flatten these differences into a single explanatory schema, we risk obscuring the very discontinuities that give rise to levels of organization in the first place. A constraint regime that defines interactions at the level of international economies does not simply scale down to describe the behaviour of individual humans, just as the principles governing macroscopic physics do not translate seamlessly into the quantum realm. The boundaries between levels are not arbitrary, they are points at which different sets of constraints dominate, creating qualitatively distinct modes of organization.

In other words, the conceptual frameworks we develop function as constraint regimes in their own right, stabilizing certain patterns of thought while closing off others. The language of *cyclical synthesis*, *constraint cycles* and *regimes* for instance, has enabled us to see the interdependencies between *self-organization* and *self-constraint* in living systems, but it also necessarily directs our focus in particular ways, reinforcing certain regularities while bracketing the always messy complexities of how differently these processes play out.

This recursive dynamic is an essential feature of how knowledge evolves. A theory remains useful not because it absolutely *represents* reality, but because it effectively constrains our thinking in ways that make certain patterns intelligible. Instead of seeking a single, overarching framework, we need to recognize the necessity of different constraint regimes for different domains of inquiry. Like living systems operating at multiple nested levels, each with its own constraints, models must be tailored to the scale and nature of the systems they describe. The logic of *cyclical synthesis* we applied to biological organization is equally relevant to epistemology: knowledge itself is maintained through the interplay of enabling and limiting constraints, with each conceptual cycle reinforcing some patterns while allowing periodic transformations.

Understanding modelling as constraint selection reshapes how we approach it, forcing us to accept the partiality of knowledge, not as a failure, but as an intrinsic feature of structured inquiry. It reorients how we meet anomalies and theoretical gaps, not as problems to be eliminated, but as markers of discontinuities where existing constraint regimes no longer hold. Such discontinuities are invitations to reframe the problem, to reconsider the boundaries that have structured our thinking thus far. Modelling becomes both a *practice*, in an ongoing process of constraint negotiation, and the model itself *a tool*, a means for us to interact with specific aspects of our environment.

Thus, we arrive at the perspective of modelling as an act of co-creation, where the constraints we impose foreground patterns of reality while simultaneously shaping the very vantage points from which we perceive and act. To model is to cultivate a conceptual landscape, imposing limits that make knowledge possible while knowing that those limits themselves will, in time, be revised[308]. The structure of thought relates to the structure of life, a dynamic interplay of growth, disruption, reorganization, where constraints, once selected, both enables and limit the possibilities of what comes next.

The dream of a frictionless order—everything aligned, everything at once—is always the dream of tyranny. It is the fantasy of erasing what does not fit. Life has another grammar. It holds together by letting tensions live, by organising them across time. To accept this is costly, we yield the comfort of a single truth, but it is also the only way minds stay open and societies remain free. The practice of partiality is nothing less than the practice of keeping a world alive.

Appendix

Overview

The chapter *Degrees* developed Distributed Integration (DI) as a framework for understanding how living systems maintain organizational coherence through energetic investment in collective constraints. This appendix formalizes that intuitive framework using tools from Causal Emergence 2.0 (CE), a mathematical method for quantifying how causal power distributes across a system's hierarchy of scales[309].

CE was developed as a domain-general formalism applicable to any system describable as a Markov chain, from computer architectures to neural networks. It is adopted here because it provides mathematical precision for concepts DI articulates energetically: Where DI measures the costly work organisms invest in maintaining collective constraints ($E_{\text{self-constraint}}/E_{\text{total}}$), CE measures the structural consequences: non-redundant causal contribution (ΔCP) at each scale, the distribution of causal strength across levels (S_{path}), and differentiation within levels ($\overline{S}_{\text{row}}$).

The appendix proceeds in three steps. First, CE core concepts and measurement framework are introduced, focusing on how emergent hierarchies form through coarse-graining. Second, linking propositions connecting DI dynamics to shifts in causal architecture are articulated, namely, how investment in coordination raises ΔCP at integrative scales, how efficiency-driven streamlining creates hierarchical brittleness, and how maintained alternatives preserve reorganization capacity under perturbation. Third, methodological limitations, particularly the gap between CE's intervention-based ideal and biological reality are discussed.

This appendix is exploratory and programmatic rather than methodologically complete. CE is presented as one possible formalization of DI concepts, recognizing that substantial operationalization challenges remain.

Introduction to Causal Emergence 2.0

Every complex system contains not one description but many. A computer operates simultaneously at the scale of transistor circuits, machine code, and operating system functions. The brain can be described via molecules, neurons, cortical columns, or psychological states.

This multiplicity creates what Hoel & Jansma call the *paradox of scale*. For any given system, there exists an astronomically large set of possible scales of description. For a system with just 8 discrete states, the number of possible partitions (ways of grouping states into coarser categories) is 4140. These partitions form a lattice structure, ordered from the microscale (finest-grained description) at the bottom to the macroscale (coarsest-grained) at the top, depicted in Figure 1. The number of possible partitions grows as the n^{th} Bell number, making exhaustive analysis computationally prohibitive for large systems.

The paradox is that either all scales are equally valid (relativism, which cannot explain why some descriptions work better than others), or only the microscale truly matters (reductionism, rendering most of science epiphenomenal), or we need a principled way to identify which scales *genuinely contribute* to a system's causal workings. CE pursues the third option.

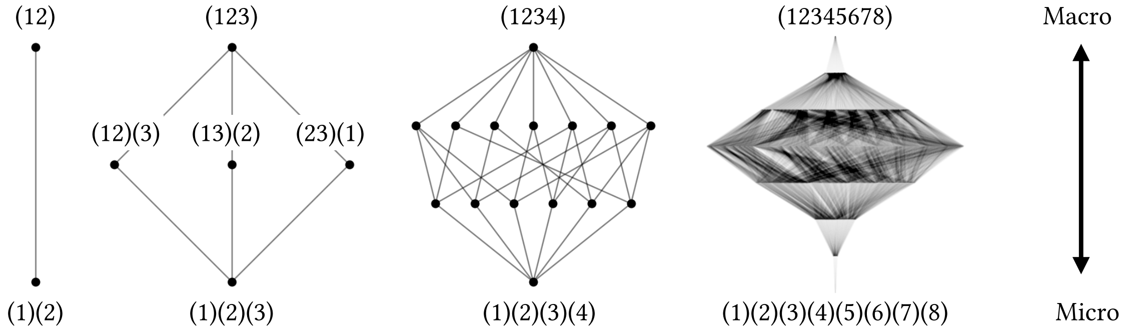

Figure 1: Partition lattices showing all possible scales of description for systems with 2, 3, 4, and 8 states (containing 2, 5, 15, and 4,140 partitions respectively). Each partition represents a coarse-graining from microscale (bottom) to macroscale (top). Reproduced from Jansma and Hoel [2025].

To make the analysis tractable, CE treats systems as Markov chains, collections of discrete states with specified transition probabilities. Think of a board game where your current position (state) determines the probability of moving to each possible next position. Formally, a Transition Probability Matrix (TPM) T captures these dynamics: $T_{ce} = p(\text{effect } e|\text{cause } c)$.

Critically, this framework treats states not merely as observations but as *causal relationships* verified through interventions. Using Pearl's do-operator notation, one can reach into the system and set it to state c (written $\text{do}(c)$), then observe which effects e occur[310]. This intervention-based approach distinguishes genuine causation from mere correlation.

Measuring Causal Strength: Determinism and Degeneracy

How do we score a scale's causal contribution? Hoel & Jansma identify two information-theoretic primitives:

Determinism measures how reliably causes produce effects. Mathematically:

$$\text{determinism}_T = 1 - \frac{H(E|C)}{\log_2 n} \tag{1}$$

where $H(E|C)$ is the conditional entropy of effects given causes, and n is the number of states. Determinism is maximal ($= 1$) when each cause leads to a single, certain effect; minimal ($= 0$) when transitions are uniformly random.

Degeneracy measures how much effects overlap—whether multiple causes lead to the same effects. Its inverse, specificity, captures how unique each cause's effects are:

$$\text{specificity}_T = 1 - \text{degeneracy}_T = \frac{H(E)}{\log_2 n} \tag{2}$$

The combined score, CP (causal primitives), sums them:

$$\text{CP} = \text{determinism} + \text{specificity} - 1 \tag{3}$$

CP ranges from 0 (random, noisy transitions) to 1 (perfectly deterministic transitions to unique effects). Intuitively, CP is high when causes are *sufficient* for their effects (high determinism) and *necessary* (low degeneracy/high specificity).

These primitives connect to classical notions of probabilistic causation. High determinism captures sufficiency: given cause c, effect e reliably follows. Low degeneracy captures necessity: effect e requires cause c (few alternative causes exist).

Coarse-Graining

To see how coarse-graining changes causal quality, consider the six-state system in Fig. 2. The microscale transition-probability matrix (center) encodes the chance of moving from each state at time t to each state at t+1 (darker cells/arrows indicate higher probability). From this same microscale we can construct different macroscales by grouping microstates into macrostates.

The two panels at the top show *two* three-state coarse-grainings of this very system:

(i) **Partition** $\pi_\ell = (15)(34)(26)$**.** This grouping mixes microstates that do not form a stable dynamical unit. The resulting macro-TPM exhibits smeared transition structure across macrostates, i.e., lower determinism and higher overlap of effects. Intuitively, the coarse-graining averages together behaviors that should remain distinct, producing a noisy, low-quality macroscale.

(ii) **Partition** $\pi_r = (13)(24)(56)$**.** This grouping aligns with the system's natural modules (as seen in the graph at the bottom): each macrostate collapses a pair of microstates that behave similarly and largely cycle among themselves. The macro-TPM becomes near-permutation (clear self-loops, minimal cross-talk), i.e., higher determinism and lower overlap. This coarse-graining *improves* causal quality by removing nuisance variation while preserving the dynamics that matter.

Crucially, both macroscales have the *same dimensionality* (three macrostates), yet only one increases causal quality. In the next subsection, we make this precise by asking whether a coarse-graining *adds anything* beyond what finer scales already provide. We quantify that with the non-redundant gain ΔCP and use it to build the emergent hierarchy.

The key innovation of CE 2.0 is systematic comparison across *all* possible scales. For each partition π, CE 2.0 calculates its CP score and the maximum CP of any scale *below* it in the lattice (i.e., any partition π' that refines π). The non-redundant contribution is:

$$\Delta \mathrm{CP}_\pi = \mathrm{CP}_\pi - \max_{\pi' < \pi}(\mathrm{CP}_{\pi'}) \tag{4}$$

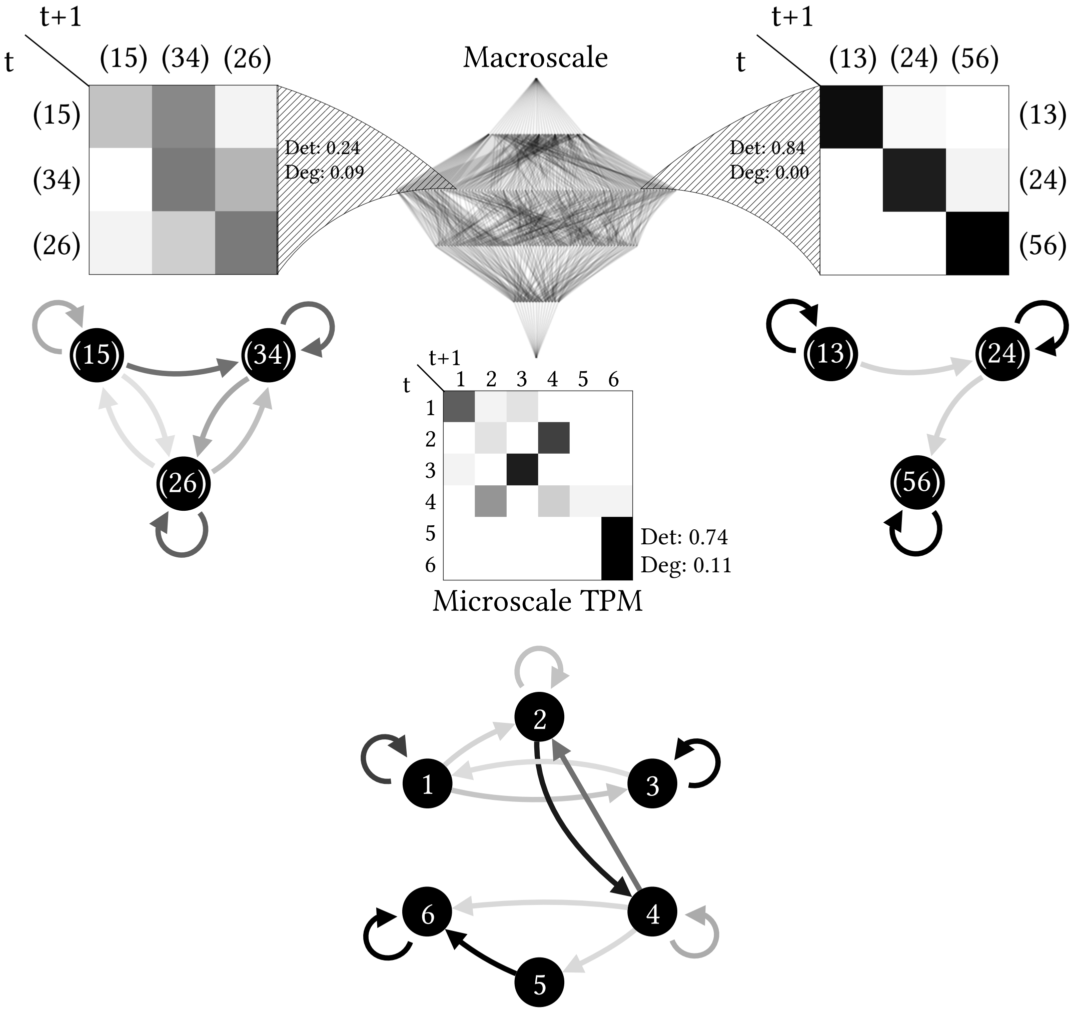

Figure 2: **Coarse-graining the same system in two ways.** Bottom: microscale graph of a six-state system (arrow thickness/gray encodes transition probability). Center: microscale TPM (darker = higher $p(e \mid c)$). Top left: the partition $\pi_\ell = (15)(34)(26)$ mixes unrelated states, yielding a noisy macroscale with smeared transitions. Top right: the partition $\pi_r = (13)(24)(56)$ respects dynamical modules, producing a near-permutation macro-TPM with clear, reliable transitions. Both macroscales have the same dimensionality, yet only the right coarse-graining captures stronger causal organization. Reproduced from Jansma and Hoel [2025].

Partitions with $\Delta\text{CP} > 0$ contribute causally in a way not reducible to lower scales. The collection of all such scales forms the hierarchy. Hoel & Jansma describe a taxonomy of emergence patterns which is where things get relevant for us:

1. **Top-heavy systems**: Most ΔCP concentrated at coarse scales (few large macrostates). These systems exhibit strong collective organization where high-level descriptions dominate.
2. **Bottom-heavy systems**: Most ΔCP concentrated at fine scales (many small components). Causal power remains primarily at the microscale.
3. **Mesoscale-peaked systems**: ΔCP peaks at intermediate levels. Neither micro nor macro dominates; middle-scale modules drive behavior.
4. **Complex (scale-free) systems**: ΔCP distributed evenly across levels. Causation operates simultaneously at many scales with no single level dominating.

The diversity of these hierarchical shapes can be reproduced by varying how systems grow. Hoel & Jansma simulated networks whose connectivity follows a preferential-attachment rule, where the probability that a new node links to an existing one is proportional to that node's degree raised to an exponent α. By tuning α, they generated structures ranging from sparse, tree-like networks to dense, star-like hubs[311]. Figure 3 shows how this single parameter reshapes the distribution of causal power across levels: bottom-heavy hierarchies emerge for sublinear growth ($\alpha < 1$), complex multiscale hierarchies peak near the scale-free regime ($\alpha \approx 1$), and top-heavy patterns appear under superlinear attachment ($\alpha > 1$).

Two measures, indicated in the figure above capture the *shape* of emergence. First, Path entropy (S_{path}), asking how uniformly is ΔCP distributed across levels?

For a path through the lattice from microscale to macroscale, normalize the ΔCP values to form a probability distribution $\{p_1, \ldots, p_L\}$ where:

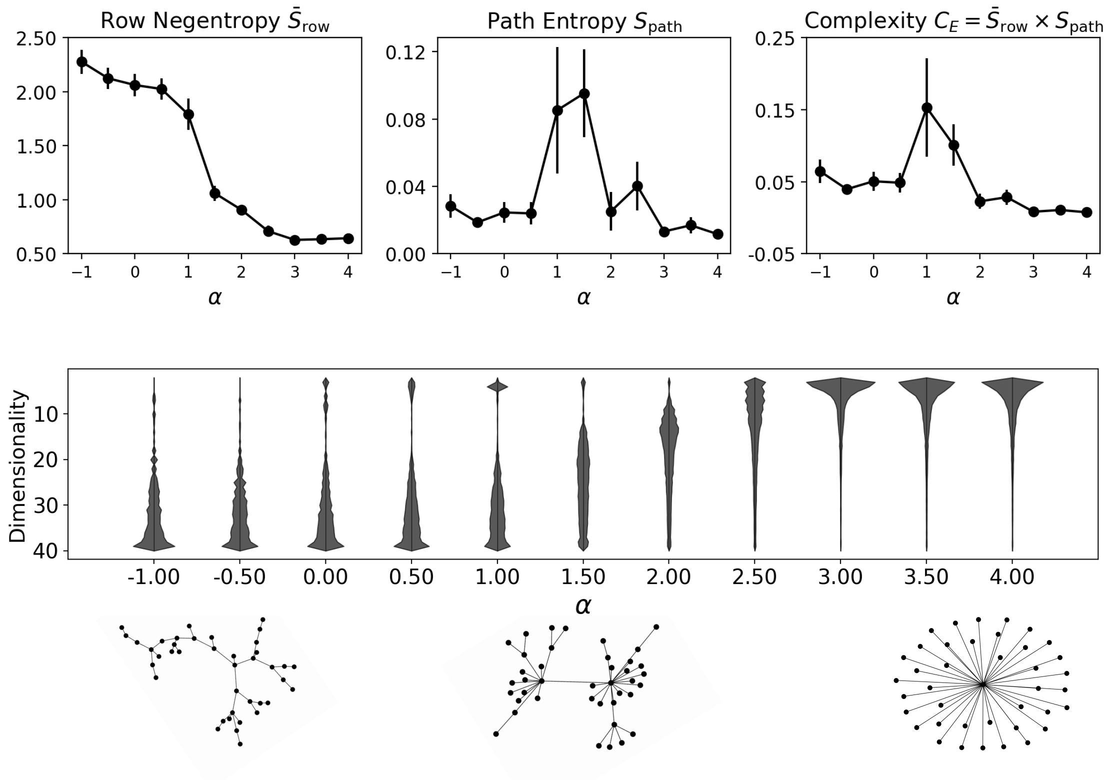

Figure 3: **Emergent hierarchy shapes under preferential attachment.** Varying the attachment exponent α during network growth alters how causal contributions distribute across scales. Top: average values of row negentropy $\overline{S}_{\text{row}}$, path entropy S_{path}, and their product $C_E = \overline{S}_{\text{row}} \times S_{\text{path}}$ (a measure of emergent complexity) across networks of 40 nodes. Complexity peaks near $\alpha \approx 1$, the regime of classical scale-free networks. Middle: distributions of dimensionality (number of effective levels) for each α. Bottom: example networks generated at representative α values. Sub-linear growth ($\alpha < 1$) yields branching, bottom-heavy hierarchies; the scale-free regime ($\alpha \approx 1$) produces rich, multilevel structure; super-linear growth ($\alpha > 1$) collapses into top-heavy, hub-dominated networks. Reproduced from Jansma and Hoel [2025].

$$p_i = \frac{\Delta\mathrm{CP}_i}{\sum_{j=1}^{L} \Delta\mathrm{CP}_j} \tag{5}$$

The path entropy is:

$$H(p) = -\sum_{i=1}^{L} p_i \log_2(p_i) \tag{6}$$

Averaging over all paths through the emergent hierarchy yields S_{path}. High S_{path} means causation spreads across many scales.

Second, row negentropy ($\overline{S}_{\text{row}}$), asking, within each level, how differentiated are the useful coarse-grainings? At each level ℓ with multiple partitions, normalize their ΔCP values to form a distribution, calculate its entropy S_ℓ, then take the negentropy:

$$\overline{S}_{\text{row}} = \log_2 L - \frac{1}{L}\sum_{\ell} S_\ell \tag{7}$$

High $\overline{S}_{\text{row}}$ means many distinct, valuable macroscales coexist at the same dimensionality, heterogeneity within levels rather than convergence to a single "correct" coarse-graining.

Systems with high S_{path} and high $\overline{S}_{\text{row}}$ exemplify *scale-freeness*: no single scale dominates causal organization. Remarkably, Hoel & Jansma show this corresponds to the regime of classical scale-freeness in network theory (power-law degree distributions around preferential attachment parameter $\alpha = 1$), suggesting deep connections between topological and causal notions of complexity.

Distributed Integration & Causal Emergence 2.0

DI and CP are not equivalent but complementary. DI quantifies the *work* devoted to maintaining collective self-constraint while CP quantifies the *structural causal strength*. In living systems, high DI targeted at collective constraints tends to produce high CP at macroscales that black-box those constraints, however this relationship is neither necessary nor sufficient in all cases.

Hierarchies Over Time

Let $\Delta\text{CP}_\ell(t)$ denote the **total** non-redundant contribution at level ℓ at time t (the sum of ΔCP for all partitions of dimensionality ℓ). With this, we define the level weights $w_\ell(t)$ and the **Hierarchy Centroid** $L_{\text{com}}(t)$ (Level Center of Mass):

$$w_\ell(t) = \frac{\Delta\text{CP}_\ell(t)}{\sum_k \Delta\text{CP}_k(t)}, \qquad L_{\text{com}}(t) = \sum_\ell \ell \cdot w_\ell(t) \tag{8}$$

This centroid tracks *where* in the hierarchy causal power concentrates. For example, $L_{\text{com}} = 2$ in an 8-level hierarchy indicates causation remains bottom-heavy, concentrated near microscales, while $L_{\text{com}} = 7$ signals top-heavy organization where coarse-grained macroscales dominate the system's causal workings.

The time-dependence of these quantities proves crucial for connecting CE to DI. Living systems don't simply possess static emergent hierarchies, they actively construct, maintain, and sometimes dismantle them through energetic investment. When organisms increase distributed integration by channeling energy into collective constraint maintenance, this work should manifest as measurable shifts in causal architecture. Specifically, we expect activation of latent potential or sustained increases in DI to produce three characteristic signatures in the emergent hierarchy.

First, the hierarchy centroid should shift upward. $L_{\text{com}}(t)$ increases as causal power migrates toward coarser scales that integrate the coordinated components. Second, path entropy often increases, $S_{\text{path}}(t)$ rises as causal power spreads more evenly across multiple levels rather than concentrating at a single scale. This reflects the system developing genuine multiscale organization where micro, meso, and macro levels all contribute meaningfully. Third, row negentropy sometimes increases, $\overline{S}_{\text{row}}(t)$ grows when the system maintains multiple distinct useful coarse-grainings at each level, preserving organizational flexibility through redundant pathways.

These expectations are tendencies rather than invariant laws. The relationship between energetic investment and causal structure depends on how work is allocated—whether it targets coordination at specific scales, whether it maintains alternatives or streamlines toward efficiency, and whether organizational context makes collective constraints appropriate for present challenges. The linking propositions that follow articulate these relationships more precisely, specifying mechanisms and predictions that connect DI dynamics to observable shifts in emergent hierarchies.

Linking Propositions

The DI→CE Hypothesis

The first linking proposition establishes a direct relationship between energetic investment and organizational structure. When living systems increase their distributed integration by targeting a coalition of components through signaling, policing, or boundary maintenance, this investment should manifest as increased causal power at the scales that integrate those components. Specifically, we expect to see higher ΔCP at integrative scales and an upward shift in the hierarchy centroid $L_{\text{com}}(t)$.

When organisms invest energy in coordination work, they reduce the conditional entropy $H(E|C)$ by limiting the range of possible effects that follow from any given cause. Simultaneously, this increases the entropy $H(E)$ across the full state space by generating more distinctive outcomes. Together, these changes raise both determinism (causes more reliably produce their effects) and specificity (effects become more unique to their causes).

To test this hypothesis empirically, one would plot the hierarchy centroid $L_{\text{com}}(t)$ along with path entropy $S_{\text{path}}(t)$ and row negentropy $\overline{S}_{\text{row}}(t)$ before and after an intervention that increases distributed integration. The prediction is straightforward: monotonic increases in these causal measures should correlate with increases in ΔDI. The null case proves equally informative. If distributed integration increases without corresponding gains in causal power, this indicates misallocation—energy flowing into activities that fail to reduce uncertainty or overlap in causal relationships.

The Efficiency-Brittleness Hypothesis

Selection for efficiency often reduces the distributed integration devoted to seemingly redundant features: backup pathways, partner diversity, or alternative coordination mechanisms. This second hypothesis predicts that such efficiency-driven streamlining produces a characteristic signature in the emergent hierarchy. Specifically, we expect **increased** row negentropy $\overline{S}_{\text{row}}$ (fewer distinct useful macroscales), increased top-heaviness (the centroid L_{com} shifts upward while path entropy S_{path} compresses), and the emergence of degradation dependencies that become visible as rapid collapse of causal power under perturbation.

The underlying mechanism reflects a fundamental trade-off. Maintaining alternative pathways requires ongoing energetic work, contributing to higher distributed integration. These alternatives provide multiple pathways through the state space, preserving organizational flexibility. When efficiency pressures eliminate these alternatives, the system saves energy in the short term but reconfigures into what might be called a single-path topology. States increasingly funnel toward a single dominant macrostate, and the emergent hierarchy becomes a spike: high ΔCP concentrated at the ultimate macroscale, with little causal contribution elsewhere.

The coral-zooxanthellae system offers a natural test case. Under thermal stress, we predict that efficiency-optimized pairings characterized by high specificity and **high** $\overline{S}_{\text{row}}$ (brittle concentration) should exhibit greater losses in hierarchical complexity compared to more flexible partnerships that maintain multiple symbiont strains and thus **lower** $\overline{S}_{\text{row}}$ (higher entropy). The observable signature becomes clear: pre-perturbation row negentropy should predict post-perturbation recovery of hierarchical structure, measured through the restoration of S_{path} and $\overline{S}_{\text{row}}$.

The Resilience Hypothesis

The third linking proposition connects pre-shock distributed integration to post-shock reorganization capacity. Systems with higher distributed integration before a perturbation should show both smaller drops and faster rebounds in $\sum \Delta\text{CP}$ and $L_{\text{com}}(t)$ after disruption. The explanation lies in how maintained alternatives preserve multiple high-CP pathways through the state space, providing the system with organizational options when its primary configuration becomes untenable.

Bacterial biofilms provide an ideal system for testing this hypothesis. One would measure pre-antibiotic distributed integration as the fraction of metabolic budget allocated to extracellular polymeric substances (EPS) and quorum sensing signals. High-DI biofilms should then demonstrate three characteristic responses during and after antibiotic exposure: maintained hierarchical structure as measured by S_{path} during the stress period, faster return to baseline hierarchy centroid post-treatment, and preservation of mesoscale peaks (intermediate ΔCP at the module level).

Connections to Panarchy Theory

The Panarchy model developed by Gunderson and Holling describes cross-scale interactions in complex systems through two key dynamics: revolt, where fast scales disrupt slow ones, and remember, where slow scales stabilize fast ones. These concepts map naturally onto CE dynamics, providing another avenue for linking distributed integration to causal emergence.

Revolt manifests when a localized pulse of high distributed integration at fine scales increases ΔCP at those scales. This change cascades upward as macroscale transition probability matrices absorb the perturbation, producing an upward shift in $L_{\text{com}}(t)$ and a temporary spike in path entropy S_{path}. The system's causal organization temporarily spreads across more scales as the fine-scale disruption propagates through the hierarchy.

Remember operates through a complementary mechanism. Slow-scale distributed integration maintains memory scaffolds that preserve persistent high ΔCP at coarse levels. When disruption occurs, these maintained coarse-scale structures guide the reconstruction of fine-scale organization post-shock, enabling rapid recovery of $L_{\text{com}}(t)$ to its pre-shock baseline. The system's investment in maintaining these slow-scale patterns pays off precisely when fast-scale organization must be rebuilt.

Slime mold aggregation offers a concrete prediction. Blocking cyclic AMP signaling eliminates fast-scale distributed integration. Under this intervention, we predict that L_{com} fails to rise from the microscale—aggregation cannot proceed because the fine-scale coordination mechanism has been disabled. More generally, in recovering systems, high pre-shock ΔCP at coarse scales (Panarchy's remember function) should predict shorter reconstruction time for fine-scale ΔCP, as the preserved macroscale organization provides a template for rebuilding.

These dynamics suggest that the value of distributed integration may lie not in maintaining a static hierarchical configuration, but in enabling flexible oscillation between scales. Different contexts demand causal organization at different levels—fine-scale coordination for rapid local response, macroscale integration for collective challenges. Systems with high distributed integration maintain the capacity to shift $L_{\text{com}}(t)$ contextually, keeping the entire hierarchy causally relevant across varying timescales rather than optimizing for a single, fixed organizational level. This temporal flexibility, rather than any particular centroid position, may constitute the deepest advantage of sustained investment in collective self-constraint.

Integration Efficiency: A Joint Metric

To assess how effectively energetic investments translate into causal organization, we define integration efficiency over a temporal window $[t_1, t_2]$:

$$\eta = \frac{\sum_{\ell} \max(0, \Delta\text{CP}_{\ell}(t_2) - \Delta\text{CP}_{\ell}(t_1))}{\text{DI}(t_2) - \text{DI}(t_1)} \tag{9}$$

This ratio captures the return on investment for coordination work, revealing how much *new* causal organization is generated per unit of metabolic investment.

The metric's interpretation depends on its magnitude. When $\eta \gg 1$, we observe efficient constraint-building: small increases in distributed integration yield disproportionately large gains in causal power. This pattern appears in signaling cascades with amplification, where modest energetic investments in signal production create substantial downstream coordination. When $\eta \approx 1$, the relationship becomes linear—work translates proportionally into organization without amplification or inefficiency.

Values of $\eta \ll 1$ indicate inefficient or misallocated work. Distributed integration rises but causal power does not, suggesting energy spent on signals that fail to reduce uncertainty or coordination mechanisms that don't actually constrain the system's dynamics. The most pathological case occurs when $\eta < 0$: rising distributed integration paradoxically reduces causal power. This can happen in over-constrained systems where additional coordination creates noise rather than organization, or where competing coordination mechanisms interfere with each other.

However, interpreting integration efficiency requires careful attention to context and appropriate organizational scales. Consider baboon grooming behavior, where females invest approximately DI $\approx$ 0.05–0.15 of their time budget in social maintenance. An increase in grooming time by ΔDI $= 0.05$ might yield $\eta > 1$ when the troop faces inter-group conflict (where collective cohesion matters), but $\eta \ll 1$ during resource-abundant periods when individuals benefit more from independent foraging. The efficiency metric thus reveals not whether distributed integration is universally beneficial, but rather how well energetic investments match current organizational demands. High values of η indicate that coordination work is appropriately targeted—generating causal organization at scales that matter for present challenges. Low or negative values suggest misallocation, where energy flows into maintaining constraints that either fail to reduce uncertainty at any scale or actively constrain the system at scales irrelevant to its current context. This context-dependence reinforces that distributed integration's value lies in enabling flexible hierarchical reorganization rather than maximizing causal power at any single scale.

Methodological Considerations and Synthesis

Using CE as Formal Infrastructure

Causal Emergence 2.0 provides a mathematical framework for quantifying how causation distributes across a system's hierarchy of scales. The framework was developed as a domain-general tool, applicable to any system describable as a Markov chain, from computer architectures to neural networks to social dynamics. It makes no special commitments to biological organization. We adopt it here not because it was designed for living systems, but because its formalism offers precision for concepts that DI articulates more intuitively.

Where DI describes the work organisms invest in maintaining collective constraints, CE provides tools to measure the structural consequences of that work: the non-redundant causal contribution ΔCP at each scale, the distribution of causal strength across levels S_{path}, and the differentiation within levels $\overline{S}_{\text{row}}$. The framework quantifies organizational morphology—whether systems exhibit top-heavy, bottom-heavy, mesoscale-peaked, or complex hierarchical structure, giving mathematical expression to patterns we observe in biological systems. The hierarchy centroid $L_{\text{com}}(t)$ tracks where causal power concentrates, while path entropy and row negentropy capture how that power distributes and differentiates across the emergent hierarchy.

However, CE remains fundamentally static. It analyzes snapshots of causal structure rather than explaining how that structure came to be or how it changes. The framework tells us where causation operates within a given system configuration but says nothing about the energetic flows that construct or maintain those configurations. This is where DI becomes necessary, not as a complement to CE, but as the biological reality that CE can help formalize. Living systems actively build their hierarchies through sustained energetic work. The slime mold doesn't simply have an emergent hierarchy, it constructs one through cAMP investment during aggregation, maintains it through coordinated movement, then dismantles it during sporulation. CE's mathematics can track these transitions as shifts in $L_{\text{com}}(t)$ and changes in the emergent hierarchy, but the explanation for why these shifts occur comes from tracking distributed integration.

Fundamental Limitations

Several constraints limit how we apply this formalization to living systems. The most fundamental concerns CE's intervention-based causal analysis. The framework assesses causation through Pearl's do-operator, which works by determining a system's causal structure by asking what would happen under direct interventions that set the system to specific states. For living organisms, this creates profound practical problems. We cannot literally intervene to place a baboon troop in a precise social configuration or force a biofilm into a specific metabolic state with experimental control comparable to manipulating digital systems. Living systems resist such manipulation, and forcing them into unnatural states may fundamentally alter the causal relationships we aim to measure.

This limitation doesn't invalidate the theoretical value of CE's formalism. The mathematics provides conceptual clarity that helps us reason about multi-scale organization, even when complete experimental validation remains impossible. We can construct transition probability matrices from observational data, tracking how systems naturally move between states, then analyse their emergent hierarchies. Such analyses reveal organizational structure relevant for understanding persistence and adaptation. But we must recognize the gap between theoretical precision and experimental accessibility. The formalization disciplines our intuitions about how distributed work creates organizational coherence without requiring that every mathematical quantity become empirically measurable.

State space construction introduces additional challenges. Building TPMs requires discretizing continuous biological variables into discrete states, and these boundary decisions affect all subsequent analysis. For bacterial biofilms, we might define states by cell density, molecular markers, or spatial architecture where each choice yields different transition matrices and different emergent hierarchies. The most principled approach aligns state definitions with boundaries that organisms themselves maintain through active regulation: quorum sensing thresholds, developmental switches, social rank transitions. Where organisms invest significant energy to maintain sharp transitions, these should correspond to state boundaries in our analysis and to high-ΔCP partitions in the emergent hierarchy.

Computational tractability limits the systems we can fully analyse. Exhaustive CE calculation scales with the Bell number of states, a super-exponential explosion making analysis beyond roughly ten states prohibitive. This mathematical barrier means we cannot apply the full framework to complex systems without aggressive simplification. The framework works best for clearly bounded subsystems where relevant states remain tractable.

CE also assumes Markovian dynamics where a system's next state depends only on its current state, ignoring memory and path-dependence that characterize biological processes. While we can partially address this by expanding state definitions to include recent history, this increases state space size and may not capture all relevant memory effects. These assumptions mean the framework provides greater insight within individual systems under different conditions than across fundamentally different systems with incomparable state space structures.

Toward Formalized Understanding

Despite these limitations, CE's formalism helps express what DI articulates about living organization. The slime mold's metamorphosis becomes a trajectory where the hierarchy centroid rises from microscale (independent amoebae, DI near zero) to mesoscale organization (coordinated slug, DI approximately 0.3) with peak ΔCP at the partition grouping cells into functional units. Inuit seasonal cycles appear as oscillations in causal organization where summer's minimal collective coordination yields family-level dominance, while winter's intensive communal maintenance shifts causal power to collective scales. Degradation dependencies manifest as hierarchical brittleness. Coral systems with low $\overline{S}_{\text{row}}$ exhibit sharp CP collapse under thermal stress, while populations maintaining higher row negentropy through partner diversity show graceful degradation.

The key insight recognizes that organizational levels are achieved through work rather than simply discovered through analysis. The emergent hierarchy reflects how organisms allocate energy toward constraints, actively constructing their causal architecture. Multicellularity represents cells investing DI in adhesion and signaling that creates high-CP organization at tissue scales. Social insect colonies achieve their distinctive organization through investments in communication and policing that create mesoscale ΔCP peaks. Human institutions emerge when groups invest DI in memory scaffolds and enforcement mechanisms that shift causal power to macro-social scales.

The framework equally reveals organizational failures. Systems can waste work through coordination mechanisms that don't reduce uncertainty, over-integrate into brittle monocultures, under-integrate and fail to achieve collective capabilities, or mistime their investments relative to environmental change. The integration efficiency metric η (defined above) captures whether energetic investments translate into causal organization, though measuring both quantities rigorously remains challenging.

The question both frameworks address cuts to what makes collective organization scientifically real, do organization becomes real when work invested in constraints generates non-redundant causal contributions at scales beyond the microscale. The emergent hierarchy serves as formalized evidence that collective levels accomplish something irreducible to component parts. When causal power distributes across many scales with high path entropy and row negentropy, we observe literal scale-freeness—systems where no single level dominates, achieved through persistent work against entropy's pull toward the microscale.

Life, viewed through this lens, becomes the art of sustaining emergent hierarchies through continuous energetic investment. The study of life correspondingly examines how work, channelled through constraint maintenance, sculpts causation's landscape across scales—concentrating causal power at some organizational levels while leaving others causally inert, creating a topography that shows where systems operate as agents in their worlds.

Glossary

This glossary provides definitions for key technical terms, specialized concepts, and concepts introduced in this thesis. The definitions are tailored to their specific usage within the text.

Adaptive Cycle A model, developed by C.S. Holling and Lance Gunderson, describing how complex adaptive systems persist and evolve through a recurring four-phase cycle: exploitation (r-phase), conservation (K-phase), release (Ω-phase), and reorganization (α-phase). The cycle illustrates how systems manage the inherent tension between opposing tendencies like growth and stability versus change and variety. (*See also: Panarchy*)

Autocatalytic Set A collection of chemical entities where each member of the set is produced by a reaction that is catalyzed by at least one other member of the same set. This creates a self-sustaining network of chemical reactions, serving as a foundational model for the emergence of self-maintaining patterns at the origin of life.

Autogen A theoretical model of a minimal, self-constraining system, proposed by Terrence Deacon, that serves as a conceptual bridge between non-living and living dynamics. An autogen consists of two coupled self-organizing processes: (1) a reciprocal catalytic network that produces molecular components, and (2) the self-assembly of these components into a container (e.g., a capsid shell) that encloses the network. By containing its own catalysts, the system constrains its own runaway reaction, thereby achieving a minimal form of self-maintenance.

Autopoiesis A term coined by Humberto Maturana and Francisco Varela, meaning "self-creation." It describes a system organized as a network of processes that continuously produce the very components that create and maintain the network itself. Autopoietic systems are characterized by their *operational closure* and their ability to actively produce their own boundaries, distinguishing them from their environment.

Closure A property of a system where its organizational or operational processes form a self-referential or self-sustaining loop. Key distinctions include:

- **Operational Closure:** As used in *autopoiesis*, the processes of the system form a closed network where each process is produced by other processes within the network. The system is materially and energetically open, but organizationally closed.
- **Closure to Efficient Causation:** As defined by Robert Rosen, a system in which every component that acts as an efficient cause (e.g., a catalyst) is itself a product of processes within the system. This property is argued to be a defining characteristic of life that distinguishes it from machines.

Constraint A local, context-dependent condition that limits a system's degrees of freedom, guiding its dynamics along specific pathways while excluding others. Constraints are distinct from universal physical laws, which provide the general framework within which constraints operate. In this thesis, the term primarily refers to *non-holonomic* constraints, those that are dependent on the system's history and path, making them crucial for describing the dynamic, adaptive nature of living systems. While constraints limit degrees of freedom, this thesis emphasizes their generative role, by excluding non-functional states, constraints enable the emergence of higher-order complexity and reliable macro-phenomena.

Constraint Cycle A concept introduced in this thesis to describe a recurring macroscopic pattern of dynamic organization fundamental to the persistence and adaptation of complex systems. The cycle is composed of three temporally distinct phases or dominant tendencies: a **Growth Phase** (accumulation of resources and structure), a **Disruption Phase** (breakdown of structure and release of resources), and a **Reorganization Phase** (reconfiguration and incorporation of novelty). This pattern is identified as a scale-invariant feature of living systems.

Constraint Alignment A concept introduced in this thesis to describe the principle that the internal constraints of an effective descriptive system (such as a formal logic, a scientific model, or a language) can functionally align with or track the inherent constraints of the phenomenon it is intended to describe.

Constraint Regime A concept introduced in this thesis for a stable and persistent set of constraints that defines a distinct level of organization by systematically excluding a range of possible configurations or behaviors. These regimes create the observable *discontinuities* ("lumps" and "gaps") in scaling distributions. In biotic systems, constraint regimes are actively maintained through energetic work. A constraint regime is characterized by three criteria: **emergence**, **discoverability**, and **stability**.

Cyclical Synthesis A concept introduced in this thesis to name a philosophical thought figure and a form of processual logic derived from the dynamics of *constraint cycles*. It describes how seemingly opposing or contradictory processes are integrated not through static resolution (as in classical logic) or dialectical sublation (as in Hegelian logic), but through their temporal sequencing and coordination within a cycle. Coherence emerges from this dynamic interplay over time, allowing constituent processes to maintain their distinct identities within their respective phases.

Degradation Dependencies A vulnerability that arises when a system, through a process of specialization, loses alternative organizational pathways and becomes dependent on a narrow set of external inputs or relationships. This loss of flexibility and redundancy often remains hidden during periods of stability but can lead to failure when the system is perturbed.

Diachronic A perspective that examines phenomena as they develop and change through time. A diachronic view emphasizes process, history, and temporal evolution as central to understanding a system. It is contrasted with the *synchronic* perspective.

Discontinuity (Lumps and Gaps) A concept from the work of C.S. Holling and Craig Allen describing the non-uniform distribution of entities along a continuous scaling relationship (e.g., body mass). Organisms and other entities tend to cluster into aggregations ("lumps") separated by regions with few or no entities ("gaps"). In this thesis, these discontinuities are interpreted as the observable manifestation of underlying *constraint regimes*.

Distributed Integration (DI) A concept introduced in this thesis for the degree to which the components of a composite system actively coordinate and invest energy to maintain their collective constraints and organizational coherence. DI exists on a spectrum, from loose aggregations with minimal collective investment to tightly integrated, organism-like wholes with high investment in shared maintenance. It can be operationalized as the ratio of energy spent on collective self-constraint to the system's total energy throughput.

Emergence The arising of novel properties, patterns, or behaviors in a system that are not present in its individual components. These emergent features result from the interactions and relationships between the parts. The thesis refers to a taxonomy of emergence: Computational, Material, Organizational, and Reflexive, representing increasingly complex forms of organization.

Finite Game / Infinite Game A distinction introduced by James P. Carse. A **finite game** is played for the purpose of winning and has fixed rules, boundaries, and an end point. An **infinite game** is played for the purpose of continuing the play, and its rules and boundaries must change during the course of play. The thesis maps this distinction onto computational systems (*finite games*) versus living systems (*infinite games*).

Hardware-Software Distinction The fundamental separation in conventional computing between the physical machinery that performs operations (hardware) and the set of instructions or programs that direct these operations (software). The thesis argues that this rigid distinction is a defining feature of non-living computational systems and is absent in living systems, where information and material substrate are recursively intertwined.

Impredicativity A logical property where a definition or process references the totality to which it belongs. In the context of Robert Rosen's work, living systems are described as impredicative because they generate their own efficient causes from within; the system defines its own components, which in turn define the system. This circularity prevents the rigid separation of hardware and software found in machines (which are predicative).

Modelling as Constraint Selection A concept introduced in this thesis that reframes the act of scientific modelling. Instead of viewing models as representations of reality, this perspective treats modelling as the deliberate selection, imposition, and stabilization of a set of constraints to make a particular phenomenon intelligible and tractable. This act both enables certain insights and necessarily excludes others.

Open System / Closed System A distinction from thermodynamics. A **closed system** is an idealized system that exchanges energy (e.g., heat, work) but not matter with its surroundings. An **open system**, which characterizes all living organisms, continuously exchanges both energy and matter with its environment. This exchange is a prerequisite for maintaining a far-from-equilibrium state.

Organizational Approach (OA) A theoretical framework in biology that seeks to explain the distinctive properties of life, including teleology, through the specific organization of living systems. Its central concept is the "closure of constraints," where a set of mutually dependent constraints collectively maintains the system as a whole, allowing it to be both self-determining and thermodynamically open.

Panarchy A conceptual framework, developed by Holling and Gunderson, that models the dynamic interactions between nested adaptive cycles (or *constraint* *cycles*) operating at different spatiotemporal scales. Panarchy explains how different organizational levels influence one another through two primary connections: **"revolt"** (when rapid changes in a smaller, faster cycle cascade upward to disrupt a larger, slower cycle) and **"remember"** (when a slower, larger cycle provides memory, resources, and stability to guide the reorganization of a disrupted faster cycle).

Scaling Law A mathematical relationship where one quantity varies as a power of another (e.g., Kleiber's law, where metabolic rate scales to body mass to the 3/4 power).

Self-Constraining Dynamics A dynamic regime in which coupled self-organizing processes mutually constrain one another, preventing the system from reaching thermodynamic equilibrium. Unlike simple self-organization, which accelerates entropy production, self-constraining dynamics perform work to preserve the system's organization and energy gradients.

Self-organization A process where local interactions between components lead to the emergence of global patterns without external control (e.g., convection cells, whirlpools). In this thesis, it is strictly distinguished from *self-constraining dynamics*. While self-organizing systems spontaneously generate order to dissipate energy gradients *faster*, they lack the internal constraints to preserve those gradients or their own organization in the absence of external flow.

Small World / Large World A distinction from decision theory, used here to contrast computational and living systems. A **small world** is a bounded context where all states, acts, and consequences can be fully described and analyzed in advance (characteristic of computational systems). A **large world** is an open-ended environment that cannot be exhaustively described, where novelty can emerge and rules can change (characteristic of living systems).

Synchronic A perspective that examines a system at a single point in time, focusing on its structure, components, and relationships as they exist in a static snapshot. It abstracts away from temporal development and historical processes. It is contrasted with the *diachronic* perspective.

Teleodynamics A term coined by Terrence Deacon for a dynamic regime that arises when self-organizing processes become coupled in such a way that they mutually constrain one another. This opposition prevents the system from reaching a state of maximum entropy, generating a higher-order, self-maintaining dynamic that gives rise to emergent, intrinsic purpose (or *telos*). In this thesis, it is used interchangeably with *self-constraining dynamics*.

Teleology The study of purpose, goals, or end-directedness in natural processes.

- **Extrinsic Teleology:** Purpose that is imposed on a system from an external agent or designer (e.g., the purpose of a clock).
- **Intrinsic Teleology:** Purpose that arises from the internal organization of a system itself. In this thesis, intrinsic teleology is argued to be a naturalized, emergent property of *self-constraining* systems that actively work to maintain their own existence.

Notes

1. Paraphrased from Heraclitus, fragments DK 22B12 and 22B49a; cf. Plato, *Cratylus* 402a. See [Reeve, 1998, Heraclitus, 1952].

2. Liz Harper. Understanding earth. *Geological Magazine*, 139(1):101–102, 01 2002. ISSN 0016-7568. doi: 10.1017/S0016756802286513. URL `https://doi.org/10.1017/S0016756802286513`

3. For a comprehensive articulation of this perspective, see [Nicholson and Dupré, 2018]. The volume consolidates the *processual turn* in the philosophy of biology, arguing that living systems are best understood not as static substances (things) that undergo change, but as stabilized processes (flows) whose persistence is a dynamic achievement. By shifting the ontological priority from matter to process, they demonstrate how biological identity is maintained not despite metabolic turnover, but through it. Their work has been instrumental in grounding process philosophy in the concrete empirical realities of physiology, development, and ecology.

4. Aristotle. *Metaphysics: Books I-IX.* Harvard University Press, Cambridge, 1979

5. See Pasteur [1861]

6. Darwin [2012]

7. See Oparin [1924], Haldane [1929]

8. Miller [1953]

9. The discovery of ribozymes (catalytic RNA molecules) in the 1980s by Thomas Cech and Sidney Altman, work that earned them the Nobel Prize, demonstrated this was a biochemical realityGuerrier-Takada et al. [1983], Kruger et al. [1982]. Additionally, the ribosome, the universal cellular machinery that translates genetic information into proteins, has RNA at its catalytic core.

10. The technical challenges facing RNA World are substantial. First, RNA's building blocks (ribonucleotides) are intricate structures that are exceedingly difficult to synthesize under plausible prebiotic conditions. Second, catalytic activity is rare among random RNA sequences—it typically requires long, specifically folded polymers that are statistically improbable to arise by chance. Third, even if a self-replicating ribozyme formed, the resulting RNA duplex (two strands stuck together) would be so stable that the strands wouldn't separate without sophisticated protein machinery. This leads to what's called product inhibition—after a single replication cycle, the system gets stuck and can't continue copying.

11. One version is Günter Wächtershäuser's Iron-Sulfur World theory Wächtershäuser [1988]. Rather than a dilute oceanic soup, Wächtershäuser proposed that life originated on mineral surfaces near deep-sea hydrothermal vents. These environments are rich in chemical reactants, and crucially, the charged surfaces of iron-sulfide minerals could concentrate organic molecules and catalyze reactions. This surface metabolism would be powered by the chemical energy released from mineral formation, creating an autocatalytic cycle where some products enhance the catalytic activity of the surfaces, allowing the network to grow and complexify.

12. See Orgel [2008]

13. E. Schrödinger. *What is Life? the Physical Aspect of the Living Cell & Mind and Matter*. University P., 1967. URL `https://books.google.at/books?id=9RHbAAAAMAAJ`

14. H. Bergson. *Creative Evolution*. Modern library. University Press of America, 1984. ISBN 9780819135537. URL `https://books.google.at/books?id=Cx4H9wTF8mkC`

15. E.D. Schneider and D. Sagan. *Into the Cool: Energy Flow, Thermodynamics, and Life*. University of Chicago Press, 2005. ISBN 9780226739366. URL https://books.google.at/books?id=u4wb1EQWQDOC

16. As an illustrative example, consider how physicists handle atoms of increasing complexity. The hydrogen atom—the simplest atom—comprises just two bodies: one proton and one electron. Because of this simplicity, the quantum mechanical equations (the Schrödinger equation with the hydrogenic Hamiltonian) are exactly solvable, allowing for precise calculation of all its properties. However, adding just one more electron, as in the helium atom, makes the situation vastly more complicated. In helium, we have three quantum particles to consider: the nucleus (treated as a single, nearly stationary mass) and two electrons. Unlike hydrogen, the equations for helium cannot be separated into simple, independently solvable parts. This is not merely a technical limitation; it reflects the fundamental mathematical complexity of three-body systems, well-known both in classical mechanics (as famously studied by Poincaré) and in quantum mechanics. Since the electrons' motions are intertwined and cannot be decoupled, approximation methods are required. This underscores why analysing large, biologically relevant systems with billions of interacting particles is orders of magnitude more complex than even the helium atom—a problem that is already too complex for exact analytical solutions.

17. S.J. Gould. *Wonderful Life: The Burgess Shale and the Nature of History*. W. W. Norton, 1990. ISBN 9780393245202. URL https://books.google.at/books?id=q8BUAgAAQBAJ

18. A.S. Eddington. *The Nature of the Physical World*. Gifford lectures. Macmillan, 1940. URL https://books.google.at/books?id=GCfcBqVFQSIC

19. For an excellent discussion on the field of thermodynamics and its history the work of ([Brush, 1976]) is recommended. As well as Julian Barbour's work, both his recent book The Janus Point ([Barbour, 2020]), as well as the extended discussion that can be found on his website: http://platonia.com/index.html

20. Rudolf Clausius. On different forms of the fundamental equations of the mechanical theory of heat and their convenience for application. *The second law of thermodynamics*, pages 162–193, 1865

21. This mechanism aligns with Erwin Schrödinger's concept of *negative entropy* (later shortened to *negentropy* by Léon Brillouin). In *What Is Life?*, Schrödinger argued that organisms avoid rapid

decay into thermodynamic equilibrium by *feeding on negative entropy*, continuously extracting order from their environment to offset the entropy they naturally produce. Although he later acknowledged that the term was physically imprecise (noting that *free energy* is the rigorous quantity), it captures the essential thermodynamic transaction: life maintains its organization by importing low-entropy energy and exporting high-entropy waste.

22. CGP Grey. Hexagons are the bestagons. YouTube, 2020. URL https://www.youtube.com/watch?v=thOifuHs6eY&t=326s. Accessed: 2024-06-13

23. Ilya Prigogine and Isabelle Stengers. *Order out of chaos: Man's new dialogue with nature*. Verso Books, 2018

24. The term was coined by W. Ross Ashby [1947]. Many people are not comfortable with the term self-organization because there is no 'self' there. However, since the term is well established by now, it is the term we will continue to use in this thesis. It is also important to note that Alan Turing independently explored concepts related to emergent order and pattern formation in biological systems in his paper [Turing, 1990]. Turing and Ashby were contemporaries who corresponded and were both members of the Ratio Club, a group of British scientists and engineers interested in cybernetics and related fields. Their interactions suggest a mutual awareness of each other's work, although their contributions to the understanding of self-organizing systems were developed independently.

25. The appeal to *closed systems* here is a deliberate idealization in the sense analyzed by Weisberg: models simplify, omit, or distort features of target systems to isolate explanatory structure and manage trade-offs among accuracy, tractability, and generality [Weisberg, 2007, 2013]. Morgan likewise emphasizes that models are not mirrors of reality but constructed, idealized *little worlds* that mediate between theory and phenomena, letting us explore how certain organizing principles would behave under controlled assumptions [Morgan, 2012, 1999]. On this view, talk of closed systems functions as a map-like device, empirically unattainable in the strict sense, yet indispensable for identifying the dynamical roles of openness, gradients, and far-from-equilibrium self-organization.

26. Free energy refers to the portion of energy in a system that is available to perform work under specific conditions, such as constant temperature and pressure. It plays a crucial role in driving spontaneous and ordered processes on Earth by utilizing the Sun's energy to maintain and develop complex structures. While no energy is entirely *free* in an absolute sense, within Earth's open

system, it effectively serves as free energy by sustaining the necessary energy gradients. This concept, rooted in the principles of thermodynamics and developed by scientists like Gibbs and Helmholtz, is fundamental to understanding how energy flows enable the emergence and persistence of complexity in natural systems. Importantly, the sun is not the only gradient driving the evolution of complexity on earth, other gradients like the hot core of the earth or the tidal pull of the moon, play a role as well.

27. The cosmological background assumed here is the standard CDM + inflationary reconstruction: tiny, nearly scale-invariant primordial density perturbations (as constrained by CMB anisotropies) are generated in the early universe and later amplified by gravitational instability into a long-lived, hierarchically clustered cosmic web of overdense and underdense regions [Planck Collaboration, 2020, Peebles, 1980, Bernardeau et al., 2002]. Within that framework, large-scale gradients persist because equilibration is limited by finite propagation speeds and dissipative transport (radiation cannot exceed c, matter flows face viscosity/friction) and by the competition between local collapse and global Hubble expansion, so contrasts are not smoothed out on the largest scales even over cosmic times [Peebles, 1980, Bernardeau et al., 2002]. Alternative accounts of the deep origin of these initial inhomogeneities exist. Among the most intriguing is Smolin's *cosmological natural selection* (an evolutionary cosmology in which universes reproduce via black holes and inherit slightly varied constants), recently re-popularized and creatively extended by Julian Gough in his ongoing public book project *The Egg and the Rock* [Smolin, 1997, Gough, 2022]. For present purposes, however, the specific micro-dynamics of how the unsmoothness first arose are secondary. What matters is the robust consequence shared by these approaches and by CDM alike, the universe begins with and continues to sustain non-vanishing gradients, and these enduring asymmetries provide the free-energy differences that drive subsequent self-organization (including life's).

28. Image sources and details: (1) MODIS satellite image acquired June 6, 2002, by NASA Terra satellite. Image credit: Jacques Descloitres, MODIS Land Rapid Response Team, NASA/GSFC. Available from NASA Earth Observatory (https://earthobservatory.nasa.gov/). The image shows closed Rayleigh-Bénard convection cells in marine stratocumulus clouds over the South Atlantic Ocean, where air rises in cell centers and sinks at edges. Scientific interpretation provided by Robert Cahalan, NASA-GSFC. (2) Salar de Uyuni salt flat (20°19'57.27"S, 66°59'51.09"W) patterns formed during seasonal water evaporation at 3,656m elevation. These large-scale Bénard cells form during the December-March rainy season as water evaporates under still conditions. (3) Basalt columns in Ásbyrgi, Vatnajökull National Park, Iceland. Photo © Catherine Christopoulou, July 15, 2014, 11:48:00, licensed under Creative Commons Attribution-Share Alike 4.0 International license. (4) Swedish 1-meter Solar Telescope (SST) at La Palma, Spain, observation at Ca II K-line wavelength. The field of view spans 46,900 × 33,500 km, showing solar granules—the visible tops of convection cells. Data courtesy of the Institute for Solar Physics, Stockholm University. (5) Lab-

oratory visualization from NOAA/OAR/Physical Sciences Laboratory, Boulder, Colorado, USA. Ten-second exposure showing aluminum particle trails in Rayleigh-Bénard convection cells. Data and image retrieved from https://psl.noaa.gov/. (6) Honeycomb structure photograph from PickPik (https://www.pickpik.com/), licensed under Creative Commons Zero (CC0) license. While not a convection pattern, the honeycomb demonstrates optimal hexagonal surface division as proven by Hales's honeycomb conjecture

29. The honeycomb conjecture was first conjectured as early as 36 BC by Marcus Terentius Varro and discussed by Pappus of Alexandria (c. 290–350 AD). Partial proofs were later provided by László Fejes Tóth [1943], who proved the case for convex polygons, and by Jaigyoung Choe [1989], who addressed square-grid constraints. The full proof was announced in 1999 and published in 2001 by Thomas C. Hales [2001]. The conjecture relates closely to the densest circle packing problem, which also favors hexagonal arrangements Weisstein [2010]. For broader context, see discussions in Quanta Magazine [Cepelewicz, 2020].

30. Interestingly for the case of the honeycomb conjecture has been proven in two dimenions, for three dimensions an alternative structure (based on the truncated octahedron) could reduce material use slightly (roughly 0.35%). So while the honeycomb conjecture explains the geometric efficiency of hexagons, some have argued that the actual construction of honeycomb cells involves more than geometric optimality. For example, Räz [2017] highlights alternative hypotheses including physical mechanisms like liquid equilibrium shaping and acoustic resonance tuning that bees might use, indicating that structural optimality alone may not fully explain how hexagonal combs are formed. Specifically, studies suggest bees may initially build circular cells, which then transform into hexagons due to surface tension when the bees heat the wax, similar to how a raft of bubbles self-organizes. Thus, the optimization goal may be achieved not through cognitive calculation, but by harnessing physical laws of surface tension.

31. The distinction between the bottle and bathtub vortices reveals an important subtlety about self-organizing systems: not all emerge spontaneously under the same conditions. The bathtub vortex forms automatically because its geometry—a large, flat surface converging toward a small drain—creates ideal conditions for amplifying any tiny seed rotation through conservation of angular momentum. As water spirals inward over this long distance, rotation naturally amplifies into a stable vortex. Crucially, the bathtub is open to the atmosphere, so air can freely replace draining water from above without creating pressure conflicts, allowing continuous organized flow. The inverted bottle presents a fundamentally different scenario: its geometry creates *plug flow* where the water column moves downward as a unit, and critically, water and air must compete for passage through the same narrow opening. When water drains, the expanding air inside creates a partial

vacuum, triggering a fast-growing Rayleigh-Taylor instability that causes violent *glugging*—large air bubbles forcefully intrude upward, disrupting any nascent rotation before it can amplify. The system becomes kinetically trapped in this inefficient cycle of intermittent flow. By spinning the bottle before inverting it, we manually impose angular momentum strong enough to establish a stable air column down the center before the glugging instability takes over. This vortex provides separate channels for water (spiraling down around the edges) and air (flowing up through the core), eliminating the pressure conflict and allowing the system to access its more efficient dissipative pathway. This illustrates a broader principle. Self-organizing structures represent optimal solutions for dissipating gradients, but whether they emerge spontaneously or require an initial energy investment depends on the specific constraints of the system—its geometry, the presence of competing instabilities, and the timescales on which different processes can develop. The right conditions must be aligned for nature's most efficient solutions to manifest.

32. See [Odum, 2007, Lotka, 1922, Schneider and Kay, 1994]

33. While the Maximum Entropy Production Principle focuses on dissipation, systems ecology offers a complementary orientor: the maximization of eco-exergy (work capacity). Fath and Jørgensen argue that as ecosystems mature, they do not just process more energy but increase their capacity to store information and structure, effectively moving further from thermodynamic equilibrium. This suggests that the "goal" of self-organization is not just throughput, but the retention of organization." Jørgensen et al. [2007].

34. This self-referential moment of observation echoes John Wheeler's concept of the "universe's eye" - his notion that the universe achieves self-observation through the emergence of conscious observers. Wheeler suggested that quantum mechanics implies a universe that creates observership, while these observers in turn give reality to the universe through their observations[Wheeler, 1980].

35. It is interesting to consider why this type of relaitonship is so hard for us to grasp. Lakoff [1999] suggest that we fundamentally think about the world through what they call the "containment doctrine" - we conceptualize things as discrete containers with clear inside/outside boundaries. This works well for understanding mechanical systems but breaks down when confronting self-referential processes. Hofstadter's concept of "strange loops" helps us understand these cases: they are patterns where, by moving up or down through some hierarchical system, we paradoxically arrive back where we started Hofstadter [2006, 2008]. Consciousness observing itself is a prime example of such a loop.

36. Gallup conducted the original work on this in the 1960s, and olfactory and tactile versions of the test have been suggested; see [Gallup, 1970] and [Horowitz, 2017] as well as [De Waal, 2008]

37. While discussing the intricacies of what precisely sets humans apart from other species is beyond the scope of this discussion, it is clear that within the Darwinian paradigm, we are firmly situated within the kingdom of animals. However, if we are to explore potential differences, our capacity for language and symbolic communication provides a promising avenue of inquiry. These abilities, particularly in their complexity and unique elements, are perhaps the most distinguishing features of our species. Human language is distinct from animal communication in its reliance on symbolic representation, a feature that enables abstract thought, planning, and storytelling. Unlike the indexical or iconic communication seen in other species, symbolic systems allow humans to transcend immediate contexts and create interconnected networks of meaning. Words in human language simultaneously reference objects and form relationships with other words, creating a higher-order combinatorial system. This unique cognitive capability gave rise to a "virtual world" accessible only to humans, where possibilities, hypothetical scenarios, and complex ideas can be explored. Moreover, the evolution of this symbolic capacity has profoundly shaped the structure and function of the human brain, particularly the prefrontal cortex, which supports this advanced mode of communication. For an in-depth exploration of this topic see [Deacon, 1998, Tomasello, 1999, Pinker, 1994].

38. The history of biology has seen various attempts to explain life's distinctive properties by positing additional forces or principles beyond physical law—like an élan vital [Bergson, 1984]. However, the organizational approach suggests we can understand life's special characteristics through careful analysis of how living systems organize matter and energy, without requiring any extra-physical principles.

39. The transition from mechanical to informational metaphors in biology is thoroughly documented by historian Lily Kay in *Who Wrote the Book of Life?*, where she traces how cybernetics and information theory reshaped molecular biology in the 1950s and 1960s [Kay, 2000]. The specific view of the cell as a computational entity is extensively developed by Dennis Bray, who shows how protein signaling networks form biochemical circuits that implement logic-like operations and distributed information processing [Bray, 2009, 1995]

40. When discussing the perspective of "the largest," it's crucial to clarify what is meant by this term in contrast to the "smallest." The smallest refers to reductionist approaches—those that begin from the most fundamental units of matter and attempt to explain complex systems through the

interactions and properties of these basic components. In contrast, the largest is not referring to the macrocosmic scale of galaxies, planets, or large structural hierarchies of the universe. While these, too, might represent grand, large-scale phenomena, the perspective invoked here is of a different kind entirely. The "largest," in this context, refers to the entirety of reality as it is funneled through and mediated by subjective experience. It is, in essence, an idealistic perspective that takes our conscious experience—the immediate, lived phenomenon of being—as the most fundamental or "largest" frame of reference. From this viewpoint, the universe does not begin with particles, galaxies, or even physical laws, but with the raw, undeniable reality of experience itself.

41. David Deutsch's The Fabric of Reality presents an argument for the computational nature of the universe, integrating ideas from quantum physics, epistemology, and computation to propose that the universe operates as a vast quantum computer [Deutsch, 1997b]. Stephen Wolfram's earlier work, A New Kind of Science, explores the principle of computational equivalence, suggesting that even simple systems can produce computational behavior as complex as any known system [Wolfram, 2002b]. This foundational work provided the basis for his more recent Physics Project, which radicalizes these ideas by proposing that the universe is computational from the ground up [Wolfram, 2020]. The Physics Project investigates how simple computational rules can give rise to the intricate structures and laws observed in fundamental physics, including space, time, and quantum mechanics.

42. Rolf Landauer's seminal article Information is Physical established a critical link between computation and physical systems, asserting that information is not abstract but intrinsically tied to its physical representation. Landauer argues that computation and information processing are constrained by the laws of thermodynamics and quantum mechanics, fundamentally tying the nature of information to the behavior of physical systems. His work laid the groundwork for understanding information as a central concept in physics, with implications for fields ranging from computational theory to the study of entropy and energy[Landauer, 1991].

43. Panpsychism has a long and diverse history, with early traces found in ancient Greek philosophy, including the work of Thales, Anaxagoras, and the atomists, who grappled with the idea of mind as a fundamental aspect of reality. In modern philosophy, panpsychism has been refined by thinkers like Spinoza and Leibniz, and in the 19th century by Fechner, Wundt, and James. Contemporary panpsychism, as advanced by Philip Goff [2019] in *Galileo's Error* and Galen Strawson [2006] in *Realistic Monism: Why Physicalism Entails Panpsychism*, argues for consciousness as an intrinsic and ubiquitous feature of the physical world, offering a middle path between reductive physicalism and dualism. For a comprehensive history, see Skrbina's *Panpsychism in the West* [Skrbina, 2005] and the entry on panpsychism in the Stanford Encyclopedia of Philosophy[Goff et al., 2022].

44. Thompson [2007]

45. Thomas Nagel critiques panpsychism by highlighting the combination problem: how proto-psychic properties of an organism's elements combine to produce unified individual consciousness[Nagel, 2012].

46. For an extended discussion of Descartes's position in the philosophy of mind, see Descartes [1641], Rozemond [2002], and Almog [2001].

47. Descartes placed the soul in the pineal gland because, unlike the brain's bilaterally duplicated parts, it is a single midline structure.

48. On more contemporary dualisms, see Foster [1989], Lowe [2006], Swinburne [1997], and Taliaferro [1996]. Some defenders of these newer dualisms propose that while physical science explains many aspects of cognition and consciousness, certain qualitative or intentional features (*what it is like* to experience or the *aboutness* of thought) cannot be captured by purely physicalist accounts. Others, often termed *emergent dualists*, suggest that mental properties emerge from physical systems but remain ontologically irreducible. In either case, the mental is not subsumed by the purely physical.

49. For helpful surveys of the interaction problem and causal closure, see Kim [2003] and Robinson [2003]. The core question: if the physical realm is closed under physical laws, can the non-physical exert genuine causal influence without violating energy conservation?) attempt to resolve these issues.

50. While this organizational pattern is typically exemplified by organisms, it can also manifest in other biological systems like colonies, symbioses, and ecosystems. The third part of this book *Discontinuity* will spend more time with these questions.

51. Kant [1797]

52. Cuvier et al. [1859]

53. Ironically, while this principle powerfully demonstrated the deep interconnections within living systems, Cuvier used it to argue against evolutionary change. He reasoned that since all parts were so intimately connected, any significant modification would disrupt the whole system, making evolutionary transitions impossible. History has proven him wrong on this point—we now know that evolution can indeed modify multiple correlated traits gradually over time. Nevertheless, one could argue that his arguments of relational co-dependence resonates to some extent with insights into gene regulatory networks. These networks function in a similarly interconnected manner, where changes to one gene or regulatory element often cascade through the system, producing wide-ranging effects. This co-regulatory dynamic mirrors Cuvier's observations: small changes within an organism's structure or function must align with broader systemic compatibility, therefore also aligning with the punctuated equilibria theory in evolutionary biology, introduced by paleontologists Niles Eldredge and Stephen Jay Gould Eldredge and Gould [1971]. Their work on the fossil record suggests that evolutionary change often occurs in rapid bursts, punctuating long periods of stability—perhaps a reflection of the same systemic interdependence Cuvier observed. When changes do occur, they often require simultaneous shifts across multiple traits to maintain functional coherence.

54. John Locke famously distinguished between primary qualities (like shape, size, and motion) which he considered inherent in objects themselves, and secondary qualities (like color, taste, and smell) which he saw as dependent on the observer [Locke, 1753]. Gregory Bateson [2000] later challenged this view by emphasizing that all properties emerge from relationships. His insight that even supposedly primary properties reveal their relational nature in different contexts (consider how mass and weight depend on gravitational field strength) helps us understand why organization itself, rather than just the organized components, must be considered primary.

55. This understanding of biological organization as fundamentally different from mechanical arrangement continued to evolve, particularly through the German organicist tradition. Johannes Müller's physiology and Karl von Baer's embryology demonstrated how this perspective could reveal new insights about both the development and functioning of organismsMüller [1835], von Baer [1828]. It is no coincidence that many of these insights emerged from embryologists like von Baer, who witnessed first-hand how through cell differentiation and growth, an organism maintains its wholeness even as it increases in size and complexity—a vivid demonstration of how parts and whole co-develop during embryogenesis, with each part's development intimately dependent on and influenced by the development of other parts and the whole organism.

Claude Bernard would later emphasize, even when we must decompose an organism for study, we can only truly understand its parts by reference to the whole they constitute. In his words, "The physiologist and the physician must never forget that the living being comprises an organism and an individuality... If we decompose the living organism into its various parts, it is only for the sake of experimental analysis, not for them to be understood separately. Indeed, when we wish to ascribe to a physiological quality its value and true significance, we must always refer to this whole and draw our final conclusions only in relation to its effects in the whole." Bernard [1865]

The German-speaking world, in particular, developed a sophisticated organicist tradition, including Goethe's morphological studies von Goethe et al. [1994] and Alexander von Humboldt's holistic approach to natural phenomena von Humboldt and Gröls-Verlag [2022]. This tradition continued into the 20th century with the formation of the Theoretical Biology Club in the 1930s by Joseph Henry Woodger, Joseph Needham, Conrad Hal Waddington, John Desmond Bernal, Dorothy Needham, and Dorothy Wrinch Peterson [2016]. Influenced by Alfred North Whitehead's philosophy, these scientists and philosophers developed a systematic opposition to mechanical reductionism biology, though their ambitious research program ultimately failed to secure institutional support from the Rockefeller Foundation. Their work represents just one example of the broader intellectual movements that have sought to understand biological organization in non-reductionist terms.

56. In collaboration with my colleague Kevin Purkhauser, we have developed a taxonomy of emergence that distinguishes four fundamental types, each representing a distinct ontological stratum: (0) Computational Emergence, (1) Material Emergence, (2) Organizational Emergence, and (3) Reflexive Emergence. While various taxonomies of emergence have been proposed (see Stephan [1999]), this classification offers a novel perspective that bridges physical and biological emergence. This framework builds upon previous discussions of emergence in physics [Hüttemann and Terzidis, 2000] and addresses long-standing questions about how emergent properties can be both dependent on and autonomous from their bases, a central concern in philosophical discussions of emergence [Morgan, 1923]. Like other taxonomies, it remains primarily descriptive rather than explanatory, it helps identify and categorize different manifestations of emergence without fully elucidating the generative dynamics that enable transitions between levels. Nevertheless, it has proven valuable as an analytical framework, particularly for understanding the bridges between non-living and living systems.

Level	**Type**	**Characteristics**	**Examples**
0	Computational Emergence	Idealized emergence in closed systems; deterministic rules; perfect repeatability	Cellular automata; agent-based models; digital simulations
1	Material Emergence	Open physical systems; actual material interactions; non-computable complexity; environmental coupling	Whirlpools; crystal formation; weather patterns
2	Organizational Emergence	Recursive self-stabilizing patterns; multi-level process coupling; self-maintenance; normative constraints	Metabolic networks; biological cells; ecosystems
3	Reflexive Emergence	Mental representations; symbolic processing; self-modelling; anticipatory behaviour; value-driven choices	Consciousness; cognitive processes; cultural systems

Table 1: Taxonomy of Emergent Forms

Computational Emergence represents idealized emergence in closed systems, where complex patterns arise from deterministic rules, as seen in cellular automata and computer simulations. Material Emergence occurs in open physical systems where emergence manifests through actual material interactions that cannot be perfectly computed due to complexity and non-computability, characterized by real-world constraints and environmental coupling. Organizational Emergence builds upon these lower forms, characterized by recursive, self-stabilizing patterns that emerge from the coupling of multiple lower-level processes, exemplified by biological systems and their capacity for self-maintenance through component turnover. Finally, Reflexive Emergence introduces a higher order of virtuality through mental representations, enabling agents to distinguish and choose between different environmental paths based on their internal models, leading to genuinely anticipatory behaviour and value-driven choices.

This framework naturally subsumes traditional distinctions like synchronic versus diachronic emergence [Sartenaer, 2015]. Beyond idealized computational emergence, all higher forms necessarily involve both immediate (synchronic) and historical (diachronic) aspects, which become increasingly pronounced at higher organizational levels. Material emergence already introduces temporal irreversibility, organizational emergence adds historical contingency through adaptive processes, and reflexive emergence incorporates both individual and collective historical dimensions.

The framework's progression from software (Level 0) through hardware (Level 1) and wetware (Level 2) to cogware (Level 3), reflects increasingly sophisticated forms of causation and constraint. Each level builds upon and incorporates the dynamics of previous levels while introducing qualitatively new forms of organization and behaviour. While this taxonomy doesn't fully explain the mechanisms of transition between levels, it provides a systematic way to understand the different manifestations of emergence in nature, from physical self-organization to conscious experience.

57. See Maturana and Varela [2012], Kauffman [1993], Rosen [1991]

58. Maturana and Varela [2012]

59. The concept has also influenced fields far beyond biology. Niklas Luhmann applied autopoiesis to social systems, arguing that social systems too are self-producing networks (of communications rather than molecules) Luhmann [1982]. Cognitive scientists have used it to understand consciousness as a self-producing pattern of neural activity. Even some approaches to artificial life draw on autopoietic principles.

60. Kauffman [1993]

61. Since these illustration are used to build up the model of the Autogen illustrated by Terrence Deacon, these diagrams and the shape of the molecules is inspired by his work

62. Kauffman simulated the dependency of autocatalytic set emergence on N (number of polymer types) using mathematical methods, specifically random graph theory, showing that catalytic closure becomes probable as N increases, without empirical lab experiments.

63. See Vaidya et al. [2012] for experimental demonstration that small RNA fragments form mutually catalytic networks capable of collective self-replication, with network-level persistence not reducible to a single autonomous replicator.

64. In ecological network analysis, this circularity generates a phenomenon known as network synergism. Fath and Patten demonstrated mathematically that as energy cycling increases within a network, the system retains energy longer, and indirect positive effects begin to outweigh direct competitive interactions. This provides a formal quantitative basis for the claim that the organization of the whole generates a surplus utility that stabilizes its components. See Fath and Patten [1998]. For a review of the foundational methodology, see Fath and Patten [1999].

65. Gabora and Steel [2020] develops the idea that cultural evolution can be modelled as autocatalytic networks of mental representations/ideas that mutually evoke and generate further ideas, enabling self-sustaining, cumulative cultural change.

66. Rosen [1991]

67. Hofmeyr [2021]

68. **Extended Mathematical Discussion of Rosen's (M,R)-Systems**

Overview. Robert Rosen's (M, R)-systems offer a rigorous mathematical formalism to demonstrate why living systems exhibit a type of organization that fundamentally differs from that of machines. The key is *closure to efficient causation*, which Rosen formalized using category theory. In what follows, we address:

- A succinct primer on category-theoretic constructs relevant to Rosen's argument.
- The classical (M, R)-system notation (metabolism, repair, and replication).
- How *impredicativity* underwrites the impossibility of fully modeling organisms as machines.

1. Category Theory in Brief. A category $\mathcal{C}$ consists of:

- **Objects** (which may be sets, spaces, etc.).

- **Morphisms** or **arrows** between these objects, closed under composition, and each object has an identity arrow.

In Rosen's biological modeling:

- *Objects* (e.g., A, B) represent *material causes* such as sets of substrates or products.
- *Morphisms* (e.g., f) represent *efficient causes*, typically transformations like enzymatic catalysis.

Hence, a morphism $f : A \to B$ can be read as "a metabolic function transforming inputs A into outputs B."

2. The (M, R)-System Notation. Rosen's archetypal diagrams often appear as:

$$A \xrightarrow{f} B \xrightarrow{\Phi} H(A, B),$$

where

- A is the set of *inputs* (nutrients, substrates, etc.).
- B is the set of *outputs* (e.g. proteins, cellular components).
- $f \in H(A, B)$ is the *metabolic function*.
- $H(A, B)$ denotes the set (or space) of all morphisms from A to B, i.e. $\mathrm{Hom}(A, B)$.
- $\Phi : B \to H(A, B)$ is a "repair" mapping, which uses elements of B to generate or sustain the metabolic transformations f.

The main takeaway is that products (B) feed back to create or repair the very transformations that produced those products, *closing* an otherwise open causal chain.

3. Closure to Efficient Causation and Impredicativity. A defining hallmark of organisms is that for *every* efficient cause (every $f \in H(A, B)$), there exists a $b \in B$ inside the system such that

$$\Phi(b) = f.$$

Since Φ itself is in $\mathrm{Hom}\big(B,\ H(A,B)\big)$, the system is *impredicative*: Φ is both part of and a mapping into the space of all possible transformations it helps generate or maintain.

In conventional machines, efficient causes (drivers, catalysts, instructions) originate externally (an engineer, operator, or external supplier). By contrast, in a living system, all catalysts, templates, and repair functions are *produced internally*, thus the closure.

4. The Emergence of Replication β. Under certain category-theoretic conditions, the presence of Φ entails an additional morphism often called "replication" or an *evaluation map*:

$$\beta : H(A,B) \ \longrightarrow\ H\big(B,\ H(A,B)\big).$$

Though not equivalent to full biological reproduction, it reflects how the system can replicate the very functions Φ and f from within—answering "Who repairs the repair function?" The solution arises from an "evaluation" operation, typically $\hat{b}(\Phi) = \Phi(b)$, whose invertibility in certain contexts underpins β.

5. Why Such Closure Defies Mechanical Modeling. Rosen's central theorem is that such closure to efficient causation cannot be captured by a classical machine model. Machines, no matter how complex, rely on external inputs for building, designing, or updating their components. They cannot *internally* originate all their own efficient causes. By trying to replicate an (M,R)-system with a purely mechanical system, one inevitably leaves out some causal factor or places it "outside," thereby violating closure.

Thus, Rosen's work reveals fundamental limits to strictly reductionist or mechanistic approaches in biology: living systems exhibit an irreducible circular (impredicative) organization that machines inherently lack.

69. This of course is tautological given Rosen's definitions. A machine is defined in a such a way that it is different in organization from a living system - hence if we were to ever build things that behaved like living processes, those would then be considered living systems in Rosen's modelling.

70. For an excellent discussion on Convention and how we might interpret it here, the work by David [Lewis, 1969] is recommended, which argues that Convention can be understood as the solution of

game-theoretic coordination problems and hence need not imply explicit or implicit agreement by those who come to establish a given convention. This is generally how I understand convention when using it, as it helps point to a pragmatic dimension of already established conventions that needn't be explicit but that have emerged and stabilized as coordination mechanisms.

71. Contrary to popular belief, Johannes Gutenberg did not invent the concept of a font as we understand it today—a standardized set of characters produced from a common mould. Instead, his innovation involved individually casting each letter, resulting in unique variations of the same character across printed materials. This approach reflects a more artisanal and less mechanized method than previously thought, emphasizing the craftsmanship involved in early printing techniques rather than the mass production capabilities often attributed to him.

72. Humans are interestingly unique in the extent of their lateralization and asymmetry of preference for the right hand. Paleontological evidence (the negative hand prints found in caves worldwide (mostly left hands since the right was used to hold the pigment Steele and Uomini [2005])) shows that this tendency to use the right hand has been with us for at least 40.000 years. This is a further link to the strong reliance on the left hemisphere in the brain, that will be briefly talked about later in the context of the work of [McGilchrist, 2021].

73. The concept of *constraint alignment* suggests a deep relationship between the structure of our descriptive tools and the phenomena they describe. This relationship becomes evident when we consider how different fields have developed specialized languages that match their subject matter: engineering employs precise mathematical formulations that track physical constraints, while poetry uses fluid language structures that resonate with emotional and experiential states. This alignment isn't just related to matching quantity or complexity, it's about aligning the constraints of our descriptive system with the inherent constraints of what we're describing (or to be more precise, the model we develop of what we are trying to describe). For example, quantum mechanics required the development of new mathematical formalisms because classical physics language couldn't adequately map quantum-level constraints. Similarly, the emergence of process philosophy brought with it new linguistic constructions to better track the dynamic, processual nature of reality. Furthermore, consider how programming languages evolve to match the constraints of different computational paradigms, or how legal language develops to correspond to the constraints of social and juridical systems. The effectiveness of constraint alignment suggests a fundamental principle about the relationship between epistemology and ontology, where our most successful descriptions of reality tend to be those whose structural constraints best align with the constraints of the phenomena they describe - more on that in in our discussion on logic.

74. There is a somewhat loose parallel in self-referential statements that cross logical types across hierarchies, like strange loops, the liar's paradox or Gödel's incompleteness theorems. This will be discussed in significantly more detail in Part 2 on Paradox. Why there is still a fundamental difference between a strange loop and living systems will also be shown there, which leads to a re-imagination of Hofstadter's book title and claim "I Am a Strange Loop" to something like "I am a self-constraining system that can symbolically produce a Strange Loop"—not so catchy, sadly.

75. The connection between symmetry and conservation laws, as formalized in Emmy Noether's seminal work, is one of the most profound insights in theoretical physics and mathematics. Her paper, "Invariante Variationsprobleme" (1918), established that every differentiable symmetry of the action of a physical system corresponds to a conservation law. This result not only unified diverse physical phenomena under a common mathematical framework but also demonstrated the deep interplay between symmetry and the fundamental principles governing the universe Noether [1983].

76. Howard Pattee and Robert Rosen shared a close intellectual partnership, spanning two decades from the late 1960s to the late 1980s, centered around understanding the modeling relation and its implications for life and epistemology (see [Pattee, 2007]). Both scholars explored the nature of models as internal representations fundamental to the persistence and evolution of living systems. However, their approaches diverged significantly, creating a complementary dynamic. Rosen, with a formalist perspective, emphasized the logical and relational conditions necessary for life, seeking abstract, mathematically rigorous models that captured the organization underlying biological systems. His focus was on the theoretical limitations of formal systems, such as Gödel's incompleteness and Turing undecidability, and their implications for understanding biological complexity. For Rosen, the genetic system functioned as an internal predictive model, with evolution tied to the proliferation of inequivalent models—a cornerstone of his "anticipatory systems" framework.In contrast, Pattee, grounded in a materialist and experimental physicist's perspective, concentrated on the physical realization of these abstract models. He was particularly interested in the physical constraints and laws that enable symbolic control within organisms, facilitating open-ended evolution. Pattee introduced the concept of non-integrable constraints—mathematical representations of informational processes that are not state-determined—to bridge formal systems and their physical embodiments. The synergy between their perspectives was evident in their shared rejection of Newtonian determinism as adequate for describing life. They instead proposed that complex systems required multiple, inequivalent models to account for phenomena like measurement, control, and prediction. Despite their shared foundations, Rosen's later work in Life Itself marked a departure, advocating for a pure relational biology that eschewed formal and physical models, and critiquing the limits of contemporary science, including physics and evolutionary biology. Pattee found this shift away from their collaborative foundations puzzling and emphasized the value of Rosen's earlier contributions within the broader landscape of biological and physical modeling.This

relationship illuminates the epistemic necessity of integrating formal and material perspectives in understanding the interplay of constraints and laws in life systems, a theme central to holonomic and non-holonomic constraints discussed in this text.

77. While the primary focus of our discussion has been on the foundational aspects of laws and constraints as explored through classical and contemporary theories, it is worth briefly mentioning the contributions of Alicia Juarrero [2023] to the understanding of constraints in complex systems. Juarrero's typology, although not central to our thesis offers an interesting perspective on the multiplicity of constraints that can influence system behaviour. Juarrero distinguishes between context-independent and context-dependent constraints. Context-independent constraints are those that establish background conditions that skew probabilities away from equilibrium, such as gradients, polarities, and processes of repetition or replication. These set the stage for events by making certain outcomes more likely than others, independent of specific system interactions. On the other hand, context-dependent constraints involve the intertwining and mutual conditioning of entities within a system, exemplified by catalysts, feedback loops, and the closure of constraints. These are integral in forming coherent wholes and demonstrating emergent properties that transcend the capabilities of isolated components. Additionally, Juarrero discusses sedimentation and entrenchment as constraints that reinforce and stabilize existing constraints but at the potential cost of reduced system flexibility. Sedimentation serves as a form of memory, preserving influences from the past, while entrenchment solidifies constraints, making them more resistant to change. Lastly, scaffolds, frameworks, templates, and affordances are described as enabling constraints. These do not directly transfer energy but facilitate the completion of processes and guide the realization of outcomes, thus shaping the pathways through which systems evolve.

78. The terms "holonomic" and "non-holonomic" go back to the work of Henrich Herz who introduced them in his work on mechanics [Hertz and von Helmholtz, 1910]. The application of those terms to the theory of modelling complex biological systems was developed by Pattee (see [Pattee and Rączaszek-Leonardi, 2012] for a collection of his essential papers)

79. "Standing there, the edifice rests upon the rocky ground. This repose of the work draws forth from the rock the darkness of its vast yet unforced bearing. Standing there, the edifice withstands the storm raging above it, and in doing so first reveals the storm itself in its force. The gleaming and shining of the stone, seemingly only by grace of the sun, first brings forth into appearance the light of day, the expanse of sky, and the darkness of night. The steadfast rising makes visible the invisible space of air." ([Heidegger, 2012], Translation by author)

80. This pervasive bias is famously diagnosed by Jacques Derrida as the *metaphysics of presence*, the tendency in Western thought to privilege immediate existence and substance over relation and difference [Derrida, 1976]. This critique builds upon Martin Heidegger, who he argues that the western tradition has focused on entities as static objects present-at-hand (*Vorhandenheit*) while neglecting the temporal processes that allow them to appear [Heidegger, 1962]. In the philosophy of science, Alfred North Whitehead identifies a parallel error as the *fallacy of misplaced concreteness*, where abstract spatial configurations are mistaken for concrete reality, obscuring the processual nature of the universe [Whitehead, 1925]. This Western predilection for the tangible is further illuminated by sinologist François Jullien, who contrasts the Greek obsession with static Being against the Chinese philosophical emphasis on *propensity* and the dynamic potential of the void [Jullien, 1995].

81. This concept resonates with the phenomenon of protein misfolding in prion diseases, where the loss of precise constraints on protein conformation leads to a pathological state. The misfolded prion protein (PrPSc) exemplifies how the absence of proper structural exclusion—the failure to maintain a functional configuration—results in the propagation of disease through aberrant molecular interactions. Thus, the biological importance of "absence" or exclusion is evident not only in the generation of functional structures but also in the prevention of pathological ones.

82. All M.C. Escher works © 2025 The M.C. Escher Company - the Netherlands. All rights reserved. Used by permission. `www.mcescher.com`

83. This implies that at any given point in time, one perspective or enacted position has to be taken. This inability to synchronically superimpose two outcomes of an interpretive process will become relevant in the second part of the thesis focusing on paradox, where this will be linked to the difference in processual nature of organisms. You can experience that in a direct sense with the tessellations. The process of focusing on one form inherently excludes the other from being fully perceived. While it is possible to squint or glance in such a way as to perceive both contours simultaneously, the act of concentrating on a specific form—such as the duck—brings it into the foreground of awareness and relegates the fish-form to the background. This shift occurs because focused observation is inherently selective, allowing one interpretation to dominate consciousness momentarily, while the alternative recedes

84. As are all composite perspectives, since no limited being in our universe could possibly hold all of them.

85. Nisbett and Masuda's work is emblematic of broader research on cultural cognition. Nisbett's *The Geography of Thought* explores how cultural differences shape analytical and holistic thinking styles Nisbett [2003], while Masuda and Nisbett's experimental study on underwater scenes demonstrated that Western participants focus on salient objects, whereas East Asians emphasize contextual and relational elements Masuda and Nisbett [2001]. This research aligns with related work by Hofstede, who identified individualism and collectivism as key dimensions of cultural difference in *Culture's Consequences* Hofstede [2001], and Markus and Kitayama, whose seminal paper "Culture and the self" examined how independent and interdependent self-construals influence cognition and motivation Markus and Kitayama [1991]. Hall's *Beyond Culture* provides foundational insights into high-context and low-context communication, further illuminating East-West divides in perception and interaction Hall [1976]. Together, these studies underscore how culture shapes not only what we see but also how we think about and interpret the world around us.

86. The hypothesis that brain asymmetry evolved to solve the computational problem of simultaneous attention—balancing focal *grasping* with global *vigilance*—is currently the dominant framework in comparative neuroscience. The most compelling empirical evidence comes from Rogers et al. [2004], who compared the performance of strongly lateralized chicks against weakly lateralized ones in a dual-task scenario. They found that while lateralized chicks could successfully peck at food (a left-hemisphere discrimination task) while monitoring a simulated predator (a right-hemisphere vigilance task), non-lateralized chicks failed to perform both simultaneously. This suggests that lateralization segregates functionally incompatible neural circuits, thereby increasing parallel processing capacity. For the mathematical modelling of this population-level asymmetry as an "Evolutionarily Stable Strategy" (ESS), see Ghirlanda and Vallortigara [2004]; for a broader review of survival advantages, see Vallortigara and Rogers [2005].

87. For the seminal comparative review establishing lateralization in lower vertebrates (fish, reptiles, and amphibians), see Bisazza et al. [1998]. For the extension of these findings into invertebrates (such as honeybees), see Frasnelli et al. [2012]. These studies collectively argue that brain asymmetry increases cognitive efficiency by allowing parallel processing of contradictory tasks—categorizing objects for feeding (typically a left-hemisphere bias) and monitoring spatial novelty for survival (typically a right-hemisphere bias).

88. See McGilchrist [2009, 2021]. It is important to note that McGilchrist's broad mapping of cultural history onto hemispheric differences has faced criticism for potential reductionism and for oversimplifying the complex, distributed nature of neural networks. For a critical scientific review distinguishing the valid asymmetries from the pop-psychology myths of "left vs. right" dominance, see Corballis [2014]. However, even if one treats McGilchrist's hemispheric hypothesis as

a metaphorical heuristic rather than a strict biological determinism, it offers a helpful framework for understanding the tension between two competing modes of attention, namely, the focal and the global.

89. It is important to acknowledge that there are several other examples adjacent to what we have investigated that, while not explored in depth here, are nonetheless crucial to understanding the role of absence and incompleteness in various domains. In mathematics, the concept of zero serves as a paradigmatic example of a thing that is not, but whose existence allows for the birth of modern mathematics [Kaplan, 2000]. Various spiritual and religious traditions, such as negative theology and apophatic knowledge, emphasize the importance of understanding the divine or ultimate reality through negation, by describing what it is not rather than what it is [Turner, 1995]. The idea of "wu wei" in Taoist philosophy emphasizes the power of non-action and letting things unfold naturally, rather than forcing them through direct intervention [Lao-tzŭ, 1993]. In Japanese aesthetics, the concept of "ma" refers to the empty space or interval between objects, and is considered essential to the overall composition and meaning of an artwork [Davis, 2020]. On a personal level, the role of absence and loss in shaping identity and narrative is a rich area of exploration in literature, psychology, and philosophy [Freud, 2005]. The importance of rest and recovery in physical and mental health highlights how the absence of activity is crucial to the body and mind's ability to heal and regenerate. These often lesser-explored topics, spanning from mathematics to the intimate domain of personal experience, underscore the potential for further investigation into the generative power of absence and incompleteness across diverse fields.

90. Kauffman [2020]

91. Montévil and Mossio [2015]

92. For the comprehensive introduction of teleodynamics and the relationship between homeodynamics, morphodynamics, and teleodynamics, see Deacon [2011].

93. For a more extensive discussion see [Deacon, 2011, 2021]

94. There has been some ongoing debate and accusations of academic dishonesty from Alicia Juarrero against Terrence Deacon, alleging that he plagiarized her work [Riddell, 2013]. While these claims seem unfounded after reviewing the primary literature, there is a tendency not to cite other thinkers widely in the study of the origins of living systems. E.g.: Moreno Mossio did not cite Deacon in

their 2015 work on the closure of constraints [Moreno and Mossio, 2015a]. My personal take is that the research field of the origin of life, by its very nature, bridges physics, chemistry, and biology, aiming to answer foundational questions that defy categorization within a single discipline. This interdisciplinary character, while necessary, makes it challenging for researchers to organize their work in a cohesive manner. Many thinkers approach the subject from specific disciplinary traditions, such as systems theory, biochemistry, or evolutionary biology, and often devise theories that echo others in structural ways without immediately recognizing these overlaps. Consequently, this has led to a field where significant works, such as those on hypercycles, the chemoton, autopoiesis, and autocatalytic sets, exhibit limited cross-referencing and few explicit connections. This lack of integration is a symptom of the broader difficulty in synthesizing perspectives across such a vast and diverse intellectual landscape, which I believe contributes to the fragmented nature of the field. The clear upside is dynamic is, that it provides the research community with diverse models trying to do similar work. Hence interesting structural isomorphisms between models surface that can point to robust and relevant features (or collective biases, of course).

95. I have discussed the implementability of the autogen with chemists and have received the response that all the "parts" are there, however, finding one system that actually works would likely require years, luck, and a small army of PhD students and post-docs.

96. The oldest dated intact rocks on Earth are found in the Acasta Gneiss (4.031 ±0.003 billion years ago) Bowring and Williams [1999], with even older individual zircon crystals from the Jack Hills of Western Australia (4.404 ±0.008 billion years ago) Wilde et al. [2001], Valley et al. [2014]. The earliest evidence for life appears in microbialites from the Nuvvuagittuq Greenstone Belt (at least 3.77 and possibly 4.32 billion years ago) Dodd et al. [2017], with additional evidence from biogenic graphite in Greenland (3.7 billion years ago) and the Pilbara region of Australia (3.49 billion years ago). This suggests that life emerged relatively early in Earth's history, with some evidence pointing to its existence even before the formation of the oldest surviving rocks.

97. Terrence Deacon discusses the work cycles of Autogenic systems in multiple texts, see Deacon [2011, 2021]

98. Not all new species are however necessarily beneficial. Some may be lethal or degrade system efficiency, underscoring that autogenic systems operate on a precarious balance and can easily fail if disruptions are too severe.

99. The concept of proto-evolution in autogenic systems merits careful consideration in understanding life's emergence. While lacking the sophisticated machinery of modern cells, autogens exhibit primitive evolutionary features through their unique dynamics. Their ability to maintain organizational patterns across cycles represents a rudimentary form of inheritance, though accomplished through system dynamics rather than dedicated information storage. Furthermore, they demonstrate selective incorporation of beneficial molecules while excluding detrimental ones—a primitive form of selection, albeit far less efficient than Darwinian processes. When disrupted, autogens can give rise to new instances maintaining the original organizational constraints, constituting an accidental form of reproduction.

However, several critical elements separate this proto-evolution from true Darwinian evolution. Modern life relies on RNA/DNA as dedicated carriers of heritable information, while autogens lack any such system for reliable information storage and transmission. Their accidental propagation provides no mechanism for ensuring accurate copying of constraints, severely limiting the accumulation of beneficial changes. As systems become more complex, they encounter a combinatorial catastrophe, where the likelihood of conflicting molecular interactions increases without templates to manage these interactions.

The transition to evolutionary systems likely involved a series of crucial developments. Initially, nucleotide-like molecules might have emerged serving dual roles, perhaps as energy carriers or catalytic aids, while gradually developing template-like properties. This would enable a progressive offloading of system constraints from direct molecular interactions to more stable, reproducible patterns. As these templates evolved copying mechanisms, the first true hereditary systems would emerge, allowing faithful transmission of organizational patterns across generations. The integration of such template-based inheritance with autogenic self-maintenance would bridge the gap between simple self-organization and biological evolution.

This progression illuminates how chemical systems might develop evolutionary capabilities, though the precise historical pathway remains speculative. The contribution of autogenic systems lies not in their status as primitive living things, but in demonstrating fundamental principles upon which life could build, particularly in the crucial transition from pure chemistry to primitive biology.

100. Viruses provide an interesting parallel to autogenic systems, particularly in their fundamental organization. Like autogens, viruses exhibit self-assembly of capsid structures, maintain specific molecular organizations, and cycle between assembled and disassembled states. Both systems also demonstrate a form of self-containment through their capsid structures and show capabilities for maintaining their organizational patterns across cycles. These similarities make viruses appear, at first glance, like a natural example of autogenic-like systems in nature.

However, this comparison ultimately reveals more about what autogens lack than what they share with viruses. Modern viruses are far more sophisticated systems that possess crucial features absent in autogens: they contain RNA or DNA-based inheritance systems and rely on hijacking the complex biosynthetic machinery of host cells for reproduction. Unlike the accidental propagation of autogens, viruses have evolved precise regulatory mechanisms for replication and inheritance. The organizational complexity of even simple viruses illustrates a crucial distinction from autogenic systems. While autogenic systems exhibit basic self-constraining properties, viruses demonstrate sophisticated gene regulatory networks where multiple genes interact in hierarchical patterns of mutual regulation. These networks create a level of organizational stability that transcends simple information storage and transmission. Viral gene regulation involves complex co-stabilizing mechanisms where networks of genes regulate each other's expression in intricate feedback loops. This comparison thus helps us understand what's missing for "true" evolution, reliable information storage and transmission systems, accurate replication mechanisms, and the capacity for systematic inheritance of variations.

101. A parallel can be drawn with Formula One driving technique. In high-performance racing, drivers often apply light throttle while initiating braking—primarily to stabilize the car's dynamics and maintain turbo pressure. This trade-off, like the cellular example, sacrifices a small amount of energy to gain disproportionately fine control. Both systems—biochemical and mechanical—highlight how continuous, opposing processes enable rapid responsiveness and stability, despite a slight energetic cost.

102. Different membrane components turn over at distinct rates: while some structural lipids persist for days, certain membrane proteins are replaced every few hours.

103. Protein degradation rates typically match synthesis rates, maintaining stable concentrations despite continuous turnover. This dynamic equilibrium enables rapid adaptation while preserving cellular function.

104. see Gell-Mann [1994], Holland [1992], Carmichael and Hadžikadić [2019]

105. See [Gunderson and Holling, 2002]

106. Regarding the origin of the adaptive cycle and the choice of variables Holling writes: *The concept of the adaptive cycle and the observation that key variables operate at distinct and separate scales*

*emerged from a synthesis of empirical studies; it was not deduced from first principles. Theory did not determine what observations were made; rather, observations of nature and the practice of ecological management dictated theory.*Gunderson and Holling [2002]

107. Levin [1992]

108. Buchakjian and Kornbluth [2010]

109. Morgan [2007]

110. The comparison with the autogenic work-cycle is particularly striking because the autogen remains an entirely theoretical model with no empirical evidence whatsoever. Unlike the adaptive cycle, which emerged from observations of real-world systems, the autogenic cycle is a purely conceptual construct designed to explore how chemical systems might have transitioned to living systems; we have never observed such a system in nature, and it remains an untested hypothesis. The argument is that as life evolved from its posited autogenic beginnings, it would have incorporated additional molecules during reorganization phases, developing adaptive capabilities to survive in diverse environments. This process would have continued, eventually integrating more complex mechanisms like RNA and DNA replication. The transition from autogenic systems to modern cells involves a profound shift in how molecular constraints are managed. In autogenic systems, constraints are inherent in the direct chemical interactions between molecules – their catalytic relationships and self-assembly properties are determined by immediate physical affinities. This intrinsic constraint management faces a severe limitation: as the number of interacting molecules increases, the system encounters a combinatorial explosion of possible interactions, most of which are non-productive. The solution emerged through constraint offloading – the transfer of constraints from direct molecular interactions to template-based systems, primarily nucleic acids. Templates like DNA and RNA provide a stable substrate for storing and managing constraints, their linear structure allowing for sequence-specific information that can guide molecular interactions with unprecedented precision. This template-based organization transforms the system's fundamental nature. Instead of relying solely on direct molecular affinities, the system develops semiotic functionality – the ability to use molecular sequences as representations that "stand for" or "encode" particular constraints. The emergence of genetic regulatory systems exemplifies this process, where nucleotide sequences not only encode functional proteins but also control their timing and location of expression. Modern cells thus represent a profound transformation in how biological organization is achieved and maintained. They have evolved elaborate template-based systems for managing molecular constraints, enabling them to achieve levels of complexity and adaptability far beyond what's possible

in purely autogenic systems. This transition fundamentally altered the nature of biological evolution, enabling the emergence of increasingly complex forms of life. Evolution has since overwritten these early, less stable forms with more robust systems, making the discovery of a simple, "early" autogenic system on Earth today highly improbable.

The adaptive cycle, therefore, holds particular significance for our exploration. It represents an empirically derived model of cyclical patterns observed across diverse biological and ecological systems, spanning multiple scales. This scale-invariant nature of the adaptive cycle resonates deeply with our investigation of the autogenic cycle. While the autogenic cycle remains a theoretical construct for life's origins, the adaptive cycle provides tangible evidence of similar cyclical dynamics in observable, complex systems. This empirical grounding lends credence to the idea that fundamental cyclical processes may indeed underpin life and complex systems at all levels. Reflecting on both, the parallel contours in their described cyclical processes of growth, disruption, and reorganization become apparent, albeit operating at different scales (proto-molecular for autogenic, various complex system scales for adaptive). This supports the endeavor to graft these models together to yield a deeper understanding of fundamental principles governing living systems.

111. The idea echoes von Uexküll's Beobachtungsbegriff: what counts as a phase is partly an artefact of the observer's sampling rate and the variables tracked. Nothing in the formalism prevents stacking many out-of-phase micro-cycles to obtain a macro-cycle, or decomposing a macro-cycle into finer, non-synchronous threads.

112. Jim Brown makes precisely this point from the standpoint of desert ecosystems, he argues: *I remain skeptical about Buzz's [Holling's] ideas about "Panarchy" and the "figure 8" [adaptive cycle] as a general model. I agree that some systems may (self-?)organize until they reach some critical state and then have a revolutionary, resetting collapse. But I think that this is far from universal. Many forest ecosystems with regular fire conform to this model, but I am not convinced that the deserts that I study—or many other systems—do this. I think that they continually reorganize in response to internal and external perturbations, but more or less gradually and without collapse or revolution. I am concerned that Buzz wants the community to buy into his complete worldview. That view is very creative, but I don't think that it will stand the test of time as an integrated whole. I also feel strongly that if he wants us to buy in and start pursuing these ideas, he has to make concepts like resilience, Panarchy, innovation, etc., more precise and operational—i.e., expressed in mathematical terms and much more firmly rooted in the fundamentals of physics, chemistry, and biology*Allen and Holling [2008].

113. Shields [2006]

114. See Kohn [1997] for a discussion of how the Paradox is apperently resolved by posing a dialogue between four distinct archetypes: the A'mel (builder/worker) who seeks meaning through creation and material achievement, the Hakham (wise man) who pursues understanding through philosophy, the Neheneh (pleasure-seeker) who advocates for enjoying life's pleasures, and the Yare' (God-fearer) who finds purpose through religious devotion. Rather than contradicting itself, the text presents these competing viewpoints in dialogue, ultimately suggesting that the Yare's perspective - fearing God and keeping His commandments - provides the most coherent answer to the book's central question about life's meaning.

115. The biological principle of opponent processing provides a compelling example of how seemingly contradictory processes can be integrated to maintain adaptive function. In neural systems, opponent processing occurs when two mechanisms work in opposition while remaining functionally interdependent - like the sympathetic and parasympathetic nervous systems that regulate arousal levels. Rather than representing a design flaw, this opposition creates a dynamic balance that allows organisms to maintain optimal states through continuous self-organizing adjustment. As detailed in Vervaeke's work on relevance realization, opponent processing emerges as a fundamental principle of biological and cognitive organization, where the tension between opposing forces (such as efficiency versus resilience) generates adaptive behavior through their very opposition. This mirrors the broader pattern we see in cyclical synthesis, where apparent contradictions across time reveal themselves as necessary complementarities in maintaining complex living systems. See Vervaeke [2024] for an extensive discussion of opponent processing in cognitive systems.

116. For a good overview of the subject the Stanford Encyclopedia of Philosophy entry on *Teleological Notions in Biological Systems* is recommended [Allen and Neal, 2020]

117. see Aristotle [1979]

118. The suggestion to more aptly translate the Aristotelian causes as *becauses* comes from Hofmeyr [2021]

119. As already indicated, the school of thinkers around Robert Rosen with concepts of *closure to efficient causation* have been using these ideas very recently to address the same questions we are thinking about here, see [Rosen, 1991, Hofmeyr, 2021]

120. [McDonough, 2020]

121. In the *General Scholium* to the *Principia*, Newton famously declared, "I have not been able to discover the cause of those properties of gravity from phenomena; and I frame no hypotheses." Yet in the same text he insisted that "This most beautiful System of the Sun, Planets, and Comets, could only proceed from the counsel and dominion of an intelligent and powerful Being." This juxtaposition shows how Newton stripped teleology out of natural philosophy itself, restricting science to mathematical laws, while relocating purpose to the divine order that underpins the cosmos .[Newton, 1999]

122. see the foreword by Michael Ghiselin in Darwin [1984]. While Darwin did use teleological language his work can be interpreted in a way that it helps get rid of teleology. For our discussion we will follow this interpretation

123. For a recent, similar investigation that focuses on attempts to naturalize teleology through cybernetics, Jonas and the OA, see Sachs [2023]

124. For more on this, see Wiener [1948], Ashby [2015], McCulloch [1965], von Foerster [2003], and Pickering [2010].

125. As Lawrence Frank argued in his 1948 foreword to a special issue of the Annals of the New York Academy of Sciences, cybernetics represented *an attempt to escape from these older mechanistic formulations that now appear inadequate, and to provide new and more fruitful conceptions and more effective methodologies for studying self-regulating processes, self-orientating systems and organisms, and self-directing personalities. But these new concepts carry no psychic or vitalistic assumptions, nor do they imply that any mysterious, supernatural powers or psychic forces or final causes are operating the system or guiding the organism-personality. The idea of purposive behavior is not a regressive movement to an earlier stage in the history of ideas, but a forward movement toward a more effective conception of the problems we face today.* - marking a decisive break from both traditional mechanistic and vitalistic approaches. See Frank [1948].

126. He writes that cybernetics is: *the newer study of automata, whether in the metal or in the flesh, is a branch of communication engineering, and its cardinal notions are those of message, amount of disturbance or 'noise'—a term taken over from the telephone engineer—quantity of information, coding technique, and the like. In such a theory, we deal with automata effectively coupled to the external world, not merely by their energy flow, their metabolism, but also by a flow of impressions, of incoming messages, and of the actions of outgoing messages.*[Wiener, 1948]

127. Jonas developed this critique most fully in his 1952 essay *A Critique of Cybernetics* where he argued that cybernetic systems fundamentally lack the intrinsic purpose and self-concern characteristic of living things. See JONAS [1953].

128. The distinction between cybernetic feedback mechanisms and autopoietic self-production is crucial. As Sachs [2023] notes, *the circular causality is not informational but constitutional. Thus, while the organizational approach shares with cybernetics the emphasis on circular causation or causal loops as the key to naturalizing teleology, it is a very different kind of circular causation.*

129. This concept of operational closure has its roots in the work of Maturana and Varela. As Weber and Varela [2002] explain, *The living system must distinguish itself from its environment while at the same time maintaining its coupling; this linkage cannot be detached since it is against this very environment that the organism's self-distinction takes place.*

130. see Weber and Varela [2002]

131. see [Montévil and Mossio, 2015, Moreno and Mossio, 2015b, Mossio and Bich, 2017]

132. See García-Valdecasas [2022]. It's worth noting that this criticism might depend on how one interprets the closure of constraints, whether it's seen as a static pattern or as inherently including the dynamic aspects of biological organization. Different readings of the OA might lead to different assessments of this limitation.

133. This critique is voiced by Sachs [2023] in a footnote to his article where he writes: *Arguably the organizational approach at best specifies the kind of nonlinear or circular causality that grounds teleology, but not how teleology itself emerges from non-teleological systems. It is here that I think Deacon's (2012) distinctions between homeodynamics, morphodynamics, and teleodynamics should be considered an advance upon the organizational approach.*

134. Rosen's distinction between conditional and teleological explanations, articulated in *Life Itself*, provides a crucial lens for evaluating the limitations of the Organizational Approach. He demonstrates that formal systems typically only support three *traditional* causal categories operating within what he calls *formal time*: material cause (axioms), efficient cause (production rules), and formal cause (specific algorithms) [Rosen, 1991]. Final causation, however, requires additional

modes of entailment not typically present in formal systems. As Rosen notes, *in purely formal terms, a concept of final causation requires modes of entailment that are simply not generally present in formalisms.* Most crucially, final causation appears to violate the forward flow of formal time - it *cannot fit within the same temporal sequence in which the other causal categories harmoniously operate.* This temporal anomaly points to a fundamental limitation in approaches that attempt to reduce teleology to networks of mutual constraints: they risk mistaking spatial patterns of dependency for genuine purposiveness. The OA effectively describes how various processes conditionally depend on each other in space, but as Rosen's analysis suggests, genuine teleology requires something more - namely, modes of organization that can operate across time rather than just space. This distinction that Rosen highlighted decades before OA's development remains pertinent: without incorporating how constraints are coordinated across time (not just space), explanations based purely on organizational closure may capture only conditional dependencies rather than true biological purposiveness. As he writes, *Final cause clearly cannot fit within the same temporal sequence in which the other causal categories harmoniously operate*, suggesting that genuine teleology requires us to understand how living systems integrate opposing processes across time rather than just resolving their dependencies in space.

135. see Weber and Varela [2002]

136. see DiFrisco and Mossio [2020]

137. [Laplace, 1951]

138. It's worth noting that while Newtonian physics appears explicitly dynamic in its treatment of motion and forces, it paradoxically embodies a deeply static worldview. Consider Newton's *physics in a box*: while the particles within may move and interact according to mechanical laws, the organizational principles, the laws themselves and the nature of the interactions, remain frozen in time. There is motion within the system, yes, but the system itself, its fundamental structure and possibilities, remains unchanging. This is precisely what makes it synchronic: despite describing dynamic processes, it assumes a timeless, predetermined framework within which all these processes must occur. Any Newtonian system, no matter how complex its motions, is essentially *frozen* in its organizational possibilities from the start.

139. see Habermas [2019]

140. For further exploration of the ideas discussed, see Whitehead's works on process philosophy, particularly *Process and Reality* Whitehead [2010] and *Science and the Modern World* Whitehead [1967]. Bergson's concept of *durée* (duration) is elaborated in *Time and Free Will: An Essay on the Immediate Data of Consciousness* Bergson [2001]. Heidegger's temporal analysis can be found in *Sein und Zeit* Heidegger and von Herrmann [1977], while Merleau-Ponty examines the temporal nature of perception in *Phänomenologie der Wahrnehmung* Merleau-Ponty and Boehm [1966]. For a discussion of becoming and the critique of static being, see Deleuze's *Difference and Repetition* Deleuze and Patton [2004]. Also see Lorenz [1963], Waddington [1942, 1957], Smolin [2013]

141. While process thinking does not necessarily entail perspective in all its forms (as can be argued with Whitehead's non-perspectivist metaphysics), the specific conception of process thinking developed through constraint cycles necessarily leads to a form of perspectivism. This necessity emerges from the fundamental properties of constraint cycles: their hierarchical self-constraint and environmental coupling create a system that must maintain its organizational integrity while adapting to environmental changes. This maintenance occurs through self-constraining processes where the system actively works against its own dissolution through dynamically interlocked processes. The emergence of a clear beneficiary (the system whose integrity is maintained) and its intrinsic normative dimension (what the system must do to maintain itself) creates a reference point. This point, can constitute something akin to a proto-perspective unique to each constraint cycle. Thus, while not all process thinking entails perspectivism, the specific form developed here through constraint cycles necessarily generates perspectival beings due to the intrinsic requirements of maintaining organizational coherence through time.

142. [Peirce et al., 1931](1.351)

143. or cycles-depending on how many co-emerged or managed to marry themselves into the lineages we see currently expressed.

144. Although her derivation of that inseparability comes from quantum physics and the integral role the observer plays in any such system, see Barad [2007]

145. As defined by Maturana and Varela [2012]: "*We speak of structural coupling whenever there is a history of recurrent interactions leading to the structural congruence between two (or more) systems.*" This is precisely the relation that ontology and epistemology must have in an organism adapted to a given environment.

146. Philip K. Dick once quipped, *Reality is that which doesn't go away when you stop believing in it.* I will not linger here on the perennial debate over how to define reality, that very effort would simply reprise the book's core theme, a subject recursively trying to describe the field in which it is entangled. It is enough to acknowledge that something continually constrains our interpretations without ever being exhausted by them. We live, inevitably, in the in-between of the purely subjective and the purely objective.

147. The concept of "corrupted replay" is central to the "overfitted brain hypothesis," which posits a specific function for the bizarre and hallucinatory nature of dreams. While memory consolidation is often associated with the relatively faithful "replay" of waking experiences to strengthen neural traces, this hypothesis suggests that dreams serve a complementary function analogous to regularization techniques in machine learning. By presenting the brain with novel, varied, and often strange sensory and narrative scenarios—a form of stochastic data augmentation—dreams may actively counteract the tendency of the brain to "overfit" to the statistically repetitive and often mundane experiences of daily life. This process is thought to enhance cognitive flexibility and the ability to generalize knowledge to new situations, rather than simply reinforcing specific memories. See Hoel [2021]

148. During REM sleep, high acetylcholine levels suppress bottom-up sensory input while amplifying internal feedback processes, creating stochastic neural activity that simulates warped, hallucinatory experiences. See Nir and Tononi [2010]

149. The sparseness and hallucinatory nature of dreams differs markedly from waking experiences, suggesting they serve functions beyond just memory consolidation. This is supported by evidence that dream-related neuronal activity often involves novel patterns rather than simple replay of previous sequences. See Gupta et al. [2010]

150. This is evidenced by behavioural studies showing that sleep improves performance on cognitive tasks requiring generalization but has limited effects on tasks relying solely on memorization, supporting the distinction between memory consolidation and generalization enhancement. See Wagner et al. [2004]

151. This oversight stems from the historical influence of Malthusian thinking on evolutionary theory, which emphasized resource limitation and competition while downplaying the generative aspects of biological processes. Malthus of course was a central inspiration for Darwin: "In October 1838,

that is fifteen months after I had begun my systematic enquiry, I happened to read for amusement Malthus on Population, and being well prepared to appreciate the struggle for existence which everywhere goes on from long-continued observation of the habits of animals and plants, it at once struck me that under these circumstances favourable variations would tend to be preserved, and unfavourable ones to be destroyed. The result of this would be the formation of a new species." Darwin [1887]

152. The tension between these processes parallels a deeper philosophical question about the relationship between constraint and creativity in biological systems. Selection acts as a constraint, narrowing possibilities, while generation expands them. Both are necessary for evolution to occur.

153. Recent work in evolutionary developmental biology suggests that these generative processes are far more structured and constrained than previously thought, with variation itself being shaped by developmental and systemic factors. For example, see Uller et al. [2018]

154. The temporal aspect of this coordination is crucial. These processes must occur in a particular sequence and rhythm, much like the sleep-wake cycle discussed earlier. This temporal organization appears to be a fundamental feature of biological systems. See Goodwin [1963].

155. Recent research has revealed that this process of neural selection involves multiple waves of generation and selection that are precisely coordinated with different developmental stages, suggesting a more sophisticated temporal organization than previously recognized.

156. See Ohno [1970]. Also for an analysis of how self-referential feedback loops in gene regulatory networks can mimic logical paradoxes (e.g., the liar paradox) when viewed statically, but resolve into productive dynamic behaviors through temporal coordination, see Isalan [2009]. This provides a molecular-level illustration of how apparent contradictions in network topology enable adaptive evolution through cyclical synthesis.

157. Thompson's framework builds on Mary Douglas's grid-group typology but extends it to explicitly consider the dynamic relationships between different social solidarities. See Douglas [1970], Thompson et al. [1990]

158. This understanding has significant political implications that connect cyclical synthesis to deliberative democratic theory. The framework strongly argues for institutional designs that can accommodate and coordinate between different organizational logics rather than imposing any single mode as dominant. This aligns with metamodern approaches to governance, such as those explored by Hanzi Freinacht, and with Luke Kemp's analysis in *Goliath's Curse* of how thriving societies require deliberative coordination between differences Freinacht [2017, 2019], Kemp [2025]. What this suggests is that effective social-ecological systems may need to maintain institutional space for the cyclical expression of different organizational approaches, resisting the temptation to optimize for any single logic. Rather than seeking to eliminate competing modes of organization—viewing hierarchy as necessarily oppressive, individualism as inherently destructive, or egalitarianism as inevitably inefficient—the task becomes one of finding meta-institutional structures that enable their productive temporal coordination. This doesn't mean all modes are equally appropriate at all times, but rather that the capacity to move between them, guided by deliberative processes that themselves embody egalitarian principles, may be essential for long-term societal adaptability and resilience.

159. While this section focuses on logic as structures of reasoning, it's important to acknowledge the diverse foundational interpretations of logic in philosophical discourse. These interpretations can be broadly categorized into four main conceptions: (1) The mathematical study of formal languages and their properties through fields like proof theory and model theory, examining systems such as predicate calculus and modal logics; (2) The analysis of valid inference patterns and logical consequence, focusing on how conclusions necessarily follow from premises based on formal structure rather than specific content; (3) The investigation of logical truths as the most general truths that underlie all other sciences - propositions that hold purely in virtue of their logical form; and (4) The Kantian approach studying the fundamental forms of judgment and thought, examining how we structure our thinking independent of specific content. While these interpretations often overlap and inform each other (for instance, logical truths can be understood as conclusions that follow from empty premise sets), they represent distinct philosophical perspectives on the nature and purpose of logic. For a comprehensive discussion of these interpretations and their interrelations, see Hofweber [2023]

160. The term *constraint alignment* is chosen over alternatives like *structure mirroring* to emphasize a specific aspect of effective modelling. While models certainly represent structure, their explanatory power often derives more fundamentally from capturing the constraints that govern a system's dynamics and possibilities. Constraints represent the specific reductions in degrees of freedom, the boundaries, rules, dependencies, or limitations that channel the system's behaviour away from randomness and towards organized patterns. Consider the bridge example: a useful model doesn't just replicate its structure, but aligns with the physical constraints (material stress limits, force distri-

bution laws) that determine its crucial function, how much weight it can bear without collapsing. It captures the limits on the system's degrees of freedom, specifically, the freedom to break. This principle extends beyond physical systems. Modelling a biological process involves tracking biochemical constraints (enzyme specificity, pathway regulation), modelling a social system involves capturing constraints like norms or laws. In each case, the model's utility lies in representing the key limitations that shape the system's actual behaviour out of its potential possibilities. Focusing on constraint alignment highlights this active process of limitation and enablement inherent in organized systems, offering a more dynamic and functionally relevant perspective than focusing solely on static structure.

161. This division is a heuristic, not a rigid categorization. These traditions are not monolithic, they contain immense internal diversity, and there are many points of overlap and cross-pollination between them.

162. The survey presented here is necessarily selective. The landscape of non-classical logic is vast, and several sophisticated frameworks approach the dynamics of living systems closely. Three in particular warrant mention for their structural resonance with Cyclical Synthesis: *Non-Monotonic Logic* [Reiter, 1980, McCarthy, 1980], which allows for the retraction of conclusions (defeasibility) and thus models how organisms must update behavioral commitments in changing contexts; *Linear Logic* [Girard, 1987], which treats logical truths as finite, consumable resources rather than eternal facts, offering a syntax remarkably suited to metabolic economy; and *Polycontextural Logic* [Günther, 1979], a cybernetic framework that explicitly models living systems as distributing truth-values across multiple, interacting logical domains.

However, a crucial distinction regarding genealogy remains. These frameworks generally arose from the requirements of computation, proof theory, or abstract cybernetics. The *logic of life* proposed in this text distinguishes itself by its origin: it is derived directly from the thermodynamic and morphological constraints of physical self-organization. It seeks to formalize not how we might compute information about life, but how living matter itself maintains coherence through the temporal coordination of opposing physical constraints.

163. Classical logic's commitment to bivalence (that every proposition must be either true or false) and monotonicity (that adding premises cannot invalidate previous conclusions) reflects its underlying assumption of a stable, unchanging reality. These properties make it particularly suited for mathematical reasoning but create challenges for modeling dynamic systems. Consider how living systems often exhibit non-monotonic behavior - new information can invalidate previous conclusions, and the addition of new elements can fundamentally alter system dynamics. This tension

between classical logic's rigid framework and the fluid nature of biological systems points to why alternative logical frameworks became necessary.

164. The principle of explosion (*ex contradictione quodlibet*) represents more than just a technical feature of classical logic. It embodies a particular view of contradiction itself. In formal terms, it follows from the rules of natural deduction: from a contradiction ($P \wedge \neg P$), one can derive any proposition Q through the following steps:

1. From ($P \wedge \neg P$), derive P [conjunction elimination]
2. From P, derive ($P \vee Q$) [disjunction introduction]
3. From ($P \wedge \neg P$), derive $\neg P$ [conjunction elimination]
4. From ($P \vee Q$) and $\neg P$, derive Q [disjunctive syllogism]

This seemingly result has implications for how we think about systems that appear to contain contradictions. Either the contradictions must be eliminated, or we need logical frameworks that can contain them without trivializing the system.

165. While Cyclical Synthesis describes a biological dynamic, it can be rendered as a conservative embedding in standard temporal/hybrid semantics. In such a model, the Law of Non-Contradiction holds at every state s (i.e., $s \not\models A \wedge \neg A$), preserving local consistency. However, the system enforces phase-separated constraints, such as $G(\text{Phase}(P_i) \rightarrow A) \wedge G(\text{Phase}(P_j) \rightarrow \neg A)$, where $i \neq j$. The identity of the system is not fixed by the truth of any instantaneous predicate, but by an invariant $G(I)$ preserved across the cycle. This is one convenient formalization and others are possible. Its purpose here is only to show that CS can preserve classical reasoning at instants without requiring paraconsistency.

166. [Garson, 2024]

167. see Kripke [1959, 1963]

168. The formal apparatus of modal logic consists of adding these new operators to the classical framework. The basic modal operator $\Box$ (necessary) can be prefixed to any well-formed formula to create a new formula expressing a modal claim. Its dual operator $\Diamond$ (possible) is typically defined as

$$\Diamond P \equiv \neg\Box\neg P$$

capturing the intuition that something is possible if its negation is not necessary. Different modal systems are then characterized by different axioms governing these operators. The minimal normal modal logic K includes all classical tautologies plus the distribution axiom

$$\Box(P \to Q) \to (\Box P \to \Box Q).$$

Stronger systems add axioms like

$$\Box P \to P \quad \text{(what is necessary is actual)},$$

$$\Box P \to \Box\Box P \quad \text{(if something is necessary, it's necessarily necessary)},$$

$$\Diamond P \to \Box\Diamond P \quad \text{(if something is possible, it's necessarily possible)}.$$

.

169. Zadeh [1965]

170. Goguen [1969]

171. Hájek [2013]

172. Prior [1967]

173. Prior [2003]

174. Pnueli [1977]

175. Clarke et al. [1986]

176. Emerson [1990]

177. Jaskowski [1948], Da Costa [1974]

178. see Priest [2007]

179. Priest [1979]

180. Anderson et al. [2017]

181. This characterization of the analytic-continental divide necessarily simplifies complex historical and philosophical developments. The distinction itself has been challenged and problematized by scholars. However, it serves here as a useful heuristic for understanding different approaches to logical structure and reasoning.

182. Hegel et al. [1978]

183. Gadamer [1976]

184. Deleuze [1994]

185. Bryant [2008]

186. Somers-Hall [2012]

187. Whitehead and Russell [1910]

188. Whitehead [1978]

189. Theodor Adorno's critique of dialectical synthesis offers an important perspective on these issues. In *Negative Dialectics* [Adorno, 1973] rejected the idea of final synthetic resolution, arguing that forced synthesis often sacrifices the particular and non-identical, potentially covering up real contradictions and suffering. His *negative dialectics* deliberately holds contradictions open without resolving them into higher identity. This represents a self-critical turn in continental logic, emphasizing what systematic thought cannot fully capture. While we won't fully explore this trajectory, it highlights important limitations in any attempt at complete logical synthesis.

190. Peirce [1867]

191. Peirce [1878]

192. Peirce [1877]

193. Dewey [1938]

194. Brier [2008]

195. Hoffmeyer [2007]

196. Sebeok [1968]

197. Sebeok [2001]

198. Hoffmeyer [2008]

199. My reservation for following Varela's Calculus for Self-Reference further, centres on the ontological status of the *autonomous state*. While Varela positions this state as a third element alongside Spencer Brown's marked and unmarked states within the calculus of indications, it can be argued that autonomy is not just a third state in this triad, but rather a precondition for the very possibility of distinction that underpins the entire system. Spencer Brown's original calculus begins with the act of distinction itself, positing the unmarked void as primary, and the marked state as emerging from the drawing of a boundary, the making of a difference. This act of distinction is logically prior to the states it delineates.

200. This relationship between paradox and time finds its formal proof in the seminal work of Spencer-Brown [1969], *Laws of Form*. Spencer-Brown sought to reconstruct logic not from propositional truths, but from the primary act of making a distinction. Cleaving a space into an *inside* and an *outside*. His critical insight concerns what he termed *re-entry*, where a form enters into its own indication. In classical logic, a self-negating equation is treated as a static contradiction. However, Spencer-Brown demonstrated that if we allow the calculation to proceed, the system automatically introduces a new dimension: *time*. The indeterminate value becomes an oscillation, a frequency rather than a state. This offers a formal corollary to Cyclical Synthesis, where logical paradoxes are only paradoxical when viewed synchronically (as a snapshot), but unfold into viable temporal dynamics when viewed diachronically.

However, a crucial limitation remains in applying Spencer-Brown's logic directly to the origins of life. His system is fundamentally idealist. It begins with the axiom of drawing a distinction, which tacitly presupposes the existence of an entity or observer capable of performing that act. His calculus describes the *behavior* of a boundary once established but cannot explain the *genesis* of the boundary-maker itself. In the material world, distinctions are not drawn by abstract fiat but must be generated by physical processes. This brings us back to the necessity of the thermodynamic constraint cycles discussed in earlier chapters, they provide the physical mechanism that *draws* the first distinction (self/non-self) against the equilibrium environment. Thus, while Spencer-Brown provides the syntax for how a self-referential boundary operates in time, the thermodynamic theory of life is required to explain how such a boundary, and the *observer* it constitutes,can emerge from the void in the first place.

201. This text was generated with Anthropic's 3.5 Sonnet, a model known for its flowery writing style, other models would have generated significantly more academic sounding text.

202. Proponents of the scaling hypothesis argue that continually increasing model size and training data leads to quantitative improvements that eventually trigger qualitative leaps in capability.

Kaplan et al. [2020] demonstrated that language model performance follows smooth power-law improvements as model parameters, data, and compute grow, fueling optimism that sufficient scale could yield something like general intelligence.

203. Savage [1972] introduced the distinction in *The Foundations of Statistics*. A *small world* is any decision problem whose acts, states, and consequences can be fully listed, so one can analyse it in advance. A *large world* cannot be exhaustively described beforehand; Bayesian decision theory then fails unless we carve out a tractable small-world approximation while acknowledging what lies outside.

204. In this text, I follow the traditional sense of *computation* as historically used by Turing, Church, and others who studied formal symbolic manipulation under well-defined rules. This framework presupposes discrete states, a systematic transformation of symbols, and, crucially, an interpretive act that assigns meaning to the symbols and their transformations. By contrast, certain contemporary physicists and philosophers apply the term *computation* far more broadly, sometimes to all physical processes, on the grounds that anything unfolding in accordance with natural laws may be seen as *computing* some function. While one can label all physical or dynamical systems as *computations*, I find that this broad usage dilutes the concept. If everything is computation, then *computation* risks becoming an empty descriptor. A catch-all that no longer distinguishes anything special about formal, symbol-processing systems versus physical processes that merely instantiate or evolve under certain constraints. In other words, it is not that I deny the possibility of calling everything *computation* (one can do that), but rather I see it as unhelpful for precise analysis. It obscures the fruitful distinction Turing originally made, namely, that a computational system is one whose states and transitions *must be interpreted* to yield meaning or utility. Without this interpretive component, the concept of *computation* collapses into a trivial restatement of *things happen in the universe*, and we lose an essential analytical tool for differentiating intentionally constrained, symbol-based operations (like a universal Turing machine) from raw physical processes (like a waterfall or a cloud). Consequently, when I say *computation* in these pages, I refer to Turing computation proper. That is processes designed or harnessed to manipulate symbols according to a program, rather than the more generic, everything-is-computing usage found in some corners of physics or digital philosophy.

205. The *Church-Turing sense* of computation refers to the Church-Turing thesis, which posits that any function that can be computed by an algorithm can be computed by a Turing machine. This thesis emerged from the independent work of Alonzo Church and Alan Turing in 1936 [Turing, 1937, Church, 1936]. Church, in his paper on the lambda calculus, proposed that the notion of an *effectively calculable* function could be identified with that of a recursive function. Concurrently

and independently, Turing introduced his abstract *universal computing machine* (now known as the Turing machine) as a formal model of a human computer following a set of rules. The two formalisms, Church's lambda calculus and Turing's machine, were proven to be computationally equivalent, strengthening the claim that they captured the universal limits of what is mechanically computable.

206. see [Bostrom, 2003]

207. see [Zuse, 1969, Landauer, 1961, Wheeler, 1990, Deutsch, 1997a, Lloyd, 2006, Wolfram, 2002a]

208. Not only do they quickly become highly technical, but more importantly, my comrade Yogi Jäger provides a very informative and engaging discussion of these ideas in his forthcoming book *Beyond the Age of Machines.*

209. Your inner punk has probably challenged this now by stating that you associate *green* with jealousy. That was my pick at least. According to the study W*hat Color Is Your Anger?* in 2016, English speaker tend to associate *green* with jealousy the most [Fugate and Franco, 2019]. So synesthesia is a real phenomena and the idea of a pristine clear cut category error belongs into the platonic realm of forms, as we humans easily connect seemingly disparate realms with ease. Not the least because all our concepts and metaphors are based in our embodied ways of knowing connected by a matrix of sensations, as argued by Lakoff and Johnson [2008]. So while I have used the concept of the *category error* I would prefer to think of it as a *categorical distance*, and if this distance is very far between two concepts, they asymptotically converge towards a *category error*. In our case, with the question of computation and physics, the difference between those two categories is quite stark, as computation in the Church-Turing sense requires someone to interpret the computation for it to be a computation. This is not the case for the physical laws.

210. See Mould [2019]

211. In all fairness this is for now mostly a metaphysical questions if the universe is discrete or continuos, or both, or neither at the lowest level and to what extend this can be simulated in a computer. As of now, our progress in answering this question seems to be only moderately promising.

212. [Carse, 1986]

213. A necessary caveat here is that modern day computational hardware has many modules that are not designed for the most *general* purpose but specified for specific types of computation, perhaps best known is that this is why your computer has a CPU and a GPU. So there is clearly a path dependence with the types of programs that we run on computers and how we optimize our hardware to run these types of programs efficiently.

214. By *putting an LLM in a robot body* I do expect a major shift: embodiment moves interaction beyond disembodied symbol manipulation, making the technology more felt and accelerating adoption. It will also bring social side-effects—e.g., habituation to slave-like relationships that lack the friction and feedback of human ties. However, this advance does not resolve the architectural gap at issue, namely, more sensors, actuators, or clever integration do not yield systems that repair themselves, resist entropy, or reproduce. In short, embodied AI will transform user experience and uptake, but it leaves the fundamental computational architecture unchanged, which is precisely where our attention must turn.

215. see von Neumann [1966]

216. As an example see Chan [2019]

217. An interesting albeit speculative question concerns the possibilities of communication between small- and large-world systems. While the fundamental distinction between these worlds would persist, we might imagine interfaces or translation mechanisms that enable meaningful interaction. This would be analogous to how we might conceptualize communication with alien life forms - entities operating under different constraints. Just as we would need to establish common ground with extraterrestrial intelligence despite potentially radical differences in physical embodiment or cognitive architecture, the challenge of small-large world communication would require finding shared patterns or regularities that both domains could recognize and respond to. This communication would necessarily be bounded by the constraints of the computational system.

218. Wittgenstein [1922]

219. The process of breaking down raw text into discrete units for processing represents a crucial step in how LLMs interact with language. Through techniques like Byte-Pair Encoding, this process identifies and preserves meaningful subword units, allowing the model to capture linguistic patterns at multiple scales. This discretization creates what we might think of as a symbolic substrate -

a finite vocabulary of tokens that serves as the basic building blocks for the model's pattern recognition. [Sennrich et al., 2016]

220. The training of LLMs can be conceptualized as transforming an informational gradient into an energy-like gradient within the model's parameter space, analogous to physical systems like Bénard cells. In Bénard cells, a temperature difference between a heated base and a cooler surface creates an energy gradient, driving the self-organization of hexagonal convection patterns that minimize this gradient. Similarly, LLMs are exposed to a vast corpus of text, where statistical differences between predicted and actual tokens define an informational gradient, encapsulated by the loss function. The training process minimizes this loss, akin to reducing an energy differential, by iteratively adjusting the model's weights to better capture linguistic patterns. This adjustment organizes the high-dimensional weight matrix into structured configurations—akin to attractors—that encode recurring statistical regularities, from syntax to semantics. Unlike open-ended physical systems, however, this organization occurs within the closed, deterministic bounds of the training process, converging toward local minima in the loss landscape. This analogy highlights how informational disparities, like energy gradients, can drive emergent complexity.

221. The training process of LLMs can be understood through the lens of *constraint consolidation*. Each training step represents a form of constraint fixation. The model's weights are progressively shaped to capture statistical regularities in the training data. These constraints then influence how the model processes subsequent inputs, creating a cascade of increasingly refined pattern recognition capabilities. However, LLMs operate within the bounded space of their training distribution, and their constraints, once established, remain fixed. This differs fundamentally from living systems, which can evolve new constraints through interaction with their environment. The training process finds an attractor states in the network's parameter space - configurations that effectively capture recurring patterns in the data. These attractors represent the statistical regularities in the language that the model has learned to recognize and reproduce.

222. The sheer dimensionality of these models' parameter spaces, often containing hundreds of billions of weights, creates room for compressing vast amounts of data. Like a hologram can encode three-dimensional information in a two-dimensional surface, the high-dimensional geometry of a Transformer's weight space can encode complex linguistic patterns in surprisingly efficient ways. This enables the model to capture not just surface-level statistics but deep structural regularities in language. Recent research in mechanistic interpretability has begun to uncover how specific neurons or directions in this parameter space encode particular linguistic features, from syntactic categories to semantic relationships, showing how this dimensionality enables the emergence of structured representations. See [Bricken et al., 2023, Templeton et al., 2024]

223. Vaswani2017Attention

224. The Transformer architecture is an example of how some structures can give rise to emergent computational capabilities. The architecture consists of multiple layers of self-attention mechanisms coupled with feed-forward neural networks. Each attention layer allows the model to dynamically establish connections between different positions in a sequence, effectively creating a context-sensitive representation of each token. The multi-head attention mechanism enables the model to simultaneously attend to information from different representation subspaces at different positions, creating an "understanding" of contextual relationships within text.

225. see [Ouyang et al., 2022]

226. see [Christiano et al., 2017]

227. Reinforcement Learning (RL) has traditionally been viewed as the *alignment* phase—a relatively cheap polishing step compared to the massive computational cost of pre-training. However, this paradigm is rapidly shifting with the emergence of "reasoning" models (such as OpenAI's o-series). In these systems, RL is no longer just for stylistic alignment but is used to train the model to generate and critique its own "chains of thought" before answering. This has turned the post-training RL phase into a massive computational undertaking. As noted in recent industry discussions, by figures like Ilya Sutskever, we are witnessing the emergence of a new scaling law. Intelligence no longer scales solely with the size of the model (pre-training compute) but also with the amount of thinking time and RL optimization applied during post-training and inference. Consequently, the costs associated with this RL phase are exploding, potentially rivaling or exceeding pre-training costs as models are trained to simulate *System 2* thinking processes.

228. see [Cuthbertson, 2025]

229. see [Kegl et al., 1999]

230. [Bickerton, 1981]

231. The innatist position originates with Chomsky's proposal of an innate *language-acquisition device* and Universal Grammar (UG) [Chomsky, 1965, 1980]. In its minimalist reformulation, UG is reduced to a single computational operation—recursion [Hauser et al., 2002]. Empirical challenges come from several fronts. Neuro-imaging finds no isolable, language-specific module. Putative Broca's area, intermingles language-selective and domain-general circuits [Gibson and Fedorenko, 2013]. Field research on Pirahã shows a living language whose grammar apparently lacks recursive embedding [Everett, 2005]. Usage-based accounts therefore explain recurrent linguistic patterns as products of general cognitive and social learning mechanisms rather than a specialized neural organ [Tomasello, 2003].

232. [Kirby et al., 2008]

233. see [Saffran et al., 1996]

234. see [Reber, 1967]

235. see [Kallini et al., 2024]

236. see [Peng and Søgaard, 2024, Jha et al., 2025]

237. Though high-quality sentence-level translation typically requires additional training

238. see [Peñaloza et al., 2022]

239. See [MacSweeney et al., 2008, Emmorey, 2002]

240. Nonetheless, it is clear that the constant stream of text produced by LLMs will have significant influence on our cultural and linguistic landscape. Based on data from we are likely going to hit the infleciton point where more online data is generated from AI than from humans is before 2030.

241. LLMs can map the terrain of language with extraordinary fidelity, but they do not expand the terrain on their own, they optimize over an already-stabilized manifold of patterns. Scaling improves coverage, fluency, and retrieval over that manifold, yet delivers diminishing returns on abilities that require crossing from small-world symbol management into large-world constraint creation—new instruments, new norms, new kinds of work, new couplings to matter and energy.

In this light, *super-intelligence by more tokens and parameters* mistakes depth for dimensionality. Depth (more of the same gradients) polishes performance within fixed constraints, dimensionality (new gradients, new couplings, new rules) alters the game itself. The former saturates, the latter is what living systems do when they modify their own boundary conditions. Absent genuine routes to second-order constraint change - embodied access to energy and substrates, autonomy over objectives, and the capacity to make new rules rather than only infer them—LLMs will continue to get better at language about the world while remaining bounded in action on the world. To move beyond the ceiling, architectures must couple to processes that generate and revise constraints in the large world, not merely approximate them in text.

242. see [Merrill and Oremus, 2021]

243. see [Cheng, 2023]

244. Kauffman [1995] introduced the concept of the *adjacent possible* to describe the set of immediate possibilities available to a system at any given moment - those states or configurations that are one step away from the current state. Language dramatically expanded humanity's *adjacent possible* by allowing us to model and communicate potential futures, creating shared maps of possibilities beyond immediate physical reality. While AI systems could theoretically further expand this surface area of exploration, there's an important caveat. Strong path dependence in optimization can actually constrain the direction of exploration. Much like a force vector, once a system commits to a particular trajectory of development, the *adjacent possible* tends to become increasingly confined to variations along that same direction. This geometric metaphor, while speculative, helps illustrate how optimization-driven AI systems might paradoxically reduce rather than expand our effective adjacent possible by intensifying existing trajectories rather than enabling genuine exploration of alternative paths.

245. Of course inspired by Frankfurt's idea of second-order volition, the ability to want what you want [Frankfurt, 1971].

246. A very clear version of this argument can be found in a discussion between Schmachtenberger and Boeree, see [Schmachtenberger and Boeree, 2023]

247. This points towards a design philosophy grounded in *Cyclical Synthesis*, diverging sharply from conventional optimization paradigms. Instead of maximizing static metrics, CS design focuses on cultivating the dynamic health and resilience of systems over time. Its core tenets suggest:

1. **Temporal Dynamics:** Design interventions that understand and modulate the essential rhythms, phases, and transitions inherent in living or adaptive processes. The goal is healthy flow, not a fixed endpoint.
2. **Productive Tension:** Recognize opposing forces (e.g., growth/consolidation, exploration/exploitation, certainty/doubt) not as flaws to be eliminated, but as vital tensions to be managed. Design to orchestrate their interplay, ensuring neither dominates pathologically.
3. **Disrupt Stasis:** Identify points of *lock-in*–stagnation, runaway loops, brittle optimization–where a system becomes stuck in one phase. Design interventions that introduce catalytic disruption or productive friction to enable necessary shifts.
4. **Resilience:** Act as a dynamic modulator, providing countervailing force against dominant trends when needed. The aim is not peak performance in one dimension, but the adaptive capacity to navigate change across multiple dimensions.
5. **Systemic Adaptability:** View the technology (e.g., AI) not as a solution-provider in isolation, but as a participant in a coupled system (AI-human, AI-society). Success lies in enhancing the adaptive potential and long-term viability of the whole interacting system.

This approach reframes design from engineering predictable outcomes to coordinating differing processes, demanding an understanding of temporal patterns and the creative potential found within managed opposition.

248. The term *race to the bottom* describes a competitive dynamic where entities (such as states or corporations) progressively lower standards—in areas like taxation, environmental regulation, or labour laws—to undercut competitors and attract investment or market share. While each individual move may seem rational, the collective result is a degradation of standards and a worse outcome for all. The concept has roots in early 20th-century U.S. corporate law, where states competed to attract incorporations by offering increasingly lax regulations. In game theory, this situation is often modelled as a multi-player Prisoner's Dilemma, where the dominant strategy for each player (defect by lowering standards) leads to a suboptimal collective equilibrium.

249. Wimsatt [2007]

250. See Nietzsche [1882]. Also, through the lens of attachment theory, Europe's taste for fundamentality looks less like courage than compensation. Break the mother-tether early, then canonize the craving for an unbreakable base as philosophy.

251. This is what I like to term *scale-variant* theory building. Recognizing that explanatory principles must change as we move across different scales or domains. This contrasts with the *scale-invariant* approach we pursued earlier in developing universal concepts like constraint cycles. The tension between these approaches reflects a deeper challenge in scientific thought, namely, our drive to discover universal principles while continually encountering the stubborn specificity of different levels of organization.

252. For a good introduction to this field, the physicist Geoffrey West's "Scale" is the go-to place [West, 2017].See also [Pareto, 1896, Zipf, 1949]

253. See [Kleiber, 1932]. Importantly, exponents can vary by clade, ecology, and measurement context. While the 3/4-power has been influential, alternative derivations and data support 2/3 or mixed exponents in some taxa and circumstances. None of this undermines the value of scaling analysis. It clarifies that universality lives alongside scale-variant deviation, which will matter later when we examine discontinuities.

254. See [West et al., 1997]

255. [Allen and Holling, 2008]

256. [Allen et al., 1999]

257. see [Allen and Saunders, 2002]

258. Lambert and Holling [1998]

259. [Holling, 1992]

260. I do not want to argue or support the view that our positing at the edge of body-mass aggregations is the sole or even most important reason for our expansion into many parts of the Earth, but compelling arguments can be made that our specific body-mass forced us to adapt to new niches (we left the jungle for the planes of Africa) forcing us to develop more sophisticated symbolic communication mechanisms because we were not inhabiting a stable aggregation. See Deacon [1998] for the discussion on the emergence on symbolic communication mechanism.

261. Some human size distributions are often well-described by heavy-tailed, near-Zipf distributions that look continuous rather than clumpy. Even there, discontinuities can appear at policy-driven thresholds or through *club convergence* in macroeconomic growth. Where clumps do not appear, the broader point still holds, scale relationships alone do not fix occupancy, institutional and ecological constraints shape which size ranges are viable and persistent.

262. See Allen and Holling [2008]

263. [Baumol, 1986]

264. Wimsatt [2007] develops robustness as a criterion for ontological commitment: something is robust if it is invariant under multiple independent derivations, measurements, or observational perspectives.

265. Understanding these regimes requires recognizing three constraint types operating simultaneously to define possibility spaces. Internal constraints emerge from entities within the level, membrane selectivity, genetic regulation, metabolic requirements. External constraints are imposed by environment or adjacent levels like resource availability, physical forces, chemical gradients. Inter-level constraints arise from interactions with higher or lower organizational levels. Consider how a cell membrane operates fundamentally through selective exclusion by actively preventing specific molecules from entering while allowing others through based on size, charge, or transport mechanisms. This selective constraint is as definitional of cellular organization as any positive feature we might catalogue.

266. The quantum mechanical "origin" of these constraints, solving Schrödinger's equation in the atom's electric field, yields only certain allowed energy states, with vast ranges of intermediate energies systematically excluded. New electrons entering the system must align with these pre-existing levels according to Pauli's exclusion principle. Similar constraints appear throughout other domains in physics. Electrons in crystals organize into energy bands separated by forbidden gaps (the foundation of all semiconductor technology); superconductors expel magnetic fields and eliminate electrical resistance below critical temperatures; the quantum Hall effect creates plateaus of exactly quantized electrical conductance; even white dwarf stars exist because quantum mechanical constraints prevent further gravitational collapse until a critical mass limit.

267. The constraint operates through electron delocalization. When the number of mobile electrons fits the 4n+2 pattern, quantum mechanical wave functions reinforce to create extra stability. With 4n electrons, they interfere destructively. Similar chemical constraints appear throughout molecular organization. Woodward-Hoffmann rules forbid certain reactions under normal conditions but allow them under UV light; micelles only form above critical concentrations where geometric packing constraints force collective assembly; ionic crystals adopt specific geometries based on size ratios that exclude other arrangements; metal complexes strongly favor 18 valence electrons; even acid-base chemistry faces limits where water prevents any acid from being stronger than HO. Each shows the same pattern: gaps in what can exist, multiple detection methods revealing the same boundaries, and systematic success or failure based on constraint alignment.

268. Gunderson and Holling [2002]

269. Gunderson and Holling [2002]

270. Gunderson and Holling [2002]

271. See [Brand, 1999, 1994]

272. This example comes from Hazen et al. [2008] work on mineral evolution, demonstrating how the progressive layering of constraintsnfrom basic chemistry to complex biological processes, enabled an explosion in mineral diversity. This shows that constraint-driven complexity isn't unique to living systems but represents a more general principle of natural organization.

273. Upon starvation, individual amoebae begin to synthesize and release cAMP in periodic pulses, approximately every 5-6 minutes. Neighboring cells detect this signal via G-protein-coupled receptors, causing them to move toward the source and, in turn, relay the signal by releasing their own pulse of cAMP. This highly organized, self-propagating relay system creates the characteristic spiral or concentric waves of communication that coordinate the aggregation of up to 100,000 individual cells. This process is a classic example of how a simple molecular signal can generate complex, emergent spatiotemporal patterns, enabling a transition from independent units to a coordinated collective. See Bonner [1947]

274. See Strassmann et al. [2000]

275. See Mauss and Beuchat [1979]

276. Plants under biotic or abiotic stress emit complex blends of volatile organic compounds (VOCs), such as terpenes and green leaf volatiles. This signaling has mediates both plant-insect and plant-plant interactions. The evolutionary framework for this communication is sometimes described as "eavesdropping". A plant under attack releases VOCs primarily for its own benefit—for example, to signal other parts of itself or to attract predators of the herbivores attacking it. Neighboring plants that have evolved the ability to detect and interpret these airborne cues gain a significant fitness advantage by preemptively activating their own defense systems. The emitter may not directly benefit from warning a competitor, but it cannot prevent the public broadcast of its distress signal. See Heil and Karban [2010] as well as Howes and Simmonds [2014]

277. The molecule even directly influences behaviour. In some alpine lakes, Daphnia water fleas modulate their astaxanthin content based on UV exposure and predation risk. Under high UV but low predation, they become red with protective astaxanthin. If fish are present, they reduce pigmentation to avoid being seen, accepting higher UV damage to escape predation. See Loboda et al. [2016], Zhao et al. [2022]

278. See Niu et al. [2018]

279. Similar signals appear in other contexts. The mTOR pathway, which regulates cell growth and metabolism, responds to nutrient availability in organisms from yeast to humans. The mechanistic Target of Rapamycin (mTOR) pathway, particularly the mTOR Complex 1 (mTORC1), is a highly conserved signaling nexus that functions as a master regulator of cell growth and metabolism. It

integrates a wide range of environmental cues, most notably the availability of nutrients such as amino acids and glucose, as well as growth factors and cellular energy status. When nutrients are abundant, mTORC1 promotes anabolic processes (e.g., protein and lipid synthesis) and inhibits catabolic processes (e.g., autophagy). This positions mTORC1 as a fundamental constraint-management system at the cellular level, coupling the cell's growth machinery to the availability of external resources, a logic conserved from single-celled eukaryotes to humans. See Zoncu et al. [2011]. Plant volatile compounds can trigger immune responses in insects and even mammals that consume them. These molecular dialogues create webs of shared signals across species, allowing entire ecosystems to exhibit degrees of coordinated response to environmental challenges.

280. See Ostrom [1990]

281. This energetic trade-off mirrors the distinction Fath and colleagues draw between growth (quantitative expansion of biomass or throughput) and development (qualitative increase in organization and complexity). They argue that while early-stage systems prioritize growth, long-term persistence requires a phase transition where energy is redirected toward development—maintaining the complex constraints and information structures that ensure resilience. See Jørgensen et al. [2015].

282. see [Pellegrino et al., 2015, Merckx and Gomes, 2020, Smith and Read, 2008, Aug'e, 2001]

283. see [Flemming and Wingender, 2010, Stewart and Costerton, 2001, Mah and O'Toole, 2001]

284. see [Alberts, 2019, Silk et al., 2009]

285. This corresponds closely to Terrence Deacon's framework of homeodynamics, morphodynamics, and teleodynamics and its recursive application discussed in his work on Dynamical DepthDeacon and Koutroufinis [2014]. We've chosen to emphasize *self-constraining* and *constraint* to differentiate self-organizing from self-constraining dynamics, which we find didactically helpful and conceptually clearer, though this is mostly a matter of preference. The fundamental organizational principles remain similar.

286. Taken again from Doolittle [2024]

287. Compare Peter Godfrey-Smith's reproduction-centered account of evolutionary individuality. On his view, *Darwinian individuals* are members of Darwinian populations, with individuality graded by three reproduction-related parameters—Bottleneck (B), Germ-line specialization (G), and overall Integration (I). Paradigm cases are high on B/G/I, while metabolically integrated but multi-lineage complexes (e.g., squid–*Vibrio* symbioses) can count as organisms yet not as *Darwinian individuals* because the collective does not reproduce as a unit. In the present framework, such high-closure cases appear as a limiting point on a broader spectrum: constraint regimes can become selectable before organism-like reproduction, with *fitness* tracked by persistence and reconstructability rather than offspring number. See Godfrey-Smith [2013]

288. This preservation operates through distinct but complementary memory channels. First are physical and institutional scaffolds (like soil structures, terraces, or technical standards) that provide durable arrangements that guide pattern re-creation. Second are codes and protocols which are transmissible instructions, whether as genes, epigenetic markers or migration timing cues. Third are partner pools that maintain reliable access to necessary collaborators, from symbiotic bacteria to seed banks to professional communities. When these memory channels work together effectively, patterns can reconstruct themselves even when their current carriers change completely.

289. See Lewontin [1970]

290. This concept of evolved dependency shares a direct conceptual parallel with the Black Queen Hypothesis (BQH) in microbial ecology [Morris et al., 2012]. The BQH posits that natural selection can favour the loss of a gene for a costly, leaky public good when its product can be reliably sourced from other community members. Both frameworks explain how interdependence arises through the adaptive loss of a function. However, they differ in scope and emphasis. The BQH is primarily gene-centric, focusing on the immediate energetic benefits of genome streamlining via the loss of public functions. Degradation dependency, as used here, is a broader organizational principle that includes the loss of private, internal functions (e.g., entire metabolic pathways) as a consequence of specialization. Its central focus is not on immediate efficiency but on the creation of a hidden vulnerability, an addiction-like lock, in to a specific partner or environmental condition that only becomes apparent during a perturbation test.

291. Terrence Deacon has explored similar principles through the example of vitamin C synthesis in primates. Humans and other primates lost the capacity to produce ascorbic acid internally—our genomes retain the pseudogene for L-gulonolactone oxidase, but it has accumulated mutations that render it non-functional. This created a dietary dependency that proved catastrophic during ocean

voyages (scurvy), yet the loss coincided with enhanced color vision that enabled efficient identification of vitamin C-rich fruits. Deacon frames this as "Inverse Darwinism"—evolutionary processes that create dependencies reflecting adaptive trade-offs within specific environmental contexts. His work explores how such degradation dependencies operate across biological and cultural evolution, creating addiction-like lock-ins that constrain future evolutionary possibilities.

292. See Ashby [2015]. The law ($V_r \geq V_d$) dictates that if the variety of environmental disturbances (V_d) exceeds the variety of the system's regulatory responses (V_r), the system cannot maintain stability. Efficiency mechanisms often work by reducing V_r to match the historical average of V_d, stripping away "redundant" capacity. This leaves the system mathematically incapable of compensating for novel or extreme variations in V_d.

293. Animal products amplify this imbalance further. Beef production demands roughly twenty-five fossil calories per food calorie, pork fifteen, and chicken nine. The Haber-Bosch process for synthetic nitrogen production alone consumes 26-40 gigajoules per tonne of ammonia, accounting for 1-2% of global energy use and roughly one-fifth of industrial natural gas demand.

294. The theoretical framework that dominated evolutionary biology for much of the twentieth century was the Modern Synthesis, established through a series of foundational works in the 1930s and 1940s Dobzhansky [1937]Mayr [1942]Simpson [1944] The term *Modern Synthesis* was coined by the English biologist Julian Huxley in his influential 1942 book, *Evolution: The Modern Synthesis.* Huxley's work was less a presentation of novel research and more a comprehensive review and popularization that aimed to demonstrate the vast explanatory power of the new, unified theory. He synthesized the findings from genetics, systematics, paleontology, ecology, and developmental biology, arguing that they all converged on a Darwinian framework updated with the mechanisms of heredity. See Huxley [1942]. The focus on the gene as the central unit of reproduction can be traced back to Dawkins [1976]

295. Multi-level selection (MLS) theory posits that natural selection can operate on a nested hierarchy of biological units, such as genes within cells, cells within organisms, and organisms within groups. The modern formulation of this theory, developed largely from the 1970s onward by theorists such as David Sloan Wilson, represents a significant refinement of the earlier, discredited "naive" group selection theories [Sober and Wilson, 1998]. Crucially, modern MLS theory is not a return to vague "good of the species" arguments. It employs a mathematical framework, often using tools like the Price equation, to partition the total evolutionary change into separate components of within-group selection and between-group selection. Within-group selection typically favors selfish

traits (e.g., an individual who shirks a costly cooperative duty), while between-group selection favors traits that enhance the collective success of the group (e.g., cooperation leading to greater group productivity or success in conflict with other groups). A group-beneficial trait can evolve if, and only if, the force of between-group selection is stronger than the opposing force of within-group selection.

296. In recent decades, a growing number of theorists have argued that the Modern Synthesis, while foundational, is incomplete and must be supplemented by a range of additional concepts, collectively termed the Extended Evolutionary Synthesis. See Laland et al. [2015]. For a specific focus on niche construction see Odling-Smee et al. [2003]

297. One of the more radical rethinking of biological individuality comes from the proposal that the unit of selection is not the singular organism, but the *holobiont*. The host plus its entire community of symbiotic microbes. See Zilber-Rosenberg and Rosenberg [2008]

298. This broader view of evolution also considers the large-scale historical events that have fundamentally restructured the biological hierarchy, known as the major evolutionary transitions. See Maynard Smith and Szathmáry [1995]

299. see Deleuze [1994]

300. Early organisms developed simple nervous systems to coordinate movement, vertebrates evolved systems for learning from experience, and mammals developed the ability to simulate possible futures. Each stage builds upon and incorporates earlier achievements while enabling new forms of adaptive behavior. Bennett traces five evolutionary *breakthroughs* that scaffold humanlike cognition while preserving earlier capacities. (1) Steering in early bilaterians: centralized, valenced control for taxis that integrates competing inputs into a single movement policy. (2) Reinforcing in vertebrates: basal ganglia–mediated, prediction-error learning that retunes action policies through experience. (3) Simulating in mammals: a neocortical generative model with hippocampal replay and vicarious trial-and-error, enabling counterfactual planning. (4) Mentalizing in primates: granular prefrontal circuitry that models one's own and others' mental states, supporting imitation, deception, and prospective choice. (5) Speaking in humans: language systems that externalize and transmit learned policies, accelerating cumulative culture. Each layer reorganizes, rather than replaces, the prior ones, yielding increasingly flexible, compositional control over perception and action. For a full account, see Bennett [2023].

301. For Bateson, information is not a substance or a quantity but a relational event. A *difference* (e.g., a change in light intensity, a sound) only becomes *information* when it is perceived by a system in a way that *makes a difference* to that system's subsequent state or behavior. This definition places the emphasis on the active, selective process of the perceiving system, which must distinguish meaningful signals from a background of undifferentiated reality. I do not think that this is a good definition of information in general as difference in itself can already be Shannon information (which from a semiotic perspective is also a bad definition of information), but for cognitive processes, it is an excellent definition as those differences that come to make a difference, that come to be perceived make up our cognition. See Bateson [2000]

302. See Winawer et al. [2007]

303. See Lovejoy [1936]

304. Neo-Piagetian theorists like Case [1992] moved beyond Piaget's original stage theory by integrating insights from information processing. Case proposed that cognitive development proceeds through a sequence of *central conceptual structures*. These are not just new logical abilities but are more like complex, integrated networks of concepts and procedures that function as executive control structures. Each new structure reorganizes the child's understanding within a domain (e.g., number, narrative) and increases the complexity of problems they can solve. Development is thus seen not as the acquisition of isolated facts but as the construction of progressively more powerful and integrated constraint systems for thought. A very clear articulation of this concept supplied by ample evidence can also be found in the work of Michael Lamport Commons and colleagues called the Model of Hierarchical Complexity [Commons, 2007]. It describes the the trajectory of different constraint regimes that build in a non-arbitrary manner on each other.

305. This stance dovetails with the artifactual view that treats models as epistemic tools (or artifacts): purpose-built, materially realized devices whose *constrained design*, manipulability, and outcome-orientation enable inquiry. See especially Knuuttila's account of models as epistemic tools [Knuuttila, 2011], the early *parser as epistemic artifact* case [Knuuttila and Voutilainen, 2003], and Knuuttila & Loettgers' work on synthetic biology and modelling practice [Knuuttila and Loettgers, 2013, 2017].

306. To further develop this metaphor of the adaptation of organism-niche construction, consider how modelling itself proceeds through cycles. As Kuhn has famously argued, history of scientific

progress follows a recurring pattern: An exploratory phase, where new observations and ideas generate multiple competing models. A consolidation phase, where the most effective constraints stabilize into an accepted framework. A disruption phase, where accumulating anomalies expose the cracks in the existing constraints, and a reorganizational phase, where older models are either refined or overturned in favour of new ones better suited to the data. See Kuhn [1962]. This dynamic pattern of stability and revision aligns with the logic of *Cyclical Synthesis* operating at the level of knowledge creation.

307. Importantly, this is not epistemological relativism. Models remain accountable to the systems they describe, tested by their capacity to generate reliable inferences and satisfy Wimsatt's robustness criterion discussed earlier. The argument is for irreducibility, not equivalence: different levels require different conceptual architectures because reality itself is structured in levels, not because our methods are inadequate.

308. This perspective shifts the understanding of modelling away from naive realism (representation as mirroring) toward an artifactual or inferential view. Drawing on Roman Frigg's analysis of models as fictions, we can understand models not as direct descriptions of physical reality, but as imagined systems defined by specific descriptions and constraints. These *fictions* possess properties only within the context of their generation but allow for a key of translation to the target system Frigg [2010]. Furthermore, as Mauricio Suárez argues in his inferential conception of scientific representation, the primary function of a model is not isomorphism (structural identity with the world), but its capacity to license valid inferences about its target. In this light, *modelling as constraint selection* is the act of constructing a specific inferential architecture, one imposes constraints to generate a system that, while not identical to the life process, allows to reason effectively about it Suárez [2004].

309. The central papers that articulated the idea of causal emergence are Hoel [2017, 2025], Jansma and Hoel [2025]. For a more detailed treatment of CE these are the primary sources.

310. Pearl [2009]

311. Jansma and Hoel's simulation varies the attachment exponent α to demonstrate how different network growth rules generate distinct causal hierarchies (e.g., sublinear $\alpha < 1$ yields bottom-heavy hierarchies, superlinear $\alpha > 1$ yields top-heavy ones). This simulation directly engages with the *classical* model of scale-free networks, where $\alpha \approx 1$ is the canonical regime for generating

networks with power-law degree distributions. For the seminal paper introducing this *preferential attachment* mechanism, see Barabási and Albert [1999].

Bibliography

Theodor W. Adorno. *Negative Dialectics*. Seabury Press, New York, 1973. Original work published 1966 as *Negative Dialektik*.

Susan C. Alberts. Social influences on survival and reproduction: Insights from a long-term study of wild baboons. *Journal of Animal Ecology*, 88(1):47–66, January 2019. doi: 10.1111/1365-2656.12887. Epub 2018 Aug 21.

Colin Allen and Jacob Neal. Teleological Notions in Biology. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2020 edition, 2020.

C.R. Allen and C.S. Holling. *Discontinuities in Ecosystems and Other Complex Systems*. Complexity in Ecological Systems. Columbia University Press, 2008. ISBN 9780231516822. URL https://books.google.at/books?id=tJa0iGlcLeIC.

Craig R. Allen and Denis A. Saunders. Variability between scales: Predictors of nomadism in birds of an australian mediterranean-climate ecosystem. *Ecosystems*, 5(4):348–359, jun 2002. doi: 10.1007/s10021-001-0079-z. URL https://doi.org/10.1007/s10021-001-0079-z.

Craig R. Allen, Elizabeth A. Forys, and C. S. Holling. Body mass patterns predict invasions and extinctions in transforming landscapes. *Ecosystems*, 2(2):114–121, 1999. doi: 10.1007/s100219900063.

Joseph Almog. *What Am I? Descartes and the Mind-Body Problem*. Oxford University Press, Oxford, 2001.

Alan Ross Anderson, J Michael Dunn, and Nuel D Belnap. Entailment: The logic of relevance and necessity. 2017.

Aristotle. *Metaphysics: Books I-IX*. Harvard University Press, Cambridge, 1979.

W Ross Ashby. Principles of the self-organizing dynamic system. *The Journal of general psychology*, 37(2):125–128, 1947.

W.R. Ashby. *An Introduction to Cybernetics*. Martino Publishing, 2015. ISBN 9781614277651. URL https://books.google.at/books?id=9tT0rQEACAAJ.

R. Michael Aug'e. Water relations, drought and vesicular–arbuscular mycorrhizal symbiosis. *Mycorrhiza*, 11:3–42, 2001. doi: 10.1007/s005720100097.

Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999. doi: 10.1126/science.286.5439.509.

K. Barad. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Titolo collana. Duke University Press, 2007. ISBN 9780822339175. URL https://books.google.at/books?id=U8frDwAAQBAJ.

J. Barbour. *The Janus Point: A New Theory of Time*. Random House, 2020. ISBN 9781473548862. URL https://books.google.at/books?id=w5g0DwAAQBAJ.

G. Bateson. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. University of Chicago Press, 2000. ISBN 9780226039053. URL https://books.google.at/books?id=HewJbnQmn1gC.

William J. Baumol. Productivity growth, convergence, and welfare: What the long-run data show. *The American Economic Review*, 76(5):1072–1085, 1986.

M.S. Bennett. *A Brief History of Intelligence: Evolution, AI, and the Five Breakthroughs That Made Our Brains*. HarperCollins, 2023. ISBN 9780063286368. URL https://books.google.at/books?id=tymCEAAAQBAJ.

H. Bergson. *Creative Evolution*. Modern library. University Press of America, 1984. ISBN 9780819135537. URL https://books.google.at/books?id=Cx4H9wTF8mkC.

Henri Bergson. *Time and Free Will: An Essay on the Immediate Data of Consciousness.* Courier Corporation, 2001.

Claude Bernard. *Introduction á l'étude de la médecine expérimentale.* Baillière, Paris, 1984 edition, 1865.

Francis Bernardeau, Stéphane Colombi, Enrique Gaztañaga, and Román Scoccimarro. Large-scale structure of the universe and cosmological perturbation theory. *Physics Reports*, 367(1-3):1–248, 2002. doi: 10.1016/S0370-1573(02)00135-7. Review of perturbation theory and gravitational amplification of primordial perturbations into the cosmic web.

Derek Bickerton. *Roots of Language.* Karoma Publishers, Ann Arbor, MI, 1981.

Angelo Bisazza, Lesley J. Rogers, and Giorgio Vallortigara. The origins of cerebral asymmetry: a review of evidence of behavioural and brain lateralization in fishes, reptiles and amphibians. *Neuroscience & Biobehavioral Reviews*, 22(3):411–426, 1998. doi: 10.1016/S0149-7634(97)00050-X.

John Tyler Bonner. Evidence for the formation of cell aggregates by chemotaxis in the development of the slime mold dictyostelium discoideum. *The Journal of Experimental Zoology*, 106(1):1–26, 1947. doi: 10.1002/jez.1401060102.

Nick Bostrom. Are you living in a computer simulation? *The Philosophical Quarterly*, 53(211):243–255, 2003.

S. A. Bowring and I. S. Williams. Priscoan (4.00-4.03 ga) orthogneisses from northwestern canada. *Contributions to Mineralogy and Petrology*, 134(1):3–16, 1999. doi: 10.1007/s004100050465.

Stewart Brand. *How Buildings Learn: What Happens After They're Built.* Viking Press, New York, NY, 1994.

Stewart Brand. *The Clock of the Long Now: Time and Responsibility.* Basic Books, New York, NY, 1999.

Dennis Bray. Protein molecules as computational elements in living cells. *Nature*, 376:307–312, 1995. doi: 10.1038/376307a0. Seminal paper proposing that phospho-proteins act as logical elements akin to transistors in a circuit.

Dennis Bray. *Wetware: A Computer in Every Living Cell*. Yale University Press, New Haven, 2009. Argues that cells possess an internal representation of their environment and perform computations via protein signaling networks.

Trenton Bricken, Adly Templeton, Joshua Batson, Brian Chen, Adam Jermyn, Tom Conerly, Nick Turner, Cem Anil, Carson Denison, Amanda Askell, Robert Lasenby, Yifan Wu, Shauna Kravec, Nicholas Schiefer, Tim Maxwell, Nicholas Joseph, Zac Hatfield-Dodds, Alex Tamkin, Karina Nguyen, Brayden McLean, Josiah E Burke, Tristan Hume, Shan Carter, Tom Henighan, and Christopher Olah. Towards monosemanticity: Decomposing language models with dictionary learning. *Transformer Circuits Thread*, 2023. https://transformer-circuits.pub/2023/monosemantic-features/index.html.

Søren Brier. *Cybersemiotics: Why Information Is Not Enough*. University of Toronto Press, Toronto, 2008.

Stephen G Brush. *The kind of motion we call heat*, volume 2. North-Holland Amsterdam, 1976.

Levi R. Bryant. *Difference and Givenness: Deleuze's Transcendental Empiricism and the Ontology of Immanence*. Northwestern University Press, Evanston, IL, 2008.

Marilyn R. Buchakjian and Sally Kornbluth. The engine driving the ship: metabolic steering of cell proliferation and death. *Nature Reviews Molecular Cell Biology*, 11(10):715–727, 2010. doi: 10.1038/nrm2972.

Ted Carmichael and Mirsad Hadžikadić. *The Fundamentals of Complex Adaptive Systems*, pages 1–16. Springer International Publishing, Cham, 2019. doi: 10.1007/978-3-030-20309-2_1. URL https://doi.org/10.1007/978-3-030-20309-2_1.

James P. Carse. *Finite and Infinite Games: A Vision of Life as Play and Possibility*. Free Press, New York, 1986.

Robbie Case. *The Mind's Staircase: Exploring the Conceptual Underpinnings of Children's Thought and Knowledge*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1992.

Jordana Cepelewicz. Mathematicians complete quest to build 'spherical cubes'. Quanta Magazine, 2020. URL https://www.quantamagazine.org/mathematicians-complete-quest-to-build-spherical-cubes-20200416/.

Bert Wang-Chak Chan. Lenia: Biology of artificial life. *Complex Systems*, 28(3): 251–286, October 2019. ISSN 0891-2513. URL http://dx.doi.org/10.25088/ComplexSystems.28.3.251.

Michelle Cheng. A startup founded by former google employees claims that users spend two hours a day with its ai chatbots. *Quartz*, October 2023. URL https://qz.com/a-startup-founded-by-former-google-employees-claims-tha-1850919360. Accessed 2025-07-08.

Jaigyoung Choe. On the existence and regularity of fundamental domains with least boundary area. *Journal of Differential Geometry*, 29(3), 1989. ISSN 0022-040X. doi: 10.4310/jdg/1214443065.

Noam Chomsky. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA, 1965.

Noam Chomsky. *Rules and Representations*. Columbia University Press, New York, 1980.

Paul F. Christiano, Jan Leike, Tom B. Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, pages 4299–4307. Curran Associates, Inc., 2017.

Alonzo Church. An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 58(2):345–363, 1936. doi: 10.2307/2371045.

Edmund M Clarke, E Allen Emerson, and A Prasad Sistla. Automatic verification of finite-state concurrent systems using temporal logic specifications. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 8(2):244–263, 1986.

Rudolf Clausius. On different forms of the fundamental equations of the mechanical theory of heat and their convenience for application. *The second law of thermodynamics*, pages 162–193, 1865.

Michael Lamport Commons. Introduction to the model of hierarchical complexity. *Behavioral Development Bulletin*, 13(1):1, 2007.

Michael C. Corballis. Left brain, right brain: Facts and fantasies. *PLOS Biology*, 12(1):e1001767, 2014. doi: 10.1371/journal.pbio.1001767.

Anthony Cuthbertson. Chatgpt is pushing people towards mania, psychosis and death – and openai doesn't know how to stop it. *The Independent*, jul 2025. URL https://www.independent.co.uk/tech/chatgpt-psychosis-ai-therapy-chatbot-b2781202.html. Published 6 July 2025, 06:00 BST.

G. Cuvier, W.B. Carpenter, J.O. Westwood, E. Blyth, R. Mudie, and G. Johnston. *The Animal Kingdom: Arranged After Its Organization; Forming a Natural History of Animals, and an Introduction to Comparative Anatomy*. A. Fullarton, 1859. URL https://books.google.at/books?id=bcI47mu8GlsC.

Newton CA Da Costa. On the theory of inconsistent formal systems. *Notre dame journal of formal logic*, 15(4):497–510, 1974.

Charles Darwin. *The Life and Letters of Charles Darwin: Including an Autobiographical Chapter*, volume 1. John Murray, London, 1887.

Charles Darwin. *The Various Contrivances by Which Orchids are Fertilised by Insects*. University of Chicago Press, Chicago, 2nd edition, 1984. Originally published in London by John Murray in 1862. Second edition published in 1877. Reprint of the second edition with a new foreword by Michael Ghiselin.

Charles R. Darwin. Letter to j. d. hooker, 1 february 1871. In Frederick Burkhardt, James A. Secord, Janet Browne, Samantha Evans, Shelley Innes, Alison M. Pearn, and Paul White, editors, *The Correspondence of Charles Darwin, Volume 19: 1871*. Cambridge University Press, Cambridge, 2012. Letter no. 7471.

B.W. Davis. *The Oxford Handbook of Japanese Philosophy*. Oxford handbooks in philosophy. Oxford University Press, 2020. ISBN 9780199945726. URL https://books.google.at/books?id=kmCpDwAAQBAJ.

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

Frans B. M. De Waal. *The Age of Empathy: Nature's Lessons for a Kinder Society*. Harmony Books, 2008.

Terrence Deacon and Spyridon Koutroufinis. Complexity and dynamical depth. *Information*, 5:404–423, 07 2014. doi: 10.3390/info5030404.

Terrence W. Deacon. *Incomplete Nature: How Mind Emerged from Matter*. W. W. Norton, United Kingdom, 2011.

Terrence W Deacon. How molecules became signs. *Biosemiotics*, 14(3):537–559, 2021.

T.W. Deacon. *The Symbolic Species: The Co-evolution of Language and the Brain*. W. W. Norton, 1998. ISBN 9780393343021. URL https://books.google.at/books?id=umppEX7O6oIC.

G. Deleuze and P. Patton. *Difference and Repetition*. Continuum Impacts Series. Bloomsbury Academic, 2004. ISBN 9780826477156. URL https://books.google.at/books?id=xoHsexDBgmEC.

Gilles Deleuze. *Difference and Repetition*. Columbia University Press, New York, 1994. Original work published 1968 in French as *Différence et répétition*.

Jacques Derrida. *Of Grammatology.* Johns Hopkins University Press, Baltimore, 1976. Originally published as *De la grammatologie* (1967).

René Descartes. *Meditations on First Philosophy.* Cambridge University Press, Cambridge, 1641. J. Cottingham (trans.) 1996 Edition.

David Deutsch. *The Fabric of Reality.* Penguin Books, New York, 1997a.

David Deutsch. *The Fabric of Reality: The Science of Parallel Universes–and Its Implications.* Penguin Books, 1997b.

John Dewey. *Logic: The Theory of Inquiry.* Henry Holt and Company, New York, 1938.

James DiFrisco and Matteo Mossio. Diachronic identity in complex life cycles: An organisational perspective. In *Biological identity*, pages 177–199. Routledge, 2020.

Theodosius Dobzhansky. *Genetics and the Origin of Species.* Columbia University Press, New York, 1937.

M. S. Dodd, D. Papineau, T. Grenne, J. F. Slack, M. Rittner, F. Pirajno, J. O'Neil, and C. T. S. Little. Evidence for early life in earth's oldest hydrothermal vent precipitates. *Nature*, 543:60–64, 2017. doi: 10.1038/nature21377.

W.F. Doolittle. *Darwinizing Gaia: Natural Selection and Multispecies Community Evolution.* Vienna Series in Theoretical Biology. MIT Press, 2024. ISBN 9780262549523. URL https://books.google.at/books?id=Hff4EAAAQBAJ.

Mary Douglas. *Natural Symbols: Explorations in Cosmology.* Pantheon Books, New York, 1970.

A.S. Eddington. *The Nature of the Physical World.* Gifford lectures. Macmillan, 1940. URL https://books.google.at/books?id=GCfcBqVFQSIC.

Niles Eldredge and S. Gould. *Punctuated Equilibria: An Alternative to Phyletic Gradualism*, volume 82, pages 82–115. 11 1971.

E Allen Emerson. Temporal and modal logic. In *Formal models and semantics*, pages 995–1072. Elsevier, 1990.

Karen Emmorey. *Language, cognition, and the brain: Insights from sign language research*. Lawrence Erlbaum Associates, Mahwah, NJ, 2002.

Daniel L. Everett. Cultural constraints on grammar and cognition in pirahã: Another look at the design features of human language. *Current Anthropology*, 46 (4):621–646, 2005.

Brian D. Fath and Bernard C. Patten. Network synergism: Emergence of positive relations in ecological systems. *Ecological Modelling*, 107(2-3):127–143, 1998. doi: 10.1016/S0304-3800(97)00213-5.

Brian D. Fath and Bernard C. Patten. Review of the foundations of network environ analysis. *Ecosystems*, 2(2):167–179, 1999. doi: 10.1007/s100219900067.

Hans-Curt Flemming and Jost Wingender. The biofilm matrix. *Nature Reviews Microbiology*, 8(9):623–633, 2010. doi: 10.1038/nrmicro2415.

John Foster. A defence of dualism. In J. Smythies and J. Beloff, editors, *The Case for Dualism*, pages 59–80. University of Virginia Press, Charlottesville, 1989.

Lawrence Kelso Frank. *Teleological mechanisms*, volume 50. New York Academy of Sciences, 1948.

Harry G. Frankfurt. Freedom of the will and the concept of a person. *The Journal of Philosophy*, 68(1):5–20, 1971. doi: 10.2307/2024717.

Elisa Frasnelli, Giorgio Vallortigara, and Lesley J. Rogers. Left–right asymmetries of behaviour and nervous system in invertebrates. *Neuroscience & Biobehavioral Reviews*, 36(4):1273–1291, 2012. doi: 10.1016/j.neubiorev.2012.02.006.

Hanzi Freinacht. *The Listening Society: A Metamodern Guide to Politics, Book One*. Number 1 in Metamodern Guides. Metamoderna ApS, March 2017. ISBN 978-8799973903.

Hanzi Freinacht. *Nordic Ideology: A Metamodern Guide to Politics, Book Two*. Number 2 in Metamodern Guides. Metamoderna ApS, May 2019. ISBN 978-8799973927.

S. Freud. *On Murder, Mourning and Melancholia*. Penguin Modern Classics. Penguin Books Limited, 2005. ISBN 9780141915517. URL https://books.google.at/books?id=J-Kg1DahJ2UC.

Roman Frigg. Models and fiction. *Synthese*, 172(2):251–268, 2010. doi: 10.1007/s11229-009-9505-0.

Jennifer Marie Binzak Fugate and Courtny L. Franco. What color is your anger? assessing color-emotion pairings in english speakers. *Frontiers in Psychology*, 10, 2019. ISSN 1664-1078. doi: 10.3389/fpsyg.2019.00206. URL https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2019.00206.

Liane Gabora and Mike Steel. Modeling a cognitive transition at the origin of cultural evolution using autocatalytic networks. *Cognitive Science*, 44(9):e12878, 2020. doi: 10.1111/cogs.12878. Develops the idea that cultural evolution can be modeled as autocatalytic networks of mental representations/ideas that mutually evoke and generate further ideas, enabling self-sustaining, cumulative cultural change.

Hans-Georg Gadamer. *Hegel's Dialectic: Five Hermeneutical Studies*. Yale University Press, New Haven, 1976.

Gordon G. Gallup. Chimpanzees: Self-recognition. *Science*, 167(3914):86–87, 1970. doi: 10.1126/science.167.3914.86.

Miguel García-Valdecasas. On the naturalisation of teleology: self-organisation, autopoiesis and teleodynamics. *Adaptive Behavior*, 30(2):103–117, 2022.

James Garson. Modal Logic. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2024 edition, 2024.

M Gell-Mann. Complex adaptive systems. *Complexity: Metaphors, Models and Reality/Addison-Wesley*, 1994.

Stefano Ghirlanda and Giorgio Vallortigara. The evolution of brain lateralization: a game-theoretical analysis of population structure. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1541):853–857, 2004. doi: 10.1098/rspb.2003.2669.

Edward Gibson and Evelina Fedorenko. The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes*, 28(1-2):88–124, 2013. doi: 10.1080/01690965.2010.515080. URL https://www.tandfonline.com/doi/full/10.1080/01690965.2010.515080.

Jean-Yves Girard. Linear logic. *Theoretical Computer Science*, 50(1):1–102, 1987. doi: 10.1016/0304-3975(87)90045-4.

Peter Godfrey-Smith. Darwinian individuals. In Frédéric Bouchard and Philippe Huneman, editors, *From Groups to Individuals: Evolution and Emerging Individuality*, number 16 in Vienna Series in Theoretical Biology, pages 17–36. MIT Press, Cambridge, MA, 2013. ISBN 978-0-262-01872-2. URL https://direct.mit.edu/books/edited-volume/3386/chapter/112115/Darwinian-Individuals.

Philip Goff. *Galileo's Error: Foundations for a New Science of Consciousness.* Pantheon Books, New York, 2019.

Philip Goff, William Seager, and Sean Allen-Hermanson. Panpsychism. Stanford Encyclopedia of Philosophy, 2022. URL https://plato.stanford.edu/entries/panpsychism/. Substantive revision Fri May 13, 2022.

Joseph A Goguen. The logic of inexact concepts. *Synthese*, 19(3/4):325–373, 1969.

B. C. Goodwin. *Temporal organization in cells. A dynamic theory of cellular control processes.* London and New York: Academic Press., 1963.

Julian Gough. The egg and the rock. Substack book-in-progress / serial, 2022. URL https://theeggandtherock.com/. Ongoing public non-fiction project on evolutionary cosmology inspired by Smolin; launched 2022 and continuing.

S.J. Gould. *Wonderful Life: The Burgess Shale and the Nature of History.* W. W. Norton, 1990. ISBN 9780393245202. URL https://books.google.at/books?id=q8BUAgAAQBAJ.

CGP Grey. Hexagons are the bestagons. YouTube, 2020. URL https://www.youtube.com/watch?v=thOifuHs6eY&t=326s. Accessed: 2024-06-13.

Cecilia Guerrier-Takada, Katheleen Gardiner, Terry Marsh, Norman Pace, and Sidney Altman. The rna moiety of ribonuclease p is the catalytic subunit of the enzyme. *Cell*, 35(3, Part 2):849–857, 1983.

L.H. Gunderson and C.S. Holling. *Panarchy Synopsis: Understanding Transformations in Human and Natural Systems*. Island Press, 2002. ISBN 9781559633307. URL https://books.google.at/books?id=PXon0AEACAAJ.

Gotthard Günther. *Beiträge zur Grundlegung einer operationsfähigen Dialektik. Band 2: Wirklichkeit als Poly-Kontexturalität*, volume 2. Felix Meiner Verlag, Hamburg, 1979. ISBN 978-3-7873-0460-6.

Anoopum S. Gupta, Matthijs A. A. van der Meer, David S. Touretzky, and A. David Redish. Hippocampal replay is not a simple function of experience. *Neuron*, 65 (5):695–705, 2010. doi: 10.1016/j.neuron.2010.01.034.

Jürgen Habermas. *Auch eine Geschichte der Philosophie: Band 1: Die okzidentale Konstellation von Glauben und Wissen. Band 2: Vernünftige Freiheit. Spuren des Diskurses über Glauben und Wissen*. Suhrkamp Verlag, 2019.

Petr Hájek. *Metamathematics of fuzzy logic*, volume 4. Springer Science & Business Media, 2013.

J. B. S. Haldane. The origin of life. *The Rationalist Annual*, 3:148–153, 1929.

Thomas C. Hales. The honeycomb conjecture. *Discrete and Computational Geometry*, 25(1):1–22, January 2001.

Edward T. Hall. *Beyond Culture*. Anchor Books, Garden City, NY, 1976. ISBN 978-0385124744.

Liz Harper. Understanding earth. *Geological Magazine*, 139(1):101–102, 01 2002. ISSN 0016-7568. doi: 10.1017/S0016756802286513. URL https://doi.org/10.1017/S0016756802286513.

Marc D. Hauser, Noam Chomsky, and W. Tecumseh Fitch. The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598):1569–1579, 2002.

Robert M. Hazen, Dominic Papineau, Wouter Bleeker, Robert T. Downs, John M. Ferry, Timothy J. McCoy, Dimitri A. Sverjensky, and Hexiong Yang. Mineral evolution. *American Mineralogist*, 93(11-12):1693–1720, 2008. doi: 10.2138/am.2008.2955.

G.W.F. Hegel, F. Hogemann, and W. Jaeschke. *Wissenschaft der Logik*. Number Bd. 1 in Gesammelte Werke. Meiner, 1978. ISBN 9783787303823. URL https://books.google.at/books?id=vqj2Qm08tiEC.

Martin Heidegger. *Being and Time*. Harper & Row, New York, 1962. Originally published as *Sein und Zeit* (1927).

Martin Heidegger. *Der ursprung des kunstwerkes*. Klostermann, 2012.

Martin Heidegger and Friedrich-Wilhelm von Herrmann. *Sein und zeit*, volume 2. M. Niemeyer Tübingen, 1977.

Martin Heil and Richard Karban. Explaining evolution of plant communication by airborne signals. *Trends in Ecology Evolution*, 25(3):137–144, 2010. doi: 10.1016/j.tree.2009.09.010.

Heraclitus. *Die Fragmente der Vorsokratiker*, volume 1. Weidmann, Berlin, 6 edition, 1952. Fragments DK 22B12 and DK 22B49a.

Heinrich Hertz and Hermann von Helmholtz. *Die Prinzipien der Mechanik in neuem Zusammenhange dargestellt*, volume 3. JA Barth, 1910.

Erik Hoel. The overfitted brain: Dreams evolved to assist generalization. *Patterns*, 2(5):100244, 2021. doi: 10.1016/j.patter.2021.100244.

Erik Hoel. Causal emergence 2.0: Quantifying emergent complexity. *arXiv preprint*, 2025. doi: 10.48550/arXiv.2503.13395.

Erik P. Hoel. When the map is better than the territory. *Entropy*, 19(5), 2017. ISSN 1099-4300. doi: 10.3390/e19050188. URL https://www.mdpi.com/1099-4300/19/5/188.

Jesper Hoffmeyer. Semiotic scaffolding of living systems. In Marcello Barbieri, editor, *Introduction to Biosemiotics: The New Biological Synthesis*, pages 149–166. Springer, Dordrecht, 2007.

Jesper Hoffmeyer. *Biosemiotics: An Examination into the Signs of Life and the Life of Signs*. University of Scranton Press, Scranton, PA, 2008.

J.S. Hofmeyr. A biochemically-realisable relational model of the self-manufacturing cell. *Biosystems*, 207:104463, Sep 2021. doi: 10.1016/j.biosystems.2021.104463.

D.R. Hofstadter. *Gödel, Escher, Bach: ein endloses geflochtenes Band*. Klett-Cotta, 2006. ISBN 9783608944426. URL https://books.google.at/books?id=4tJGrzR7bEAC.

D.R. Hofstadter. *I Am a Strange Loop*. Basic Books, 2008. ISBN 9780465030798. URL https://books.google.at/books?id=78S6swEACAAJ.

Geert Hofstede. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. SAGE Publications, Thousand Oaks, CA, 2nd edition, 2001. ISBN 978-0803973244.

Thomas Hofweber. Logic and Ontology. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2023 edition, 2023.

John H Holland. Complex adaptive systems. *Daedalus*, 121(1):17–30, 1992.

C. S. Holling. Cross-scale morphology, geometry, and dynamics of ecosystems. *Ecological Monographs*, 62(4):447–502, 1992. doi: 10.2307/2937313.

Alexandra Horowitz. Smelling themselves: Dogs investigate their own odours longer when modified in an "olfactory mirror" test. *Behavioural Processes*, 143:17–24, 2017. doi: 10.1016/j.beproc.2017.08.001.

Melanie-Jayne R. Howes and Monique S. J. Simmonds. The role of phytochemicals as micronutrients in health and disease. *Current Opinion in Clinical Nutrition & Metabolic Care*, 17(6):558–566, 2014. doi: 10.1097/MCO.0000000000000115.

Andreas Hüttemann and Orestis Terzidis. Emergence in physics. *International Studies in the Philosophy of Science*, 14(3):267–281, 2000. doi: 10.1080/09672550010011962.

Julian S. Huxley. *Evolution: The Modern Synthesis*. Allen & Unwin, London, 1942.

Marcelo Isalan. Gene networks and liar paradoxes. *BioEssays*, 31(10):1110–1115, October 2009. doi: 10.1002/bies.200900072.

Abel Jansma and Erik Hoel. Engineering emergence. *arXiv preprint*, 2025. doi: 10.48550/arXiv.2510.02649.

Stanisław Jaskowski. Rachunek zdan dla systemów dedukcyjnych sprzecznych. *Studia Societatis Scientiarum Torunensis, Sectio A*, (5):57–77, 1948.

Rishi Jha, Collin Zhang, Vitaly Shmatikov, and John X. Morris. Harnessing the universal geometry of embeddings, 2025. URL `https://arxiv.org/abs/2505.12540`.

HANS JONAS. A critique of cybernetics. *Social Research*, 20(2):172–192, 1953. ISSN 0037783X. URL `http://www.jstor.org/stable/40969483`.

Sven E. Jørgensen, Brian D. Fath, Simone Bastianoni, João C. Marques, Felix Müller, Søren N. Nielsen, Bernard C. Patten, Robert E. Ulanowicz, and Enzo Tiezzi. *A New Ecology: Systems Perspective*. Elsevier, Amsterdam, 2007. ISBN 9780444531605.

Sven E. Jørgensen, Brian D. Fath, Søren Nors Nielsen, Federico M. Pulselli, Daniel A. Fiscus, and Simone Bastianoni. *Flourishing Within Limits to Growth: Following Nature's Way*. Routledge, London, 2015. ISBN 9781315731445.

A. Juarrero. *Context Changes Everything: How Constraints Create Coherence*. MIT Press, 2023. ISBN 9780262374781. URL `https://books.google.at/books?id=fYyKEAAAQBAJ`.

François Jullien. *The Propensity of Things: Toward a History of Efficacy in China.* Zone Books, New York, 1995. Translation of *La propension des choses* (1992).

Julie Kallini, Isabel Papadimitriou, Richard Futrell, Kyle Mahowald, and Christopher Potts. Mission: Impossible language models, 2024. URL https://arxiv.org/abs/2401.06416.

I. Kant. *Kritik der Urteilskraft.* Number Bd. 1 in Kritik der Urteilskraft. 1797. URL https://books.google.at/books?id=Nu9ZAAAAcAAJ.

Jared Kaplan, Sam McCandlish, Tom Henighan, Tom Brown, Benjamin Chess, Rewon Child, et al. Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*, 2020.

Robert Kaplan. *The Nothing That Is: A Natural History of Zero.* Penguin Books, London, 2000.

S.A. Kauffman. *The Origins of Order: Self-organization and Selection in Evolution.* The Origins of Order: Self Organization and Selection in Evolution. Oxford University Press, 1993. ISBN 9780195079517. URL https://books.google.at/books?id=lZcSpRJz0dgC.

Stuart Kauffman. Answering schrödinger's "what is life?". *Institute for Systems Biology*, 2020.

Stuart A Kauffman. *At home in the universe: The search for laws of self-organization and complexity.* Oxford University Press, USA, 1995.

Lily E. Kay. *Who Wrote the Book of Life? A History of the Genetic Code.* Stanford University Press, Stanford, CA, 2000. The definitive historical account of how information theory, cybernetics, and military cryptography metaphors shaped the development of molecular biology.

Judy Kegl, Ann Senghas, and Marie Coppola. Creation through contact: Sign language emergence and sign language change in nicaragua. In Michel DeGraff, editor, *Language Creation and Language Change: Creolization, Diachrony, and Development*, pages 179–237. MIT Press, Cambridge, MA, 1999.

Luke Kemp. *Goliath's Curse: The History and Future of Societal Collapse*. Viking, London, July 2025. ISBN 978-0241741238.

Jaegwon Kim. Lonely souls: causality and substance dualism. In T. O'Connor and D. Robb, editors, *Philosophy of Mind: Contemporary Readings*, pages 65–78. Routledge, London, 2003.

Simon Kirby, Hannah Cornish, and Kenny Smith. Spontaneous evolution of linguistic structure: an experimental study of cultural transmission. *Proceedings of the National Academy of Sciences*, 105(31):10681–10686, 2008. doi: 10.1073/pnas.0707835105.

Max Kleiber. Body size and metabolism. *Hilgardia*, 6(11):315–353, 1932. doi: 10.3733/hilg.v06n11p315.

Tarja Knuuttila. Modelling and representing: An artefactual approach to model-based representation. *Studies in History and Philosophy of Science Part A*, 42: 262–271, 2011.

Tarja Knuuttila and Andrea Loettgers. Modeling and experimenting: The combinatorial strategy in synthetic biology. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(2):158–169, 2013. doi: 10.1016/j.shpsc.2013.03.011.

Tarja Knuuttila and Andrea Loettgers. Modelling as indirect representation? the lotka–volterra model revisited. *British Journal for the Philosophy of Science*, 68 (4):1007–1036, 2017. doi: 10.1093/bjps/axv055.

Tarja Knuuttila and Atro Voutilainen. A parser as an epistemic artifact: A material view on models. *Philosophy of Science*, 70(5):1484–1495, 2003. doi: 10.1086/377424.

E. Kohn. *Book of Ecclesiastes-Kohelet: A Living Dialogue*. Thesis, University of Cape Town, Faculty of Humanities, Kaplan Centre for Jewish Studies, 1997. URL http://hdl.handle.net/11427/19639.

Saul A Kripke. A completeness theorem in modal logic1. *The journal of symbolic logic*, 24(1):1–14, 1959.

Saul A Kripke. Semantical analysis of modal logic i normal modal propositional calculi. *Mathematical Logic Quarterly*, 9(5-6):67–96, 1963.

Kelly Kruger, Paula J. Grabowski, Arthur J. Zaug, Julie Sands, Daniel E. Gottschling, and Thomas R. Cech. Self-splicing rna: Autoexcision and autocyclization of the ribosomal rna intervening sequence of tetrahymena. *Cell*, 31(1): 147–157, 1982.

Thomas S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, IL, 1962.

G. Lakoff. *Philosophy In The Flesh*. Basic Books, 1999. ISBN 9780465056743. URL https://books.google.at/books?id=iKG1G3UfAOMC.

George Lakoff and Mark Johnson. *Metaphors we live by*. University of Chicago press, 2008.

Kevin N. Laland, Tobias Uller, Marcus W. Feldman, Kim Sterelny, Gerd B. Müller, Armin Moczek, Eva Jablonka, and John Odling-Smee. The extended evolutionary synthesis: its structure, assumptions and predictions. *Proceedings of the Royal Society B: Biological Sciences*, 282(1813):20151019, 2015.

W. David Lambert and C. S. Holling. Causes of ecosystem transformation at the end of the pleistocene: Evidence from mammal body-mass distributions. *Ecosystems*, 1(2):157–175, 1998. doi: 10.1007/s100219900015.

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

Rolf Landauer. Information is physical. *Physics Today*, 44(5):23–29, 1991. doi: 10.1063/1.881299. URL https://pubs.aip.org/physicstoday/article/44/5/23/406580/Information-is-PhysicalThere-are-no-unavoidable.

Lao-tzŭ. *Tao Te Ching*. Penguin Books, London, 1993.

Pierre-Simon Laplace. *A Philosophical Essay on Probabilities*. Dover Publications, New York, 6th edition, 1951. Translated into English from the original French.

Simon A. Levin. The problem of pattern and scale in ecology: The robert h. macarthur award lecture. *Ecology*, 73(6):1943–1967, 1992. doi: 10.2307/1941447.

David Kellogg Lewis. *Convention: A Philosophical Study*. Wiley-Blackwell, Cambridge, MA, USA, 1969.

Richard C Lewontin. The units of selection. *Annual review of ecology and systematics*, pages 1–18, 1970.

Seth Lloyd. *Programming the Universe: A Quantum Computer Scientist Takes On the Cosmos*. Alfred A. Knopf, New York, 2006.

Agnieszka Loboda, Milena Damulewicz, Elzbieta Pyza, Alicja Jozkowicz, and Jozef Dulak. Role of the nrf2/ho-1 system in development, oxidative stress response and diseases: An evolutionarily conserved mechanism. *Cellular and Molecular Life Sciences*, 73(17):3221–3247, 2016. doi: 10.1007/s00018-016-2223-0.

J. Locke. *An Essay Concerning Human Understanding: In Four Books*. Number Bd. 2 in An Essay Concerning Human Understanding: In Four Books. S. Birt, 1753. URL https://books.google.at/books?id=N-4pAAAAYAAJ.

Edward N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, 1963. doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Alfred J. Lotka. Contribution to the energetics of evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 8(6):147–151, 1922. ISSN 00278424, 10916490. URL http://www.jstor.org/stable/84174.

Arthur O. Lovejoy. *The Great Chain of Being: A Study of the History of an Idea*. Harvard University Press, Cambridge, MA, 1936.

E. J. Lowe. Non-cartesian substance dualism and the problem of mental causation. *Erkenntnis*, 65(1):5–23, 2006.

Niklas Luhmann. Autopoiesis, handlung und kommunikative verständigung. *Zeitschrift für Soziologie*, 11(4):366–379, 1982. doi: doi:10.1515/zfsoz-1982-0403. URL https://doi.org/10.1515/zfsoz-1982-0403.

Mairéad MacSweeney, Cheryl M Capek, Ruth Campbell, and Bencie Woll. The signing brain: the neurobiology of sign language. *Trends in Cognitive Sciences*, 12(11):432–440, 2008.

Thomas-F. Mah and George A. O'Toole. Mechanisms of biofilm resistance to antimicrobial agents. *Trends in Microbiology*, 9(1):34–39, 2001. doi: 10.1016/S0966-842X(00)01913-2.

Hazel Rose Markus and Shinobu Kitayama. Culture and the self: Implications for cognition, emotion, and motivation. In *Psychological Review*, volume 98, pages 224–253. American Psychological Association, 1991. doi: 10.1037/0033-295X.98.2.224.

Takahiko Masuda and Richard E. Nisbett. Attending holistically versus analytically: Comparing the context sensitivity of japanese and americans. *Journal of Personality and Social Psychology*, 81(5):922–934, 2001. doi: 10.1037/0022-3514.81.5.922.

H.R. Maturana and F.J. Varela. *Autopoiesis and Cognition: The Realization of the Living*. Boston Studies in the Philosophy and History of Science. Springer Netherlands, 2012. ISBN 9789400989474. URL https://books.google.at/books?id=iOjVBQAAQBAJ.

Marcel Mauss and Henri Beuchat. *Seasonal Variations of the Eskimo: A Study in Social Morphology*. Routledge Kegan Paul, London, 1979.

John Maynard Smith and Eörs Szathmáry. *The Major Transitions in Evolution*. W. H. Freeman/Spektrum, Oxford, 1995.

Ernst Mayr. *Systematics and the Origin of Species from the Viewpoint of a Zoologist*. Columbia University Press, New York, 1942.

John McCarthy. Circumscription—a form of non-monotonic reasoning. *Artificial Intelligence*, 13(1–2):27–39, 1980. doi: 10.1016/0004-3702(80)90011-9.

Warren S. McCulloch. *Embodiments of Mind*. MIT Press, Cambridge, MA, 1965. URL https://mitpress.mit.edu/9780262529617/embodiments-of-mind/.

J.K. McDonough. *Teleology: A History.* Oxford philosophical concepts. Oxford University Press, 2020. ISBN 9780190845711. URL https://books.google.at/books?id=TsnhDwAAQBAJ.

I. McGilchrist. *The Matter with Things: Our Brains, Our Delusions, and the Unmaking of the World.* Perspectiva Press, 2021. ISBN 9781914568251. URL https://books.google.at/books?id=lKrgzwEACAAJ.

Iain McGilchrist. *The Master and His Emissary: The Divided Brain and the Making of the Western World.* Yale University Press, New Haven, CT, 2009.

Vincent S. F. T. Merckx and Sofia I. F. Gomes. Symbiosis: Herbivory alters mycorrhizal nutrient exchange. *Current Biology*, 30(10):R437–R439, 2020. doi: 10.1016/j.cub.2020.04.016.

M. Merleau-Ponty and R. Boehm. *Phänomenologie der Wahrnehmung.* De Gruyter Studienbuch. Verlag Max Niemeyer, 1966. ISBN 9783110068849. URL https://books.google.at/books?id=GuqeAAAAIAAJ.

Jeremy B. Merrill and Will Oremus. Five points for anger, one for a ‘like’: How facebook’s formula fostered rage and misinformation. *The Washington Post*, October 2021. URL https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji-algorithm/. Accessed 2025-07-08.

Stanley L. Miller. A production of amino acids under possible primitive earth conditions. *Science*, 117(3046):528–529, 1953.

Maël Montévil and Matteo Mossio. Biological organisation as closure of constraints. *Journal of theoretical biology*, 372:179–191, 2015.

Alvaro Moreno and Matteo Mossio. *Biological Autonomy.* Springer Netherlands, Dordrecht, 2015a.

Alvaro Moreno and Matteo Mossio. Biological autonomy. *A philo*, 2015b.

C. Lloyd Morgan. *Emergent Evolution.* Williams & Norgate, London, 1923.

David O. Morgan. *The Cell Cycle: Principles of Control*. New Science Press, London, 2007. ISBN 9780878935086.

Mary S. Morgan. Models as mediators: Perspectives on natural and social science, 1999. Develops the view of models as idealized mediating instruments that connect theory with phenomena.

Mary S. Morgan. *The World in the Model: How Economists Work and Think*. Cambridge University Press, Cambridge, 2012. Argues that scientific models function as idealized, constructed "little worlds" that mediate between theory and the real world, enabling exploration, inference, and intervention.

J. Jeffrey Morris, Richard E. Lenski, and Erik R. Zinser. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *mBio*, 3(2):e00036–12, 2012. doi: 10.1128/mBio.00036-12. URL https://journals.asm.org/doi/10.1128/mBio.00036-12.

Matteo Mossio and Leonardo Bich. What makes biological organisation teleological? *Synthese*, 194(4):1089–1114, 2017.

Steve Mould. I made a water computer and it actually works. YouTube video, 2019. URL https://www.youtube.com/watch?v=IxXaizglscw. Accessed: 2025-09-21.

J. Müller. *Handbuch der Physiologie des Menschen: für Vorlesungen*. Number Bd. 2,Teil 1 in Handbuch der Physiologie des Menschen: für Vorlesungen. Verlag von J. Hölscher, 1835. URL https://books.google.at/books?id=87K90AEACAAJ.

Thomas Nagel. *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False*. Oxford University Press, Oxford, 2012. ISBN 978-0-19-991975-8.

Isaac Newton. General scholium. In *The Principia: Mathematical Principles of Natural Philosophy*, pages 939–944. University of California Press, Berkeley, 1999. Includes: "But hitherto I have not been able to discover the cause of those properties of gravity from phenomena; and I frame no hypotheses.".

D.J. Nicholson and J. Dupré. *Everything Flows: Towards a Processual Philosophy of Biology*. OAPEN Library. Oxford University Press, 2018. ISBN 9780198779636. URL https://books.google.com/books?id=OBJbDwAAQBAJ.

Friedrich Nietzsche. *Die fröhliche Wissenschaft ("La gaya scienza")*. Ernst Schmeitzner, Chemnitz, 1882. Erstausgabe.

Yuval Nir and Giulio Tononi. Dreaming and the brain: From phenomenology to neurophysiology. *Trends in Cognitive Sciences*, 14(2):88–100, 2010. doi: 10.1016/j.tics.2009.12.001.

Richard E. Nisbett. *The Geography of Thought: How Asians and Westerners Think Differently...and Why*. Free Press, New York, 2003. ISBN 978-0743255356.

Tingting Niu, Rongrong Xuan, Ligang Jiang, Wei Wu, Zhanghe Zhen, Yuling Song, Lili Hong, Kaiqin Zheng, Jiaxing Zhang, Qingshan Xu, Yinghong Tan, Xiaojun Yan, and Haimin Chen. Astaxanthin induces the nrf2/ho-1 antioxidant pathway in human umbilical vein endothelial cells by generating trace amounts of ros. *Journal of Agricultural and Food Chemistry*, 66, 01 2018. doi: 10.1021/acs.jafc.7b05493.

Emmy Noether. Invariante variationsprobleme. In *Gesammelte Abhandlungen-Collected Papers*, pages 231–239. Springer, 1983.

F. John Odling-Smee, Kevin N. Laland, and Marcus W. Feldman. *Niche Construction: The Neglected Process in Evolution*. Princeton University Press, Princeton, NJ, 2003.

Howard T. Odum. *Environment, Power, and Society for the Twenty-First Century: The Hierarchy of Energy*. Columbia University Press, 2007. URL http://www.jstor.org/stable/10.7312/odum12886.

Susumu Ohno. *Evolution by Gene Duplication*. Springer-Verlag, Berlin, Heidelberg, 1970. doi: 10.1007/978-3-642-86659-3.

Aleksandr I. Oparin. *Proiskhozhdenie Zhizni*. Moskovskii Rabochii, Moscow, 1924. Translated as 'The Origin of Life' in J. D. Bernal (1967), 'The Origin of Life', pp. 199–234, Weidenfeld & Nicolson.

Leslie E. Orgel. The implausibility of metabolic cycles on the prebiotic earth. *PLOS Biology*, 6(1):e18, 2008. doi: 10.1371/journal.pbio.0060018. Critiques metabolism-first origin-of-life scenarios and emphasizes the unresolved requirement for the later emergence of a replicating informational polymer (genetic information), highlighting a central weakness of purely metabolic beginnings.

Elinor Ostrom. *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge University Press, Cambridge, 1990.

Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Paul Christiano, Dario Amodei, and Jan Leike. Training language models to follow instructions with human feedback. *arXiv preprint arXiv:2203.02155*, 2022.

Vilfredo Pareto. *Cours déconomie politique.* F. Rouge, Lausanne, 1896. Vol. I 1896; Vol. II 1897.

Louis Pasteur. Mémoire sur les corpuscules organisés qui existent dans l'atmosphère: Examen de la doctrine des générations spontanées. *Annales des Sciences Naturelles*, 16:1–98, 1861. Reports the swan-neck flask experiments showing sterilized broths remain sterile unless contaminated by airborne dust/microbes, disproving spontaneous generation and supporting biogenesis.

H.H. Pattee. Laws, constraints, and the modeling relation–history and interpretations. *Chemistry & Biodiversity*, 4(10):2272–2295, Oct 2007. doi: 10.1002/cbdv.200790187.

Howard Hunt Pattee and Joanna Rączaszek-Leonardi. *Laws, language and life: Howard Pattee's classic papers on the physics of symbols with contemporary commentary*, volume 7. Springer Science & Business Media, 2012.

Judea Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, Cambridge, UK, 2 edition, 2009. ISBN 978-0-521-89560-6.

P. J. E. Peebles. *The Large-Scale Structure of the Universe.* Princeton University Press, Princeton, NJ, 1980. ISBN 9780691082400. Classic treatment of gravitational instability and the growth and persistence of cosmic density contrasts in an expanding universe.

Charles Sanders Peirce. On a new list of categories. *Proceedings of the American Academy of Arts and Sciences*, 7, 1867. Presented 14 May 1867 to the American Academy of Arts and Sciences. Published 1868.

Charles Sanders Peirce. The fixation of belief. *Popular Science Monthly*, 12:1–15, 1877.

Charles Sanders Peirce. How to make our ideas clear. *Popular Science Monthly*, 12: 286–302, 1878.

C.S. Peirce, C. Hartshorne, and P. Weiss. *Collected Papers*. Number Bd. 1. Harvard University Press, 1931. ISBN 9780674138001. URL https://books.google.at/books?id=HoRfcRUtpnEC.

Elisa Pellegrino, Maarja Öpik, Enrico Bonari, and Laura Ercoli. Responses of wheat to arbuscular mycorrhizal fungi: A meta-analysis of field studies from 1975 to 2013. *Soil Biology and Biochemistry*, 84:210–217, 2015. ISSN 0038-0717. doi: https://doi.org/10.1016/j.soilbio.2015.02.020. URL https://www.sciencedirect.com/science/article/pii/S0038071715000656.

Claudia Peñaloza, Nadine Martin, Matti Laine, and Antoni Rodríguez-Fornells. Language learning in aphasia: A narrative review and critical analysis of the literature with implications for language therapy. *Neuroscience & Biobehavioral Reviews*, 141:104825, 2022. doi: 10.1016/j.neubiorev.2022.104825.

Qiwei Peng and Anders Søgaard. Concept space alignment in multilingual llms, 2024. URL https://arxiv.org/abs/2410.01079.

E. Peterson. *The Life Organic: The Theoretical Biology Club and the Roots of Epigenetics*. University of Pittsburgh Press, 2016. ISBN 9780822981985. URL https://books.google.at/books?id=WeEZDgAAQBAJ.

Andrew Pickering. *The Cybernetic Brain: Sketches of Another Future*. University of Chicago Press, Chicago, IL, 2010. URL https://press.uchicago.edu/ucp/books/book/chicago/C/bo8169881.html.

Steven Pinker. *The Language Instinct: How the Mind Creates Language*. William Morrow and Company, New York, 1994.

Planck Collaboration. Planck 2018 results. x. constraints on inflation. *Astronomy & Astrophysics*, 641:A10, 2020. doi: 10.1051/0004-6361/201833887. Precision CMB anisotropy results supporting inflationary, nearly scale-invariant primordial perturbations that seed later gravitational structure formation.

Amir Pnueli. The temporal logic of programs. In *18th annual symposium on foundations of computer science (sfcs 1977)*, pages 46–57. ieee, 1977.

Graham Priest. The logic of paradox. *Journal of Philosophical logic*, pages 219–241, 1979.

Graham Priest. Paraconsistency and dialetheism. In Dov M. Gabbay and John Woods, editors, *The Many Valued and Nonmonotonic Turn in Logic*, volume 8 of *Handbook of the History of Logic*, pages 129–204. North-Holland, 2007. doi: https://doi.org/10.1016/S1874-5857(07)80006-9. URL https://www.sciencedirect.com/science/article/pii/S1874585707800069.

Ilya Prigogine and Isabelle Stengers. *Order out of chaos: Man's new dialogue with nature*. Verso Books, 2018.

Arthur Prior. *Past, present and future*. Oxford University Press, 1967.

Arthur N Prior. *Time and modality*. Oxford University Press, 2003.

Arthur S. Reber. Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6(6):855–863, 1967. doi: 10.1016/S0022-5371(67)80149-X.

C.D.C. Reeve. *Cratylus*. Hackett Classics. Hackett Publishing Company, 1998. ISBN 9780872204164. URL https://books.google.at/books?id=6XHyqdh2AgkC.

Raymond Reiter. A logic for default reasoning. *Artificial Intelligence*, 13(1–2): 81–132, 1980. doi: 10.1016/0004-3702(80)90014-4.

Roger Riddell. Uc berkeley finds anthropologist did not commit plagiarism, jan 2013. URL https://www.highereddive.com/news/

uc-berkeley-finds-anthropologist-did-not-commit-plagiarism/95409/. Accessed: 2025-01-10.

Howard Robinson. Dualism. In S. Stich and T. Warfield, editors, *The Blackwell Guide to the Philosophy of Mind*, pages 85–101. Blackwell, Oxford, 2003.

Lesley J. Rogers, Paolo Zucca, and Giorgio Vallortigara. Advantages of having a lateralized brain. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(Suppl 6):S420–S422, 2004. doi: 10.1098/rsbl.2004.0200.

R. Rosen. *Life Itself: A Comprehensive Inquiry Into the Nature, Origin, and Fabrication of Life*. Complex Ecological Systems (COUP) Series. Columbia University Press, 1991. ISBN 9780231075640. URL https://books.google.at/books?id=Q1lABK_C9X8C.

Marleen Rozemond. *Descartes's Dualism*. Harvard University Press, Cambridge, MA, 2002.

Tim Räz. The silent hexagon: explaining comb structures. *Synthese*, 194(5):1703–1724, 2017. ISSN 00397857, 15730964. URL http://www.jstor.org/stable/26748789.

Carl Sachs. Naturalized teleology: Cybernetics, organization, purpose. *Topoi*, 42 (3):781–791, 2023. ISSN 1572-8749. doi: 10.1007/s11245-022-09851-9. URL https://doi.org/10.1007/s11245-022-09851-9.

Jenny R. Saffran, Richard N. Aslin, and Elissa L. Newport. Statistical learning by 8-month-old infants. *Science*, 274(5294):1926–1928, 1996. doi: 10.1126/science.274.5294.1926.

Olivier Sartenaer. Synchronic vs. diachronic emergence: A reappraisal. *European Journal for Philosophy of Science*, 5(1):31–54, 2015. doi: 10.1007/s13194-014-0097-2.

Leonard J Savage. *The foundations of statistics*. Courier Corporation, 1972.

Daniel Schmachtenberger and Liv Boeree. Daniel schmachtenberger & liv boeree on AI, moloch & capitalism. YouTube video, "Win-Win with Liv Boeree" channel, April 2023. URL https://www.youtube.com/watch?v=KCSsKV5F4xc. Runtime 1 h 31 min.

E.D. Schneider and J.J. Kay. Life as a manifestation of the second law of thermodynamics. *Mathematical and Computer Modelling*, 19(6):25–48, 1994. ISSN 0895-7177. doi: https://doi.org/10.1016/0895-7177(94)90188-0. URL https://www.sciencedirect.com/science/article/pii/0895717794901880.

E.D. Schneider and D. Sagan. *Into the Cool: Energy Flow, Thermodynamics, and Life*. University of Chicago Press, 2005. ISBN 9780226739366. URL https://books.google.at/books?id=u4wb1EQWQD0C.

E. Schrödinger. *What is Life? the Physical Aspect of the Living Cell & Mind and Matter*. University P., 1967. URL https://books.google.at/books?id=9RHbAAAAMAAJ.

Thomas A. Sebeok. "zoosemiotics". *American Speech*, 43(2):142–144, 1968.

Thomas A. Sebeok. *Global Semiotics*. Indiana University Press, Bloomington, IN, 2001.

Rico Sennrich, Barry Haddow, and Alexandra Birch. Neural machine translation of rare words with subword units. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1715–1725, Berlin, Germany, August 2016. Association for Computational Linguistics. doi: 10.18653/v1/P16-1162.

Martin A. Shields. *The End of Wisdom: A Reappraisal of the Historical and Canonical Function of Ecclesiastes*. Penn State University Press, 2006. ISBN 9781575061023. URL http://www.jstor.org/stable/10.5325/j.ctv1bxh3j0.

Joan B. Silk, Jacinta C. Beehner, Thore J. Bergman, Catherine Crockford, Anne L. Engh, Liza R. Moscovice, Roman M. Wittig, Robert M. Seyfarth, and Dorothy L. Cheney. The benefits of social capital: close social bonds among female baboons enhance offspring survival. *Proceedings of the Royal Society B: Biological Sciences*, 276(1670):3099–3104, 2009. doi: 10.1098/rspb.2009.0681. Epub 2009 Jun 10.

George Gaylord Simpson. *Tempo and Mode in Evolution.* Columbia University Press, New York, 1944.

David Skrbina. *Panpsychism in the West.* MIT Press, Cambridge, MA, 2005.

Sally E. Smith and David J. Read. *Mycorrhizal Symbiosis.* Academic Press, London, 3 edition, 2008. ISBN 978-0-12-370526-6.

Lee Smolin. *The Life of the Cosmos.* Oxford University Press, New York, 1997. ISBN 9780195126648. Develops cosmological natural selection (fecund universes): universes reproduce via black holes with slightly varied constants, yielding an evolutionary account of cosmic initial conditions.

Lee Smolin. *Time Reborn: From the Crisis in Physics to the Future of the Universe.* Houghton Mifflin Harcourt, Boston, 2013.

Elliott Sober and David Sloan Wilson. *Unto Others: The Evolution and Psychology of Unselfish Behavior.* Harvard University Press, Cambridge, MA, 1998.

Henry Somers-Hall. *Hegel, Deleuze, and the Critique of Representation: Dialectics of Negation and Difference.* State University of New York Press, Albany, 2012. doi: 10.1515/9781438440101.

George Spencer-Brown. *Laws of Form.* George Allen & Unwin, London, 1969.

James Steele and Natalie Uomini. Humans, tools and handedness. In *Stone knapping: the necessary conditions for a uniquely hominin behaviour.* 2005.

Achim Stephan. Varieties of emergentism. *Evolution and Cognition*, 5(1):49–59, 1999.

Philip S. Stewart and J. William Costerton. Antibiotic resistance of bacteria in biofilms. *The Lancet*, 358(9276):135–138, 2001. doi: 10.1016/S0140-6736(01)05321-1.

Joan E. Strassmann, Yong Zhu, and David C. Queller. Altruism and social cheating in the social amoeba dictyostelium discoideum. *Nature*, 408(6815):965–967, 2000. doi: 10.1038/35050087.

Galen Strawson. Realistic monism: Why physicalism entails panpsychism. *Journal of Consciousness Studies*, 13(10-11):3–31, 2006. URL https://philpapers.org/rec/STRRM-2.

Mauricio Suárez. An inferential conception of scientific representation. *Philosophy of Science*, 71(5):767–779, 2004. doi: 10.1086/421415.

Richard Swinburne. *The Evolution of the Soul*. Clarendon Press, Oxford, revised edition, 1997.

Charles Taliaferro. *Consciousness and the Mind of God*. Cambridge University Press, Cambridge, 1996.

Adly Templeton, Tom Conerly, Jonathan Marcus, Jack Lindsey, Trenton Bricken, Brian Chen, Adam Pearce, Craig Citro, Emmanuel Ameisen, Andy Jones, Hoagy Cunningham, Nicholas L Turner, Callum McDougall, Monte MacDiarmid, C. Daniel Freeman, Theodore R. Sumers, Edward Rees, Joshua Batson, Adam Jermyn, Shan Carter, Chris Olah, and Tom Henighan. Scaling monosemanticity: Extracting interpretable features from claude 3 sonnet. *Transformer Circuits Thread*, 2024. URL https://transformer-circuits.pub/2024/scaling-monosemanticity/index.html.

Evan Thompson. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Harvard University Press, Cambridge, MA, 2007. ISBN 9780674057517. URL https://www.hup.harvard.edu/catalog.php?isbn=9780674057517.

Michael Thompson, Richard Ellis, and Aaron Wildavsky. *Cultural Theory*. Westview Press, Boulder, CO, 1990.

Michael Tomasello. *The Cultural Origins of Human Cognition*. Harvard University Press, Cambridge, MA, 1999.

Michael Tomasello. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press, Cambridge, MA, 2003.

A. M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1):230–265, 1937. doi: 10.1112/plms/s2-42.1.230.

Alan Mathison Turing. The chemical basis of morphogenesis. *Bulletin of mathematical biology*, 52:153–197, 1990.

Denys Turner. *The Darkness of God: Negativity in Christian Mysticism.* Cambridge University Press, Cambridge, 1995.

László Fejes Tóth. Über das kürzeste kurvennetz, das eine kugeloberfläche in flächengleiche konvexe teile zerlegt. *Math. Naturwiss. Anz. Ungar. Akad. Wiss.*, 62:349–354, 1943.

Tobias Uller, Armin P. Moczek, Richard A. Watson, Paul M. Brakefield, and Kevin N. Laland. Developmental bias and evolution: A regulatory network perspective. *Genetics*, 209(4):949–966, 2018. doi: 10.1534/genetics.118.300995. URL https://doi.org/10.1534/genetics.118.300995.

Nilesh Vaidya, Michael L. Manapat, Irene A. Chen, Ramon Xulvi-Brunet, Eric J. Hayden, and Niles Lehman. Spontaneous network formation among cooperative RNA replicators. *Nature*, 491(7422):72–77, 2012. doi: 10.1038/nature11549. Experimental demonstration that small RNA fragments form mutually catalytic networks capable of collective self-replication, with network-level persistence not reducible to a single autonomous replicator.

J. W. Valley, A. J. Cavosie, T. Ushikubo, D. A. Reinhard, D. F. Lawrence, D. J. Larson, P. H. Clifton, T. F. Kelly, S. A. Wilde, D. E. Moser, and M. J. Spicuzza. Hadean age for a post-magma-ocean zircon confirmed by atom-probe tomography. *Nature Geoscience*, 7:219–223, 2014. doi: 10.1038/ngeo2075.

Giorgio Vallortigara and Lesley J. Rogers. Survival with an asymmetrical brain: advantages and disadvantages of cerebral lateralization. *Behavioral and Brain Sciences*, 28(4):575–589, 2005. doi: 10.1017/S0140525X05000105.

John Vervaeke. Relevance realization meets dynamical systems theory. *Awakening from the Meaning Crisis*, 30, 2024. URL https://meaningcrisis.co/episode-30-relevance-realization-meets-dynamical-systems-theory/. Lecture series on cognitive science and meaning.

K.E. von Baer. *Über Entwickelungsgeschichte der Thiere: Beobachtung und Reflexion ...* Über Entwickelungsgeschichte der Thiere: Beobachtung und Reflexion. Bei den Gebrüdern Bornträger, 1828. URL https://books.google.at/books?id=ev7OAAAAMAAJ.

Heinz von Foerster. *Understanding Understanding: Essays on Cybernetics and Cognition*. Springer, New York, NY, 2003. URL https://link.springer.com/book/10.1007/b97451.

J.W. von Goethe, D. Kuhn, R. Wankmüller, and C.F. von Weizsäcker. *Goethes Werke: Naturwissenschaftliche Schriften I*. Goethes Werke / hrsg. von Erich Trunz. Beck, 1994. ISBN 9783406084935. URL https://books.google.at/books?id=7XeRmyR8pSUC.

A. von Humboldt and R. Gröls-Verlag. *Kosmos Band 1: Entwurf einer physischen Weltbeschreibung*. Edition Werke der Weltliteratur. Gröls Verlag, 2022. ISBN 9783966376655. URL https://books.google.at/books?id=2rCTEAAAQBAJ.

John von Neumann. *Theory of Self-Reproducing Automata*. University of Illinois Press, Urbana, IL, 1966.

Günter Wächtershäuser. Before enzymes and templates: theory of surface metabolism. *Microbiological Reviews*, 52(4):452–484, 1988.

C. H. Waddington. The epigenotype. *Endeavour*, 1:18–20, 1942.

Conrad Hal Waddington. *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. George Allen & Unwin, London, 1957.

Ullrich Wagner, Steffen Gais, Hartmut Haider, Rolf Verleger, and Jan Born. Sleep inspires insight. *Nature*, 427(6972):352–355, 2004. doi: 10.1038/nature02223.

Andreas Weber and Francisco Varela. Life after kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1:97–125, 01 2002. doi: 10.1023/A:1020368120174.

Michael Weisberg. Three kinds of idealization. *The Journal of Philosophy*, 104(12): 639–659, 2007. doi: 10.5840/jphil20071041240. Distinguishes multiple forms of idealization used to simplify and manage tradeoffs in scientific modeling.

Michael Weisberg. *Simulation and Similarity: Using Models to Understand the World*. Oxford University Press, Oxford, 2013. Systematic account of scientific modeling and idealization, emphasizing how models simplify, omit, and distort aspects of target systems for specific epistemic aims.

Eric W. Weisstein. Honeycomb conjecture. MathWorld, 2010. URL `https://mathworld.wolfram.com/HoneycombConjecture.html`. Retrieved December 27, 2010.

G.B. West. *Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies*. Penguin Press, 2017. ISBN 9781594205583. URL `https://books.google.at/books?id=Pf0jDwAAQBAJ`.

Geoffrey B. West, James H. Brown, and Brian J. Enquist. A general model for the origin of allometric scaling laws in biology. *Science*, 276(5309):122–126, 1997. doi: 10.1126/science.276.5309.122.

John A Wheeler. Beyond the black hole. *Some strangeness in the proportion*, pages 341–375, 1980.

John Archibald Wheeler. Information, physics, quantum: The search for links. In Wojciech H. Zurek, editor, *Complexity, Entropy, and the Physics of Information*. Addison-Wesley, Redwood City, CA, 1990.

Alfred North Whitehead. *Science and the Modern World*. Macmillan, New York, 1925.

Alfred North Whitehead. *Science and the Modern World*. Simon and Schuster, 1967.

Alfred North Whitehead. *Process and Reality: An Essay in Cosmology*. Free Press, New York, corrected edition edition, 1978. Original work published 1929.

Alfred North Whitehead. *Process and Reality*. Simon and Schuster, 2010.

Alfred North Whitehead and Bertrand Russell. *Principia Mathematica*. Cambridge University Press, Cambridge, 1910. 3 vols., published 1910–1913.

Norbert Wiener. *Cybernetics: Or Control and Communication in the Animal and the Machine*. MIT Press, Cambridge, MA, 1948. URL https://mitpress.mit.edu/9780262537841/cybernetics-or-control-and-communication-in-the-animal-and-the-machine/.

S. A. Wilde, J. W. Valley, W. H. Peck, and C. M. Graham. Evidence from detrital zircons for the existence of continental crust and oceans on the earth 4.4 gyr ago. *Nature*, 409:175–178, 2001. doi: 10.1038/35051550.

William C. Wimsatt. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Harvard University Press, Boston, MA, USA, 2007.

Jonathan Winawer, Nathan Witthoft, Michael C. Frank, Lisa Wu, Alex R. Wade, and Lera Boroditsky. Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, 104(19):7780–7785, 2007. doi: 10.1073/pnas.0701644104.

L. Wittgenstein. *Tractatus Logico-philosophicus*. International library of psychology, philosophy, and scientific method. Harcourt, Brace, Incorporated, 1922. ISBN 9781989708682. URL https://books.google.at/books?id=w-PWAAAAMAAJ.

Stephen Wolfram. *A New Kind of Science*. Wolfram Media, Inc., Champaign, IL, 2002a.

Stephen Wolfram. *A New Kind of Science*. Wolfram Media, Champaign, IL, 2002b. URL https://www.wolframscience.com/nks/.

Stephen Wolfram. A class of models with the potential to represent fundamental physics. *arXiv preprint arXiv:2004.08210*, 2020. URL https://arxiv.org/abs/2004.08210.

Lotfi Asker Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.

Wei Zhao, Yu-Cai Guo, Ming-Yan Huai, Lily Li, Chi Man, Wolf Pelletier, Han-Lin Wei, Rong Yao, and Jin Niu. Comparison of the retention rates of synthetic and natural astaxanthin in feeds and their effects on pigmentation, growth, and health in rainbow trout (oncorhynchus mykiss). *Antioxidants*, 11(12), 2022. ISSN 2076-3921. doi: 10.3390/antiox11122473. URL https://www.mdpi.com/2076-3921/11/12/2473.

Ilana Zilber-Rosenberg and Eugene Rosenberg. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiology Reviews*, 32(5):723–735, 2008.

George Kingsley Zipf. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Addison-Wesley Press, Cambridge, MA, 1949.

Roberto Zoncu, Alejo Efeyan, and David M. Sabatini. mtor: from growth signal integration to cancer, diabetes and ageing. *Nature Reviews Molecular Cell Biology*, 12(1):21–35, 2011. doi: 10.1038/nrm3025.

Konrad Zuse. *Rechnender Raum*. Friedrich Vieweg & Sohn, Braunschweig, 1969.