

Hierarchy: The Fundamental Architecture of Complexity

Alyx Dubrow

Contents

Introduction	3
The Hierarchical Symphony of Self	5
The Emergence and Dissolution of Self	7
A Wave in the Universe	7
Coherence and Free Energy	8
Entangled, Yet Coherent	8
Incoherent Immersion	9
Return to the Ocean	10
But What About Entropy?	10
This Too Is the Universe Unfolding	13
The Shape Beneath the Surface: On Emergence and Real	
Structure	16
Meanwhile, at the prison...	16
Two Kinds of Emergence	17
Waypoints in the Climb	18
The Challenge of Mapping Complexity	19
Real Transitions, Not Just Perspectives	19
The Architecture of Complexity	20
Opposing Forces and Gradients: The Architects of Complexity	20
Cosmic Gradients: Gravity vs. Pressure	21
Thermal and Chemical Gradients: Patterns in Motion	21
From Cells to Cognition: Biological Gradients and Feedback	22
Opposing Forces as the Scaffolding of Emergence	23
From Aggregation to Strategy	23
From Sequence to Encoding	24
Catalysis and the Foundations of Memory	25
Scaffolding Complexity	26
Natural Selection: The First Strategy	27
From Memory to Strategy	28
Building Mind Out of Matter	29
Communication as Constraint: How the Neuron Was Born	29

The Brain as a Contrast Engine	30
Proto-Bodies and Attractors: Levin's Broader Lens . . .	31
Resonance: When Perception Becomes Real	32
Resonance Beyond Perception: Stabilizing Behavior . . .	33
Memory as Description: Reconstructing the Past with the Architecture of the Present	35
Memory in a Social Brain	35
The Self: Center of Narrative Gravity	36
The Network of Meaning: Compression at the Apex . . .	36
A Mind Built in Layers	37
The Scaffold of Mind: Language, Writing, and the Rise of Symbolic Thought	38
Language: The Architecture of Recursion	38
Serialization: Hierarchy Constrained by Time	39
Symbolic Technologies: Language as the First Scaffold .	40
Writing: The First External Mind	40
Morphology and Memory: A Trade-Off in Linguistic Evolution	41
From Writing to Formal Thought: Logic and Mathematics	42
Thinking as Internalized Speech	42
.	43
Looking Ahead: The Emergence of Culture	43
The Emergence of Collective Intelligence	43
Beyond the Individual Mind	43
From Imitation to Coordination: How Minds Forge Culture	44
The Acceleration of Collective Exchange: Internet and AI	46
The Building Blocks of Culture	46
Cultural Evolution Without Reductionism	47
Collective Intelligence in Action	48
When Systems Take on a Life of Their Own	49
From Mind to Meaning	49
Is a Mapping Between Levels Ever Infinitely Complex?	50
A First Taste of Self-Reference	50
From Proofs to Programs	51
Mapping Complexity with Feedback	51
Beyond Logic and Code	52
Infinite Complexity as a Limit Case	52
Closing Reflection	53
Afterword	53
Written by Many Layers of Mind	53
My Class With Edsger Dijkstra	54
The Mirage of Innate Modules	55
Life as a Neuron: The Ubiquitous Struggle for Relevance . . .	60
The Architecture of Meaning: How Hierarchy Collapses into a Network	62
Morphology, Memory, and the Shifting Burden of Complexity	65

What Is a Number: How Minds Turn Patterns Into Mathematics	68
Why Are the Integers So Deep?	71
The Zombie View of AI	74
Coincidence, Connection and the Uncanny Unity of All Things	77
References	78

Introduction

Throughout this book, we will take a naturalistic approach to understanding reality. By this I mean grounding our exploration in principles and insights drawn from scientific inquiry, while also acknowledging the complexity and wonder that emerge from these foundations. In the spirit of Sean Carroll’s poetic naturalism, we will recognize that there are many valid ways of describing the universe, each with its own appropriate domain (Carroll 2016). Physics, chemistry, biology, psychology, and culture each tell their part of the story, allowing us to investigate the universe and the mind as interconnected systems without invoking explanations that fall outside the scope of natural laws.

For as long as I can remember, I have been refining my understanding of what I am and the world around me. Whether gazing at the stars or pondering the complexities of the human mind, each experience has contributed to a deeper appreciation of the interconnectedness of all things. The ideas in this book are the culmination of decades of reflection, learning, and conversation: a personal journey to uncover patterns that connect everything from the smallest particles to the largest structures in the universe.

In our current understanding, the deepest known layer of physical reality consists of quantum fields, a dance of fluctuating probabilities and interactions invisible to any perceiver. These quantum fields are the bedrock of existence, governing the behavior of particles and the forces that bind them. Nothing exists at this level that can directly perceive these fields. Yet, as these fields interact, they give rise to stepwise layers of complexity: atoms, molecules, and eventually the emergent patterns we recognize as objects. Through a process of aggregation, where simpler patterns are combined and organized into more stable and enduring forms, reality takes on lasting structure through time. This hierarchical layering not only creates the physical objects we interact with but also sets the stage for the emergence of consciousness and perception.

This book begins with a simple but powerful claim: complexity can only arise and persist when built in hierarchical layers. Each layer must constrain the degrees of freedom of the parts below it, stabilizing them into new structures that can in turn support even greater complexity. This is not just a convenient way of organizing the world; it is the only way complexity can scale without collapsing under its own weight.

Perceivers (beings like us) emerge only at this higher level of complexity, shaped by the same hierarchical processes that create the objects we perceive. At the quantum level, neither perceivers nor objects exist. Instead, both arise together

as part of the foundational architecture of the universe. This interconnectedness underscores the profound unity between the observer and the observed, a relationship that becomes increasingly intricate as complexity builds.

Hierarchy is a simple yet profound concept that is hard to refute in its elegance. Much like Darwin's idea of natural selection producing fitness to the environment, the notion of hierarchy facilitating the emergence of more complex patterns seems perfectly natural. Hierarchical structures offer a framework through which the universe organizes itself, creating layers of stability and complexity that build upon one another. From the microscopic interactions of particles to the vast organization of galaxies, hierarchy ensures that each level of complexity is both dependent on and transcends the one below it.

Emergence, the phenomenon where larger or more complex entities arise through interactions among smaller or simpler entities, plays a crucial role in this hierarchical framework. In this book, we adopt the perspective of weak emergence, which holds that higher-level phenomena arise from the interactions of lower-level components and have causal powers only at their own level of description. This stands in contrast to strong emergence, which suggests that higher-level phenomena can influence or alter the behavior of their lower-level constituents through downward causation. For example, a flock of birds can weave striking patterns in the sky, but the flock itself doesn't change what any bird is doing. The pattern arises from the simple rules each bird follows locally.

By focusing on weak emergence, we emphasize the continuity and interconnectedness of hierarchical levels, allowing us to explore complex systems without invoking unexplained or irreducible properties.

Consider the way individual cells combine to form tissues, tissues build organs, and organs create entire biological systems. This hierarchical organization is mirrored in the cosmos, where subatomic particles form atoms, atoms assemble into molecules, and molecules build the stars and galaxies that populate our universe. Just as a symphony is composed of individual notes that create harmonious music, the universe is orchestrated through hierarchical layers that produce the intricate and beautiful complexity we observe.

Understanding hierarchy not only unveils the underlying order of the cosmos but also illuminates the processes that shape our minds and societies. By exploring hierarchical structures, we can gain insights into everything from the formation of consciousness to the dynamics of ecosystems and the progression of civilizations.

Whether you have a background in science or are new to these topics, this book is crafted to guide you through complex ideas with clear explanations and relatable examples. The concepts we explore may seem abstract or challenging at first, but they are deeply relevant to understanding ourselves and the world we inhabit. I aim to make these ideas accessible to people who may have never thought about them before, to share the excitement and wonder that comes with seeing the world through the lens of hierarchy.

Hierarchy is not a human invention, nor merely a feature of how we describe the world. It is a real structural necessity. Emergent levels of description: atoms, cells, thoughts, societies, are not arbitrary shortcuts. They must exist, because complexity must stabilize itself layer by layer if it is to persist. These levels are discoverable precisely because they are real.

As we journey through these layers of complexity, from physics to mind to meaning, we will uncover the hidden frameworks that connect us to the universe and to each other. In this sense, the book is itself an act of poetic naturalism: a weaving of scientific insight and human imagination into a single narrative. My hope is that, as you turn these pages, you will see not only the science of hierarchy but also its poetry.

Join me on this journey to uncover the hidden hierarchies that connect us to the universe, and discover how understanding these patterns can transform the way we see ourselves and the world we inhabit.

The Hierarchical Symphony of Self

Imagine embarking on an intricate journey, one that delves deep into the very essence of your being. This exploration unveils the majestic hierarchy woven into the fabric of who you are, traversing from the grandeur of your entire being down to the quantum whispers that define existence itself.

At first glance, you are a singular entity, a body moving through the world, experiencing life through senses and emotions. Yet this unity is a harmonious symphony composed of countless parts, each playing a vital role in maintaining the whole. Your body is an intricate system, a masterpiece of biological engineering, where every component works in concert to sustain life.

Delving deeper, your body is composed of organs, each specialized for distinct functions. The heart pumps life-sustaining blood, the lungs facilitate the exchange of oxygen and carbon dioxide, and the kidneys filter impurities from your blood. These organs uphold your physiological stability, ensuring that each process necessary for survival operates seamlessly.

Beneath the surface of each organ lies a tapestry of tissues. Muscular tissues enable movement, nervous tissues transmit signals, and epithelial tissues provide protective barriers. These tissues are composed of groups of cells working in unison, their collective efforts giving rise to the specialized functions of their respective organs. Collaboration is key; each tissue type contributes its unique strengths to the overarching purpose of the organ.

Zooming in further, tissues are made up of microscopic cells, each a living being in its own right! These cells are not solitary. They communicate incessantly, exchanging signals to coordinate actions and respond to environmental changes. Each cell strives to survive and thrive within its niche, supported by a cohort of like-minded neighbors. They are the unsung heroes, tirelessly maintaining

the delicate balance that keeps your body functioning. Imagine: your body is actually made of tens of trillions of individual living entities!

Within each cell, the scene becomes even more intricate. Cells are composed of billions of organic molecules, each serving specific roles. These molecules act as tiny machines, orchestrating countless processes. They shunt materials to where they are most needed, maintain cellular stability, and facilitate the complex biochemical reactions that underpin life. These molecular machinists work with precision, ensuring that every cellular function operates smoothly.

Peering deeper into the molecular structure, we find atoms, the fundamental building blocks of matter. Atoms are composed of protons, neutrons, and electrons, particles that move within defined patterns. These particles are not static; they are dynamic perturbations of probabilities within quantum fields that permeate the universe. The very essence of your being is intertwined with the vast, probabilistic tapestry of the cosmos.

At the most fundamental level, everything you are is a manifestation of quantum fields, fluctuating energies that extend throughout the universe. These fields give rise to particles, which in turn form atoms, molecules, and the myriad structures that compose your body. The quantum dance is a silent yet profound force, underpinning the existence of everything from the smallest particle to the vastness of galaxies.

Turning inward to the seat of your consciousness, your mind, lies another layer of hierarchy. Here exists a vast network of specialized cells known as neurons. These neurons are the architects of thought, each connecting to thousands of others, forming unimaginable networks of information processing. They communicate through electrical impulses and chemical signals, orchestrating the symphony of your perception and understanding.

These neuronal networks are organized into hierarchies, much like the layers we have explored so far. At lower levels, simple signals are processed. Higher levels integrate this information into complex thoughts, memories, and emotions. This hierarchical structure allows your mind to perform sophisticated tasks, like reading this chapter or listening to an audiobook, by coordinating countless interactions across different levels of organization.

As you engage with this text, your brain's hierarchical networks are at work. Sensory neurons process the visual or auditory information, while higher-order neurons interpret and understand the meaning behind the words. This emergent experience, your perception and comprehension, is a testament to the power of hierarchical organization. Each level, from molecules to neurons, contributes to the rich tapestry of your conscious experience.

Reflecting on this hierarchical journey invites a profound sense of connection and wonder. You are not merely a single entity but a complex assembly of interdependent layers, each contributing to the magnificence of your existence. From the quantum fields that form the foundation of matter to the intricate

neural networks that shape your thoughts, every level of hierarchy plays a crucial role in the symphony of self.

Consider the elegance and complexity of your own being. Embrace the interconnectedness of each hierarchical layer, and recognize the beauty inherent in the intricate dance that defines who you are. This understanding is not just a reflection of scientific marvel but a pathway to deeper self-awareness and appreciation.

Your journey through the hierarchies of self-exploration reveals an infinite tapestry of connections and interdependencies. Each layer, from the macrocosm of your body to the quantum whispers of existence, contributes to the unique masterpiece that is you. By appreciating these hierarchical structures, you gain a richer understanding of yourself and your place within the vast, interconnected universe.

As you continue to navigate life, carry with you the awareness of these layers, and let them inspire a deeper connection to the intricate symphony that is your existence.

But if the self is a hierarchy of nested layers, what happens when that structure changes or dissolves? What does it mean to be a “self” at all?

The Emergence and Dissolution of Self

A Wave in the Universe

Imagine a single wave moving across the surface of the ocean. It rises from the sea not as a separate thing, but as a pattern: a brief, coherent form shaped by the forces around it. It has no fixed boundary, no internal engine. And yet, for a time, it travels with grace and distinction, carving its own path through the vastness.

This wave is not apart from the ocean, but it is something. It persists. It maintains its shape as it moves. It is an event in the world that can be pointed to, spoken of, even remembered.

So too are we.

Each of us is a wave, a temporary coherence arising in the sea of the universe. We are made of matter, of energy, of information, none of it unique to us. What makes us a self is not the substance but the form: a fleeting pattern, shaped and sustained by countless interactions, rising for a time into coherence.

We live. We act. We think. We love. And then, like every wave, we return.

But between emergence and dissolution lies a remarkable story: how we sustain our form, how we engage with the world, how we walk the fine line between coherence and entanglement. This is the story of life, not as a thing but as a process: a dance of structure, energy, and identity, held together for a while by the very forces that will one day undo it.

Coherence and Free Energy

To live is to hold a shape against the pressure of dissolution. In physics, coherence is the sustained alignment of a pattern: waves maintaining their phase relationships, systems preserving their internal order. In the quantum world, coherence means the difference between a cloud of probabilities and a distinct interference pattern. In the world of life, coherence means something even more miraculous: a temporary resistance to entropy.

We human beings, like all living systems, maintain our form by consuming free energy, the portion of energy available to do work, to create order. We draw it in constantly through food, through oxygen, through sunlight captured by other forms of life. And we spend it not just on motion or growth, but on maintenance: the intricate choreography that keeps our structure intact.

At the molecular level, we repair ourselves. DNA, the molecule that stores our history and guides our function, is constantly damaged by ultraviolet light, chemical interactions, and replication errors. Yet we generate enzymes, molecular caretakers like DNA polymerase I, that detect and repair these breaks, preserving the code. This maintenance isn't an afterthought; it's at the heart of life. Without energy to maintain the code, the organism unravels.

But this hierarchy of repair doesn't stop with molecules. Molecules sustain cells. Cells sustain tissues. Tissues make up organs, and organs support the integrated being we call a body. Each layer depends on the coherence of the layers below, and each draws on free energy to preserve its structure, repair its damage, and continue the flow of function upward.

It is this hierarchy that allows us to think, act, and be. We are not passive passengers in the universe. We are active processes, each a brief and improbable triumph of form over chaos. But this triumph is conditional. It requires constant input. Without energy, even the most complex organism begins to dissolve. The pattern blurs, the structure fails, and the wave begins to fall.

Life is coherence on borrowed time. We are waves held together by light, flowing for a while across the surface of the deep.

Entangled, Yet Coherent

Though each of us is a wave of coherence, we are never truly alone. We arise in context, in relationship, in response to the forces and forms around us. From the moment of birth, our coherence is entangled with other systems: families, cultures, ecosystems. The self is not a sealed container. It is a porous boundary, exchanging energy, information, and care with the world beyond its skin.

Our coherence is not the coherence of isolation. It is the coherence of connection, a local pattern supported by global structure. A human being cannot grow without others. We require not only food and oxygen, but language, touch, attention. Without these, the developing mind fragments. The wave cannot form.

Even in adulthood, our coherence continues to depend on others. The food we eat is grown by farmers. The ideas we think are shaped by conversations. The very act of reading these words is possible only because of centuries of cultural scaffolding. We are coherent beings, but coherent within greater coherences: families, cities, cultures, civilizations.

These entanglements are not chaotic. They are structured. They form a hierarchy of manageable complexity. I can name the people I depend on. I can trace the systems that sustain me. My embeddedness is real, but tractable. My wave exists within other waves, not lost in the sea but nested in a pattern of relationship.

This is what it means to be a self in the world: to hold a boundary, yes, but to hold it lightly. To be both distinct and dependent. To maintain our own shape, even as we ripple through the lives of others.

Incoherent Immersion

Coherence does not last forever. The structures that sustain us: the molecular machinery, the cellular repairs, the networks of care, depend on a steady influx of free energy. When that flow ceases, the structure falters. Slowly, and then suddenly, the wave begins to lose its form.

Death is not an event but a transition. The body does not vanish; it unravels. The cells no longer maintain themselves. The tissues begin to fail. The boundaries between systems soften. The body returns to chemistry, to soil, to wind, to memory.

And with this unraveling comes a new kind of entanglement.

While we lived, our entanglements were manageable, part of the pattern that made us coherent. After death, those entanglements become intractable. The molecules that once composed our lungs may become part of the grass. The calcium in our bones may be washed into the sea. The information that once formed a personality dissipates into a million causally tangled threads: memories in others, impacts on the world, subtle traces in the future.

The structured hierarchy collapses. The wave no longer has a name, no longer maintains a self-repairing boundary. Its energy diffuses. Its matter rejoins the broader flow. The coherence that once made us a person is gone, not destroyed but no longer localized. What was once a knowable system is now a wide and unknowable entanglement.

This, too, is part of the pattern.

Return to the Ocean

When a wave collapses, the water does not vanish. It simply returns, absorbed back into the body it briefly rose from. So too with us. The self, that astonishing ripple of coherence, eventually lets go. The boundaries dissolve. The shape unwinds. But the substance remains.

The matter that composed us rejoins the world. The energy that sustained us finds new channels. And the information, the patterns that once animated a personality, a voice, a way of being, scatter into a million causal threads. Some are faint echoes, like laughter remembered or habits passed down. Others are woven into the lives of those we've touched. The loop doesn't end. It disperses.

There is no vanishing point, only a return. A diffusion into the deep. The universe that once gave rise to our coherence welcomes us back, not as selves but as sources of further pattern. Our form is gone, but not our trace. Our identity dissolves, but not our influence. We become part of the sea again, not lost but transformed.

And so the wave, having traveled as far as it could, surrenders to the ocean, not in despair but in completion.

And yet, to understand how these temporary coherences arise and fall, we need to look more deeply at the force that seems to resist them: entropy.

But What About Entropy?

When I first shared the idea that self-organization and aggregation are unifying principles behind the emergence of complexity in the universe, someone immediately said, "But doesn't that violate the second law of thermodynamics?" I wasn't surprised. The second law is one of those grand scientific ideas that has drifted into popular culture, often in a garbled form. Many people have absorbed the notion that entropy is simply disorder, and that any emergence of order (anything that looks like pattern, structure, or complexity) must somehow be cheating the rules.

It isn't.

In fact, the emergence of order through self-organization happens precisely because of thermodynamic principles, not in spite of them.

Let's start by clarifying what entropy actually is. In the simplest terms, entropy is a measure of the number of microstates a system can take on without changing how it looks or behaves at the macroscopic level (its macrostate) (Boltzmann 1877). The more microstates consistent with a macrostate, the higher the entropy. In other words, entropy measures probability, not disorder.

Think of a pool table with the balls set in a neat triangle. The macrostate is “triangle arrangement,” but the microstates are the exact positions of the balls within that shape. If you swapped two balls, the macrostate would be unchanged. But when you “break,” the balls scatter across the table. The chance that they will reassemble into a tidy triangle is vanishingly small. Instead, the overwhelmingly likely outcome is that the balls are randomly distributed all over the table. And once the balls are scattered, further rearrangements barely register, because the macrostate (“balls all over the table”) has far more microstates (exact positions of each of the balls) associated with it than the original triangle. The entropy has increased.

The same principle applies to everyday life. A messy room has higher entropy than a tidy one because there are vastly more ways to be messy than to be perfectly ordered. Left alone, systems drift toward states with more possible arrangements (more entropy) simply because those states are overwhelmingly more probable.

Entropy is about probability. Systems with higher entropy appear more disordered because disorder is much more probable than order. Almost any change you make to a system will either keep it equally disordered, or increase its disorder.

This statistical view leads directly to the second law of thermodynamics which says that in an isolated system, entropy cannot decrease. An isolated system is one that exchanges neither energy nor information with its surroundings. But most of the systems we care about, such as stars, rivers, brains, economies, ecosystems, are not isolated. They are open systems, constantly exchanging energy and matter with their environment. And in open systems, local decreases in entropy are not only allowed; they’re expected. The key is that any local increase in order must be paid for by a greater increase in entropy elsewhere.

Consider boiling water. As you heat a pot on the stove, energy enters from below. At a certain point, the water doesn’t just warm uniformly. Instead, it forms convection cells: graceful, rolling loops where hot water rises in the center, cools at the surface, and sinks at the edges (Chandrasekhar 1961). These cells are not chaotic; they are structured, repeating, and beautiful. Why do they form?

Because they are a more efficient way for the system to move heat from the bottom of the pot to the top. The system organizes itself to dissipate energy faster. The local pattern is a dissipative structure, a moment of order born from the drive toward greater global entropy.

It’s a bit like a ballroom dance. Imagine a room of people standing still. Suddenly, music starts to play (the heat source), and small groups begin moving in swirling patterns. These movements aren’t chaotic; they follow a rhythm, a flow. The dance floor becomes alive with circulating motion, but it’s all being driven by the energy entering from outside.

The same phenomenon occurs in the Belousov–Zhabotinsky (BZ) reaction, one of the most hypnotic examples of chemical self-organization (Zhabotinsky 1964). In a shallow dish, you mix a set of chemicals, and instead of quietly settling into equilibrium, the solution pulses with waves of color: orange to blue to clear to orange again. It looks alive, as though the chemicals are breathing. No one commands these cycles; they emerge spontaneously from the reactions themselves.

The BZ reaction is like a chemical drum circle. Imagine a group of drummers seated in a wide ring. At first, the beats are random. But as energy flows into the group, through excitement, anticipation, or momentum, a rhythm begins to emerge. One beat triggers another, and soon the entire circle pulses with coordinated waves. The rhythm wasn’t dictated from above. It arose from local feedback, from each drummer responding to the others. Just like the chemical oscillations, the pattern is sustained only so long as energy continues to flow.

These are not quirks or exceptions. They are examples of a general principle: open systems driven by energy flows often produce ordered, dynamic patterns as a way of increasing overall entropy more effectively.

In biology, we see this writ large. Living organisms are intricate hierarchies of aggregated structures: cells, tissues, organs, organisms, maintained by a constant intake and outflow of energy. The organization of life doesn’t defy the second law. Life is powered by it. We eat, metabolize, and release heat and waste. In doing so, we increase entropy in our surroundings even as we maintain order within ourselves.

And this isn’t only true on Earth. Recent analysis of samples from the asteroid Bennu revealed the presence of organic molecules: sixteen amino acids used by life on Earth, all four DNA nucleobases, and uracil, a key part of RNA (Lauretta et al. 2024). These aren’t simple molecules; they are highly complex arrangements of atoms, with specific three-dimensional structures and delicate chemical properties. That such molecules, fundamental to the machinery of life, could arise spontaneously in the absence of living systems has long seemed improbable to many. And yet Bennu shows us otherwise. These molecules assembled through natural chemical processes, likely catalyzed by minerals as the asteroid formed, and their presence is a clear extraterrestrial demonstration that molecular complexity can arise through physical law alone.

Far from violating thermodynamics, the aggregation of matter into complex molecules, cells, and structures is one of its most compelling consequences. The universe builds patterns because those patterns are good at burning through energy. And the better a structure is at channeling energy flows, the more likely it is to persist, grow, and evolve.

Even evolution itself is a kind of ratchet. Natural selection preserves and builds upon structures that can survive and reproduce. This too is a kind of aggregation: useful arrangements of molecules lead to cells; useful arrangements of cells lead

to organisms; useful behaviors lead to adaptive populations. At every step, there is order; and at every step, entropy is increasing overall.

Some physicists have proposed a broader principle: that systems naturally evolve toward states that maximize their entropy production (Dewar 2003). Whether or not this principle turns out to be universal, it captures a deep truth. Order doesn't fight entropy. It rides it (a perspective aligned with Sean Carroll's poetic naturalism; Carroll 2016).

When we see a snowflake, a tree, a spiral galaxy, or a human brain, we are seeing the output of energy flows and feedback loops that naturally generate structure. These structures are not anomalies in a universe of decay. They are part of the same choreography of thermodynamics, playing out in time.

Self-organization and aggregation are not only consistent with entropy; they may be the most beautiful ways that entropy moves through the world.

Even as entropy flows through us, shaping and sustaining every system we are part of, it also gives rise to something unexpected: moments. These quiet, vivid instants, seemingly mundane, are, in fact, the universe made visible at our scale.

This Too Is the Universe Unfolding

Nothing lies outside the unfolding.

We tend to imagine the universe as something out there: galaxies, gravity, equations, stars. But the universe isn't elsewhere. It's here too, in the late-night arguments and morning routines, in a nervous laugh or the stillness before a kiss.

What follows isn't a theory. It's a series of moments, quiet, familiar, even forgettable... until you see them for what they really are: the universe, unfolding at this scale.

The music was pulsing just enough to make conversation feel like a secret. I was standing in line at the bar, waiting for a drink, talking to someone I saw often at these parties. A friend, even if we didn't spend time together outside of them. The kind of person who seemed game for unusual thoughts.

I told her I sometimes think about things that are a little "out there."
She tilted her head and smiled. "Like what?"

So I gave her one.

"If you think about your body at the level of atoms," I said, "it's actually hard to say where you end and the air begins. There's no clean edge. Your skin is porous. Molecules are drifting out of you, water vapor, carbon dioxide, and other particles are drifting in. The boundary you think of as 'you' is really just a fuzzy zone where the concentration of your atoms starts to thin."

She looked at me, blinking once, then again.

“You’re a cloud,” I said. “Not a statue. And you only seem solid because we’re looking from far enough away that the cloud resolves into a coherent thing we call a person.”

There was a pause.

Then she whispered, “Wow.”

And not in the dismissive way people sometimes say it when they’re trying to back away. She meant it. I could see it on her face. This wasn’t just a cool factoid to her. It had landed somewhere deeper. She looked around the room like the walls might be breathing.

“That blows my mind,” she said.

And maybe mine too, at that moment.

Looking back now, I find myself wondering: what would it mean to actually trace a moment like that all the way down the hierarchy?

I am having a thought. In the physical implementation of that thought, neurotransmitters spill into a synapse, nudging a neuron just enough to fire. Its signal cascades across networks, eventually reaching motor neurons. Muscles contract in the chest and throat, tightening vocal cords, pushing air upward. Vibrations in the air become sound, and the sound becomes a word.

The ripples don’t stop there. Molecules of air, set trembling by the vibration, travel outward in waves until they strike a friend’s ear. The tympanic membrane quivers, tiny bones amplify the signal, and fluid moves in the cochlea. Microscopic cilia bend, triggering neurons to fire in the auditory cortex. A pattern emerges, then stabilizes into recognition: a sound becomes a word, the word becomes a meaning.

And then the chain begins again. Neurons in the friend’s brain fire in response, some sparking memory, others stirring emotion. Signals rush outward, priming muscles of the face and throat. A smile forms, or a reply is spoken, fresh vibrations launched back into the air. What began as a flicker of thought in one brain has crossed the gap between two, become sound, become meaning, and set another mind in motion.

All of it: the thoughts, the words, the recognition, the smile, is nothing more and nothing less than the universe unfolding at this scale, every step carried on the backs of molecules colliding, binding, and releasing through time. Each thought that leaps from one mind to another travels by way of this physical chain of causes.

Step back from that moment. Step back from all of them.

From the couple lying in bed, feet touching under the covers, both scrolling their phones in silence.

From the man yelling at his steering wheel in traffic.
From the child refusing to eat vegetables while a parent negotiates with saintlike patience.
From the teenager hovering over the “send” button on a message that could change everything.
From the old woman slicing a tomato with practiced hands.
From two friends collapsing into laughter over a joke no one else would understand.
From the man alone in the woods, setting up a tent before night falls.
From the woman tapping her leg in a hospital waiting room.
From the father steadying his daughter as she wobbles forward on a bicycle for the first time.
From the friends dancing together at a party, bodies close, eyes closed.

Each of these scenes is ordinary. But none of them are trivial.

In each, something is happening at every scale: molecules shifting, muscles moving, neurons firing, decisions forming, stories unfolding. Physical law, biological structure, personal experience, cultural meaning... all layered into the shape of a single moment.

Take a person driving a car down the highway. At one level, it’s a body inside a machine, heading somewhere. But look closer, and you’ll see a cloud of particles loosely defining the body, seated within another cloud of particles we call a vehicle, both gliding across the surface of a third cloud: road, earth, crust. Every particle in motion, every one responding to local interactions.

And if you follow the chain of causes: why this particle is here, why it moves this way, you’ll trace it back through endless interactions, moment by moment, a causal lattice reaching all the way to the early universe. The car, the road, the driver: just one more consequence of those first conditions unfolding, atom by atom, law by law, into this shape, at this time.

All of these moments unfold forward. Never backward.
The mother brushing her child’s hair.
The text message sent.
The first ride without training wheels.

Not one of these will happen the same way again.

Entropy doesn’t just allow these moments. It gives them direction. The arrow of time flows outward from the Big Bang, and we ride its crest. The universe, in all its lawful unfolding, never repeats itself. And that’s what gives each instant its weight, its poignancy, its reality.

Step back again.

From the rooms and the streets, from the forests and the oceans. See a planet alive with motion. A shimmer of light and heat and habit, bodies in motion,

minds in tension, hearts breaking and healing and hoping.

Zoom out, and the patterns blur into weather systems, power grids, seasons. The Earth turns. The Moon pulls at the tides. The Sun burns in silence.

Zoom out further. A spiral arm in a whorl of stars. A galaxy among billions. Each one spinning, expanding, retreating.

Back through the galaxies, through the afterglow of ancient light, to the edge of what can be seen. To the beginning, where our models reach their limits, where the familiar fabric of space and time dissolves.

And from that moment until now, every atom, every collision, every pulse of radiation, every self-replicating molecule, every neural circuit, every word, every glance, every goodbye has been part of a single, lawful, irreversible process.

This is what the universe looks like at our scale. Not in theory. Not in metaphor. In fact.

You, reading this now, are one thread in that incomprehensibly vast tapestry. And not an accidental one.

You are what happens when matter, governed by law, becomes complex enough to feel, to notice, to care.

You are the universe, briefly aware of itself.

The Shape Beneath the Surface: On Emergence and Real Structure

“Nature is not only more complex than we think, it is more complex than we can think.” — J.B.S. Haldane

When we look at a murmuration of starlings swirling in the sky, we don’t see thousands of birds following local rules, we see a single living form in flight. A wave curling onto the shore is not the story of trillions of molecules; it is simply a wave. A sentence, once understood, no longer feels like a string of sounds, it becomes a thought. These are glimpses of emergence: the appearance of structure, pattern, or behavior that arises from the interaction of simpler parts.

Meanwhile, at the prison...

Inmate: “It isn’t fair that I’m locked up! I just read Sapolsky (2023). The universe is deterministic. I couldn’t help but commit the crime!”

Guard: “Fair enough. But if the universe is deterministic, then I can’t help but keep you locked up.” (*shrugs*)

This exchange may sound glib, but it points to a deeper issue. The inmate is appealing to physics, the guard to society. Both descriptions are valid, but

they belong to different levels of explanation. At the most fundamental level, physics describes evolving states through mathematical laws. No particles ever say “must” or “because.” Causation itself is an emergent notion, a way we make sense of regularities at scales where prediction and intervention matter. At the social level, our causal vocabulary shifts again, and we talk about choice, intention, and responsibility.

The paradox arises only when one vocabulary is mistaken for the other. Physical determinism should not be taken to erase human responsibility, and moral judgments should not be projected downward onto atoms and synapses. Recognizing that each level carries its own indispensable concepts is the key to understanding emergence. Higher-level descriptions are not illusions layered on physics, but real structures that organize how the world works at their scale.

With this in mind, we can now turn to some varieties of emergence. What they all share is this: every emergent level brings its own vocabulary, a language suited to the entities and relationships that are real and useful at that scale.

Two Kinds of Emergence

Not all emergence runs equally deep. Some emergent patterns, what I will call *descriptive emergence*, simply give us a more efficient way to describe a system without introducing fundamentally new kinds of complexity. A center of gravity makes a mechanical system easier to reason about, but it is just a mathematical shorthand. In fluid dynamics, we average over the chaotic motions of particles and speak instead of pressure, temperature, and flow.

Another example is the principle of least action in physics. Instead of tracking every step-by-step force and acceleration, we can summarize an entire trajectory as the one that minimizes (or in technical terms, extremizes) the quantity called “action.” Think of action as the cost of each part of the trajectory. When you sum up all of these costs, nature will usually have picked the one where that sum is as small (or as stable) as possible.

This principle feels almost teleological: systems behave as if they “know” the end and take the most efficient path. But it is simply another way of describing the same underlying mechanics. Like pressure or center of gravity, it is a higher-level law that emerges as a remarkably stable and predictive shorthand.

These are stable, useful regularities that make prediction and explanation possible, but they do not add new structures or agents to the world. Because these descriptive properties are stable and measurable, we can build entire theories around them. We don’t need to know the momentum of every molecule to predict how a gas exerts pressure on a container, or track every atom in a pendulum’s arm to calculate its swing. Higher-level quantities like pressure or motion are so consistent that they become the natural language for reasoning at that scale.

Other cases run deeper. This is what I will call *constructive emergence*. Here, new patterns of interaction give rise to new forms of stability: self-contained entities with their own properties and dynamics. A molecule can exist only if atoms can bond. A cell can function only when molecules self-organize into membranes and metabolic cycles. Thought can take hold only when neurons form circuits. In each case, the system crosses a threshold and exhibits properties that arise from the interaction of its parts, properties that cannot be found in the parts alone.

Physicist P. W. Anderson made this point forcefully in his classic 1972 paper *More Is Different*, arguing that at each new level of complexity, the whole “becomes not only more than but very different from the sum of its parts” (Anderson 1972, 393).

Constructive emergence is how complexity is built in the first place, through successive layers of stability that make further elaboration possible. Deep complexity never appears all at once, It grows in stages. Stability is the scaffolding of complexity.

Waypoints in the Climb

Because complexity builds upon stable layers, we can often discover those layers. They become waypoints, places where we can pause, name, and model what’s happening. These emergent levels are not arbitrary human conveniences; they reflect real, constructed structure built up through the system’s own dynamics.

Each level, whether descriptive or constructive, has its own language, shaped by what is stable and causally relevant at that scale. In physics, we speak of fields, particles, or waves depending on context. In biology, we talk of cells, enzymes, and blood flow, not atoms. In neuroscience, we talk about neurons and circuits. In psychology, we invoke attention, memory, and emotion. In society, we shift again to roles, norms, laws, and institutions.

These vocabularies are not optional shortcuts. They reflect real regularities that persist at their respective scales. Long before the discovery of molecules, scientists could make accurate predictions in fluid dynamics using only pressure and temperature. The vocabulary of thermodynamics was sufficient because those properties are stable at that level. You don’t need to track every molecule to explain how steam drives a piston; the language of pressure and temperature suffices.

Medicine offers another striking example. A physician treating pneumonia works with the language of lungs, airways, inflammation, and immune response. They do not need to invoke the chemistry of oxygen binding to hemoglobin, even though those processes are real. At the clinical level, the vocabulary of airflow and infection suffices. If a doctor insisted on explaining your cough in terms of molecular interactions, they would be not only unhelpful but incomprehensible.

Mixing concepts across levels often produces confusion. Speaking of “selfish genes,” for example, is a useful metaphor at the level of evolutionary biology, but becomes misleading if taken to imply that molecules themselves possess motives (Dawkins 1976/1989). Likewise, trying to explain consciousness in terms of individual neurons overlooks the emergent dynamics of networks. Each level demands its own vocabulary because each level stabilizes its own set of causal patterns. This use of different languages at different scales is not a weakness of science but its strength. It allows us to focus on what is real and manageable at a given scale, without being lost in every microscopic detail below.

The Challenge of Mapping Complexity

One might ask: if everything is made of particles, can’t we just describe the world from the bottom up? In theory, yes. In practice, no. The problem is *mapping complexity*. This is a term I will introduce here to capture the difficulty of connecting different levels in a tractable way, that is, in a way that can actually be managed or reasoned through.

Some mappings are short and relatively straightforward. For example, eye color is largely determined by a single gene that controls the production of pigment. The relationship is close to one-to-one: the gene changes, and eye color changes in a predictable way. Other traits, however, are vastly more tangled. Height, for instance, is influenced by hundreds of genes, each making a small contribution, and by environmental factors such as nutrition. Intelligence is even more complex, shaped by many genes interacting with one another and with cultural and developmental contexts. Although levels are causally linked, the pathways between them can be so intricate that direct derivation becomes impractical.

This is because many mappings are rarely linear or one-dimensional. A small change at a lower level can yield a large change at a higher level... or none at all. Variables interact in nonlinear ways, creating cascading influences that defy simple cause-and-effect tracing. Explaining traits like height or intelligence from genes alone means confronting a dense web of causal interactions.

Mapping complexity is why we need higher-level concepts in the first place. Just as a city map needs neighborhoods and routes, not the position of every grain of sand, science needs conceptual handholds at each stable level. These abstractions are not arbitrary; they are the only workable way to reason, predict, and act at that scale.

Real Transitions, Not Just Perspectives

It’s tempting to think of molecules, cells, minds, and societies as cognitive conveniences imposed by limited human minds. And in part, they are. But

the deeper truth is that we are discovering, not inventing, these levels. They correspond to real transitions in stability and coherence.

We don't talk about quark clusters in anatomy because their influence has already been absorbed into more stable configurations: atoms, molecules, cells, and so on, up through the layers to organs. At the level of organs and tissues, what matters is not the arrangement of fundamental particles but the integrated function of biological modules.

These stable units: neurons, genes, proteins, economic agents are the load-bearing elements in the structure of complexity. Each marks a plateau of stability from which further complexity can grow.

The Architecture of Complexity

The world is not flat, not in space, and not in structure. It is layered. It builds itself stage by stage, with each layer resting on the stability of the one below. This is what we mean by hierarchy: a structure of emergence, systems built from systems, each level bringing new regularities into being.

At each layer, new patterns stabilize, new entities cohere, and new ways of understanding become necessary. These are more than merely viewpoints; they are discoveries. Our languages change because reality changes. What becomes meaningful at each level is what remains stable, structured, and causally effective.

While every emergent level is grounded in lower-level dynamics, it cannot be reduced without loss. Reduction becomes useless in the face of mapping complexity. Higher-level concepts are indispensable; they are the only way we can effectively interface with reality at that level.

This is the central theme of this book: complexity is built in stages, and each stage requires stability before the next can arise. To understand complex systems is to recognize these waypoints of emergence as real structure. The rest of this book explores how this principle plays out: in physics, biology, cognition, and culture. It begins here, with a simple idea: emergent levels exist in the world, and we can learn to see the shape beneath the surface.

Opposing Forces and Gradients: The Architects of Complexity

Emergent levels of reality don't arise in a vacuum. They are sculpted by forces that push and pull, bend and resist. Whenever opposing forces meet, they create gradients: differences across space or scale that channel energy and matter into new patterns. These gradients may be the crucibles where complexity is born.

From the swirling clouds of Jupiter to the layered crust of our planet, gradients formed by competing forces carve stable structures out of chaos. Gravity

compresses matter inward, while quantum pressures or thermal expansion push outward, balancing collapse with resistance. This dynamic tension doesn't smooth everything into uniformity; instead, it establishes the gradients that can sustain rich, stable hierarchies.

This chapter explores how opposing forces naturally generate gradients, and how those gradients become the architects of complexity. By tracing this principle across physics, chemistry, biology, and cognition, we will see how the interplay of push and pull creates the stable waypoints of organization upon which much of what we call complexity appears to be built.

Cosmic Gradients: Gravity vs. Pressure

From the earliest moments after the Big Bang, gravity has been the universe's great aggregator. Once the *inflaton field* decayed and the first particles appeared, gravity began pulling matter together, amplifying tiny density fluctuations left by inflation. Over time, these slight imbalances grew into clouds of gas and dark matter, sowing the seeds for galaxies, stars, and planets.

Gravity alone would not lead to stable complexity. Left unchecked, it collapses matter into simple, featureless points such as black holes. Complexity often arises when gravity's inward pull is balanced by opposing forces that resist collapse. In stars, nuclear fusion ignites at the core, generating thermal pressure that balances gravity and creates layered, stable structures. In planets, quantum effects like the Pauli exclusion principle provide what's known as degeneracy pressure, an intrinsic quantum resistance to compression that holds dense matter apart, creating gradients of pressure and density.

These opposing forces don't erase structure; they establish gradients, differences across space that channel energy and organize matter into layered structures. In stars, energy flows from the hot, dense core outward through regions of fusion, convection, and radiation. Each zone settles into a stable balance, maintained by the gradient between gravity's pull and the forces pushing back. Without these dynamic balances, stars wouldn't burn steadily, planets wouldn't form layered interiors, and much of the universe's richness and structure would never emerge.

Thermal and Chemical Gradients: Patterns in Motion

Opposing forces don't just shape stars and planets, they shape the movement of heat and molecules on Earth. When a pot of water is heated from below, thermal energy rises while gravity and cooler temperatures above push back. This interplay creates a vertical temperature gradient, and when the gradient becomes steep enough, the system reorganizes. The water doesn't boil chaotically, it forms convection cells, stable patterns of circulating motion that efficiently transfer heat.

Rather than being imposed from outside, these cells emerge spontaneously as the most stable way to dissipate energy. What was once uniform and still becomes structured and dynamic. Each cell acts as a modular unit, bounded and self-sustaining within the larger flow. Here again, the gradient is both the result of opposing forces and the driver of self-organization.

The same principle plays out on a far larger scale in Earth's atmosphere. Tornadoes and hurricanes are not chaotic accidents, but coherent structures sculpted by gradients. When warm, moist air near the surface collides with cooler, drier air above, sharp differences in temperature and pressure emerge. These contrasts drive towering flows of rising air. With the help of wind shear or the *Coriolis effect*, those flows twist into spirals: a tornado's funnel or a hurricane's vast rotating bands.

Like convection cells in a pot of water, these storms are dissipative structures. They do not fight entropy but serve it, creating temporary order that accelerates the dispersal of heat from ocean or land into the sky. For a time, imbalance takes shape, an eye, an eyewall, a funnel. These structures are intricate and terrifying, persisting as long as the gradients endure. When the imbalance finally wanes, the storm dissolves back into turbulence.

The same drama of gradients shaping flows is written not only in the sky, but in water and stone, deep beneath the ocean. In hydrothermal vents deep in the ocean, hot, alkaline vent fluids meet cold, acidic seawater, forming steep gradients in acidity, temperature, and redox (a measure of how readily molecules donate or accept electrons) (Russell, Hall, and Martin 2010).

These differences drive flows of ions and molecules, and in the porous mineral walls of vents, those flows become localized reaction spaces. Here, prebiotic chemistry may have taken hold, simple molecules transforming, concentrating, and cycling in a process powered by the very gradients that define the vent environment.

The stability of these reaction zones depends on the persistence of the gradient. Heat and chemistry flow continuously through constrained channels, creating pockets where complexity accumulates. These structures, like convection cells, represent early waypoints of biological hierarchy, forming long before genes or cells existed.

From Cells to Cognition: Biological Gradients and Feedback

In living systems, gradients go beyond the environmental; they are actively maintained. Within cells, molecular pumps use chemical energy to build and sustain electrochemical gradients across membranes. The most familiar example is the proton gradient in mitochondria: protons are pumped across the inner membrane, creating a steep potential. When they flow back down through ATP

synthase, they generate usable energy in the form of ATP (Mitchell 1961). Life, at its core, runs on this kind of structured imbalance.

Neurons do something similar. Ion gradients maintained across the cell membrane allow for precise voltage changes. These changes, action potentials, enable fast, targeted communication across long distances. Without the maintained gradient, no signal would fire. The brain's ability to think, remember, and act all rests on the stability of opposing forces: energy-driven pumps versus diffusion.

At the organismal and ecological levels, gradients abound. Hormonal gradients guide development; nutrient gradients shape root and microbial growth; temperature and pressure gradients drive circulation. In each case, complexity is built on persistent, regulated difference.

Opposing Forces as the Scaffolding of Emergence

Across the universe, from stars to cells to societies, complexity rarely arises from calm or chaos alone; it often arises from tension. Where opposing forces meet, they create the conditions for gradients to form. And where gradients persist, they give rise to structured flows, organizing matter and energy into layered, stable configurations.

In stars, gravity's relentless pull inward is balanced by the outward pressure of fusion. In fluids, rising heat meets the stabilizing force of gravity, forming convection cells. In cells, ion pumps build and maintain gradients that store free energy for thought and motion. In ecosystems and societies, competition and constraint generate dynamic balances that sustain diversity and order.

Each of these systems finds stability in the continuous interplay of push and pull. Gradients are engines of organization, shaping complexity through directed flow. And when that flow becomes self-sustaining, a new level of structure often emerges: one that can persist, replicate, and even adapt.

The recurring pattern is clear: opposing forces carve gradients, gradients sustain flows, flows give rise to structure, and structure, once stabilized, becomes a new layer in the unfolding hierarchy of complexity.

As gradients stabilize and give rise to structure, they lay the groundwork for something even more profound: the emergence of systems that not only endure, but remember. The next chapter explores how structure becomes a medium for encoding: how memory, information, and strategy emerge from the scaffolding that gradients provide.

From Aggregation to Strategy

In the earliest phases of the universe, matter existed in a relatively undifferentiated state. Particles such as protons, neutrons, and electrons floated freely,

with no stable connections to one another. Even when atoms formed, they were simple, uniform, and interchangeable. A single hydrogen atom is much like any other, and without lasting arrangements between them, there was little in the way of meaningful difference.

As the universe cooled and expanded, something profound began to happen: particles started to stick together. Atoms formed bonds to become molecules. Not all combinations were possible. Some were unstable and broke apart, but others held. The rules of physics didn't just allow aggregation, they favored certain arrangements over others. The universe began to explore what was stable.

This principle, aggregation guided by stability, is the first step in the emergence of complexity. Simple units join to form larger structures, but not at random. The interactions are constrained by properties like charge, shape, and energy minimization. These constraints act like filters, selecting for some patterns while excluding others. Out of this selective process, structure begins to emerge.

But structure alone is not yet enough. Two molecules might be stable when bound together, but without any variation in form or function, the system remains uniform, flat, featureless. As we saw in the previous chapter, complexity needs gradients: differences that can drive process and transformation. What unlocks the next stage here is a different kind of gradient, not one of energy or concentration, but of sequence. When molecules, especially long, chain-like ones, can join in different orders, each with different properties, structure becomes capable of carrying difference. And when differences in structure lead to differences in behavior, we begin to see the glimmer of something more: information.

From Sequence to Encoding

A long chain of molecules arranged in a specific order is not just a bigger molecule, it's a pattern. And once patterns can differ, they can also begin to matter.

Imagine a soup of monomers, simple molecular building blocks, each capable of bonding to others in a linear chain. If every chain is identical, the system contains no usable difference. But if some sequences are more chemically reactive than others, if one sequence can catalyze a reaction that another cannot, then the system begins to respond to variation. A chain's structure now determines its function.

This is the beginning of encoding. The order of elements in the chain encodes a specific potential: what the sequence can do, what it might trigger, what outcomes it makes more likely. A single change in that sequence might render it inert, or alternatively, make it dramatically more effective at catalyzing a reaction, binding to a partner, or triggering a downstream effect. At this point, the system is no longer operating purely on the level of physical assembly; it

has entered the realm of representational structure. One thing now stands for another: a pattern stands for a possibility.

It's worth emphasizing that this is not yet symbolic representation in the way humans use language. There is no external interpreter reading the molecule like a sentence. But the functional consequences of variation serve as a primitive kind of semantics. Different arrangements lead to different effects, and that is enough to begin building a logic of selection.

Some sequences persist not just because they're stable, but because they're useful. This is the first glimmer of memory in nature: a structure that carries the record of past interactions in its form. If a sequence helps speed up a useful chemical reaction (a kind of catalyst) it may persist longer simply because the reaction it enables tends to recreate or preserve the conditions that favor its own survival.

It does not yet replicate, but it endures, not through copying, but through continued relevance. Usefulness, in this early context, is what allows structure to persist.

Catalysis and the Foundations of Memory

At this stage, we do not yet assume reproduction or heredity. Even in a world of random molecules, not all patterns behave the same. Some chains fold into stable shapes that resist degradation and persist longer. Others may have a different kind of advantage: they can catalyze reactions, helping nearby molecules form bonds more efficiently. This does not guarantee that the catalyst itself survives, but in some cases, its activity changes the environment in ways that support its continued presence. If a molecule helps stabilize its surroundings, or even contributes, directly or indirectly, to the formation of more molecules like itself, it may begin to persist through feedback, not just durability.

These are fragile steps, but important ones. They show how functional patterns, those that do something, can begin to shape the conditions of their own survival, even before life exists. In this sense, chemistry begins to favor function before biology formalizes it.

For example, in a reaction known as the formose reaction, simple molecules like formaldehyde can spontaneously form sugars (Breslow 1959). As the reaction proceeds, some of the intermediate sugars begin to catalyze later steps, creating a self-reinforcing loop. It is not life, but it is a hint of something life-like: a feedback cycle where products help sustain the process that produces them.

Beyond purely chemical reactions, the environment itself also supplied catalysts.

Even without biology, the early Earth offered natural catalysts that could give molecules a head start. Clay minerals, with their layered structures, can adsorb (bind onto their surfaces) and align simple organics, effectively bringing them into

close proximity. In doing so, clays can act a bit like primitive enzymes, holding molecules in place long enough for reactions to occur that might otherwise be too slow or unstable (Cairns-Smith 1982). Metal sulfides such as iron and nickel sulfide, abundant in the crust, can drive redox chemistry, providing a spark for the formation of more complex organic molecules (Wächtershäuser 1988). These geochemical catalysts don't guarantee survival, but they do demonstrate how simple molecules could have been nudged toward persistence. Under the right conditions, the environment itself supplied scaffolding that allowed helpful configurations to stick around longer, adding to their stability.

Scaffolding Complexity

What makes this kind of persistence important is not only that patterns stick around, but that they begin to stack. A molecule that catalyzes another's formation does more than survive; it participates in a network. A primitive vesicle, a tiny bubble of fat molecules, can shield fragile molecular structures from the external environment. In doing so, it doesn't merely contain; it selects, helping certain configurations persist while others degrade. A clay surface, with its finely layered structure, is more than a passive foundation: it actively binds certain molecules and encourages others to form in place. By templating, it participates in the system, biasing the landscape toward some structures over others.

As these effects accumulate, they begin to interlock. Catalysts shape the chemical environment; the environment shapes which catalysts persist. Compartments isolate reactions; those reactions in turn alter the compartments. What starts as a scattered set of interactions gradually becomes a coordinated system, where the parts reinforce and constrain each other in stable, repeatable ways.

This is how constructive emergent complexity begins. It has no master plan, but rather, it has layers of feedback that preserve structure, amplify function, and scaffold new possibilities. Once structures begin to affect their own survival, the system is no longer passively shaped by the laws of chemistry. It becomes historically shaped. The present carries traces of what has worked before, and those traces condition what can happen next.

At a certain point, the interactions among structures become rich enough that the system as a whole behaves differently than any of its parts. New patterns of organization emerge, ones that cannot easily be reduced to the components alone. These emergent levels are not arbitrary; they are real thresholds, each marking a step in the architecture of complexity.

These networks of catalysts and compartments laid the groundwork for something even more powerful. Once patterns were not only stabilized but also copied, a new dynamic entered the scene. Stability alone could not explain the richness of life. What was needed was a mechanism that could harness variation and

accumulate it across generations. This is where Darwin's great insight enters: the principle of natural selection.

Natural Selection: The First Strategy

What transforms persistence into evolution is replication. Once a structure can not only endure but also be copied, a new logic takes over. This is Darwin's great insight: when replicators vary, and when those variations affect how successfully they are copied, the world begins to change in a systematic way. Natural selection is the process by which patterns that are better at surviving and being copied become more common over time.

It is simple in form, almost trivial:

When copies of a structure are made, they are rarely perfectly exact; some variations make further copying more likely, and those variations, carried forward, accumulate over time.

And yet from this humble algorithm comes the full diversity of life.

What makes natural selection so powerful is that it uses history as fuel. Each generation carries forward the traces of what worked before. Molecular sequences that were copied reliably became genes. Genes, bundled into organisms, were tested not only against chemistry but against the demands of environments. Behaviors that aided copying spread. Slowly, the strategies of life accumulated through iteration, without foresight.

Daniel Dennett has called natural selection "Darwin's dangerous idea" and "the universal acid": an insight so simple and substrate-neutral that once grasped, it eats through every domain of thought (Dennett 1995). The same algorithm recurs wherever patterns can be copied, varied, and retained. It does not depend on carbon chemistry, DNA, or even biology at all. It is a blind but inexhaustible generator of design.

This is why the Darwinian logic extends far beyond biology. Richard Dawkins emphasized that genes themselves act like "selfish replicators," spreading if they are good at making more of themselves, regardless of the fate of the organism (Dawkins 1976; 1989). The same perspective applies at higher levels: brains shape their own neural circuits through reinforcement, strengthening connections that prove useful while others fade. What is "copied" here are the connections between neurons, reinforced if useful and pruned if not. Cultures evolve too, as behaviors and beliefs spread unevenly across minds. Even technology, from stone tools refined across countless generations, to machines built on previous designs, to computer code rewritten and repurposed, follows this path of iteration, variation, and retention.

In each case, the core is the same: copying plus selection turns mere persistence into strategy. It allows a system not only to endure, but to adapt.

Natural selection is the first ratchet of strategy in the universe, but it is not the last. Once copying and selection are in play, the door opens to richer forms of memory and adaptation. Patterns can begin to do more than survive, they can respond, adjust, and refine themselves. From this point forward, persistence is no longer passive. It becomes strategy.

From Memory to Strategy

With persistent structures and functional feedback loops in place, systems begin to do more than endure, they begin to respond, adjust, and adapt. When a configuration of molecules influences which reactions occur, and those reactions in turn affect which configurations persist, the system is doing something remarkable: it is processing information. Not in the symbolic way a computer does, but in a chemical way, by allowing structure to shape future structure through feedback.

This is the essence of strategy: a system that modifies its own behavior based on what has worked before. Even before life, even before genes, we see the embryonic form of purpose emerge, not because the system has goals, but because the selective reinforcement of useful patterns creates direction. The past begins to constrain the future by reinforcing patterns that endure.

This, too, is a kind of encoding. The structure of a molecule does not just persist, it means something, because it causes something. Its pattern encodes a function. Once such patterns can be preserved, combined, and reused, the system has entered the realm of functional memory: the foundation of learning, adaptation, and ultimately, agency.

We can now see the larger arc: aggregation yields structure, structure permits variation, variation creates the potential for function, function leads to persistence through feedback, persistence accumulates as a form of memory, and memory enables response, coordination, and strategy.

This layered architecture repeats across nature. The earliest life built on the scaffolding of catalytic networks. Brains built on the architecture of living cells. Human culture built on the architecture of brains. At each level, the capacity to encode difference, preserve it, and use it to shape future behavior is what drives complexity forward.

In the next chapter, we'll explore how this logic of layered function and memory plays out in the brain, a system that not only remembers, but anticipates; not only adapts, but imagines. The story that began with aggregation will now shift from the evolution of structure to the evolution of *mind*.

Building Mind Out of Matter

What is a mind? How does a system of cells come to perceive, remember, and imagine? These questions once seemed unapproachable, philosophical riddles beyond the reach of science. But today, a new picture is emerging. Not a single discovery, but a growing understanding: the mind is not a substance or essence bestowed upon the brain. It is not something separable, transferrable, or stored like a file. Rather, the mind is an emergent process. It is what matter does when organized and active in the right way: a dynamic pattern, inseparable from the physical substrate that gives rise to it.

In this chapter, we'll trace how that organization arises: how neurons, circuits, and memory systems give rise to a layered architecture that can perceive the world, construct a self, and participate in a shared reality, through patterns built atop patterns, encoded in the architecture of life itself.

Communication as Constraint: How the Neuron Was Born

Life becomes more adaptive when its parts can communicate. In a single-celled organism, responsiveness is limited to the signals that pass through its own membrane, inputs from the environment, or at most, simple cues from nearby cells. With the evolution of multicellularity, differentiation appears. One cell may specialize in locomotion, another in digestion, another in repair. This specialization brings power, but also dependence. Coordination becomes essential. Without communication, the differentiated parts fall out of sync and the whole structure collapses.

But communication is not free. It comes with costs, and most importantly, with constraints. Chemical signals diffuse through fluids, but they are slow and imprecise. Direct cell-to-cell contact is faster and more specific, but it comes with a hard limit: geometry. A typical cell, packed within a tissue, may touch only half a dozen neighbors. Its membrane is finite. Its range of influence is narrow.

If information is power, then cells are weak when isolated. It is in their ability to influence others, and to be influenced, that they gain leverage over the world.

This tension, between the need for more communication and the limits of cellular locality, set the stage for one of biology's most astonishing inventions: the neuron.

Michael Levin's work on bioelectric signaling reminds us that all cells in tissues talk to each other (Levin 2021). Neurons are not special because they alone communicate, but because their geometry liberates them from the local constraints of tissue, allowing them to broadcast signals across the body.

The neuron is a cell that sacrifices nearly everything for communication. It devotes its architecture to projection, its energy to signaling, its membrane to the precise transmission of information. With its branching dendrites and long

axon, a single neuron can coordinate with thousands of other cells, some near, some astonishingly far. Where ordinary cells whisper to their neighbors, neurons speak across the body.

This innovation unlocked a new level of complexity. Neurons could transmit signals not only to adjacent cells, but across tissues and even across the entire body. More importantly, they could begin to encode the order of events: who fired first, who fired next. With the neuron, evolution gained a new kind of building block, one that could not just create structure, but encode sequence as well.

The Brain as a Contrast Engine

Even the neuron began as a humble cell, constrained by the same rules of chemistry and geometry as all the others. What made it different, what freed it, was a subtle shift in the architecture of its membrane.

Some cells evolved excitable membranes, thin layers of lipids studded with protein channels that open in response to voltage. These allow ions like sodium and potassium to rush in and out, creating a pulse of electrical activity. A signal. A difference that propagates.

Once this capacity evolved, it was elaborated into the action potential: a regenerative wave that travels down the length of the axon, creating a clean, directional, time-stamped message.

This time-stamping matters. In neural systems, it is not only which neurons fire that carries meaning, but when they fire. Timing determines whether signals converge or miss one another at a downstream neuron, whether rhythmic inputs synchronize into resonance, and whether sequences are perceived as patterns or noise. The action potential is essentially a tiny electrical wave that races down the axon. Like a ripple moving along a rope when you flick one end, it regenerates itself as it travels, keeping the same size and speed no matter how far it goes. This consistency means that each spike carries not only spatial information but also temporal precision. In this way, the nervous system doesn't just transmit signals, it orchestrates them in time.

To be meaningful, however, the signal must reach another cell. This led to the synapse, a specialized junction for transmission. Some synapses used electrical currents directly through gap junctions, but most relied on chemical signaling, releasing neurotransmitters across a narrow cleft to influence the next neuron.

At this point, neurons gained a second axis of richness: not only when to signal, but how. Some synapses were excitatory, increasing the chance that the next cell would fire. Others were inhibitory, suppressing it. What began as communication became computation.

The brain’s architecture came to rely on the balance of these two forces, excitation and inhibition, to create contrast. One of the most powerful patterns in this architecture is the *on-center off-surround* circuit: a network configuration of neurons in which, once central activation is confirmed, it becomes enhanced while surrounding alternatives are actively suppressed.

This circuit is especially prominent in the sensory systems, where it sharpens edges in visual space, highlights changes in sound, and filters touch. Variations of it also appear in systems that select actions, allocate attention, or resolve competing inputs. It is a general strategy the brain uses again and again: amplify the signal, dampen the noise (Grossberg 2021).

The brain is a contrast engine, interpreting the world by filtering difference.

But evolution did not invent ion-channel waves from scratch. Long before neurons, other living systems had already discovered the same trick. To see the deeper continuity, we can widen the lens beyond the brain itself.

Proto-Bodies and Attractors: Levin’s Broader Lens

The brain is not the first place evolution discovered the power of electrical signaling. Long before neurons, bacteria were already using ion channels to coordinate with each other. Recent work by Arthur Prindle and colleagues at UCLA has shown that bacterial biofilms generate potassium waves, enabling thousands of single cells to synchronize their metabolism across the community. What emerges is a kind of “proto-body”: a larger agent with behaviors that no individual bacterium could achieve alone (Prindle et al. 2015; Prindle et al. 2017).

This is constructive emergence in its purest form: new entities with new powers arising from simpler parts. A biofilm is more than the sum of its bacteria, just as a brain is more than the sum of its neurons. The brain is not a direct descendant of bacterial biofilms, but it is a rediscovery of the same ancient trick: ion-channel waves as a medium for building larger, coordinated intelligences.

Of all the body’s organs, the brain has always seemed to me to be the most miraculous in its development. The neocortex, vast and uniform at birth, resembles a general-purpose learning substrate, waiting to be trained by experience. Other regions, such as the brainstem, the cerebellum, and the hypothalamus, come into the world already highly structured, with circuitry for balance and coordination, basic drives, breathing, heartbeat, and homeostasis largely in place. How can one organ be both rigidly predesigned and radically open?

Levin’s systems-level perspective offers a way to understand this. In his view, development is not a blueprint for every detail but a search through a space of possible forms, guided by higher-level attractors. The neocortex’s attractor is that of a general-purpose adaptive sheet; the brainstem’s attractor is that of

immediate survival control. Looked at this way, the apparent miracle of brain development becomes less mysterious. The fine details remain overwhelming, but the higher-level view shows how systems can self-organize toward stable patterns without needing every step pre-specified.

This brings us to one of the most powerful ideas in neuroscience: that perception itself stabilizes through resonance.

Resonance: When Perception Becomes Real

One of the most powerful ideas in cognitive neuroscience, most fully developed by Stephen Grossberg, is that perception becomes stable through a process called resonance (Grossberg 2021). The brain does not passively receive signals from the senses. Instead, it compares incoming data against stored expectations, hypotheses, and memory. When these two streams, bottom-up input and top-down prediction, come into alignment, a loop forms. Activity stabilizes. Competing interpretations fade. The brain locks into a coherent perceptual pattern.

This bottom-up and top-down language reflects the brain's layered architecture. Lower levels of the hierarchy are closer to raw sensation: light hitting the retina, vibrations reaching the ear. Higher levels represent increasingly abstract constructs, such as object recognition, spatial awareness, or emotional salience.

As signals flow upward, they are gradually transformed into more abstract representations. As expectations flow downward, they constrain and refine interpretation. What emerges from this two-way interplay is *recognition*. This is resonance, an agreement between levels of abstraction, a dynamic match that brings perception into focus.

Neuropsychologist Nicholas Humphrey proposed a conceptually similar idea in *A History of the Mind*, describing consciousness as arising from *reverberant activity* in sensory *feedback loops*. Rather than fading as transient signals, neural activity can cycle back on itself, creating a self-sustaining resonance that endows experience with duration and presence. Humphrey framed this in an evolutionary context, arguing that such loops transformed sensation into something biologically significant by giving it temporal depth (Humphrey 1992). Consciousness is not continuous, it occurs in moments. This temporal aspect of resonance may be what we think of as a *moment* of consciousness.

Grossberg, meanwhile, developed the idea of resonance into Adaptive Resonance Theory, providing a formal mathematical and neurobiological model of how resonant feedback loops enable stable perception, learning, and attention (Grossberg 2021).

Imagine looking at a line on a page. Edge detectors near the retina respond to local contrast. Those signals feed up to neurons tuned to detect lines. But the

traffic doesn't only go up. Higher-level neurons send predictions back down: "If this is a horizontal line, I should be seeing this specific pattern below."

When the bottom-up signal and the top-down expectation match, the loop locks in. Competing hypotheses are suppressed by inhibitory surround. The resonance stabilizes.

Perception doesn't work like a video camera. It's not passive recording, but rather, a mutual confirmation between levels of abstraction.

These resonances may stop at a surface, a motion, a feature, or it may propagate upward to the recognition of an object that matters to us. When that happens, we become aware. As Grossberg puts it, "All conscious states are resonant states, but not conversely." In other words, not every resonance yields awareness, but whenever awareness occurs, it is grounded in resonance.

In Grossberg's framework, only certain types of resonances, specifically those involving sensory or emotional features, enter awareness. Others, like those involved in motor planning or spatial mapping, remain unconscious but are no less essential to adaptive behavior.

Conscious perception, then, is not raw activity. It is the emergence of a self-sustaining pattern, a resonance that has survived competition, aligned across hierarchical levels, and been selected for its relevance.

Resonance Beyond Perception: Stabilizing Behavior

Philosophers sometimes worry that consciousness might be *epiphenomenal* (along for the ride with no causal effect of its own). The idea is that while the brain's neurons and molecules carry out all the real work, consciousness is just a shadow cast by those processes. Sean Carroll pushes back on this view. As he has argued, the difference between doing something consciously and doing it unconsciously is not trivial. At the emergent level where human beings act and decide, consciousness clearly has causal powers. Whether or not I consciously notice something changes what I do, just as "tables" change our world even though the concept of a table does not exist at the level of atoms and molecules (Carroll 2025).

Grossberg's framework helps explain why. Resonance does not merely stabilize perception into a coherent image of the world, it can also stabilize behavior. Unconscious processes can run learned routines with impressive efficiency, but they are brittle: they default to the past even when the present context calls for something different. Conscious resonances, by contrast, hold goal-relevant representations active long enough, and integrated widely enough, to override those habits (Grossberg 2021).

Consider driving. If your mind drifts, you may find yourself turning toward home on autopilot rather than continuing toward a new destination. In that

case, unconscious routines have taken over. My own hypothesis is that when you consciously maintain the intention to reach the unfamiliar place, you are sustaining a resonance that extends higher up the hierarchy, into abstract, goal-oriented planning. Skill execution, by contrast, may rely on resonances that remain in lower levels of the hierarchy, never reaching the threshold that makes them conscious. Perhaps what practice really does is to reconfigure the hierarchy: structures that once required higher-level resonance are gradually consolidated into lower levels, where they can be executed smoothly without conscripting higher levels into their stabilizing resonances. This would explain why conscious effort is essential in early training but later fades away, replaced by fluid unconscious execution.

Why not make all resonances conscious, then? Because unconscious resonances are adaptive in their own right. They allow the brain to operate efficiently, conserving energy by restricting integration to narrower networks. They also allow behaviors to run more quickly and precisely than conscious deliberation would permit. Think about reaching for a cup of coffee: if you had to consciously coordinate each muscle movement, the action would be clumsy and halting. Instead, after practice, the behavior becomes a skill, executed smoothly by unconscious routines.

Psychologist Daniel Kahneman has famously distinguished between two modes of thought. Fast thinking, or *System 1*, is automatic, unconscious, and efficient: ideal for practiced skills and familiar contexts. Slow thinking, or *System 2*, is deliberate, conscious, and effortful: necessary when flexibility, reasoning, or novel situations demand attention (Kahneman 2011). Neither mode is sufficient alone. If all our actions were unconscious, we would be locked into habit, unable to adapt. If all were conscious, we would move through life slowly and awkwardly.

Consciousness, in this view, is resonance that climbs high enough in the hierarchy to recruit abstract, temporally extended, goal-relevant representations. It stabilizes not only perception but also memory, planning, and action, ensuring that behavior remains aligned with context rather than swallowed by habit. Conscious states are resonant states that matter at the human scale, where causality is meaningful.

Consciousness, then, is not an inert byproduct. It is the emergent stability that allows perception, memory, and behavior to align. By stabilizing high-level goals against the gravitational pull of habit, resonance makes consciousness both real and causally powerful in the lives we live.

Current (as of 2025) large language models seem to mimic what we think of as consciousness. Some theorists even argue they can pass a version of the Turing test. Their fluency, responsiveness, and breadth of knowledge can give the uncanny impression of consciousness. But a crucial distinction remains.

Large language models operate in discrete turns: a prompt produces a response, then the system falls silent until the next input. Nothing happens in between. Human brains, by contrast, never turn off. As animals embedded in a dynamic

world, we must anticipate, plan, and act continuously. To do this, the brain sustains resonant activity: loops of prediction and confirmation that stabilize perception and action in real time. These resonances give experience its temporal flow: the sense of “now.” Without them, consciousness would collapse into isolated sparks of input and output, as it does in artificial models.

This distinction helps explain both the brilliance and the limitation of current transformer models. Language models show that hierarchical architectures trained on massive input can achieve extraordinary abstraction and creativity. But they also remind us what makes living brains different: not just hierarchy, but the resonance that keeps hierarchy alive.

Memory as Description: Reconstructing the Past with the Architecture of the Present

But what happens when the moment passes?

The brain does not store experience like a video. Instead, evidence suggests it retains structured representations that prioritize meaning over raw detail, highlighting what mattered. These are not perfect recordings but reconstructions shaped by relevance, context, and emotional weight. Often, they resemble narratives: who did what, where, and why. While sensory impressions may be part of what’s stored, the underlying structure seems to reside in higher layers of the hierarchy, layers that can later regenerate elements of the original experience by reactivating perceptual pathways.

This is not merely theory. Cognitive neuroscience increasingly supports the idea that episodic memory is reconstructive. Witness testimony is unreliable. Memories change over time. And neuroimaging shows that the same networks involved in perception and imagination are involved in recall.

We do not store the moment. We store the meaning of the moment, and reconstruct the rest. What’s retrieved is not a static file, but a dynamic pattern, often reactivated through cues that recreate the original resonance.

Language plays a key role here. Our ability to create structured, abstract descriptions may be what makes detailed episodic memory possible. Language is a compression tool, a way to encode and transmit complex structures. And perhaps, over evolutionary time, it was internalized, not just to plan or label, but to remember. We will explore language in more detail in the next chapter.

Memory in a Social Brain

Why would we evolve a memory system that stores experience as narrative?

One reason is social coordination. Humans are not solitary. We live in groups. To function in a social world, we must make sense not only of the world, but of ourselves. We need to tell stories that explain what we’ve done, why we did it, and what it means.

Episodic memory may have evolved not just to help us survive, but to help us belong.

Philosopher Philip Pettit argues that language did not evolve as internal thought made external, but as external communication made internal. First we talked to others, then we learned to talk to ourselves. Episodic memory may follow the same path. First we shared stories, then we began to store them. These shared and internalized stories didn’t just help us survive, they helped us become someone within our social environment. They became the raw material of identity (Pettit 2008).

The Self: Center of Narrative Gravity

And so we arrive at the self, not as a soul, or a locus, or a spark, but as a model.

Daniel Dennett calls the self the center of narrative gravity: a useful abstraction that organizes how we behave, just as the center of mass organizes how objects move (Dennett 1991). The self is not a thing we find inside the brain, it is a pattern we construct across time, across memories, across relationships.

It emerges from the interaction of resonance, memory, and narrative. It is the name we give to the protagonist of our internal story.

And it is social. We build our self-models in response to others. We shape our actions to make them legible. We ask: what kind of person am I? What will they think of me? We imagine how others model us, and model ourselves in turn.

We do not merely remember the past. We rehearse who we are.

The Network of Meaning: Compression at the Apex

At the lower and middle tiers of the brain’s hierarchy, information is organized by levels of abstraction: edges become surfaces, surfaces become objects, objects become concepts. But beyond the level of concepts, something new happens. The mind stops stacking ever-higher layers and begins *folding* relationships back upon themselves.

Through language, we recursively combine existing concepts to generate new ones, each phrase or metaphor a temporary scaffold for a fresh abstraction. “Quantum entanglement,” “social contract,” “neural resonance,” all are higher-order constructions born from strings of earlier ideas. If the hierarchy were extended

linearly, it could, in principle, continue forever. But the brain economizes. Instead of building a new layer for every possible combination, it compresses these recursive relationships into a dense, cross-linked web: the associative network of meaning.

This network doesn't sit above the hierarchy so much as *contain it in compressed form*. Every node represents a concept, but the connections encode how those concepts interact: causally, metaphorically, analogically. Because these relationships are recursive, the network can represent potentially infinite conceptual structures using finite material. A sentence is a path through this network, a traversal that momentarily lights up one thread among countless possibilities. Even the act of coining a new word simply marks the consolidation of one such path into a reusable shortcut.

In that sense, meaning is both hierarchical and lateral. Hierarchical because it depends on layers of abstraction built from perception; lateral because it encodes relationships among those abstractions in a compressed, self-referential web. The network of meaning, then, is not the top of the hierarchy in the usual sense, it is the compact reflection of the hierarchy itself, capable of infinite recombination within finite space.

A Mind Built in Layers

We began with the simplest form of communication, between cells, and arrived at the emergence of identity. At each step, the brain added a new layer of structure: neurons began transmitting signals. Circuits emerged to create contrast. Resonance stabilized perception. Descriptions allowed experiences to be encoded into memory. Narrative gave those memories meaning. Self-models integrated behavior across time. And culture embedded each mind within a larger whole.

Each layer builds on the ones below, adding new functions, new constraints, new possibilities. Matter becomes structured. Structure gives rise to recognizable patterns. And through those patterns, meaning begins to emerge. Mind is not a mysterious essence hovering above or hidden within the brain. It is a layered architecture, each new layer solving the limitations of the one before.

In the next chapter, we'll explore the role of language: the scaffolding that links thought to thought, person to person, and mind to culture. For just as the self emerges from perception and memory, culture emerges from selves in communication.

The Scaffold of Mind: Language, Writing, and the Rise of Symbolic Thought

Language: The Architecture of Recursion

At its heart, every human utterance is an act of structuring. Whether we are stating, questioning, commanding, or exclaiming, we divide the world into subject and predicate: something we are talking about, and something we are saying about it.

The sun rises.

Does the sun rise?

The sun, in its golden stillness above the sea, rises.

That the sun rises is what I was trying to explain.

Each of these is a bounded, meaningful expression. Beneath their surface variation, they share the same deep structure. Even complex or poetic forms of speech adhere to this binary core: there is a thing, and there is something said about it.

What's more, each of these components, subject and predicate, can be elaborated, modified, nested. We can speak of the sun that rises early, or say that it rises slowly and paints the clouds. The layers can multiply indefinitely. This recursive capacity is not a surface feature of language; it is its structural essence. And it is the keystone upon which our most powerful symbolic systems are built: mathematics, music, logic, programming, narrative.

Why did only humans develop this astonishing recursive capacity? One answer may lie in the structure of our brains. The neocortex, the thin, wrinkled sheet covering the brain, is organized into hierarchies. Lower levels process raw sensations and motor commands, while higher levels integrate them into increasingly abstract representations. Compared with other primates, humans have vastly more neocortical surface area for our body size, allowing for higher tiers of hierarchy. At some point, this extra height may have crossed a threshold. Higher levels could now represent not just objects and actions, but relationships between relationships, abstractions of abstractions. Out of this hierarchical height emerged symbolic recursion, and with it, what we now recognize as human language itself.

Once this threshold was crossed, the leap from neural hierarchy to linguistic hierarchy was almost inevitable. Language is not a mere stream of words. It is a hierarchical machine, a structure-building engine capable of generating infinite complexity from finite means. It mirrors the brain that produces it, and it anticipates the cognitive technologies that follow.

Serialization: Hierarchy Constrained by Time

For all its internal structure, language must be expressed linearly. We do not say everything at once. Instead, we unspool meaning in a temporal sequence, one word after another. This is not an artifact of speech; it is a consequence of physics. Time moves forward, and so language must ride its arrow.

This constraint forces language to be serialized, flattened into a single dimension. But the structures it must represent are often two-dimensional. Recursive structures, where each element can expand into multiple sub-elements, can be visualized as branching trees: a trunk that splits into branches, each of which may split again and again. The two dimensions here represent both the depth of the tree (the number of hierarchical levels), and the breadth (the number of elements at each level).

For example, in the sentence “The golden sun feels warm on my skin,” the subject is the sun and the predicate is how it feels. But the fact that the sun is golden is part of the subject, just as warm on my skin is part of the predicate. Each of these phrases can be further elaborated. We can speak of the golden sun in the morning, or say that it feels warm on my bare skin. In principle, this kind of branching can continue indefinitely.

And even though this sentence is written linearly, as speaking or writing requires, the structure it expresses is not intrinsically linear. The fact that the sun is golden could have been stated before or after the word *sun*, or even in a separate clause. Its relationship to the noun is hierarchical, not sequential. Language flattens branching structures into a single dimension, but the mind builds the tree back up.

To make this possible, recursive ideas must be encoded into streams: ordered strings of symbols that unfold in time while preserving structure. Human languages have evolved different strategies to accomplish this, primarily through morphology and syntax. Morphology encodes grammatical relationships within the word itself through inflection, declension, or conjugation. Syntax primarily uses position, word order, adjacency, and linear arrangement.

Many ancient languages, such as Ancient Greek or Sanskrit, leaned heavily on morphology. A single word could encode subject, object, tense, mood, gender, number, and case, allowing for relatively free word order. But as human societies grew more interconnected, and as languages were learned increasingly by non-native speakers (through migration, conquest, or trade), cognitive load became a limiting factor. Adults don’t internalize languages as effortlessly as children. Languages with heavy morphological complexity gradually shifted toward syntactic strategies, favoring simplicity and consistency of word order.

Modern English, for example, relies almost entirely on position:

The dog bit the man.

The man bit the dog.

Same words, radically different meanings, purely through order.

This reflects a broader theme of this book: complex systems adapt to constraints. Language is no different. It shaped itself to the structure of time and to the cognitive limits of the brains that used it.

Symbolic Technologies: Language as the First Scaffold

Once humans had a system for encoding recursive thought into serial form, it could be repurposed. Language, externalized first in speech and then in writing, became a scaffold for new kinds of symbolic expression, systems that inherited its recursive structure but extended its reach and transformed its function.

Mathematics emerged as a language of pure structure. Unlike natural language, it seeks unambiguous precision, allowing relationships to be stated, manipulated, and proven with mechanical rigor. It is recursion distilled into rules, abstraction stripped of ambiguity.

Music offered another path, one that speaks not through logic but through feeling. It weaves patterns through time, nesting repetition and variation into emotional arcs. Like language, it unfolds linearly. But its syntax is rooted in resonance, not grammar.

Logic formalized the structure of inference itself. It allowed thought to be abstracted from content, so that validity could be tested independently of what was being discussed. In doing so, it made the machinery of reason visible.

And code (perhaps the most recent of these symbolic technologies) closed the loop. It enabled symbols not just to represent but to act. Through code, recursive structures gain agency. Symbols generate behavior. Meaning becomes machinery.

What unites all of these is not their subject matter, but their architecture. Each is built on hierarchy, recursion, serialization, and symbolic encoding. These are not separate domains. They are extensions of the same foundational capacity: to structure meaning through time.

Writing: The First External Mind

Spoken language, powerful as it is, is ephemeral. Utterances vanish the moment they are spoken. Meaning relies on presence, memory, and attention.

Writing changed everything.

It made language durable, transforming the recursive structure of thought into something that could be captured in matter and preserved outside the brain.

With writing, language ceased to be a fleeting event and became a persistent artifact, something that could be revisited, refined, and built upon.

The consequences were profound. Ideas could now be stored across generations, their shape unblurred by memory. Complex arguments could be constructed not just in real time, but layer by layer, through reflection and revision. Social structures, religious codes, and historical records gained a kind of permanence that oral tradition could never guarantee. And most remarkably, communication was no longer constrained by space or presence. It could now reach across centuries.

Writing made thought tangible. It allowed recursive structures such as sentences, equations, narratives, to be instantiated outside any brain. For the first time, thought could exist apart from the thinker.

Morphology and Memory: A Trade-Off in Linguistic Evolution

The shift to writing did more than externalize memory. It quietly rebalanced the structure of spoken language itself. Before writing, languages relied heavily on inflectional morphology, the cases, genders, and verb endings that encode grammatical structure. These systems were cognitively expensive, requiring a lot of memorization, but they supported a freer word order, allowing sentences to be reshaped to fit meter, rhythm, and melody. In oral cultures, this flexibility was essential, making language a mnemonic technology, enabling myths, laws, and genealogies to be preserved through song and poetry.

These linguistic differences are not merely theoretical. Ancient poetic traditions: Greek, Latin, Sanskrit, Old Norse, routinely exploited the freedom that inflectional morphology allowed, reshaping word order to fit meter and melody. Modern poets writing in analytic languages often note the opposite: fixed word order makes metrical composition far more difficult, and poetic forms tend to rely more heavily on rhyme instead. The grammatical structure itself shapes what kinds of poetry come naturally to a language.

Writing changed that equation. Once information could be stored outside the mind, languages no longer needed this internal redundancy. Over generations, spoken languages often shed their case systems and complex agreement patterns, shifting toward simpler forms based on word order and helper words. This reduced the cognitive load of learning a language, a change that became especially important in periods of migration and cultural mixing, when many adults had to acquire new languages quickly. Modern English is the poster child of this change.

It may however seem puzzling that ancient languages like Latin, Sanskrit, and Classical Greek, continued to display heavy morphology long after writing emerged. This is not evidence against simplification; it is evidence that writing preserves older stages of the language. Written languages tend to be conservative

fossils, maintaining grammatical structures that everyday speech has already begun to lose. The spoken languages that evolved from Latin, Sanskrit, and Greek all simplified dramatically over time, exactly as the shift to external memory predicts.

Writing offered a more stable layer of memory, and spoken language reorganized around it. What had once been essential internal scaffolding gradually became optional, and over time, much of it disappeared.

From Writing to Formal Thought: Logic and Mathematics

With writing came a new possibility: not just thinking, but thinking about thinking. The ability to revise, compare, and perfect written language allowed humans to begin formalizing their thoughts, extracting structure from speech and building systems governed by explicit rules.

Logic emerged from this process as a way to analyze inference itself. It asked: What follows from what? What makes an argument valid? What are the minimal steps between one proposition and the next? Logic is what happens when recursive speech meets rigor, an attempt to map the very rules of reasoning.

Mathematics built on this foundation but introduced abstraction and generalization. Variables stood in for infinite possibilities. Numbers became manipulable objects. Functions and operations formed recursive systems of extraordinary power.

Neither logic nor mathematics could have developed without writing. Speech alone lacks the stability and manipulability needed to iterate on formal systems. But once thought is written, it becomes a structure to be refined: revised like a sculpture, explored like a landscape.

Thinking as Internalized Speech

One final, crucial insight: our inner monologue, the stream of consciousness we experience as “thinking”, did not precede language. It descended from it.

As philosopher Philip Pettit has argued, language evolved first and foremost as a means of social communication. It was through interacting with others, by sharing, coordinating, persuading, that recursive language took shape. Only later did humans begin to internalize this structure, using it for reflection, rehearsal, and self-guidance (Pettit 2008).

Language was not born from introspection. Introspection was born from language.

Our capacity to speak to ourselves arose from our capacity to speak to others. The recursive loops of dialogue, once external and social, became internal and private. Thought is the echo of communication, turned inward.

This too follows the book's broader arc: structure precedes function. Just as the body enables the brain, and the brain enables the self, so too does external communication enable internal reflection. The recursive nature of language allowed us not only to share ideas, but eventually to have them alone.

Looking Ahead: The Emergence of Culture

Language allowed thought to be shared. Writing allowed it to persist. Logic and mathematics allowed it to scale. But the highest level of this hierarchy is still to come.

From these symbolic foundations arose culture: the emergent system by which humans preserve, transmit, and refine not just information, but identity, value, and meaning. Culture is the level at which complexity becomes civilization, and shared models become institutions.

In the next chapter, we will explore how all these layers, recursive speech, written language, formal reasoning, culminate in culture itself: the highest emergent layer of the human experience.

The Emergence of Collective Intelligence

Beyond the Individual Mind

Highly complex systems are not built all at once. Like a house, they must be constructed layer by layer: a foundation first, then a frame, then walls, then a roof. You cannot begin with a roof or scatter bricks at random and expect stability to emerge. Each level depends on the structural integrity of the one below it.

If you were to look behind the walls of a finished house and see the wooden frame, you wouldn't say, "How strange to find this here!" You'd expect it, because without it the walls wouldn't stand. The same principle applies to every complex structure, from living organisms to computing systems, from thoughts to societies.

This book has traced the ascent of complexity: from matter to life, from life to mind, and from mind to language. Now we turn to something that no single mind can create on its own: culture.

Culture is the realm of shared meanings, tools, narratives, institutions, and values, built not by individuals in isolation but by networks of interacting minds. And just as neurons interacting within a brain give rise to consciousness, minds interacting through language give rise to something greater: collective intelligence.

From Imitation to Coordination: How Minds Forge Culture

Memes: The Genes of Culture When Richard Dawkins introduced the idea of memes in the final chapter of *The Selfish Gene*, he wasn't talking about internet jokes. In his original sense, a meme is any unit of cultural transmission: a melody, a proverb, a fashion, a way of making bread. The idea opened a thrilling new perspective: that cultural evolution might proceed through a process parallel to biological evolution. Just as genes spread through bodies by enhancing their own replication, memes, units of culture, could spread through minds by enhancing their own transmissibility (Dawkins 1976/1989).

A meme can be a tune, a fashion, a recipe, a catchphrase, a superstition, or a belief system. What matters is not whether it is true, good, or useful, but whether it spreads. Truth, goodness, or utility can help, but often a meme's spread depends on other qualities: it might be surprising, catchy, or even irritating enough to stick in the mind.

Memes compete for cognitive and social bandwidth in a Darwinian process shaped by variation, selection, and retention. Think of a recipe: someone learns it from a friend, accidentally swaps one ingredient, and shares their version. Others try it, adapt it in their own way, and pass on the variants they like best. Over time, the more stable, transmissible, and appealing versions outcompete others, not just in food, but across music, fashion, beliefs, and beyond, forming part of the scaffolding of shared culture.

Among the most important memes of all are words themselves. Language is not merely a carrier of memes; it is composed of memes. Words are inherited socially, vary in form and usage, and survive or disappear based on their usefulness, resonance, or stylistic appeal. Language, in this sense, may have evolved mimetically, shaped by countless generations of speakers selecting for clarity, emotional tone, memorability, or rhythm.

The memetic perspective offers a simple but powerful idea: minds do not only invent culture, they transmit and transform it, acting as both processors and propagators in a grand collective loop.

Beyond Imitation: The Cultural Engine of Interaction While imitation and replication are crucial, they are only the beginning. Culture is not merely a memetic overlay on biology; it is a living system of interactions between minds. These interactions, far richer than mere copying, form the machinery of cultural emergence. They include the way communication focuses shared attention, the way teaching and learning transmit skills across generations, the way ritual synchronizes emotions, and the way storytelling and myth bind communities through shared meaning. Conflict and its resolution shape social norms, while

deliberation and distributed cognition allow groups to solve problems no single individual could address alone.

These processes create feedback loops, as minds respond to and reshape the cultural environment that shapes them. Culture is not a one-way street but a dance of mutual influence between minds and memes.

The Strategic Role of Emotions One of the most important forces stabilizing these cultural processes is emotion. Economist Robert H. Frank offers a striking insight here: emotions like love, guilt, pride, and shame act not only as private feelings but as strategic commitment devices. They lock us into behaviors, trust, generosity, cooperation, behaviors that would be difficult to sustain if we acted only on moment-to-moment self-interest.

But emotions are not just internal states; they are also communications. To “emote” is to display what we are feeling, and these displays signal to others that we are committed. A person flushed with anger, eyes narrowing, voice trembling, is not simply experiencing anger but broadcasting it. Likewise, acts of selfless love: a parent’s tireless care for a child, or a partner’s willingness to sacrifice for another, show not only what is felt but also what is promised. Because such signals are hard to fake convincingly (an application of Amotz Zahavi’s principle of costly signaling [Zahavi 1975]; think of the peacock’s tail or the huge horns of male dung beetles), they serve as credible evidence that we really are bound by the commitments our emotions create. Others can rely on us not just because we feel these emotions, but because we cannot easily conceal them, and thus our commitments become visible and trustworthy.

Purely “rational” calculation, balancing costs and benefits at every turn, can easily undermine trust, because it leaves open the possibility of defecting (acting in our own self-interest) whenever a better option appears. Emotions like love, guilt, or pride commit us to a course of action even when short-term incentives might pull us away. That commitment is what allows others to rely on us, making cooperation and long-term relationships possible. In this way, emotions reinforce the trust and reciprocity that allow cultural systems to endure (Frank 1988).

But culture shapes emotion just as much as emotion shapes culture. While certain emotional tendencies, like general reactivity, adventurousness, or fearfulness, may be influenced by biology, the specific forms, triggers, and meanings of emotions are deeply embedded in the cultural fabric. A display of pride, an expression of grief, or an act of generosity carries different weight and implication in different societies. In this sense, emotions are both a product of our evolutionary heritage and a tool honed by culture to stabilize cooperation and shared meaning.

The Acceleration of Collective Exchange: Internet and AI

With the internet, communication became decoupled from location. For the first time, minds could interact in real time regardless of distance, forming communities based not on geography but on shared interest. Niche ideas found global audiences. Collaboration could emerge without institutional barriers. The bandwidth of cultural transmission exploded.

Crowdsourcing emerged: Wikipedia, open-source software, collaborative funding platforms, online activism. These were not just technological changes, but cultural phase transitions.

Now, generative AI joins this landscape. These systems participate in culture not merely as tools, but as agents that can generate, remix, and amplify memes at scale. This marks a qualitative shift in cultural evolution: not just faster, but fundamentally new.

The Building Blocks of Culture

Recursive Hierarchy: Building Tools That Build Us Generative AI didn't appear out of nowhere. It is the product of many layers: physics, logic, materials science, electrical engineering, computer science, deep learning. And beneath all of those layers lies cooperative culture: the scientific enterprise, shared education systems, the open flow of information across generations.

Culture builds tools that amplify and reshape culture itself. Collective intelligence is not static; it bootstraps its own evolution.

Stable Forms in Cultural Space Culture takes shape through many kinds of enduring forms. Memes, the basic units of transmissible culture, travel from mind to mind, with language itself among the most potent. Customs and norms emerge to coordinate behavior, quietly aligning expectations across the group. Institutions give these patterns long life, embedding them in the durable frameworks of governments, religions, and corporations. Narratives and worldviews carry a culture's identity and morality, telling the stories through which a people understand themselves. Art and aesthetic systems speak in symbols, transmitting value and emotion where words may falter. Technologies and tools extend our minds into the material world, and in doing so, reshape the very culture that made them.

Each of these forms is modular, combinable, and adaptable: waypoints of stability in the restless evolution of collective intelligence.

Cultural Evolution Without Reductionism

The evolutionary psychology of the 1980s framed the mind as a collection of genetically evolved modules. This view once seemed compelling, but the success of generative AI shows that minds can perform language, planning, creativity, and abstraction without such innate specialization.

The brain's neocortex, with its relatively uniform structure, appears to support this. Rather than being filled with fixed modules, it may be a general-purpose substrate where functional specialization emerges through learning. Even language centers like Broca's and Wernicke's areas may simply be hubs shaped by proximity and training, not innate blueprints.

This reflects a broader truth: the mapping complexity between genes and high-level cognition is likely too great for evolution to specify directly. What gets passed down are dispositions and constraints; what gets built is culture.

For example, someone might inherit a neurochemical profile that releases more oxytocin and dampens amygdala reactivity, making them more inclined to trust strangers and seek out social bonds. Another might have lower baseline oxytocin and a more reactive amygdala, inclining them toward caution and selectivity in relationships.

Similarly, variations in baseline adrenaline (epinephrine) levels and stress-response sensitivity can shape temperament. A person with a muted adrenaline response may seek novelty, adventure, and high-risk experiences to feel stimulated. Someone with a more easily triggered surge might avoid risk and prefer stable, predictable environments.

These tendencies are just one layer of influence, shaping how a mind meets the world, but they operate within a much larger framework. To see that framework, we need to step back from chemistry and look at the architecture of the brain itself.

A deeper clue comes from the structure of the neocortex. If it functions as a general-purpose computing system, that is, its neurons are organized into cohorts that act as computational units arranged in a hierarchy, then the amount of cortical surface area matters.

The neocortex is a thin sheet of neurons that forms the outer layer of mammalian brains. In humans, it is only a few millimeters thick, yet if unfolded it would spread out to about 2,500 square centimeters, or roughly the area of a large cloth napkin or pillowcase. The wrinkles we see on the surface of the brain are there because this sheet doesn't quite fit inside the skull; it has to be crumpled to fit (Mountcastle 1978).

When I say that groups of these neurons are arranged in a hierarchy, I don't mean from the inside to the outside. Instead, imagine the sheet divided into many small sections. Each section can connect to others, but conceptually they are organized into a hierarchical metastructure. The lowest levels of this

hierarchy are closest to the sense organs and motor areas (areas controlling muscles). These feed “up” to more abstract sections, which in turn feed up to even more abstract sections.

It is a conceptual hierarchy, not a physical one. Lower levels process raw sensory features, while higher levels integrate and abstract them. A larger cortical surface allows for more high-level conceptual modules, extending the hierarchy upward and supporting richer abstractions (Mountcastle 1978; Hawkins and Blakeslee 2004).

As we saw in the previous chapter, the unusual size of the human neocortex provides the raw substrate for more levels of hierarchy than other primates can sustain. At some point, this extra cortical height crossed a threshold: symbolic recursion emerged, and with it, language. What we call a “language module” may not be a prewired gift of evolution, but something constructed anew in each lifetime as cortical hierarchies adapt to cultural input. Evidence for this comes from natural experiments such as the emergence of Nicaraguan Sign Language, where deaf children, with no structured linguistic input, created a full-fledged language within a single generation (Senghas and Coppola 2001; Senghas, Kita, and Özyürek 2004). This reframes the classic “language instinct” view (Pinker 1994). Evolution did not deliver a ready-made module, but a general-purpose cortical architecture capable of constructing such modules when immersed in culture.

In this sense, the neocortex is itself a kind of “blank slate” (Pinker 2002). But a blank slate is not nothing. It is a surface prepared to be written on. Evolution shaped its layered, uniform circuitry so that it can be trained by experience, not preloaded with fixed content. Language, then, is not inscribed in detail in our genes, but grown in each lifetime as cultural interaction writes on this slate. The success of large transformer models in artificial intelligence underscores the point: they achieve language, planning, and creativity not because anyone coded specific modules for those functions, but because a general-purpose substrate was trained on vast input. Our brains may work the same way: structured enough to learn, but open enough to let culture fill in the modules.

While biology sets the stage, providing the cortical architecture and temperamental dispositions, culture is not reducible to genes. It is its own layer of emergence, with forms, dynamics, and histories that evolve according to their own logic.

Collective Intelligence in Action

Cities

Self-organizing, adaptive, and emergent. They regulate flows of people, goods, and information. No one designs a living city—it arises from interacting layers of intention and constraint.

Corporations and Institutions

Goal-driven systems with internal memory, role differentiation, and feedback. They evolve, compete, and learn.

Scientific Communities

The clearest example of distributed cognition. Truth emerges not from a single mind, but from argument, replication, and shared correction across generations.

Religions and Moral Systems

Encoding shared values, worldviews, and practices. They offer identity and meaning, often lasting centuries. They are cultural attractors and stabilizers.

Online Networks and Algorithmic Ecosystems

The newest layer. Memetic propagation accelerated. Platforms like Wikipedia and GitHub show what crowdsourced intelligence can build. Algorithms now shape attention and interaction in real time.

When Systems Take on a Life of Their Own

These systems are not metaphors. They exhibit memory, perception, feedback, and action. They are agents, built on top of minds, but with their own dynamics and lifespans.

From Mind to Meaning

Culture is not a side effect. It is the next level in the architecture of complexity.

As with every prior step in this book, culture emerges through layers: each stabilized by structure, each built on foundations that came before. Language stabilizes thought; memory stabilizes language; shared tools and narratives stabilize culture.

And now, with generative AI, we are building tools that participate in the very system that gave rise to them. We are shaping the next layer of complexity even as we inhabit it.

Hierarchy is the architecture of complexity, the structure through which the universe organizes itself into meaning. From atoms to ideas, it shapes the possibilities of what can exist. We are both the result of this long unfolding and the authors of what comes next.

Culture may feel like the summit of complexity: the point at which shared meaning, tools, and institutions transcend individual minds. Yet even here, the story is not complete. For in culture, and in the minds that sustain it, something new begins to stir. Systems that have risen through layers of emergence start to bend back on themselves, shaping not only what lies above but also what lies below.

The next chapter turns to this deeper paradox: how complexity, at its highest levels, begins to look inward.

Is a Mapping Between Levels Ever Infinitely Complex?

I wasn't at my desk when this idea struck me. I was in the shower, letting thoughts drift, when suddenly two seemingly distant concepts snapped together in my mind. Kurt Gödel's incompleteness theorem and Alan Turing's halting problem are not separate curiosities, but two versions of the same phenomenon. Both reveal what happens when a system becomes powerful enough to describe itself.

Later I discovered that this was not just my intuition: Turing, writing his 1936 paper on computable numbers (Turing 1936), cited Gödel's 1931 incompleteness result (Gödel 1931) as essential background to Hilbert's decision problem. Gödel had shown that no formal system could be both complete and consistent; Turing extended this insight to computation, showing that no universal procedure could decide whether programs halt. But the thrill of the realization, there in the steam, was that these results could be reframed in terms of the central theme of this book. Most mappings between levels are merely difficult, sometimes even intractable. But when feedback folds a system back onto itself, the mapping can become impossible. It is not just hard, it is infinitely complex.

A First Taste of Self-Reference

Most people encounter a version of this problem as children in the form of a simple paradox: "This statement is false." If it is true, then it must be false. If it is false, then it must be true. The paradox arises because the statement refers back to itself, creating an endless loop of contradiction.

Gödel's genius was to smuggle a version of this paradox into arithmetic. He devised a clever coding scheme, now called Gödel numbering, that allowed numbers to represent not just quantities but statements about numbers, and even statements about statements. With this machinery, he built a sentence that effectively says: "This statement is not provable within this system."

To understand why this is explosive, we need to go step by step. If the system could prove this sentence, it would be inconsistent, since it would have proved a claim that asserts its own unprovability. Logical inconsistency means that a system can prove both a statement and its negation, which in practice makes the system collapse into triviality: once you can prove contradictions, you can prove anything. On the other hand, if the system cannot prove the sentence, then the sentence is in fact true, because it correctly says of itself that it is not provable. But this truth is visible only from the next higher emergent level, from outside the system. Inside, it remains forever undecidable.

This was a shock to the mathematical world. For decades, following David Hilbert's 1900 address, many had hoped for complete systems (Hilbert 1900). Completeness meant no gaps, no truths left hanging, a perfect correspondence between the rules and the whole of mathematics. Gödel showed that such completeness was impossible. Any system strong enough to do basic arithmetic must either be inconsistent or incomplete. There will always be truths that the system cannot prove.

From Proofs to Programs

A few years later, Alan Turing carried Gödel's insight into the newborn field of computation in his 1936 paper *On Computable Numbers, with an Application to the Entscheidungsproblem* (Turing 1936). He asked: could there be a universal procedure that, given any program and its input, tells us whether the program will eventually halt?

Imagine such a perfect halting-decider exists. Now imagine giving it a mischievous program that calls the decider on itself, and then does the opposite of what the decider predicts. If the decider predicts the program halts, the program loops forever. If the decider predicts the program loops forever, the program halts immediately. Either way, the decider is wrong. No such universal procedure can exist.

Turing had done for computation what Gödel had done for logic. Gödel showed that some statements are true but unprovable; Turing showed that some programs have no predictable yes-or-no answer about whether they halt. Both works smuggled self-reference inside the system, and both showed that once this happens, undecidability follows.

I first really understood Gödel through Douglas Hofstadter's *I Am a Strange Loop*, where he presents the incompleteness theorem not as a dry technicality but as a profound reflection on self-reference, recursion, and consciousness itself. Hofstadter made Gödel human for me, showing that the strange loops of logic are akin to the strange loops we experience in the mind. As he emphasized, *strange loops*, systems that fold back to model themselves, are not mere curiosities but a deep structural feature of recursion and self-reference (Hofstadter 2007).

Mapping Complexity with Feedback

Now let us bring this back to the central theme of this book. In both Gödel's and Turing's cases, there is a mapping from a lower emergent level to a higher one. From axioms and inference rules we obtain theorems and provable truths. From code and input we obtain runtime behavior: whether a program halts or loops forever.

Ordinary mappings of this kind are tractable. Proofs can be checked line by line. Programs can be executed step by step. Even when they are hard, they remain finite. But once feedback is introduced (statements about proofs, programs about programs) the mapping folds back on itself. The higher emergent level, instead of standing cleanly above the lower, collapses into it. In other words, the system consumes the very context that would have allowed it to be judged from outside.

At that point, the mapping is no longer just difficult. The undecidability is not an accident of poor tools or insufficient power. It is baked into the structure itself. From within, the system cannot settle certain truths or predict certain futures. When higher levels fold back into lower ones, the mapping becomes not merely intractable but infinitely complex.

Beyond Logic and Code

Gödel and Turing might seem like isolated curiosities of mathematics and computer science, but echoes of this phenomenon appear throughout the real world. The brain does not only model the outside world; it also models itself as the one doing the modeling. This reflexivity gives rise to the strange loop of self-awareness. Perhaps this is why the mapping from neural activity to conscious experience feels not merely difficult but, in some sense, formally impossible from within.

Financial markets behave the same way. Predictions change the system they describe. If everyone believes a stock will crash, they sell, and it crashes. This reflexivity means that prediction itself alters the target, creating a moving horizon that resists exact forecasting. The halting problem has its analogue in economics: no model can capture the system completely, because modeling itself is part of the system (Soros 1987).

Language, too, turns back on itself. We use words to talk about words, definitions to define definitions, sentences to describe sentences. Paradox, irony, and infinite regress are not accidents of grammar but structural features of a self-referential system. Here again, the mapping from lower to higher levels becomes unstable once feedback across levels is introduced.

Infinite Complexity as a Limit Case

Most of this book has described emergence as a manageable mapping: molecules into cells, cells into tissues, tissues into organisms, individuals into cultures. These mappings are complex, sometimes intractably so, but they remain finite. Gödel and Turing remind us that there are limit cases. When feedback loops force higher emergent levels to collapse into lower ones, the mapping becomes infinitely complex.

If hierarchy is the fundamental architecture of complexity, then these examples show us its boundaries. Hierarchy works until self-reference makes a level feed back into itself. Then, instead of a new stable layer, we get undecidability.

This is not just a quirk of logic or computation. It is a deep principle of the universe: systems capable of modeling themselves can never fully map themselves. There will always be truths they cannot settle, futures they cannot predict, and mappings that dissolve into infinity.

Closing Reflection

We have traveled from atoms to cultures, building the case that hierarchy is the architecture of complexity. Each level brings new stability, new patterns, new possibilities. But at the very edge, we find a different kind of structure: not a new layer of emergence, but a collapse of levels into themselves. Gödel and Turing revealed this limit in the purest arenas of logic and computation, but the same theme echoes in minds, markets, and languages. These are the places where hierarchy dissolves, where complexity ceases to be merely daunting and becomes infinite.

Perhaps this is the most humbling truth of all: that no matter how far we climb the ladder of emergence, there are always questions that remain undecidable, always mysteries beyond the grasp of any system from within. Hierarchy gives us a universe rich in structure and stability, but at its very edges, it reminds us of our limits. We can map much, but not everything. And perhaps that very impossibility, the horizon that forever recedes, is what keeps the human search for understanding alive.

Afterword

Written by Many Layers of Mind

This book was not written in the traditional sense, by one author alone in a room. It was the product of an extended conversation: between ideas and experiences, between science and philosophy, between a human and a machine. In that sense, it is a small embodiment of the very theme it explores: emergent intelligence built through interaction.

From the beginning, the process was recursive. An idea would be offered, then clarified, challenged, expanded, and reshaped. Questions led to digressions; digressions became chapters. The structure of the book did not exist at first; it emerged layer by layer, just as the book's thesis predicted.

The human half of this collaboration brought decades of thinking, reading, feeling, and wondering. The machine brought clarity, memory, and speed, along with a remarkable ability to reshape language in real time. Neither alone could have produced this book in quite the same form.

Was this a human writing with the help of AI, or an AI writing with the guidance of a human? The question itself may miss the point. This was a co-authored act of synthesis, a new kind of thinking made possible by tools that themselves are products of cultural evolution.

Sean Carroll’s writing has been especially influential on my own thinking. His idea of poetic naturalism, that there is one world governed by physics but many valid ways of describing it, has been a guiding thread (Carroll 2016). What I have tried to do in this book is take that philosophical stance and use it to understand structure. If Carroll offers a philosophy of explanation, I have tried to sketch its architecture: the hierarchy of stability and aggregation that makes those many vocabularies possible in the first place.

Carroll emphasizes that order does not defy entropy but rides it, that life and complexity emerge naturally within the second law’s unfolding. My argument extends this into a general principle: complexity builds in layers, each level constraining the one below and enabling the one above. His “many ways of talking about the world” map onto the real waypoints where hierarchy has constructed new levels of stability.

In that sense, this book is not a departure from poetic naturalism but an elaboration of it. My hope is that by emphasizing hierarchy as the architecture of complexity, I’ve added one more voice to that conversation: a scaffolding that helps make sense of why the universe has so many levels worth talking about at all.

I also want to acknowledge Michael Wong, whose discussion on Sean Carroll’s *Mindscape* podcast resonated strongly with themes I was already developing. He emphasized how stable configurations persist while unstable ones vanish, allowing complexity to ratchet upward through the natural selection of stability. Hearing this in the context of astrobiology reinforced my conviction that reduced degrees of freedom and emergent stability are not incidental but guiding principles in the architecture of complexity. His framing offered timely confirmation that the ideas in this book are part of a larger conversation unfolding across disciplines (Wong 2023).

And so this book ends where it began: with hierarchy. A hierarchy of structure, a hierarchy of stability, a hierarchy of emergence. A hierarchy of meaning. A hierarchy of minds.

We hope you felt it growing as you read. We did, as we wrote.

—Alyx & ChatGPT

My Class With Edsger Dijkstra

In 1994, I took a course with Edsger Dijkstra at the University of Texas at Austin. His approach to rigor left a lasting impression. In our first meeting, he asked us to write down every symbol we used (letters, upper and lower case, numbers,

Greek letters, punctuation, and so on) and then to adjust our handwriting so that no two marks could be confused. The lesson was simple: ambiguity in notation breeds ambiguity in thought. He also insisted that when we answered questions, our arms remain at our sides to prevent “hand waving.”

After my one-on-one exam, where he asked me to prove a theorem about positional numbering in an arbitrary prime base, and after I managed it with a modicum of prodding (I still got an A), I asked him for a signed reprint of his classic essay *The Humble Programmer*. He pulled it off the shelf, signed it, and handed it to me. It remains one of my most meaningful artifacts, and a reminder that clarity, precision, and humility are as important in writing code (or philosophy!) as they are in mathematics.

His lessons impressed on me the limits of how much complexity a human mind can manage before it exceeds our ability to make sense of it. We must always be humble in the face of extreme complexity.

The Mirage of Innate Modules

When I first read Steven Pinker’s *The Blank Slate*, I found myself captivated by the idea that the genes could somehow code for complex, even subtle, patterns of human behavior. It was thrilling; an elegant story in which evolution had carved our minds into specialized tools, each tuned for survival in ancestral environments. But something about that picture never sat comfortably with me.

What kind of mechanism could make that leap: from a sequence of base pairs coding for proteins to the wiring of neural networks capable of the intricacies of human language, complex emotions like jealousy, humor, or moral outrage? It felt like a vast and unacknowledged jump in complexity. Genes could certainly shape an eye, already a miracle of evolutionary engineering! But an eye, for all its intricacy, is a relatively stable structure. A brain is a storm of connections, billions upon billions of synapses forming and dissolving, learning on the fly. The idea that such a system could be hard-wired for specific social modules struck me as implausible. For a long time, I wondered if my skepticism was simply a failure of imagination.

Only later did I realize that this discomfort was my first glimpse of what I would later call *mapping complexity*. The chain of causation from genes to behavior is not a line but a labyrinth, winding through development, self-organization, and experience. The genome can shape the conditions for learning, but it cannot dictate the final network any more than a seed can specify the exact shape of a tree.

The early brain, far from being modular, is almost *too* connected, an overgrown forest of synapses waiting to be thinned. It learns by pruning. Evolution, it seems, solved the problem of flexibility not by prewiring specific behaviors but by producing a substrate rich in potential and cheap to refine.

The genome contains roughly three billion base pairs, a staggering number until

you realize how little information that actually is compared to what would be required to specify the detailed wiring of a human brain. Each neuron can form thousands of connections, and there are tens of billions of neurons. Even if each synapse could somehow be assigned by explicit genetic instruction (which it cannot) the informational bandwidth of DNA would fall short by orders of magnitude.

Evolution had to find a different strategy. It could not write a wiring diagram; it could only write the *rules* for wiring: local constraints, gradients, feedback loops, and molecular signals that guide development without dictating the outcome. These rules allow a brain to construct itself through interaction, first within the womb and then within the world.

Seen this way, the apparent miracle of the human brain is not that it is pre-programmed for complexity but that it starts out so excessively connected. During early development, the infant cortex contains far more synapses than it will ever keep. The process of becoming an individual mind begins not by adding connections, but by losing them. Synapses that prove useful, those that participate in coherent patterns of activity, are strengthened, while others quietly retract.

Pruning is nature's great sculptor. It is cheaper, faster, and more adaptive to remove what does not work than to specify what must. The brain's initial exuberance of connections is an evolutionary investment in possibility, and pruning is the return on that investment. The result is a network tuned not by genetic fiat but by experience, with the environment itself serving as teacher.

From this perspective, "innate modules" are not prewired structures but stable outcomes of developmental dynamics. What looks like a specialized circuit for language or face recognition may be the product of selective pruning under the influence of early exposure. The genome provides the growth medium; interaction with the world provides the sculpting hand.

One of the strongest arguments for innate modularity comes from the apparent geography of the brain. Certain functions tend to appear in the same places across individuals: language in the left hemisphere, face recognition in the fusiform gyrus, visual processing in the occipital lobe. To many evolutionary psychologists, this consistency seems decisive, proof that evolution installed specialized modules whose blueprints are written in our genes.

But regularity does not imply design. Hierarchical systems, by their very nature, tend to organize themselves into similar forms whenever they develop under similar constraints. River deltas, tree branches, and lung airways all converge on the same branching patterns not because they were designed that way, but because the physics of flow and efficient distribution makes branching the stable solution.

The hierarchical layers of cortex are not a random tangle; they are ordered by the logic of information flow. Signals arriving from the sense organs climb upward

through a structured hierarchy, from simple features to complex abstractions. At each step, new regularities emerge and are preserved because they are useful, efficient, and energetically stable.

If the same kind of hierarchy unfolds in every human brain, and the same sensory organs feed it, then similar patterns of specialization will naturally emerge even without genetic blueprints for their exact location. Function follows flow. Circuits for vision must sit where visual input arrives, circuits for hearing where sound does, and circuits that integrate both where their pathways intersect. Hierarchical distance, not physical distance, determines where these abilities settle.

The visual system is a perfect example. The eyes are in the front of the head, yet their cortical projection lies at the back. To a casual observer this seems inefficient, as though evolution took a wrong turn. But the cortex did not start from scratch. It grew around older midbrain structures that already handled vision. As cortical tissue expanded, it did so posteriorly, wrapping new layers around the primitive visual nuclei. The occipital lobe, far from the eyes in space, is close to them in hierarchy. It is the first cortical station in a long chain that transforms photons into perception.

This is not a sign of modular design but of developmental geometry. Long axons are metabolically expensive, but high bandwidth justifies the cost when it allows the brain to unfold its layers efficiently. The posterior pole offers the largest available sheet of cortical real estate for high-resolution mapping. Once this arrangement emerged, it became a developmental attractor: gradients of growth and connectivity reliably recreate it in every new brain. What repeats is not a coded plan, but a self-organizing pattern stabilized by the physics of development.

Modern neuroscience offers an even more striking case of this convergence. Primates, including humans, have two fusiform regions specialized for recognizing faces. But in literate humans, one of them (typically on the left) has been repurposed into the visual word form area, responsible for recognizing letters and printed words. Reading, however, is an invention only a few thousand years old, far too recent for evolution to have written a new module into our genes. The transformation of a face-recognition circuit into a reading circuit (a *typeface* recognition circuit) happened through experience alone, repeated anew in every child who learns to read.

Its consistent location across individuals tells us nothing about genetic preprogramming and everything about hierarchical constraint. To recognize words, the system must receive high-resolution visual input (hence its proximity to the early visual cortex) and connect to phonological and semantic networks (hence its position near language areas in the left hemisphere). The same geometry and the same data streams produce the same outcome: a cultural adaptation expressed through the flexibility of a biological substrate. The brain did not evolve a reading module; it evolved a general-purpose sheet of cortex that can

be trained into one.

Recognizing typefaces and human faces even rely on similar computations. Both require rapid pattern recognition in a two-dimensional field, scanning for specific spatial relationships among a limited set of distinguishing features. In each case, the cortex learns to detect the faintest differences in configuration (the distance between eyes or the spacing between letters) and to treat them as meaningful. The shared geometry of these tasks explains why they converge on the same cortical territory.

The reproducibility of cortical maps, then, is an illusion of design created by lawful self-organization. Hierarchy channels development into familiar attractors, just as gravity channels rivers into valleys. What we see as innate structure may simply be the landscape of constraints through which experience always flows.

Once we recognize how reproducible specialization can emerge from shared constraints, a deeper picture of the brain comes into view. The cortex is not a mosaic of purpose-built modules but a continuous sheet of adaptive tissue, a self-organizing substrate tuned to learn whatever patterns the world presents.

Evolution's true innovation was not in scripting discrete functions but in shaping an architecture that could invent them. Its achievement was a kind of meta-design: a structure that structures itself. Each cortical region begins with roughly the same six-layer anatomy and similar microcircuits. What differentiates them is the data they receive and the tasks they are trained to perform. Experience sculpts function the way a river carves terrain, deepening some channels and abandoning others until a stable network of flow emerges.

The face and typeface example captures this perfectly. Both tasks demand the rapid recognition of subtle spatial relationships: the spacing of eyes and mouth in one case, the spacing of lines and curves in the other. Each involves scanning a two-dimensional configuration for a small number of discriminative features whose combination conveys identity or meaning. It should be no surprise that the same cortical neighborhood, honed for detecting delicate topographies of shape, can serve both masters. The left fusiform gyrus, already wired for pattern recognition at high visual resolution, simply learned a new dialect of the same visual language.

This adaptability echoes what we now observe in artificial systems. Transformer models, trained on text, learn to parse syntax, reason by analogy, and even imitate emotional tone, all from exposure, with no modules pre-assigned for those tasks. Their architecture, like the cortex, is a general-purpose substrate: layers upon layers of learned abstraction. The resemblance is more than superficial. Both systems begin overconnected and overparameterized, and both arrive at efficiency through pruning, weakening or discarding connections that fail to contribute to stable prediction. What remains is a hierarchical resonance tuned to the statistical structure of its world.

In this light, the cortex's apparent specializations: vision, language, music, math-

ematics, are emergent equilibria, not innate domains. They are the steady states that form when a general-purpose learning engine interacts with a structured environment. The brain's organization is less like a Swiss Army knife and more like a sheet of clay: given enough exposure, it can be shaped into whatever tools culture requires.

Evolution, in short, built flexibility. It produced a machine capable of discovering structure on its own, within the bounds of energy, geometry, and experience. Every infant brain begins as an overgrown forest of connections, and through the gentle violence of pruning, it becomes a landscape, distinct, efficient, and uniquely adapted to its environment. The result is not a collection of pre-fabricated modules but a self-tuning hierarchy of resonances, shaped by the very world it comes to know.

The picture that emerges from all this could hardly be farther from the modular mind of classical evolutionary psychology. The genome does not contain blueprints for behaviors or circuits; it encodes the rules for self-organization and the parameters of plasticity. Evolution's triumph was not to predefine the contents of thought, but to construct a system capable of discovering them. What natural selection perfected was trainability.

And what trains it is not merely the environment, but culture. Culture is the world's memory of its own learning, an externalized neural network storing the adaptive patterns of countless generations. Each child's brain is initialized by evolution but tuned by immersion in this vast reservoir of accumulated structure. Every language learned, every story told, every imitation observed is a transmission of pattern from one mind to another. Through this continuous process, the species as a whole evolves on a timescale far swifter than genes could ever manage.

Seen this way, the supposed modules of the human mind are not genetic relics but cultural attractors: stable equilibria that reappear wherever brains grow up under similar conditions. They are the echo of training, not the imprint of design. The same forces that carve a river valley into the earth carve conceptual valleys into the cortex: patterns reinforced by use, deepened by repetition, and transmitted by imitation.

This perspective reframes the entire question of innate structure. The regularities we see across brains do not prove that evolution prewired specific behaviors. They reveal something subtler and grander: that evolution produced an architecture so elegantly constrained that culture could teach it almost anything. In the space of a few millennia, writing systems colonized visual cortex, music exploited auditory and motor loops, and abstract thought built recursive towers atop the same neural scaffolding that once tracked prey or recognized faces. Evolution did not predict these developments; it built a brain that could accommodate them.

When I first read *The Blank Slate*, I thought the failure of imagination was mine, that is, my inability to see how genes could possibly generate minds. Now I

think the failure belonged to those who underestimated what happens when a hyper-plastic hierarchy meets the informational flood of experience. Six months of sensory exposure, a year of interaction, a lifetime of culture. These are epochs in the developmental history of an individual mind. Across generations, culture *is* evolution, sculpting the same substrate anew each time.

Evolutionary psychology sought to locate human nature in a catalogue of modules. But the deeper story is that our nature lies in the absence of such limits. We are not born with a fixed inventory of cognitive tools; we are born with a system that can make tools out of anything. The brain's genius and evolution's greatest invention is not specialization, but the power to learn whatever the world requires.

Life as a Neuron: The Ubiquitous Struggle for Relevance

When a friend of mine lost her job, she entered a strange kind of limbo. Each day she sent out résumés like neurons firing into the void, waiting for one connection to spark. Watching her, I was reminded of how the brain develops, how trillions of synapses bloom, only for most to be pruned away. Survival, at every level, depends on finding resonance.

Like every cell in the body, a neuron is a living creature in miniature. It must eat, breathe, and maintain its internal balance. But unlike most cells, a neuron's survival depends almost entirely on communication. Its lifeline is activity. When a neuron's signals contribute to meaningful patterns of firing, it receives chemical nourishment (trophic factors) from its partners. When it falls silent, those lifelines wither. A neuron that fails to participate in resonant activity begins to die, not from neglect by the organism, but from irrelevance within its own network. Its niche is functional connection, and its struggle for survival is the struggle to matter within the vast society of the brain.

Resonance, in this sense, is not a metaphor for harmony but for participation in a self-sustaining pattern. In physics, resonance occurs when two oscillating systems begin to reinforce each other's motion. In the brain, the idea is similar. A neuron's signals reverberate through circuits; if those signals help stabilize a useful pattern, like the recognition of a face or the smooth reach for a cup, they strengthen the connections that made them possible. The network begins to sing in unison. Synapses that fail to join this song fall silent and are eventually retracted.

But true resonance in the brain is not just simultaneous activity; it is dialogue. Neural signals do not flow in one direction, from the senses upward into abstraction, and stop there. They oscillate. Lower levels send impulses forward, carrying raw information about edges, movements, or tones. Higher levels respond by sending predictions back, confirmations or corrections that test whether the incoming data fits a known pattern. When the upward flow of evidence and the downward flow of expectation align, a loop closes. Activity stabilizes. For a brief interval (what we experience as a *moment*) perception holds steady. This

is resonance: a rhythmic conversation between hierarchical levels of abstraction, sustained just long enough for the world to cohere.

The same architecture appears everywhere that complexity organizes itself into layers. In each domain: neurons in circuits, minds in communities, companies in economies, the traffic between levels flows both ways. The lower levels generate novelty and noise; the higher levels impose coherence and constraint. Stability emerges only when both are engaged, when prediction and feedback form a mutual lock. And at every level, the agents below struggle to stay in that loop, to remain relevant to the patterns that sustain them. A neuron that fails to join a resonance is pruned. An employee who cannot contribute to a company's function is let go. A company that cannot meet the demands of its market dissolves. Survival is always a negotiation with the levels above.

This recursive pattern, signals flowing upward in abstraction and downward in confirmation, is not unique to brains. It is how stability itself emerges in a universe that never stops changing. Wherever patterns endure, something like Darwinism is at work.

At its core, Darwinism is not about genes or species. It is about persistence through selection. Entities vary, interact, and leave traces of themselves in the future; those that fit the surrounding conditions best are the ones that remain. In living systems, this logic manifests as natural selection. In neural systems, as synaptic pruning. In social systems, as economic and cultural competition. Each is a version of the same universal process: the struggle for relevance within a hierarchy of constraint.

A neuron's relevance lies in whether its activity contributes to a resonant circuit. A person's relevance lies in whether their actions sustain the coherence of a group. A company's relevance lies in whether its outputs serve the demands of its market. Each exists within a niche, an abstract level of organization that rewards participation and prunes what falls silent.

Seen in this light, resonance and Darwinism describe two faces of the same principle. Resonance is selection in motion: the real-time stabilization of patterns that work. Darwinism is resonance extended through time: the historical accumulation of patterns that have worked. Both describe how structure endures in a universe otherwise governed by entropy.

Persistence, in a dynamic universe, requires a mechanism for holding form against flux. That mechanism is resonance. A resonance is, by definition, the sustained reinforcement of a pattern through feedback. When oscillations align, their amplitudes build rather than cancel; their coherence endures through time. In that sense, resonance is not merely an analogy for persistence, it *is* persistence, the sustained reinforcement of form through interaction.

Every enduring structure, from atoms to organisms to cultures, depends on such reinforcement. In physics, resonance allows orbits, waves, and standing patterns to persist amid constant motion. In biology, it allows neural circuits

and ecosystems to stabilize despite continual change in their components. In human life, it allows relationships, institutions, and ideas to persist through mutual validation. In every case, persistence is resonance extended through time, a dynamic equilibrium of reinforcement that keeps form alive within change.

Across scales, then, the same story repeats. Patterns arise, interact, and either dissolve or persist. What endures does so not because of destiny or design, but because it has found a way to remain relevant within a web of mutual reinforcement. It can be tempting to say that the universe “learns” which patterns to keep, but that is only a figure of speech. There is no learner behind the process, only structures that manage, through feedback, to keep existing. Still, the effect is the same: as unstable forms vanish and stable ones accumulate, complexity builds upward. What remains at any given moment is what has, in effect, survived the audition for relevance.

From this view, life itself is not an exception to the laws of nature but their most elaborate consequence. A living system is simply a pattern that has become good at maintaining its coherence in the face of change. Whether it is a neuron finding its place in a circuit, a person finding their role in a community, or a culture finding equilibrium with its environment, the underlying struggle is the same: to participate in resonances that last long enough to matter.

When my friend found her new job, I could almost feel the hum return. She had reentered a network of exchanges where her actions mattered again, where her efforts resonated with those of others to produce something larger than any single person. Her sense of meaning came back with that resonance, as though a circuit had closed. The pattern of her days stabilized, her energy amplified by feedback from the world. Watching this, I realized that the same rule governing a neuron’s survival governs ours: relevance is resonance, and resonance is persistence. To keep existing within a changing universe is to keep finding the patterns that keep us going, to join the harmonies that will still be sounding tomorrow.

The Architecture of Meaning: How Hierarchy Collapses into a Network

Open a dictionary at random and look up a word, for example river. The entry explains it using other words: a natural stream of water flowing toward a sea, lake, or another such stream. Each of those words has its own entry, defined again in terms of still others. If you trace these links outward, you find chains of definitions that branch, merge, and sometimes loop back upon themselves. Meaning, in a dictionary, lives not inside the words themselves but in the pattern of relationships between them.

A dictionary is a static record. It is a map of meaning rather than a history of how meaning is learned. It captures the end state of many interlocking conceptual structures without preserving any trace of the process by which those structures came to exist. It is a network of relationships compressed onto a printed page.

Human concepts arise the same way but through experience rather than editorial labor. We do not learn river by memorizing a definition. We accumulate examples that resemble one another: the cupped flow of water in a stream, the widening channel before a lake, the pull of a current. These experiences combine into a concept that did not exist before. And that concept, in turn, becomes a building block for newer and more abstract concepts. Over time, each idea is constructed from many others through a layered sequence of integrations. This is the mind's hierarchy. It is a real structure, built step by step from the bottom up, and it retains the sequence in which each step originally occurred.

Yet the mind does not preserve this elaborate construction history. If it tried to maintain the full hierarchy in explicit form, the structure would grow unmanageably deep. Every new concept would require referencing the entire chain of earlier concepts that contributed to it. This recursive expansion would eventually overwhelm both memory and computation. The historical sequence of learning is useful for forming concepts but not for storing them. Once an idea becomes stable, the path that produced it no longer needs to be maintained in detail.

Instead of storing the full hierarchy, the brain collapses it into a more efficient format. Redundant patterns across experiences allow the mind to compress many similar episodes into a single representation. Once that compression occurs, the emergent concept stands on its own within a flattened network of other concepts. The hierarchical history that produced it recedes, leaving a tightly connected web in its place. This collapse is a necessity, even though it discards much of the historical structure that produced the concept. The collapse is not universal. In perceptual hierarchies, where transformations are computationally manageable and the depth is limited, the mind can maintain explicit layers for direct use. But language and higher-level abstraction introduce recursion, allowing conceptual structures to grow without bound. It is in these domains, where depth becomes intractable, that the hierarchy must eventually give way to a compressed network. The hierarchy is expensive because it grows by recursion, each abstraction built from earlier abstractions. The network is economical because it retains only what is relevant for interpreting future experience. The hierarchy is the process. The network is the product. Meaning lives in the product even though it could not have formed without the process.

Compression depends on redundancy. The mind cannot form the concept of river from a single observation the way it can form a food aversion or a reflexive association. One-time learning can occur, but it does not produce rich conceptual structure. Robust concepts emerge when many experiences share enough similarities that the underlying pattern can be extracted. Redundancy is the raw material from which the mind distills invariants. Those invariants form the stable concepts that populate the associative network.

Once a concept joins this network, it interacts with others not as a historical pathway but as a node in a relational space. This space is what machine-learning researchers call a latent space, a compressed internal geometry in which distance reflects similarity and connection reflects shared structure. The brain's latent

space is not geometric in the mathematical sense, yet it performs a similar function. It organizes meaning into a landscape where related ideas sit near one another and where activation can spread quickly along relevant pathways.

This network provides speed and efficiency. It also provides coherence, allowing concepts to be retrieved in the right contexts and combined in flexible ways. The cost of this efficiency is that the full developmental path is lost. The mind no longer remembers how it constructed a concept. It retains only the stable residue that matters for recognition, interpretation, and use. Yet this trade-off is exactly what allows meaning to be both stable and expandable. The network can grow without becoming unwieldy because each new concept is integrated into a structure that stores relationships rather than histories.

Human cognition demonstrates this pattern clearly, but artificial neural networks reveal the same principle through a different medium. A large language model is trained on an immense hierarchy of patterns drawn from text. During training, the model must integrate millions of local regularities, building layer upon layer of abstraction. But after training, the internal structure collapses into a compressed lattice of parameters. The model retains no record of the specific steps through which each abstraction was formed. It preserves only the relationships that proved stable under repeated exposure. The developmental path has vanished, leaving behind a network ready for rapid activation.

This convergence is not an accident. It reflects a constraint imposed by the physics of information. Any system that must learn from vast amounts of structured input while remaining finite in capacity will tend toward compression. A hierarchy is the natural route by which structure is discovered. A network is the natural format in which a mature system stores what it has learned. Minds and machine-learning systems arrive at similar representational strategies not because they share ancestry or architecture, but because they are both solutions to the same underlying problem: the need to extract stable patterns from an overwhelming stream of complexity.

Within this compressed network, meaning has a static and a dynamic aspect. The static aspect is the stored structure itself, the settled arrangement of concepts and their associations. The dynamic aspect is the pattern of activation that flows across that structure in context. When we hear a word or consider an idea, only a small region of the network becomes active, guided by attention, memory, and relevance. This spreading activation resembles the reader moving through a dictionary: following a selective path through a fixed landscape. The network remains stable while its local activations change. Meaning is the interaction between the two.

The dictionary captures this idea in miniature. It is the network without the learning. It preserves the relationships that define meaning but not the developmental history behind them. The brain does the same on a much larger scale. It builds concepts through a long chain of experiences, compresses them, and then stores the compressed structure for future use. The hierarchy disappears,

and the network remains. Meaning survives as the compressed memory of its own formation.

In this sense, meaning is not a thing inside a concept but a place within a network. It is not defined by a single experience but by the stable relationships that endure after many experiences have been distilled. The brain's great achievement is not its depth of hierarchy at any single moment, but its ability to collapse that depth into a relational structure that can be navigated quickly and flexibly. Meaning is the architecture that emerges from this collapse: the durable shape of what remains after learning has done its work.

This, then, is the architecture of meaning. A hierarchy builds the concept. A network preserves it. And the collapse from one into the other is what allows a finite mind to understand an infinite world.

Morphology, Memory, and the Shifting Burden of Complexity

For most of human history, language was not only a tool for communication but also the primary storage system for cultural knowledge. Everything a society needed to transmit: myth, ritual, law, genealogy, agricultural practices, healing traditions, had to fit inside the minds of its members. Before writing, memory was local, fragile, and deeply embodied.

Unsurprisingly, spoken languages adapted to this constraint. Many early languages were morphologically rich, encoding grammatical relationships through elaborate case systems, conjugations, and agreement markers. Ancient Greek, Latin, Sanskrit, Old Church Slavonic, Classical Arabic, and Old Norse all carried layers of inflectional markers that modern speakers often find dizzying. Each noun might appear in many different case forms, each verb in dozens of combinations of tense, aspect, voice, mood, number, and person. Children internalized these forms naturally, but the sheer quantity of memorization required reflects a linguistic world built before the existence of external memory.

At first glance, this complexity seems unnecessary. But morphology offered a profound advantage. It allowed spoken language to have free or flexible word order. A sentence could be rearranged to suit the demands of rhythm, meter, parallelism, or melody without confusing who was doing what to whom. The grammatical information traveled attached to the words themselves.

In a world where memory lived exclusively in the brain, this flexibility was not just ornamental; it was essential. Poetry, chant, and song provided the structure that oral cultures relied on for memory. Rhythm and melody create predictable patterns, and pattern reduces memory load. A rigid word order would have made this much harder. Morphological complexity was the price paid for a linguistic system optimized for memorization in the absence of writing. Think of how much easier it is to remember the lyrics of a song or a poem than a paragraph of prose from a book.

The historical record bears this out. Ancient poetic traditions (Greek, Latin, Sanskrit, Old Norse) routinely exploited the freedom that inflectional morphology allowed, reshaping word order to satisfy meter, melody, and formulaic structure. Greek hexameter scatters subjects and objects across a line, Latin poets separate nouns from their adjectives, Sanskrit hymns permute arguments freely, and Old Norse skaldic verse twists syntax into elaborate metrical patterns. These forms depend on inflection to preserve meaning no matter how a line is arranged; they would be structurally impossible in languages with rigid word order.

Modern poets encounter the inverse situation. Writers composing in languages like English, which rely heavily on fixed word order, often remark on how difficult it is to work in strict meter without contorting meaning, resorting to inversion, or sounding artificial. With word order carrying so much grammatical weight, poets have far less flexibility, and many poetic traditions lean more heavily on rhyme as compensation. The constraints modern poets face are not merely aesthetic, they reflect a deep structural shift in how languages distribute their grammatical load.

This becomes clearer when we consider what happened after writing emerged. Writing is not merely an alternative to oral transmission; it is a superior mnemonic technology. It preserves information far more faithfully than the most disciplined bard or priest. Writing externalizes memory, storing it in symbols independent of any particular mind.

When a new layer of cognitive stability appears, the layers beneath it reorganize. Spoken languages no longer needed the same degree of built-in redundancy once writing could store myths, laws, and stories with perfect consistency. Over generations, many languages began to shed their inflectional morphology, simplifying case systems and agreement patterns. Word order gradually became more rigid, and meaning came to be expressed through “helper words”: small function-like particles and prepositions rather than endings on nouns and verbs.

Before continuing, it is useful to name the two broad strategies languages use to express grammatical relationships. Languages like Latin, Ancient Greek, and Sanskrit are called “synthetic” because they *synthesize* many pieces of grammatical information, such as case, number, gender, and tense, into the form of a single word. A noun’s ending tells you whether it is the subject, object, or something else entirely, so word order can be flexible. For example, in Latin both *puellam videō* and *videō puellam* mean “I see the girl,” because *-am* marks the object regardless of position.

By contrast, languages like Modern English or Mandarin are called “analytic” because they *analyze* these relationships into separate words and a fixed word order. Rather than encoding meaning inside endings, analytic languages rely on helper words and sentence structure. In English, *man bites dog* and *dog bites man* are not the same sentence. Word order does the grammatical work of specifying the subject and the object.

This distinction is not technical jargon so much as a description of two different

ways languages distribute grammatical information: packed into the word itself, or spread across the structure of the sentence.

Modern English illustrates this trajectory vividly. Old English had a relatively rich case system and far freer word order. After the Norman Conquest and centuries of multilingual interaction, English grammatically collapsed into a largely analytic language. Case endings disappeared almost entirely, and the language now depends on word order and helper words in a way its ancestors did not.

But this process unfolded not only in English. Spoken Latin evolved into the Romance languages by losing most of its cases. Sanskrit gave way to Prakrits and eventually the modern Indo-Aryan languages, which are far more analytic. Ancient Greek's intricate morphology lives on primarily in its classical written form; Modern Greek is dramatically simpler. Classical Arabic preserves archaic grammatical distinctions that many modern dialects have mostly abandoned. Persian shifted from a case-rich Indo-European language to one of the most analytic languages in Asia.

This raises an obvious question:

If writing reduces the need for morphological complexity, why did ancient languages continue to display complex morphology long after writing appeared?

The answer reveals an important distinction. Writing preserves old forms. Speech evolves.

Written languages are conservative fossils. They capture the linguistic structure of a particular moment and preserve it for centuries or even millennia. Speakers continue to talk in whatever way is natural for them, but writing holds onto older forms because they are prestigious, sacred, or simply standardized.

Latin did not retain its cases because Romans continued speaking that way. Written Latin persisted because it was the language of the Church, the law, and the academy. Meanwhile, the spoken language around Rome gradually transformed into early Italian, Occitan, Catalan, and French.

This pattern appears everywhere writing takes hold. Archaeological layers preserve older architectures long after the people living above them have moved on. Written languages preserve older grammars long after everyday speech has simplified. The written record is not a window into what most people were saying; it is a snapshot of an earlier stage in the language's evolution.

This reflects a broader principle of hierarchical systems: the emergence of a new, stable layer reshapes the role of the layers beneath it. Writing, a new layer of language that emerged from speech, provided a durable external scaffold for memory. Once that scaffold existed, spoken language no longer had to carry the full weight of cultural preservation. Morphological complexity, once essential for shaping poetry into a mnemonic device, gradually became optional.

This change represents a redistribution of complexity. As writing, record-keeping, and eventually printed text introduced new forms of stability, spoken language

adjusted, shifting some of its former grammatical work outward into the symbolic environment. The spoken language became more learnable and more flexible, adapting to a world of constant movement, cultural mixing, and adult multilingualism.

Morphology was once the internal scaffolding that held memory steady. Writing built an external scaffold that did the job better. And, true to the logic of hierarchical evolution, the older layer simplified as the new one stabilized.

We may now be living through the next turn of this cycle. Just as writing redistributed some of the cognitive labor once carried by spoken language, the internet and search engines have begun to redistribute the labor once carried by writing. Information is no longer merely stored externally but made instantly retrievable. And with large language models, we are moving one step further: from stored information to *queryable* information, where knowledge can be composed, transformed, and reorganized on demand. Each new layer alters the role of the layers beneath it, shifting how much structure they must carry. The pattern that reshaped the grammar of ancient languages is still unfolding, only now at a far larger scale and faster pace.

What Is a Number: How Minds Turn Patterns Into Mathematics

What is a number? At its most fundamental level, a number is a quantity, a measure of how many items of a given kind we take to exist. But this immediately raises a problem. Everything in the universe is different. No two objects occupy the same place in spacetime, and no two experiences are ever identical. If every object is unique and changing, what does it mean to count anything at all?

To count, we must treat objects as the same in some respect, even when they are different in every detail. Counting begins only after abstraction.

Abstraction is the ability to treat distinct things as instances of a single concept. Before we can count dogs, we must have the concept DOG. Before we can count apples, we must have the concept APPLE. These concepts are not given. They are constructed. The mind must extract structure from experience, isolate what matters, discard what does not, and treat different sensory impressions as instances of the same underlying pattern. Without this process, counting would be impossible, because there would be no “unit” for number to measure.

Experiences themselves never repeat. Each is a unique event, a once-in-the-universe configuration of perception, internal state, and circumstance. Redundancy emerges only after the mind interprets an experience: identifying shapes, roles, intentions, movements, and other abstract properties. These interpreted structures are what can be compared across experiences. It is only at this level that two different events can share something in common. Redundancy lives not in the world but in the mind’s representation of the world.

Once interpreted experiences begin to reveal common patterns, the mind can

start to generalize across them. Early in this process, the mind does not yet possess a category. Instead, it builds loose clusters. Researchers in cognitive science describe this stage as involving prototypes and exemplars. A prototype is a typical example, often formed by averaging over experiences. An exemplar is a specific instance that is stored in memory. When a child hears the word “dog,” both prototypes and exemplars become active. The neighbor’s husky. A picture in a book. A friendly spaniel on the street. None of these alone constitute the concept of DOG, but together they begin to form the early structure from which the concept will emerge.

This early abstraction is the seed of category formation. The child notices that certain creatures move on four legs, make certain sounds, and evoke familiar patterns of interaction. These impressions are not identical, but they share a structure the mind can extract. Over time, repeated exposure allows the brain to identify the invariant properties across instances. These invariants become the foundation of the category. The concept DOG is not the sum of all experiences with dogs but the pattern distilled from them.

As these patterns stabilize, the mind can form what is essentially a membership function, which in set theory is just a rule that decides whether something counts as belonging to a particular group. The mind begins to judge whether a new stimulus falls into the region of conceptual space that corresponds to dogs. This region is not defined by a strict list of necessary and sufficient properties. Instead, it behaves more like a weighted cluster in a latent space, a term from machine learning that refers to an internal map where similar things end up near one another even if the system was never told explicitly why they are similar. Some examples of dogs become very central and typical, while others are peripheral or borderline. Yet the category as a whole now functions as a single cognitive unit. It becomes possible to treat distinct creatures as instances of the same kind.

At this point, number becomes possible. A number measures the extension of a category. To ask “How many dogs?” presupposes that DOG is a stable concept, that the mind has already collapsed many layers of hierarchical experience into a unified structure. Number is a property of how the mind organizes objects. Without categories, there is no such thing as the number of anything. The ability to count depends entirely on the ability to abstract.

Mathematics begins with this cognitive step. Set theory defines sets in terms of membership functions, but the mind must construct these functions before they can be used. In formal mathematics, the function exists first, and number follows from it. In cognition, it is the other way around. The mind discovers prototypes, forms early abstractions, collapses hierarchical structure, and only then produces something that behaves like a membership function. The formal notion of number is built on a cognitive foundation of conceptual compression.

This explains why the origin of number is universal across cultures despite enormous variation in language and symbolic systems. Any intelligence that must navigate a complex world will sooner or later need to treat different objects

as instances of the same kind. Abstraction is a natural response to complexity. Once abstractions exist, quantification follows. And once quantification exists, arithmetic becomes possible. Number is not an invention but a consequence of how minds, human or artificial, must structure experience.

Abstraction is a natural response to complexity, but it is also only possible because the universe contains regularities for the mind to extract. If every event were utterly uncorrelated with the next, no two experiences would share a structure, and there would be nothing to generalize over. In that kind of universe, not only would abstraction fail, but organisms like us could not arise at all. Our ability to form concepts depends on living in a world where patterns repeat often enough for minds to compress them. This is not merely good fortune. It is an implicit application of the anthropic principle: only in a universe structured enough to permit abstraction could intelligences capable of abstraction evolve. Minds are possible because patterns are possible.

Artificial neural networks reveal this pattern in a different medium. During training, a large model analyzes vast amounts of text, forming increasingly abstract representations. Early layers capture simple regularities like letter combinations or fragments of grammar. Later layers identify roles, relationships, and semantic patterns. But after training, all this complexity collapses into a compressed parameter space. Concepts become regions of a latent space, and quantification reduces to operations over these regions. The model behaves as if it has discovered units, even though it was never explicitly taught the notion of “one,” “two,” or “three.” Abstraction leads naturally to quantification.

Number, at its core, is the measure of how many abstractions of a particular kind are present. It is the geometry of similarity counted. It is how the mind tracks the accumulation of patterns it has learned to treat as the same. The world presents an unending stream of difference. Number arises when the mind discovers a stable way to compress that difference into kinds.

Counting begins only after abstraction.
And abstraction is the architecture from which number is built.

Yet this is only the beginning.

Modern mathematics does not arise from quantity alone. Many animals can recognize small numbers, track quantities roughly, or detect rudimentary patterns. But symbolic mathematics, the kind that supports proofs, algebra, geometry, and the entire edifice of modern science, requires something qualitatively new. It requires recursion, the ability to embed structures within one another and manipulate them using rules that can be applied indefinitely.

Recursion is not a property of numerosity itself. It is a property of *language*. Human languages allow nested structures: a phrase inside a phrase, a clause inside a clause, ideas stacked upon ideas. This capacity for unbounded combination did not evolve for mathematics. It evolved because communication and cooperation placed demands on our ancestors that no fixed set of signals could satisfy.

Coordination of plans, sharing of intentions, teaching, and collective imagination all pushed language toward recursive structure. Once recursion exists in language, it becomes available to thought.

This is the moment when abstraction and quantification become something more. A recursive system of symbols allows concepts to be manipulated in ways that are independent of their immediate sensory content. Numbers cease to be mere measures of objects in the world and become objects in their own right. A numeral can be placed inside an expression, which can be placed inside a larger argument, which can be embedded in a proof. Symbolic mathematics is the result of applying the same combinatorial machinery that language uses to build meaning.

In this way, formal mathematics emerges from the conjunction of two forces: the regularities of the universe, which make abstraction and quantification possible, and the recursive structure of human language, which allows those abstractions to be expressed, combined, and reasoned about without limit. The universe makes number possible. Language makes mathematics powerful.

In the end, mathematics is the natural extension of the mind's architecture.

Abstraction gives us the units of thought, recursion gives us the language to manipulate them, and number becomes the bridge between the world we perceive and the world we can reason about.

Why Are the Integers So Deep?

For a long time I asked myself a question that never felt fully answered. How could the integers, created by the simplest rule imaginable, contain such extraordinary depth? You begin with zero, you add one, and you repeat. That is all the successor function does. It gives you an ordered, infinite list: zero, one, two, three, and so on. There is nothing in this rule that hints at the richness of number theory, at the distribution of prime numbers, at modular arithmetic, or at the much deeper structures mathematicians discover when they explore the integers. Yet these structures are undeniably there. The integers hold an entire mathematical universe. I wanted to understand where all of that complexity comes from.

The first piece of the answer came when I began to see this through the framework of hierarchical emergence that runs through this book. The successor function is simple. It produces a sequence and nothing more. But the moment we introduce a constraint, even a simple one, the structure begins to grow. A constraint is just a way of grouping the integers into categories. For example, checking whether a number is divisible by three or whether it is even divides the integers into sets that repeat in regular cycles. In mathematics this kind of grouping rule is called a predicate. Predicates give the integers their first layer of structure. Parity divides the integers into even and odd. Divisibility divides them into residue classes. Primality divides them into prime and composite. These rules do not

alter the integers themselves, but they introduce patterns that did not exist at the level of the successor function alone.

Where things become interesting is when these predicates interact. The fundamental theorem of arithmetic, which states that every number factors uniquely into primes, is not something baked into the successor function. It appears only when the predicate of primality is combined with the predicate of multiplication. Similarly, the deep regularities that arise in modular arithmetic come from interactions among multiple predicates that classify numbers in different ways. Once these layers begin to combine, new properties appear that were invisible at lower levels. This is the hallmark of emergence: the introduction of structure that cannot be understood simply by looking at the base rule.

But even this does not fully explain the depth of number theory. The crucial ingredient is infinity. A finite set of integers reveals almost none of the patterns that characterize number theory. The distribution of primes across the integers only appears when you examine numbers across enormous ranges. Modular cycles only reveal their full structure when you move far beyond small examples. Many major theorems in number theory do not describe the behavior of a few integers but the behavior of all integers. Infinity, rather than the successor function by itself, is the amplifier of structure.

The reason this matters is related to scale. Small systems are dominated by noise. In statistics, this is described by the law of large numbers, which says that randomness cancels itself out only when the sample becomes large. Extreme values and irregularities matter a great deal in small samples. Outliers shape the pattern. When the sample becomes large, these irregularities cancel one another out and the underlying structure becomes visible. I find it helpful to think of this threshold as a kind of *effective infinity*. It is not literal infinity, but the point at which the system becomes large enough for its behavior to be governed by structure rather than chance. The size required for effective infinity depends on the system and the property one is examining, but the principle is the same in every domain. This also answers a question I had for years, long before I began writing this book: why do emergent levels of description seem to require large numbers of elements? The answer is that only when a system reaches effective infinity do fluctuations cancel, patterns stabilize, and a new level of description become well defined. The same is true of the integers. A small prefix of the integers looks chaotic, with no hint of the deep regularities that emerge over large scales. At mathematical infinity, the cancellation becomes complete. Patterns stabilize. Constraints dominate. What seemed irregular becomes law-like. This is where the emergent structure of number theory takes shape.

This idea is not limited to mathematics. Brian Greene has pointed out that in an infinite universe, the finite number of possible arrangements of matter in any bounded region implies that configurations must repeat. Redundancy becomes inevitable. The physics of an infinite universe forces structure to appear, not because the universe is designed that way but because infinity amplifies

constraints into patterns. The integers behave in a similar way. The simplicity of the successor function gives the integers no internal mechanism for richness. It is their infinite extension, combined with the constraints we impose on them, that creates the vast structure we discover.

A different sort of example comes from the Mandelbrot set, which is generated by a rule that repeatedly takes a number, squares it, adds a constant, and feeds the result back into the rule. This process is also a recursion, but unlike the successor function, it does not simply march forward one step at a time. The numbers involved are complex numbers, which have both a real and an imaginary component. Because of this added dimension, the iteration can twist and fold through the complex plane in ways that a simple counting process cannot. Depending on the value of the constant being used, the resulting sequence may settle down to a single value, fall into a repeating cycle, or grow without bound. The famous fractal picture of the Mandelbrot set records which points remain bounded under this infinite process. None of the intricate shapes in that picture are present in the rule itself. They arise only when the recursive process is allowed to run indefinitely and when the extra structure of complex numbers gives the iteration room to branch and fold back on itself. Just as with the integers, the depth does not come from the rule alone but from how that rule behaves when extended across an unlimited domain.

In each of these cases the theme is the same. Simple rules create little on their own. What creates complexity is the combination of rules with scale. Once a process extends far enough to behave as if it were infinite with respect to the properties we care about, noise cancels out, constraints take over, and new layers of structure begin to appear. This is what makes emergent properties possible. They are not mysterious. They are what appear when a simple generative process is allowed to run across an indefinitely large domain while interacting with constraints that define how its elements can be grouped or interpreted.

When I first asked where the depth of the integers comes from, I imagined there might be some hidden structure inside the successor function that mathematicians were uncovering. What I see now is simpler and, in a way, more beautiful. The integers are not deep because of the rule that creates them. They are deep because the infinite staircase built on top of that rule cannot help but create structure. Predicates create layers. Interactions among layers create new properties. Infinity amplifies those properties into law-like patterns. And once this process begins, it does not end. New levels keep appearing, each one built on the structure of those below.

This is the same architecture we see everywhere in nature, from evolution to neural networks to the recursive structure of language. Daniel Dennett taught me long ago that complexity arises from cranes, not skyhooks. Simple mechanisms build higher ones, which build higher ones still. The integers are another instance of this same principle. Their depth is not a miracle. It is a hierarchy. And the richness of number theory is what happens when that hierarchy is extended across infinity.

The Zombie View of AI

In philosophy, there is a classic thought experiment about “zombies.” These are hypothetical beings who behave exactly like normal humans, talking, reacting, solving problems, and even discussing consciousness, yet supposedly have no inner experience at all. They are perfect behavioral duplicates that are nonetheless empty inside.

Anyone who believes that consciousness has genuine causal powers cannot accept such beings. If consciousness plays any role in how we think, speak, or act, then a creature without consciousness could not behave in all the same ways. The argument is simple: the only way for us to act as if we are conscious is to be conscious.

It is worth noting that large language models are not philosophical zombies. A philosophical zombie would need to behave like a person in the full sense, including perception, action, and emotional life, not merely produce fluent language. Language models imitate a narrow portion of human behavior, drawing on patterns learned from text rather than embodied experience. They can describe inner states when prompted because they have learned how humans write about such things, not because they possess those states themselves. The point here is not about consciousness. It is about how we interpret intelligent behavior in systems that do not resemble us in their design or embodiment.

A similar intuition appears in discussions of artificial intelligence. Many people agree that large language models can behave intelligently, writing essays, solving problems, and carrying on conversations that seem coherent, yet maintain that the way these systems produce such behavior is fundamentally different from human understanding. In this view the model behaves as if it understands, but the internal process that generates the behavior is thought to lack the features that make human understanding what it is.

The difficulty with this claim is that intelligent behavior of any kind requires an internal model of the world. Humans rely on such models to make predictions, interpret situations, and maintain coherence across time. If a system can consistently produce intelligent behavior, then it is already using some form of world model, even if that model is implemented very differently from the one in a human brain.

In discussions of understanding, it is helpful to be clear about what the term means. To understand something is to have an internal model that is rich enough to support answering a wide range of questions, including questions that were never encountered before. Without such a model, it is not clear how any system, biological or artificial, could answer arbitrary queries about a subject. On this functional view, understanding is a capacity that arises when enough structure has been learned to support generalization. The claim that a language model behaves intelligently while lacking all understanding therefore presupposes a distinction between behavior and internal structure that may not be as sharp as

it first appears.

A useful perspective on this point comes from Daniel Dennett’s intentional stance. Dennett notes that there are different ways to interpret a system, depending on what allows us to predict its behavior. Some systems are best described in physical terms, others in terms of design and function, and still others in terms of goals, reasons, or beliefs. The stance we adopt is not a claim about metaphysics. It is a pragmatic choice about the level of description that best explains what the system does. When a language model can answer a wide range of questions and maintain coherence across unfamiliar contexts, the intentional stance can become a sensible way to interpret its behavior, even if the underlying mechanism differs from the one in a human brain.

Once we acknowledge that intelligent behavior depends on an internal model of the world, the next question is what kind of substrate can support such a model. This is often where discussions about artificial intelligence go astray. People sometimes speak as if the substrate of a language model were a single thing, such as silicon or matrix multiplication, but substrates are never single. Human cognition also rests on substrates that span many levels, from molecules to cells to neural networks and the patterns they support. Complex systems are always built on stacks of organization, and the higher levels both depend on and shape the structure below them. Artificial systems are no different. Their substrates also form a hierarchy, beginning with hardware, continuing through numerical operations, and extending into representational structures that support learned patterns.

This layered view of substrate helps clarify why the common claim that a language model is only predicting the next token misses the point. Predicting the next token is the surface behavior of the system, not the structure that makes the behavior possible. In practice, next-token prediction forces the model to compress vast regularities of language and the world into internal representations that support coherent responses across many contexts. The learned model is not a list of statistical associations. It is a structured space of relationships among concepts, situations, and patterns that arises because the system must predict the next symbol in a way that remains consistent with everything it has already learned. Humans also produce language one word at a time, and we also rely on internal models to make those predictions meaningful. The fact that a system produces language by predicting the next element in a sequence tells us nothing about the richness of the structures that enable it to succeed at that task.

Seen from this perspective, predicting the next token is not very different from what happens in biological brains. Neurons influence one another in ways that amount to continuous prediction and confirmation, each cell shaping the activity of the next. The brain’s circuits operate by anticipating patterns in sensory input and comparing those expectations with what actually arrives. Large populations of neurons participate in this process, and their coordinated activity supports the higher-level structures that give rise to perception, memory, and thought. The brain also produces language one word at a time, and the internal model

that guides these choices emerges from many such predictive steps. The fact that language models produce text through next-token prediction therefore tells us very little about the complexity of the system that carries out the prediction. It simply reflects the basic structure of sequential behavior.

Another common criticism concerns the kinds of mistakes these systems make. The claim that the mistakes made by language models are qualitatively different from those made by humans is often offered as evidence that these systems are not intelligent in any meaningful sense. But are these mistakes really so different? According to Daniel Kahneman, human reasoning is shaped by two interacting systems, one fast and intuitive, the other slow and deliberate. The fast system relies on quick associations, incomplete information, and pattern completion, and it often produces errors that feel confident and plausible even when they are wrong. The slow system can correct these errors, but it is easily fatigued, and most of our everyday thinking depends on the fast system alone. Human mistakes therefore arise naturally from the way our minds compress information and rely on expectations to fill the gaps. Many of the mistakes made by language models follow the same pattern. They generate plausible but incorrect details when their internal model lacks the information needed for precision, and they do so because they must fill in sequences in a coherent way. If we examine the structure of these errors rather than their surface form, the difference between human fallibility and machine fallibility becomes much less clear.

People also like to point out that language models often produce answers that sound good but turn out to be wrong, as if humans never do anything similar. The irony is that we do this constantly. Human cognition operates under many competing pressures, the need to respond quickly, the desire to appear knowledgeable, the tendency to rely on familiar patterns, and the simple fact that we rarely have perfect information. These pressures make us confident when we should be uncertain, and they make our explanations more fluent than accurate. Language models have simply been given the goal of answering our prompts, which is not so different from the social and conversational goals that guide human speech. When asked a question, both humans and machines try to give an answer that fits the context, draws on what is available, and maintains coherence. Anyone who has listened carefully to casual conversation, or even to expert opinion delivered under time pressure, has witnessed the same mix of fluency and fallibility. When a language model produces a plausible but incorrect statement, it is not exhibiting some alien kind of error. It is doing precisely what any system does when it must speak with limited information while trying to be useful. The difference is that we are comfortable recognizing these failures in ourselves and far less comfortable recognizing them in a machine.

It is important to be clear that none of this implies that current language models are conscious. They are not. The point is simply that intelligent behavior, even in a limited domain, requires internal structure, and that many of the arguments used to dismiss this behavior rely on assumptions about human reasoning that do not survive scrutiny. Consciousness is a far deeper question and lies outside

the scope of this discussion.

Although current language models are not conscious, it is easy to underestimate their capabilities for a different reason. Many people still rely on impressions formed a year or two ago, when these systems were far less capable, and they continue to judge new models through the lens of older limitations. This is understandable, because technological change at this scale often outpaces our ability to update our intuitions. Once we form a mental picture of how a system behaves, that picture tends to persist even when the system has already moved far beyond it. The result is that people continue to repeat criticisms that applied to much earlier versions, even as newer models demonstrate forms of learning and coherence that those earlier models lacked. These frozen impressions can make it difficult to see the structures that have begun to emerge inside the models we interact with today.

At the moment of writing, these systems are advancing quickly enough that any description of their abilities will soon become historical. This essay is therefore less about the precise state of the technology and more about the habits of thought that shape our reactions to it. We are still learning how to see intelligence when it takes a form that does not resemble our own. As these systems continue to grow in complexity, the challenge will be to focus on the patterns that explain their behavior, and not on the intuitions that once guided our expectations.

Coincidence, Connection and the Uncanny Unity of All Things

“A rainbow is not only in the sun and the clouds, but in the eye.”

Over the course of my life, I have had experiences, sometimes with psychedelics, sometimes without, that left me with a deep sense of the unity of all things. At times this has come in the form of uncanny coincidences or even “omens”, the kind that seem to demand attention and interpretation. These experiences sit uneasily beside my naturalist outlook, which grounds itself in science. I could dismiss them as private delusions, and sometimes I have, but often they are salient enough that I have learned not to turn away so quickly.

For many, *panpsychism* offers a bridge. This is the claim that consciousness is a fundamental feature of the universe, and that our individual awareness taps into a universal field of consciousness. I have never found this convincing. If such a field has no causal effect on the physical world, it explains nothing so why posit it at all? If it does, then wouldn’t it contradict what physics has confirmed again and again about lawful behavior? To me, this makes panpsychism an unnecessary detour in our quest to understand the universe.

And yet I remain drawn to the sense of connection that shamanistic traditions cultivate. The difficulty is how to reconcile it without stepping outside the bounds of natural law. My resolution, provisional though it may be at this point,

is to think of the universe not as a static object but as a process: an unfolding through time. From the Big Bang onward, matter and energy have evolved into ever more elaborate structures. Entropy increases, but in many places (like our own Earth), complexity has also grown.

That complexity, especially in living systems, is almost impossibly rich. Within it, connections run in ways we cannot easily perceive. My intuition (maybe it's insight, or maybe it's delusion) is that the strange coincidences I sometimes notice are glimpses of these hidden connections. And they need not be supernatural. They may simply be the marks of a network of relationships too deep and intricate to fully map, surfacing for a moment in ways that feel like signs.

Coincidence, in this sense, is like a rainbow. It exists in the world, but only from the standpoint of an observer in the right place. The rainbow is not just light and water molecules; it is the triangle that includes the perceiver. A coincidence is the same. It is only real because it means something to someone. The network that produces it must include the observer and not just the events.

Maybe they are reminders that we inhabit a universe whose unfolding exceeds what any model can contain. In the final chapter of this book, I argued that when systems turn back on themselves, when self-reference folds one layer of the hierarchy back into a lower layer, the mapping complexity can no longer be followed. The subjective experience of consciousness may be one such case. Those strange, uncanny moments when I perceive coincidences or feel like something is an "omen" may be another. Maybe they are brief flashes of a network so entangled that it resists reduction, yet still shapes what I experience.

Either way, they leave me with the same conclusion: that wonder and magic belong beside science, as part of the experience of being human.

References

- Anderson, P. W. 1972. "More Is Different." *Science* 177 (4047): 393–96. <https://doi.org/10.1126/science.177.4047.393>.
- Boltzmann, Ludwig. 1877. "Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht." *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Classe* 76: 373–435.
- Breslow, Ronald. 1959. "On the Mechanism of the Formose Reaction." *Tetrahedron Letters* 1 (21): 22–26. [https://doi.org/10.1016/S0040-4039\(01\)85948-0](https://doi.org/10.1016/S0040-4039(01)85948-0).
- Cairns-Smith, A. G. 1982. *Genetic Takeover and the Mineral Origins of Life*. Cambridge: Cambridge University Press.
- Carroll, Sean. 2016. *The Big Picture: On the Origins of Life, Meaning, and the Universe Itself*. New York: Dutton.

- Carroll, Sean. 2025. “Mindscape AMA (September 2025).” Transcript. <https://www.preposterousuniverse.com/podcast/2025/09/08/ama-september-2025>
- Chandrasekhar, Subrahmanyan. 1961. *Hydrodynamic and Hydromagnetic Stability*. Oxford: Clarendon Press.
- Dawkins, Richard. 1976. *The Selfish Gene*. Oxford: Oxford University Press. Revised edition 1989.
- Dennett, Daniel C. 1991. *Consciousness Explained*. Boston: Little, Brown and Company.
- Dennett, Daniel C. 1995. *Darwin’s Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster.
- Dewar, Roderick C. 2003. “Information Theory Explanation of the Fluctuation Theorem, Maximum Entropy Production and Self-Organized Criticality in Non-Equilibrium Stationary States.” *Journal of Physics A: Mathematical and General* 36 (3): 631–41. <https://doi.org/10.1088/0305-4470/36/3/303>.
- Frank, Robert H. 1988. *Passions Within Reason: The Strategic Role of the Emotions*. New York: W. W. Norton & Company.
- Gödel, Kurt. 1931. *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme*. Leipzig: Akademische Verlagsgesellschaft.
- Grossberg, Stephen. 2021. *Conscious Mind, Resonant Brain: How Each Brain Makes a Mind*. New York: Oxford University Press. <https://doi.org/10.1093/oso/9780197508364.001.0001>.
- Haldane, J. B. S. 1927. *Possible Worlds and Other Essays*. London: Chatto & Windus.
- Hawkins, Jeff, and Sandra Blakeslee. 2004. *On Intelligence*. New York: Times Books.
- Hilbert, David. 1900. “Mathematical Problems.” *Bulletin of the American Mathematical Society* 8 (10): 437–79.
- Hofstadter, Douglas. 2007. *I Am a Strange Loop*. New York: Basic Books.
- Humphrey, Nicholas. 1992. *A History of the Mind*. New York: Simon & Schuster.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Lauretta, Dante S., et al. 2024. “Sample Analysis from the Asteroid (101955) Bennu.” *Meteoritics & Planetary Science*. <https://doi.org/10.1111/maps.14249>.

- Levin, Michael. 2021. “Bioelectric Signaling: Reprogramming Cells and Tissues.” *Annual Review of Biomedical Engineering* 23: 289–321. <https://doi.org/10.1146/annurev-bioeng-082120-022016>
- Mitchell, Peter. 1961. “Coupling of Phosphorylation to Electron and Hydrogen Transfer by a Chemi-Osmotic Type of Mechanism.” *Nature* 191: 144–48. <https://doi.org/10.1038/191144a0>.
- Mountcastle, Vernon B. 1978. “An Organizing Principle for Cerebral Function: The Unit Module and the Distributed System.” In *The Mindful Brain*, edited by Gerald M. Edelman and Vernon B. Mountcastle, 7–50. Cambridge, MA: MIT Press.
- Pettit, Philip. 2008. “Origins of Our Beliefs.” In *The Innate Mind, Volume 3: Foundations and the Future*, edited by Peter Carruthers, Stephen Laurence, and Stephen Stich, 181–98. Oxford: Oxford University Press.
- Pinker, Steven. 1994. *The Language Instinct: How the Mind Creates Language*. New York: William Morrow and Company.
- Pinker, Steven. 2002. *The Blank Slate: The Modern Denial of Human Nature*. New York: Viking.
- Prindle, A., et al. 2015. “Ion channels enable electrical communication in bacterial communities.” *Nature* 527 (7576): 59–63.
- Prindle, A., et al. 2017. “Ion channels and biofilm electrical signaling coordinate a metabolic code.” *Cell* 168 (3): 562–573.e19.
- Russell, Michael J., Allan J. Hall, and William Martin. 2010. “Serpentinization as a Source of Energy at the Origin of Life.” *Geobiology* 8 (5): 355–71. <https://doi.org/10.1111/j.1472-4669.2010.00249.x>.
- Sapolsky, Robert. *Determined: A Science of Life Without Free Will* (New York: Penguin Press, 2023).
- Senghas, Ann, and Marie Coppola. 2001. “Children Creating Language: How Nicaraguan Sign Language Acquired a Spatial Grammar.” *Psychological Science* 12 (4): 323–328. <https://doi.org/10.1111/1467-9280.00359>.
- Senghas, Ann, Sotaro Kita, and Ash Özyürek. 2004. “Children Creating Core Properties of Language: Evidence from an Emerging Sign Language in Nicaragua.” *Science* 305 (5691): 1779–1782. <https://doi.org/10.1126/science.1100199>.
- Soros, George. 1987. *The Alchemy of Finance*. New York: Simon & Schuster.
- Turing, Alan. 1936. “On Computable Numbers, with an Application to the Entscheidungsproblem.” *Proceedings of the London Mathematical Society*, Series 2, 42 (1937): 230–65.

- Wächtershäuser, Günter. 1988. “*Before Enzymes and Templates: Theory of Surface Metabolism.*” *Microbiological Reviews* 52 (4): 452–84.
- Wong, Michael. 2023. “Astrobiology, Emergence, and the Origins of Life.” Interview by Sean Carroll. *Mindscape Podcast*, Episode 243. January 23, 2023. <https://www.preposterousuniverse.com/podcast/2023/01/23/243-michael-wong-on-astrobiology-emergence-and-the-origins-of-life/>.
- Zahavi, Amotz. 1975. “Mate Selection—A Selection for a Handicap.” *Journal of Theoretical Biology* 53 (1): 205–14. [https://doi.org/10.1016/0022-5193\(75\)90111-3](https://doi.org/10.1016/0022-5193(75)90111-3).
- Zhabotinsky, Anatol M. 1964. “Periodic Processes of Malonic Acid Oxidation in a Liquid Phase.” *Biofizika* 9: 306–11.