The Etiology of Progress

Admissibility Across Brains, Molecules, Energy, and Institutions

Flyxion

August 24, 2026

Contents

Preface

This book began with a question simple enough to state in one line and answer badly in one line more. Six words recur throughout it: possible, admissible, reachable, stable, legible, reconstructible. The first temptation, on encountering six related terms, is to arrange them in a sequence, each narrowing the one before it, so that a state's fate is decided by how far down the funnel it manages to fall. Chapter 1 wrote that sequence out explicitly and then declined to defend it. The decision to postpone judgment, rather than assume the funnel and spend the rest of the book illustrating it, is the single methodological choice everything else here follows from.

The chapters that make up Parts I through IV were not written to confirm the funnel. They were written as tests of it, applied to four substrates chosen because they share essentially nothing at the level of implementation: a nervous system learning and being decoded, a chemical system being synthesized and manufactured, an energy system being harnessed and sustained, and an institution deciding what gets attempted and what gets kept. A neuron, a molecule, a plasma, and a grant review panel have no mechanism in common. If a single structure kept appearing across all four regardless, that would be evidence the structure was real rather than an artifact of any one domain's particular physics.

Something did keep appearing, and it was not the funnel. A directional manifold-learning result in Chapter 2 showed reachability could not be a property of a destination at all, only of a specific transition from a specific starting point. A decoding result in Chapter 3 showed legibility varying with the observer applying it, not with the state being observed. A civilizational case in Chapter 10 showed the funnel's final link failing in both directions independently, a specification surviving for two thousand years without producing reconstruction, and reconstruction being achieved routinely with no specification ever having existed. By the time Chapter 11 returned to the question Chapter 1 had postponed, the funnel was no longer a live candidate. It had been tested, chapter by chapter, against material it was never asked to anticipate, and it had not survived.

What replaced it was not decided in advance either. The book's working method was to keep a running ledger, at the close of every chapter, of which relations among the six original terms that chapter's cases had actually established, and to resist naming a final structure until the evidence gave one shape. Chapter 11 proposes a candidate object, a directed, indexed transition geometry, and then does something a synthesis chapter is not obligated to do: it tests the candidate against the accumulated ledger rather than declaring victory

once the candidate is stated. The test found the candidate wanting in two specific places. A result from the neural chapters forced the object to depend on trajectory rather than present state alone. A result appearing independently in both the neural and institutional chapters, in substrates sharing no mechanism whatsoever, forced the addition of a seventh element the original six-part sketch had no room for. Both additions have identifiable empirical causes. Neither was anticipated when the book began.

This is offered as the strongest evidence available for the book's approach, stronger than any single argument the finished chapters make. The framework governing this book was itself subjected to an admissibility constraint. It was not permitted to keep a structure the evidence could no longer support, and when the evidence forced a change, the change is documented rather than smoothed over. A reader who wants to see exactly where and why should look first at Chapter 11, Section 5, where the failure is recorded in the same terms as every other result in the book, and not treated differently for being a failure of the book's own apparatus rather than of someone else's.

What follows, then, is not a history of progress and not a celebration of it. It is closer to an etiology: an attempt to trace, case by case and substrate by substrate, the specific and often non-obvious conditions under which a change becomes durable, rather than assuming those conditions are simply present wherever a change can be imagined. Progress, treated this way, stops being a single undifferentiated capacity a system either has or lacks, and becomes a phenomenon with traceable, substrate-specific causes, some of which turn out to be counterintuitive once actually traced.

The chapters can be read in the order they were written, following the argument's actual construction, or read by part, treating each of the four substrates as a self-contained case study that happens to share a vocabulary with its neighbors. Either way, the book's central claim survives intact: capability is not the size of a possibility space. It is the structure of what can be traversed, sustained, observed, and recovered from where a system actually stands, and that structure, this book argues, has a geometry worth taking as seriously as the systems it governs.

Chapter 1

The Admissibility Problem

1.1 The Unit of Analysis

Progress is usually described as an expansion of capability. A civilization, a laboratory, or a learning system is said to advance when something that was previously impossible becomes possible: a new material can be synthesized, a new behavior can be produced, a new disease can be treated. This book proposes a different unit of analysis. A system advances not merely when some state becomes conceivable, but when a path to that state can be traversed, sustained, recognized, reproduced, and incorporated into what the system can subsequently do. The unit of analysis is not the state itself but the admissible transformation: a change that a given system, from its actual present condition, using its actual resources and tools, can carry out, keep, and later draw on again.

Clarke's Moonwatcher, from the "Dawn of Man" sequence that opens *2001: A Space Odyssey*, is the cleanest fictional staging of this claim available. The bones scattered around the waterhole were not new. What changed was not the set of objects present to the hominid but the set of transformations available to him: a bone could now be picked up, swung, and used to break another bone, and that transformation, once performed once, was retained and repeated. Nothing in the environment became newly possible that morning. A path through the space of admissible actions became newly reachable, and, crucially, stable: the behavior persisted past the first execution and propagated to others. The scene is often read as being about intelligence or about violence. Read through the vocabulary this book is building, it is more precisely about the moment a system crosses from a state in which a transformation is merely admissible to one in which it is reachable, and then retains it long enough for it to become a stable part of what the system can subsequently do. That three-part crossing, admissible to reachable to stable, is the smallest complete unit the rest of this book will analyze at increasing scale.

This distinction sounds pedantic until it is applied. A molecule may be permitted by every law of chemistry and still be beyond the reach of any known synthesis route. A neural population may encode a distinction that no decoding algorithm can currently recover, even in principle, from the recorded signal. A technology may have existed once, in com-

plete working form, and be permanently lost, not because the physics stopped applying but because the people, tools, and institutional memory required to rebuild it no longer exist anywhere. In each case, capability talk collapses several very different situations into a single word. The purpose of this chapter is to pull them apart.

The four parts that follow this one apply the same question to four different substrates. Brains search over states they can learn. Molecular systems traverse configurations they can physically realize. Energy systems determine which of those traversals can be sustained. Innovation systems, which is to say institutions, search over ways of altering the admissibility conditions of all three at once, deciding which paths get attempted, financed, remembered, and handed forward. That four-domain formulation is the promise of the book. This chapter exists to earn the vocabulary the rest of the book will spend.

1.2 A First Attempt at an Ordering

The most natural first move is to propose a chain. A state is first *possible*: consistent with the basic laws governing the substrate, ruling out only what those laws forbid outright. Among the possible states, some are *admissible*: consistent not merely with the abstract laws but with the actual constraint set that defines the system under discussion, including its materials, its available operations, and the rules that govern how one configuration may give rise to another. Among the admissible states, some are *reachable*: there exists an actual path, made of the system's real transformation steps, from where the system currently stands to that state. Among the reachable states, some are *stable*: once reached, they persist for some meaningful interval rather than immediately relaxing back to where they came from. Among the stable states, some are *legible*: an observer with some specified set of tools can detect, distinguish, or read out that the state obtains. And among the legible states, some are *reconstructible*: a process exists, or could be assembled from available knowledge and materials, that would allow the state to be produced again, independently of whatever produced it the first time.

Written this way, the six terms look like a funnel, each stage a strictly narrower subset of the one before it. That is the picture to test, and it does not survive contact with real cases. The chain is a useful first pass precisely because working through where it breaks reveals the actual structure faster than starting from definitions alone.

1.3 Testing the Chain: Six Counterexamples

1.3.1 Possible but Inadmissible

Physical law permits a diamond the size of a house. Nothing in thermodynamics or in the chemistry of carbon forbids such an object; the bonding rules that make diamond possible at all place no upper bound on how large a single crystal could in principle grow. The object is possible. It is inadmissible, however, given the actual constraint set of any manu-

facturing process presently available: no known combination of pressure vessel, catalyst, growth rate, and defect-control method can carry a diamond lattice to that scale without the accumulated strain and impurity gradients fracturing it long before completion. The gap here is not between what is imaginable and what is achievable in some vague sense. It is a specific mismatch between a state permitted by the base laws of the substrate and a state permitted by the narrower operating constraints of any actual process that acts on that substrate. Admissibility, unlike possibility, is always relative to a named set of operations, not to the substrate's laws taken in the abstract.

1.3.2 Admissible but Unreachable

Consider a trained neural network sitting at some point in its parameter space, partway through optimization. Elsewhere in that same space, admissible under the network's architecture and loss function, sits a configuration that would solve the task far better: nothing about the architecture forbids those weights, and if the network were initialized there directly, it would perform well. From the network's actual current position, however, gradient descent may never find a path to that configuration, because the loss surface between here and there passes through a region no local, incremental update rule can cross. The destination is admissible. It is unreachable, not because it violates any constraint of the substrate, but because reachability is a property of the pair formed by a starting state and a transformation rule, not a property of the destination alone. Two systems with identical admissible regions can differ completely in what is reachable from them, simply because they started in different places or because their update rules are shaped differently. This is the clearest demonstration that admissibility and reachability are genuinely distinct axes rather than adjacent points on a single scale: an entire region can be uniformly admissible while remaining almost entirely unreachable from where a particular system happens to stand.

1.3.3 Reachable but Unstable

Rapid quenching can carry a metal alloy into a metastable crystalline phase that is both admissible, in that it does not violate the alloy's chemistry, and reachable, in that the quenching process is a real, executable transformation. Left alone, however, that phase decays back toward a lower-energy configuration over hours, days, or years, depending on temperature and composition, unless a further process, such as controlled annealing, actively maintains it. Something structurally similar happens in induced pluripotent stem cell reprogramming: a differentiated cell can be driven into a pluripotent state that is reachable through a defined protocol, but the state does not persist on its own and requires continuous maintenance conditions, specific factors, media, and substrate, or the cell relaxes back toward differentiation or toward death. In both cases the destination was reached. Reaching it and keeping it are different accomplishments, and a great deal of engineering effort in materials science and cell biology is spent entirely on the second problem after the first

has already been solved.

1.3.4 Stable but Illegible

A synaptic weight change encoding a memory trace can be genuinely stable, persisting in the physical structure of a dendrite or a set of connections for years, and still be entirely illegible to any presently available reading technology. No noninvasive imaging method can currently recover the specific content encoded by a specific configuration of synaptic strengths in a specific circuit; even highly invasive methods, applied to animal models, can only decode particular, deliberately engineered signals rather than arbitrary stored content. The state is not hidden in the sense of being suppressed or encrypted. It is simply outside the resolving power of any observer's current instruments. This decouples legibility from stability in a way the naive chain does not allow for: legibility is a property of the relationship between a state and an observer's toolkit, not a further-narrowed property of the state itself, and a state can therefore be maximally stable while being read by nobody at all.

1.3.5 Legible but Unreconstructible

The Saturn V's F-1 engine is a case NASA itself confronted directly. The engine's design was documented: blueprints existed, engineering drawings existed, and the vehicles themselves survived in museums. The state was legible in the strongest sense, fully observable and fully recorded. Decades later, engineers attempting to build a new F-1-class engine discovered that the drawings alone were not sufficient to reconstruct the object. The original welding techniques, the specific tolerances achieved by specific machinists on specific tooling, and a great deal of process knowledge that had never been written down because it lived in the hands and judgment of individual workers, had been lost along with the workforce and the factory floor that held it. Reconstruction required not retracing the original path but discovering, essentially from scratch, a new admissible path to a state that had already been fully legible for fifty years. Legibility, in other words, does not entail reconstructibility even when it is complete. What is missing is not information about the destination but a currently admissible route to it, and that route can vanish independently of the documentation describing the state it once produced.

1.3.6 Reconstructible but Prohibitively Costly

A vaccine whose synthesis pathway is fully published, whose reagents are all commercially available, and whose process any qualified laboratory could in principle carry out, can still be effectively unreconstructible in practice, not because any physical or informational barrier stands in the way but because maintaining a certified pilot plant, sourcing specialized reagents at the required purity, and completing the regulatory approval process again from zero costs more, in money and in time, than any actor is willing or able to spend. Here the constraint has moved outside the physical and informational domains entirely and become institutional and economic. This case matters because it shows that the chain of distinctions

built so far cannot stop at the boundary of a laboratory or a substrate. Cost, certification, and organizational capacity are themselves constraints that determine what is admissible, in exactly the same structural sense that a catalyst's operating temperature determines what is admissible in a chemical process. They simply operate at a different scale and are enforced by different mechanisms.

1.4 Why the Chain Is Not a Chain

The six cases were chosen to be clean, but taken together they do more than fill in six boxes. They show that the six terms are not stages of a single funnel narrowing toward reconstructibility. They are, instead, largely independent axes, each capturing a different way a transformation can fail, and a system can fail along any one of them without failing along the others.

Legibility in particular cuts across the sequence rather than sitting inside it. The synaptic case showed a state that is reachable and stable while remaining illegible. The Saturn V case showed a state that is legible while remaining unreconstructible, and it is worth noting that legibility there long preceded any question of reconstruction: the drawings existed for decades before anyone tried to rebuild the engine, so legibility was not merely a way station on the path to reconstruction but a separate, freestanding property that reconstruction depended on without following from. A state can even be legible without ever having been stable in the ordinary sense: the light from a supernova is fully legible to an observer today, in that its properties can be measured with precision, even though the event itself was a one-time transient with no persisting physical structure to speak of and no path back to it under any transformation, since the star that produced it no longer exists to be acted upon again. Legibility, then, is best treated not as the fifth link in a chain but as an independent relation between a state, however it arose, and an observer's toolkit at a given time.

Reconstruction has a further complication that the chain obscures. The naive picture treats reconstruction as the terminal stage of the original path, as though rebuilding a state simply meant retracing, in reverse, the steps that produced it the first time. The F-1 engine case shows this is false in general. What NASA's engineers needed was not the original path, which no longer existed as an executable sequence of actions available to anyone, but a new admissible path, built from currently available tooling, materials science, and labor, that happened to converge on a previously achieved state. Reconstruction, in other words, is not merely a memory operation performed on the terminus of an old path. It is itself a fresh search for an admissible path to a specified target, and that search is subject to exactly the same possible, admissible, reachable, and stable distinctions as any other search, only now indexed to whatever constraints happen to hold at the later time. This is why reconstructive capacity, which becomes central in the fourth part of this book, cannot be treated as a passive residue of documentation. It is an active capability that a system either retains, rebuilds, or permanently loses, and losing it is a distinct event from losing the documented knowledge of what was once achieved.

1.5 Admissibility Is Indexed

The cases above share a structural feature worth making explicit, because it is what allows the same vocabulary to be reused across four substrates as different as neurons, molecules, energy grids, and institutions without collapsing into an empty analogy. Nothing is admissible in an unqualified sense. Admissibility is always indexed to a substrate, a starting state, a set of available operations, and a set of resource and institutional constraints that bound what those operations can actually accomplish. The diamond case indexed admissibility to a manufacturing process rather than to the laws of carbon chemistry. The neural network case indexed reachability to a specific starting point and a specific update rule, not to the architecture's admissible region as a whole. The vaccine case indexed reconstructibility to an economic and regulatory environment that has nothing to do with the chemistry of the synthesis itself.

This indexing is what makes the framework transportable rather than merely metaphorical. When the index changes, meaning the substrate, the tooling, the energy budget, the institutional structure, or the historical moment, the admissible region changes with it, sometimes without any change at all to the underlying laws governing the substrate. A configuration inadmissible in vivo can be admissible in vitro, because the constraint set defining what counts as an allowed transformation has changed even though the target molecule has not. A technological capability lost to one generation of workers and factories can become admissible again to a later generation with different tools, different materials science, and a different willingness to spend resources rebuilding a lost path, as the F-1 case eventually showed when a later reconstruction effort succeeded using contemporary methods rather than the original ones. This is also what licenses the book's central architectural claim: brains, molecules, energy systems, and institutions are not secretly the same kind of thing, and nothing in this framework requires them to be. What they share is not substance but structure, namely that each defines, at any given moment, an admissible region relative to its own substrate, state, operations, and constraints, and each therefore raises the identical question of which transformations across that region are actually possible, reachable, stable, legible, and reconstructible, even though the concrete answers differ completely from one substrate to the next.

1.6 What the Later Chapters Must Demonstrate

The distinctions established here are not decorative and are not meant to be taken on faith once stated. Each of the four parts that follow is required to reproduce, in its own substrate, a version of the same load test just performed in the abstract. The NeuroAI part must show a concrete case of a representation that is admissible to a learning system but unreachable from its current trained state, and a concrete case of a stable neural state that remains illegible to current decoding methods, echoing the counterexamples above but drawn from the specific empirical literature on brains and learning systems rather than from illustra-

tive analogy. The molecular manufacturing part must show, with actual chemistry, the gap between a molecule being conceivable, being synthesizable by some known pathway, and being manufacturable at industrially useful yield, since this is precisely where the possible, admissible, and reachable distinctions were shown above to do real separating work. The energy part must show cases where abundant energy expands the admissible region without specifying which path within it gets taken, since energy's role in this framework is to loosen a constraint rather than to select a destination, and the chapter must resist collapsing into the energy-determinist reading this book explicitly sets out to avoid. The innovation paradigms part must show, at the institutional scale, cases in which a state was legible but not reconstructible, and cases in which reconstruction required a genuinely new admissible path rather than a retracing of an old one, extending the F-1 pattern from a single engineering artifact to entire technological lineages and the institutions that carry or fail to carry them forward.

If any of the four parts cannot produce clean cases of this kind in its own domain, that is not a failure of the domain. It is a signal that the vocabulary constructed in this chapter needs revision before it is allowed to propagate any further, and the chapter should be treated as provisional until each transplant has been attempted and has held.

What this chapter has attempted, in other words, is not a definition of progress but a first pass at its etiology: a search for the specific conditions, rather than a single undifferentiated capacity, that determine whether a given change becomes durable. The four parts that follow are that search, conducted independently in four substrates that share nothing at the level of implementation.

Part I

Neural and Cognitive Substrates

Chapter 2

What a Brain Can Learn

2.1 An Ordinary Question

Can this nervous system acquire this distinction. The question sounds like the opening line of an undergraduate cognitive science lecture, and for most of its history it has been treated that way, answered by pointing to a capacity, a critical period, an amount of training data, or an inductive bias built into the architecture of a particular circuit. This chapter argues that the ordinary phrasing conceals the book's actual claim, because the word doing the real work in the sentence is not "distinction" and not "nervous system" but "can." Chapter 1 argued that capability talk collapses several structurally different situations into a single word. This chapter exists to show that a brain is the first substrate in which that collapse is not a philosophical nicety but an empirically documented, measurable failure of description, one with a substantial experimental literature behind it that has already been forced, independently of any framework proposed here, to draw distinctions that look very much like the ones Chapter 1 introduced in the abstract.

The chapter's task is narrower than a survey of what is currently known about learning and the brain, and the narrowing is deliberate. Its only job is to provide the book's first empirical stress test of admissibility: to show, with real cases rather than illustrative analogy, that a representational state can be imaginable by an external theorist without being learnable by a particular nervous system, and that even among learnable states, some are reachable only from particular developmental histories, training regimes, bodily configurations, or previously acquired representations. A companion question, whether a state that has in fact been learned can subsequently be read out by an observer, belongs to Chapter 3 and is deliberately excluded here. Keeping that boundary intact is not a matter of tidiness. It protects one of the book's most useful distinctions from collapsing at the first opportunity, and the reasons for that will become clear by the chapter's final section.

2.2 Learning as Constrained State-Space Transformation

The default picture of learning, inherited largely from information-theoretic treatments of perception and from early connectionist enthusiasm about universal function approximation, treats a nervous system as a receiver that acquires information from its environment and stores it. On this picture, what a brain can learn is bounded mainly by exposure: enough examples, enough repetition, enough time, and in principle any distinction present in the environment's structure becomes learnable. The picture is not wrong so much as it is dangerously incomplete, because it treats the learner's own current organization as a passive container rather than as an active constraint on which transformations of that organization are available at all.

An early and still striking demonstration comes from research on ocular dominance and binocular vision in kittens. Closing one eye during a specific early window of development produces a permanent loss of function in that eye, even though the eye itself remains structurally intact and even though the same closure imposed before or after the window produces little or no lasting effect [1]. The information the closed eye would have delivered was, in an ordinary sense, available in the environment throughout the animal's life. What had changed was not the information but the admissible region of synaptic reorganization available to the visual cortex at that developmental moment. Outside the window, the same input no longer drives the same category of structural change, because the tissue's own developmental trajectory has already committed itself to a narrower set of admissible states. This is not a story about information acquisition failing. It is a story about a transformation becoming inadmissible to a specific substrate at a specific point in its own history, independent of what the environment continued to offer.

A related and more directly relevant case comes from work on constrained learning within populations of cortical neurons engaged in brain-computer interface control. When monkeys are trained to control a cursor through a neural interface, the mapping from neural activity to cursor movement can be altered by the experimenter in two structurally different ways. One class of alteration keeps the new mapping within the same low-dimensional manifold of neural activity patterns the animal's cortex was already capable of producing, only reassigning which patterns correspond to which outputs. The other class of alteration requires activity patterns that fall outside that manifold entirely. Animals learn the first kind of alteration quickly, often within a single session. They struggle profoundly, and in some reported conditions fail outright, with the second kind, even when given extensive additional training [3]. The distinction the experimenters were forced to draw, between mappings that stay inside the existing manifold of producible neural states and mappings that require states outside it, is precisely the distinction Chapter 1 called admissible versus inadmissible relative to a substrate's actual constraint set. The cortex here is not failing to acquire information about the new mapping. It is being asked to enter a region of its own state space that its current organization does not make available as a target of learning at all, regardless of how much training time is provided.

A useful terminological correction follows from these two cases together. Learning

should not be described as the acquisition of information from an environment. It should be described as a constrained transformation of a state space, where the constraints are supplied jointly by the substrate's current organization and by whatever developmental or synaptic mechanisms are available to alter that organization further. Some transformations are available at every stage of that trajectory. Others are available only during specific windows, and still others may never become available at all under ordinary developmental or training conditions, not because the relevant information is absent from the environment, but because no admissible path from the system's actual current organization reaches the state that would encode it.

2.3 Path Dependence: Same Capacity, Different Futures

If learning is constrained transformation rather than information acquisition, a further consequence follows immediately, and it is one the naive information-acquisition picture has no natural way to express. Two systems that appear, by any present behavioral measure, to have equivalent capacities can nonetheless have entirely different sets of states reachable from where they currently stand, because their histories have constructed different admissible transitions out of what looks, from the outside, like the same starting point.

The clearest demonstration is a natural experiment rather than a laboratory manipulation: the restoration of vision, through surgery, to congenitally blind individuals whose sight becomes correctable later in life. Project Prakash, a program that identified and surgically treated congenitally blind children and adults in India, produced a striking and repeated finding. Newly sighted individuals could see in every low-level sense immediately after surgery, with intact retinas, intact optic nerves, and functioning early visual cortex, and yet they could not perform object recognition tasks that require matching a shape seen for the first time to the same shape previously known only by touch [5]. The capacity for basic visual processing was present. The information needed for cross-modal object recognition was, in an ordinary sense, available to the visual system for the first time. What was absent was a developmental history that had constructed the specific admissible transitions, built during the ordinary sighted infant's early months of correlated visual and tactile experience, that allow a newly seen shape to be bound to an already known felt shape. A sighted infant of a few months old and a newly sighted adult from this population can, on many low-level visual measures, look similar in what their visual system can currently register. They do not have the same reachable futures, because one of them arrived at the present moment along a path that built certain transitions and the other did not.

The same structure appears, in a gentler form, in second-language acquisition research. Learners who acquire a language in adulthood commonly reach very high levels of vocabulary and even syntactic competence, and behavioral testing shows that the earlier in life second-language exposure begins, the more native-like the eventual outcome tends to be across a range of grammatical structures, with the effect strongest in exactly those structures whose native-language mastery depends on the same early sensitive-period mecha-

nisms implicated in other domains [2]. Two adult learners can enter a language class with identical prior exposure to the target language, identical general cognitive ability by any standard measure, and identical motivation, and still diverge substantially in which grammatical distinctions become reliably learnable for them, depending on subtle features of each learner's own linguistic developmental history that have nothing to do with the target language itself. The relevant constraint is not a difference in present capacity. It is a difference in which transitions out of the present state are admissible, given a specific history that already shaped the substrate before the target language was ever encountered.

Path dependence of this kind is not a minor caveat to be appended to an otherwise flat account of learnability. It is direct evidence against treating the set of learnable states as a fixed property of a substrate type, independent of history. The same cortex, the same general architecture, and even the same present behavioral profile can sit at the entrance to very different admissible regions depending on the specific sequence of prior transformations that produced the present state, and no measurement of present capacity alone can reveal which region a given system currently occupies.

2.4 Continuation Geometry: The Present as Constraint on What Follows

The cases above establish that history matters and that admissibility, not raw capacity, is the operative constraint. A further and more structural point remains to be made, because path dependence by itself could still be read as a claim only about starting conditions: different histories produce different starting points, and different starting points admit different futures, in the same way that a chess position determines which moves are legal next. The stronger claim this book needs is that a nervous system's present organization does not merely occupy a state. It actively shapes the geometry of what can follow from that state, in a way that goes beyond simply specifying a starting point on an otherwise fixed map.

Evidence for this stronger claim comes from work on the population-level dynamics of prefrontal cortex during context-dependent decision-making. When animals perform a task that requires attending to one sensory feature and ignoring another, depending on a contextual cue, the relevant computation cannot be recovered by looking at any single neuron in isolation. It becomes visible only at the level of trajectories through the joint activity space of the recorded population, where the same sensory input drives the population along different trajectories depending on which context is currently active, and where the geometry of the space itself, not merely the momentary point occupied within it, determines which subsequent trajectories are available in response to a given input [6]. The finding that matters for this chapter is structural rather than task-specific: the population's current organization defines a landscape of admissible subsequent trajectories that is not reducible to the state currently occupied within it, in the same way that a river's channel constrains which paths downstream water can take without the channel itself being any single point the water passes through.

The brain-computer interface manifold findings introduced earlier extend the same point directly into the domain of learning itself, and this is where continuation geometry becomes indispensable rather than merely descriptive. Later work in that same experimental tradition showed that when animals are given repeated opportunities to learn mappings outside their existing manifold, the manifold itself can gradually be reshaped through extended training, expanding what is subsequently learnable, but the reshaping proceeds along specific directions determined by the manifold's own current geometry rather than uniformly in all directions [4]. This is the clearest available demonstration that a substrate's present organization functions as a constraint on its own future admissible region, not merely as a location within a fixed one. Learning something new does not just add a new state to an unchanged space. It can alter which further states subsequently become reachable, and it does so unevenly, favoring extensions along directions already partially supported by the existing organization over directions that are not.

This is precisely the object Chapter 1 called continuation geometry when it distinguished reachability, a property of a specific starting point and transformation rule, from admissibility considered in the abstract. The brain gives that abstract distinction concrete, measured content: the manifold is the admissible region, its current shape is the continuation geometry, and the experimentally documented fact that learning reshapes the manifold unevenly is a direct empirical instance of a system's present organization constraining not just its current output but the future trajectory of its own admissible region.

2.5 The Objection: Is Admissibility Doing Any Work

The strongest challenge available to this chapter, and the one it has to answer honestly before proceeding to Chapter 3, is that nothing said so far requires a new concept at all. A skeptical reader could grant every case above and respond that developmental critical periods, path dependence, inductive bias, and manifold constraints are already well-established, well-named concepts in their respective literatures, each with its own mechanisms and its own predictive successes, and that "admissibility" is simply a redundant umbrella term draped over an existing and perfectly functional vocabulary.

The objection deserves to be taken seriously rather than deflected, because much of what makes a unifying concept worth introducing is exactly its vulnerability to this kind of challenge. The answer this chapter proposes is that the unified concept earns its place only if it generates a prediction that the piecemeal vocabulary, used separately, does not naturally generate. Here is that prediction, stated plainly. If a system fails to acquire a given distinction, the piecemeal vocabulary offers no principled way to decide, in advance, whether additional training of the same kind will eventually succeed. Admissibility does offer such a way, because it distinguishes two structurally different failure modes that the case studies above have already shown to be empirically separable. A failure that occurs near the interior of a system's current admissible region, where the target state differs from currently producible states only along directions the system's own manifold already par-

tially supports, should be recoverable with continued training of the same general kind, because the target lies within the reachable extension of the existing organization. A failure that occurs at or beyond the actual boundary of the admissible region itself, where the target state requires activity patterns, synaptic configurations, or cross-modal bindings that the current organization does not support along any direction, will not be recoverable by more of the same kind of training, no matter how extensive, because no admissible path connects the current organization to the target regardless of how much time is allowed to traverse it.

This is not merely a restatement of the manifold-versus-outside-manifold finding, because that finding was reported as a single empirical result about one task. The claim here is that the admissible-versus-reachable-versus-boundary distinction, once made explicit as a general property of a substrate's organization rather than a fact about one experiment, predicts in advance which future experiments should show recoverable failure and which should show persistent failure, based on a measurable property of the system, namely its manifold geometry relative to the target, rather than on trial and error with additional training. The ocular dominance case and the Project Prakash case, read this way, are not merely illustrations of the same abstract point already made by the manifold studies. They are independent tests of the same prediction in entirely different substrates and time scales, developmental windows measured in weeks for kittens, irreversible adult outcomes measured in years for formerly blind humans, and population dynamics measured in milliseconds for behaving monkeys, and all three come out the same way: failures near an admissibility boundary do not yield to additional exposure of the same kind, while failures within an admissible region do. A vocabulary of separately named mechanisms, critical periods here, inductive bias there, manifold constraints somewhere else, has no obligation to make this cross-case prediction align, because each mechanism was defined locally to its own literature. Admissibility, defined once and applied across all three, is obligated to align, and in these cases it does. That is what the unified concept is for, and it is also the standard the rest of this book is committed to meeting in every subsequent domain: not a relabeling exercise, but a generator of predictions that would fail if the underlying structure were not genuinely shared.

2.6 What This Chapter Deliberately Leaves for Chapter 3

It is worth stating the boundary with Chapter 3 explicitly rather than letting it emerge by omission, because the two questions are easy to blur in ordinary usage and the blurring is exactly what this book is designed to resist. Every case in this chapter concerned whether a nervous system's own organization permits a given transformation to occur and to persist within that system. None of them concerned whether a state, once present, could be read out by an external observer using some measurement technology. The ocular dominance case is about whether the cortex reorganizes, not about whether an experimenter can detect that it has. The manifold-learning case is about whether the population's own activity can

be driven into a new pattern, not about whether a decoding algorithm applied after the fact could recognize that pattern as meaningful. The Project Prakash case is about whether cross-modal binding develops, not about whether an outside test could verify it. Keeping these separate matters because the two failures have different diagnostic signatures and, on the account developed here, different remedies. A learning failure, addressed in this chapter, is a failure of the substrate to reach or retain a state. A decoding failure, reserved for the next chapter, is a failure of an external system to recover a state the substrate has already reached and is already retaining. Confusing the two would make it impossible to ask, of any given case, which of two very different problems is actually present, and Chapter 3 depends on that question remaining answerable.

2.7 Ledger

The following relations among the terms introduced in Chapter 1 were established by the cases examined here and are recorded for use in the synthesis chapter, without yet committing to what overall structure they imply.

Possible states exist that are not admissible to a given substrate at a given developmental stage: the closed-eye case shows a fully specifiable, physically ordinary state, namely a normally wired visual cortex responsive to both eyes, that becomes inadmissible to a specific cortex once its developmental window closes, despite being freely admissible earlier in the same animal's life.

Admissible states exist that are unreachable from a given present state: the outside-manifold cursor mappings were, by the researchers' own later demonstration, admissible in the sense that extended training could eventually reshape the manifold to include them, yet unreachable from the animal's initial trained state within any ordinary single-session or short-term training regime.

Reachability is asymmetric with respect to direction relative to a system's current organization: the manifold-reshaping findings show that some directions of extension are reachable through training and others are not, from the identical starting manifold, meaning reachability is not a single scalar quantity attached to a target state but a property that depends on the direction of approach relative to the substrate's current geometry.

Two systems with matching present capacity can have different reachable futures: the Project Prakash comparisons and the second-language sensitive-period findings both demonstrate present-capacity equivalence, or near equivalence, coexisting with substantial divergence in which further states are subsequently reachable, tracing the divergence to differences in developmental history rather than to differences in current measured capacity.

A substrate's present organization constrains the shape of its own future admissible region, not merely its current state: the manifold-reshaping results show learning altering which further states are subsequently reachable, and doing so unevenly, which is a property of continuation geometry rather than of any single admissible or reachable state considered in isolation.

Failure near an admissibility boundary and failure within an admissible region are empirically distinguishable by whether continued exposure of the same kind resolves the failure: this relation, rather than any single case, is the chapter's primary contribution to the ledger, since it is the one novel prediction generated by treating admissibility as a unified concept rather than as a bundle of separately named mechanisms.

Chapter 3

Decoding and the Legibility Boundary

3.1 The Apparent Triumph

Neural decoding has produced some of the most publicly striking results in contemporary neuroscience. Individuals with paralysis have controlled robotic arms and computer cursors using signals recorded from motor cortex, translating an intention to move into an executed movement of an external device [7]. More recently, a person unable to write by hand had imagined handwriting movements decoded directly into text at rates approaching ordinary typing, with a trained model recovering individual letters from population activity in real time [8]. Results of this kind are routinely, and understandably, described as instances of a state within the brain becoming readable. The temptation these results create for a book like this one is exactly the temptation Chapter 2 warned against from the other direction: having established that a nervous system's own capacity to learn a state is a separate question from whether that state, once present, can be recovered by an outside observer, it becomes easy to assume that a sufficiently successful decoder settles the second question definitively, that legibility, once demonstrated in one striking case, is simply a property a neural state either has or lacks. This chapter exists to show that assumption is false in ways that matter for the rest of the book, and that the apparent triumph of decoding conceals a legibility boundary at least as intricate as the admissibility boundary Chapter 2 documented on the learning side.

3.2 Stable but Illegible: A Working Case

Chapter 1 introduced the possibility of a state that is stable yet illegible using the hypothetical case of a synaptic memory trace no current instrument can read. Chapter 3 needs a working case rather than a hypothetical one, and representational drift supplies it directly. Recordings from posterior parietal cortex in mice performing a well-learned, stably executed task show that the specific population-level activity patterns associated with each behavioral variable change substantially from day to day, even while the animal's behavior

itself remains constant and well within its trained competence [9]. A decoder trained on one recording session, applied without modification to activity recorded days later during behaviorally identical performance, degrades and eventually fails, not because the relevant information has left the population, since the animal's behavior shows the information is still being used correctly, and not because the population has become disorganized, since the drift itself follows a structured, low-dimensional trajectory rather than random noise, but because the fixed mapping the decoder relies on no longer corresponds to how the population currently represents the task.

This is the chapter's first working instance of stability without legibility, and it is stronger than the hypothetical version in one important respect. The hypothetical synaptic trace was illegible because no instrument existed that could read it at all. The drifting population is different: it was legible, by a specific decoder, at one point in time, remained just as stable and just as behaviorally load-bearing afterward, and became illegible to that same decoder purely as a function of elapsed time, without any change to the animal's competence. Legibility, in this case, is shown to have its own temporal dynamics, independent of whether the underlying state remains present, stable, and functionally intact.

3.3 Legibility Is Relational, Not Intrinsic

The natural response to representational drift is to ask whether a better decoder could simply track the drift and remain accurate, and the existence of methods built specifically to do this confirms the point this section needs to establish. Techniques developed to stabilize brain-computer interfaces across days work by aligning the low-dimensional space a new day's recordings occupy with the space the original decoder was trained on, effectively recovering legibility not by reading a fixed representation more carefully but by continuously re-establishing a relation between two different, non-identical representational geometries [10]. The success of these alignment methods is itself evidence against treating legibility as a property attached to a neural state in isolation. What is being restored is not access to an unchanged object but a working correspondence between two objects, the population's current geometry and the decoder's assumed geometry, that must be actively reconstructed because the underlying state did not stay put.

A separate line of work makes the same point from a different angle. Populations of motor cortical neurons engaged in different behavioral tasks, tasks that look unrelated at the level of individual muscles and movements, have been shown to share a common low-dimensional manifold, such that a mapping trained to decode one behavior transfers, with modest adjustment, to decode a structurally related but superficially different behavior [11]. Individual-neuron-level decoding, by contrast, transfers far less readily across the same behaviors, because individual neurons participate in the population code in ways specific to each task even when the population-level geometry is shared. Legibility here depends heavily on which level of description, single neuron or population manifold, the decoder is built to read, and a state can be robustly legible at one level while being nearly

illegible at the other, using recordings drawn from the identical tissue at the identical moment. Put together with the drift result, this establishes that legibility is a three-place relation among a state, a decoder, and a level of description, not a two-place relation between a state and an observer in general, and certainly not a one-place property of the state alone.

3.4 Decoder Dependence and the Asymmetry of the Boundary

Once legibility is granted to be relational, the question Chapter 2 raised about reachability transfers naturally to the decoding side: does the legibility boundary show the same kind of directional asymmetry that reachability showed, where some distinctions become legible readily while others, drawn from what looks like the same underlying representational format, remain resistant.

The evidence available points toward yes, though more unevenly than the reachability case. In the handwriting decoding work, individual letters were not recovered with uniform fidelity. Certain letter pairs, particularly those sharing similar initial strokes in the imagined handwriting movement, were confused by the decoder substantially more often than others, producing an error structure shaped by the geometry of the movements themselves rather than distributed evenly across all possible confusions [8]. A closely related pattern appears in work decoding continuous speech from cortical activity during attempted speech production, where certain phonetic contrasts were reconstructed with high intelligibility while others, generally those involving finer articulatory distinctions, were recovered far less reliably, again in a way that tracked the underlying motor representation's own geometry rather than appearing as uniform noise across all phonemes [12]. In both cases, the same decoding pipeline, applied to the same participant's cortex, produced a legibility boundary with clear internal structure: some distinctions well inside it, others sitting close enough to its edge that performance degraded sharply and unevenly rather than falling off as a flat function of some single difficulty measure.

This is not yet as clean a demonstration as the manifold-versus-outside-manifold result from Chapter 2, where the asymmetry was established by deliberately manipulating the direction of a required transformation and observing a stark difference in outcome. What is available here is closer to a naturally occurring, uneven boundary, shaped by which distinctions the underlying motor representation makes cleanly separable and which it represents in overlapping, harder-to-disentangle form. The chapter records this as a genuine but weaker echo of Chapter 2's asymmetry rather than as an identical result, and leaves open, deliberately, whether a more targeted experiment, analogous to the manifold manipulation, would produce the same kind of stark, direction-dependent legibility boundary that reachability displayed. That is a question for the ledger, not for this chapter to settle by assertion.

3.5 What Has Actually Been Recovered

The preceding sections were preparation for the question this chapter is ultimately organized around. When a decoder correctly predicts an intended movement, a spoken sentence, or the category of an image a sleeping subject is dreaming, what exactly has been established.

A result that makes the question unavoidable comes from work decoding the content of visual imagery during sleep from fMRI activity recorded just before subjects were woken and asked to report what they had been experiencing. A classifier trained on activity patterns associated with specific visual object categories during a separate waking task was able to predict, above chance and with considerable reliability, the general category of content subjects reported dreaming about [13]. The result was widely and not unreasonably described in public discussion as a step toward reading dreams. What the experiment actually demonstrated was narrower and more precise: a trained mapping from a specific set of activity patterns to a specific set of category labels, validated under specific experimental conditions, predicted a specific reported variable above chance. It did not establish that the classifier had reconstructed the dream as subjectively experienced, that the categories used by the classifier correspond to whatever the dreaming brain's own representational format actually consists of, or that every distinction present in the recorded activity had become available to the classifier, since the classifier was trained only to detect the specific categories the experimenters had selected in advance.

This gives the chapter its sharpest available distinction, and it is one that recurs, unacknowledged, throughout public discussion of brain decoding generally. Prediction is evidence of a recoverable relation between a measured signal and a target variable, under the conditions the relation was established and tested. Legibility, as this chapter has used the term, is the broader claim that a decoder can recover which of the substrate's own distinctions currently obtain, not merely which of an experimenter's pre-specified labels currently applies. Reconstruction is stronger still, and closer to the word's use in Chapter 1: a claim that the original state, or something functionally equivalent to it, has been regenerated, independently of the process that produced it the first time, rather than merely predicted from it. A successful decoder of the kind reviewed in this chapter routinely establishes the first. It does not automatically establish the second, and only in narrow, carefully constructed cases, image reconstruction pipelines that regenerate a picture resembling what a subject viewed being the clearest existing examples, does it approach the third, and even then the regenerated image is a reconstruction of what the decoder's own trained mapping implies about the signal, not an unmediated readout of the brain's internal representational format.

Collapsing these three into a single word, reading, is precisely the move this chapter has been built to resist, and it is also, not incidentally, the same collapse Chapter 2 warned against from the learning side, where a failure to decode was shown not to imply a failure to learn. Here the danger runs the other way: a success at prediction should not be allowed to imply success at legibility in the fuller sense, let alone at reconstruction, and the gap

between these three is not a matter of degree on a single scale but a matter of what kind of claim each one actually licenses.

The distinction argued for here is not merely this book's own interpretive imposition on the decoding literature. Contemporary practitioners inside NeuroAI have begun stating a closely related worry from within the field itself, cautioning that a model which accurately predicts neural activity while providing no mechanistic insight into how that activity is produced offers limited scientific value [14]. That caution, arriving independently of this book's framework, is corroboration rather than authority: it does not establish the prediction, legibility, and reconstruction separation drawn above, but it shows working neuroscientists converging on the same fault line from the opposite direction, worried about exactly the collapse this chapter has spent its length resisting. The concern has, if anything, intensified rather than eased as decoding performance has improved. Foundation models now trained to predict neural population responses to entirely novel stimulus categories achieve striking predictive accuracy on data the model never saw during training [15], and the temptation to read this as the model having recovered something like the underlying representational code is at least as strong as it was for dream-category decoding a decade earlier, if not stronger, precisely because the scale and fluency of the prediction has grown.

3.6 Ledger

Chapter 2 recorded reachability as direction-dependent: from a given state, some transitions are readily traversable and others are not, in a way not reducible to a single admissibility ranking of destinations. Chapter 3 records an independent asymmetry on the legibility side. Legibility is relational rather than intrinsic, depending jointly on the state, the decoder, and the level of description at which the decoder operates, as shown by manifold-level transfer succeeding where single-neuron-level transfer does not, using recordings from the identical tissue at the identical moment. Legibility also has its own temporal dynamics, independent of the stability of the underlying state: representational drift produces states that remain stable and behaviorally load-bearing while becoming illegible to a fixed decoder purely as a function of elapsed time, and recovering legibility in that case required actively reconstructing a relation between two changed geometries rather than reading a fixed object more carefully. A weaker, naturally occurring asymmetry in which distinctions are more or less legible within a single decoding pipeline was also observed, tracking the underlying representation's own geometry, though this chapter did not attempt the more targeted manipulation that would establish it as cleanly as Chapter 2 established reachability's asymmetry. Finally, prediction, legibility, and reconstruction were shown to be three distinct claims with three distinct evidentiary requirements, routinely collapsed into a single word in ordinary usage, with successful decoding results in the existing literature reliably establishing the first, only sometimes approaching the second, and rarely establishing the third in any strong sense.

These entries are recorded without resolving whether they indicate the same underly-

ing structure as Chapter 2's asymmetry or a structurally different one. Chapter 11 inherits both.

Chapter 4

Neuro-for-AI

4.1 Three Claims Wearing One Adjective

The phrase brain-inspired covers so much ground in machine learning writing that it has effectively stopped discriminating between claims of very different strength. This chapter needs a version of the phrase narrow enough to be useful to the book's argument, and the narrowing has to happen before any specific mechanism is examined, or every example will be judged by a different implicit standard.

The direction this chapter pursues is, by the field's own recent account, the underdeveloped one: contemporary NeuroAI is described as having entered an "AI-exploitation phase," in which the traffic of ideas runs overwhelmingly from artificial intelligence toward neuroscience rather than the reverse [14]. That asymmetry is the disciplinary backdrop this chapter responds to, and it sharpens rather than answers the question the field leaves open: if renewed neuro-for-AI work is worth pursuing, what would count as genuinely transferring something from neuroscience, as opposed to merely borrowing its vocabulary.

Three claims are routinely compressed into the single adjective. An artificial mechanism can be historically inspired by neuroscience while doing nothing structurally biological, borrowing a name, a diagram, or a loose analogy without importing any constraint that changes how the artificial system actually learns. An artificial mechanism can reproduce a biological computational function using different underlying machinery, arriving at a comparable input-output capability through a route the brain does not obviously take. Or an artificial mechanism can import a constraint that changes which learning trajectories are available to the artificial system, altering not just what the system can eventually compute but which paths through its own parameter space are traversable at all.

Only the third claim is directly interesting for this book, because only the third claim is a claim about admissibility rather than about capability, function, or vocabulary. The first two claims still matter to the history and practice of machine learning, and neither should be dismissed as worthless. They should simply not be allowed to masquerade as evidence for this book's argument, which is not that artificial systems can be made to resemble brains, but that a constraint operating on one substrate can, under specific and testable conditions,

be translated to another substrate in a way that measurably reshapes what that second substrate can and cannot reach. The governing question for this chapter is therefore not the informal one usually asked in surveys of biologically inspired learning. It is narrower and more demanding: what would it mean to transfer an admissibility structure from one substrate to another.

4.2 Credit Assignment as a Test Case

Credit assignment, the problem of determining how much each part of a learning system contributed to an error and adjusting it accordingly, is an unusually good test case for this question, because backpropagation and biological synaptic plasticity confront a recognizably related problem while operating under constraints that appear, on the surface, almost entirely incompatible. Standard backpropagation requires that each synapse have access to a precise error signal computed using the exact transpose of the forward weight matrix used elsewhere in the network, a requirement with no known anatomical correlate, since biological synapses have no evident mechanism for accessing a symmetric copy of forward connection strengths at a distant layer [16]. This mismatch, sometimes called the weight transport problem, has motivated a substantial body of work proposing alternative credit assignment mechanisms constrained to operate under conditions closer to what a biological synapse could plausibly satisfy, using only locally available signals, biologically bounded timing, or bounded forms of feedback.

The credit assignment literature is exactly the right size of test case for this chapter because it contains both kinds of example the argument needs: cases where a biologically motivated constraint, once tested, turns out to change almost nothing about the artificial system's reachable solutions, and cases where it changes a great deal. Distinguishing these requires an explicit criterion rather than an appeal to how biologically plausible a mechanism sounds when described in prose.

4.3 An Ablation Criterion

The criterion this chapter uses is the one demanded of the whole book's admissibility claims generally, applied here to a specific comparison. If removing a supposedly imported biological constraint leaves essentially the same learning trajectories, failure modes, and accessible solutions as keeping it, the admissibility interpretation has lost support, and the mechanism belongs in the first or second category above rather than the third. If introducing the constraint systematically makes some trajectories easier to reach, others harder or impossible, and changes how the system recovers from perturbation, that is evidence that something stronger than inspiration or functional reproduction has occurred.

Feedback alignment supplies a clean case on the weaker side. In feedback alignment, the backward error signal is computed using a fixed, random matrix instead of the exact transpose of the forward weights, directly relaxing the weight transport requirement that

motivated the whole line of biologically motivated research in the first place. The empirical result is that networks trained this way learn nearly as well as networks trained with exact backpropagation on a range of standard tasks, with the random feedback weights becoming effectively aligned with the forward weights over the course of training [17]. The result is scientifically important, and it genuinely shows that exact weight symmetry is not required for effective learning. But by the criterion this chapter is using, it is closer to a negative result for the admissibility claim than a positive one: relaxing the constraint that motivated the mechanism did not produce a network with substantially different reachable trajectories, different failure modes, or different accessible solutions from ordinary backpropagation. The biologically motivated relaxation, in other words, turned out not to be doing much admissibility work at all, since the unconstrained and constrained versions occupy essentially the same functional territory.

Networks trained under Dale's law supply a case on the stronger side. Dale's law states that a biological neuron's outgoing connections are overwhelmingly either excitatory or inhibitory, not a mixture of both, a constraint with no analog in an ordinary artificial network where a single unit's outgoing weights can carry either sign without restriction. When recurrent networks trained on cognitive tasks are constrained so that units are assigned a fixed excitatory or inhibitory identity, matching the biological constraint, the resulting networks reliably require more units, different connectivity patterns, and sometimes qualitatively different dynamical solutions to achieve comparable task performance relative to unconstrained networks of the same nominal size, and certain solution types readily found by unconstrained networks become substantially harder to reach or are not found at all under the sign constraint [18]. Here, removing or imposing the constraint visibly reshapes the space of reachable solutions rather than leaving it essentially intact. By the chapter's criterion, this is a case of genuine constraint transfer: the biological restriction is doing real admissibility work on the artificial substrate, not merely lending it a biologically flavored vocabulary.

4.4 Constraint-Preserving Translation

The Dale's law case succeeds without requiring anyone to claim that an artificial recurrent network is implemented the way a biological circuit is implemented. The units are still simple weighted sums passed through a nonlinearity, nothing like an actual neuron's dendritic integration, and the learning rule used to train them, typically some form of gradient descent, has no direct biological analog either. What transferred was not an implementation but a constraint, and the constraint reshaped the artificial substrate's admissible region despite the two substrates sharing almost nothing at the implementation level.

This motivates a more careful formulation than transplant. A biological mechanism and an artificial mechanism need not share implementation for a relevant constraint to survive translation between them; there may instead exist some constraint under which the two substrates' transformation spaces acquire genuinely comparable structure, even while

everything else about how each substrate carries out its computations remains completely different. Predictive coding networks illustrate this directly. A predictive coding architecture can be shown to approximate the credit assignment computed by ordinary backpropagation while using only local, Hebbian-style synaptic updates and no global error signal transmitted across layers, achieving a mathematically demonstrable relationship to backpropagation's gradient under specific conditions despite operating through an entirely different mechanism, iterative settling toward a local energy minimum rather than a single explicit backward pass [19]. Equilibrium propagation achieves a related result through a different route again, computing something closely related to a backpropagated gradient using a single type of local computation applied in two phases of a network settling to equilibrium, with no separate forward and backward computational pathway of the kind ordinary backpropagation requires [20]. In both cases, the constraint that survives translation is locality itself, no unit's update may depend on information not available to it through its own local connections, and the surviving constraint measurably reshapes what these networks can be built to do and how they fail, even though neither network is built from anything resembling an actual cortical microcircuit.

This is the working sense the chapter proposes for constraint-preserving translation: not that a biological structure is copied into an artificial one, but that a constraint identified in the biological substrate, locality of computation, sign-restricted connectivity, bounded access to global signals, is deliberately imposed on the artificial substrate, and its effect on that substrate's own reachable trajectories is then tested rather than assumed. The scientific content of a neuro-for-AI proposal, on this account, is entirely carried by which constraints survive this kind of translation and which turn out, once tested, to leave the artificial system's admissible region essentially unchanged.

4.5 The Reversal: Constraint as Productive

Chapters 2 and 3 treated biological constraint primarily as a limit, a boundary marking what a nervous system could not readily learn or what an observer could not readily recover. This chapter's strongest conclusion runs in the opposite direction, and it is worth stating plainly because it changes how every constraint discussed so far in the book should be read. An admissibility boundary can be productive. Restricting which transitions a learning system is permitted to make can improve, rather than diminish, its effective capability, by eliminating enormous regions of otherwise reachable but useless search.

Convolutional architectures make the point with unusual force because their biological motivation is well documented and their computational advantage is uncontroversial. The receptive field structure of simple and complex cells in visual cortex, the same experimental tradition that supplied Chapter 2's account of ocular dominance plasticity, motivated an architecture in which units are restricted to local receptive fields and weights are shared across spatial locations, a severe restriction relative to a fully connected network in which any unit may connect to any input with an independently learned weight [21, 22]. This

restriction removes almost the entire space of representations a fully connected network could in principle reach, since translation-variant, spatially unstructured filters become inadmissible by construction. The restricted network does not merely fail to lose much by this exclusion. It learns image recognition tasks far more effectively than an unconstrained network of comparable size, because the excluded region was overwhelmingly composed of representations that were reachable in principle but never useful in practice, and removing it lets the remaining search concentrate entirely on the region worth exploring. The admissibility boundary imposed by the constraint is not a cost paid for biological plausibility. It is the mechanism by which the system becomes more capable, not less.

Read against Chapters 2 and 3, this reframes what a constraint is for. A brain's inability to learn certain distinctions outside a developmental window, examined in Chapter 2 as a limitation, and an observer's inability to decode certain distinctions without the right alignment, examined in Chapter 3 as a limitation, are not simply costs the nervous system or the observer pays. In at least some cases, the same restriction that produces those limitations is what makes the remaining, narrower space of learnable or legible distinctions tractable to search or to read at all. Neuro-for-AI, understood this way, becomes an attempt, sometimes deliberate, as in the convolutional and Dale's law cases, and sometimes closer to an accidental byproduct of pursuing biological plausibility for its own sake, to discover which of a brain's restrictions are worth deliberately inheriting because they are productive rather than merely limiting.

4.6 Toward the Molecular Coda

Part I closes on a single structural observation that the molecular part of this book is built to extend. Throughout Chapters 2 through 4, violating a constraint has meant producing a representation the system in question cannot readily learn, cannot readily read out, or cannot readily search toward without wasting most of its effort on useless trajectories. The representation remains, in an abstract sense, describable. It simply becomes unreachable, illegible, or intractable to find, relative to the substrate and observer under discussion. The coda that follows this chapter asks what happens to the same three-part structure, possible, admissible, reachable, when the substrate stops being a matter of learning trajectories through a parameter space and becomes a matter of literal physical trajectories through a molecular configuration space. The transition can be stated almost bluntly. In Part I, violating a constraint gives a system a representation it cannot readily learn. In Part II, violating one may give no molecule at all.

4.7 Ledger

Chapter 4 contributes a new kind of entry to the ledger, one about how constraints relate to each other rather than about a single admissibility relation observed directly. A constraint transferred from one substrate to another can leave reachable trajectories, failure modes,

and accessible solutions essentially unchanged, as in feedback alignment's relaxation of weight transport, in which case the transfer should be classified as inspiration or functional reproduction rather than admissibility transfer. A constraint transferred from one substrate to another can instead measurably reshape the receiving substrate's reachable region, as in Dale's law's effect on recurrent network solutions, in which case the transfer counts as genuine constraint-preserving translation even where implementation is entirely dissimilar, as further shown by predictive coding and equilibrium propagation achieving comparable functional relationships to backpropagation through mechanisms sharing no implementation detail with it beyond the transferred constraint of locality. Finally, and against the pattern of Chapters 2 and 3, a constraint that shrinks a substrate's admissible region is not thereby shown to have diminished that substrate's effective capability, as convolutional weight-sharing demonstrates directly: an admissibility boundary can be the mechanism of increased capability rather than merely its limit, and this relation should be treated as a live possibility, not an exception, whenever a boundary is examined in the chapters that follow.

Part II

Molecular Substrates

Chapter 5

Conceivable, Synthesizable, Manufacturable

5.1 The Bluntness of the Second Substrate

Part I closed on a deliberately blunt transition. In the substrate of learning and decoding, violating a constraint produces a representation a system cannot readily learn, or cannot readily read out, or cannot readily search toward without wasting effort on useless trajectories. The representation remains describable throughout; it simply becomes unreachable, illegible, or intractable to find. Molecular substrates remove that softness. Violating certain constraints here does not yield a hard-to-learn representation. It yields no molecule at all, or a molecule that exists for a fraction of a second before decomposing, or a molecule that exists indefinitely on a laboratory bench but can never be produced in a quantity larger than a few milligrams. Chapter 5 exists to show that the possible, admissible, and reachable distinctions introduced in Chapter 1, tested so far only against cognitive and computational substrates, apply with at least as much force, and considerably less ambiguity, to chemistry, where the boundary between what a theorist can draw and what a laboratory can produce is measured and published rather than inferred from behavior.

5.2 Possible Is Almost Unimaginably Larger Than Admissible

The starting fact any account of molecular admissibility has to reckon with is scale. Enumerations of small organic molecules that satisfy nothing more than basic valence rules, atom connectivity constraints, and chemical stability heuristics, without any requirement that a synthesis route exist, produce numbers in the hundreds of billions even when restricted to molecules of seventeen heavy atoms or fewer, the well known GDB-17 database being one systematic example, cataloguing on the order of 166 billion such structures [23]. The overwhelming majority of these molecules have never been synthesized and, on any realistic accounting of global synthetic chemistry capacity, never will be. This is possibility

in the sense Chapter 1 used the term: consistent with the base laws governing the substrate, in this case valence and elementary structural stability, and nothing more. It says nothing about whether any known or discoverable chemical transformation could actually produce a given member of that set from available starting materials, which is precisely the gap this chapter needs to make visible rather than assume away.

5.3 Admissible but Unreachable: The Retrosynthesis Gap

A molecule can be admissible, in the sense that its structure is not merely valence-legal but consistent with known classes of chemical reactivity, meaning some sequence of established reaction types could plausibly connect it to available starting materials, while remaining unreachable in practice because no such sequence has been identified, verified, or in some cases even searched for. This gap is the entire justification for the field of computer-assisted retrosynthesis, in which algorithms search backward from a target molecule through a graph of known reaction transformations toward purchasable starting materials. One influential system combined a neural network trained on millions of published reactions with a Monte Carlo tree search to explore this backward space, and was able to propose synthesis routes for target molecules that professional chemists, in a blinded evaluation, could not reliably distinguish from routes taken from the published literature [24]. The existence of a system built specifically to search for a path through admissible-but-unfound territory is itself evidence that admissibility and reachability are separate achievements in chemistry: a molecule's reactivity class does not hand a chemist the sequence of steps required to reach it, any more than a neural network's architecture, in Chapter 2's terms, hands it the training trajectory required to reach a particular admissible configuration of weights.

A related and more quantitative demonstration of the same gap comes from synthetic accessibility scoring, a family of methods that attempt to estimate, from a molecule's structure alone, how difficult it would likely be to synthesize, based on the frequency with which its substructural fragments appear in known, previously synthesized compounds and on measures of the molecule's overall structural complexity [25]. That such a score is useful at all, and that it correlates reasonably well with expert chemists' independent difficulty judgments, confirms that reachability varies continuously and substantially even within the admissible region, and that this variation is estimable in advance of attempting synthesis, exactly as Chapter 2's manifold geometry allowed some directions of learning to be predicted as easier than others before training was attempted.

5.4 Reachable but Not Manufacturable: The Taxol Case

The clearest available demonstration that reachability itself fractures into further, more demanding distinctions comes from the history of taxol, a complex diterpenoid with potent anticancer activity, isolated originally from the bark of the Pacific yew. The first full laboratory synthesis of taxol from simple starting materials was achieved after intense effort

by multiple research groups, requiring, in one completed route, dozens of sequential steps with declining yield compounding at each stage, so that even a successful execution of the entire published sequence delivers only a minute fraction of the starting material's mass as final product [26]. The synthesis is a genuine scientific achievement and settles, unambiguously, that taxol is reachable: an explicit, verified, step-by-step path exists from ordinary starting materials to the target molecule. It settled nothing about whether taxol could be manufactured in the quantities a cancer treatment program actually requires. A total synthesis whose overall yield across dozens of steps is a small fraction of a percent is reachable in the sense that matters for a research demonstration and entirely inadmissible as an industrial production route, since no plausible scale of investment in reagents, solvent, and purification capacity converts a milligram-scale laboratory achievement into a metric-ton pharmaceutical supply chain.

The route that actually supplies the pharmaceutical industry with taxol illustrates why this distinction is not merely a matter of degree. Rather than starting from petrochemical building blocks and constructing the entire diterpenoid skeleton from scratch, the industrial route begins from a related natural compound, 10-deacetylbaccatin III, extractable in useful quantity from the needles of a renewable, non-endangered yew species, and completes the synthesis with a comparatively short sequence of high-yielding steps [27]. The final target molecule is identical in both cases. What differs is the starting state and the set of available operations relative to that starting state, which is to say the index against which admissibility and reachability are being evaluated. Total synthesis from petrochemical precursors and semisynthesis from a natural precursor are two different admissible regions, indexed to two different starting materials, converging on the same target through routes with radically different yields, step counts, and industrial viability. This is Chapter 1's claim that admissibility is indexed, not an abstract property of a target state, demonstrated with a real molecule whose industrial supply chain depended entirely on discovering the more favorably indexed of the two available approaches.

5.5 The Contested Frontier: Mechanosynthesis and Molecular Manufacturing

Every case examined so far concerns molecules whose reachability, once established, was not seriously disputed by domain experts, even when manufacturability at scale remained a separate and much harder problem. The molecular manufacturing literature contains a case where the admissibility question itself, not merely the manufacturability question, remains genuinely contested among qualified chemists, and the book is better served by presenting that contest honestly than by resolving it in either direction.

Proposals for molecular manufacturing built around mechanically guided, positionally controlled synthesis, in which a mechanism analogous to a robotic arm would place individual reactive groups at precise locations to build complex structures atom by atom, generated a sustained and pointed exchange between advocates who argued the approach was

a natural, if extremely demanding, extension of known chemistry, and critics who argued that positional control of the kind proposed cannot substitute for the close, many-body electronic interactions that actually govern whether a given bond-forming step proceeds, meaning the core mechanism was not merely difficult to engineer but chemically inadmissible as described [28]. Neither side's position, as stated in that debate, has since been settled by a demonstrated working assembler of the proposed kind, and neither has been definitively closed off by a proof that no version of the proposal could work. What is available, more than two decades later, is a genuine open admissibility question rather than a merely open manufacturability question of the taxol kind, and the distinction matters for how this chapter's framework is applied honestly. The taxol case involved uncertainty about which of two known-admissible routes would prove more manufacturable. The mechanosynthesis case involves uncertainty about whether the proposed route is admissible at all, at the level of chemistry rather than engineering, and treating the two kinds of uncertainty identically would blur exactly the distinction this book depends on.

5.6 Ledger

Part II's opening chapter contributes several relations, some sharpening Part I's findings and one that resists easy classification.

The gap between possible and admissible, measured only impressionistically in Part I, is directly quantifiable in the molecular substrate: systematic enumeration shows possible structures outnumbering ever-synthesized structures by many orders of magnitude, confirming that possibility, understood as mere consistency with base structural laws, is a vastly weaker condition than admissibility relative to any actual chemistry.

The gap between admissible and reachable, established more informally through the neural network case in Chapter 2, is directly instrumented in the molecular substrate: retrosynthesis-planning systems and synthetic-accessibility scores exist specifically because that gap is wide enough, and its internal variation regular enough, to be worth predicting and searching computationally, echoing Chapter 2's finding that reachability varies continuously and predictably even within a fixed admissible region.

Reachability itself was shown to fracture further, into a reachable-as-demonstrated category and a reachable-at-useful-yield category that do not coincide, with the taxol total synthesis satisfying the first without approaching the second. This sharpens, rather than merely repeats, Chapter 1's diamond-manufacturing counterexample, by showing the fracture occurring within a single successfully completed synthesis rather than as a gap between the possible and the admissible.

Admissibility's indexed character, proposed abstractly in Chapter 1, receives its clearest demonstration yet in the taxol semisynthesis route: an identical target molecule was inadmissible at industrial scale relative to one starting material and set of operations, and comfortably admissible relative to a different starting material and a shorter set of operations, with no change whatsoever to the target itself.

Finally, the mechanosynthesis case is recorded not as a resolved relation but as an open question about which axis a persistent failure belongs to. The book takes no position on whether molecular manufacturing of the proposed kind is admissible, reachable-but-undiscovered, or genuinely inadmissible at the level of chemistry, and flags this explicitly as a case Chapter 11 should treat as unresolved rather than silently assume has been settled by either side of the original debate.

Chapter 6

Yield and Recursion

6.1 Beyond the Single Event

Chapter 5 established that a transformation can be performed: taxol can be synthesized, by more than one route, and the routes differ enormously in how efficiently they convert starting material into product. That chapter's task ended at the demonstration that a pathway exists and can be executed. This chapter's task begins at the harder question the demonstration leaves open. What has to be true for a transformation that succeeded once to become part of a persistent productive system, one that can be run again, scaled, depended upon by other processes, and, in the strongest cases, extended or repaired using resources the system itself produces. Chapter 5 measured reachability at the level of a single event. Chapter 6 measures it at the level of a process that has to survive repetition, and the two turn out to fail in different ways.

6.2 Yield as the Hinge

A pathway that succeeds once and a pathway that converts inputs to product reliably, batch after batch, at high fractional yield, are both technically reachable in Chapter 5's sense. They occupy radically different positions in any account of admissibility that cares about more than a single demonstration. Yield introduces the question of repetition directly: can essentially the same traversal survive another attempt, and another, without accumulating error, waste, purification burden, or resource demand to the point where the process becomes unaffordable well before it becomes physically impossible. This is the pivot from what Chapter 1 called reachability, a property of a single path from a single starting state, to something closer to process viability, a property of whether that path can be traversed repeatedly under real operating conditions without degrading or collapsing.

The distinction matters because the two properties can diverge sharply, as Chapter 5's taxol case already showed in miniature: a completed total synthesis is reachable in the strongest sense, an explicit, verified route exists, while its yield, compounding a small loss

at each of dozens of sequential steps, makes repeated execution at any economically meaningful scale unaffordable regardless of how many times a laboratory is willing to attempt it. Yield is therefore not one manufacturing constraint among many. It is the specific property that converts a one-time reachability result into a question about whether a process can be relied upon at all.

6.3 From Repeatability to Dependency

A process that clears the yield hurdle produces something further worth asking about: what does running that process depend on, and how much of what it depends on can be supplied, maintained, or replaced without leaving the system's own admissible operations. A single high-yield chemical step depends on reagents, which depend on suppliers, which depend on their own upstream processes, each with its own yield and reliability profile. As soon as a process is examined not as an isolated event but as a node embedded in a larger dependency structure, the relevant question stops being can this step be performed and becomes what has to remain intact, upstream and downstream, for this step to keep being performable.

This reframing is what makes the microfactory and distributed-industrial-cloud material genuinely load-bearing for this chapter rather than a speculative aside. Read as a proposal about which particular objects a distributed network of small producers might manufacture, it is one design among many. Read as a hypothesis about where manufacturing admissibility actually resides, it becomes a sharper and more general question. Is productive capability located in a single machine, in a factory, in the network of tooling that supplies and services that factory, in the supply chain of raw materials feeding it, in the specification and metrology infrastructure that certifies its output meets requirements, or in the trained workforce that operates and repairs it. Asking what would have to be reconstructed after destroying one factory, rather than what the factory alone currently produces, is what forces the answer to expand beyond the factory's walls, and it is the question this chapter treats as primary.

6.4 Systemic Recursion: The Semiconductor Case

Semiconductor manufacturing supplies an unusually clear historical case because it makes this expanded dependency structure visible rather than merely asserted. Producing a modern integrated circuit requires photolithography equipment capable of patterning features at a scale of a few nanometers, and that equipment, extreme ultraviolet lithography systems in particular, depends in turn on an extraordinarily narrow, globally concentrated set of upstream capacities: specialized optics manufactured by a small number of firms, high-power plasma light sources produced by an even smaller number, ultra-high-vacuum engineering, precision metrology, and control software, much of which is itself designed and verified using semiconductor devices produced by the very industry the equipment serves [29]. No

single company, machine, or facility in this chain constitutes a self-contained productive unit. A lithography manufacturer does not itself mine the raw materials, does not itself fabricate every precision component, and does not itself operate the fabrication plants that use the finished equipment. Yet the industrial ecology as a whole participates in reproducing and extending its own productive capacity, generation after generation of chip technology enabling the design of better lithography tools, which in turn enable more advanced chips, in a loop distributed across dozens of specialized firms and multiple countries rather than concentrated in any single reproducing unit.

This is the sense in which recursion in this chapter should be understood, and it is deliberately weaker, and more industrially realistic, than literal self-replication. A recursively productive manufacturing system is one that produces outputs which expand, repair, reproduce, or reconfigure some portion of the capacity by which further outputs are produced. Semiconductor fabrication satisfies this description at the level of an entire industrial ecology without any individual machine in that ecology reproducing itself. The recursion is systemic, distributed across a dependency network, rather than local to a single self-copying unit, and treating these as the same phenomenon would understate how much of the world's actual recursive productive capacity looks like the semiconductor case rather than the stricter case examined next.

6.5 The Stricter Case, and Why This Chapter Does Not Need It

The strict version of productive recursion, a single system capable of reproducing itself in its entirety from raw or generally available materials, has a long formal history, beginning with von Neumann's theoretical treatment of self-reproducing automata, in which he specified the logical requirements a system would need to satisfy to construct a complete copy of itself, including a copy of its own construction instructions [30]. That formal result establishes an important reference point: self-replication in the strict sense is not merely difficult but was shown, in principle, to be logically achievable given the right architecture, decades before any physical system approached it.

The closest sustained physical attempt is a distributed hobbyist and academic project to build a three-dimensional printer capable of manufacturing a substantial fraction of its own replacement parts, with published accounts explicitly tracking what proportion of a given machine's components could be produced by an existing machine of the same design, as distinct from parts that still had to be purchased or fabricated by other means [31]. This project is valuable to the chapter not because it succeeded at full self-replication, which it did not, but because it is one of the few manufacturing efforts to have explicitly measured and published a fractional self-production capacity rather than simply claiming success or failure at an all-or-nothing standard. That measurement practice, a system tracking what percentage of itself it can currently produce, is the seed of a more general and more useful concept than strict self-replication, and the chapter adopts it directly rather than chasing the stronger, largely unrealized standard von Neumann's formalism describes. Molecular

assemblers of the kind debated in Chapter 5 would, if admissible, represent an attempt at the strict case at a much smaller scale; this chapter deliberately stays with the weaker, industrially demonstrated case, because conflating the two would import Chapter 5's unresolved dispute into a chapter that does not need it.

6.6 The Recursive Closure Ratio

The measurement practice suggested by the self-replicating printer project generalizes naturally to the distributed-industrial-cloud proposal, and gives it a genuinely testable form rather than a qualitative aspiration. Rather than asking only whether a distributed network of microfactories can manufacture a broad catalogue of objects, a sharper question can be asked of any manufacturing system, existing or proposed: what fraction of the infrastructure required to maintain and extend the network can the network itself regenerate, using its own outputs and its own available operations, without drawing on capacity outside the system under examination.

This fraction, a recursive closure ratio, is not a claim of self-sufficiency and is not proposed here as an already-solved measurement problem. It is a way of stating precisely what would have to be measured to compare two manufacturing systems that might otherwise look similarly capable. A system could exhibit high output diversity, manufacturing many different kinds of objects, while having low recursive closure, depending heavily on externally supplied tooling, precision components, or raw materials it cannot itself produce, as most current distributed manufacturing proposals do relative to the semiconductor-grade components they typically still need to purchase. A different system could have narrow output diversity, manufacturing only a small catalogue of objects, while achieving surprisingly high recursive closure, because that narrow catalogue happens to include most of what the system needs to sustain and repair itself. These are not the same accomplishment, and treating high output diversity as evidence of high recursive closure, or the reverse, would conflate two properties this chapter has now shown to vary independently.

6.7 Output Reproducibility Is Not Productive-Capacity Reproducibility

The distinction the recursive closure ratio makes measurable has a sharper, qualitative form worth stating directly for the ledger, because it is the concept this chapter is best positioned to hand forward. A civilization, or a firm, or a factory, might retain full knowledge of how to make an object while having lost the capacity to reproduce the machinery, tooling, and specialized dependency structure that made producing it possible. Conversely, a system might lose the specific process that once produced a given object while retaining enough general productive infrastructure, precision manufacturing capacity, materials science knowledge, and skilled workforce, to eventually rediscover a working process for that object or something functionally equivalent to it. Output reproducibility, whether the

specific artifact can be made again by retracing the original steps, and productive-capacity reproducibility, whether the broader system that made those steps possible can itself be regenerated or extended, are different achievements, measured differently, and a system can score high on one while scoring low on the other.

This is not a new claim so much as a generalization of one already established. The Saturn V's F-1 engine, examined in Chapter 1 as a case of a legible but unreconstructible state, is precisely an instance of output knowledge persisting, complete engineering drawings, published specifications, surviving physical examples, while the productive capacity that generated those outputs, the specific welding techniques, machine-tool calibrations, and embodied judgment of a particular trained workforce operating within a particular factory, was not retained and had to be substantially rebuilt from a different starting point decades later. Chapter 6 gives that earlier case its general name: the engine's documentation was output reproducibility, essentially complete; the factory's capability was productive-capacity reproducibility, substantially lost. Chapter 10 inherits both halves of this distinction at civilizational scale, where the same question, has the object been remembered, or has the capacity to make it been remembered, has to be asked about entire technological lineages rather than a single engine.

6.8 Toward the Energy Coda

Chapter 5 asked whether matter can be brought into a desired configuration at all. Chapter 6 has asked a harder question of the same material: whether the capacity to do so can persist, survive repetition, and in the strongest cases extend or repair itself. Every element of that persistence has a physical cost that has so far been treated only implicitly. A repeated traversal consumes energy. A purification step consumes energy. A failed batch, a replaced machine tool, a shipment moved between two nodes of a supply chain, all consume energy, and none of this chapter's yield, dependency, or recursion concepts can be evaluated without asking what budget they draw against. The coda this chapter hands to Part III is accordingly sharper than a general question about resource cost. It is a direct extension of the recursive closure ratio just introduced: what determines how much traversal a productive system can afford before its admissible manufacturing network collapses. That is an energy question stated in exactly the vocabulary this chapter has built, and it is the question Part III opens with.

6.9 Ledger

Yield is recorded as the property that separates single-event reachability, established in Chapter 5, from process viability, the capacity of a reachable transformation to survive repeated execution without accumulating unaffordable loss. Dependency is recorded as a further, distinct property: a viable process can still fail to be a productive system if the upstream and downstream capacities it depends on are not themselves reliably available,

which is why manufacturing admissibility was shown to reside in a distributed dependency structure, illustrated concretely by semiconductor lithography, rather than in any single machine or facility. Productive recursion was defined weakly, as the expansion, repair, reproduction, or reconfiguration of some portion of a system's own productive capacity by its own outputs, deliberately distinguished from the stricter, largely unrealized standard of complete self-replication formalized by von Neumann and partially approached by self-printing hobbyist systems, because the weaker standard is the one actually satisfied by real industrial ecologies. The recursive closure ratio was introduced as a way of making that weaker standard measurable, and was shown to vary independently of output diversity, meaning neither property can be inferred from the other. Finally, output reproducibility and productive-capacity reproducibility were separated as distinct forms of reconstructive capacity, with Chapter 1's Saturn V case reinterpreted as an instance in which the two diverged sharply, a distinction Chapter 10 is expected to inherit and apply at the scale of entire technological lineages rather than single artifacts.

Part III

Energetic Substrates

Chapter 7

Energy as Region, Not Direction

7.1 From Recursion to Cost

Chapter 6 closed by naming what every element of a recursive manufacturing system quietly depends on. Synthesis, separation, purification, transport, computation, temperature control, pressure maintenance, error correction, and machine replacement all draw against a resource budget, and a recursive closure ratio that looks favorable on paper can still be unaffordable in practice if the energy required to sustain it exceeds what the surrounding system can supply. Give the same recursive system access to cheaper or more abundant usable energy, and some transformations that were previously unaffordable become affordable. This chapter exists to state that observation precisely enough to be useful, and then to show, with as much force as the opposite claim deserves, that it is not the whole story.

The central task of this chapter is to establish a claim already latent in its title. Energy changes the extent of what is admissible without determining the internal structure of that extended region. A terawatt does not supply a missing catalyst. It does not discover a synthesis route. It does not create an absent neural representation, manufacture a missing lithography system, or tell an institution which experiment to perform. More energy changes the set of transformations that can be paid for. It does not specify the transformations themselves. Energy can remove an energetic prohibition without removing a structural prohibition, and the distinction between these two kinds of prohibition, easy to state and easy to blur in practice, is what this chapter spends its length defending.

7.2 The Region Is Not Uniform

The chapter's title risks a metaphor stronger than the evidence supports, and it is worth complicating that metaphor before relying on it. Energy does not expand admissibility uniformly, the way inflating a balloon expands its interior volume in every direction at once. Additional energy typically makes one family of transformations dramatically cheaper while

barely affecting another. Raising temperature can accelerate a desired reaction while simultaneously accelerating an undesired decomposition pathway competing for the same starting material, so that more thermal energy narrows, rather than widens, the useful operating window for a process sensitive to that competition. Greater computational power enlarges the space of searches that can be attempted without making every search strategy within that space equally productive, since some strategies scale efficiently with added computation and others do not. Additional grid capacity built in one location does not create usable energy somewhere else entirely, however abundant the local supply becomes.

This anisotropy is worth naming explicitly because it is not a new phenomenon appearing for the first time in this chapter. Chapter 2 found that reachability within a neural population's admissible manifold depends on direction, some extensions readily trainable, others not, from the identical starting geometry. Chapter 3 found that legibility depends on which decoder and which level of description is applied to an identical underlying state. Chapter 5 found that molecular reachability depends heavily on which starting materials and reaction classes are available, the same target molecule wildly differing in accessibility depending on the route attempted. This chapter can now report a fourth, independent appearance of the same pattern: resource relaxation itself has direction-dependent effects, expanding some regions of a system's admissible space considerably more than others, and sometimes contracting a region even as it expands the space nominally containing it. The pattern is recorded in the ledger rather than named, consistent with the book's stated method of letting each domain damage or extend the provisional ontology without prematurely declaring what final structure the accumulated damage implies.

7.3 A Case Where Energy Clearly Expands the Region: Aluminum

Aluminum is an unusually clean historical case for energy genuinely expanding an admissible region, and it is worth walking through carefully because the easy version of the story, more energy made aluminum available, is subtly wrong in a way that matters for the rest of the chapter. Aluminum is the third most abundant element in the Earth's crust, yet for most of the nineteenth century it remained a scarce, expensive novelty metal, more valuable per unit weight than silver, because no economical process existed for separating it from the tightly bound oxide ore in which it naturally occurs. What changed this was not simply the arrival of abundant electricity. It was the discovery, achieved independently and almost simultaneously by two chemists, of an electrochemical pathway, dissolving aluminum oxide in molten cryolite and passing an electric current through the resulting solution, that made electrolytic reduction of aluminum oxide practically achievable at industrial scale [32]. Electricity alone, applied to aluminum oxide directly, does not efficiently separate the metal; the cryolite solvent step is what makes the electrochemistry tractable at a workable temperature and voltage. The process that resulted, still the basis of essentially all primary aluminum production today, required both an available energy supply and a specific chemical pathway capable of converting that supply into the desired transformation. Electricity

did not make aluminum admissible on its own. A pathway capable of exploiting electricity did, and the pathway had to be discovered, not merely funded.

This gives the chapter a formulation precise enough to reuse throughout the rest of the book. Energy does not substitute for a path. It changes the cost of traversing one that already exists, or that has already been discovered to exist, and no quantity of energy supplied to a process lacking such a pathway converts that process into a productive one.

7.4 A Case Where Energy Conspicuously Fails to Solve the Problem: Desalination

Desalination supplies a contemporary counterpart, useful precisely because it shows energy abundance failing to dissolve a structural constraint rather than succeeding against one. Converting seawater to fresh water is, in the most general sense, an energy-intensive separation problem, and a substantial technical literature has focused on reducing the specific energy consumption of both thermal and membrane-based desalination processes, since energy cost is a major component of the total cost of produced water [33]. But the same literature is equally clear that energy availability alone does not resolve the practical constraints that determine whether a given desalination strategy is viable at a given site: membrane fouling by organic matter and scale-forming minerals degrades performance regardless of how much energy is supplied to drive flow across the membrane, intake infrastructure has to be engineered around local marine ecology and sediment conditions independently of the energy budget, and brine disposal imposes environmental and regulatory constraints that a larger power supply does nothing to relax. Pouring additional energy into an unsuitable membrane, an inadequately designed intake system, or a fouled process train does not erase these constraints. It can, in some configurations, worsen them, by driving more water through a membrane that fouls faster under higher flux. Energetic and structural admissibility interact in this case without becoming identical, and the direction of causation runs only one way: solving the structural constraints can make a process more responsive to added energy, but adding energy does not, by itself, solve the structural constraints.

7.5 Gross Energy Versus Usable Transformation Capacity

The aluminum and desalination cases together motivate a distinction essential to the rest of this book, and indispensable to Chapter 8. A system does not merely need energy to exist somewhere in its vicinity. It needs energy available in an appropriate form, location, timing, concentration, and degree of controllability, coupled to an actual admissible transformation capable of using it. A petroleum deposit, a quantity of solar flux striking a rooftop, a charged battery, a geothermal gradient, a national electrical grid, and a confined fusion plasma can all be described, in a loose accounting sense, as containing or transmitting energy, while making radically different sets of transformations available to a system trying

to draw on them. This is close to what thermodynamics separately formalizes as exergy, the portion of a system's energy that is actually available to perform useful work given the constraints of the surrounding environment, as distinct from the system's total energy content, which includes portions that are thermodynamically present but practically unusable for driving a given process [34].

The theoretically relevant quantity for this book's argument, in other words, is not simply joules. It is something closer to energy coupled to an admissible transformation, a quantity that depends jointly on how much energy is available and on whether a pathway exists that can convert that specific form of energy, at that specific concentration and location, into the specific transformation a system needs. This prevents the book's account of energy from drifting toward the simplest and least useful version of an energy-return argument, in which every problem is treated as fundamentally a matter of insufficient joules waiting to be supplied.

7.6 Efficiency as an Admissibility Technology

The same distinction gives the chapter a precise way of handling efficiency, one that treats it as doing real admissibility work rather than merely conserving a fixed resource. If a process requires one-tenth the energy of a competing process to achieve the same useful transformation, that improvement can enlarge effective admissibility without any increase in gross energy supply whatsoever, by converting a transformation that was previously unaffordable, given the available budget, into one that is repeatedly affordable within the same budget. Efficiency, on this account, is an admissibility technology in exactly the sense Chapter 5's synthesis routes and Chapter 6's yield improvements were admissibility technologies: it does not create new physical law, but it moves a transformation from the inadmissible side of a cost boundary to the admissible side, for a fixed resource endowment. Better insulation, more effective catalysts, heat recovery systems, power electronics, more efficient algorithms, improved motors, and tighter process integration all function this way, enlarging the effective energetic region available to a system just as surely as additional generation capacity would, and in many documented cases considerably more cheaply.

There is a converse to this that the chapter cannot responsibly omit, because omitting it would let efficiency's admissibility-expanding effect silently collapse into a claim about reduced total resource consumption, which is a different and considerably more contested proposition. Greater efficiency can also make entirely new demand patterns admissible that were not previously worth pursuing at the old cost, and a substantial body of evidence associated with the rebound effect documents cases in which improved efficiency is followed by increased total consumption of the resource being economized, partially or in some studied cases more than fully offsetting the efficiency gain [35]. The book does not need an extended treatment of this literature to make its point. It needs only the distinction the literature already supports: efficiency enlarging the admissible region and efficiency reducing total resource consumption are separate claims, empirically dissociable, and this

chapter commits only to the first.

7.7 Attacking the Strongest Version of Energy Determinism

The chapter has so far argued for a moderate position: energy matters, expands admissibility unevenly, and interacts with but does not replace structural constraints. A stronger and more dangerous version of the claim deserves to be confronted directly rather than left to accumulate credibility by default. Suppose someone argues that every admissibility boundary examined so far in this book ultimately reduces to an energy boundary, that given sufficient energy, a system could brute-force its way through any of them: exhaustively search molecular configuration space, run arbitrarily large artificial learning systems, manufacture arbitrary tooling from raw materials, perform an unlimited number of experiments, and reconstruct any lost industrial capability, purely by throwing enough power at the problem.

The answer to this argument should not simply be no. It should show why the brute-force strategy fails on its own terms, independent of any energy limit. Brute-force search of any kind requires, as a precondition for the strategy to make sense at all, a defined search space over which to search, some means of discriminating a successful outcome from an unsuccessful one, a way of retaining and building on results once found rather than losing them, accessible intermediate states through which the search can actually proceed, and transformations capable of reaching the candidates under consideration from wherever the search currently stands. An unlimited energy budget multiplies how many executable possibilities a system can attempt within a defined space. It does not define the space. It cannot search distinctions that have never been represented in the first place, cannot traverse a transition that is physically forbidden regardless of how much energy is applied to attempting it, and cannot infer information that a substrate has genuinely erased rather than merely made expensive to recover.

This is where the chapter can reach backward across the entire book and show the pattern holding at every substrate examined so far. Hubel and Wiesel's developmental window for ocular dominance plasticity does not reopen because an animal receives more calories; the constraint governing when that plasticity is available is not an energy constraint and is not relaxed by one. A decoder's level-of-description problem, established in Chapter 3, is not solved by supplying a recording electrode array with more electrical power; the problem is which relation between decoder and state has been established, not how much current drives the recording apparatus. Chemical space does not become synthesizable because a laboratory's power budget becomes unlimited, since an unlimited budget still cannot search reaction pathways that have not been discovered or represented, as Chapter 5's retrosynthesis gap made explicit. Taxol's original low-yield pathway problem, examined in the same chapter, was never fundamentally an energy shortage; it was a yield and selectivity problem that a different, more favorably indexed pathway solved, not a problem more electricity supplied to the original route would have solved. In every case examined

so far, energy interacts with an independently existing geometry. It does not generate that geometry, and no accumulation of energy alone converts an unrepresented, physically forbidden, or undiscovered transformation into an executable one.

The chapter's title claim can therefore be stated in a stronger form than it opened with. Energy does not define the admissible world. It determines which portions of an already structured world can be repeatedly traversed at acceptable cost.

7.8 Toward Chapter 8

That stronger formulation is precisely what Chapter 8 is positioned to attack, using a case chosen because it resists the easy version of this chapter's own argument. A fusion reaction is physically possible, has been experimentally achieved, and is energetically productive in the sense that the reaction itself releases far more energy than a comparable chemical process. None of that establishes a viable reactor, because the harder question is not whether the reaction can occur but whether the entire coupled set of conditions required to sustain, contain, and productively extract energy from that reaction can be maintained together, continuously, without any one of them collapsing and taking the others down with it. Chapter 7 has asked how much of an already structured world an energy budget allows a system to traverse. Chapter 8 asks a different question of the same structured world: whether a system, having reached a productive state, can remain there.

7.9 Ledger

This chapter records energetic expansion as anisotropic, a fourth independent instance, after Chapters 2, 3, and 5, of a resource or capacity relaxation affecting different directions of a system's admissible region unevenly rather than uniformly, and in some documented cases narrowing a functionally relevant window even while nominally expanding the space containing it. It records energetic prohibition and structural prohibition as distinct kinds of constraint, demonstrated by the aluminum case, where a discovered pathway rather than raw electrical supply converted an inadmissible transformation into an admissible one, and by the desalination case, where energy abundance conspicuously failed to relax structural constraints on fouling and intake design. It records gross energy and energy coupled to an admissible transformation as distinct quantities, the second being the one with genuine explanatory work to do, following the thermodynamic concept of exergy. It records efficiency as an admissibility technology capable of enlarging a system's effective energetic region without any increase in gross supply, while explicitly separating that claim from the empirically distinct and sometimes contrary claim that efficiency reduces total resource consumption, following the documented rebound effect. Finally, it records the brute-force objection to the book's entire argument as answerable in principle, not merely by assertion: unlimited energy multiplies executable possibilities within a search space without defining that space, discriminating successful outcomes within it, retaining results, or creating pas-

sage through transitions the substrate itself forbids, a conclusion the chapter tested against every substrate examined in the book so far rather than asserting in the abstract.

Chapter 8

Viability Under Constraint

8.1 From Reaching to Remaining

Chapter 7 asked how much of an already structured world an energy budget allows a system to traverse. This chapter asks a different question of the same structured world: whether a system, having reached a productive state, can remain there. The distinction sounds slight until a concrete case forces it apart, and fusion energy is used here not because it is the chapter's real subject but because it is an unusually well-instrumented experimental object through which the underlying claim becomes visible. The claim is that energetic admissibility, examined closely enough, is itself a coupled systems property: not a single threshold a system either clears or does not, but the simultaneous, continuous satisfaction of several independently constrained conditions, any one of which can fail without the others failing, and whose joint failure mode looks nothing like a simple insufficiency of energy. Fusion is the case that makes this visible with unusual clarity, because a fusion reaction can be, and has been, achieved as a single physical event without a viable reactor existing anywhere in the world, and the gap between the two is not a gap in energy at all.

8.2 The Lawson Criterion as an Admissibility Threshold

The condition under which a fusion reaction produces more energy than is required to sustain it was formalized in the 1950s as a relationship among plasma density, temperature, and confinement time, now generally known as the Lawson criterion: the product of these three quantities has to exceed a threshold value before the reaction can, even in principle, deliver net energy [36]. This criterion is worth reading in the vocabulary this book has built rather than treating it as a mere engineering specification. It identifies an admissibility threshold, a boundary separating configurations of plasma density, temperature, and confinement time that are consistent with net energy production from configurations that are not, in exactly the sense Chapter 1 used the term for a substrate's own constraint set, here applied to a plasma rather than to a nervous system or a molecular reaction network.

What the Lawson criterion does not specify, and was never intended to specify, is whether a system built to satisfy it can be operated continuously, safely, and repeatedly, using materials and control systems capable of surviving the environment the reaction itself creates. Crossing an admissibility threshold once, as an isolated demonstration, is a different achievement from occupying the admissible region persistently as an operating condition, and the difference between these two achievements is exactly what the rest of this chapter is built to formalize.

8.3 Viability Theory: Reachable Versus Sustainable

A useful formal vocabulary for this difference already exists outside this book's own framework, developed originally for problems in control theory and mathematical economics under the name viability theory. Viability theory distinguishes the set of states a system can reach at all from the narrower set of states, called a viability kernel, from which the system can be kept indefinitely within a specified constraint set using controls available to it, without ever being forced by the system's own dynamics to leave that constraint set regardless of how it is steered [37]. A state can belong to the reachable set without belonging to the viability kernel, if reaching it places the system on a trajectory that, however it is subsequently controlled, eventually forces a violation of some required constraint. This is precisely the formal shape of the distinction fusion research has had to confront empirically, years before most of that research described its own problem in this vocabulary: reaching net energy output once is a reachability result; building a reactor is a claim about the viability kernel, the set of operating conditions from which net energy production can be sustained indefinitely without violating any of the constraints, material, thermal, or otherwise, that the surrounding system also has to satisfy.

8.4 Reached but Not Remained

Two documented cases show this gap directly, and it is worth examining both because they fail in structurally different ways.

In December 2022, an inertial confinement experiment using powerful lasers to compress a small fuel capsule achieved, for the first time, a fusion energy output exceeding the energy delivered by the laser system to the target, a result reported as scientific breakeven [38]. This is a genuine, historically significant crossing of the Lawson-type admissibility threshold. It is also, definitionally, a single event: one capsule, compressed and ignited once, producing net energy for a duration measured in a fraction of a second before the plasma disperses. No claim was made, and none is warranted, that the facility constitutes a power plant, or that the demonstrated configuration could be repeated on a timescale, at a repetition rate, and at a net system-level energy balance, accounting for the energy cost of the laser system itself rather than only the energy delivered to the target, that would make continuous power production viable. The result establishes reachability in the strongest

available sense. It says nothing directly about the viability kernel.

Magnetically confined plasmas in tokamak reactors face the complementary failure mode, one that shows even a system already operating inside its intended admissible region can be violently ejected from it. Tokamak plasmas are subject to disruptions, sudden, large-scale instabilities that can terminate confinement within milliseconds, releasing the plasma's stored thermal and magnetic energy onto the reactor's internal surfaces and structural components in a burst intense enough to cause serious material damage, and survey studies of disruption events at major operating tokamaks document these events as a recurring, only partially predictable feature of operation rather than a rare edge case confined to early or poorly tuned experiments [39]. A tokamak experiencing a disruption was, up until the moment of the disruption, operating inside a configuration that satisfied the Lawson-type admissibility condition. The disruption shows that satisfying that condition at a single moment does not establish that the state is viable in the stronger, dynamical sense viability theory requires: the system's own internal dynamics, under realistic and only partially controllable perturbations, can carry it out of the admissible region even when no external change to the applied energy budget has occurred. This is the coupled-systems failure mode Chapter 7's argument anticipated but could not itself demonstrate, since Chapter 7's cases concerned whether a transformation could be reached or afforded at all, not whether a reached state could be dynamically retained against the system's own tendency to leave it.

8.5 The Coupled Constraint Set

Reaching and momentarily sustaining net energy output, even repeatedly and without disruption, still falls short of a viable reactor, because a working power plant has to jointly satisfy several further constraints that are only loosely coupled to the plasma physics itself, and a design that improves performance against one of these constraints can straightforwardly worsen performance against another.

The reactor's internal structural materials face continuous bombardment by high-energy neutrons produced by the deuterium-tritium reaction, and this neutron flux progressively damages the crystal structure of most candidate materials, degrading their mechanical properties and inducing radioactivity, over a lifetime that current materials cannot sustain for the multi-decade operational periods a commercially viable power plant would require, making materials development, rather than plasma confinement, one of the field's most persistent bottlenecks [40]. A reactor design that increases plasma performance by operating at higher power density typically increases the neutron flux striking the surrounding structure, directly shortening the material lifetime the design can sustain, which is exactly the pattern this book has repeatedly found: improving one admissibility margin can tighten another, rather than the two varying independently.

A second, largely separate coupled constraint concerns fuel supply. The deuterium-tritium reaction favored by most reactor designs because of its comparatively lower ignition

threshold requires tritium, an isotope not available in nature in anything close to the quantity a fleet of power plants would consume, so a viable reactor has to breed its own tritium fuel in situ, typically by surrounding the plasma with a lithium-containing blanket that absorbs escaping neutrons and produces new tritium as a byproduct, and detailed physics and engineering analyses of this requirement show that achieving tritium self-sufficiency, breeding at least as much tritium as the reactor consumes, once realistic neutron losses, blanket coverage limits, and processing efficiencies are accounted for, is a demanding constraint in its own right, only loosely related to whether the core plasma physics satisfies the Lawson criterion [41]. A reactor could in principle satisfy the Lawson criterion, avoid disruption, and use materials capable of surviving its operational lifetime, while still failing as a viable power plant because its blanket geometry cannot breed enough tritium to fuel its own continued operation.

Viability, at the level of an actual reactor, is therefore not membership in a single admissible region defined by plasma physics alone. It is membership in the intersection of several separately constrained regions, plasma confinement, structural material lifetime, disruption avoidance, and fuel self-sufficiency among them, each governed by different physics, different engineering constraints, and in several cases different and only partially compatible design incentives. A configuration optimized aggressively against any one of these constraints in isolation has no guarantee of remaining inside the others, and the field's decades-long difficulty is best read, in this book's vocabulary, not as a single admissibility boundary proving stubborn but as the intersection of several boundaries proving small relative to the individually admissible regions each one bounds.

8.6 Margin, Not Just Membership

The disruption case adds one further refinement that the coupled-constraint picture alone does not capture. A system can occupy a state that satisfies every relevant constraint at a given instant and still not belong to the viability kernel in the fuller, dynamical sense, if the available controls are not powerful enough to keep the system inside the constraint set against its own subsequent evolution and against realistic perturbations. This means the naive admissible region computed from steady-state physics, the set of density, temperature, and confinement-time combinations satisfying the Lawson criterion at one moment, systematically overstates the true viability kernel, the narrower set of states from which the system can actually be held inside every required constraint continuously, using only the control authority the reactor's actual engineering provides. The gap between these two regions is not an error in the physics. It is what viability theory predicts whenever a system's own dynamics can carry it toward a constraint boundary faster than its available controls can correct for, and disruption research is, in this book's terms, largely the study of how much smaller the real viability kernel is than the nominally admissible operating window computed without accounting for that gap.

8.7 What Fusion Generalizes To

Fusion is the chapter's experimental object, not its true subject, and it is worth stating plainly what the case generalizes to before closing Part III. Any system in which sustained operation requires the simultaneous, continuous satisfaction of several independently constrained conditions, rather than a single threshold crossed once, exhibits this same structure: a reachable region wider than a sustainable one, a coupled constraint set in which improving one margin can tighten another, and a true viability kernel narrower than the naive admissible region computed from any single subsystem's physics in isolation. Chapter 6's recursive manufacturing systems already showed a version of this, a process viable at the level of a single reachable event failing to remain viable under repeated operation once yield, dependency, and resource cost were accounted for jointly rather than one at a time. Fusion simply makes the same structure visible with unusual experimental precision, because its constraint set, plasma physics, materials science, and fuel cycle chemistry, is unusually well characterized and unusually resistant to being satisfied all at once.

This closes Part III on a question Part IV is built to answer. Every constraint examined in this chapter, disruption avoidance, materials lifetime, tritium self-sufficiency, could in principle be relaxed by further research, further investment, or further time, and different research programs currently make different bets about which of these coupled constraints is most tractable to relax first. Deciding which of several coupled, expensive, multi-decade traversals a civilization actually attempts, sustains funding for across the setbacks a project of this length inevitably produces, and does not abandon before the relevant viability kernel has been found, is not a question plasma physics, materials science, or fuel-cycle engineering can answer on its own. It is a question about institutions, and it is exactly the question Part IV takes up next: who chooses which traversals get attempted, and what determines whether a civilization's institutions can sustain the attempt long enough to find out whether a given viability kernel is empty or merely difficult to reach.

8.8 Ledger

This chapter records the distinction between a reachable region and a viability kernel as a formal sharpening of a gap the book has approached informally since Chapter 1: reaching a state once, as the NIF ignition result demonstrates, establishes membership in the reachable set without establishing membership in the narrower set of states from which a system can be held indefinitely against its own dynamics and realistic perturbation, as tokamak disruption events demonstrate from the opposite direction, a system already inside its nominal admissible region being ejected from it by dynamics no external energy-budget change caused. It records energetic admissibility as a coupled systems property rather than a single threshold, with plasma confinement, structural material lifetime, and fuel self-sufficiency identified as three separately governed constraint regions whose intersection, not any one of them alone, determines whether a reactor design is viable, and with

the pattern, familiar from Chapter 7's aluminum and desalination cases, that improving performance against one constraint can measurably tighten another, appearing here at the level of an entire coupled engineering system rather than a single process. Finally, it records margin as distinct from mere momentary membership: a state satisfying every constraint at one instant can still lie outside the true viability kernel if the system's available controls cannot hold it there against its own subsequent evolution, meaning the size of the viability kernel depends on control authority in addition to the constraint set itself, a relation this book has not previously had occasion to state and that Chapter 11 should treat as a genuinely new entry rather than a restatement of reachability or stability alone.

Part IV

Institutional Substrates

Chapter 9

Search Over Admissibility Itself

9.1 A Different Kind of Substrate

Chapters 2 through 8 examined three substrates, each with its own admissible region, its own reachable subset of that region, and its own patterns of stability, legibility, and reconstructibility. In every case, the agents doing the searching, a nervous system learning, a chemist synthesizing, a reactor operator controlling a plasma, were themselves taken as given, and the chapter's task was to characterize what the substrate under their control would and would not permit. Part IV changes the object under examination. Institutions do not occupy an admissible region defined by chemistry or plasma physics. They determine which parts of those regions ever get searched, funded, attempted, retained, and built upon, by agents who are themselves shaped by the institutions that trained, employed, and evaluated them. Chapter 8 closed by asking who chooses which multi-decade, coupled-constraint traversals a civilization actually attempts, and what determines whether the attempt is sustained long enough to discover whether a given viability kernel is empty or merely difficult to reach. This chapter takes up that question directly, and its governing claim can be stated plainly: institutions do not merely operate within an admissible region inherited from some other substrate. They determine which regions become searchable in the first place.

9.2 Lock-in: Choosing One Admissible Path Among Several

A civilization's institutions confront, at any given moment, multiple physically admissible paths toward a given goal, and choosing to build the sustained expertise, supply chains, regulatory familiarity, and trained workforce around one of them can foreclose the others, not because the alternatives become physically inadmissible but because no institutional path toward them is ever built. The history of nuclear reactor design supplies an unusually well documented case. Several reactor architectures, cooled and moderated by different combinations of water, graphite, gas, and liquid metal, were physically viable candidates

during the early development of civilian nuclear power, and computational and engineering assessments at the time did not uniformly favor the light-water design that came to dominate global reactor construction. The light-water design's dominance is traceable in large part to its earlier adoption for naval submarine propulsion, where compactness mattered more than the criteria that would later govern civilian power generation, and to the resulting accumulation of specialized engineering expertise, supply chains, and regulatory experience that made light-water designs progressively cheaper and less risky to build relative to alternatives, independent of whether those alternatives remained physically superior along other dimensions [42]. This is technological lock-in in the strict economic sense, and it is worth reading in this book's vocabulary rather than treating it as a separate phenomenon. The reactor designs that lost out were not rendered chemically or physically inadmissible by this history. They became institutionally unreachable, in the specific sense that no accumulated expertise, licensing precedent, or supply chain existed to make attempting them affordable relative to the increasingly well-supported alternative, echoing Chapter 6's finding that manufacturing admissibility resides in a distributed dependency structure rather than in any single component, now applied to an entire national industry's choice of which physically admissible design to build that structure around.

9.3 The Searched Region Is Narrower Than the Admissible Region

Lock-in shows institutions selecting among paths already identified as viable candidates. A more fundamental institutional filter operates earlier, determining which candidate ideas are proposed, funded, and tested at all, and this filter has been measured directly rather than only inferred from historical outcome. A large-scale study of a major grant-funding program found that proposals combining knowledge from more distant, less conventionally associated fields, a reasonable proxy for genuine novelty, were rated systematically less favorably by expert review panels than proposals recombining more familiar, closely associated knowledge, even when the more novel proposals went on, when they were funded, to produce outcomes at least as valuable as the conventional ones [43]. This is a direct, quantified demonstration that the space of research a scientific community actually attempts is narrower than the space of research that community could productively attempt, and that the narrowing is not random but systematically biased against exactly the proposals most likely to explore admissible territory the field has not yet searched.

This motivates a distinction the book has not previously needed. Every substrate examined in Parts I through III had an admissible region and a narrower reachable subset of it, determined by the substrate's own dynamics and the transformations available to whatever agent was acting on it. Institutions introduce a further, narrower region still: the searched region, the subset of the admissible and reachable space that institutional processes, funding review, regulatory approval, editorial judgment, actually permit an agent to attempt. The gap between reachable and searched is not a physical or informational constraint of the

kind Chapters 1 through 8 examined. It is a constraint imposed by an institution's own selection process, and Boudreau and colleagues' finding shows this gap is neither negligible nor evenly distributed across the reachable space, but concentrated specifically at its less familiar edges.

9.4 A Constraint Layered on Top of a Substrate: Pharmaceutical Development

The pharmaceutical development pipeline shows this institutional layer operating directly on top of the molecular substrate examined in Part II, and shows how sharply it narrows what was already, at the chemistry level, a demonstrated admissible and reachable candidate. Estimates of the cost and attrition rate of bringing a new drug from initial discovery to market approval show that only a small fraction of compounds entering formal preclinical development, themselves already a highly filtered subset of the synthesizable molecules examined in Chapter 5, survive the sequential institutional gates of preclinical safety testing and the multiple phases of human clinical trials required for regulatory approval, with the cumulative attrition and cost of the surviving path running into the billions of dollars and well over a decade of elapsed time per approved compound [44]. Every one of the compounds that fails this pipeline was, in Chapter 5's sense, chemically admissible and, in most cases, already demonstrated as synthesizable at laboratory scale. What determines whether it proceeds further is not a further chemistry problem but an institutional admissibility gate, safety and efficacy standards, funding continuity across a multi-year trial program, regulatory judgment, layered on top of a substrate that had already been shown to admit the molecule in question. The searched region here is a specific, formally defined subset of the reachable region, narrowed by a sequence of institutional checkpoints each of which can terminate a candidate that the underlying chemistry never disqualified.

A related institutional mechanism, intellectual property law, shows the same layering operating on recombination rather than on initial testing. Legal analysis of how patent scope is set and interpreted shows that a patent granted with broad claims over a foundational technique can restrict which follow-on uses of that technique other researchers and firms are legally permitted to attempt, independent of whether those follow-on uses are technically or scientifically admissible, meaning the institutional definition of a patent's boundaries can narrow the searched region for an entire subsequent field of inquiry, in ways that vary considerably depending on how courts and patent offices interpret scope in a given jurisdiction and era [45]. Here the institutional constraint is not a funding or approval gate applied to a single candidate but a legal boundary applied to an entire technique, illustrating that the searched-region concept operates at multiple institutional scales at once, from a single grant panel's judgment on a single proposal to a legal system's judgment on an entire class of follow-on innovation.

9.5 The Reversal: Institutional Constraint Can Be Productive

Chapter 4 showed that a constraint narrowing a substrate's admissible region is not automatically a cost, and in the case of convolutional weight-sharing, was shown to be the specific mechanism responsible for making an otherwise intractable search tractable. The same reversal applies at the institutional scale examined in this chapter, and it is worth stating explicitly rather than leaving the pharmaceutical and grant-review cases to read only as costs imposed on otherwise unconstrained search. A field in which every proposed idea, however implausible, and every synthesized compound, however untested, received unlimited institutional resources to pursue to completion would not search its admissible space more effectively than a field with well-designed gates. It would exhaust its resources across an enormous number of unproductive traversals, echoing Chapter 4's convolutional case directly: the exclusion of most of the technically admissible region was what made the remaining, narrower search tractable at all, not merely cheaper. Peer review's systematic bias against distant recombination, examined critically above as a cost, coexists with peer review's role in preventing a research community from attempting every combinatorially possible pairing of existing findings, most of which would be unproductive. The pharmaceutical pipeline's attrition, examined above as a narrowing of the searched region relative to the chemically admissible one, is simultaneously the mechanism by which a healthcare system avoids exposing patients to the overwhelming majority of chemically admissible compounds that would, if tested directly in humans without preceding safety filtration, cause harm. Institutional constraint, exactly like the substrate-level constraint examined in Chapter 4, is not distinguishable from productive search discipline by its narrowing effect alone. It is distinguishable only by whether the specific region it excludes was, on balance, worth excluding, a judgment this chapter's cases show institutions sometimes making well, as in trial-phase safety gating, and sometimes making poorly, as in systematic novelty aversion, using the same underlying mechanism, a narrowed searched region, in both cases.

This is a stronger echo of Chapter 4's finding than a loose analogy. The same structural relation, a constraint that shrinks a searchable or learnable region while increasing the effective productivity of what remains searchable or learnable, recurs here at an entirely different substrate, institutions rather than artificial neural networks, governed by entirely different mechanisms, legal and administrative process rather than architectural connectivity. That recurrence, across substrates sharing no implementation detail whatsoever, is exactly the kind of evidence Chapter 11 is positioned to weigh most heavily, since a structural relation appearing independently in a learning system's weight-sharing and in a funding agency's review panel is considerably harder to dismiss as coincidence than a relation observed only once.

9.6 Toward Reconstructive Capacity

This chapter has focused on institutions as gatekeepers of what gets attempted going forward, from lock-in's selection among physically admissible paths to peer review's and regulatory review's narrowing of the searched region relative to the reachable one. A further institutional function has been assumed throughout without being examined directly: institutions do not only decide what gets attempted next. They determine what is retained from what has already been attempted, which findings, techniques, and capabilities survive institutionally, and which are lost even when the underlying chemistry, physics, or biology that once produced them remains, in principle, exactly as admissible as it ever was. Chapter 6 introduced the distinction between output reproducibility and productive-capacity reproducibility using the Saturn V's F-1 engine, a single artifact. Chapter 10 extends that distinction from a single engine to entire technological lineages, asking what it means for a civilization, rather than a single factory, to remember or forget how something was done.

9.7 Ledger

This chapter introduces institutional lock-in as a mechanism that selects among multiple substrate-level admissible paths and renders the unselected paths institutionally, rather than physically, unreachable, extending Chapter 6's finding that manufacturing admissibility resides in a distributed dependency structure to the scale of an entire national technology choice. It introduces the searched region as a category distinct from the admissible and reachable regions examined in Parts I through III, narrower than both and shaped by institutional selection processes rather than substrate dynamics, and shows this gap measured directly and shown to be systematically biased against novel recombination in the case of grant review, and formally structured as a sequence of checkpoints layered on top of an already-established chemical admissibility in the case of pharmaceutical development. It records patent scope as a further institutional mechanism narrowing the searched region, operating on legal permission to attempt a follow-on use rather than on funding or approval for an initial one, showing the searched-region concept applies at multiple institutional scales simultaneously. Finally, and most significantly for the synthesis chapter, it records a second, independent occurrence of Chapter 4's reversal: a constraint that narrows a searchable region is not thereby shown to have reduced a system's effective capability, and may instead be the mechanism responsible for making the remaining search tractable, a relation now observed at two substrates, artificial learning architectures and institutional review processes, that share no implementation detail whatsoever.

Chapter 10

Reconstructive Capacity

10.1 What This Chapter Has to Settle

Chapter 1 proposed, without yet demonstrating it twice over, that reconstruction is not legibility plus enough time. It is a second reachability problem, conditioned on whatever traces survive from an earlier state, and satisfying it requires a genuinely new admissible path rather than a retracing of an old one. Chapter 6 gave that claim its first full demonstration, using the Saturn V's F-1 engine: complete documentation, fully legible, and a productive capacity, the specific tooling and trained workforce that once built the engine, that had nonetheless been lost and had to be substantially rebuilt from a different starting point. That is one direction of the separation between legibility and reconstructibility. This chapter's task is to establish the direction Chapter 6 did not reach, at the scale Chapter 6 did not attempt. A single engine is one artifact. This chapter needs a case operating at civilizational scale, and it needs the converse of the F-1 pattern: a case in which reconstruction succeeds, or productive capacity survives, without ever passing through a stage of complete explicit specification at all. Only with both directions in hand does reconstructibility stop looking like a shadow of legibility and start looking like what Chapter 1 proposed it to be, an independent operation on the same evolving substrate.

10.2 Direction One, at Civilizational Scale: Roman Concrete

Roman marine concrete supplies an unusually clean civilizational-scale instance of the F-1 pattern, and it is worth working through carefully because the details matter more than the headline. Structures built from this material, harbor installations and breakwaters exposed continuously to seawater for roughly two thousand years, remain standing today in a condition that ordinary modern concrete, exposed to the same marine environment, would not survive for more than a few decades. The material's basic recipe was not lost in the sense of vanishing without any trace. A description of the process, using volcanic ash, lime, and seawater, survives in Vitruvius's first-century architectural treatise, giving

later readers a legible, textual specification of the general method. What Vitruvius's text does not supply, and what was not reconstructed successfully for many centuries afterward, is an understanding of why the material behaves the way it does, which particular volcanic sources produce the effect, and what proportions and curing conditions reliably reproduce ancient performance rather than an inferior modern approximation. Contemporary materials science, using electron microscopy and mineralogical analysis unavailable to any Roman builder or any subsequent reader of Vitruvius, eventually identified the specific chemistry responsible for the material's durability: seawater percolating through the concrete over centuries drives an unusual, ongoing mineral-forming reaction, producing aluminous tobermorite and related crystals that continue to reinforce the material's internal structure long after initial curing, a self-reinforcing process with no analog in ordinary hydraulic cement chemistry [46].

The case reproduces the F-1 pattern with unusual precision, and at a civilizational rather than single-artifact scale. A textual specification survived and was, in an ordinary sense, legible to anyone who could read Latin architectural treatises across the intervening centuries. That legibility did not, by itself, constitute reconstructibility, and it did not for a very long historical stretch, because the productive capacity required to translate a general recipe into a reliably durable material depended on chemistry no available theory could explain and no available instrument could detect. When reconstruction of the material's actual performance finally succeeded, it did not succeed by a more careful reading of Vitruvius. It succeeded by an entirely new admissible path, modern analytical chemistry applied to surviving samples, confirming Chapter 1's claim that reconstruction is a fresh search for a path to a target state rather than a recovery of the original path. Vitruvius's text was necessary to know what to look for. It was never sufficient to produce it.

10.3 Direction Two: Reconstruction Without a Legible Specification

The converse case requires a different kind of evidence, because it asks the chapter to show something structurally harder to observe: capability regenerated, or transferred, or reconstructed, without a stage of explicit specification ever mediating the process at all. The clearest theoretical grounding for why this should even be possible comes from outside the historical record entirely, in the study of tacit knowledge, which begins from the observation that skilled practitioners routinely know how to perform a task correctly without being able to give a complete verbal account of what they know, a phenomenon summarized in the frequently cited formulation that people can know more than they can tell [47]. If this is correct, then a body of productive knowledge can exist, be genuinely operative, and be transmitted from one practitioner to another, without ever passing through a stage at which it becomes a legible, complete specification available to an outside observer, which is precisely the structural possibility Chapter 10 needs to demonstrate rather than merely assert.

Empirical support for this possibility, at an industrial rather than a craft scale, comes from research on how manufacturing capability actually transfers between organizations attempting to adopt an established process. Studies of technology transfer repeatedly find that written documentation, however detailed, is frequently insufficient on its own to reproduce a manufacturing capability in a new location, and that successful transfer typically requires direct, extended contact between practitioners, engineers physically working alongside the personnel who already possess the capability, because a substantial portion of what makes a process work reliably consists of what has been called sticky information, knowledge that resists costless transfer through documentation and can only be moved efficiently by moving the people who hold it, or by extended co-located practice [48]. This is the mechanism by which a productive capacity can survive, or extend itself to a new location, or in principle be reconstructed after a partial interruption, through channels that never require the underlying knowledge to become a complete legible specification at any point in the process. A firm receiving a fully documented manufacturing specification and a firm receiving direct apprenticeship from experienced practitioners are not equally positioned to reproduce the same capability, and the sticky-information literature's finding is precisely that the second route succeeds where the first frequently does not.

A related methodological field makes the same point at civilizational scale, applied specifically to technologies for which no written specification ever existed in the first place, because the societies that developed them had no writing system, or did not record technical processes in writing even when literate. Experimental archaeology is organized around the premise that certain categories of ancient technological knowledge, stone tool manufacture, early metallurgy, textile production, boat construction, cannot be recovered by reading anything, because nothing describing the process in explicit terms survives, and can only be recovered, to whatever extent they are recovered at all, by attempting to physically reproduce the artifacts and processes under conditions approximating the originals, using the surviving artifacts themselves, rather than any written account, as the only available guide [49]. When such reconstructions succeed, and they frequently do, to the point of producing artifacts functionally indistinguishable from originals, the success is achieved entirely without the mediation of a legible specification, since none exists to mediate through. Reconstructibility, in these cases, is demonstrated directly, while legibility of the original process, in the sense of an explicit account an observer could read and follow, was never available and never becomes available even after the reconstruction succeeds.

10.4 The Floor: When Neither Route Survives

A third case is worth including briefly, not because it demonstrates a new relation but because it establishes what the other two cases are implicitly measured against. Byzantine incendiary weaponry, commonly known as Greek fire, was deployed with significant military effect for several centuries and was, according to the historical record available to modern scholarship, closely guarded as a state secret, its precise composition apparently

restricted to a small number of individuals connected to the imperial court. Unlike the Roman concrete case, no textual specification detailed enough to constitute a legible recipe appears to have survived in the historical record. Unlike the experimental archaeology cases, no continuous craft lineage or tacit apprenticeship tradition carried the specific formulation forward either. The result, as far as the historical and scholarly record can establish, is that the exact composition has never been reconstructed with confidence, and modern proposals remain informed speculation rather than demonstrated reconstruction. This is the case in which neither of the two routes examined above was available, and it is included here as the floor the other two cases should be read against: reconstruction is not guaranteed merely because enough time and modern science have since become available, when neither a legible specification nor a surviving productive lineage gave that later science anything concrete to work from.

10.5 Legibility Is Neither Necessary Nor Sufficient

The three cases together settle the question this chapter was assigned. Roman concrete shows that legibility, a genuinely available textual specification, is not sufficient for reconstruction, since the specification survived for the entire period during which the material's actual performance could not be reproduced. Experimental archaeology's routine successes with technologies that left no textual trace show that legibility is not necessary for reconstruction either, since capability has been recovered, repeatedly and demonstrably, through routes that never passed through an explicit specification at any stage. Greek fire shows what happens when neither route is available, which is not automatic recovery through the mere passage of time and the accumulation of general scientific capability, but continued, unresolved loss. Reconstructibility, on this evidence, is exactly what Chapter 1 proposed it to be and what the reframed objective for Chapter 11 requires it to be demonstrated as: an independent operation on the substrate, formally a second reachability problem conditioned on whatever traces, textual, physical, or tacit, happen to survive, and irreducible to legibility considered on its own.

This has a direct consequence for the institutional retention question Chapter 9 deferred to this chapter. An institution, or a civilization, seeking to preserve its own capacity for future reconstruction cannot satisfy that goal through documentation alone, however thorough, because the Roman concrete case shows documentation can survive fully intact while the reconstructible capability it describes is lost for centuries. Nor is documentation dispensable, because most technological knowledge does not have the good fortune of surviving through an unbroken experimental-archaeology-style physical record either. The institutional lesson, generalizing Chapter 6's recursive closure ratio from a single manufacturing network to an entire civilization's retained capability, is that reconstructive capacity depends on maintaining both channels, legible specification and living productive lineage, because either one alone has been shown, in a real historical case, to fail on its own.

10.6 Closing Part IV

Chapter 9 examined institutions as the entity that determines which admissible transformations, across the neural, molecular, and energetic substrates examined in Parts I through III, ever get attempted at all. This chapter has examined institutions, and civilizations more broadly, as the entity that determines what survives from what was already attempted, and has shown that survival takes at least two structurally distinct forms, neither reducible to the other, that must both be maintained for a civilization's reconstructive capacity to be robust against the specific kind of loss each form is vulnerable to on its own. With this in place, all four parts of the book have now supplied concrete, cross-domain evidence for the three coupled problem-clusters proposed as the book's candidate topology, traversal, continuation, and recovery, and Chapter 11's task is no longer to propose that topology speculatively but to test it formally against the accumulated ledger from all nine preceding chapters and determine what mathematical object would have to exist for every entry in that ledger to hold simultaneously.

10.7 Ledger

This chapter records legibility as neither necessary nor sufficient for reconstructibility, demonstrated jointly by two independent civilizational-scale cases: Roman marine concrete, where a genuinely legible textual specification survived for roughly two thousand years without enabling reconstruction of the material's actual performance, which was eventually achieved through an entirely new admissible path rather than a more careful reading of the surviving text; and the routine practice of experimental archaeology, where reconstruction of ancient technological capability is repeatedly achieved for artifacts and processes that left no legible specification of any kind, through direct physical reproduction guided only by surviving artifacts. It records a third case, Greek fire, as the floor against which both are measured, a technology for which neither a legible specification nor a surviving productive lineage was available, and for which reconstruction has correspondingly not succeeded, confirming that at least one of the two routes is required rather than reconstruction being achievable through the general accumulation of later scientific capability alone. It records institutional reconstructive capacity, generalizing Chapter 6's recursive closure ratio and Chapter 9's retention question to civilizational scale, as dependent on maintaining both legible documentation and living productive lineage jointly, since each has now been shown, independently, to be capable of surviving while the other is lost. With this entry, the ledger's Recovery cluster is closed for the purposes of Part IV, and Chapter 11 inherits it, along with the Traversal and Continuation clusters accumulated across Parts I through III, as the complete evidentiary base against which the book's candidate topology is to be tested.

Part V

Synthesis

Chapter 11

The General Theory of Admissibility

11.1 The Chain That Did Not Survive

Chapter 1 proposed, and immediately began testing, a tempting first-pass ordering of its six terms:

$$P \to A \to R \to S \to L \to C,$$

possible narrowing to admissible, admissible narrowing to reachable, reachable narrowing to stable, stable narrowing to legible, legible narrowing to reconstructible. Ten chapters of case material have now had the opportunity to kill this ordering, and it is worth stating plainly, one clause at a time, that they did.

Chapter 2 killed the reachable-as-a-property-of-a-state reading directly. The manifold-learning results showed that from a single starting configuration, some directions of extension were readily trainable and others were not, meaning reachability could not be a ranking attached to destinations but had to be a relation attached to a specific transition. Chapter 3 killed the legible-as-a-property-of-a-state reading in two independent ways: decoder and level-of-description dependence showed legibility varying with the observer applying it to an unchanged state, and representational drift showed legibility varying over time for a state that remained stable and behaviorally load-bearing throughout. Chapters 5 and 6 killed any reading in which reachable and manufacturable at scale could be treated as the same achievement, the taxol case and the semiconductor dependency structure both showing a molecule or a process reachable in the strongest demonstrated sense while remaining entirely unreachable as a repeatable, affordable, industrially embedded capability. Chapter 8 killed stable as a binary property outright, replacing it with a viability margin, the finding that a state satisfying every relevant constraint at a single instant can still lie outside the set of states a system can be dynamically held within against its own perturbations. Chapter 10 killed the ordering's final link from both directions at once, showing a case in which legibility survived for two thousand years without producing reconstructibility, and a body of practice in which reconstructibility is routinely achieved with no legibility, in the sense of an explicit specification, ever having existed at all.

This section is deliberately forensic rather than diplomatic about the result. The chain proposed in Chapter 1 is false. It was useful precisely because testing it, rather than assuming it, is what generated every distinction the book has since relied on, but nothing built from this point forward should retain any trace of a six-stage funnel.

11.2 Traversal

The first object to survive the chain's collapse can now be reconstructed properly. A state y is never simply reachable. It is reachable from a specific state x, under a specific set of operative constraints $\mathcal{C}$, using a specific class of available transformations $\mathcal{T}$, given a specific resource budget $\mathcal{B}$:

$$x \xrightarrow[\mathcal{C},\mathcal{B}]{\mathcal{T}} y.$$

This single relation is sufficient to integrate the book's neural, molecular, energetic, and institutional evidence without requiring any claim that the underlying mechanisms are the same. Chapter 2's manifold results specify $\mathcal{T}$ as the set of update rules available to a learning system and show reachability varying by direction from a fixed x. Chapter 5's retrosynthesis results specify $\mathcal{T}$ as the set of known reaction classes and show the same directional asymmetry in a completely different substrate. Chapter 7's energy results show $\mathcal{B}$ expanding $\mathcal{C}$'s effective bite without ever specifying $\mathcal{T}$ directly, exactly the distinction that chapter needed and could not have stated this precisely at the time. Chapter 9's institutional results show that $\mathcal{T}$ itself, the set of transformations an agent is permitted or resourced to attempt, is set by an entity, an institution, external to the substrate undergoing transformation, which is the clearest sense in which admissibility at one substrate is governed by a structure external to it.

The chapter's central result can now be stated in the form the whole book has been building toward. Reachability belongs to paths, not states. Every attempt across ten chapters to treat it as a property of a destination broke on contact with a real case, and every successful account of those cases required specifying the transition rather than the target.

11.3 Continuation

A second question survives independently of the first: what happens after a traversal succeeds. Chapter 8 supplies this section's dominant result, because stable, examined honestly, turned out to be too weak a category to carry the weight Chapter 1 initially assigned it. What matters is not whether a system currently satisfies a constraint set but whether its trajectories can continue to satisfy that constraint set under its own subsequent dynamics and under realistic perturbation, and with what margin. Formally, continuation asks whether x belongs to the viability kernel of a constraint set $\mathcal{C}$ under the controls actually available, a narrower condition than membership in $\mathcal{C}$ at a single instant [37].

This yields the section's central result. Arrival and continuation are different problems,

solved by different means, and a system can solve the first while failing the second, as the NIF ignition and tokamak disruption cases showed from opposite directions.

Chapters 4 and 9 fit into this section more precisely than either chapter alone could show. A constraint that excludes certain trajectories, sign-restricted connectivity in a recurrent network, a funding panel declining to support an implausible proposal, does not merely subtract from the space of what a system could do. It can increase effective continuation, by excluding specifically those trajectories that would have carried the system out of a viable region, whether that region is a solvable weight configuration or a productive, fundable research program. Constraint, in other words, cannot be represented merely as subtraction from possibility. In at least these cases, it is closer to a filter selectively removing the trajectories least likely to remain viable, which is a claim about continuation, not merely about the size of an admissible region.

11.4 Recovery

Chapter 10 has left this section unusually little work to do, because it already supplied the cleanest formal result available anywhere in the book. Legibility is an observation relation, $O(x; D, L)$, holding or failing to hold between a state, a decoder, and a level of description, established in Chapter 3 and confirmed rather than merely repeated in Chapter 10. Reconstruction is a second, independent traversal problem, conditioned by whatever traces survive from an earlier state x_{t_0}, asking whether a present system x_t, together with its surviving records and productive capacities, admits a path to some state $\tilde{x}_{t_0}$ functionally equivalent to the original:

$$x_t \rightsquigarrow \tilde{x}_{t_0}.$$

Chapter 10's two civilizational cases now state the relation between these two operations with a precision the earlier chapters approached but did not achieve. Roman concrete supplies $L \not\Rightarrow C$: legibility, a genuinely available textual specification, did not imply reconstructibility, for roughly two thousand years. Experimental archaeology and the sticky-information literature together supply the converse, $C \not\Rightarrow L$: reconstructibility, demonstrated directly and repeatedly, does not imply that legibility, in the sense of a complete external specification, was ever achieved at any point in the process. Together these two results are probably the cleanest formal demolition of the original chain to be found anywhere in the book, since they show the final link failing in both directions independently rather than merely weakly.

The principle this section leaves standing is stronger than either case alone. A record can constrain reconstruction without containing the reconstructed capability. Vitruvius's text narrowed the space of plausible reconstruction attempts for two thousand years without itself supplying the missing chemistry. This is not a narrow claim about ancient engineering. It is a claim about what documentation of any kind can and cannot be expected to do, and it resonates well beyond the technological cases this book has examined.

11.5 The Object That Survived

Sections 11.2 through 11.4 have now reconstructed three objects independently: a directed, indexed traversal relation, a viability kernel governing continuation, and a pair of operations, an observation relation and a conditioned reconstruction relation, governing recovery. This section asks what single mathematical object would have to exist for all three, and for every ledger entry accumulated across the preceding ten chapters, to hold simultaneously, and then tests a candidate against that full ledger rather than merely asserting that it fits.

The candidate is a directed, indexed transition geometry: a state space, a set of admissible transformations between states, directed paths through that space rather than an undirected adjacency, viability regions within the space together with margins measuring their robustness to perturbation, observation maps relating states to what specified observers can recover from them, and reconstruction operators relating present states and surviving traces to candidate regenerations of past states. Every element of this object is indexed, to substrate, to historical state, to available resources, to a specified observer, and to a specified level of description, since Chapters 2 through 10 showed repeatedly that dropping any one of these indices from a claim about admissibility, reachability, or legibility produces a false generalization.

Tested against the ledger, most entries fit directly. Chapter 2's directional reachability is exactly the directed-path structure. Chapter 3's decoder and level-of-description dependence is exactly the observation map, indexed as required. Chapter 5's indexed admissibility, the same target molecule inadmissible relative to one starting material and admissible relative to another, is exactly the substrate-and-resource indexing built into the object from the start. Chapter 8's margin-not-membership result is exactly what the viability regions were built to capture, and captures it without requiring any further primitive. Chapter 10's two-directional dissociation between legibility and reconstructibility is exactly the separation between the observation map and the reconstruction operator, which the object treats as two distinct maps rather than stages of one process, precisely because Chapter 1 warned against the opposite reading and Chapter 10 confirmed the warning was correct.

Two entries resist this treatment cleanly enough to be worth stating honestly rather than folding in by force. Chapter 2's continuation-geometry result, that a system's present organization constrains which further transitions are subsequently reachable, requires the admissible-transformation structure itself to depend on the path already taken to reach the current state, not merely on the current state considered in isolation. A transition geometry in which admissibility depends only on the present point, the way a simple Markov process would, cannot represent this finding; the object has to be history-dependent, with admissible transformations conditioned on trajectory rather than location alone, which is a real addition forced by the evidence rather than a decorative refinement.

The second gap is sharper. Chapter 4 and Chapter 9's shared reversal, that a constraint can increase effective capability by excluding unproductive trajectories, is not fully captured by an admissible-transformations set alone, because that set only distinguishes what

is permitted from what is forbidden, and says nothing about how densely a permitted region is actually explored. The productive-constraint results in both chapters are claims about search density or attention allocated across an admissible region, not merely about the region's boundary. Representing them properly requires a further primitive beyond the six already listed: some measure of how a system's available effort is distributed over its own admissible transformations, since the entire content of Chapters 4 and 9's central finding is that this distribution, not merely the admissible set it ranges over, determines effective capability. The object, tested honestly against its own evidence, needs a seventh element, a search or attention measure over the admissible transformation set, layered on top of the six already established. This is exactly the kind of incompleteness this section was built to expose rather than paper over, and the book is better for finding it here than for pretending the six-part object was complete.

11.6 What Constraints Do

The gap just identified deserves its own section, because Chapters 4 and 9 independently produced the same reversal from substrates sharing no implementation detail whatsoever, and the reversal is significant enough to state as a general principle rather than leave attached to either chapter alone. Constraints have at least two distinguishable effects on a system's admissible transformation space. They exclude certain transformations outright, narrowing the set. But, as the search-density primitive just introduced makes precise, they can also organize search within what remains, concentrating a system's finite effort on the region of the remaining space most likely to be productive, which convolutional weight-sharing and disciplined grant review both accomplish through entirely different mechanisms.

This has a consequence worth stating as one of the book's deepest conclusions. Relaxing a constraint does not monotonically increase a system's effective capability. Removing a restriction can expand the raw space of what is technically possible while simultaneously making the useful states within that expanded space harder to find, learn, sustain, or reproduce, because the search effort that constraint had been concentrating is now spread across a larger, mostly unproductive region. Capability is not maximized by maximizing possibility. With this conclusion earned rather than assumed, a title like Admissibility Before Capability stops being a slogan borrowed from the book's own vocabulary and becomes a compressed statement of its central, evidenced finding.

11.7 The Boundary We Cannot Yet Place

Chapter 5 left the Drexler-Smalley mechanosynthesis dispute deliberately unresolved, and this section keeps it that way, because the object developed in Section 11.5 is better used here to classify the disagreement than to adjudicate it. The dispute can now be stated as a question about which specific relation within the transition geometry remains contested.

Advocates and critics of positionally controlled molecular assembly are not necessarily disagreeing about physical possibility in the crude sense; both sides generally accept that individual mechanosynthetic reactions are not forbidden by fundamental physics. The dispute is better read as targeting several distinct, more specific questions at once: whether the proposed mechanism is admissible under the actual many-body electronic interactions governing bond formation at close range, rather than merely under a simplified positional model of it; whether, granting admissibility, an actual reachable assembly trajectory from available starting materials to a working assembler has been identified rather than merely postulated; whether error rates in positional placement can be held within a viable margin under repeated operation rather than accumulating catastrophically, which is precisely Chapter 8's continuation question transplanted to a molecular substrate; and whether, granting all of the above, the resulting process could achieve the kind of scalable, recursively productive manufacturing examined in Chapter 6, as opposed to remaining a laboratory curiosity.

Read this way, the persistence of the debate becomes more informative than either side's declared conclusion. Different participants in the dispute appear, on inspection, to be arguing about different boundaries within the same object, admissibility, reachability, continuation, and recursive productive capacity respectively, while using a single word, possible, to refer to all four at once. This is exactly the kind of information loss the book's entire vocabulary was built to prevent, and the fact that a live, technically sophisticated scientific dispute exhibits it is stronger evidence for the vocabulary's usefulness than a resolved case would have been. Greek fire, introduced in Chapter 10 as the book's clearest case of genuine, unresolved loss, echoes quietly here for the same reason: the theory can sometimes identify precisely which relation remains unknown, admissibility, reachability, or something else, without being able to supply the missing evidence that would settle it, and knowing which question remains open is itself a nontrivial result even when the question stays open.

11.8 The General Theory of Admissibility

The book's central claim can now be stated in its final form. For any system in a given state, capability is determined not by the set of states that are possible in principle, but by the structured relations among the transformations it can admissibly traverse, the regions within which it can continue, the distinctions that can be made legible, and the configurations that can be reconstructed from what survives. These relations are directed, indexed, historically contingent, and altered, sometimes productively, by the constraints under which the system operates.

Returned one final time to the four substrates that supplied its evidence, compressed rather than re-argued: a brain does not explore every representation it could conceivably instantiate, constrained by developmental history and the geometry its own prior learning has already carved into its representational manifold. A molecular system does not traverse every chemically possible configuration, constrained by which reaction pathways have been discovered and which starting materials happen to be available. An energy system does not

make every transformation equally affordable merely by supplying more energy, because energy expands an admissible region without ever specifying its internal structure. An institution does not search the whole technological future available to it, because its own review processes, legal boundaries, and accumulated path dependence determine which fraction of that future is ever attempted at all. Each of these systems inhabits not an open field of possibility but a structured future, one whose shape this book has spent eleven chapters trying to make visible rather than assumed.

This was, from Chapter 1 onward, an etiology rather than a celebration: not an account of what progress is capable of, but of what specifically produces it, chapter by chapter, and of the many identifiable ways that production can fail even where nothing was ever, in the crudest sense, impossible.

The future is not the set of things that can happen. It is the geometry of the paths that remain open from here.

Bibliography

[1] Hubel, D. H., & Wiesel, T. N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *Journal of Physiology*, 206(2), 419–436.

[2] Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, 14(1), 11–28.

[3] Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., Yu, B. M., & Batista, A. P. (2014). Neural constraints on learning. *Nature*, 512(7515), 423–426.

[4] Golub, M. D., Sadtler, P. T., Oby, E. R., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., Batista, A. P., Chase, S. M., & Yu, B. M. (2018). Learning by neural reassociation. *Nature Neuroscience*, 21(4), 607–616.

[5] Held, R., Ostrovsky, Y., de Gelder, B., Gandhi, T., Ganesh, S., Mathur, U., & Sinha, P. (2011). The newly sighted fail to match seen with felt. *Nature Neuroscience*, 14(5), 551–553.

[6] Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, 503(7474), 78–84.

[7] Hochberg, L. R., Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., Branner, A., Chen, D., Penn, R. D., & Donoghue, J. P. (2006). Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, 442(7099), 164–171.

[8] Willett, F. R., Avansino, D. T., Hochberg, L. R., Henderson, J. M., & Shenoy, K. V. (2021). High-performance brain-to-text communication via handwriting. *Nature*, 593(7858), 249–254.

[9] Driscoll, L. N., Pettit, N. L., Minderer, M., Chettih, S. N., & Harvey, C. D. (2017). Dynamic reorganization of neuronal activity patterns in parietal cortex. *Cell*, 170(5), 986–999.

[10] Degenhart, A. D., Bishop, W. E., Oby, E. R., Tyler-Kabara, E. C., Chase, S. M., Batista, A. P., & Yu, B. M. (2020). Stabilization of a brain-computer interface via the alignment of low-dimensional spaces of neural activity. *Nature Biomedical Engineering*, 4(7), 672–685.

[11] Gallego, J. A., Perich, M. G., Naufel, S. N., Ethier, C., Solla, S. A., & Miller, L. E. (2018). Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nature Communications*, 9, 4233.

[12] Anumanchipalli, G. K., Chartier, J., & Chang, E. F. (2019). Speech synthesis from neural decoding of spoken sentences. *Nature*, 568(7753), 493–498.

[13] Horikawa, T., Tamaki, M., Miyawaki, Y., & Kamitani, Y. (2013). Neural decoding of visual imagery during sleep. *Science*, 340(6132), 639–642.

[14] Sadeh, S., & Clopath, C. (2025). The emergence of NeuroAI: bridging neuroscience and artificial intelligence. *Nature Reviews Neuroscience*, 26, 583–584.

[15] Wang, E. Y. et al. (2025). Foundation model of neural activity predicts response to new stimulus types. *Nature*, 640, 470–477.

[16] Crick, F. (1989). The recent excitement about neural networks. *Nature*, 337(6203), 129–132.

[17] Lillicrap, T. P., Cownden, D., Tweed, D. B., & Akerman, C. J. (2016). Random synaptic feedback weights support error backpropagation for deep learning. *Nature Communications*, 7, 13276.

[18] Song, H. F., Yang, G. R., & Wang, X.-J. (2016). Training excitatory-inhibitory recurrent neural networks for cognitive tasks: A simple and flexible framework. *PLOS Computational Biology*, 12(2), e1004792.

[19] Whittington, J. C. R., & Bogacz, R. (2017). An approximation of the error backpropagation algorithm in a predictive coding network with local hebbian synaptic plasticity. *Neural Computation*, 29(5), 1229–1262.

[20] Scellier, B., & Bengio, Y. (2017). Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Frontiers in Computational Neuroscience*, 11, 24.

[21] Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4), 193–202.

[22] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4), 541–551.

[23] Reymond, J.-L. (2015). The chemical space project. *Accounts of Chemical Research*, 48(3), 722–730.

[24] Segler, M. H. S., Preuss, M., & Waller, M. P. (2018). Planning chemical syntheses with deep neural networks and symbolic AI. *Nature*, 555(7698), 604–610.

[25] Ertl, P., & Schuffenhauer, A. (2009). Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of Cheminformatics*, 1, 8.

[26] Holton, R. A. et al. (1994). First total synthesis of taxol. 1. Functionalization of the B ring. *Journal of the American Chemical Society*, 116(4), 1597–1598.

[27] Denis, J.-N., Greene, A. E., Guénard, D., Guéritte-Voegelein, F., Mangatal, L., & Potier, P. (1988). A highly efficient, practical approach to natural taxol. *Journal of the American Chemical Society*, 110(17), 5917–5919.

[28] Baum, R. (2003). Nanotechnology: Drexler and Smalley make the case for and against 'molecular assemblers'. *Chemical & Engineering News*, 81(48), 37–42.

[29] Miller, C. (2022). *Chip War: The Fight for the World's Most Critical Technology*. Scribner.

[30] von Neumann, J. (1966). *Theory of Self-Reproducing Automata* (A. W. Burks, Ed.). University of Illinois Press.

[31] Jones, R., Haufe, P., Sells, E., Iravani, P., Olliver, V., Palmer, C., & Bowyer, A. (2011). RepRap: the replicating rapid prototyper. *Robotica*, 29(1), 177–191.

[32] Grjotheim, K., & Kvande, H. (1993). *Introduction to Aluminium Electrolysis: Understanding the Hall-Héroult Process* (2nd ed.). Aluminium-Verlag.

[33] Elimelech, M., & Phillip, W. A. (2011). The future of seawater desalination: Energy, technology, and the environment. *Science*, 333(6043), 712–717.

[34] Szargut, J., Morris, D. R., & Steward, F. R. (1988). *Exergy Analysis of Thermal, Chemical, and Metallurgical Processes*. Hemisphere Publishing.

[35] Sorrell, S. (2009). Jevons' Paradox revisited: The evidence for backfire from improved energy efficiency. *Energy Policy*, 37(4), 1456–1469.

[36] Lawson, J. D. (1957). Some criteria for a power producing thermonuclear reactor. *Proceedings of the Physical Society. Section B*, 70(6), 6–10.

[37] Aubin, J.-P. (1991). *Viability Theory*. Birkhäuser.

[38] Abu-Shawareb, H. et al. (Indirect Drive ICF Collaboration). (2024). Achievement of target gain larger than unity in an inertial fusion experiment. *Physical Review Letters*, 132, 065102.

[39] de Vries, P. C. et al. (2011). Survey of disruption causes at JET. *Nuclear Fusion*, 51, 053018.

[40] Zinkle, S. J., & Was, G. S. (2013). Materials challenges in nuclear energy. *Acta Materialia*, 61(3), 735–758.

[41] Abdou, M. et al. (2021). Physics and technology considerations for the deuterium–tritium fuel cycle and conditions for tritium fuel self sufficiency. *Nuclear Fusion*, 61, 013001.

[42] Cowan, R. (1990). Nuclear power reactors: A study in technological lock-in. *The Journal of Economic History*, 50(3), 541–567.

[43] Boudreau, K. J., Guinan, E. C., Lakhani, K. R., & Malone, T. W. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*, 62(10), 2765–2783.

[44] DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, 47, 20–33.

[45] Merges, R. P., & Nelson, R. R. (1990). On the complex economics of patent scope. *Yale Law Journal*, 90(4), 839–916.

[46] Jackson, M. D. et al. (2017). Phillipsite and Al-tobermorite mineral cements produced through low-temperature water-rock reactions in Roman marine concrete. *American Mineralogist*, 102(7), 1435–1450.

[47] Polanyi, M. (1966). *The Tacit Dimension*. University of Chicago Press.

[48] von Hippel, E. (1994). "Sticky information" and the locus of problem solving: Implications for innovation. *Management Science*, 40(4), 429–439.

[49] Coles, J. (1979). *Experimental Archaeology*. Academic Press.