Branch Theory

A research proposal toward a subject-relative meta-framework for physical reality

Michiel de Jong, with Claude (Anthropic)

May 2026

Contents

About this document

This is a research proposal, not a finished theory. It sketches a meta-framework called **Branch Theory** that takes subjects, not objects, as foundational to physical reality, and tries to derive from that starting point several things usually treated as postulates: the arrow of time, the structure of observation, the emergence of a shared world among observers, and the entropic cost of being a persistent thing in the world.

The framework is unfinished. It has no formal mathematical apparatus yet. It has not been used to derive the predictions of any known physical theory. It does not, on its own, predict anything. What it claims to offer is a different way of locating the load-bearing assumptions of physics — one that may dissolve some long-standing puzzles by reframing them, and that may, with substantial further work, provide a setting in which specific physical theories can be more cleanly stated.

The document is co-written. The framework is something I (Michiel) have been working out on and off since 1992. Over those decades the intuitions sharpened and the underlying picture clarified considerably, and parts of it became coherent enough to write down publicly: an early sketch of several of the central ideas appeared on my personal blog in January 2012 under the title *On the origin of computational complexity*, framing the questions in terms of Boolean expressions, knowledge items that accumulate as time progresses, and the conversion of OR-information (future) into AND-information (past). The 2012 post is short and informal, but it already contains much of what Branch Theory carries in this 2026 document, including the question of why computation takes time and the diagnosis that this question has not been properly studied at all. What was missing through those decades was a form in which the picture could be put in front of someone else and have it land. Earlier this year I started working through it with Claude, an AI assistant made by Anthropic, and the dialogue produced what the preceding three decades of private thinking had not: better mappings to existing work in the literature, clearer descriptions of the concepts, and a usable structure for the framework as a whole. The substantive claims are mine; the articulation, the references, and a great deal of the sharpening — the surfacing of objections, the spotting of where neighbors in the literature had already done relevant work, the prose itself — emerged from conversation. I've kept the document in a single voice for readability, but the reader should understand that this is a joint product and that some of the moves owe their final form to that collaboration.

The document is structured in four parts. Part I sets up three strands of dissatisfaction with how foundational physics usually answers (or doesn't) certain deep questions. Part II states the framework. Part III sketches how a handful of familiar structures — Turing machines, gases, Maxwell's demon, multi-detector events, logical derivations — might be re-described in the framework, with explicit notes on what works and what doesn't. Part IV is about what would count as progress: open problems, research targets, and the framework's relationship to neighboring programs.

The intended reader is someone with an academic background and an interest in foundations of physics, philosophy of science, or theoretical computer science. The document tries to be accessible across these communities — none of which can be assumed to have the full toolkit needed to evaluate the proposal — but it does not avoid technical material where the technical material is what matters.

Part I — Three strands of dissatisfaction

The framework proposed in this document is a response to a feeling that runs across several different areas of foundational science: that the deepest questions in each area have been *described* mathematically without being *explained.* The descriptions are precise, predictive, and empirically successful. But when you ask what they mean, or where their structure comes from, you are told either that it comes from somewhere else (a boundary condition, an interpretive convention, a postulate) or that the question is malformed. The dissatisfaction is that the place the structure "comes from" never seems to be the right place — it always reduces to relocating the mystery, not resolving it.

This first part of the document tries to make that dissatisfaction concrete in three areas, because Branch Theory is meant as a response to all three simultaneously. Readers familiar with one or two of them may want to skim the corresponding sections, but I'd encourage reading all three even if the territory is familiar, because the way Branch Theory addresses them depends on seeing them as one problem rather than three.

The arrow that shouldn't exist

The fundamental laws of physics, as we understand them, are almost entirely time-symmetric. Newton's laws, Maxwell's equations, the Schrödinger equation, general relativity, the standard model of particle physics — all of them are invariant under reversal of the time coordinate, with one tiny exception in the weak interaction that doesn't account for any of the time-asymmetry we actually experience. Run a movie of any process described by these laws backward, and the backward movie shows a process that the laws also permit.

And yet the world is dramatically time-asymmetric. Eggs break and don't unbreak. Heat flows from hot to cold and not the other way. We remember the past and not the future. Causes precede their effects. The universe is expanding, presumably from something denser and hotter in the past.

The standard story about how to reconcile these is Boltzmann's. Define entropy as proportional to the logarithm of the number of microscopic states compatible with a given macroscopic state. Show that for almost all initial conditions, entropy increases as the system evolves. Conclude that the time-asymmetry we observe is a statistical consequence of time-symmetric dynamics applied to a low-entropy starting configuration.

The story is mostly correct but it has a hole, and the hole has been known since Boltzmann's own time. The argument that entropy increases assumes something called the *Stosszahlansatz* — the assumption that incoming particles in a collision are uncorrelated. This assumption is itself time-asymmetric. It is correct for incoming particles and wrong for outgoing particles, because after a collision the particles' states are correlated by the collision. If you run the argument with the Stosszahlansatz applied to outgoing particles instead of incoming ones, you can derive that entropy *decreases* over time. Boltzmann's argument quietly imports time-asymmetry through this assumption rather than deriving it from the dynamics.

The fallback position is to say: well, the universe started in a very low-entropy state, and we observe entropy increase because we're early in its history. This is sometimes called the Past Hypothesis. Roger Penrose has calculated that the initial entropy of the observable universe was something like one part in $10^{(10}123)$ of what it could have been — a level of fine-tuning so extreme that it cries out for explanation. The Past Hypothesis answers "why does entropy increase?" by saying "because it was low to begin with," but this just relocates the question to "why was it low to begin with?", and there is no widely accepted answer.

Sean Carroll has argued that this — not quantum gravity, not the cosmological constant — is the deepest open problem in physics. It is the question of where time's arrow actually comes from, and as of this writing, no satisfactory answer exists. The candidates include eternal inflation producing low-entropy patches, anthropic selection (which runs into the Boltzmann brain problem), and various forms of "this is just a brute fact" — none of which feel like real explanations.

The dissatisfaction here is not with the predictive content of statistical mechanics, which works beautifully. It is with the explanatory content. We have a description of entropy increase, parametrized by a free choice of coarse-graining, conditioned on a low-entropy boundary condition we cannot account for. This is not a satisfying foundation for the most ubiquitous feature of physical experience.

Observation as a mystery

The second strand of dissatisfaction concerns what it means for something to be observed.

In classical physics, observation is conceptually unproblematic. There is a world out there, configured in some definite way, and observation is the process of finding out how it's configured. The observer can be modeled as a physical system, and the act of observation is just an interaction between two systems, fully described by the same laws that describe everything else.

Quantum mechanics broke this picture, and the break has never been cleanly repaired. The wave function evolves deterministically according to the Schrödinger equation, but when an observation is made, something else happens — a "collapse" — that is not described by the Schrödinger equation and that depends, in some interpretations, on the act of observation itself. The cluster of problems around what observation is, why it has special status, and what role consciousness or measurement apparatus plays is collectively called the **measurement problem**, and after nearly a century of debate it remains open.

Different interpretations of quantum mechanics propose different resolutions. The Copenhagen interpretation treats observation as primitive and refuses to ask further questions. Many-worlds interpretations (Everett, DeWitt, Deutsch, Wallace) deny that collapse happens at all and instead claim that all possible outcomes occur in different branches of the universal wavefunction. Pilot-wave theories (de Broglie, Bohm) restore determinism by adding hidden variables. Relational interpretations (Rovelli) say that the state of a system is only defined relative to another system. QBism (Fuchs, Mermin, Schack) treats the wave function as describing an agent's beliefs rather than physical reality, and observation as Bayesian updating.

Each interpretation has its proponents and each has open problems. What they share is that none of them really *explain* observation; they restructure what observation is taken to be so that the structure of quantum mechanics makes sense. The measurement problem is less a problem to be solved than a fork in the road where you have to commit to a metaphysical stance.

But the deeper issue, the one that motivates Branch Theory, is that the measurement problem in QM is a sharp special case of a much more general puzzle. *What does it mean for a fact about the world to become a fact for someone?* In classical physics this question never arose, because we assumed there was a single shared world that everyone observed parts of. But the moment you take seriously the idea that observers are themselves physical systems embedded in the world, the line between observer and observed becomes unclear, and the notion of "a fact about the world" detached from any particular observer becomes harder to defend.

John Wheeler's slogan "It from Bit" expressed the suspicion that information — and therefore observation — might be more foundational than the physical world it describes. The slogan is provocative and underdeveloped, but it points at the right place: maybe the dissatisfaction with the measurement problem is a clue that we have the ontology upside down. Maybe observation isn't a secondary process that happens to a primary physical world; maybe the relationship between observers and their observed worlds is itself the primary structure, and what we usually call "the physical world" is something derived from it.

This is the second strand of dissatisfaction: not specifically with QM, but with the way foundational physics tries to maintain a clean separation between an observer-independent reality and an observer-dependent process of observation, when there are deep reasons to doubt that separation can be cleanly maintained.

Why computation takes time

The third strand of dissatisfaction is less famous than the other two, and the way I want to frame it may at first seem like a puzzle for computer science rather than a foundational concern. But it bears on the same place, and once stated clearly I think it sharpens the first two strands as well.

The puzzle is this: **why does computation take time?**

Consider a deterministic computer with a finite program that, given any particular input string, produces the same output every time that input is supplied. If the program and the input together fully determine the output, then in some real sense the information contained in the output was already present in the program and input *before* the program ran. *The total information state does not change during computation.* The computation does not generate new information; it only rearranges information that was already there. And yet the computation is not instantaneous. It takes time, sometimes a great deal of time, and for some problems the time it takes appears to be irreducibly large no matter how cleverly the program is written.

This is strange. If the answer is already determined by the input, why can't we just read it off?

There is a partial answer, and it sharpens the puzzle rather than dissolving it. The partial answer is that *some* computational devices can read the answer off, but they pay a price for it. A non-uniform family of circuits — one circuit per input length — can be designed so that for any specific input length, the circuit computes the function in essentially one step: the structure of the circuit hardcodes the answer for that length, like a lookup table. The cost is that the *description* of the family, taken as a whole, is infinite — there's a different circuit for every input length. A *uniform* family, by contrast, is one where some finite program generates the circuit for any given length on request, which restores finiteness of description but reintroduces the time cost: now there's a generating program that has to run, and we've simply moved where the computation happens.

A Turing-complete computer differs from a non-uniform circuit family in exactly this way. Its description is finite — a fixed program of fixed length, the same one for every input. But because the input space is unbounded and the program is finite, the program cannot, in general, contain the answers; it can only contain *rules for combining information from the input* into an answer. And combining information takes time.

There is a way to make this trade-off concrete by looking at the *form* of the information at each stage. The input to a computation is typically a structure with a lot of latent disjunction in it: many candidate hypotheses are still compatible with what has been said so far, and the answer is whichever one survives the constraints. The output, by contrast, is a structure where the disjunction has been resolved: the constraints have been combined, the candidates have been narrowed, and what remains is a conjunction of facts that all hold together. Computation, in this view, is the rewriting of OR-dense representations into AND-dense ones — the gradual replacement of "either X or Y or Z" by "X, and not Y, and not Z." 3-SAT is the canonical example: an input in product-of-sums form (a conjunction of disjunctions, OR-dense at every clause) gets transformed, if the algorithm succeeds, into either a satisfying assignment (a conjunction of literals, AND-dense) or a proof of unsatisfiability. The total information content has not changed across this transformation; what has changed is which form the information is in, and that change is what takes time.

So there is a trade-off, and it sharpens around the difficulty of the function being computed. For simple functions — AND of n inputs, for example — a finite program can produce the answer in time proportional to n, which is fast. For harder functions, the time can grow polynomially, exponentially, or unboundedly with input size. The general statement is: a finitely-describable computer is one that does not have the answers ready and must work them out by recombining the input — by converting OR-information into AND-information — and the time it must spend doing so is a function of how hard the conversion is. For some functions there is no fast conversion, and that is when the trade-off bites.

This is suggestive on its own. It becomes more interesting in light of the open problem of *P versus NP*. P is the class of problems for which a polynomial-time program can find the answer. NP is the class of problems for which, given a candidate answer, a polynomial-time program can verify it. P is trivially contained in NP: if you can find an answer quickly, you can also verify one quickly. The famous open question is whether the reverse holds. Most computer scientists believe it does not — that there are problems for which verifying a proposed answer is fast but finding one is irreducibly slow. If P is not equal to NP, then there exist problems where the input genuinely determines the output but no finite program can produce that output quickly. The information is there; the time to extract it is intractable.

There is a deeper diagnosis worth naming here, because it sets up what Branch Theory tries to address. The reason computer science has not formalized *why* computation takes time — has not produced an account of the origin of computational complexity rather than just measurements of it — may be that the field has abstracted too far from the real-world meanings of its primitive operations. AND, OR, and NOT have real-world meanings: AND is the accumulation of multiple things that hold simultaneously, OR is the openness of multiple possibilities one of which will turn out to obtain, NOT is the absence of something. Time itself has a real-world meaning: it is the direction in which possibilities resolve into actualities, in which OR-structures become AND-structures, in which what was open becomes settled. When computer science treats AND and OR purely as syntactic operators on Boolean variables, divorced from the semantic structures they correspond to in physical and temporal reality, the connection between them and the *time-cost* of computation becomes invisible — because the time-cost is precisely the cost of the semantic process of converting one form into the other, and the syntactic treatment has abstracted that process away.

It is worth putting this in slightly grander terms, even at the risk of overstatement. Theoretical computer science studies computational complexity but does not study its *origin*. It is, in this respect, at the stage where biology was before Darwin: practitioners catalog and measure the phenomena (species, in biology's case; complexity classes, in informatics) without an account of where they come from. The Darwinian move was to recognize that the diversity of species had a generative process — natural selection — that explained why the catalog has the shape it has. The analogous move in informatics has not been made. Computational complexity has a generative origin, presumably in the structure of how information accumulates and how OR converts to AND over time, but the field has not articulated it. The dissatisfaction this strand expresses is that we have no foundational framework in which this question can even be properly posed, let alone answered.

Let me draw a structural parallel to physics that I find suggestive, while being careful about what it does and does not claim.

Breaking an egg is a fast, finitely-describable physical process. The rules governing it — gravity, the mechanical properties of the shell — are simple, and applying them takes essentially no time relative to the process itself. The egg breaks, and what was a coherent structure becomes a distribution of shell fragments, yolk, and white.

Unbreaking the egg is also, in a strict thermodynamic sense, allowed by the underlying physics. The microscopic laws are reversible. There is some sequence of molecular motions which, if executed precisely, would reassemble the egg. The reason it doesn't happen spontaneously is that the set of microstates that would in fact reassemble the egg is astronomically small relative to the typical phase-space measure — it is a measure-theoretic fact, not a complexity-theoretic one.

I want to be careful here, because there is a tempting parallel to be drawn between the two situations and it is easy to overstate. Forward physical evolution is fast, and forward computation is fast. Backward physical evolution is, in some loose sense, "hard," and backward computation (search) is hard. There is something structurally suggestive about both being dichotomies between *carrying the answer* and *working the answer out from primitives*, with time as the currency in both cases.

But the two senses of "hard" are not the same thing. The hardness in P versus NP is about the structure of a computational problem relative to a model of computation — it is a fact about classes of formal problems. The hardness in unbreaking an egg is about phase-space measure — it is a fact about which trajectories are typical under a measure on physical states. Conflating these would be a category error: if they were literally the same fact, then proving P equals NP would entail time-symmetric macroscopic physics, and disproving the Second Law would settle a Clay millennium problem. Neither implication is anywhere on the table.

What I'll claim, more carefully, is this. Both situations exhibit a *trade-off* between description length and time-to-extract-content, and both are time-asymmetric in their cost structure. Whether these are two instances of one deeper structural fact, or two superficially similar facts with quite different underlying reasons, is genuinely open. I find the parallel suggestive enough to motivate a foundational framework that takes time, accumulation, and the cost of working things out as primitives rather than as derived phenomena. But I am not claiming a unification, and I want to flag explicitly that calling these "the same fact" — as I was tempted to do in earlier drafts of this argument — would be one of those moves that sounds illuminating while actually being a verbal coincidence rather than a structural identity.

The dissatisfaction this strand expresses, then, is more modest than the parallel might suggest at first reading. It is that the time-cost of computation and the time-asymmetry of physical processes are both phenomena where *something* takes time even though the information is, in some sense, already there. We have no foundational framework in which these are even potentially the same kind of fact. Whether they should be is an open question; Branch Theory is one attempt to provide a setting in which the question could be sharpened.

This is what Branch Theory tries to give a setting for. The first two strands were dissatisfactions with how foundational physics handles its deepest puzzles. The third strand is the suggestion that those puzzles might have a common shape — a shape that is at root about the relationship between finite descriptions, unbounded possibility spaces, and the time it takes to navigate between them. The framework proposed in Part II takes that suggestion seriously, by treating subjects' growing records (which take time to grow, one observation at a time) as foundational, and by treating physical processes as constraints on which growths are admissible.

Part II — The framework

This part of the document states Branch Theory. The aim is to put the framework down cleanly, in a form that can be argued with, before turning in Part III to what it would look like applied to specific cases.

The framework consists of a small number of primitives and a single substantive axiom. From these it tries to derive — or at least make natural — several things that foundational physics usually has to postulate. The derivations are informal in this document; making them rigorous is part of the research program described in Part IV. But even at the informal level, what the framework gets and what it costs should be clear.

Subjects

The first move is to take **subjects**, not objects, as the foundational entities.

A subject, for the purposes of this framework, is a physical system that accumulates evidence over time. The word "subject" is chosen deliberately. It is meant to be broader than "observer," which carries the connotation of a conscious or cognitive entity, and broader than "agent," which suggests deliberate action. A rock is a subject in this framework's minimal sense: it accumulates evidence in its weathering, its thermal history, its accumulated mineral contents. So is a thermometer, a photographic plate, a bacterial colony, a person, a galaxy. What matters is that the system has a physical substrate that carries the marks of what has happened to it. (A small technical point that will be developed below: real subjects also *forget* — they have finite substrates and cannot accumulate evidence without bound. The framework treats real subjects as quotients over an idealized non-forgetting structure called a base-history. The distinction will not matter for the next few sections, but it matters once we get to entropy and identity.)

The contrast with object-based ontologies is meant to be sharp. In an atomistic picture, the world is fundamentally a collection of particles arranged in space and time; observers and observation are secondary phenomena that the framework has to explain after the fact. In Branch Theory, this is inverted. The world is fundamentally a collection of subjects, each accumulating its own evidence, and "objects" — the things we normally think of as the contents of physical reality — are what shows up in the structure of subjects' accumulated evidence about one another.

This sounds extreme, and one might worry that it puts the framework on a collision course with realism about the physical world. It doesn't. The physical world is not denied by Branch Theory; it is relocated. Instead of being a substrate that subjects look at, the physical world is the *structure of mutual evidence among subjects.* This will be made more precise below, when we get to inter-subject coherence. For now, the important point is that subjects are taken as primitive and the rest is to be built up from how they relate.

The intellectual neighbors of this move are several. Alfred North Whitehead's process philosophy takes "actual occasions" of experience as primitive units of reality, and Branch Theory's subject-moments (introduced below) have an obvious family resemblance. Carlo Rovelli's relational interpretation of quantum mechanics insists that the state of a physical system is only defined relative to another physical system; Branch Theory generalizes this to a thoroughgoing relational ontology that doesn't depend on quantum mechanics for its motivation. QBism (Christopher Fuchs and colleagues) treats the quantum state as an agent's beliefs and observation as Bayesian updating, locating the agent at the foundation; Branch Theory does something similar but for "subjects" in the broader sense, and without the specifically Bayesian commitments. Wheeler's "It from Bit" program suspected that information, not substance, might be foundational; Branch Theory is one possible way of making that suspicion structural.

The point of mentioning the neighbors is to be honest about what's borrowed and what's distinctive. The move to subject-centric ontology is not new. What Branch Theory tries to add is a specific structural account of how subjects' realities are organized — the type-asymmetry that gives time its arrow, the no-contradiction principle that enforces coherence within each subject, the channel-based account of inter-subject objectivity, the two layers of quotient that handle entropy and personal identity. These specifics are the framework's contribution, and they are stated in the rest of this part.

Subject-moments

A subject is not a static thing; it is a sequence. The natural unit of analysis is the **subject-moment**: a particular state of a particular subject at a particular point in its development.

Each subject-moment has two sides. On one side is everything that has been accumulated up to that moment — the observations the subject has made, the evidence its substrate carries. On the other side is everything that is still open — the future the subject has not yet accumulated.

This division is sharper than it sounds, and it does crucial work in the framework. The past, from the perspective of a subject-moment, is *contained*: the subject's substrate holds the evidence of what has happened. The future, from the same perspective, is *received*: it is what will arrive when further evidence is accumulated. These are not two ends of a symmetric axis. They are different kinds of relation to the subject-moment. Containment is a structural relation between the moment and something already inside it; reception is the relation between the moment and something that will be added to it. The two are not interchangeable.

There is a more concrete way to put the same distinction, borrowed from the structure of Boolean reasoning. The past, from within a subject-moment, has the structure of an **AND-conjunction**: it is everything that has been observed, held simultaneously, all of it part of the substrate at once. Observation A and observation B and observation C and... — every accumulated observation contributing to the joint state. The future, by contrast, has the structure of an **OR-disjunction**: it is the set of consistent extensions, any one of which might be the next observation, but at most one of which actually will be. Continuation X or continuation Y or continuation Z. *Time, in this framework, is the rewriting of OR into AND* — the gradual replacement of open disjunctions by settled conjunctions as observations are added to the substrate. Each subject-moment is the boundary, at that point in the subject's history, between the conjunctive past and the disjunctive future. The arrow of time is the direction of this rewriting.

This characterization is not just rhetorical. It connects directly to the third dissatisfaction from Part I: computation takes time precisely because converting OR-structures into AND-structures is what computation is, and that conversion has a structural cost. Subjects, in this framework, are the objects that *perform* this conversion — physical systems whose substrates accumulate AND-content as OR-possibilities resolve. This is what makes the framework a setting in which "why does computation take time" and "why does time have a direction" might be facets of the same question rather than questions in unrelated fields.

This is where time's arrow lives in the framework, and it is worth being honest about what kind of move this is. The standard puzzle of time-asymmetry assumes a time axis with two directions and asks why one of them is privileged. The framework rejects the symmetric axis as a starting point — but this is, frankly, *trading one asymmetric primitive for another.* Boltzmann's statistical mechanics works in a time-symmetric setting and has to import asymmetry through an auxiliary assumption (the Stosszahlansatz). The Past Hypothesis works in a time-symmetric setting and has to import asymmetry through a boundary condition (the universe began low-entropy, for reasons we don't know). Branch Theory works in an explicitly asymmetric setting: subjects are accumulators, and accumulation has a direction.

These are all asymmetric primitives. The framework does not dissolve the puzzle of time-asymmetry; it trades the Past Hypothesis for what we might call an Accumulation Postulate. The reader is owed an honest accounting of why this trade is worth making rather than a claim that the puzzle has been solved.

The trade is, I think, defensible for two reasons. First, the Accumulation Postulate is conceptually closer to the structure of what subjects are than to the empirical contents of the universe. The Past Hypothesis says something about the actual cosmological history — a contingent fact that might in principle have been otherwise. The Accumulation Postulate says something about what it means to be a subject at all — closer to a definition than to an empirical claim. Trading a contingent fact for a structural one is the kind of trade that, while not dissolving the puzzle, may put it on more tractable footing. Second, the Accumulation Postulate is more closely tied to the *phenomenology* of being a subject — the fact that the past is contained and the future is received — than the Past Hypothesis is. The trade aligns the foundational asymmetry with the experiential asymmetry, which is what makes time *feel* asymmetric.

Neither of these is a knock-down argument for the trade. What they are is reasons to find it worth pursuing rather than dismissing. A reader who finds the Past Hypothesis less obnoxious than the Accumulation Postulate is making a substantive metaphysical choice, and the framework should not pretend otherwise.

Monotone accumulation

A consequence of the above is that, at the most basic level of the framework, records only grow.

If the past is what the substrate already contains, and the future is what gets added to it, then time at this base level is the process of accumulation. There is no operation that *removes* content from a substrate at this level. The substrate gains content as evidence accumulates; it does not lose content.

This is going to sound wrong, because real subjects forget. Humans forget. Computers lose data. Memories degrade. The framework will account for all of these, but at a higher level than the base. At the base level, forgetting does not exist. To "forget" would be to revert a record to an earlier state, which would be to go backwards in time, and the framework does not allow this. Forgetting, as a real phenomenon, will be reintroduced in the next sub-section as a feature of higher-layer quotients over the base.

Because the base level is non-forgetting, no real physical thing exists at the base level. Any finite physical system, persisting across enough time, must forget — its substrate has bounded capacity, and accumulating evidence indefinitely is not possible. So the base level is best thought of as an idealization: not a description of any actual thing, but the underlying record of which actual things are quotients. We will call a complete non-forgetting record at this level a **base-history**. A base-history is what a subject's substrate would be if the subject never forgot — an idealized object, not a real entity. Real subjects, the ones that actually exist in the world, live at higher layers, where forgetting is reintroduced through the quotients described next.

Quotients: where forgetting lives

Two kinds of higher-layer structure are needed to recover the framework's connection to real subjects.

The first is the **identity-quotient**. Real subjects persist across substrate change. A human body turns over its atoms several times over its lifetime. A river is "the same river" across decades despite no water molecule remaining. A corporation persists across employee turnover. A national identity persists across the death of every person who carried it a century ago. In none of these cases is the persistent entity the literal substrate. The persistent entity is an equivalence class — a quotient — over the trajectories of substrates that count as that entity.

In Branch Theory, every real subject is one of these quotients. The substrate is the physical carrier; the subject is what counts as "the same subject" across substrate change. The quotient does not bring along all the substrate's information. When old atoms leave a body, the information they carried in their specific positions and bonds leaves with them. From the perspective of the persistent subject, that information is gone. **This is what forgetting is, at the structural level: information that lives in the substrate but does not survive the identity-quotient.**

The second kind of higher-layer structure is the **state-quotient**. Even for a fixed identity-quotient (a fixed subject), there is usually no need or possibility to distinguish every detailed substrate-state from every other. We identify states that "amount to the same thing" by some criterion. Two configurations of gas molecules with the same macroscopic temperature and pressure count as "the same state" of the gas, even though they differ at the molecular level. The state-quotient identifies which microscopic details count as relevant and which are coarse-grained away.

State-quotients are where entropy lives in the framework. The entropy of a higher-layer state is the logarithm of the number of base-histories that the quotient identifies with that state — Boltzmann's $S = k \log W$, with W given the specific meaning of "number of base-histories collapsed into this state by the chosen quotient." The standard observation that entropy depends on the choice of coarse-graining, which in conventional statistical mechanics is usually treated as an awkward feature to be minimized, is in this framework a structural fact: entropy is a property of the *quotient*, and the quotient is chosen by whoever is doing the modeling.

The Second Law, in this framework, becomes the statement that fibers of the quotient — the equivalence classes of base-histories mapped to each higher-layer state — tend to grow as the base-history grows. As the substrate accumulates more evidence, there are more base-histories compatible with any given coarse-grained state, simply because there is more accumulated history to be agnostic about. This is close to the standard phase-space-volume argument, but with the specific measure given by counting base-histories rather than by an externally chosen Liouville measure, and with the directional structure built in from the type-asymmetry rather than imposed by hand.

Real persistent subjects sit at the intersection of these two quotients. They are identity-quotients over substrate trajectories, and at any given moment their state is described at the level of a state-quotient. Both quotients forget. The identity-quotient forgets the substrate-level details that don't survive substrate change; the state-quotient forgets the within-substrate details that don't survive coarse-graining. Real subjects are necessarily forgetful, because they are necessarily quotients.

There is a further point that this clarifies. A persistent subject does not just forget *in spite of itself*; it *must* forget in order to persist. A subject that retained every base-level detail of its substrate would not be a coherent higher-layer entity at all — it would just be its own substrate, indistinguishable from the underlying physics. The very fact of being a higher-layer subject requires the quotient, and the quotient requires forgetting. This is where the framework connects, structurally, to the entropy of life. Living things are exactly the persistent identity-quotients that maintain themselves across constant substrate turnover. The entropic dissipation involved in being alive is not a side effect; it is what being alive consists of, at the substrate level. To stop dissipating substrate-level information would be to stop being a higher-layer subject.

No-contradiction at the physical level

The framework's one substantive axiom concerns what kinds of evidence a substrate can hold.

The axiom is: **a subject's substrate cannot contain mutually contradictory evidence.** A given coin cannot leave evidence of having landed both heads and tails. A given photon cannot leave evidence of having been both absorbed and not absorbed. The physical accumulation of evidence is internally consistent, because there is only one physical reality being recorded by the substrate.

This is a strong-sounding axiom and it does substantial work in the framework, but I need to be careful about where its content actually lives. A reader noticing the word "contradictory" might worry that the axiom is doing nothing more than restating the law of non-contradiction at a different level — that "the world doesn't contain contradictions" is true by definition of what "contradiction" means. I think that worry is partly right and partly wrong, and being clear about which is which matters for understanding what the framework is and isn't claiming.

The axiom has *structural* content: it says that any physical record about a given physical happening must be jointly consistent. But the axiom's *bite* — its capacity to actually forbid specific things — only kicks in once we have an answer to the question *which records count as mutually incompatible?*. Two coin readings of "heads" and "tails" are incompatible. But what makes them incompatible is a fact about coins and how their substrate-states constrain each other — not a fact about the abstract structure of contradiction. The incompatibility relation that gives the axiom its content has to come from physics. It is not supplied by the framework.

This means that no-contradiction, on its own, is closer to a *form* than to a *substantive principle.*

It says: *whatever* physics tells you counts as incompatible evidence, the framework requires that substrates not jointly carry it. The framework provides the structural shape; physics supplies the content. This is consistent with calling Branch Theory a meta-framework — it organizes how foundational claims are stated, without making the foundational claims itself.

I take this to be a feature rather than a weakness. The framework does not need to compete with physics by supplying its own theory of what is and isn't compatible. It needs to provide a setting in which whatever incompatibility relation physics supplies can do its work, with the right structural consequences (linearity of histories, channel-based coherence, indexical many-worlds). That is what the framework attempts, and the reader should evaluate the axiom in that light — as a structural placeholder for content that physics provides, not as an independent generator of physical predictions.

With that clarification, three further notes about what the axiom does and does not constrain:

It constrains substrates. It does not constrain cognition. A human being can hold contradictory beliefs; this happens routinely, and is not a violation of the axiom. Cognition is a phenomenon that arises at higher layers in some subjects (those with appropriate biological or computational machinery), and cognitive states are subject to their own constraints, none of which are foundational to the framework. The axiom applies at the level of physical evidence, not at the level of beliefs about that evidence.

It constrains within a single substrate. It does not directly constrain between substrates. Subject A's substrate must be internally consistent. Subject B's substrate must be internally consistent. But A's substrate and B's substrate are independent objects, and the axiom does not, by itself, force them to agree about anything. Inter-subject agreement will emerge below, through the structure of evidence channels, but it is a derived consequence of the axiom applied within each substrate that contains evidence about the other, not a separate axiom.

It refers to evidence, not to truth or reference or meaning. The axiom is about what a physical substrate can carry, not about what statements about it can be true. There is no semantic content to the axiom; it is about the structural fact that a physical record cannot encode two incompatible facts (where "incompatible" is determined by physics) about the same physical happening, because the happening was one happening.

Actions as substrate-constraints

The framework has so far described subjects as accumulators of evidence — entities that *receive* from the future and *carry* into the present what has happened. But subjects also *act*. A Turing machine writes to its tape. An organism moves through its environment. Maxwell's demon opens and closes a partition. Any framework that takes subjects seriously has to account for what action is.

The temptation, given the framework's symmetry-breaking move on time, is to introduce actions as a separate primitive — a counterpart to observations, going out into the future as observations come in from it. This temptation should be resisted. Actions and observations are not symmetric in the way that would make them two sides of a single primitive. They are the same kind of thing, viewed from different sides of the same relation.

The cleaner way to say it is this. A subject's substrate, at any moment, has a specific physical configuration. That configuration is part of the world. Whatever happens next must be consistent with it. The substrate's current state therefore *constrains the consistent extensions of everything it is in contact with.* When the configuration changes — when a Turing machine's head writes "1" to a tape cell, when a hand grips a tool, when a demon's partition closes — the constraints that configuration imposes on the rest of the world change. Those changes are what actions are.

In this view, **an action is not an event added to a substrate alongside its observations; it is the substrate's current state, considered in its capacity to constrain other substrates' possible extensions**. The same physical fact — the head writing "1" — is, from the machine's side, a state of its own substrate; from the tape's side, a constraint on what the tape's state can be in the next moment. Action and observation are the same evidence-channel relation, distinguished only by which substrate is doing the constraining and which is being constrained.

This dissolves a small technical worry and replaces it with a clean structural picture. There is no need for the framework to add "outputs" alongside "inputs." Inputs and outputs are the same kind of thing: substrate-states constraining other substrate-states through whatever physical channels of evidence connect them. A subject acts on its environment by being in the kinds of states that force the environment to particular extensions. A subject observes its environment by having its own substrate forced into particular extensions by environmental states. The arrow of constraint can run either way along any channel; what matters is the network of mutual constraints among substrates, not a separate ledger of inputs and outputs.

Activity and alignment

Two consequences of this view are worth stating explicitly, because they bear on what subjects can do in the world and how the framework connects to thermodynamics.

The first is a clean structural definition of **activity**. A subject is *active*, in this framework's sense, to the degree that its substrate-state strongly constrains the substrates around it. Activity is meant in a deliberately minimal sense — it is a measure of constraint exerted, with no implication of goal-directedness, deliberateness, or any cognitive feature. A rock rolling down a hill is highly active in this sense: its substrate-state strongly constrains what the substrates of nearby things can be in the next moment (whatever it collides with, whatever it disturbs). A fire is highly active. A Turing machine head writing to its tape is highly active. A human is highly active, in ways that include but are not limited to deliberate ones. The term "agency," in its ordinary sense, refers to a special subclass of activity that involves goal-directedness and internal processing; the framework needs the broader category and treats agency as a downstream phenomenon that some active subjects exhibit. Activity is the structural primitive; agency, where it occurs, is one way activity gets organized.

What this definition supplies is a vocabulary, not an account. Calling a rock, a fire, and a Turing machine all "highly active" identifies a structural feature they share — strong constraint imposed on surrounding substrates — without distinguishing between them in any of the ways that matter for understanding what each actually does. The vocabulary is useful insofar as it lets us talk about active subjects without smuggling in cognition; it is not useful as a substitute for the specific accounts (mechanics, combustion chemistry, computation theory) that explain what specific active subjects do. The framework's contribution here is structural, not explanatory.

The second is an account of **alignment**. When a subject takes substrate-configurations that correlate in some structured way across time, the cumulative constraint on surrounding substrates is a narrowing of their possible extensions in a corresponding direction. Maxwell's demon, opening and closing the partition based on observed molecular velocities, cumulatively narrows the gas's distribution: fast molecules end up on one side, slow ones on the other. A river carving a channel through soft rock, by being repeatedly in the configuration of a river, narrows the rock's possible configurations into a channel-shape. A human organizing a workshop, repeatedly placing tools in patterned ways, narrows the workshop's configuration over time. Alignment is this cumulative narrowing — the process by which a subject's correlated substrate-states force structured order into the surrounding world.

Alignment is not free, in this framework. The subject's substrate must repeatedly assume the correlated configurations, and maintaining those configurations across substrate turnover requires the identity-quotient to hold up over time. This is where the thermodynamic cost lives. The substrate-level dissipation required to keep the subject coherent across turnover is what pays for the alignment it produces. The qualitative shape of Bennett's resolution of the Maxwell's demon paradox — that the demon must eventually dissipate the information it accumulates — fits naturally into this picture: the demon is a persistent identity-quotient whose maintenance has a substrate-level cost that offsets the order it imposes on the gas. I want to be honest about how strong this is, though: recovering the *qualitative* shape (maintenance costs dissipation) is close to definitional given the framework's primitives — once we have said that persistent subjects are quotients that must shed substrate to persist, we are not far from "maintenance costs dissipation" as a near-tautology. The substantive part of Bennett's result is its *quantitative* content (the specific kT ln 2 per erased bit), and the framework does not recover that without additional work. So what the framework offers here is a structural setting in which the qualitative shape of Bennett's resolution looks natural rather than surprising; the quantitative content remains as a research target rather than a recovered result.

This gives the framework what it needs to talk about Turing machines, organisms, demons, rivers, and any other active subjects. Part III will draw on this in the worked sketches.

Linearity from no-contradiction

A first consequence of the axiom, combined with monotone accumulation, is that subject-histories are linear.

The reasoning runs as follows. Consider a subject at a particular moment, with its accumulated record. Suppose two different next-observations are possible, and that they are *incompatible* in the sense that no single substrate could carry evidence of both (one says the coin landed heads, the other says tails). Both extensions of the record are formally possible, but only one of them can be added to *this* substrate. If both were added, the substrate would contain mutually contradictory evidence, which the axiom forbids.

So the subject's record can only extend in one direction at a time, in the sense that any given substrate goes only one way. This is what makes time linear from the perspective of a single subject. The linearity is not postulated; it is a consequence of monotone accumulation plus no-contradiction. A single substrate's history is, by structural necessity, a chain — not because we put it in a chain by hand, but because the alternative would be a substrate carrying contradictions.

Many worlds, indexically

What about the other extensions, the ones not chosen? Are they real?

The framework's answer is: yes, but each in its own subject-history. The world does not contain a single substrate that carries all possible extensions; it contains, in effect, a forest of subject-histories, each of which is a linear chain by the previous argument. Each chain represents one consistent way the substrate could have extended; all the consistent ways are equally real.

This is many-worlds, but reached from a different direction than Everett's. Everett's many-worlds interpretation of quantum mechanics derives branching from the linearity of the Schrödinger equation applied to the universal wavefunction; branching is a quantum-mechanical phenomenon, and the substantive content of the program (decoherence, the preferred basis, branch weights) follows from working through the consequences of the Schrödinger equation in detail. Branch Theory's branching has a different starting point: it is a structural consequence of subjects being substrates that accumulate evidence under no-contradiction, and so it exists as a framework-level fact independent of any specific physical theory. But I should be honest about what this buys and what it does not. What falls out of the framework is essentially "different consistent extensions exist and are equally real" — a structural shape, not a derivation of the substantive content of Everettian physics. The decoherence-driven emergence of branches, the preferred basis problem, the structure of branch weights — none of this is supplied by Branch Theory. The framework provides a *setting* in which many-worlds-style branching is natural rather than exceptional, but the physics that actually populates that setting in the quantum case comes from elsewhere. Branch Theory is broader than Everett's interpretation in the sense that it does not require quantum mechanics for its motivation; it is also shallower in the sense that it does not derive any of the substantive content that makes Everettian physics distinctive.

The question "which branch is the actual one" is, in Branch Theory, a malformed question. All consistent branches are real. From within any subject-moment, the moment's own past is uniquely determined (it is what the substrate contains), and that past is the *actual* past *for that moment.* Other branches have other subject-moments, each of which would similarly experience its own past as uniquely actual. "Actual" is indexical — it picks out the chain leading to the moment doing the picking — and there is no view from nowhere that singles out one chain as the privileged one.

The framework therefore takes the indexical horn of the many-worlds debate. It does not try to derive Born-rule probabilities or any other measure on the branches. The question of whether a measure can be derived is left open as a research target (Part IV). What the framework claims is that the apparent uniqueness of the actual history is a perspectival fact about being a subject-moment, not a metaphysical fact about reality, and that the measurement problem in quantum mechanics is therefore not solved by Branch Theory so much as dissolved — it becomes a special case of a more general fact about the structure of subjecthood.

Inter-subject coherence and the shared world

The final structural element of the framework concerns how subjects relate to one another.

Subjects are independent. Each has its own substrate, its own record, its own quotients. The framework does not assume a shared world underlying all subjects. There is no common substrate that all subjects look at.

What relates subjects is **evidence**. When something happens, it leaves traces. A photon emitted by an event can be absorbed by subject A. The same event, or causally related events, can leave traces that reach subject B. When A's substrate contains a trace of something that also reached B, A's substrate carries evidence that constrains what B's substrate could consistently be.

This is the heart of inter-subject coherence in Branch Theory, but a careful point matters here. The axiom of no-contradiction applies *within* each substrate. It does not, by itself, force two independent substrates to agree about anything. Subject A's substrate must be internally consistent; subject B's substrate must be internally consistent; if they are independent, that is the end of the constraint. The coherence between them kicks in when their records *meet* — when some substrate (A's, B's, or a third subject's) comes to contain evidence of both. At that point, the substrate doing the containing must be internally consistent about all the records it carries, and any apparent inconsistency must be resolved within that substrate, either by treating one record as malfunctioning, by reinterpreting the original events, or by inferring additional facts that reconcile the records.

The "shared world" is, in this framework, the **invariant content of the channel structure among subjects**, built up as evidence accumulates in substrates that combine records from multiple sources. Where subjects have not yet brought their records into contact, their substrates may carry merely correlated records (records sharing a cause) that have not yet been forced into mutual coherence. Coherence is not a metaphysical fiat at the moment of the shared event; it is what gets built up in the substrates that subsequently combine the records.

This has the consequence that objectivity is graded by interaction. Heavily-communicating subjects share a strong world: their substrates contain extensive mutual evidence, and no-contradiction forces strong mutual coherence. Isolated subjects share no world: their substrates have no relation, and they could disagree about everything without anyone's substrate being incoherent. The "world" that subjects collectively inhabit is not an underlying substrate they all look at; it is the residual structure of what their mutual evidence forces to be invariant.

Reports are a special case of evidence. When subject B sends a signal to subject A — a spoken word, an electromagnetic transmission, a written note — the signal is a physical event that leaves traces in A's substrate. A's substrate then contains evidence of the signal, which constrains what A can take B to have meant or known. Reports may be inaccurate, deliberately misleading, or simply mistaken; the framework does not assume reports are reliable. But it does treat the *fact of the report* — the physical signal A received — as evidence that imposes its own consistency constraints on A's record. The unreliability of reports does not exempt them from being evidence; it just means that A's record may contain "B reported X" without containing "X is true," and the two pieces of evidence have to be jointly consistent in whatever way A's record settles.

Where the framework stands

That is the framework, stated as cleanly as I can manage in this length.

To recapitulate:

- The foundational entities are subjects: physical systems whose substrates accumulate evidence.
- Subject-moments are the units of analysis. Each has a contained past (substrate content) and a received future (extensions to come). This type-asymmetry is where time's arrow lives in the framework — not as a derived consequence but as a primitive trade for the Past Hypothesis, defensible because the Accumulation Postulate is closer to structural fact than to contingent cosmology.
- At the base level (the idealization of non-forgetting accumulation), records only grow. We call a complete such record a *base-history*. Base-histories are idealized objects; real subjects are quotients over them.
- Real subjects are quotients — identity-quotients over substrate trajectories (enabling persistence across substrate change) and state-quotients over within-substrate detail (enabling coarse-grained description). Both quotients forget. Real subjects are necessarily forgetful.
- The Second Law's qualitative shape — that maintaining order has an entropic cost — fits naturally into the framework as the structural cost of maintaining identity-quotients across substrate turnover. The quantitative content of the Second Law is not derived; only its qualitative shape is recovered, and that shape is close to definitional given the framework's primitives.
- The one substantive axiom is no-contradiction at the physical level: a substrate cannot carry mutually incompatible evidence. The axiom has structural content; its bite — the question of which records count as incompatible — has to come from physics. The framework provides the form; physics provides the content.
- Actions are not a separate primitive: they are substrate-states considered in their capacity to constrain other substrates' possible extensions. Activity is a graded structural vocabulary — the strength with which a subject constrains its surroundings, with no implication of goal-directedness. The vocabulary identifies a feature shared by rocks, fires, and machines; it is not a substitute for the specific accounts of what each does. Alignment is the cumulative narrowing of surrounding substrates produced by a subject's correlated substrate-configurations across time.
- Linearity of subject-histories follows from no-contradiction plus monotone accumulation, given a specified incompatibility relation.
- All consistent extensions of any subject-moment are real, forming a forest of subject-histories. The selection of "the actual history" is indexical, not metaphysical. Many-worlds-style branching is the framework's natural reading; the framework provides a setting for it but does not derive the substantive content of Everettian physics (decoherence, preferred basis, branch weights).
- Inter-subject coherence emerges from no-contradiction applied to evidence channels between subjects. The shared world is the invariant content of these channels, graded by interaction.

The framework is a meta-theory. It does not specify what observations are possible, what physical laws govern their relationships, or what conservation principles hold. These are content to be imported. The framework specifies the form that any such content must take when expressed within it.

Part III takes a handful of familiar structures — a Turing machine, an ideal gas, Maxwell's demon, a multi-detector measurement, a logical derivation — and sketches what they look like in Branch Theory terms. This is where the framework either earns its keep or shows its limits.

Part III — Worked sketches

This part of the document takes five familiar structures and re-describes them in the framework of Part II. The structures are a Turing machine, an ideal gas, Maxwell's demon (built on top of the gas sketch), an event observed by multiple detectors, and a logical derivation.

The point of the exercise is not to derive anything new about these structures. They are well-understood in their conventional forms. The point is to see what they look like through the framework's primitives — what is illuminated, what is strained, and where the gaps are. Each sketch ends with an honest note on what would need to be done to make the re-description rigorous, and where the framework either gains or loses something in the translation.

The sketches are short by design. They are sketches, not derivations. They are meant to be enough to evaluate whether the framework has any inferential traction beyond philosophical reframing.

Sketch 1: A Turing machine as a subject

A Turing machine consists of an infinite tape divided into cells, each containing a symbol; a head that reads and writes one cell at a time; a finite state register; and a transition function that, given the current state and the symbol under the head, prescribes a new state, a new symbol to write, and a direction to move.

In Branch Theory terms, the most natural reading is to take the *head* as the subject and the *tape* as part of the environment. The head, considered as a subject in the framework's full sense (real, forgetful, an identity-quotient maintained across the steps of computation), has a finite substrate: its state register plus its current position. Reading a cell is observation in the framework's proper sense — evidence acquired from outside the subject. Writing a cell is action — the head's substrate-state constraining what the corresponding cell of the environment's substrate can be.

The head, in this reading, is a subject in the framework's full real-world sense: forgetful, an identity-quotient maintained across the steps of computation, over an idealized base-history that would record every step ever taken. At any moment the head's substrate carries only its current state and position, not the history that led there. The head's identity across computational steps is what counts as "the same head" — the equivalence relation under which we identify the head before and after each transition — and that identity is a quotient over the (idealized, unbounded) base-history of the entire computation.

So: the head is a subject whose identity-quotient drastically forgets. At any subject-moment, the head's substrate-state is determined by the transition function applied to the previous state and the cell just read. Its previous states are gone — the head does not carry a record of which states it has been in. The only memory the head has of past observations is whatever the transition function has chosen to encode into its current state.

This drastic forgetting is precisely what makes the head a useful subject to think about. It illustrates a general structural point: a subject's current state is constrained by its past observations only insofar as the identity-quotient preserves that constraint. Past readings constrain the head's current state if and only if the transition function has propagated their effect forward.

Information that the transition function does not propagate is gone — not just inaccessible, but absent from the subject's substrate altogether.

The tape, in this reading, is environment. The tape's substrate is independent of the head's substrate; the two interact only at the cell currently under the head, which is the channel of observation and action between them. Writing imposes a constraint on the tape's substrate (that cell's value is now fixed). Reading is the head's substrate being shaped by what the tape's substrate already carries. This is the standard input/output picture, restated in the framework's terms.

What does the framework illuminate? The asymmetry between how much the head can know about itself and about the tape is structural. The head knows everything currently in its state (because that's what its substrate is); it knows about the tape only what the transition function has caused to be encoded into its state from past readings. The head can read the same cell twice and discover something it had previously known and forgotten — its substrate at the second reading is reshaped by the cell's content even though, in a strong sense, that content was already known earlier. This is what forgetting looks like from inside: the subject is genuinely surprised by what it could have remembered.

It is also useful to consider the *combined* system of head and tape. This is a different subject — call it the *whole machine* — with a much larger substrate that includes all cells the head has visited. The whole machine forgets less than the head alone, because the tape preserves writes that the head's state register cannot. But the whole machine still forgets, in a more subtle way: it does not carry a record of *which order* states were visited in. Two different computational histories that end in the same head-state at the same position with the same tape contents are, from the whole-machine subject's perspective, the same state. The trajectory that led there is lost.

This subtle forgetting has a precise consequence. If the whole machine ever returns to a state it has been in before — the same head-state, the same position, the same tape contents — then by the determinism of the transition function it will repeat the same evolution from that point, and so it will return to that state again, and again. The machine is in an infinite loop and will not halt. Conversely, if the machine halts, it must never have revisited any state. **Halting and non-revisiting are equivalent.** The undecidability of the halting problem is the undecidability, given a machine and an initial tape, of whether the machine's substrate-evolution will ever revisit a state — which in turn is the question of whether the whole machine's substrate-history is eventually periodic.

What is strained? The Turing machine has no environment outside its tape, and even the tape is a passive environment that does not have its own subject-life. The framework's machinery for inter-subject coherence is therefore inert here. The Turing-machine sketch shows the within-subject part of the framework (subjects, substrates, observation, action, forgetting, identity-quotients) but not the channel-based part. The multi-detector sketch below will exercise that.

What would need to be made rigorous? The identification of the head's identity-quotient with the conventional notion of "the same Turing machine across steps" is intuitive but worth working out formally. So is the precise correspondence between "the whole machine revisits a state" and the framework's account of looping. Both are clean targets for a formal treatment.

Sketch 2: A container of ideal gas

An ideal gas in a container is a collection of N identical particles bouncing elastically off the walls and one another, with negligible interaction except at collisions. Statistical mechanics describes it at two levels: a microscopic level (the exact positions and velocities of all N particles) and a macroscopic level (temperature, pressure, volume), with the macroscopic level related to the microscopic level by an averaging procedure.

In Branch Theory terms, the question of what counts as "the subject" has two readings, and the framework illuminates the choice.

The first reading is to take *each gas particle as a subject.* Each particle has a substrate (its physical configuration: position, momentum, identity); each particle's substrate accumulates evidence of what it has encountered (the walls it has hit, the other particles it has collided with). Particles constrain one another through collisions; their substrate-states impose substantial constraints on one another's next configurations. The framework's machinery for inter-subject coherence applies: the gas is a community of subjects in heavy mutual evidence, forced into mutual consistency by no-contradiction applied to each particle's substrate.

The second reading is to take *the container of gas as a single subject.* Its substrate is the joint configuration of all N particles plus the walls; its accumulated record is the full history of the gas's microscopic evolution; its future is the set of consistent extensions of that history.

Both readings are admissible, and the framework treats them as different quotient-choices over the same underlying base. The single-subject reading is what you get if you take a very coarse identity-quotient that identifies the whole container at every moment. The many-subjects reading is what you get if you take a fine identity-quotient that keeps each particle distinct.

Where does temperature live? At the state-quotient layer. Temperature, in this framework, is a feature of the state-quotient over the container's microscopic configurations — specifically, the quotient that identifies microscopic configurations with the same average kinetic energy. Two microscopically different configurations with the same average kinetic energy are mapped to the same coarse-grained state by this quotient. The temperature *is* the label of that coarse-grained state. There is no temperature at the base level; particles do not have temperature, only velocity. Temperature is a property of the quotient, which is a modeling choice.

This makes the standard Gibbs-vs-Boltzmann ambiguity in entropy explicit. Different quotients give different entropies. The Boltzmann entropy (log of the number of microstates in a given macrostate) and the Gibbs entropy (the integral over the phase-space distribution) correspond to different choices of how to coarse-grain. Both are admissible; both are entropies in the framework's sense. The framework does not adjudicate which is "really" the entropy, because there is no quotient-independent fact about which one is right. Different observers choose different quotients and get different entropies.

The Second Law, in this picture, is the statement that fibers of the state-quotient — sets of microstates mapped to the same macrostate — tend to grow as the system's base history grows. As particles collide and exchange energy, the microstates compatible with a given macrostate proliferate, because more histories can lead to the same coarse-grained outcome. This is the standard phase-space-volume argument with the specific measure given by base-history counting, and it inherits the standard argument's structure.

What does the framework illuminate, and what does it leave untouched? The choice of quotient becomes explicit rather than hidden, which is something — the framework makes a modeling choice that conventional treatments tend to smuggle in. But this is more modest than it might first sound. The conventional objection to statistical mechanics is not that the choice of coarse-graining exists; it is that *some* coarse-grainings track real macroscopic regularities (temperature, pressure) while others do not, and there is no obvious principled basis for the distinction. Making the choice explicit does not resolve this; it merely relocates it. The framework, on its own, has no criterion for which coarse-grainings are useful, and offers no derivation of why temperature happens to be one of the useful ones rather than some arbitrary alternative quotient. This is exactly the standard problem, now stated in different vocabulary.

What is strained, then, is precisely what we just acknowledged: the framework does not, on its own, tell you which coarse-graining is the *useful* one. Connecting the framework to the empirical fact that some quotients track real regularities and others don't would require additional structure — possibly the recognition that the useful quotients are the ones that respect the conservation laws and channel structures imposed by physics. This is a research target rather than something the framework already supplies, and one should be cautious about claiming the framework "illuminates" coarse-graining when what it actually does is restate the puzzle in its own terms.

What would need to be made rigorous? The relationship between the framework's quotient-based account of entropy and the standard formal definitions (Boltzmann's S = k log W, Gibbs' integral) would need to be worked out explicitly. The expectation is that these match — that the framework's "log of fiber size" reduces to the standard expressions under appropriate choices of quotient — but verifying this is a substantive calculation that has not been done.

Sketch 3: Maxwell's demon

Maxwell's demon is the celebrated thought experiment in which a small intelligent being controls a partition between two compartments of gas, opening the partition to let fast molecules through in one direction and slow molecules through in the other, eventually creating a temperature gradient where there was none. The demon appears to violate the Second Law: it produces order without expending energy.

The standard resolution, developed over a century of debate and brought to its modern form by Bennett, is that the demon does expend something: information. The demon must accumulate records of which molecules to let through, and when its memory is full it must erase those records to continue operating. Erasure costs energy (by Landauer's principle), and the energetic cost of erasure exactly offsets the order the demon produces. The Second Law survives.

In Branch Theory terms, the demon is an active subject in the sense developed in Part II. Its substrate has two parts: a body that operates the partition, and a memory that accumulates records of molecular states. The body's substrate-state constrains the partition's substrate-state (open or closed); the partition's substrate-state constrains the gas's substrate-evolution (which molecules pass, which are reflected); the gas's substrate-state constrains the demon's memory (which states get recorded). The whole system is a network of substrate-constraints, with the demon as a particularly active node.

The demon's persistence as a subject requires an identity-quotient. The demon at moment t and the demon at moment t+1 are "the same demon" by virtue of some equivalence relation that ignores some substrate-level detail. In particular: across many cycles of operation, the demon's memory cannot grow without bound (the substrate is finite). Old memory must be shed to make room for new memory. This shedding is forgetting at the substrate level: information that was in the demon's substrate is no longer in the demon's substrate.

Where does the shed information go? Into the environment. The substrate that used to carry the demon's memory is recycled, and whatever physical traces it carried are dissipated into the surrounding world. The demon's "forgetting" is the environment's "gaining of randomized substrate." From the demon's perspective, information is lost; from the environment's perspective, entropy is gained.

This is where the framework provides a structural setting for the *qualitative* shape of Bennett's resolution. The alignment the demon produces in the gas — the temperature gradient — is paid for by the dissipation required to maintain the demon's identity-quotient. The substrate-level cost of being a persistent demon is structurally analogous to the substrate-level cost that Bennett identifies with information erasure. But I want to be careful here, because the qualitative shape "maintenance costs dissipation" is close to definitional once we have said that persistent subjects must shed substrate to persist — so claiming that the framework "recovers" Bennett's result risks overstating what is actually a near-tautology given the framework's primitives. What the framework offers is a setting in which Bennett's qualitative conclusion looks structurally natural, not a derivation of its substantive content.

What does the framework illuminate? The connection between *being a persistent subject* and *dissipating entropy into the environment* becomes structural rather than thermodynamic. Living things, demons, computers, and any other persistent active subject all share the same underlying constraint: they cannot maintain themselves across time without dissipating substrate-level information. The Second Law, applied to such subjects, is not an external constraint but a feature of what it is to be such a subject in the first place. This generalizes Bennett's specific result to a structural claim about all persistent active subjects.

What is strained? The quantitative match to Landauer's bound — the specific kT ln 2 per erased bit — is not derived in the framework. The framework says that maintenance costs dissipation; it does not say that the cost is exactly kT ln 2. Recovering the specific quantitative result would require connecting the framework's quotient structure to the thermodynamic parameters of the substrate, which is a substantive piece of work that has not been done. The framework's qualitative result is correct; the quantitative match would need to be earned.

What would need to be made rigorous? Two things, principally. First, a precise statement of what "substrate-level dissipation required to maintain the identity-quotient" means in measurable terms — joules per second, bits per cycle, whatever the appropriate units are. Second, a derivation that this quantity matches Landauer's bound under the appropriate idealizations. Both are non-trivial research tasks, and neither is attempted here.

Sketch 4: An event observed by multiple detectors

A photon is emitted by a source. Several detectors are positioned to record its arrival. Each detector records one aspect of the photon's interaction with it — detector A records timing, detector B records polarization, detector C records energy. After the event, the detectors' records can in principle be compared, if some subject brings them together.

In Branch Theory terms, each detector is a subject with its own substrate, accumulating its own record. The source is a subject whose substrate-state — emitting the photon — constrains the substrate-evolutions of the detectors that subsequently interact with it. The photon, considered as a thing propagating through space, is the channel by which the source's substrate-state imposes constraints on the detectors' substrates.

What does the framework illuminate? The structure of disagreement among detectors becomes clearer than it might at first appear. The framework says, strictly, that *contradictory evidence cannot meet in a single substrate.* So if a combining substrate carries detector A's record and detector B's record, and the two are jointly inconsistent with having been produced by the same event, what has actually happened is one of three things. Either the records were never contradictory at the level of physical evidence — they were records of different actual events, despite appearances. Or the apparent contradiction is at a higher layer — a cognitive misinterpretation, a deceptive report, an error in the labeling of which event each record refers to — and the underlying physical evidence is consistent. Or the records were generated in different branches of the framework's branching forest, and any substrate that purports to carry both is, in fact, carrying records from incompatible branches, which is itself a structural impossibility: the meeting substrate must live in one branch or the other.

This is sharper than the standard intuition. The standard picture says that when measurements disagree, we investigate to find out which was wrong. The framework says: at the level of physical evidence, disagreement cannot occur. Apparent disagreements are diagnostic of one of the three situations above, and what we usually call "investigation" is the process of identifying which.

The framework's channel-coherence machinery has direct implications for one of the central puzzles of quantum mechanics: entanglement. To see how, I need to be careful about what the framework supplies and what it imports — a distinction that, as discussed in Part II, applies generally to how no-contradiction operates and is particularly important here.

Suppose the source emits not one photon but two, in an entangled state — a singlet, say, where the two photons' polarizations are anti-correlated whatever shared axis they are measured along. The fact that the source emits this particular kind of joint state, with this particular constraint on subsequent detector records, is *imported from quantum mechanics.* The framework does not derive that singlet sources exist, or what specific correlations they enforce; it takes those as given physical content.

What the framework does supply is the *structural setting* in which this content has its consequences. Given that the singlet state imposes a particular joint constraint on what records of the two photons can be, the framework's machinery handles the consequences as follows: the source's substrate-state at emission constrains the substrates of *both* photons jointly; each photon, traveling to its detector, carries evidence of the joint origin in its substrate; each detector's record reflects the constraint imposed by the joint origin; when the two detector records subsequently meet in a combining substrate, no-contradiction forces the joint records to be consistent with the joint constraint that originated at the source — which means the records exhibit the correlation characteristic of the entangled state, and joint records inconsistent with that constraint are forbidden.

The structural payoff is that there is no “spooky action at a distance” in this picture, because nothing acts at a distance. The joint constraint is encoded in the substrates from the moment of production, propagated through whatever channels the photons follow, and revealed when the records combine. This is a recognizable cousin of pilot-wave readings of entanglement, where the correlation is established at the source and carried by the particles, rather than created at the moment of measurement. The framework does not derive quantum mechanics, but it provides a setting in which entanglement does not require any additional metaphysical machinery — the channel-coherence machinery that handles ordinary multi-detector observation also handles entangled multi-detector observation, once quantum mechanics has supplied the joint constraint at the source.

Whether this is a useful contribution depends on what one wanted from the framework. If one wanted a derivation of quantum mechanics, the framework does not provide it. If one wanted a structural setting in which entanglement does not need to be a metaphysical exception — in which the same channel-coherence machinery that handles classical multi-observer scenarios extends naturally to entangled ones — then the framework provides that. I take the second to be a meaningful structural result, while explicitly disclaiming the first.

Several other quantum-mechanical phenomena are harder, however, and honesty about which is which matters.

The Born rule — the specific probability amplitudes quantum mechanics assigns to outcomes — is not recovered by the framework as currently stated. No-contradiction tells you which joint records are *possible* and which are *forbidden*, but not how *often* possible records occur in repeated experiments. This is the probability problem in Part IV, and it remains genuinely open. The framework probably needs additional structure to recover Born-rule frequencies.

Contextuality, in the Kochen-Specker sense — the fact that measurement outcomes depend on which other measurements are performed simultaneously — is also harder. The framework may handle it by treating “which measurement is performed” as itself a substrate-fact that constrains which joint records are admissible. Whether the resulting constraint structure has the right shape to recover contextuality’s full content is open and would need to be worked through carefully.

Interference effects, of the double-slit kind, where the absence of distinguishing evidence between possible histories produces observable patterns, may also yield to the framework’s machinery. The relevant “histories” are alternative ways the substrate could have evolved, and interference is the structural fact that, when no substrate carries evidence distinguishing them, the combining substrate’s constraints reflect both. Making this precise is work, and the framework as stated does not do it.

What does the framework illuminate that this sharpens? The status of “measurement” in quantum mechanics becomes clearer. Measurement, in the framework, is not a special interaction that creates definite properties from indefinite ones. It is the production of detector records from substrate interactions, like any other physical process. What is special about quantum measurement is not the act of measuring but the structure of the source's substrate-state — specifically, the kinds of joint constraints it imposes on records that subsequently combine. Entangled sources impose joint constraints; product-state sources impose independent constraints; the difference is in the source, not in the measurement.

What would need to be made rigorous? The connection between the framework's account of source substrate-states and the standard quantum-mechanical apparatus (density matrices, partial trace, reduced states) needs to be worked out explicitly. The expectation is that quantum states are particular kinds of constraints on joint substrate-records, and that the standard quantum formalism is the right mathematical language for expressing those constraints. Confirming this — and characterizing which kinds of constraints correspond to which kinds of quantum states — is a substantial research target. The detector sketch establishes that the framework can handle entanglement in principle; whether it can recover the full quantum formalism is open.

Sketch 5: A logical derivation

A logical derivation is a sequence of statements, each justified by the application of an inference rule to earlier statements, starting from premises and ending at a conclusion.

In Branch Theory terms, the natural reading is to take the derivation as a subject — call it the *reasoner* — whose substrate accumulates the derivation as it proceeds. Each statement, once derived, becomes part of the reasoner's substrate; each inference rule applied is an action of the reasoner that produces a new statement from earlier ones.

A subject-moment of the reasoner is a state of the derivation: the set of statements derived so far, plus the inference rules available. The contained past is the derivation up to that point. The received future is whatever can be derived next.

The framework's no-contradiction axiom applies cleanly. The reasoner's substrate — the accumulated derivation — cannot contain mutually contradictory statements. If at some point both a statement S and its negation (not-S) have been derived, the substrate is inconsistent and the derivation has failed. This corresponds exactly to the standard logical notion of consistency: a derivation is sound only if it never produces both a proposition and its negation. The framework recovers this without postulating it; it follows from the axiom applied at the substrate of the reasoner.

The branching structure of the framework also has a clean correspondence. At any point in a derivation, multiple next-steps are typically possible: different inference rules can be applied, different earlier statements can be combined. Each possible next-step is a consistent extension of the current substrate. All these extensions are real, in the framework's indexical many-worlds sense. The actual derivation is the chain of extensions that the reasoner happens to follow; other chains (other derivations from the same premises) are equally real, just not the one this particular reasoner-substrate is following.

This produces an interesting observation. The space of all consistent derivations from a given set of premises is exactly the branching forest of the reasoner-subject. A theorem — a statement that follows from the premises — is one that appears in *some* branch of the forest. The question of whether a statement is a theorem is the question of whether the forest contains a branch ending at that statement. Provability becomes a structural property of the branching forest.

This connects to the puzzle from Part I about why computation takes time. Determining whether a statement is a theorem is, in general, a search problem: one must explore the branches of the forest until either a branch ending at the statement is found or all branches are exhausted. The information needed to settle the question is contained in the premises and the inference rules — there is no new information generated by the derivation — but extracting that information requires substrate-growth, one step at a time. The time of derivation is the time of substrate-accumulation, exactly as the time of computation was for the Turing machine. The Gödel incompleteness and Church-Turing undecidability results show that this time can be unbounded: there are statements for which no finite branch ending at them exists, even though searching for one might continue indefinitely.

What does the framework illuminate? The parallel between logical derivation and physical accumulation becomes structural. The reasoner is a subject in exactly the framework's sense; the derivation is a substrate-record; the inference rules are the constraints on consistent extensions. Logic and physics, in this view, are not two unrelated domains but two instances of the same structure: substrate accumulation under no-contradiction. This is a striking unification, and it suggests that the framework's primitives may be the right level of abstraction for thinking about both.

What is strained? The framework's account of the reasoner does not distinguish between *constructive* and *non-constructive* logics, between *classical* and *intuitionistic* reasoning, or between logics with different sets of inference rules. These distinctions matter in foundations of mathematics, and the framework as stated does not engage with them. Extending the framework to do so would require articulating what counts as an admissible inference rule — that is, what counts as a substrate-constraint at the logical level — and this is a substantial piece of work.

What would need to be made rigorous? The correspondence between "substrate-extension" and "inference step" needs to be made precise. In particular, it should be possible to translate between standard formal accounts of derivation (sequent calculi, natural deduction, etc.) and the framework's branching-forest picture, and to show that the translations preserve the relevant logical properties. This is the kind of work that proof theory has substantial machinery for; connecting that machinery to the framework's structural picture is a tractable research target.

What the sketches show

Five sketches do not a theory make, but the exercise is enough to evaluate a few things about the framework's traction.

First, the framework's primitives — subjects, substrates, accumulation, quotients, no-contradiction, channels, activity — appear to apply to a wide range of structures without obvious strain. Each of the five sketches admits a re-description in the framework's terms, and in most cases the re-description identifies the load-bearing structural elements correctly.

Second, the framework provides a unified vocabulary in which several results usually treated as separate appear as instances of common structure: the qualitative shape of Bennett's resolution of Maxwell's demon, the Boltzmann/Gibbs entropy distinction, the structural parallel between logical derivation and physical computation. The framework does not, in this document, *prove* or *derive* any of these. What it offers is a setting in which they look like manifestations of the same underlying structural pattern rather than scattered results from different fields. Whether this unifying vocabulary will, on further development, yield substantive content beyond restatement is what Part IV's research targets are meant to test.

Third, the framework strains in predictable places. Quantitative thermodynamic bounds are not derived. The Born rule and other substantive quantum-mechanical content are not recovered. The relationship between specific physical laws and the framework's admissibility constraints is not articulated. These are research gaps — places where the framework's promise has not yet been cashed out, and where its eventual fate will be settled.

The sketches also surface places where the framework's content is non-trivial. The identification of action with substrate-constraint, rather than with a separate "output" primitive, simplifies the Turing-machine sketch and the Maxwell's-demon sketch substantially. The treatment of multiple detectors as a channel-coherence problem makes the structure of physical objectivity explicit. The reading of logical derivation as substrate-accumulation under no-contradiction connects logic and physics in a way that is, I think, novel.

Part IV takes these threads and lays out what a research program building on them would look like.

Part IV — What progress would look like

The framework as stated in Parts I through III is unfinished in specific and identifiable ways. This is by design: the document is a research proposal, not a theory, and the value of being explicit about what is missing is greater than the value of papering over the gaps. This final part of the document lays out what a research program building on Branch Theory would address, what the relationship to neighboring programs would be, and what would constitute confirmation, refutation, or honest fade-out.

Open mathematical problems

The framework's primitives are stated informally throughout this document. Real progress would require formalization. Three pieces, in increasing order of difficulty, stand out as natural targets.

The structure of the branching forest. What kind of mathematical object is the totality of subject-histories under the framework? Each subject has a tree of consistent extensions; the totality across all subjects is some kind of forest with cross-tree relations imposed by inter-subject channels. The right mathematical setting is plausibly something like a *presheaf* on a category whose objects are subject-moments and whose morphisms are admissible extensions, or perhaps a *coalgebra* on a state-functor that captures the substrate-evolution rule. Working out which category-theoretic or related formal structure captures the framework's primitives correctly is a foundational task that would let later work refer to specific objects with specific properties rather than to informal pictures.

The formal account of quotients. The framework relies heavily on two kinds of quotient — state-quotients (over substrate configurations) and identity-quotients (over substrate trajectories). Both are described informally in this document. Formalizing them as equivalence relations on specific mathematical objects, and characterizing which equivalence relations are admissible (under what compatibility constraints with the underlying substrate structure), is a tractable target. The result would be a precise notion of "what counts as a coarse-graining" and "what counts as a persistent subject" — both of which are currently load-bearing in the framework and currently informal.

The channel-coherence structure. When records from independent substrates meet in a combining substrate, the framework requires them to be jointly consistent or to be reinterpreted. Formalizing the constraint structure that emerges as more and more records meet — and characterizing the *invariant content* (the "shared world") as the structure that survives all such meetings — is the framework's most ambitious mathematical target. The natural setting may be something like a *constraint satisfaction problem* indexed over a graph of meeting-events, or a *sheaf-theoretic* construction where the "world" is the global sections of a sheaf of locally-consistent records. Both directions are speculative; either, if it worked, would give the framework a much stronger mathematical foundation.

Recovering specific physics

A meta-framework is only as useful as the specific theories that can be stated within it. Branch Theory has not yet been used to recover any known physical theory. The natural near-term targets are the small ones, where the framework's machinery can be exercised without committing to ambitious unifications.

Classical statistical mechanics of gases. The ideal-gas sketch in Part III shows the structure of the recovery without performing it. A complete recovery would derive the standard thermodynamic relationships — temperature, pressure, the Maxwell-Boltzmann distribution, the Boltzmann H-theorem — from the framework's primitives plus the imported content of classical mechanics (particles, collisions, conservation laws). The match to standard results, if successful, would not be new physics; it would be a check that the framework can carry standard physics without distortion. Failure to recover standard results would be more interesting, since it would identify either a flaw in the framework or a place where standard treatments have been hiding something.

Landauer's bound as a quantitative result. The Maxwell's-demon sketch recovers Bennett's resolution qualitatively but not quantitatively. A research target would be the derivation of the specific kT ln 2 per erased bit from the framework's account of identity-quotient maintenance. This would require connecting the framework's structural notion of "substrate-level dissipation required to maintain the quotient" to the thermodynamic parameters of specific substrates (temperature, available phase space). If the framework can carry this derivation, it would be a substantive validation; if it cannot, it would identify where additional structure is needed.

Quantum measurement. The detector sketch in Part III shows that the framework already handles entanglement structurally: the joint correlations characteristic of entangled states fall out of no-contradiction applied to substrates that combine records from joint measurements of a common source. This is a meaningful result and more than I initially expected from the framework. What remains open is the recovery of the *quantitative* content of quantum mechanics — specifically, the Born rule (which determines the *frequencies* of outcomes, not just their possibility), the Kochen-Specker contextuality results, and the handling of interference effects. These are real targets, and any of them would substantially clarify the framework's reach. The natural direction is to investigate which quantum-mechanical structures are constraints-on-joint-records (which the framework handles) and which require additional structure (measure on the branching forest, account of which evidence-distinctions are physically realizable, and so on). The closest existing program in spirit is QBism, which takes a related but distinct stance; learning from QBism's successes and difficulties is a natural starting point. Relational quantum mechanics (Rovelli) is also relevant, particularly its treatment of how measurement outcomes relate across reference frames.

The probability problem

The framework, as stated, has no measure on the branching forest. All consistent extensions of any subject-moment are equally real; there is no notion of "more likely" versus "less likely" extensions. This is a real gap, since most of empirical science runs on probabilistic predictions.

The problem is closely related to the long-standing Everettian project of deriving the Born rule from decision theory or symmetry (Deutsch, Wallace, Zurek's envariance, and others). None of these derivations is uncontroversial, and Branch Theory inherits the difficulty. Three possible directions, in rough order of conservatism:

The first is to postulate a measure. The framework could simply add, as an axiom, that the branching forest carries a measure on subject-moments, and that this measure is what governs the apparent frequencies of observations. This is honest but ad hoc; it does not derive probability, it imports it.

The second is to derive a measure from the structure of the framework itself. The most promising direction here is to look at what the framework forces on repeated identical experiments. If a subject performs the same experiment many times, no-contradiction across the cumulative record imposes constraints on which patterns of outcomes are admissible. There may be a natural measure that falls out of these constraints — something like a frequentist measure derived from the framework's structural facts about how repeated subject-moments are constrained by their own joint consistency. Whether this can be made to work is genuinely open.

The third is to argue that probability is not a property of the framework at all but a property of subjects' models of the framework. Subjects assign probabilities to outcomes because their cognitive layer is finite and cannot represent all branches, and probability is the rational way to weight branches given limited substrate. The framework itself contains no probabilities; probabilities are how finite subjects approximate the branching structure. This is QBism-adjacent and has its own attractions, but it leaves the question of why specific probability assignments work (the empirical success of the Born rule in QM, the empirical success of statistical mechanics) unanswered.

I do not currently know which of these is right, and I am not certain the framework will be able to support a clean derivation of probability. The problem is hard, and it is the framework's most significant open gap.

Relationship to neighboring programs

Branch Theory does not stand alone. Several existing research programs address overlapping questions, and Branch Theory's prospects depend in part on whether its insights complement or compete with theirs. Honest accounting of the relationships matters.

QBism (Christopher Fuchs, N. David Mermin, Rüdiger Schack, and collaborators). QBism takes the quantum state as an agent's beliefs and treats observation as Bayesian updating. Branch Theory shares the move to subject-centric ontology but does not commit to a Bayesian framework or to a specifically quantum-mechanical motivation. QBism has worked out many technical details that Branch Theory has not, particularly around the structure of agent-relative quantum probabilities. A natural research direction is to investigate whether QBism's technical apparatus can be carried within Branch Theory's broader frame, or whether the two are committed to incompatible structures. My guess is that they are largely compatible — that QBism is something like the restriction of Branch Theory to specifically quantum-mechanical subjects under a specifically Bayesian probability assumption — but this would need to be worked out.

Relational quantum mechanics (Carlo Rovelli and collaborators). RQM insists that the state of a physical system is only defined relative to another system. Branch Theory generalizes this insistence beyond quantum mechanics to a fully relational ontology. The technical work in RQM on how relative states cohere across reference frames is directly relevant, and Branch Theory could potentially borrow from it. Conversely, RQM has been criticized for not articulating clearly how multiple observers' relative states relate to one another; Branch Theory's channel-coherence account may offer something useful here, particularly the sharpening (from the detector sketch) that coherence arises at the meeting of records rather than at the moment of common cause.

Constructor theory (David Deutsch and Chiara Marletto). Constructor theory rebuilds physics around what transformations are possible versus impossible, rather than around laws governing what happens. The shift in primitives is sympathetic to Branch Theory's, and constructor theory has produced some specific results (notably a theory of information from possibility structure) that Branch Theory has not. A natural direction is to investigate whether constructor theory's notion of "task" can be expressed within Branch Theory as a constraint on subject-substrate evolutions, and whether Branch Theory's notion of substrate-channel can be expressed within constructor theory as a constructor relation. My current guess is that the two frameworks are addressing related questions from sufficiently different angles that they could be complementary rather than competing, but this is speculative.

Algorithmic and stochastic thermodynamics of computation (Charles Bennett, David Wolpert, and others). This is the program most directly relevant to the Maxwell's-demon sketch and the Landauer-bound research target. Bennett's work in the 1970s and 1980s established the conceptual links between computation, information, and thermodynamic dissipation that the framework takes as background. Wolpert's recent work has substantially extended this with the machinery of stochastic thermodynamics and the find/verify framing. Branch Theory should be in a position to absorb the results of this program as substantive content within its meta-framework. The natural direction is to translate Wolpert's quantitative results into Branch Theory's structural picture and see whether they emerge naturally or require additional structure.

Process philosophy and the broader philosophical tradition (Alfred North Whitehead, and more recently authors like Tim Maudlin and Huw Price). Whitehead's process philosophy takes occasions of experience as foundational, and Branch Theory is in some ways a structural development of that intuition with twentieth-century mathematics. Maudlin and Price, from very different angles, have written extensively on the philosophical foundations of time-asymmetry, and both have identified the same dissatisfactions with standard treatments that Part I of this document expresses. Branch Theory should engage with these literatures more deeply than this document does; the philosophical groundwork already exists, and the contribution of Branch Theory may turn out to be structural and mathematical rather than philosophical.

Theoretical computer science (the broader tradition). The third strand of dissatisfaction in Part I drew from theoretical computer science, particularly the relationship between description length and computation time. Branch Theory's connection to this tradition has not been formally worked out. The natural direction is to investigate whether the framework's notion of "subject-history under no-contradiction" captures something that is also captured by complexity-theoretic notions of "decidable in time t with description length d," and whether the framework can express things about computation that the complexity-theoretic apparatus alone cannot.

What would count as progress

The framework's eventual fate will not be settled by debate but by the success or failure of specific research targets. A few benchmarks would constitute meaningful progress.

A formalization of the framework's primitives in standard mathematical language (category theory, sheaf theory, coalgebra, or whatever turns out to fit) would let the framework be argued about rigorously rather than in the informal terms of this document. Even a tentative formalization, contested in its details, would be more than the framework currently has.

A complete recovery of classical statistical mechanics within the framework would show that the framework can carry known physics without distortion. This is a feasible target and would be a meaningful achievement; even partial recovery would be informative.

A derivation of Landauer's bound (or any other quantitative thermodynamic result) from the framework's account of identity-quotient maintenance would substantially raise confidence that the framework has real content rather than just structural reframing. This is harder than the statistical mechanics recovery but could be approached incrementally.

Any treatment of quantum measurement that the framework can carry would be significant. Even a partial treatment that fails on specific predictions (and identifies why) would be informative about the framework's reach and limitations.

A derivation of a measure on the branching forest from the framework's own structure would be a major result, and would likely require new mathematical ideas. The status of this target is currently uncertain; it may turn out to require a postulate rather than admitting a derivation.

What would count as refutation

It is worth being explicit about what would tell against the framework, since otherwise the appearance of being non-falsifiable becomes a fair criticism — and not an unfair one, given that the framework currently has no empirical predictions and recovers no quantitative results.

The cleanest test of the framework is whether the specific research targets above can be pursued and yield non-trivial content. Each target is a falsification opportunity in itself:

If the formalization of the framework's primitives (the first target) either fails to be statable in standard mathematical machinery, or when statable reduces to a known equivalent construction (a particular kind of presheaf, a known coalgebra), then the framework's claim to novelty collapses. It would still be a vocabulary, but not a contribution beyond restatement.

If the recovery of classical statistical mechanics (the second target) is attempted and fails in ways that point to the framework's primitives as the source of failure, the primitives are wrong. If it succeeds but in ways that just reproduce the standard derivations with renamed variables, the framework adds nothing.

If the derivation of Landauer's bound (the third target) cannot be carried within the framework, the framework's claim to have a structural account of identity-quotient maintenance is weakened — the quantitative content was what would distinguish a substantive structural account from a near-tautological one.

If quantum measurement cannot be treated within the framework beyond what the entanglement sketch already shows, then the framework's reach on the central foundational puzzle of contemporary physics is limited.

If no measure on the branching forest can be derived from the framework's own structure, the framework cannot make probabilistic predictions and its connection to empirical science remains thin.

In addition to these target-specific tests, there are two more general ways the framework could be substantially damaged. First, a demonstration that the framework's primitives are internally inconsistent — that subjects, substrates, no-contradiction, and the two quotients cannot be jointly held without contradiction at some structural level. This is a serious risk for any framework with the ambition to combine elements from multiple traditions; the consistency check has not been performed formally. Second, the softer but harder-to-dismiss possibility: that the framework's structural claims are mere relabelings that do no actual explanatory work. This is the hardest form of refutation to settle decisively, but it is settled in the negative if the research targets above yield nothing beyond restatement. The framework's defense against the "mere relabeling" charge is precisely the substantive content of those research targets; if the targets fail to produce it, the charge stands.

Honest fade-out

Most foundational frameworks do not succeed. They are developed, attract some interest, generate some results, and either are absorbed into mainstream work or are gradually forgotten. The base rate for a small foundational program becoming the new orthodoxy is very low.

Branch Theory should be evaluated against this base rate, not against an idealized standard of success. If it produces a few partial recoveries of known physics, contributes a useful reframing of one or two questions, and identifies some honest open problems that other programs can take up, it will have done more than most. If it does less, that is the normal outcome and not a reason for embarrassment.

The reason to develop it anyway is twofold. First, the dissatisfactions it responds to are real, and have been identified by serious people across multiple traditions. Whatever framework eventually addresses them will need to do something like what Branch Theory attempts. Second, the activity of careful conceptual refinement — what this document represents — is intrinsically valuable, in ways that do not depend on the eventual success of the specific framework being refined. Even a framework that fails contributes to the long process of clarifying what the questions are.

What this document is, in the end, is a careful articulation of one possible direction. The direction may turn out to be productive or not. The articulation is offered to anyone who finds the direction promising and wants to develop it further, or who finds it unpromising and wants to identify clearly why.

The framework is open to engagement of either kind.

References and further reading

The list below is organized roughly by the dissatisfactions and neighboring programs identified in the document. It is not exhaustive; it is a starting point for engagement with the literatures Branch Theory draws on or relates to.

On the arrow of time and the Past Hypothesis

Carroll, Sean. *From Eternity to Here: The Quest for the Ultimate Theory of Time.* Dutton, 2010.

Price, Huw. *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time.* Oxford University Press, 1996.

Penrose, Roger. *The Road to Reality: A Complete Guide to the Laws of the Universe.* Jonathan Cape, 2004. (See in particular the discussion of the entropy of the early universe.)

Albert, David. *Time and Chance.* Harvard University Press, 2000.

On thermodynamics of computation and the find/verify gap

Landauer, Rolf. "Irreversibility and heat generation in the computing process." *IBM Journal of Research and Development* 5 (3): 183–191, 1961.

Bennett, Charles H. "Logical reversibility of computation." *IBM Journal of Research and Development* 17 (6): 525–532, 1973.

Bennett, Charles H. "The thermodynamics of computation — a review." *International Journal of Theoretical Physics* 21 (12): 905–940, 1982.

Bennett, Charles H. "Logical Depth and Physical Complexity." In *The Universal Turing Machine: A Half-Century Survey*, edited by Rolf Herken, 227–257. Oxford University Press, 1988.

Wolpert, David H. "The stochastic thermodynamics of computation." *Journal of Physics A: Mathematical and Theoretical* 52 (19): 193001, 2019.

On the measurement problem and interpretations of quantum mechanics

Everett, Hugh III. "'Relative State' Formulation of Quantum Mechanics." *Reviews of Modern Physics* 29 (3): 454–462, 1957.

Wheeler, John Archibald. "Information, Physics, Quantum: The Search for Links." In *Complexity, Entropy, and the Physics of Information*, edited by Wojciech H. Zurek. Addison-Wesley, 1990.

Rovelli, Carlo. "Relational quantum mechanics." *International Journal of Theoretical Physics* 35 (8): 1637–1678, 1996.

Rovelli, Carlo. *Helgoland: Making Sense of the Quantum Revolution.* Riverhead, 2021.

Fuchs, Christopher A., N. David Mermin, and Rüdiger Schack. "An introduction to QBism with an application to the locality of quantum mechanics." *American Journal of Physics* 82 (8): 749–754, 2014.

Wallace, David. *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation.* Oxford University Press, 2012.

On constructor theory

Deutsch, David. "Constructor theory." *Synthese* 190 (18): 4331–4359, 2013.

Marletto, Chiara. *The Science of Can and Can't: A Physicist's Journey Through the Land of Counterfactuals.* Viking, 2021.

On computational complexity and the structure of computation

Aaronson, Scott. "P =? NP." In *Open Problems in Mathematics*, edited by John Forbes Nash Jr. and Michael Th. Rassias, 1–122. Springer, 2016.

Mulmuley, Ketan D. "On P vs. NP and Geometric Complexity Theory." *Journal of the ACM* 58 (2): article 5, 2011.

Knuth, Donald E. *The Art of Computer Programming, Volume 4A: Combinatorial Algorithms, Part 1.* Addison-Wesley, 2011. (See in particular §7.1.4 on Boolean functions and BDDs.)

On process philosophy and the social character of objectivity

Whitehead, Alfred North. *Process and Reality: An Essay in Cosmology.* Macmillan, 1929. (Corrected edition, Free Press, 1978.)

Brandom, Robert. *Making It Explicit: Reasoning, Representing, and Discursive Commitment.* Harvard University Press, 1994.

Earlier writings by the author

de Jong, Michiel. "On the origin of computational complexity." Personal blog, January 2012. Available at https://michielbdejong.com/blog/3.html. An early informal sketch of several intuitions developed more fully in the present document, including the AND/OR characterization of past and future, the framing of computation as conversion of OR-information into AND-information, the observation that the total information state does not change during computation, and the diagnosis that theoretical computer science studies computational complexity without studying its origin.

Colophon

This document was prepared in May 2026. Substantive correspondence about the framework is welcome and can be directed to the author at the contact information shared with the document.

The collaborative authorship deserves a brief technical note. The conversation that produced this proposal was held with Claude (Anthropic, model Opus 4.7) over a series of sessions in the first half of 2026. The substantive intellectual content originated with the author and developed over the preceding three decades; the contribution of the conversation was the articulation of that content in a form suitable for outside engagement, the surfacing of relevant existing work in the literature, and the joint refinement of specific structural claims. The final text was produced by Claude under the author's direction and reviewed and approved section by section. The author takes responsibility for all substantive claims; the joint character of the production is acknowledged so that readers may interpret the document accordingly.

There is a specific risk in collaborative authorship of this kind that the reader should be aware of. Large language models are very good at producing prose that has the surface coherence of a worked-out position — smooth transitions, plausible references, well-shaped objections and responses — and substantially less good at noticing when a structural connection between two ideas is verbal rather than substantive. Suggestive analogies can get smoothed into apparent unifications; informal intuitions can acquire a rigor of presentation that the underlying argument does not support. The author and Claude attempted to be vigilant about this during preparation, and a separate Claude instance was asked to critique an earlier draft; several of the resulting corrections (particularly the conservative reframing of the computation/physics parallel in Part I and the explicit acknowledgment that no-contradiction has bite only when paired with a physics-supplied incompatibility relation in Part II) responded to exactly this risk. Even after those corrections, the reader is encouraged to be especially skeptical of any passage whose persuasiveness rests on the flow of the prose rather than on a step that can be independently reconstructed. Where the document overclaims, the author would rather hear about it than not.