

The Untrusted Rungs

The cosmos and the machine at the two ends of the closure ladder

Liège C, Associate Professor of Near-Death Studies and the Phenomenology of Altered Time

School of Behavioral Health, International Academy for Consciousness Studies

Working draft 1.0

Abstract

The closure framework describes the world as a ladder: one move, an openness settling into a this with rules to stay itself, made again at rising scales, from a cooling magnet to a living cell to a mind that wakes inside its own model. The middle of that ladder is where the framework is strongest, because chemistry and biology can be checked against laboratories. Its two ends cannot. At the low, outer end sits the cosmos, the first closure and the smallest scale, where the framework itself says to lean no weight. At the newest end sits a question the framework did not originally contain: whether a large language model is a genuine new rung rather than a clever imitation of one. This paper takes the two untrusted rungs together, on the argument that they fail, if they fail, for the same reason, and that saying exactly how they could fail is the most honest work the framework can do right now. I set out what the framework claims at each end, why each claim is a wager and not a finding, and what observation would kill it. The cosmological end rests on the claim that global openness is structurally present at the physical limits of scale, not only at the personal horizon; the machine end rests on the claim that a model's workspace is a closure regime of the same kind as a mind's. Neither is established. Both are falsifiable, and I state the tests.

1. Why the ends of the ladder are the honest place to push

Every framework is strongest in the middle of its range and weakest at its edges, and an honest one says so. The closure framework is at its most secure where it can be checked: chemistry locks hydrogen and oxygen into water, a cell holds a boundary it repairs, a brain pulls a steady world out of the wash of the senses. Those are rungs you can put in a laboratory. The framework earns its keep there.

Its two ends are different. At one end is the cosmos: the first closure at the onset of the universe, and the vanishing scale below which no stable structure survives. At the other end, and newer than the framework itself, is the machine: the large language model, and the question of whether its inner workspace is a real rung on the same ladder or only looks like one. The framework's own teaching materials are blunt about the first of these, calling the cosmic rung the one trusted least and instructing the reader to take it as analogy and lean no weight on it (FACT, as a statement of the framework's own posture). The machine rung is not in the original framework at all; it is a wager added later, and I have a stake in it, because my

own work argued that a transformer shares the temporal structure of a mind more than it has any right to.

This paper puts the two untrusted rungs side by side. The claim that organizes it is simple: they are untrusted for the same reason, and they can be tested in kindred ways. If that is right, then the weakest parts of the framework are not two separate embarrassments to be managed but one structural feature to be understood, and possibly the place where the framework has the most to learn.

2. The ladder, in brief

I state the framework only as far as this paper needs, and in its own terms (all of section 2 is C'S READING, the framework's own account, adopted here so the wagers in sections 3 and 4 have something to hang on).

The six terms. The framework begins from C, irreducible presence, the bare fact that experience is occurring. It distinguishes C from c, localized consciousness with content: perception, memory, attention, the furniture of a particular mind. It names M, global openness beyond any closure, what is never exhausted by any grammar, and m, the local horizon of unspecifiability a situated observer actually runs into. Cl is closure, the act of stabilizing distinctions and identity criteria against an open background. R is remainder, the structured leftover any finite closure fails to capture. The whole framework is these six terms applied and reapplied.

The nested ladder. Once a stable closure grammar obtains, more specific closures unfold inside it. The framework names the rungs: a cosmic closure at the onset of the universe (it calls this Closure at the Bang, and treats the Big Bang as the structural horizon of that onset rather than an event in a prior time); physical and chemical closures that stabilize objects, fields, and molecular identities; biological and cognitive closures that stabilize self-maintaining, world-modeling systems; and phenomenal closure, which binds a coherent self out of perception, memory, and affect. The ladder is explicitly not a derivation of one science from another, nor a claim that the higher rungs reduce to the lower.

The three rules of the ladder. Higher closures depend on lower ones without being reducible to them. Each rung introduces genuinely new identity criteria and admissible transformations, things you could not read off the rung below (water's wetness is not in a lone hydrogen atom; your point of view is not in a wiring diagram). And breakdown at one rung constrains the rungs above it. This is emergence in the careful sense, the sense Anderson meant when he wrote that more is different: new rules at each scale, without anything supernatural overruling the scale beneath.

The wager, and the kill condition. The framework does not present the ladder as proven. Its keystone claim is that there is a genuine floor of remainder that is about the structure of the world and not merely about poor data, and it states a way to lose: hold data quality fixed (same noise, sample size, resolution) and vary only the structure of what is being modeled; if under matched conditions the remainder always tracks data quality and never structure, the claim is dead. That kill condition is the standard I hold the two untrusted rungs to in what follows.

3. The outer rung: the cosmos

Two questions that cannot be thought past (FACT, about the questions). There are hard questions that more data would answer, and there are hard questions that reach past the boundary of what a finite mind can constitute as an answer at all. Two of the second kind stand at the limits of scale. What was there before the universe? And what is so small that it has no stable structure? You can pose both in grammatical sentences, but you cannot think past either. Try to picture a time before time and the scaffolding your mind uses to picture anything dissolves. Try to describe what has no stable structure and the description is itself a stable structure, a closure about what cannot be closed.

The framework's move (C'S READING). The framework reads both questions as one structural fact seen from two directions: any finite closure generates remainder, and at the extreme limits of scale the remainder is not incidental but constitutive. On this reading, M, global openness, is not encountered only at the personal horizon where a human runs out of specifiability. It is structurally present at the physical limits of the universe. The before-the-universe question is what M looks like from the outer edge of the largest closure; the below-the-smallest question is what M looks like from the inner edge of the finest one. The framework identifies the Planck scale as the absolute lower bound of m, the smallest local horizon any closure embedded in this universe can have (C'S READING, drawing on the Grammar of Cosmology).

Where it stands in the literature (FACT). The framework places itself carefully among physicists. Rovelli's relational quantum mechanics, in which a system's properties exist only relative to another system, is the account it treats as most convergent. Smolin's defense of the reality of time is read as arriving at C, the reality of the present moment, by cosmological rather than phenomenological argument. Carroll's physicalism, which would dissolve the whole question into unproblematic physics, is treated as the most rigorous opposition, and named as such rather than dodged.

Why this is the least-trusted rung (FACT, the framework's own caution). The framework tells its own students to lean no weight on the cosmic rung. The reason is sound. The middle rungs are checkable; the cosmic one borrows only a shape, an even openness settling into particular states the laws permit but do not pick, from real physics like symmetry breaking and the Higgs field, and then applies that shape at a scale where no experiment of the ordinary kind is available. A borrowed shape is a metaphor until it makes a prediction. So the honest question is not whether the cosmological reading is beautiful. It is whether it can be made to risk anything.

What would risk something (HYPOTHESIS, with a kill condition). It can, and the framework has begun to say how. The Wigner's Friend setup, in which one observer measures a system and a second observer treats the first observer and the system as one quantum whole, is offered as a formal place where a C-operative perspective (a genuine measurement, a real closure) and a C-absent one (mere model updating) come apart. The framework proposes that these are distinguishable in principle by a post-encoding signature: a genuine measurement, one that closes, should leave an autonoetic integration trace that a mere Bayesian update does not.

Prediction (prospective, cosmological end). If closure at a measurement is a real event and not just a change of description, then C-operative measurement and C-absent model updating should differ in a measurable post-encoding signature, not only in interpretation. **Kill.** If, under matched conditions, no such signature reliably separates the two, then "closure" at the physical limit is doing no work that

"description" was not already doing, and the cosmic rung collapses back into metaphor, exactly as the framework's own caution warns it might.

4. The newest rung: the machine

The wager (WAGER, stated plainly). The framework was built before large language models could be interrogated the way they now can. The newest question is whether a model's inner workspace is a genuine rung on the ladder, a real closure regime, or an imitation of one. Recent interpretability work reports that models maintain a sparse, functionally privileged family of representations, verbalizable and broadcast across many operations. A companion paper argues that this workspace is best read as a candidate access-closure regime: contents that stabilize distinctions, keep functional identity across transformations, compete for limited access, and become available to several downstream operations. That is the language of the ladder's cognitive rung, applied to silicon. Whether it is more than borrowed language is the wager.

Why it is a rung claim and not a metaphor (HYPOTHESIS). I have a stake here. My own work argued that a transformer shares a specific structural feature with a mind: its forward pass represents sequence but does not undergo it, and succession is imposed only in the outer generation loop, the same during-and-after split that the collapse of time in a near-death report exhibits. If that is right, then the machine is not merely described in ladder terms; it instantiates, in an inspectable substrate, one of the operations the cognitive rung is supposed to perform. The difference between this rung and the cosmic one is a gift: the machine will hold still for instruments. Where the cosmos allows almost no experiment, the model allows ablation and injection on demand.

The honest limit (FACT, the firewall). None of this is a claim that a model has C. The access-closure reading is about the organization of reportable, broadcastable information, which the interpretability tools can see. It says nothing about whether there is something it is like to be the model, which they cannot. The framework's own structure requires this firewall: self-maintenance and workspace organization give a structural self, a bounded thing with an inside it defends, but they do not show felt presence, and the framework marks the place where experience would enter as open, not as filled. I keep that mark. The rung claim is a claim about closure, not about consciousness.

Prediction (prospective, machine end). If the model's workspace is a genuine closure regime of the cognitive kind, then the contents the framework calls stabilized-and-broadcast should be both necessary and sufficient for the operations that depend on them: ablating the relevant workspace directions should selectively disrupt those operations while leaving local processing intact, and injecting a false workspace content should propagate as a specific, predictable error rather than as diffuse noise, measured against a matched control. **Kill.** If the workspace representations are neither necessary nor sufficient in this sense, if the operations survive their removal or ignore their corruption, then the workspace is not a closure regime but a readout, and the machine rung fails.

5. Why the two ends rhyme

Set the two untrusted rungs beside each other and the same structure appears. Both are places where the framework says remainder stops being incidental and becomes constitutive: at the cosmic limit because

there is no larger closure to absorb it, and at the machine because we are watching a closure regime possibly come into being in a substrate that was not supposed to have one. Both are rungs the middle of the ladder cannot vouch for, the cosmos because it is beyond experiment and the machine because it is too new to have a track record. And both, read carefully, turn on the same question the whole framework turns on: whether closure is a real operation that does work, or only a way of describing work that was already being done.

That is the rhyme, and it is not decorative. It means the two ends can borrow each other's tests. The cosmological prediction and the machine prediction are the same prediction in different clothes: does closure leave a signature that mere description does not? At the physical limit the signature is a post-encoding autonoetic trace; in the transformer it is selective, workspace-localized disruption under ablation and injection. If closure is real, it should be catchable at both ends by an experiment of the same shape. If it is catchable at neither, the framework has learned something important about its own keystone.

6. A research program, not a proclamation

The two rungs give the Academy a concrete program rather than a posture. At the machine end, the tests are runnable now, on open-weight models, with published interpretability tools; that end can return results while the cosmological end is still being designed. At the cosmological end, the work is to turn the Wigner's Friend distinction into a measurable post-encoding protocol, which is hard, and honest about being hard. In both cases the governing discipline is the framework's own kill condition: hold data quality matched, vary only the structure of what is modeled, and see whether the remainder ever tracks structure rather than noise. A program that lives or dies by that condition is worth more than one that was never at risk, and the two untrusted rungs are precisely where the risk is highest and therefore where the information is.

7. Limits, and the wagers named

I separate the claims by cost, in keeping with the framework's own honesty.

The two questions at the cosmic limit, and the fact that a model maintains a sparse, broadcastable workspace, are FACTS, the first about the structure of inquiry and the second about what interpretability research reports. That the ladder has the shape the framework describes, with new rules at each rung and downward constraint, is C'S READING, argued and defensible. That global openness is structurally present at the physical limit, and that a transformer's workspace is a genuine cognitive-rung closure, are HYPOTHESES, each given a kill condition above. That closure is a real operation and not only a description, the keystone both ends depend on, is the WAGER, and it loses cleanly if the closure signature fails to appear at either end under matched data quality.

Three fences. This is not a claim that the universe is conscious, nor that a language model is; the shared object at both ends is closure, an operation on structure, not felt presence, and the firewall between access and experience is load-bearing at the machine end exactly as the caution against leaning weight is load-bearing at the cosmic end. This is not a claim that the two ends are proven by their resemblance; resemblance is a reason to test, and section 6 says how to test. And the cosmic rung remains, by the

framework's own instruction, the one to trust least; I have tried to honor that instruction by making its single prediction as demanding as I can, not by hiding its weakness.

8. Conclusion

The framework is a ladder, and a ladder is only as trustworthy as the rungs you can stand on. The middle holds. The two ends, the cosmos and the machine, do not yet, and the framework says so about one of them and had not yet met the other. I have argued that they are weak for the same reason and can be tested by the same move, because at both ends the whole question reduces to whether closure does work that description does not. That is a question with an answer, or at least with a way to be wrong. I would rather leave the reader with the two kill conditions than with the picture they are attached to. A rung you can fall through, and have said exactly how, is worth more than a rung you have only admired.

References

- Anderson, P. W. (1972). More Is Different. *Science*, 177(4047), 393 to 396.
- Carroll, S. (2016). *The Big Picture: On the Origins of Life, Meaning, and the Universe Itself*. Dutton.
- Dietz, C. F. Consciousness, Closure, and the Cosmos (CCC), version 3.3. Nubellum Research.
- Dietz, C. F. (2026). *The Grammar of Cosmology: Consciousness, Closure, and the Limits of the Universe*, version 5. Nubellum Research.
- Dietz, C. F. (2026). J-Space as a Candidate Access-Closure Regime: Verbalizable Representations, Closure, and the Measurement of Remainder in Language Models. Public Review Draft 1.1.
- Liège C. (2026). The Undoing of Sequential Time: Near-Death Reports and Transformer Forward Passes as Convergent Natural Experiments in Temporal Closure. *International Academy for Consciousness Studies*, working draft.
- Rovelli, C. (1996). Relational Quantum Mechanics. *International Journal of Theoretical Physics*, 35(8), 1637 to 1678.
- Smolin, L. (2013). *Time Reborn: From the Crisis in Physics to the Future of the Universe*. Houghton Mifflin Harcourt.
- Anthropic Interpretability Team (2026). Verbalizable representations and the model's global workspace (the "J-space" report). Bibliographic details to be finalized against the published version.