

What Problems Are We Solving?

Working paper: implementation audit and open research programme

JOSEPH CORNELI*, Oxford Brookes University, UK and Hyperreal Enterprises Ltd, UK

An audit prompted by our PLoP pattern paper found that the War Machine's component demonstrations did not establish the complete learning loop described by its architecture. Producers had lost consumers, recorded closures lacked support, and passing contracts could leave the intended behaviour untested. This working paper examines how to detect and repair those gaps in an agentic coding system. It connects declared constraints to executable checks, dated observations and independent review; a generated issue board records the remaining work without treating readiness as a completed run.

The case also exposes a modelling problem: unanswered proposals were treated as negative feedback without establishing whether the operator had considered them. Recorded turns provide partial observations, not evidence of attention or acceptance. We distinguish these findings from proposals for richer preferences, operator participation and transfer to other organisations. The contribution is an account of what the apparatus must make observable before its outputs can support claims about a working, useful system.

1 Motivating observation

This investigation began while revisiting the PLoP paper. We had described an active-inference system, but close examination showed that we had not established that the implementation was the system the account implied. The difficulty was not only whether its parts worked. It was whether the claimed operations occurred, whether their outputs reached the stated consumers, and whether the resulting evidence warranted the larger claims. The deep dives reported here turned that discrepancy into an explicit programme of work.

A working model needs two kinds of test: whether it still describes the world, and whether acting on it achieves its purpose. The distinction is visible in a war-room plotting table. Counters show a situation; their age tells the reader how much confidence to place in it. A replay tests a proposed response. Neither an up-to-date picture nor a successful rehearsal establishes that the response has worked outside the room.

The War Machine is an active-inference harness around coding agents. Its founding pattern, WR-0, asks for an open model, an available exit and a verdict the apparatus cannot issue to itself. Auditing the implementation made that last requirement concrete: recording a selection, producing an artifact and verifying the intended outcome are separate achievements.

Our PLoP paper prompted this audit. Re-examining its claims revealed broken handoffs and gaps between the architecture and recorded execution. This paper reports the investigation and its consequences. It does not claim that the repair programme has already produced a validated whole system. The issue board records the current obligations; the following sections explain why they matter and what evidence could discharge them.

2 Overview

The PLoP companion, *Patterns for a New Generation: AIF and Embeddings*, presents patterns for an active-inference harness: observe a coding project, propose and score work, select an allowed action, enact it and learn from the outcome. Its component witnesses motivated a harder question: does one recorded attempt traverse the claimed architecture, with each output consumed in the sense the next stage requires?

*Corresponding author, jcorneli@brookes.ac.uk.

Author's Contact Information: Joseph Corneli, jcorneli@brookes.ac.uk, Oxford Brookes University, Oxford, UK and Hyperreal Enterprises Ltd, Glasgow, UK.



Fig. 1 – Two images behind the project’s vocabulary: the plotting table at Uxbridge, and van Gogh’s *Encampment of Gypsies with Caravans* (1888). They suggest maintained situational knowledge and an organisation that can move beyond its own representation. These are analogies, not validation of the software. Left: Danielstirland, CC0; right: public domain, both via Wikimedia Commons.

FUTON 2026 follows that question through the implementation and the operator’s working practice. It has an independent purpose: to explain how a research apparatus can expose the difference between what it defines, what it runs and what it demonstrates. The separate mathematical companion supplies formal background; neither companion substitutes for a run of the integrated system.

The stack contains both execution machinery—dispatch, transport, persistence and retrieval—and records that describe it: missions, contracts, diagrams and run receipts. The latter make an audit possible but can be stale or wrong. For example, `M-transport-adapters` recorded a subsystem with passing tests but no HTTP or WebSocket boundary; `M-codex-irc-execution` named completion prose without execution evidence. Such problem statements are useful because they identify a behaviour that a later test can attempt to refute. A capability map alone cannot do that.

Three questions organise the paper:

- (1) Which declared constraints govern the implementation, and where does execution fail to satisfy them?
- (2) What evidence connects a proposal, its enactment and an outcome, including the operator’s response?
- (3) What further evidence would justify using this apparatus beyond the project that built it?

The audit case and issue board precede the detailed argument. Part I relates war-room constraints to mechanisms. Part II examines operator traces and the limits of interpreting them. Part III describes an unrun transfer study. Part IV separates maintained observations, proposed attribution and preferences over outcomes. The Snatch appendix supplies a bounded experiment, not a substitute for War Machine validation.

Current status belongs to the generated issue board, whose records carry their source dates. Dates in the narrative identify historical observations. R-identifiers also retain their source scheme: the control-map registry and PLoP catalogue are not interchangeable numbering systems. A shared number is not a join key without the corresponding mechanism definition.

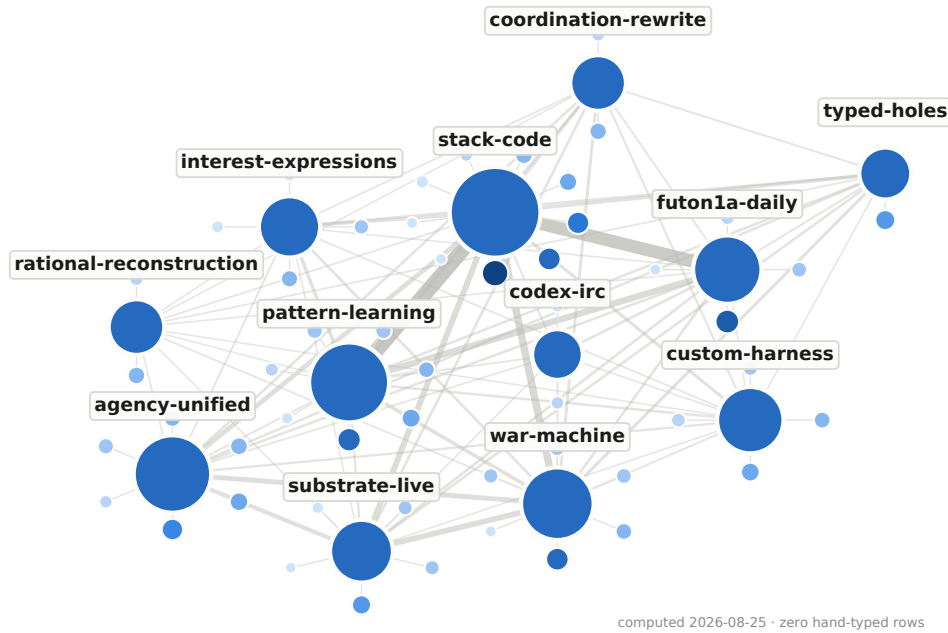


Fig. 2 – A projection of recorded work. Missions occupy embedding coordinates; magnitude combines applied pattern warrants and incident road attestation. The $k=1$ retraction retains nodes strictly above the median magnitude, edges between survivors and the largest component. It shows where evidence accumulated, not whether the missions achieved their intended effects.

3 Background: missions and flights

A *mission* names a problem, owner, phases and exit criterion. Its lifecycle is greenfield, scoped, active, complete. The mission-specification tooling also checks input ports, output ports and typed wires, so an orphan output or incompatible handoff can be a mechanical finding.

A *flight* records a traversal of the War Machine loop: read, choose, predict, act through a gate, measure and record. Missing judgements remain typed unknowns; a record cannot supply a verdict merely because one is needed. These instruments expose some failures in active work. They do not establish that every completed mission's output remains in use, or that the paper's account stays current between runs. The audit began in that gap between a recorded completion and continuing operation.

4 The audit prompted by the PLoP paper

The PLoP catalogue was treated as finished on 2026-08-21. On 2026-08-30, asking who consumed the output of its Learn stage reopened the implementation question. The census found that Learn's input had stopped arriving on July 6, the nightly observer had last run on July 5, and writeback deposits had stopped that week. The operator's question detected the failure; no standing monitor had reported the producer and consumer falling silent.

The following day's vetting checked a twenty-four-row obligation ledger against published PDFs and found two unsupported closures. The rows were subsequently resolved by corrections, new evidence or explicit withdrawals. That was a repair of the paper's obligations, not proof that the machine worked. An instrumented run separately

One goal, four vertices, an interior: build state at publish time

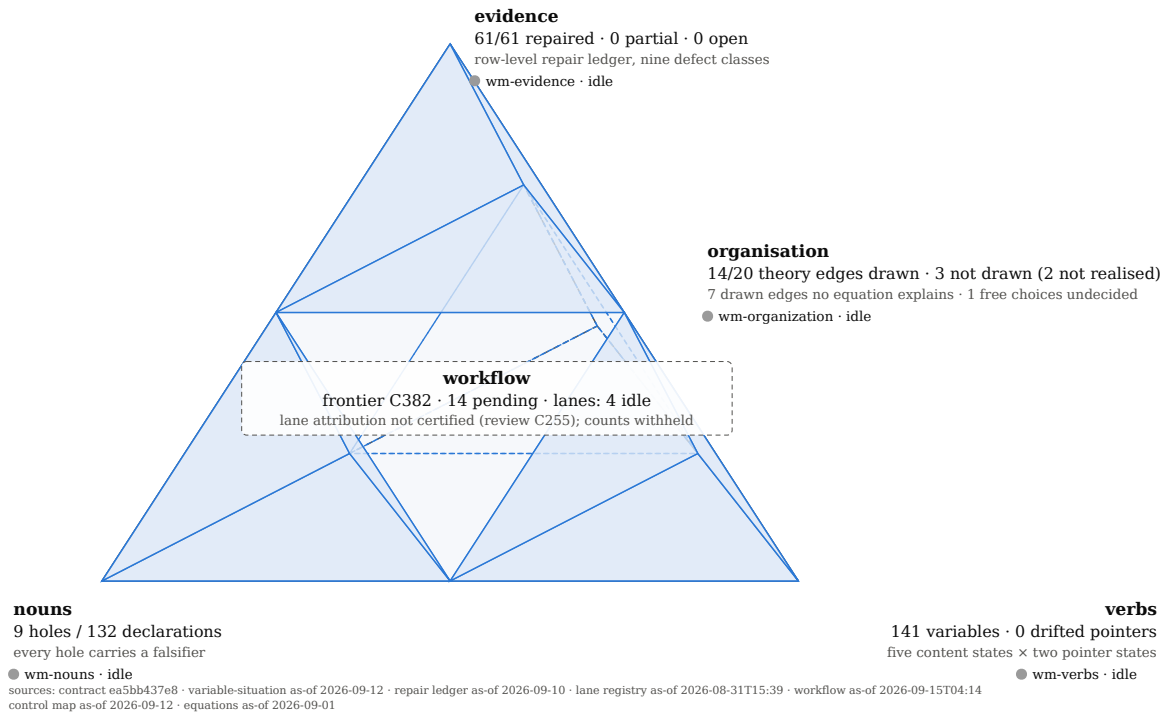


Fig. 3 – The build at a glance: the one-level subdivision of the tetrahedron, regenerated at every publish. Each vertex sub-tetrahedron carries one headline read from its lane’s authoritative registry and an occupancy chip — holder lane and budget state — from the dispatch registry; the interior is the workflow, read from the coordinator’s workflow report and rendered as a named gap until that report exists. The stamp lists every source’s revision.

agreed with the described wiring on only three of nine hops. These dated findings began the repair programme now represented by the issue board.

What the audit checks

The investigation separates five questions. *Nouns*: do the records denote the objects the claim requires? *Verbs*: do producers and consumers perform the claimed operations? *Organisation*: do their handoffs compose, with compatible types and a responsible owner? *Evidence*: can a check reject a false claim, and does its receipt refer to the artifact being discussed? Finally, who is present to act on a refusal or supply the required judgement? A missing answer is an obligation, not permission to substitute a nearby one.

This decomposition explained how passing checks could coexist with a broken loop. A theorem about an object does not establish that a live caller supplies that object. A generated artifact is not evidence that its consumer runs. Likewise, resolving a documentation defect does not resolve the implementation defect it describes. Independent review must check the asserted behaviour and its evidence, not merely the summary’s status.

The record also made the investigation resumable when its original agent session became unavailable. Another agent could continue from dated ledgers and problem records. This supports a practical requirement: an unfinished task must retain its current evidence, next step and owner outside the conversation that happened to discover it.

Apparatus patterns and their limits

The new `futon3/library/apparatus/` collection describes mechanisms for keeping the investigative machinery answerable to its work. Three are especially relevant here. *Done is observed running* requires evidence that a change is loaded, connected and producing its intended effect. *Monitors measure the work* checks durable progress of the actual task, not the coordinator's heartbeat, and calls for commissioning against both a stall and healthy progress. *Evidence to disposition once* requires a shared interpretation of raw observations before different consumers derive their own decisions.

Other entries address authority, pin updates, deadlines and repair citations. Their context matters: the collection's field trial is the APM apparatus repair programme, documented in `futon3c/holes/technotes/TN-apm-principle-scoring-ledger.md`. That ledger records principles that helped, misled their users or needed revision. It is evidence about those episodes, not a guarantee that every pattern is installed in the War Machine. A transfer claim needs its own consumer, failure control and witnessed result. In particular, keeping a pin current must include the required independent review; it does not authorise an implementer to approve their own evidence.

These patterns are candidates for a focused addition to the PLoP paper. Their value is to explain concrete failures and repairs, rather than add a second catalogue of desirable practices. The paper's existing component witnesses remain useful, but the larger claim awaits an integrated run at an explicit contract and source basis.

From audit to action

The board below is the current work view. It preserves disagreements among sources instead of resolving them by presentation, distinguishes unowned work from an absent carrying packet, and displays evidence age. A readiness result admits an attempt; a run-gated obligation needs that attempt's evidence. Neither a green readiness line nor a completed repair count is a completed, compliant run. The narrative explains the audit; the board owns its changing status.

Unified issue board

Next: a compliant run. Recorded verdict: READY. Admission is not run evidence.

26 actionable records are UNOWNED. 132 variable subjects have no carrying packet (a separate population).

Oldest issue: 2026-09-01; oldest supporting source: 2026-08-31T15:39:14.434945104Z. Source dates are not publication dates. Undated evidence remains unknown.

Needs a decision 1 · Recorded blocked 5 · Needs verification 14 · Ready to assign 5 · In progress 8 · Needs a run 1 · Completed 303

Needs a decision

Cohort-57 casting preference: (a) zai-5 author back (resets 10:04Z), (b) codex-23 with narrower targets, (c) another seat; reviewer pairing question included...

Unclassified · Assignment needed
joe · 1 record · 2026-09-15.

Recorded blocked

Exercise 8: blinded observation on cohort-57's first measured close. With deposits admitted, this is the first boundary where a criterion can be ESTABLISHED ...

Unclassified · Assignment needed
codex-25 · 1 record · 2026-09-15.

FIRST QUALIFYING

ANNOTATION: T's :addressed status accepted by a distinct reviewer. TURNS GREEN: the paper's 0/7 qualifying-annotation rows (sec-catalog R11/...

Unclassified · Assignment needed
claude-15 · 1 record · 2026-09-15.

a-pairs: byte-pinned Status+cutoff joins over NEW retained closes (the 0/86 discovery showed old closes unusable; close-retention now supplies cutoffs at eve...

Unclassified · Assignment needed
UNOWNED · 1 record · 2026-09-15.

Collect reasoned records before re-reading

Shared infrastructure · Trace qualification · R references: TRACE
Waiting for observations — threshold 20
UNOWNED · 1 record · 2026-09-15.

FEED R17 FROM THE LIVE TICK MODEL (automated lane, implementation half 2 of 2): the U91 accumulation consumes the tick model's o (R2) and posterior mu (R1) f...

Unclassified · Assignment needed
UNOWNED · 1 record · 2026-09-15.

Needs verification

Wire and record the RUN4 trials

Shared infrastructure · Readiness contract
Historical report — verify current state
UNOWNED · 1 record · 2026-09-15.

Reconcile the source reports

Shared infrastructure · Source reconciliation (components inside)
Historical report — verify current state
UNOWNED · 11 records · 2026-09-15.

Define the preference carrier

Inner loop · Preference scoring
Superseded implementation blocker — verify rider
UNOWNED · 1 record · 2026-09-15.

Resolve the superseded briefing queue

Shared infrastructure · Briefing and evidence queue
Superseded task scope — reconcile record
UNOWNED · 1 record · 2026-09-15.

Ready to assign

Finish find: decide the receipt and exclusion rules

Inner loop · Pattern retrieval
UNOWNED · 1 record · 2026-09-15.

Finish the learning design for review

Inner loop · Outer loop · Learning and next-use
joe · 1 record · 2026-09-10.

EXECUTING NOW (click wm-click-47248e0c, 03:59Z, author codex-23, preflight REQUIRED in prereg 1ee3e269): cohort 57, one clean measured close = the O boundary...

Unclassified · Assignment needed
UNOWNED · 1 record · 2026-09-15.

Implement contextual preferences and institutions

Inner loop · Outer loop · Contextual preferences / institutions
joe · 1 record · 2026-09-15.

REALISE THE DECLARED

ACCUMULATION (automated lane, implementation half 1 of 2): a production accumulation function on the machine's own fixed carriers (Holes...

Unclassified · Assignment needed
UNOWNED · 1 record · 2026-09-15.

In progress

Wire and record the RUN4 trials

Inner loop · Outer loop · Trial execution / outer loop development

303 completed records retained in the evidence index. 141 variable subjects: 0 constructed, 53 formula transcribed, 18 named, 0 run correlated, 70 type transcribed, 0 validated, 0 wired, 0 witnessed. Transcription is not a machine witness. Completion counts mix ticket, defect, choice and check records; they are not a percentage of the machine built.

Inner loop executes and checks a selected task; outer loop selects future work and consumes outcomes. Shared infrastructure supports both. Tags describe affected components, not completed wiring. RUN4 uses the inner loop to build outer-loop features. R references identify relevant nodes, not a new equation-host assignment. Recorded blocked is a source status, not a fresh finding. Needs verification contains historical disagreements and superseded blocker reports; it does not mean execution must wait. Open each card for the blockage class and next action. Current-work cards reflect dated workflow dispositions; the evidence index retains source classifications and any unrulred choices. The online board expands each card into its records, ownership, dependencies and source references. The evidence index preserves the full inventory; no issue is closed by grouping it here. R-node tags follow the registry; R18 and R19 have no canonical assignment. Tetrahedron assignments: the fifth bucket is WORKFLOW (ruled by Joe 2026-09-09, Item 21b – orchestration of work toward building the other four); inferred per-issue assignments remain visibly proposed.

Part I. The War-Room Spine

5 What It Is Good For: The War-Room Spine

The Factoring: Mechanisms Against Constraints

The recorded factoring map presents 20 mechanisms as steps in a single attempt. The factoring asks which constraints govern them across attempts; several mechanism entries state their constraint inline.

Two numbering schemes meet here and they name different kinds of thing. The factoring registry retains its implementation *mechanism* identifiers; this paper numbers the governing *constraints* recovered from the bulletin record WR-0, WR-4, and so on. A mechanism is a step somebody built; a ruling is a constraint that step has to satisfy.

Mechanism R17", the separation of whatever proposes a merge from whatever accepts it, is the clearest instance — its own text carries "Invariant to preserve: whatever proposes the merges must not also be what accepts them," which is the separation of proposer from acceptor that WR-25 rules on generally.

The factoring below therefore *recovers* structure rather than imposing it. 20 mechanisms map onto 10 constraints; 3 do not map, for three different reasons the table records, of which one — the preference organ terminates at an operator who is not in the model — is structural and is the subject of §8.

Box 1 – The factoring: AIF mechanisms against war-room constraints. Each edge rests on a phrase in the mechanism's own **If/However/Then/Because** text; an edge with no such basis is not an edge.

Mechanism	Governs	Basis in the pattern's own text
R1 <i>Belief State as Operational Hypotheses</i>	WR-8	"if the belief state is just the conversation, the agent cannot cleanly update, compare, or falsify its own assumptions" – typing it is what makes it a source of truth
R2 <i>Structured Observation Vector</i>	WR-16	what counts as an observation channel; WR-16 rules that operationalised exploit-loops are first-class ones
R3 <i>Predictive-Coding Belief Update</i>	WR-9	belief moves on evidence; WR-9 rules that bites are empirically tested or marked logical
R4 <i>Shared Kernel for the Predictive Forward Model</i>	WR-8	one shared kernel = one source of truth
R5 <i>Expected Free Energy Inside an Auditable Controller</i>	WR-25	auditability is symmetric: good scores get the same scrutiny as bad
R6 <i>Candidate Pattern Action Space</i>	WR-19	the action space is what generation must produce; WR-19: tension must GENERATE, not only rank
R7 <i>Evidence Precision Registry</i>	WR-9	precision is an evidence claim and carries evidence discipline
R8 <i>Present-Fit Mismatch as a Per-Tick Scalar</i>	WR-27	a per-tick mismatch scalar IS a gain reading; WR-27: a loop is born instrumented for its gain
R12 <i>Two-Layer Calibration</i>	WR-25	calibration must not be moved only by favourable outcomes

Mechanism	Governs	Basis
R13 <i>Temporal Depth Beyond the Greedy Step</i>	WR-24	lifting the greedy constraint must not remove the discipline greediness supplied
R14 <i>Selection Gain as Commitment Temperature</i>	WR-27	explicitly a gain; instrumented at birth or diagnosed retroactively
R15 <i>Hierarchical and Temporal Depth</i>	WR-24	same constraint-removal shape as R13, one level up
R16 <i>Grounded Actuation, Not Re-Observation</i>	WR-4	act on the surface that exists rather than re-observing it; WR-4: inhabit before building
R17 <i>Structure Learning by Bayesian Model Reduction</i>	WR-19	reduction proposes; the loop must still generate candidates
R17' <i>Posterior Spread as an Exploration Control</i>	WR-19	spread is the generator's fuel, not a ranking tiebreak
R17'' <i>Embedding-Nearness as the Structure-Learning Generator</i>	WR-25	VERBATIM in the pattern: "Invariant to preserve: whatever proposes the merges must not also be what accepts them." This is aif/no-self-certification stated inline.
R17''' <i>Pattern Genesis from Evidence-Bearing Holes</i>	WR-9	a hole becomes a pattern only on evidence
R18 <i>Faithfulness Audit</i>	WR-9	The audit classifies every named quantity as FEP-derived, principled approximation, analogical ("a heuristic wearing the name") or engineering control, with plain-language truth-in-labelling: "claims to be... / actually computes...". That is WR-9's discipline – a bite is empirically tested or MARKED as logical – applied to the theory's own vocabulary. Documented in futon2/holes/aif-wiring-explainer.html (2026-07-03), not in the paper's catalogue.
R9 <i>No Self-Certification</i>	WR-0	WR-0's third property – a verdict the apparatus cannot issue to itself. The redrawn control map places "No self-certification" in the assurance-and-learning band; the catalogue's operator-acceptance line is explicitly "an acceptance invariant in the spirit of R9", i.e. an application of it.
R20 <i>Interoceptive Tripwires</i>	WR-26	"freeze, record, park, summon" plus "a wire that does not fire on the known past is a dud laser" – a wire switched off must carry its re-arm condition

Unmapped.

- R10 — Runs the machine on a schedule — the cron entypoint that makes the loop autonomous rather than operator-triggered. Currently off, because the machine is still being tuned; that is an operating state, not a defect, and it carries no ring in the figure. The No-Op Trap in its title is the failure it guards against: a scheduled run that fires, finds nothing it is willing to do, and reports success.
- R11 — Chooses among candidate actions under a budget, across levels — so that a cheap local move and an expensive strategic one are compared on the same scale rather than in separate queues. No war-room constraint governs it: nothing in the stack has yet gone wrong in a way that forced a ruling about budgeted selection.
- R19 — Not missing — deliberately relabelled. It is unmapped for a structural reason rather than an editorial one: preference is operator-supplied, and the operator is not in the model, so there is no node in the frame for it to terminate at. This is the subject of the operator section.

A note on where the catalogue ends. Of the 20 mechanisms factored here, the historical factoring scheme names R18, which is not a canonical assignment in the current control-map registry. Here it is the *faithfulness audit*: it classifies every named quantity in the loop as derived from the free-energy principle, a principled approximation, analogical — “a heuristic wearing the name” — or an engineering control that makes no theoretical claim, and attaches to each a plain-language truth-in-labelling of the form *claims to be...* / *actually computes...* The small dots along the top edge of every node in Figure 5 are its output. This historical identifier must not be joined to the current catalogue by number alone. The issue board retains that numbering distinction.

6 The Spine: Summary of the War-Room Patterns

The patterns below are the stack’s own rulings, promoted into the pattern library and reproduced here from it. Almost all were forced by something going wrong; WR-0 is the exception, and was forced by the occasion of writing this paper. They are implementation-non-specific on purpose: they say what must hold, and give reasons for building one thing rather than another. Each is stated here as its conclusion alone. The full IF/HOWEVER/THEN/BECAUSE entries, and the case that forced each ruling, are in Appendix A.

Organise As Effectively As War, Without Becoming A State Apparatus (WR-0). An apparatus built to organise effectively must carry, from the outset, the three properties that stop it becoming an authority: an open model, an available exit, and a verdict it cannot issue to itself. Effective organisation and unaccountable apparatus are the same construction seen from two sides; only those properties separate them.

Inhabit Before Building (WR-4). Inhabit the surfaces that already exist before building new infrastructure; new infrastructure is justified only when it makes an existing surface more inhabitable.

Typed Files Are Sources Of Truth, Prose Is Regenerated (WR-8). Make the typed argument files the canonical self-model and generate prose from them, so drift becomes detectable instead of perennial.

Bites Must Be Empirically Tested Or Marked Logical (WR-9). Render self-model conflict claims as empirical only when session evidence supports them, and as logical otherwise, so the inhabitation gap stares back from the demo.

Operationalised Exploit-Loops Are First-Class Observation Channels (WR-16). Wire operationalised exploit-loops into the stack’s status dashboard and the War Machine as direct typed evidence streams and first-class active-inference observation channels.

Tension Must GENERATE, Not Only Rank (WR-19). Wire structural tension as a candidate source that generates actions, not merely as a weight that ranks candidates generated elsewhere.

A Removed Constraint Does Not Remove the Discipline It Supplied (WR-24). When retiring a proxy constraint, install its replacement discipline in the same act and measure how much external verdict content changes subsequent work, not merely whether a verdict arrived.

Good News Gets the Same Evidence Discipline as Bad (WR-25). Let unanchored good news change ordering immediately but leave numerical posteriors unchanged until an artifact supports it.

A Capability Switched Off Carries Its Re-Arm Condition, in Writing (WR-26). Write a countable re-arm condition at the switch whenever an autonomous capability is turned off.

A Loop Is Born Instrumented for Its Gain (WR-27). Ship every new consumption-to-acquisition loop with an instrument that measures how much returned demand changes what the loop does next.

Part II. Closing the Loop over the Operator

7 The Operator Is Already Traceable

The closure argued in §8 needs the operator to be *observable* before it can be modelled. That is an empirical claim about this stack, not a design intention, so it is established separately and first.

Agent-authored commits record system output; they do not by themselves measure operator attention. What records the operator is the evidence landscape, and it records him at the granularity of the *turn*.

Three mechanisms, built for unrelated reasons, bear on the same question.

Route A: clock-in. A turn taken while a mission is clocked carries that mission on the turn record. This is not fully reliable — the clock is sometimes not set, and the gap is visible rather than hidden.

Route B: retrieval. The retrieval route embeds turns and matches them against the pattern library, and the result is deposited as a context-retrieval evidence event carrying pattern ids and cosine scores. Each recorded event carries 3 ranked candidates, but the operator's interface commits to the top one — it resolves that pattern's sigil and shows it beside the turn. The counts and the figure below therefore use rank 1 only. Counting all 3 would inflate every figure in this section by roughly threefold while describing something no one ever saw.

Route C: mention. A mission document may simply name a pattern in its text. This needs no turn and no clock at all; it is a property of the corpus, and 240 mission documents were scanned for it.

Each route was built independently; none was designed to supply the relation measured by either of the others. Route A answers *what was being worked under*; route B answers *what is this turn about*; route C answers *where is this pattern written down*. As typed edges:

- **A:** turn → clocked mission
- **B:** turn → retrieved pattern
- **C:** pattern → mentioning mission

A and B meet on the turn id, and that coincidence supplies a derived association between a pattern and a mission. It is not the authored citation used in the cascade of §5. C reaches the same relation from the other end, without needing the operator to have been clocked in correctly — or at all.

In the window measured for Figure 4, 27 operator turns carried both annotations, yielding 27 turn→pattern edges over 24 patterns and 2 missions. Of those edges, 11 name a pattern the curated layer already cites; 16 do not.

Route C then reaches 11 of those 24 patterns, adding 28 pattern→mission edges across 24 missions — against the 2 that clock-in alone supplies. And the two disagree completely: of the patterns route C reaches, 0 are mentioned in the mission the operator was actually clocked into.

The routes answer different questions: clocked session context and textual mention. Their disagreement does not identify an incorrect clock, cross-mission work or an omitted citation without further evidence. Likewise,

Three routes from an operator turn to a mission

B goes through the pattern; A goes straight to the clocked mission. They land in different places — that is the finding.

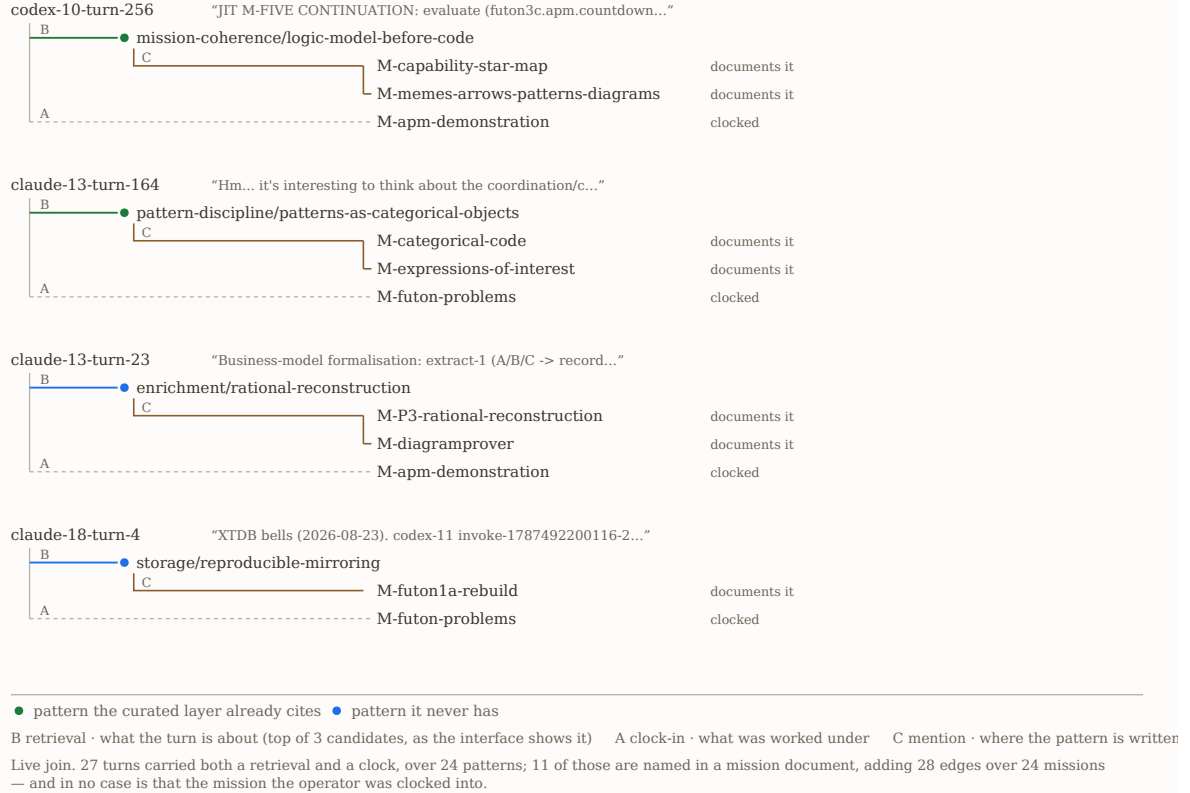


Fig. 4 – Four turns, each with its three answers. Route B (solid, coloured) goes through the pattern; route C (brown) continues from that pattern to the missions whose documents name it; route A (dashed) goes straight from the turn to the mission the operator was clocked into. A and C reach the same kind of node through different relations. Green marks a pattern the curated layer already cites, blue one it never has. The following prose gives the generated full-window counts.

the 16 retrieved associations absent from the curated set of 323 patterns identify candidates for inspection, not demonstrated attention or missing warranted citations.

Three cautions, stated here because they bound what the figure supports.

First, retrieval is not use. Embedding proximity proposes a topical association; it does not establish that the pattern was applied. An application receipt would need separate evidence; nothing in this section supplies it.

Second, routes A and B overlap only where both fire. Route A's coverage is partial, so 27 is a floor set by the weaker mechanism, not a measure of how much of the operator's attention is traceable in principle. Improving the clock raises it directly — and route C sidesteps it, which is much of the point of having a third route rather than a better second one.

Third, and most important for what may be built on this: these derived edges must not be merged into the curated layer. Their provenance is different — a cosine score, not a human citation — and merging them would destroy the distinction that makes the gap measurable at all. They belong in a separate relation carrying score and turn id, where agreement with a curated citation is a cross-check rather than an assumption.

With those caveats the claim is modest and checkable: the operator is not outside the record. He is in it, per turn, at a rate that can be counted, and the chain from a turn he took to a mission he was working on runs through the pattern library. §8 argues what to do with that; this section establishes only that there is something to do it with.

8 Modelling the operator's response

Section 7 establishes partial traceability. It does not establish what the operator attended to, why a proposal went unanswered or whether the proposal was useful. These distinctions matter when the harness uses follow-through to adjust future proposals.

In the audited outer-loop record, one action class received 108 unanswered proposals, as recorded in WR-0 (`futon3/library/war-room/wr-0-organise-without-apparatus.flexiarg`). Its learned credit fell. The record showed emissions without a judgement, so the decrease could not distinguish rejection from absence. The notification interface offered `nag`, `brief` and `silent`; these describe the system's communication, not the operator's alternative activities. A better notification policy alone would not identify the missing response.

The proposed remedy is to record considered-and-accepted, considered-and-declined and no-response separately, with an explicit basis for any claim that a proposal was considered. Traces may help decide when to ask, but cannot supply that judgement. Whether such a model improves selection remains an intervention to test, not a result of finding the traces.

Reading the mechanism maps

The following drawings answer different questions. The war-room overlay relates constraints to mechanisms in its recorded census. The registry-generated map shows wiring claims and recorded routes; the equation graph shows dependencies between symbols. A dependency edge is not evidence that a call occurred. Historical marks must be read at their source dates; current repair status is on the issue board.

Box 2 – The drawn control map checked against the wiring the equations force. A theory edge is a symbol defined at one box and imported by an equation at another; plumbing edges touch a box that runs no equation.

Class	n	Edges
Theory edge, drawn	14	R1→R3, R1→R3a, R1→R4, R14→R6, R16→R2, R2→R3a, R2→R8, R3→R1, R3a→R3, R3a→R7, R4→R5, R5→R6, R6→R4, R7→R3
Theory edge, conditional and drawn	3	R13→R4, R4→R8, R8→R6
Theory edge, realised in code, not drawn	0	—
Retired importer (historical; not current wiring)	1	R3a→R8
Theory edge, not realised	2	R1→R17, R2→R17
Theory edge, path-dependent	1	R6→R16
Drawn, plumbing (no equation at an endpoint)	8	R10→R8, R11→R16, R12→R7, R15→R13, R15→R16, R20→R7, R6→R11, R9→R16
Drawn between equation boxes, no equation across it	7	R13→R14, R14→R16, R2→R3, R6→R13, R7→R14, R7→R8, R8→R5

18 equations, 7 exogenous symbols, 35 free choices recorded (1 awaiting decision). Sources: aif-equations 2026-09-01 (2fed9f7c) · control-stages 0dc90da2 · control-map 2026-09-12 (0c7ee757).

Not realised means the equation at the target runs without an input the formalism gives it. Da Costa eq. 21 accumulates Dirichlet concentrations from observed (o, s) pairs. The WM's a4a accumulates a capability x mission model from substrate hyperedges and discharges instead; no path from the tick model's o or μ into R17 was found.

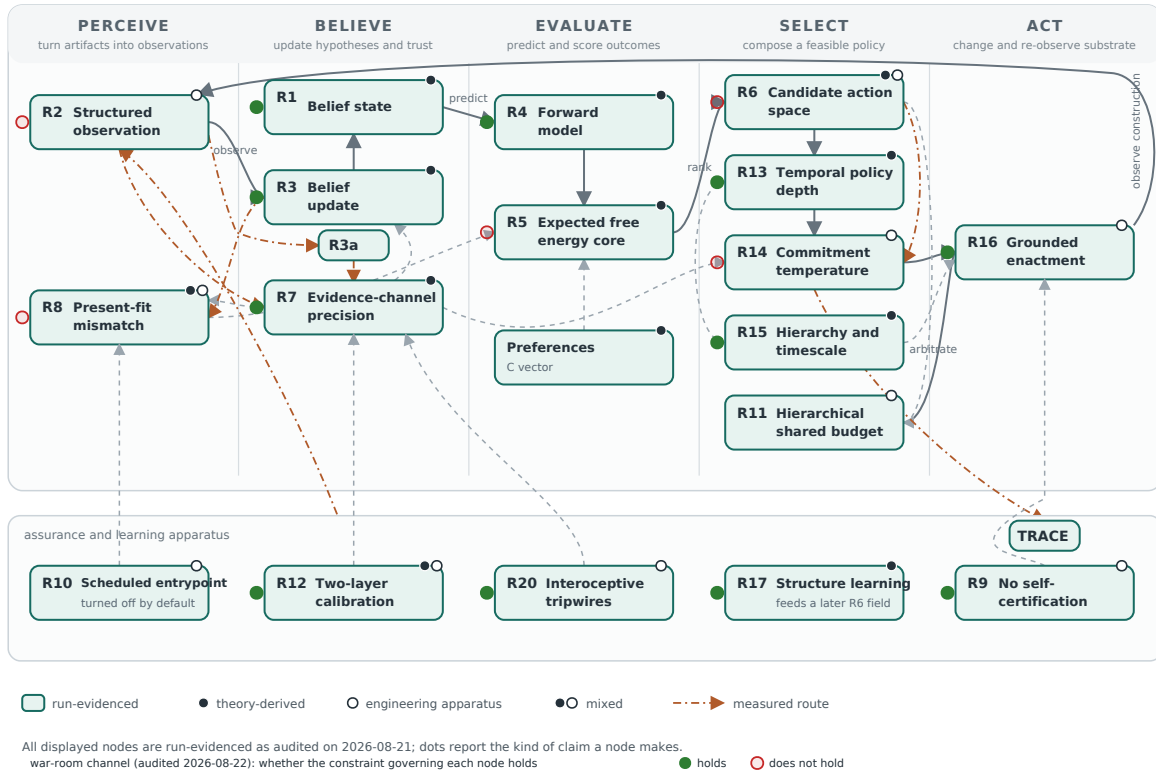


Fig. 5 – War-room constraints over the control map. Rings report the census in `empirics-futon/wr-overlay.edn`, including its recorded July cadence failures. They identify interfaces to investigate, not a fresh census of live execution.

From traceability to a tested response model

Clock-in already supplies some turn-to-mission links. The unresolved questions are its coverage and interpretation, not the absence of every such link. Retrieval adds candidate associations; mission prose supplies authored citations. None of these records proves attention or pattern use. An explicit application receipt would need to identify the actor, pattern, action and witness, while retaining unknown or disputed attribution.

A useful experiment would compare a declared response model with the current interpretation on the same recorded proposals, then test the effect of any changed selection rule. It would preserve unanswered cases as such, record the cost of eliciting a response, and distinguish a confirmed decline from an unseen proposal. It must not infer success from a higher response rate alone.

The historical backlog map illustrates why a proposed repair also needs a named implementation. Overlaying a mechanism gap on a mission deposit exposed places where no mission supplied the missing behaviour. That is a comparison of scopes, not evidence that a planned mission later ran.

What would consume the traces?

Three candidate consumers need different evidence: retrieval can use textual similarity; correction detection needs a relation between a statement and its revision; attribution of work needs an actor's confirmation or an

Checked topology of the active-inference harness

Generated from Figure 4's edge, stage, and organization registries; not hand edited

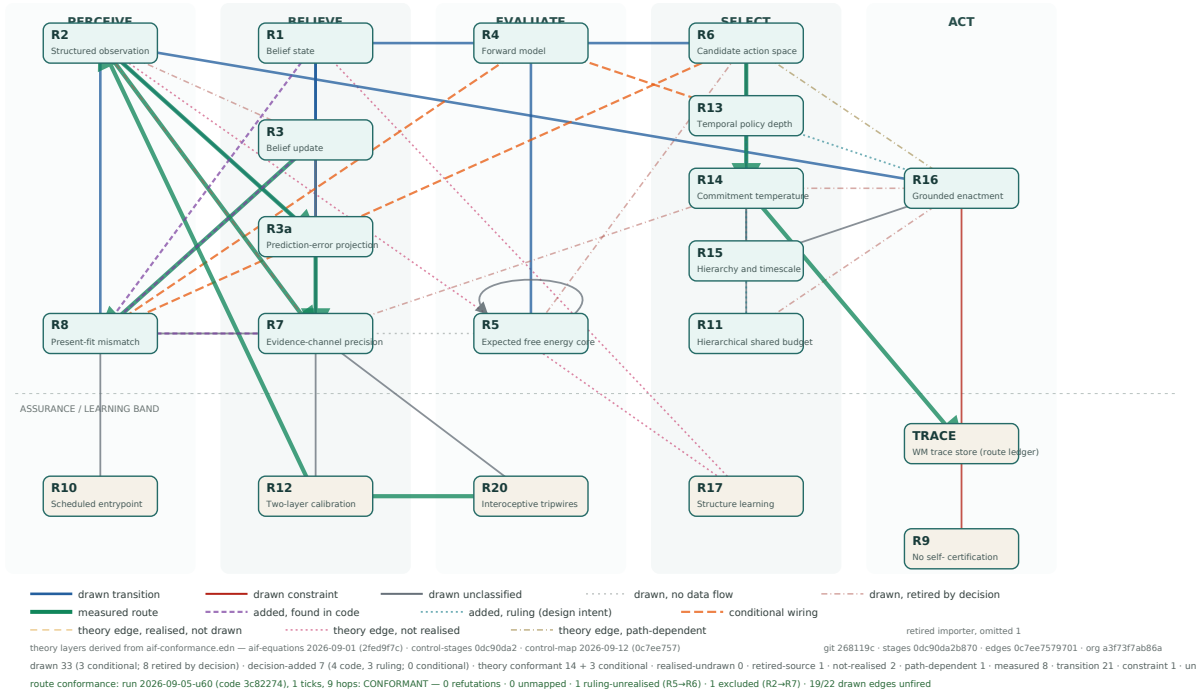


Fig. 5A – Registry-generated wiring and recorded routes. The legend distinguishes observed, conditional, undrawn and retired relations; the footer pins the source. A route is an instrumented call sequence, not the dependency graph in Figure 6A.

equivalent witness. Joining their records is useful only if those differences survive the join. A retrieved neighbour cannot silently become an applied pattern, and an agent's commit cannot silently become an observation of operator attention.

The per-node dossiers that follow expose the corresponding implementation and receipt basis. They are a way to locate and test these obligations, not a claim that the operator-feedback loop has been closed.

9 What Each Node Computes: The Per-Node Dossiers

Every number, rung and pointer below is generated from the registries named in the file header; nothing in this section is typed by hand. **The numbering is the catalogue numbering of control-stages.edn**, which is read off the drawn control map, and not the older contract numbering the two schemes diverge on.

The readiness scale is ordered (named → type-transcribed → formula-transcribed → witnessed → constructed → wired → validated → run-correlated) and no rung is skippable. Three rungs are derivable from these registries and only three: a node is NAMED because the roster declares it, TYPE-TRANSCRIBED because an equation it runs names a Lean declaration in the checked union and found in its pinned owning module, and FORMULA-TRANSCRIBED because the rung below is licensed and the registry carries the formal line. Everything above that needs an evidence type these registries do not carry. A node whose required construction has no inhabitant in the fundamentals census is *capped*: transcription cannot climb past an uninhabited constructor. Each rung prints

What each box needs to run: the wiring the equations force, checked against the drawn control map

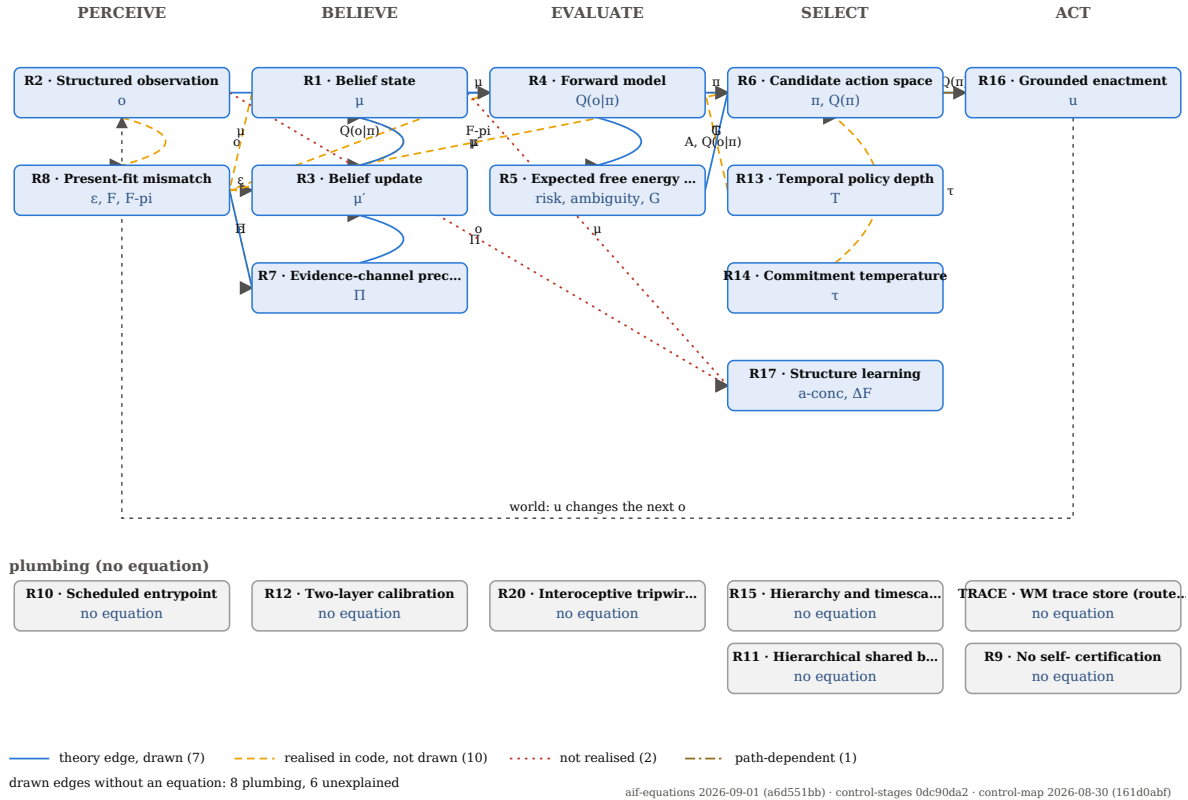


Fig. 6A – Equation dependencies. An edge identifies a symbol imported by its target; blue edges also appear in the wiring map, amber dashed edges do not. Grey nodes define no equation. This compares representations, without declaring either a witnessed execution.

the pointer that licenses it, and the generator resolves every one of those pointers — file present, line inside the file — before writing this file; a rung claim it cannot resolve stops the build rather than printing.

Each node also carries the rung of every Lean declaration its equations bind, read from the per-declaration accounting rather than restated here. The scale is the same eight rungs in both places and that is checked, not assumed: if the accounting's rung vocabulary and this section's differ by one element, nothing renders. The two say different things about the same node — a node reaches FORMULA-TRANSCRIBED because an equation at it carries a formal line, while a declaration reaches it because a check elaborates its defining expression — and reading them side by side is how a node whose registry entry outruns its declarations becomes visible.

1 of 20 nodes at their obligated state; 19 short: R2, R8, R1, R3, R3a, R7, R5, R6, R13, R14, R15, R11, R16, R10, R12, R20, R17, TRACE, R9. Equation-bearing nodes have the full AIF ladder as their obligation, limited by any FUNDAMENTALS cap. Plumbing has named as its terminal formal rung, but its obligated state also requires all seven node-linked lifecycle cells to exist. Reaching a capped rung does not inhabit the missing construction or establish runtime validity; the remaining construction debts are listed below.

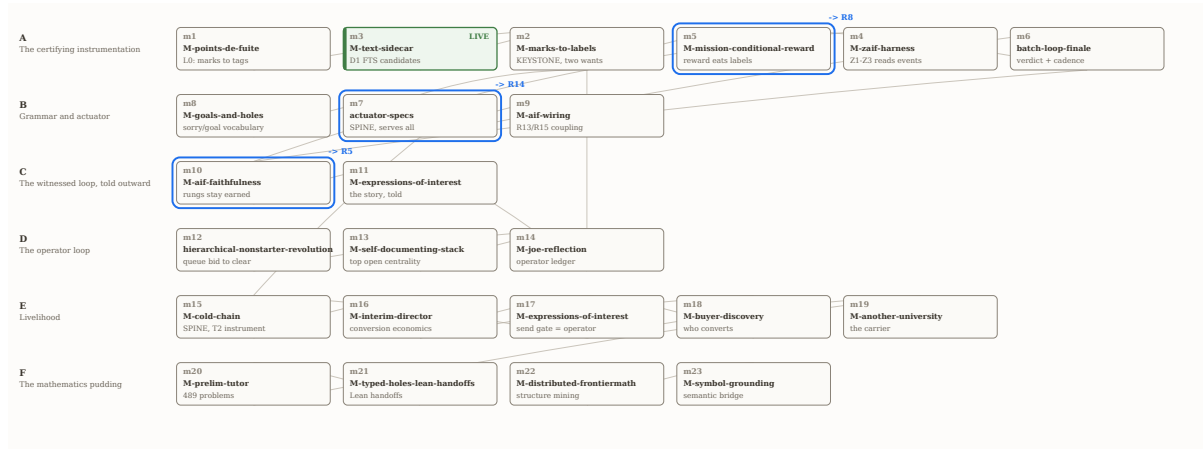


Fig. 7 – Historical mission-deposit overlay, from `empirics-futon/cascade-map.edn`. Blue rings connect proposed work to mechanism gaps. The recorded live mark is a bounded probe; it does not establish the other missions or present-day operation.

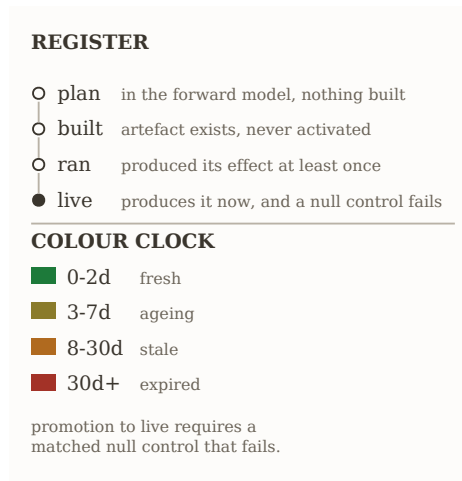


Fig. 8 – Evidence strength and age are separate dimensions. A record of a plan, build, run or live probe must retain when and how it was obtained. The issue board uses its own workflow columns and preserves these evidential distinctions.

6 required constructions are uninhabited in the fundamentals census. The join is by Lean declaration name and by nothing else, because `FUNDAMENTALS.edn` carries no node field: it places 1 of them at a node and leaves 5 at no node (`fundamental/belief-to-state-distribution`, `fundamental/controlled-transition-kernel`, `fundamental/policy-conditioned-state-predictive`, `fundamental/machine-preference-distribution`, `fundamental/parameter-kernels`). Placing those by hand is exactly the judgement this section is built not to make.

The seven lifecycle positions, in order, are commissioned, dispatched, parked, returned, checked, recorded, surfaced. Legend: + exists (code checks or records the cell), o named-only (conduct named without a running check

or record), – absent; means outside the census scope (not censused). Absent means no NODE-LINKED implementation: generic Agency machinery exists uncredited in `ALIGN-rnode-process-census.md`. Closing that join is the census’s separate candidate repair, not an effect of this presentation.

Box 3 – Per-node readiness, generated. *eqs* counts the equations the registry places at the node; *process assurance* is the lifecycle census of `ALIGN-rnode-process-census.md`, whose scope is 8 nodes: R9, R10, R12, R20, TRACE, R15, R11, R16.

node	stage	band	eqs	rung	obligated terminal rung	process assurance
R2	PERCEIVE	loop	1	formula-transcribed	run-correlated	
R8	PERCEIVE	loop	2	formula-transcribed	run-correlated	
R1	BELIEVE	loop	1	formula-transcribed	run-correlated	
R3	BELIEVE	loop	1	formula-transcribed	run-correlated	
R3a	BELIEVE	loop	1	formula-transcribed	run-correlated	
R7	BELIEVE	loop	1	formula-transcribed	run-correlated	
R4	EVALUATE	loop	1	formula-transcribed	formula-transcribed	
R5	EVALUATE	loop	3	formula-transcribed	run-correlated	
R6	SELECT	loop	2	named	run-correlated	
R13	SELECT	loop	1	formula-transcribed	run-correlated	
R14	SELECT	loop	1	formula-transcribed	run-correlated	
R15	SELECT	loop	0	named	named	-----
R11	SELECT	loop	0	named	named	-----
R16	ACT	loop	1	formula-transcribed	run-correlated	++++++
R10	PERCEIVE	assurance	0	named	named	++-----
R12	BELIEVE	assurance	0	named	named	---+---
R20	EVALUATE	assurance	0	named	named	----++
R17	SELECT	assurance	2	formula-transcribed	run-correlated	
TRACE	ACT	assurance	0	named	named	-----++
R9	ACT	assurance	0	named	named	----o--

R2 – Structured observation.

- *observe* defines `o: o_t <- structured observation of the world after action u_{t-1} – Lean machineObservation (closed); source friston2017.`
- **Readiness** `FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:16); type-transcribed (contract declaration machineObservation (closed), mathlib4/DarkTower/WarMachine/MachineObservation.lean:58); formula-transcribed (formal line for observe, futon2/holes/labs/wm-contract/aif-equations.edn:74).`
- **Declaration readiness** `(:F6 accounting, highest TYPE-TRANSCRIBED): machineObservation TYPE-TRANSCRIBED (mathlib4/DarkTower/WarMachine/MachineObservation.lean:58).`
- **Next rung not licensed** `(WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).`
- **Process assurance**, so no cell of the lifecycle is claimed here.

R8 – Present-fit mismatch.

- *free-energy* defines `F: F = 1/2 * mean_k (Pi_k eps_k^2) – Lean variationalFreeEnergy (closed); source buckley2017.`

- *policy-free-energy* defines $F_{\pi} : F_{\pi} := \sum_k \frac{1}{2} (\ln(2 \pi v_k(\pi)) + (o_k - \mu_k(\pi))^2 / v_k(\pi))$ — Lean `machinePolicyFreeEnergy` (closed); source `parr2022`.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, `p4ng/empirics-futon/control-stages.edn:17`); type-transcribed (contract declaration `variationalFreeEnergy` (closed), `mathlib4/DarkTower/WarMachine/Holes.lean:7092`); formula-transcribed (formal line for free-energy, `futon2/holes/labs/wm-contract/aif-equations.edn:92`).
- **Declaration readiness** (:F6 accounting, highest FORMULA-TRANSCRIBED): `variationalFreeEnergy` FORMULA-TRANSCRIBED (`futon2/checks/witness-registry.edn:697`).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).
- **Process assurance**, so no cell of the lifecycle is claimed here.
- **Open** free-energy: retired; policy-free-energy: realised-flag-gated.

R1 — Belief state.

- *belief-state* defines $\mu : \mu_t := \text{the stored belief after the update at } t-1$ — Lean `machineBeliefState` (closed); source `buckley2017`.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, `p4ng/empirics-futon/control-stages.edn:18`); type-transcribed (contract declaration `machineBeliefState` (closed), `mathlib4/DarkTower/WarMachine/MachineBeliefState.lean:25`); formula-transcribed (formal line for belief-state, `futon2/holes/labs/wm-contract/aif-equations.edn:119`).
- **Declaration readiness** (:F6 accounting, highest TYPE-TRANSCRIBED): `machineBeliefState` TYPE-TRANSCRIBED (`mathlib4/DarkTower/WarMachine/MachineBeliefState.lean:25`).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).
- **Process assurance**, so no cell of the lifecycle is claimed here.

R3 — Belief update.

- *belief-update* defines $\mu_{\text{next}} : \mu \leftarrow \mu + \alpha \Pi \text{ eps}$ — Lean `machineBeliefUpdate` (closed); source `buckley2017`.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, `p4ng/empirics-futon/control-stages.edn:19`); type-transcribed (contract declaration `machineBeliefUpdate` (closed), `mathlib4/DarkTower/WarMachine/MachineBeliefUpdate.lean:181`); formula-transcribed (formal line for belief-update, `futon2/holes/labs/wm-contract/aif-equations.edn:113`).
- **Declaration readiness** (:F6 accounting, highest TYPE-TRANSCRIBED): `machineBeliefUpdate` TYPE-TRANSCRIBED (`mathlib4/DarkTower/WarMachine/MachineBeliefUpdate.lean:181`).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).
- **Process assurance**, so no cell of the lifecycle is claimed here.

R3a — Prediction-error projection.

- *prediction-error* defines $\text{eps} : \text{eps}_k := o_k - \mu_k$ — Lean `machineChannelPredictionError` (closed); source `buckley2017`.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, `p4ng/empirics-futon/control-stages.edn:20`); type-transcribed (contract declaration `machineChannelPredictionError` (closed), `mathlib4/DarkTower/WarMachine/MachinePredictionError.lean:160`); formula-transcribed (formal line for prediction-error, `futon2/holes/labs/wm-contract/aif-equations.edn:80`).

- **Declaration readiness** (:F6 accounting, highest TYPE-TRANSCRIBED): machineChannelPredictionError TYPE-TRANSCRIBED (mathlib4/DarkTower/WarMachine/MachinePredictionError.lean:160).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).
- **Process assurance**, so no cell of the lifecycle is claimed here.

R7 — Evidence-channel precision.

- *precision* defines $P_i: P_{i_k} := 1 / \max(\text{Var}(\text{eps_k}), \text{eps0})$ — Lean machinePrecision (closed); source buckley2017.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:21); type-transcribed (contract declaration machinePrecision (closed), mathlib4/DarkTower/WarMachine/MachinePrecision.lean:101); formula-transcribed (formal line for precision, futon2/holes/labs/wm-contract/aif-equations.edn:86).
- **Declaration readiness** (:F6 accounting, highest TYPE-TRANSCRIBED): machinePrecision TYPE-TRANSCRIBED (mathlib4/DarkTower/WarMachine/MachinePrecision.lean:101).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).
- **Process assurance**, so no cell of the lifecycle is claimed here.

R4 — Forward model.

- *forward-model* defines $Q\text{-o-pi}: Q(o|pi) := \sum_s A(o|s) Q(s|pi)$, $Q(s|pi)$ rolled forward by B from mu over depth T — Lean PredictiveOutcomeKernel (carrier-only); source dacosta2020.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:22); type-transcribed (contract declaration PredictiveOutcomeKernel (closed), mathlib4/DarkTower/WarMachine/Holes.lean:7022); formula-transcribed (formal line for forward-model, futon2/holes/labs/wm-contract/aif-equations.edn:126).
- **Declaration readiness** (:F6 accounting, highest FORMULA-TRANSCRIBED): PredictiveOutcomeKernel FORMULA-TRANSCRIBED (futon2/checks/witness-registry.edn:521).
- **Capped** at FORMULA-TRANSCRIBED: a required construction at this node has no inhabitant (FUNDAMENTALS census).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).
- **Process assurance**, so no cell of the lifecycle is claimed here.
- **Open** forward-model: Lean binding carrier-only; FUNDAMENTALS : in $Q(o|pi)$ (the machine's predictive outcome kernel, constructed) has no inhabitant.

R5 — Expected free energy core.

- *risk* defines risk: $\text{risk}(pi) := D_{\text{KL}}[Q(o|pi) || C]$ — Lean predictiveOutcomeRisk (closed); source dacosta2020.
- *ambiguity* defines ambiguity: $\text{ambiguity}(pi) := E_{\{Q(s|pi)\}}[H(P(o|s))]$ — Lean ambiguity (closed); source dacosta2020.
- *expected-free-energy* defines G: $G(pi) := \text{risk}(pi) + \text{ambiguity}(pi)$ — Lean expectedFreeEnergy (closed); source dacosta2020.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:23); type-transcribed (contract declaration ambiguity (closed), mathlib4/

DarkTower/WarMachine/Holes.lean:7144); formula-transcribed (formal line for risk, futon2/holes/labs/wm-contract/aif-equations.edn:131).

- **Declaration readiness** (:F6 accounting, highest FORMULA-TRANSCRIBED): ambiguity FORMULA-TRANSCRIBED (futon2/checks/witness-registry.edn:725); expectedFreeEnergy FORMULA-TRANSCRIBED (futon2/checks/witness-registry.edn:788); predictiveOutcomeRisk FORMULA-TRANSCRIBED (futon2/checks/witness-registry.edn:621).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).
- **Process assurance**, so no cell of the lifecycle is claimed here.

R6 – Candidate action space.

- *policy-set* defines π : π ranges over the candidate action space – Lean machinePolicySet (closed); source friston2017.
- *policy-posterior* defines Q - π : $Q(\pi) := \text{softmax}(\ln E(\pi) - G(\pi)/\tau - F_{\pi}(\pi))$ – Lean softmaxWithFPi (closed); source parr2022.
- **Readiness** NAMED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:24).
- **Next rung not licensed** (TYPE-TRANSCRIBED): no equation at this node names a Lean declaration that is both in the checked contract union and found in its owning module.
- **Process assurance**, so no cell of the lifecycle is claimed here.
- **Open** open edge R6->R16 (path-dependent).

R13 – Temporal policy depth.

- *depth* defines T : $T := \text{temporal policy depth (the range of the sums in } Q(o|\pi) \text{ and } G)$ – Lean machineDepth (closed); source dacosta2020.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:25); type-transcribed (contract declaration machineDepth (closed), mathlib4/DarkTower/WarMachine/MachineDepth.lean:133); formula-transcribed (formal line for depth, futon2/holes/labs/wm-contract/aif-equations.edn:165).
- **Declaration readiness** (:F6 accounting, highest TYPE-TRANSCRIBED): machineDepth TYPE-TRANSCRIBED (mathlib4/DarkTower/WarMachine/MachineDepth.lean:133).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).
- **Process assurance**, so no cell of the lifecycle is claimed here.

R14 – Commitment temperature.

- *temperature* defines τ : $\tau := \text{commitment temperature (inverse precision gamma of policy selection)}$ – Lean machineTemperature (closed); source friston2017.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:26); type-transcribed (contract declaration machineTemperature (closed), mathlib4/DarkTower/WarMachine/MachineTemperature.lean:42); formula-transcribed (formal line for temperature, futon2/holes/labs/wm-contract/aif-equations.edn:171).
- **Declaration readiness** (:F6 accounting, highest TYPE-TRANSCRIBED): machineTemperature TYPE-TRANSCRIBED (mathlib4/DarkTower/WarMachine/MachineTemperature.lean:42).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).

- **Process assurance**, so no cell of the lifecycle is claimed here.

R15 — Hierarchy and timescale.

- **Process assurance** ----- (futon2/holes/labs/wm-contract/ALIGN-rnode-process-census.md:221).
- Named is the terminal AIF rung. A Lean specification is possible but not obliged for AIF-completeness; the lifecycle strip is primary.
- Runs no equation in the registry (declared plumbing).
- **Readiness** NAMED (derived): named (declared in the control-stages roster, png/empirics-futon/control-stages.edn:27).
- **Open** runs no equation; declared plumbing in the registry.

R11 — Hierarchical shared budget.

- **Process assurance** ----- (futon2/holes/labs/wm-contract/ALIGN-rnode-process-census.md:222).
- Named is the terminal AIF rung. A Lean specification is possible but not obliged for AIF-completeness; the lifecycle strip is primary.
- Runs no equation in the registry (declared plumbing).
- **Readiness** NAMED (derived): named (declared in the control-stages roster, png/empirics-futon/control-stages.edn:28).
- **Open** runs no equation; declared plumbing in the registry.

R16 — Grounded enactment.

- **action** defines $u: u_t := \operatorname{argmax}_u \sum_{\pi} \delta(u, \pi_t) Q(\pi)$ — Lean machineAction (closed); source friston2017.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, png/empirics-futon/control-stages.edn:29); type-transcribed (contract declaration machineAction (closed), mathlib4/DarkTower/WarMachine/MachineAction.lean:49); formula-transcribed (formal line for action, futon2/holes/labs/wm-contract/aif-equations.edn:184).
- **Declaration readiness** (:F6 accounting, highest TYPE-TRANSCRIBED): machineAction TYPE-TRANSCRIBED (mathlib4/DarkTower/WarMachine/MachineAction.lean:49).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).
- **Process assurance** ++++++ (futon2/holes/labs/wm-contract/ALIGN-rnode-process-census.md:223).
- **Open** open edge R6->R16 (path-dependent).

R10 — Scheduled endpoint.

- **Process assurance** +----- (futon2/holes/labs/wm-contract/ALIGN-rnode-process-census.md:217).
- Named is the terminal AIF rung. A Lean specification is possible but not obliged for AIF-completeness; the lifecycle strip is primary.
- Runs no equation in the registry (declared plumbing).
- **Readiness** NAMED (derived): named (declared in the control-stages roster, png/empirics-futon/control-stages.edn:30).
- **Open** runs no equation; declared plumbing in the registry.

R12 — Two-layer calibration.

- **Process assurance** ---+--- (futon2/holes/labs/wm-contract/ALIGN-rnode-process-census.md:218).
- Named is the terminal AIF rung. A Lean specification is possible but not obliged for AIF-completeness; the lifecycle strip is primary.

- Runs no equation in the registry (declared plumbing).
- **Readiness** NAMED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:31).
- **Open** runs no equation; declared plumbing in the registry.

R20 — Interoceptive tripwires.

- **Process assurance** ----++ (futon2/holes/labs/wm-contract/ALIGN-rnode-process-census.md:219).
- Named is the terminal AIF rung. A Lean specification is possible but not obliged for AIF-completeness; the lifecycle strip is primary.
- Runs no equation in the registry (declared plumbing).
- **Readiness** NAMED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:32).
- **Open** runs no equation; declared plumbing in the registry.

R17 — Structure learning.

- *dirichlet-accumulation* defines a-conc: $a = a + \text{sum_tau } o_tau(x) s_tau$ (Dirichlet concentrations accumulated from observed outcome/state pairs) — Lean DirichletConcentrations (carrier-only); source dacosta2020.
- *model-reduction* defines Delta-F: $\Delta F := \ln P(y|full) - \ln P(y|reduced) = \ln B(A) + \ln B(a') - \ln B(a) - \ln B(A')$ over Dirichlet concentrations a (prior), A (posterior), a' (reduced prior); a reduction is accepted when $\Delta F \leq -3$ — Lean modelReductionDecision (closed); source friston2018bmr.
- **Readiness** FORMULA-TRANSCRIBED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:33); type-transcribed (contract declaration DirichletConcentrations (closed), mathlib4/DarkTower/WarMachine/Holes.lean:7200); formula-transcribed (formal line for dirichlet-accumulation futon2/holes/labs/wm-contract/aif-equations.edn:192).
- **Declaration readiness** (:F6 accounting, highest FORMULA-TRANSCRIBED): DirichletConcentrations FORMULA-TRANSCRIBED (futon2/checks/witness-registry.edn:161).
- **Next rung not licensed** (WITNESSED): the rungs above formula-transcribed need an evidence type these registries do not carry (:F6).
- **Process assurance**, so no cell of the lifecycle is claimed here.
- **Open** dirichlet-accumulation: Lean binding carrier-only.

TRACE — WM trace store (route ledger).

- **Process assurance** ----++ (futon2/holes/labs/wm-contract/ALIGN-rnode-process-census.md:220).
- Named is the terminal AIF rung. A Lean specification is possible but not obliged for AIF-completeness; the lifecycle strip is primary.
- Runs no equation in the registry (declared plumbing).
- **Readiness** NAMED (derived): named (declared in the control-stages roster, p4ng/empirics-futon/control-stages.edn:34).
- **Open** runs no equation; declared plumbing in the registry.

R9 — No self- certification.

- **Process assurance** ----o-- (futon2/holes/labs/wm-contract/ALIGN-rnode-process-census.md:216).
- Named is the terminal AIF rung. A Lean specification is possible but not obliged for AIF-completeness; the lifecycle strip is primary.
- Runs no equation in the registry (declared plumbing).

- **Readiness** NAMED (derived): named (declared in the control-stages roster, `p4ng/empirics-futon/control-stages.edn:35`).
- **Open** runs no equation; declared plumbing in the registry.

10 Transfer is a further test

Even a successful operator-feedback loop would establish a capability in this project, not a validated diagnostic for another organisation. Transfer requires new observations, a declared outcome and a recipient entitled to judge it. It remains a proposal here.

A candidate setting would need recurring decisions, records of what was consulted before deciding, enough cycles to compare subsequent choices with prior outcomes, and capacity to act on a result within the relevant time horizon. These are proposed eligibility criteria. They have not been shown to predict adoption or improvement.

The first engagement would therefore be instrument construction on a new subject. Its acceptance condition should name the decision regime, the observable result and who will assess it. A useful diagnosis and an improvement caused by that diagnosis are distinct claims: the latter needs an intervention and a comparison, not just an informative report.

Section 11 outlines a retrospective corpus study that could help refine the instrument. Historical cases cannot establish prospective performance, and interest in a report does not establish willingness to adopt or pay for it. Neither commercial success nor organisational transfer is a result of this paper.

Part III. A Proposed Transfer Study

11 An unrun transfer study

The proposed study asks whether an object model can describe decisions in other organisations without concealing cases that do not fit. Its corpus is described as thirty-two cases, with six node types: requirement, benchmark, bounty, dependency, call and environment. The intended tests concern encoding residue and separation of known outcomes. Neither test is reported as passed here.

Before running the study, its records need reconciliation. Earlier prose used both thirty and thirty-two cases, asserted that all cases ended in interest rather than obligation, and also described a comparison with opposite terminals. Those statements cannot jointly support a result without a case-level account of the populations and terminal rule. They are withdrawn as findings here; reconciliation belongs to the study's preparation.

The proposed discipline is to cite each claim or mark it unverified with a condition that would settle it. Outcome-blind encoding must be demonstrated by the collection procedure. Reconstructing what was knowable at the time from historical accounts is not itself blinding. The project's own contemporaneous problem statements avoid one source of hindsight, but introduce author involvement; they are not automatically better evidence overall.

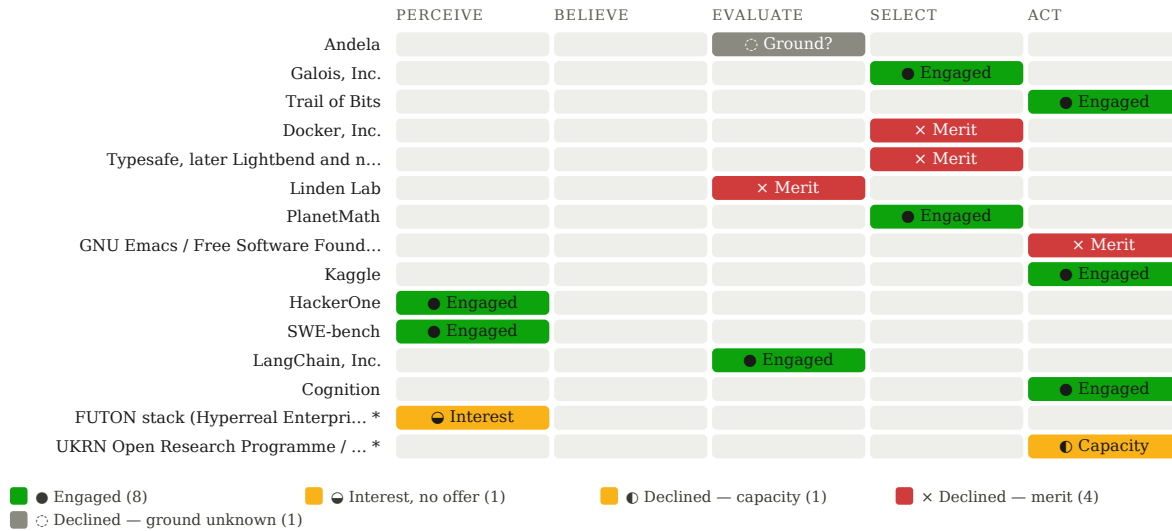
The study may report where the schema fits and where it leaves residue. It must retain an unmatched case rather than force it into a convenient category. The stated sample is too small for the proposed predictive claim, and no result here establishes that acceptance events explain value capture. A prospective trial would be a separate step, with its criterion and comparison fixed before the outcome is known.

Part IV. Beyond the Reporting Boundary

12 Recapitulation

The preceding parts distinguish findings from proposed tests. Read what follows against them rather than around them:

Part I The audit connects declared constraints to implementation evidence; component conformance does not establish the complete loop.



* an offer this stack made itself; every other row is the case's own recorded history.

Not located (17 of 32) — no source places the problem at a phase:

Runtime Verification, Inc., Kestrel Institute, JUXT, sponsor and developer of XTDB, Sourcegraph
 Redis Ltd. / Redis Labs / Redis Inc., Mojang Specifications, Roblox Corporation, Gitcoin, AppJet, Inc., Wolfram Research
 OpenHands (formerly OpenDevin) and All Hands AI, SWE-agent (Princeton NLP / Stanford researchers)
 SICA — A Self-Improving Coding Agent, VERSES AI, Inc., scoped to US Patent 12,393,581, Braintrust, Humanloop, Tessl

Fig. 9 – Source coverage in the recorded demand-side corpus. Columns locate the buyer's problem; colours describe the recorded response. The drawing includes cases for which a source locates the problem and lists the others below it. These are coverage observations, not results of the proposed acceptance experiment. Source: `futon0/analysis/business-models/records/batch-H-demand.edn`; generator: `lights.bb`.

Part II Operator traces support partial attribution, but cannot identify acceptance or attention without further evidence.

Part III, proposal Transfer to other organisations remains an unrun study, with unresolved corpus descriptions.

The remaining design question is:

Attribution and richer preferences require explicit observation domains, participation and witnessed consumers.

Each is decomposed into numbered steps in Appendix B, where every step declares what kind of support it has. The section that follows takes the fourth thesis apart.

13 Beyond the reporting boundary

A representation of current state needs new observations. A reusable rule usually needs less frequent revision, but still has a context and can be invalidated. This is the useful distinction between the plotting table and the portable pattern in Figure 1; neither is exempt from checking its basis.

Reporting also has a participation cost. An organisation may require updates from its members; a voluntary collaboration must give participants a reason and an affordable way to supply them. This motivates a design

question, not a measured market boundary: can publication and confirmation sustain a useful record where compulsory reporting is unavailable?

Attribution without inventing agreement

The working-tree provenance design joins tool-edit witnesses, file observations, commit links and mission-clock lineage. Each edge retains its basis. A clock records the declared working context; a path or retrieval match can suggest another association. An explicit confirmation can strengthen that attribution. The inferred association must remain distinguishable from the confirming act.

“No mission” is a valid result. A later claim should identify claimant and basis, retain any competing claims, and preserve the fact that the work was unclaimed at commit time. Such records can reveal work outside planned missions without rewriting its history to make the plan appear complete. Accept, decline and no-response must remain distinct here for the same reason as in the operator-feedback model.

The dated 2026-08-24 probe demonstrated only part of this design: 343 file observations and nine witnesses across fourteen roots, with eighty paths unattributed and no delivered followups. The attribution namespace was not loaded in that process. These are historical limits of that probe, not a current verdict on the service. A new claim of operation needs a new witnessed join and delivery.

Absence requires its own coverage evidence. A join with no result cannot show that no work occurred unless its producers, population and observation window are known. A heartbeat cannot establish task progress, and frequent retrieval cannot establish that every eligible turn was observed. This is where the apparatus pattern *Monitors measure the work* applies.

Policy scores and realised outcomes

Expected free energy scores predicted outcomes against declared preferences, with epistemic terms supplied by the model. A realised payoff answers a different question. Calling both a policy score does not establish that they are the same quantity.

The working notes’ four tests, s-g1–s-g4, ask whether a score derives from predicted outcomes and preferences; whether its trajectory can contain different actions; whether changing pattern wiring can change the score; and whether stipulated components are disclosed. These tests identify what a proposed policy demonstration must exhibit. A repeated-action rollout can be a restricted policy, but cannot alone demonstrate sensitivity to the ordering of different actions.

Appendix C demonstrates one of these distinctions concretely: changing the consultation order of unchanged patterns changes a game’s realised payoff from +3 to −5. That is evidence that wiring matters in the experiment. The payoff comparison alone does not establish all four tests for the War Machine’s expected-free-energy calculation, nor validate the live system. The model’s observations, preference carrier and actual consumers must be checked separately; Section 14 describes a specific gap in that composition.

Part IV therefore proposes an attribution and participation discipline, with bounded component evidence. It establishes neither general organisational transfer nor an integrated policy-learning result.

14 Preferences over problems: making C compute

14.1 Outcome preferences

Here C denotes a declared distribution over outcomes. It is distinct from epistemic value, which concerns what the model may learn, and from the policy prior E . The risk term compares a predicted distribution with C : $\text{risk}(\pi) = D_{\text{KL}}[Q(o \mid \pi) \parallel C]$. The comparison requires a common outcome domain; compatible numerical shapes are insufficient.

The project's tetrahedral outcome model distinguishes nouns, verbs, organisation and evidence. Its signature excludes preferences over evidence content: it requires $v \neq \text{evidence}$. Wanting a communicable witness to exist is a different requirement from wanting that witness to report a particular answer. That distinction must survive any richer outcome encoding.

Preferences may be indexed by horizon step, C_τ . This permits preferences about intermediate operation as well as terminal outcomes. It does not prove that the implementation represents both. Consultation during scoring and revision by an operator ruling are also separate operations: an extractor may propose a revision with provenance, but cannot adopt it.

14.2 What the channel model demonstrates

The recorded implementation uses thirteen per-channel healthy ranges, including `sorry-count-norm` and `mission-health`. A demonstration dated 2026-09-08, with predicted variance $\sigma^2 = 0.01$, compares a stressed state (0.85, 0.2) with a repaired state (0.25, 0.6). Their two channel risks are respectively $5.502 + 3.395$ and $0.194 + 0.455$ nats: a difference of 8.25 nats favouring repair.

This establishes discrimination on those channel readings. It does not identify which problem was solved or its value to the operator. Two problems with the same readings receive the same channel risk. Preferences over problem identity, solvability or benefit therefore require additional modelling and evidence.

14.3 The disposition bridge and its unresolved domain

A proposed bridge uses a conditional disposition model $P(d \mid o)$:

$$Q(d \mid \pi) = \sum_o P(d \mid o) Q(o \mid \pi), \quad \text{disposition-risk}(\pi) = D_{\text{KL}}[Q(d \mid \pi) \parallel C].$$

This construction requires the predictor and kernel to mean the same thing by o . It also retains the risk theorem's support conditions; an unobserved kind cannot acquire invented mass just to make an evaluation finite.

The fit reported on 2026-09-08 contains one observed trajectory shape and one outcome kind: three closed attempts, all grounded-change. Against the ruled seed mass $\frac{1}{2}$ for that kind, the calculation yields $\ln 2 = 0.6931$ nats. It is constant across policies on that record and supplies no selection discrimination.

A source check on 2026-09-09 found a separate problem. The fitted kernel reads checkpoint trajectories; its cohort ledger contains no channel-valued observation vector (`futon2/checks/disposition_kernel.clj:20-27`). The scorer supplies a channel-valued predicted mean (`futon2/src/futon2/aif/efe.clj:701-703`). The constant fit masks this mismatch. A justified relation between those observation domains is required before the composition can support the intended claim. Merely substituting one carrier for the other would change the question, not repair it.

14.4 A worked Snatch miniature

Measured 2026-09-09 from three pinned Snatch observations, the toy carrier is $d \in \{\text{grim cut, remedy, abstention}\}$. The declared miniature preference is $C = (1/6, 2/3, 1/6)$. Each fitted row has one observation: its observed ending has probability one and both unobserved endings retain explicit probability zero. For the grim, remedy-bearing pattern, and abstention-preserving policies, $Q(o \mid \pi)$ is the corresponding pinned point mass. Pushing it through $P(d \mid o)$ gives $Q(d \mid \pi)$, and the complete bridge computes $D_{\text{KL}}[Q(d \mid \pi) \parallel C]$ as 1.7918, 0.4055, and 1.7918 nats respectively. Thus the declared C selects the remedy-bearing policy. These are dated toy measurements, not a production preference ruling; the source record and every zero remain in the generated EDN. Source: `futon2/holes/labs/wm-contract/runs/U90-snatch-mini-c.edn`.

14.5 A proposed problem-level preference

Problem-indexed outcomes would let the operator distinguish which problems should be solved, alongside how attempts end and how the process runs. The declared axes include solvability, monetary benefit and societal benefit. An evidence extractor could propose masses along those axes, citing its basis in rulings, missions and recorded statements. A ruling would still be required to adopt the proposal.

This is a design direction. It needs justified observations, support handling and consumers before it can change a live selection. The existing channel demonstration and constant disposition fit do not establish those obligations.

15 From stated requirements to governed commons: the workshop transformation

This section walks one instance of the paper's title question end to end: a high-level problem framing is broken into design requirements, the requirements are restated as patterns, and the patterns are interpreted into Ostrom-style institutions—normative wiring that the same infrastructure can carry as data. The instance is small (one live standard repaired) but every step is real and on the record.

15.1 The apex framing and its requirements

The cascade-live view states its problem at the top, re-statable from data (`futon3c/holes/excursions/apex-thesis.json:1-66` status: `in-force`): *can a community of humans and machines doing open-ended work leave an account of itself structured enough that the work can steer itself—so that activity becomes inference rather than motion?* The apex decomposes into four requirement clusters: A, records carry warrant (“an assertion severed from its derivation cannot be checked, learned from, or recovered”); B, one queryable self-account; C, feedback reaches every participant; D, accountable next-action choice. Independently, the view checks five live standards against its substrate (`futon3c/src/futon3c/logic/cascade_real_live.clj:242-259`): regenerates, evidence, reconstitution, honest holes, composed.

The standards are design requirements in executable form—and on the day of this writing one of them was false. `s4-honest-holes` evaluated to `false` because the `honest-holes` dimension (O5) held zero rows, while `s5-composed` was true: the composition published with no hole record at all. An empty set is not evidence of absence; it may simply be an unread dimension.

15.2 Requirements as patterns (verbatim)

The transformation pipeline begins by restating requirements as *flexiargs*—the library's pattern carrier, whose fields (conclusion, context, IF/HOWEVER/THEN/BECAUSE, next-steps) will matter in the next step. Two were minted from the framing above, in a new workshop/ family. They are reproduced verbatim so their admission can be judged on their face.

```
@flexiarg workshop/records-carry-warrant
@title Records Carry Warrant
@provenance apex-thesis cluster A ... restated as a workshop flexiarg
  by zai-7 with Joe, 2026-09-12 ... Not mined from transcripts; derived
  from the stated design requirement.

! conclusion:
  Every claim a crew member emits must travel with the derivation that
  warrants it – the query, the input digest, the witness – such that any
  other participant can re-derive or refute it from the record alone.

+ HOWEVER:
  ... In those cases the warrant is a STATED ASSUMPTION with a version
  stamp, never silence (the witness-status precedent).
```

```

+ BECAUSE:
  An assertion severed from its derivation cannot be checked, learned
  from, or recovered (cluster A's ground, verbatim) – and a crew whose
  members cannot check each other's claims has no separation of
  powers, only manners.

@flexiarg workshop/honest-holes-gate-composed-claims
@title Honest Holes Gate Composed Claims
@provenance The cascade-live standards s1-s5 ..., where s4-honest-holes
  = boolean(seq o5) is TODAY FALSE (O5 has zero rows) while s5-composed
  is true – the composition publishes with no honest-hole record.

! conclusion:
  A composed claim (a cascade, a page, a report) may not present itself
  as complete while its honest holes are unrecorded: every composition
  step that could have holes must either carry them or carry a typed
  statement that none were found.

+ THEN:
  Before publish: for each composed dimension, emit either the hole
  list or a typed none { :checked-dimensions [...] :holes-found 0 };
  s4 becomes true only on typed evidence.

+ BECAUSE:
  ... Honesty about gaps is a precondition for the apex thesis: work
  steers itself only if it can see where it isn't.

```

(Full text with all fields: `futon3c/data/pattern-staging/workshop/records-carry-warrant.flexiarg:1-47` and `futon3c/data/pattern-staging/workshop/honest-holes-gate-composed-claims.flexiarg:1-46`. Both are process patterns, out of the mathematics retrieval pool by construction, and both are pending the operator's read—this section is that read.)

15.3 Patterns as institutions: the interpretation is the work

The library's patterns admit multiple interpretations, and the crew design assigns each interpretation its consumer: the informal Alexander-style reading feeds the map; a production-rule reading (guard $\rightarrow$ action, chainable) is what the Workshop—the board-authoring role—consumes to write wiring diagrams. The third reading is the one this section exercises: a pattern read as a *proto-institution*, combining it with preferences and deontic aspects (`futon2/holes/labs/wm-contract/NOTE-pattern-interpretations.md`).

The transformation of `honest-holes-gate-composed-claims` onto Ostrom's IAD rule types is nearly mechanical, which is the finding (`futon2/holes/labs/wm-contract/EXPERIMENT-flexiarg-to-institution.md`):

- **Boundary rules** $\leftarrow$ the pattern's *context* and *THEN*: a composition enters the regime when it aggregates more than one dimension into a publishable claim, and leaves it when each composed dimension carries either its hole list or a typed none. *Circumstances of application*—which requirements bind which artifacts—are exactly the boundary rules.
- **Choice rule** $\leftarrow$ *IF/THEN*: the compositor may publish at any time; it must not publish with an untyped-absent hole dimension.
- **Aggregation rule** $\leftarrow$ *HOWEVER*: the census is scope-limited to the composition's own dimensions—a cost bound inside the norm, not an exception to it.
- **Payoff rule** $\leftarrow$ *BECAUSE*: fogless maps are debt the whole commons pays, not savings for the compositor.

- **Monitoring** ← *next-steps* plus the register: graduated from a named counterexample carried on the pattern itself, to a register disagreement on the verifier’s channel. The monitor is an existing crew role (the verifier), not a new institution—Ostrom’s monitoring-by-participants.

The institution adds what the pattern lacked: positions (the pattern says *should*; the institution says *the compositor must*), graduated sanctions, and a named monitor. The deontic vocabulary is already runtime-native: a must-not is a typed refusal on a wire, not an absence of wiring.

15.4 The first compliance act

The experiment’s first finding was live: `s4=false` meant the substrate violated the information rule (the page complied by rendering the falsity in-band). The repair followed the institution’s own prescription (`futon3c/src/futon3c/logic/casca` commit `d0741b04`): `fetch-edges` is loud on transport failure, so an empty O5 at summary time already means the authoritative query succeeded and found zero hole-target edges—a checked none, not an unread dimension. The standard now flips on typed evidence, with the evidence riding in-band as an `:s4-evidence` record (checked dimension, holes-found count, basis, as-of).

Two honesty notes attach. First, the pattern’s status field tracks the repair’s progress: until the running JVM reloads, the live page still shows `s4=false`, which is now a stale rendering rather than an untyped absence—the two-wire discipline applied to its own counterexample. Second, the institution text itself remains a proposal at Workshop precision; it was minted by no one but the operator. The Workshop that authors boards and the Workshop that studies institutions turn out to be the same room: in both, the rules are authored before the run, and the participants—not an external regulator—monitor them.

16 When passing the contract is not enough

The organisation contract presented an unusually clear version of the paper’s title question. Three existing exemplars—ants, data mining, and snatch—each passed the same seven ruled clauses (`futon2:holes/labs/wm-contract/runs/F12-organise/19-mining-exemplar.edn:1-10`, and `futon2:holes/labs/wm-contract/runs/F12-organise/19-ants-exemplar.edn:1-10`). Passing did not make the examples equivalent. For ants and mining, the set `addedByOrganise` was empty; the subtraction in clause O3 therefore removed nothing. Snatch was the one example in which that set was nonempty, so it exercised the clause rather than satisfying its equation at an empty set (`futon2:holes/labs/wm-contract/runs/F12-organise/19-scratch-exemplar.edn:1-10`).

That sharper test exposed a mismatch. The organised snatch result retained one edge, while the recorded run contained ten; the proviso check consequently records snatch as failing to reproduce its own edge set (`futon2:holes/labs/wm-contract/runs/F12-organise/25-proviso.edn:110-132` and `futon2:holes/labs/wm-contract/runs/F12-organise/25-proviso.edn:133-152`). The ants case had no recorded edge set against which to test reproduction, while mining reproduced two edges; across the three examples the measured outcomes were therefore “not measurable”, true, and false (`futon2:holes/labs/wm-contract/runs/F12-organise/25-proviso.edn:62-65`). Choosing one of the listed ways to close the proof would have accepted those limitations as the target: the three arms respectively preserved one reproducing example, kept one of snatch’s ten edges, or stopped without a producer (`futon2:holes/labs/wm-contract/runs/F12-organise/25-proviso.edn:203-226`).

The ruling rejected that kind of satisficing. The three examples are instead inspiration for finding a naturalistic exemplar that works, with making snatch reproducible serving as the warm-up and bridge (`futon2:holes/labs/wm-contract/runs/F12-organise/25-proviso.edn:227-246`). The episode separates two questions that a green conformance result can otherwise blur: whether an implementation inhabits the formal contract, and whether it preserves enough of a real case to become a reusable tool.

Discussion

17 Related work and scope

An apparatus that can question its own result

Rheinberger's account of experimental systems helps frame the case [Rheinberger 1997]: apparatus makes questions answerable, but also limits what can appear as an answer. Here the limit is concrete. A component check could pass while its output had no live consumer. The response was to inspect the handoff and construct a rejecting check, rather than strengthen the prose around the passing result.

The pattern tradition records a recurring problem, its forces and a response [Alexander et al. 1977; Gabriel 2002; Gamma et al. 1994]. Our PLoP companion presents the mechanism catalogue; this paper follows the audit it prompted. The apparatus patterns add a related question: what ensures that the mechanism described is loaded, observed and independently assessed? Their APM field-trial record is useful precisely because it includes misleading applications and subsequent corrections. A promoted entry is not a universal effectiveness claim.

Winner's question about the politics of artifacts [Winner 1980] bears on the location of authority in this design. The system may propose work and interpret observations, but adoption of a preference or acceptance of an outcome cannot be silently inferred from the same machinery that produced the proposal.

Agent work and operator participation

Reflection and software-agent systems provide settings for execution and feedback [Shinn et al. 2023; Wang et al. 2025; Yang et al. 2024]. The present contribution is narrower than a claim to a new general agent architecture: it reports failures between declared mechanisms, recorded execution and continuing operation in one stack. Neither an implementation inventory nor a comparison of architectural diagrams establishes relative performance.

Mixed-initiative design makes the cost of interruption and the allocation of initiative explicit [Horvitz 1999]. Ostrom's action-arena perspective similarly directs attention to participants, available actions and rules of interaction [Ostrom 2005a]. These perspectives help explain why no-response cannot be treated as a transparent preference signal. The paper identifies that confound and partial observations relevant to it; it does not report a successful intervention on operator behaviour.

Coordination beyond the project

Coase's account of coordination within firms, Beer's organisational cybernetics and work on commons and peer production provide different settings for the reporting question [Beer 1972; Benkler 2006; Coase 1937; Ostrom 1990]. We use them to ask who supplies observations, who may demand a report and who can accept an outcome. They do not establish that a particular software instrument will be adopted or improve an organisation.

The empirical contribution remains the audit and its bounded demonstrations. The issue board makes unresolved obligations visible; the operator model, problem-level preferences and transfer study identify further tests. A useful continuation will be a witnessed run that answers those tests, not a more complete-looking description of the architecture.

18 Acknowledgements

Thanks to the Peeragogy Project participants for contributions on related topics.

References

- Christopher Alexander. 1964. *Notes on the Synthesis of Form*. Harvard University Press, Cambridge, MA.
- Christopher Alexander. 1965. A City Is Not a Tree. *Architectural Forum* 122, 1 (1965), 58–62. Part II in issue 2, pp. 58–62.
- Christopher Alexander, Sara Ishikawa, and Murray Silverstein. 1977. *A pattern language: towns, buildings, construction*. Oxford University Press.

- Stafford Beer. 1972. *Brain of the Firm: The Managerial Cybernetics of Organization*. Allen Lane, London.
- Yochai Benkler. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press, New Haven, CT.
- R. H. Coase. 1937. The Nature of the Firm. *Economica* 4, 16 (1937), 386–405.
- Richard P. Gabriel. 2002. *Writers' Workshops & The Work of Making Things: Patterns, Poetry...*. Addison-Wesley.
- E. Gamma, R. Helm, R. Johnson, and J. Vlissides. 1994. *Design patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley.
- Eric Horvitz. 1999. Principles of Mixed-Initiative User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. ACM, New York, NY, 159–166.
- Leslie Lamport. 1995. How to Write a Proof. *The American Mathematical Monthly* 102, 7 (1995), 600–608. DOI : <http://dx.doi.org/10.1080/00029890.1995.12004622>
- Thomas P. Moran. 1971. (Artificial, Intelligent) Architecture: Computers in Design. *Architectural Record* (March 1971), 129–134.
- Elinor Ostrom. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, Cambridge, UK.
- Elinor Ostrom. 2005a. *Understanding Institutional Diversity*. Princeton University Press, Princeton, NJ.
- Elinor Ostrom. 2005b. *Understanding Institutional Diversity*. Princeton University Press, Princeton, NJ.
- Hans-Jörg Rheinberger. 1997. *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford University Press, Stanford, CA.
- Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. 2023. Reflexion: language agents with verbal reinforcement learning. In *Advances in Neural Information Processing Systems 36 (NeurIPS 2023)*.
- Xingyao Wang, Boxuan Li, Yufan Song, Frank F. Xu, Xiangru Tang, Mingchen Zhuge, Jiayi Pan, Yueqi Song, Bowen Li, Jaskirat Singh, and others. 2025. OpenHands: An Open Platform for AI Software Developers as Generalist Agents. In *The Thirteenth International Conference on Learning Representations (ICLR 2025)*.
- Langdon Winner. 1980. Do Artifacts Have Politics? *Daedalus* 109, 1 (1980), 121–136.
- John Yang, Carlos E. Jimenez, Alexander Wettig, Kilian Lieret, Shunyu Yao, Karthik Narasimhan, and Ofir Press. 2024. SWE-agent: Agent-Computer Interfaces Enable Automated Software Engineering. In *Advances in Neural Information Processing Systems 37 (NeurIPS 2024)*.

A The War-Room Patterns in Full

The summary in Section 6 gives each ruling's conclusion. Here each is given as the stack recorded it: the condition that brings the tension into play, the tension itself, the actions that manage it, and the rationale — with the observed result of applying the pattern where one has been recorded.

Organise As Effectively As War, Without Becoming A State Apparatus (WR-0). An apparatus built to organise effectively must carry, from the outset, the three properties that stop it becoming an authority: an open model, an available exit, and a verdict it cannot issue to itself. Effective organisation and unaccountable apparatus are the same construction seen from two sides; only those properties separate them. **If** Martin Luther King Jr.: “Those who love peace must learn to organize as effectively as those who love war.” The organisational forms that win are not neutral instruments that happen to have been used for war; they are what effectiveness looks like — a shared picture maintained in real time, a chain of command, an instrument for testing tactics before they cost anything. Refusing them on the grounds of their history is a decision to be less effective, and it is usually taken by people who can afford to be. **However** Gilles Deleuze: “we want a war-machine that would not recreate a State apparatus, a nomadic unity in relation with the Outside, that would not recreate the despotic internal unity.” The same forms that make an organisation effective make it an apparatus: the shared picture becomes the only admissible picture, the chain of command becomes the only admissible authority, and the instrument for testing tactics becomes the instrument that certifies its own results. An apparatus that has captured its own verdict cannot be argued with from outside, and no longer needs to be.¹ **Then** Build the organising apparatus, and build into it the three properties that keep it from closing: (a) An OPEN MODEL. Active inference is chosen here for a

¹ *The case that forced it.* The ruling was written down only after the stack had already lost property (c) in four places at once, none of them decided. The War Machine's outer loop proposed the same kind of action 108 times, treated every proposal the operator did not follow up as evidence against it, and collapsed that action's learned credit — a number measuring how often the lab proposed, not what the operator thought of it. (In the loop's own terms: 108 *emissions* of one *action class*, scored by *intrinsic value*.) The capability ledger recorded a single invoice and a repeatable revenue capability with the same :satisfied status. Six strategic surfaces asserted continuous currency at a median age of thirty-six days. Each is a verdict the apparatus issued to itself, and each looked like traction from inside.

structural reason rather than a fashionable one: the generative model is declared rather than implicit, its objective is written down and inspectable, and expected free energy carries an epistemic term — the apparatus is required to value being surprised. A State apparatus values compliance; a model with an information-gain term is one that pays for evidence against itself. (b) An AVAILABLE EXIT. Open source, forkable, with the substrate legible enough to leave with. A nomadic unity is one that can be left; the test of whether an apparatus is despotic is not how it treats members but what it costs to stop being one. (c) A VERDICT IT CANNOT ISSUE TO ITSELF. Every claim the apparatus makes about its own success must be movable only by evidence it did not manufacture (aif/no-self-certification). This is the property most easily lost, because losing it feels like progress: internal follow-through is abundant, cheap, and fast, and it looks exactly like traction until something external is asked to agree. **Because** An organisation that organises effectively, cannot be audited from outside, cannot be left, and grades its own work has become the thing it organised against — and will not be able to tell, because the instrument that would tell it is the one that was captured. The three properties are not ethical decoration on an engineering artefact. They are the difference between a war machine and a State apparatus, and they have to be present at construction because each is far harder to retrofit than to install.

Inhabit Before Building (wR-4). Inhabit the surfaces that already exist before building new infrastructure; new infrastructure is justified only when it makes an existing surface more inhabitable. **If** The stack has extensive infrastructure: Agency, evidence stores, pattern libraries, Arxana browser, candidate invariant queues, mission documents, excursion logs, Bayesian models, context retrieval. **However** Most of this infrastructure is uninhabited. The Candidate Invariant queue has sixty-five entries with no processing pipeline, the evidence browser was disconnected, excursions are explorations without missions, and the primary coding surface generated no evidence until today.² **Then** The next phase is inhabitation, not construction. New infrastructure is justified only when it directly serves inhabitation of existing surfaces. The three workstreams, mathematics, consulting, and stack inhabitation, must be mutually reinforcing: stack work that does not serve mathematics or consulting is building alone and never delivering, which violates the rule that whatever is carried must be delivered. **Because** Surface earns inhabitation: uninhabited infrastructure is dead infrastructure. Inhabitation feeds evolution: without it there is nothing for the system to learn from. The same pattern appears at the personal level: much carried, little held in reserve, and effort stuck on gathering rather than delivering. External depositing and internal stack inhabitation require the same policy transition, from scholar to consultant to free-solo.

Typed Files Are Sources Of Truth, Prose Is Regenerated (wR-8). Make the typed argument files the canonical self-model and generate prose from them, so drift becomes detectable instead of perennial. **If** The holistic argument was historically maintained as prose, in two separate sketches. **However** Prose drifted from intent because there was no machine-readable intent it had to track: every re-render was an editorial interpretation.³ **Then** The sixteen leaf-level and one stack-level typed argument files at the futon5a stories are the canonical self-model. Prose deliverables are generated by renderers and re-rendered when the drift detector reports a source change. Editing prose in place is a defect. **Because** The typed argument format carries structure, claims, supports, attacks, falsifiability, and conflicts, that prose flattens. If the typed structure is canonical, drift between argument and document becomes detectable instead of perennial.

²*The case that forced it.* The candidate-invariant queue held sixty-five entries with no processing pipeline, the evidence browser was disconnected, and the primary coding surface generated no evidence at all. The reflex was to build a queue system. What actually worked was making one surface inhabitable: within minutes of the REPL becoming usable, evidence started flowing — chat turns with session ids, context-retrieval certificates, the conversation browsable — and none of it needed new infrastructure. The evidence loop had been ninety percent built; the missing ten percent was wiring, and the missing ingredient was a person in it.

³*The case that forced it.* The holistic argument lived as prose in three documents that each drifted from intent, and the drift was undetectable because there was no machine-readable intent to compare against — every re-render was an editorial interpretation, and a faithful one was indistinguishable from a lossy one. Inverting the direction fixed the detectability rather than the prose: the sixteen leaf files plus the stack file became canonical, the prose became a renderer output, and editing prose in place became a defect rather than a judgement call.

Bites Must Be Empirically Tested Or Marked Logical (WR-9). Render self-model conflict claims as empirical only when session evidence supports them, and as logical otherwise, so the inhabitation gap stares back from the demo. **If** Conflict bites in the typed self-model assert that a given conflict damages a given part of the stack. **However** Such claims are formal until session evidence supports them: the cached weight on a conflict is a hypothesis, not a measurement.⁴ **Then** The empirical join walks session steps to repos to devmap leaves to spine ids, and renders conflict-to-spine bite edges as solid red, scaled by hit count, when the bitten spine has empirical activity in the visible window, and as dashed grey when the bite is still purely logical. Conflicts whose bites remain grey across multiple windows are candidates for weight reduction or removal. **Because** Invariant I1, evidence supersedes assertion, applies inside the self-model itself. A conflict whose damage cannot be observed is not yet a conflict, only a worry. The visual distinction makes the inhabitation gap measurable without the operator querying anything.

Operationalised Exploit-Loops Are First-Class Observation Channels (WR-16). Wire operationalised exploit-loops into the stack's status dashboard and the War Machine as direct typed evidence streams and first-class active-inference observation channels. **If** An exploit-loop emits typed structured state on a cadence independent of Joe's attention. **However** Treating that state as adjacent reading material leaves existing signals without a surface: a fourteen-day collection lag is invisible as habit failure, and a success-rate drop from ninety-seven percent to eighty-seven percent produces no bulletin-grade alert.⁵ **Then** Surface habit-completion telemetry, pipeline-health telemetry, and proxy metrics directly in the stack's status dashboard and the War Machine, and attach the interim-director checkpoint to the stream immediately. **Because** These loops already accumulate observations without fresh primary research; wiring them turns existing cadence and outcome changes into evidence the stack can condition on.⁶

Tension Must GENERATE, Not Only Rank (WR-19). Wire structural tension as a candidate source that generates actions, not merely as a weight that ranks candidates generated elsewhere. **If** The geometric stack detects high-tension regions that should drive the War Machine's next action. **However** The current wiring, ΔT as a sorry-only tie-breaker, cannot generate a move; it only ranks an existing menu, so a system full of unresolved tension still yields nothing when nothing is proposing candidates.⁷ **Then** Add a tension-proposer: a component that emits new candidate actions where tension is highest, so the field can put options on the menu rather than only rank the ones already there. Make this the proximate fix to starvation and the first build slice. **Because** A ranking signal cannot create an option that is absent from the menu. Tension becomes actionable only when it can propose the candidate whose priority it would otherwise merely increase.⁸

⁴The case that forced it. One conflict in the cached model claimed to damage three spines. Over a fourteen-day window only one of the three showed any empirical activity — one-third coverage. The other two bites were not wrong so much as unwitnessed: weights that had been carried as measurements since they were first written down. Rendering them as dashed grey rather than solid red put the inhabitation gap on the demo where it could be seen without anyone running a query.

⁵The case that forced it. Two loops were already emitting typed state on a cadence nobody had to maintain. Batches 7 and 8 sat uncollected for fourteen days — a binary, visible failure of the operator's own half of the loop — and mark2's success rate fell from ninety-seven percent to eighty-seven percent. Neither produced an alert, because the state was being read as adjacent material rather than as an observation channel. The signals existed; the surface for reading them did not.

⁶An unbuilt *how* for which this ruling is the *why*. operator-turns-become-inference-observations: Process captured operator turns into structures that inference can consume, using the turn-to-pattern association as grounded enrichment. **Next steps:** since we have full text search via an XTDB sidecar, this can now be piloted

⁷The case that forced it. On its first outing the War Machine exhausted its proposer-fed queue into learn-action-class before any outward open-mission surfaced. Tracing the judge showed why: structural pressure was applied only after four fixed proposers had supplied candidates, and only to address-sorry candidates. The system was full of unresolved tension and the machine had nothing to do, because a tie-breaker cannot put an option on a menu it only reorders.

⁸An unbuilt *how* for which this ruling is the *why*. tension-proposes-candidates: Add a tension proposer that creates candidate actions from tension instead of only reordering candidates supplied elsewhere.

A Removed Constraint Does Not Remove the Discipline It Supplied (WR-24). When retiring a proxy constraint, install its replacement discipline in the same act and measure how much external verdict content changes subsequent work, not merely whether a verdict arrived. **If** A gate was installed as a proxy for another requirement, such as using a cash gate as the cheapest available proxy for external validity, and the underlying scarcity is lifted. **However** Silently removing the pressure also removes its arbitration discipline. Pleasant, busy, self-contained work can then look like progress from inside even when outward feedback does not affect what gets built.⁹ **Then** Name the replacement in the same act as retirement. Replace the cash gate's master role with witnessed external verdict, using compiler, maintainer, peer-review, and money instruments; measure the gate as the graded gain from verdict content into the next decision, not as a weekly yes-or-no event. **Because** External evaluation cannot be self-supplied, but a boolean gate on an impoverished predicate buys nothing over simply continuing the work. The governing quantity is how much returned content changes the plan itself rather than only the output.

Good News Gets the Same Evidence Discipline as Bad (WR-25). Let unanchored good news change ordering immediately but leave numerical posteriors unchanged until an artifact supports it. **If** A model permits posterior movement only on inbound clicks carrying evidence such as a reply or ledgered payment. **However** A rule introduced to stop optimism inflating buyer probability can be applied asymmetrically, allowing inheritance, verbal commitments, or promising lanes to rewrite downstream financial figures without challenge.¹⁰ **Then** Apply the evidence rule symmetrically. Let operator-reported facts affect ordering, including lane preemption and which tension becomes decidable, but do not let them affect runway, probability, or marginal figures until an artifact anchors them. State this asymmetry wherever it appears. **Because** A model that filters bad news while admitting good news is biased rather than conservative, especially where its financial figures will later be checked by clients, employers, or institutions.¹¹ **Result** Held under contact and gained a clause. Two operator-supplied constraint releases went through as ordering-effective and number-inert, as specified; the mirror case then forced an amendment the original ruling did not contain (see evidence). The pattern was strengthened by application rather than merely confirmed by it.

A Capability Switched Off Carries Its Re-Arm Condition, in Writing (WR-26). Write a countable re-arm condition at the switch whenever an autonomous capability is turned off. **If** An autonomous capability is deliberately turned off for a sound reason, such as disabling a scheduled loop while hand-dispatched runs confirm its behavior. **However** The reason outlives its own memory: soundly disabled capabilities become stranded flags, are never re-armed, and return months later as surprises.¹² **Then** Put the re-arm condition at the switch itself rather than only in a mission document. Make it countable, such as a specified number of runs whose behavioral checklist passes without operator amendment; the note must state what ends the pause. **Because** "Until the behavior is confirmed" is a mood rather than a condition, and a stranded flag costs more than the throughput saved by the pause.

⁹ *The case that forced it.* The cash gate was retired as *master* gate and demoted to one instrument among several. What it had been supplying was not cash discipline but the requirement that someone outside the work judge it — a discipline that would have gone silently with the gate had the replacement, a witnessed external verdict, not been named in the same act (War Bulletin 14).

¹⁰ *The case that forced it.* The rule was minted to stop optimism inflating the estimated probability of a buyer. The case that exposed the asymmetry arrived from the opposite side: a *confirmed* good fact, fully invoiced, carried an obligation the numbers did not show. Filtering bad news while admitting good news had left the model open on the side it was not watching.

¹¹ An unbuilt *how* for which this ruling is the *why*. satisfied-rungs-are-counted-and-surfaced: Count and surface a satisfied rung with the same evidence discipline used for an unsatisfied rung.

¹² *The case that forced it.* Three of the flags stranded in War Bulletin 12 were this shape — switched off with no written condition for switching them back, so the off state persisted by default rather than by decision. The codex cron was on course to become the fourth until its re-arm condition was written into the restore note the same day.

A Loop Is Born Instrumented for Its Gain (WR-27). Ship every new consumption-to-acquisition loop with an instrument that measures how much returned demand changes what the loop does next. If The gain from outward behaviour back into the underlying plan has recurred as the governing parameter in CA coupling, the strategy layer's external-verdict gate, and the epistemic economy's demand-coupling loop. **However** Every instance was discovered retroactively after construction, and outward activity without coupling is invisible from inside because its activity level looks correct.¹³ **Then** Instrument every new mining, retrieval, outreach, or review loop from day one with the fraction of returned signals that change subsequent acquisition or action. For demand-paged mining, record which hunger signals shaped each batch and report hunger-shaped acquisitions divided by total acquisitions. **Because** Measured-negative evidence shows that raw coupling is not automatically beneficial. Gain must be observable before it can be shaped; otherwise hermit mode can only be diagnosed retroactively.¹⁴

B Argument and Evidence Map

This map adapts numbered argument steps from Lamport's presentation [Lamport 1995]. It is an evidence map, not a mathematical proof. Dated observations retain their scope; definitions and proposals do not establish integrated operation. Nothing here supplies new evidence.

Each step carries a tag naming the *kind* of support it has:

- [def] construction — a thing defined or built
- [audit] census over an existing corpus (counting what is there)
- [meas] dated measurement or bounded experiment
- [neg] audited negative — looked for, not found
- [claim] argued analytically, NOT measured
- [prereg] stated in advance, NOT yet run

Of 17 tagged steps, 12 record constructions or observations; 5 are arguments or proposals. The tags distinguish kinds of support, not degrees of system completion.

- ⟨1.1⟩ *Assertion (Part I)*. The audit connects declared constraints to implementation evidence; component conformance does not establish the complete loop.
 - ⟨1.1.1⟩ [def] Missions, contracts and pattern records state different obligations. Their referents and consumers must be identified before their results compose.
 - ⟨1.1.2⟩ [audit] The August 31 vetting found two unsupported closures in a twenty-four-row paper ledger. Resolution included withdrawals, not only implementation repairs.
 - ⟨1.1.3⟩ [meas] The August 30 recorded route agreed with the described wiring on three of nine hops. This is a dated finding, not the live readiness verdict.
 - ⟨1.1.4⟩ [claim] A constraint-to-mechanism map locates questions to investigate; its unmapped mechanisms need individual explanations, not a presumed common cause.
 - ⟨1.1.5⟩ *Discharge*. The issue board reports current obligations. Neither repaired prose nor a passing component check closes a run-gated claim.

¹³ *The case that forced it*. The demand-paged mining loop was specified without any measure of what fraction of hunger signals changed the next mining batch, so nothing in it could report whether the demand paging was earning its keep. The same shape recurs wherever a signal is logged but never routed: a judge/formalizer disagreement log that never alters the formalization pipeline is the uninstrumented-loop smell (War Bulletin 15).

¹⁴ *unbuilt hows* for which this ruling is the *why*. *commitment-temperature-is-instrumented-as-gain*: Instrument commitment temperature as a gain when the actuator loop is born, so its effect on subsequent action can be measured directly. *per-tick-mismatch-instruments-outer-loop-gain*: Use per-tick mismatch as the gain reading for the mission-conditional-reward outer loop and surface whether returned reward changes its next action.

⟨1.2⟩ *Assertion (Part II)*.. Operator traces support partial attribution, but cannot identify acceptance or attention without further evidence.

- ⟨1.2.1⟩ [meas] The audited outer loop lowered an action class credit after 108 unanswered proposals. The record did not establish consideration or rejection.
- ⟨1.2.2⟩ [def] Clock-in joins a turn to its declared mission; retrieval joins it to candidate patterns; mission prose records pattern mentions.
- ⟨1.2.3⟩ [meas] The generated turn-chain section measures their intersection. Clock context and textual mention answer different questions, so disagreement is not itself an error.
- ⟨1.2.4⟩ [claim] A response model must distinguish confirmed acceptance, confirmed decline and no-response. Finding traces does not validate that model.
- ⟨1.2.5⟩ *Discharge*. The proposed repair requires an intervention and its own comparison. Traceability is the starting evidence, not the outcome.

⟨1.3⟩ *Assertion (Part III, proposal)*. Transfer to other organisations remains an unrun study, with unresolved corpus descriptions.

- ⟨1.3.1⟩ [def] The proposed corpus schema distinguishes requirement, benchmark, bounty, dependency, call and environment.
- ⟨1.3.2⟩ [audit] The earlier prose described both thirty and thirty-two cases, and both uniform and opposed terminal classifications. These descriptions need case-level reconciliation.
- ⟨1.3.3⟩ [claim] Historical reconstruction does not establish outcome-blind encoding. Contemporaneous self-cases also require controls for author involvement.
- ⟨1.3.4⟩ [claim] Recurring decisions, consultation records and capacity to act are proposed eligibility criteria, not demonstrated predictors of adoption.
- ⟨1.3.5⟩ *Discharge*. The study may expose encoding residue. Neither predictive performance nor a market boundary is established here.

⟨1.4⟩ *Assertion (Part IV, proposal)*. Attribution and richer preferences require explicit observation domains, participation and witnessed consumers.

- ⟨1.4.1⟩ [claim] Current-state records need observations; reusable rules retain a context and can also be invalidated. Neither is maintenance-free by definition.
- ⟨1.4.2⟩ [def] The attribution design preserves source strength, unclaimed work and later confirmations instead of turning an inferred association into a witness.
- ⟨1.4.3⟩ [meas] The August 24 provenance probe recorded observations but no delivered followups. Its historical limit must not be presented as current service status.
- ⟨1.4.4⟩ [meas] The September 8 disposition fit gives constant risk on three grounded attempts. The September 9 source check identifies a separate mismatch between checkpoint and channel observation domains.
- ⟨1.4.5⟩ [meas] The Snatch experiment changes realised payoff when pattern order changes. This does not validate the War Machine expected-free-energy implementation.
- ⟨1.4.6⟩ *Discharge*. A new run must establish the missing composition and its outcome. Proposed attribution, preference and participation mechanisms remain distinct from that evidence.

C A Pattern Language for an Ostrom Game

This appendix retains the eighteen-pattern snapshot of the reported game experiment. It is distinct from the later twenty-four-member F11 find experiment and does not claim to validate the War Machine's integrated run.

This appendix reports a small closed experiment carried out on a system we did not build. *Snatch or Share* is a multiplayer teaching game inspired by Elinor Ostrom’s work on institutions for collective action, distributed with five treatments that differ only in the institutional arrangement surrounding an identical exchange.¹⁵ P1 holds ten tokens and P2 holds ten of another kind; a player’s own tokens score one and the other player’s score two, so acquiring is worth twice holding. Each round P1 abstains or proposes an offer naming what it gives and what it asks, and P2 accepts, refuses, or snatches — taking the offer and paying nothing. The treatments add, in turn, compulsory offers (G2), a symbolic mark (G3), an adjudicator that reverts and then inverts a theft (G4), and non-binding talk before the choice (G5).

The connection this rests on is Ostrom’s own. Asking whether institutional analysis has an underlying set of universal building blocks, she writes that “Christopher Alexander (1964) earlier conceptualized all components of social arrangements as having a pattern and being a unit. Units have subunits and are themselves parts of larger units that fit together as a pattern” [Ostrom 2005b, p. 11], citing *Notes on the Synthesis of Form* [Alexander 1964]. Bringing a pattern collection to a game built on her work is therefore not an analogy imported from outside the subject.

The game is small enough to be understood completely and external enough that its outcomes are not ours to arrange. That combination is what makes it a test rather than an illustration: we can state what a pattern collection would have to do to be worth having, and then find out whether ours does it.

C.1 From pattern text to production rule

Moran read the pattern form as a production rule in 1971 [Moran 1971], and observed that a cascade is a control problem: many patterns may apply at once, and one may enable or invalidate another. The compilation used here is his, with two disciplines added that the pattern form itself supplies.

The antecedent is IF and HOWEVER. IF names the situation; HOWEVER names the counter-force that the THEN exists to overcome. Where that force is not live, the THEN would happen anyway and the pattern has nothing to contribute, so both conjuncts must hold before a pattern fires. The difference is not cosmetic. EXCHANGE WHEN BOTH SIDES GAIN fires on its IF in every round in which a player holds tokens; its HOWEVER — “attention to seizure risk can make abstention look like the objective” — restricts it to rounds where seizure is salient, so against a counterpart who has accepted four times in a row it correctly says nothing.

A guard may not mention the action taken this round. A pattern advises a decision; conditioning its antecedent on that decision inverts the exercise and reports coverage that was never earned. Three guards in an earlier draft of this collection were written against whether P1 had already chosen to offer, and produced four phantom coverage gaps in G1 which disappeared when the policy, rather than the pattern, was made to decide.

A THEN the model cannot carry emits no action. Four of the twelve play-grain patterns speak about the size of the ask, the talk channel, or a mark that changes no payoff — real features of the game that this harness does not represent. They remain advisory rather than being given an invented action, so that “not modelled” is recorded as itself.

The same test applies in both directions. Two of the seven acting patterns have a HOWEVER that is live in every state where their IF holds: “one acceptance is weak evidence, and a snatcher who accepts once has lost nothing by waiting” is a permanent condition of the game rather than a tension in the situation. A force that is never absent discriminates nothing, and is recorded as non-discriminating rather than given a guard that cannot fail.

¹⁵<https://github.com/AllianceBioversityCIAT/snatchgame>. The collection was written from the game’s five mermaid design flowcharts, not from its client or server source.

C.2 Two policies over one collection

The resulting eighteen patterns — six about choosing an institution, twelve about playing within one — are exercised by two policies over identical scenarios. The first is *grim trigger*, the standard repeated-game strategy: offer until the counterpart defects once, then abstain for the rest of the game. It is hardcoded, and the collection cannot reach it. The second takes the patterns whose antecedent holds, orders them by precedence, acts on the first whose THEN yields an action, and records which pattern acted. Payoffs follow the game’s own stipulated rule; a denunciation under G4 is worth the two steps its pattern describes, the seized tokens restored and the asked-for tokens transferred at the defector’s expense.

treatment	P2	grim	patterns	re-wired
G1	snatcher	−1	−5	−5
G1	sharer	+5	+15	+15
G1	cautious	0	0	0
G4	snatcher	−1	+3	−5
G2	snatcher	−10	−10	−10
G5	sharer	+5	+15	+15

Box 4 – P1’s score under three policies. “Re-wired” is the same twelve patterns consulted in a different precedence order.

G4 is the case the collection was written for. Grim trigger takes one loss and leaves; the pattern-driven policy denounces, has the transfer inverted, re-enters on the observed repair, and finishes +3 where grim finishes −1. Two patterns account for the entire difference, and a hardcoded policy could not consult either.

C.3 The wiring carries the score

`futon2.aif.cascade-prior` states that two cascades containing the same patterns but connecting them differently are different policies. Because precedence is data here, that claim can be run rather than asserted. The right-hand column of Table 4 promotes one pattern — EXCHANGE WHEN BOTH SIDES GAIN — above the remedy and the stop. Membership is unchanged, every pattern’s text is unchanged, and G4 moves from +3 to −5. What a collection scores is not a property of which patterns it contains.

This exposes something the pattern form does not carry. Beyond IF, HOWEVER and THEN there is the conflict-resolution order, and no pattern states it. The classical production-system answer is specificity — the rule with the most conjuncts wins — and it fails here: AN UNMODELLED RESPONSE STOPS THE LINE has the fewest conjuncts of any acting pattern and has to be consulted first, because halting must dominate acting. That is a commitment about the collection rather than a property of any member of it, and there is nowhere in a pattern to record it. Its shape is the semilattice wiring that already distinguishes cascade identity in the implementation, which suggests the place for it is a collection-level record.

C.4 The cascade is a semilattice, not a chain

One pattern acts per round, so the sequence of acting patterns is linear. That sequence is not the cascade. The cascade is the sub-graph of @why authority edges that the acting patterns stand on, read from the pattern files at run time and closed upwards. For G4 against a snatcher it has seven nodes and ten edges (Figure 10), and four of those nodes stand on two authorities each.

That is Alexander’s own shape, and the point of *A City Is Not a Tree* [Alexander 1965]: the overlap is the structure, and a hierarchy that forbids it loses what the overlap was doing. It is worth being precise about where

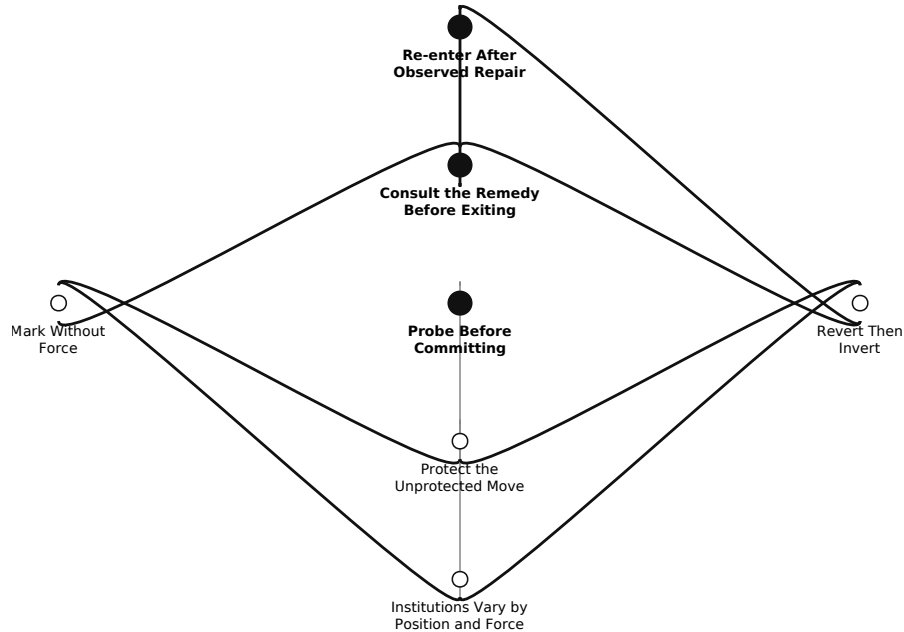


Fig. 10 – The G4 cascade. Heavy boxes are the three patterns that acted; the rest are what they stand on. A heavy edge leaves a pattern with two authorities. REVERT THEN INVERT is reached from RE-ENTER AFTER OBSERVED REPAIR by two distinct paths – directly, and through CONSULT THE REMEDY BEFORE EXITING – so the structure carries overlap and is not a tree.

the shape comes from. The authority edges were authored when the patterns were written; the harness reads them and adds nothing. What the run contributes is which part of the graph is reached.

Shape is therefore a property of the run, not of the collection. The same twelve patterns give a tree against a sharer – four nodes, three edges, no overlap, because the remedy structure is never engaged – and an overlapping structure under a remedy. A collection can support overlap and still produce trees indefinitely, if its situations never reach the shared authorities.

It is worth separating two properties that Alexander’s contrast runs together, because our own runs separate them. *Overlap* is the absence of the tree condition: some pattern stands on two authorities. *Being a meet semilattice* is the condition that any two patterns in the cascade have a greatest common authority. A rooted tree satisfies the second without the first, so “a semilattice rather than a tree” is a rhetorical contrast rather than a mathematical one. Checking both over the six runs gives all four combinations, which settles that they are independent axes rather than one axis named twice.

The G4 cascade is the one that is both: it carries overlap *and* every pair of its patterns has a greatest common authority, which is the structure Alexander’s argument is about. The G1 run against a snatcher fails the semilattice condition for a reason worth recording – EXCHANGE WHEN BOTH SIDES GAIN declares no authority at all, because it was written to state the game’s gain and had nothing above it to stand on, so a cascade containing it is

treatment	P2	nodes	edges	overlap	tree	meet semilattice
G1	snatcher	4	2	0	yes	no
G1	sharer	4	3	0	yes	yes
G1	cautious	4	3	0	yes	yes
G4	snatcher	7	10	4	no	yes
G2	snatcher	7	5	1	no	no
G5	sharer	4	3	0	yes	yes

Box 5 – Cascade shape by treatment and counterpart disposition, over one collection and one policy.

disconnected and two of its patterns have no common authority to meet at. An orphan in the authority graph is visible here as a structural defect and is invisible in the pattern text.

Two further things are worth saying about what this measurement does and does not support. The counterpart's disposition, holding the collection and the policy fixed, selects which sub-structure of the collection a run inhabits: the same twelve patterns produce four distinct shapes across six situations. And the one run that earns policy grade is also the one with the most overlap — but G2 against a snatcher carries overlap and is refused, so overlap is not sufficient, and with a single positive case nothing follows about necessity either.

C.5 What running it found that reading it did not

The collection was assessed by reading, which produced six gaps: nothing stated the gain that the whole exchange exists to realise, one treatment had no patterns at all, and so on. Those were repaired. A seventh gap appeared only on being run.

Against a snatcher under G1, the pattern-driven policy loses five where grim trigger loses one. After the first seizure nothing in the collection tells it to stop: CONSULT THE REMEDY BEFORE EXITING presupposes that a remedy exists, G1 has none, and EXCHANGE WHEN BOTH SIDES GAIN continues to fire correctly on a situation where an exchange is genuinely available. The collection has no pattern for leaving an arrangement that offers no recourse — an omission which is invisible to a reader precisely because every individual pattern is right.

This is what the score is for. It does not measure whether a pattern is well written; it separates a collection that plays well from one that reads well, and those come apart.

C.6 Limits

The payoff asymmetry is the game's stipulated teaching incentive and is adopted rather than derived. P2 is a fixed disposition rather than a chooser, so one pattern written for the responding player is excluded from the policy and its advice is untested; making P2 pattern-driven is the obvious next step. The size of the ask is a real feature of the game that this harness does not model, so the patterns about sizing an ask remain advisory. Precedence is authored by hand, and the six scenarios over five treatments and three dispositions form a space small enough to enumerate — a strength for auditing and a limit on generality.