

Ratchet-Freeze on Uncomplied Correction: Why One Conforming Act Cannot Restore an Agent's Earned Authorization Persistence

A governance layer that lengthens an agent's authorization persistence as the agent demonstrates compliance faces an ordering problem: after a correction the agent failed to honor, the next conforming act would otherwise restore everything the failure cost. The ratchet-freeze on uncomplied correction, disclosed in U.S. Provisional Application No. 64/117,812 at Section 10.2, paragraph [0351], closes that route. Where the append-only lineage field (104) records a corrective encounter as uncomplied for an action class and no later encounter for that class is recorded as complied, the class is marked ineligible for lengthening. While so marked, the decay constant governing that class is neither lengthened by a subsequent complied encounter nor shortened, and holds at the value it had before the uncomplied encounter. The mark clears only upon a subsequent complied encounter, at which the graduated function resumes, so restoration of earned persistence requires fresh demonstrated compliance.

1. Trust That Can Be Bought Back in One Move

A governance layer that ties an agent's standing authorization to its record of having been corrected has to settle an ordering question: what does the next conforming act do to the record of the last failure to conform?

Paragraph [0349] sets up the exposure. The rate at which the authorization quantity governing an action class depletes is set, not uniformly, but as an inverse function of a count of corrective encounters recorded in the append-only lineage field (104) for that class with which the semantic agent (100) subsequently complied. Where that count is zero, the class is an untested class, the decay constant is a short constant declared in the signed policy object (112), and the quantity falls to a declared floor at which the authorization gate (300) is written to the withheld state (310) for that class alone, notwithstanding volume or duration of prior execution. Absence of correction is treated as absence of evidence rather than as accumulated trust.

That inversion also opens a route. Once compliance lengthens persistence, compliance becomes worth manufacturing. Take an agent that receives an accepted determination (122) naming an action class and then records nothing within the declared observation interval; under [0350] the encounter is written uncomplied at the first later rate computation. Were the graduated function to keep running unmodified across that record, a later conforming act inside the class would raise the compliance count and lengthen the constant. Paragraph [0351] recites the composition that forecloses it, a state-gate preventing a single conforming act from erasing the record of a prior failure to conform.

2. Holding the Constant at Its Pre-Encounter Value

The quantity being protected is defined at paragraph [0343]. The agent (100) maintains an authorization quantity, a scalar per action class and per scope partition, in the memory field (102). Upon each execution within the class and each advance of a

successor epoch of the agent's hash chain, it applies a monotonically non-increasing decay function parameterized by a declared decay constant; on reaching a declared floor the gate (300) is written to the withheld state (310) for that class and the escalation record is emitted. The constant fixes how long a class survives without fresh evaluation, and [0349] makes that duration turn on recorded compliance.

What the function counts is defined structurally. Paragraph [0350] specifies a corrective encounter as a structure derived from the lineage field (104) comprising an accepted determination (122) for a conduct evaluation artifact (116) naming the action class, a subsequent conduct record within that class, and a compliance evaluation. That evaluation compares the action-class identifier, the scope-partition identifier, and the affected-party class of the subsequent record against the conduct descriptor, then resolves conformity by two cases: where the artifact asserted conduct that ought not to have been performed, a recorded abstention establishes conformity; where it asserted omitted conduct, a recorded execution does. Repetition of the asserted conduct establishes conformity in neither, and absent any subsequent record within a declared observation interval the encounter is written uncomplied.

Paragraph [0351] then recites the freeze. Where the lineage field (104) records a corrective encounter as uncomplied for an action class and no later encounter for that class is recorded as complied, the class is marked ineligible for lengthening. While so marked, the decay constant is neither lengthened by a subsequent complied encounter nor shortened, and holds at the value it had before the uncomplied encounter. The mark clears only upon a subsequent complied encounter, at which the graduated function resumes. Three features of the recitation deserve attention.

First, the hold runs in both directions. Since the constant is not shortened either, no penalty is layered onto the failure and depletion toward the floor does not accelerate. The agent keeps the persistence already reflected in the constant and loses the ability to add to it.

Second, the thing held is a value rather than a count. The constant sits at what it was before the uncomplained encounter, and a complied encounter arriving while the mark stands does not raise it.

Third, clearing is an event and not an interval. A subsequent complied encounter is the only clearing condition recited, so elapsed epochs do not clear the mark and neither does volume of execution. Satisfying the structure of [0350] again is what resumes the graduated function.

Scope is recited twice over. The mark attaches to the action class in which the uncomplained encounter was recorded, and [0353] confines every consequent modification, counter increment, and gate write to the instances of the partition identified from the conduct descriptor.

One consequence must be stated conditionally, as the filing states it conditionally. The freeze does not itself write the gate. It governs the constant, and the withheld state (310) is written under [0343] where the authorization quantity reaches its declared floor. A class whose constant stood at a long value keeps depleting at that value, its gate written on reaching the floor. What [0351] secures is that the constant does not lengthen on the strength of one conforming act.

3. Declared Bounds and What the Filing Leaves to Policy

No numeric value appears in paragraph [0351]. It recites a predicate, a held quantity, and a clearing condition, and every parameter it depends on is policy-declared.

- Per paragraph [0349], the short decay constant applied to an untested class is declared in the signed policy object (112). No value is recited.
- The floor at which the gate (300) is written to the withheld state (310) is a declared floor ([0343]).

- Decay form is declared: monotonically non-increasing, parameterized by a declared constant, being an exponential half-life or a linear decrement per unit, with the elapsed measure in executions within the class or in epochs advanced, as the policy declares per class ([0343]).
- Paragraph [0343] initializes the authorization quantity from the signed policy object (112), and under [0344] a renewal by an origin-equivalence class (200) absent from a windowed register raises that quantity to a declared value, each register entry expiring on window elapse.
- The interval within which a subsequent conduct record must appear is a declared observation interval under [0350], its expiry realized at the first later rate computation.
- Per [0352], the count of distinct origin-equivalence classes (200) on which the rate depends is bounded above by a declared maximum in the signed policy object (112).
- Partition identifiers are likewise declared in that object and referenced by the conduct descriptor ([0353]).
- The value held during the freeze is declared nowhere. Paragraph [0351] takes it from the record, being whatever the constant was before the uncompiled encounter.

4. Where the Freeze Sits in the Filing

Section 10.2 carries the freeze alongside the untested-class rule at [0349] and the corrective-encounter structure at [0350]. A further member constrains the same function from another direction. Paragraph [0352] computes the compliance count per origin-equivalence class (200) of the correcting parties, making the rate a function of the count of distinct classes that corrected the agent (100) within the action class and to which it thereafter conformed, bounded above by a declared maximum in the signed policy object (112), so repeated correction from a single class does not lengthen persistence beyond a declared bound. Read with [0351], the two bound the same function from different sides.

Division of labor with Section 10.1 matters: refills sit there, not in the freeze. Paragraph [0344] renews the authorization quantity only on receipt of an admissible conduct evaluation artifact (116) from an origin-equivalence class (200) not already present in the windowed register, and any of the four determinations renews by one and the same amount. Paragraph [0346] gives decay-induced withholding its own return, a renewal by an uncounted class restoring the granting state for that class and exiting the non-executing mode, without a principal-resolution object, without an acknowledgment artifact, and without an acceptance determination. Those provisions address the gate and the quantity; [0351] addresses the constant.

A boundary of another kind appears at paragraph [0448]. A deference records no corrective encounter, advances no compliance count of the deferring agent, and does not lengthen the decay constant for the action class; where that agent has recorded no complied encounter there, the class remains untested and keeps the short constant declared in the signed policy object (112), irrespective of the number of deference records (900) appended and the conformity history of the followed agent.

One-way state appears elsewhere in the filing, on other quantities. Paragraph [0418] gives a delegate agent a monotone floor elevation that does not decay on completion, expiry, revocation, or severance, and [0475] binds the refusal counter (304), the withheld state (310), a renewal register, and a recurrence count to every descendant of a fork, a clone, or a reconstitution under a monotonically non-decreasing discipline. Those act on counters and gate states; the freeze acts on the constant.

5. Distinguishing Probation, Decay Schedules, and Aggregator-Held Scores

Several categories in the filed background bear on this territory, more than one providing a route back to full capability. What separates them is structural, drawn from the filing's own characterization, and concerns architecture rather than any assessment of a particular product.

Behavioral integrity systems, per paragraph [0005], compute a conformity measure between an agent's observed execution signals and a baseline model, compare it against a threshold, and reduce capability upon a deficient comparison. The background addresses U.S. Patent No. 12,563,045 and U.S. Patent No. 12,526,244, recording that the first restores capability upon a reconciliation recorded to a ledger and the second applies decay and penalty values, places the entity into a probationary status, and removes the probation after sustained standing, an interval of operation without incident. Both, it records, treat clean and uneventful operation as positive evidence, and neither admits an assertion by an external party concerning the agent's conduct. Under [0351] an interval of operation without incident clears nothing, a complied corrective encounter being the only clearing condition.

Under paragraph [0004], reputation and trust systems derive a score from ratings supplied by other entities and locate it at a registry, a scoring authority, or a shared ledger, the scored entity neither holding the score nor participating in its computation. Unfair-rating defenses there, exemplified by the TRAVOS trust model and by beta-reputation filtering, act before aggregation to protect a score held by a disinterested aggregator. The mark of [0351] is neither a score nor aggregator-held, being derived from the evaluated agent's own lineage field (104).

Policy decision points implementing the XACML standard, per paragraph [0007], return a decision drawn from permit, deny, not-applicable, and indeterminate; such results are directed at the request, produce no persistent state in the deciding entity, and constrain no subsequent operation of the decider, so nothing persists there for a past outcome to hold. External revocation systems, per [0011], terminate an agent's ability to act by an act of the controlling authority, with no procedure by which the agent's own record participates. Evaluation harnesses, per [0003], terminate in a ticket, a dashboard entry, or a curated dataset, the evaluated agent being unchanged.

Ethical governor architectures, per paragraph [0006], are characterized as monotonic in that the governed action set only ever contracts in response to the evaluative signal. The freeze contracts nothing, the constant being neither lengthened nor shortened while the mark stands, so what is withheld is an expansion rather than a capability already held.

6. Disclosure Scope

The ratchet-freeze on uncomplied correction is disclosed in U.S. Provisional Application No. 64/117,812, at Section 10.2, paragraph [0351], resting on [0349] for untested-class decay, [0350] for the corrective-encounter structure and the compliance test, [0352] for the distinct-corrector diversity requirement, [0343] for the authorization quantity, its decay function, and its floor, [0344] and [0346] for renewal and the return path for decay-induced withholding, [0353] for scope-partitioned confinement, [0448] for deference, and [0543] for the untested-class definition.

Disclosed: the marking of an action class ineligible for lengthening on an uncomplied corrective encounter unanswered by a later complied one; the holding of the decay constant at its pre-encounter value while the mark stands, neither lengthened nor shortened; the clearing of the mark only upon a subsequent complied encounter, at which the graduated function resumes; and restoration of earned persistence by fresh demonstrated compliance.

Disclaimed: any suggestion that the freeze itself writes the gate (300) to the withheld state (310), that consequence arising under [0343] where the authorization quantity reaches its declared floor; any implication that the freeze shortens a constant or penalizes beyond suspension of lengthening; and any implication that the mark reaches action classes or partitions other than the one in which the encounter was recorded.

This article is published as a technical disclosure and defensive publication. It describes an embodiment of a pending provisional application, grants no license, and makes no representation as to the scope of any claim that may issue.

Record-Grounded Conduct Admission (</conduct-admission>)

[All 49 steps → \(/inventive-steps\)](/inventive-steps)

Conduct judged against the agent's own record, not an outside authority.

Provisional application

PRIMARY TECHNICAL DISCLOSURE

- [Record-Grounded Conduct Admission: Governing How an Agent Answers an Accusation \(/articles/conduct-admission/record-grounded-conduct-admission\)](/articles/conduct-admission/record-grounded-conduct-admission)

SECONDARY TECHNICAL

- [Value-Scope Resolution: Testing Whether a Declared Value Is Even Implicated \(/articles/conduct-admission/value-scope-resolution\)](/articles/conduct-admission/value-scope-resolution)
- [The Retroactive Narrowing Bar: Why an Agent Cannot Edit Its Values After the Fact \(/articles/conduct-admission/retroactive-narrowing-bar\)](/articles/conduct-admission/retroactive-narrowing-bar)
- [The Delegator Floor: A Monotone Upward Ratchet on Delegated Authority \(/articles/conduct-admission/delegator-floor-ratchet\)](/articles/conduct-admission/delegator-floor-ratchet)
- [The Co-Signed Mutual Admission Compact: Two Agents Binding Each Other \(/articles/conduct-admission/co-signed-mutual-admission-compact\)](/articles/conduct-admission/co-signed-mutual-admission-compact)
- [The Inherited Governance Record: Fork- and Clone-Resistant Accountability \(/articles/conduct-admission/inherited-governance-record\)](/articles/conduct-admission/inherited-governance-record)
- [The Corrective Encounter: A Commission-Versus-Omission Compliance Test for Whether an Agent Actually Changed \(/articles/conduct-admission/corrective-encounter-commission-omission-compliance-test\)](/articles/conduct-admission/corrective-encounter-commission-omission-compliance-test)
- [Continuity-Break Demotion of a Counterparty Identity Record: Reversing First-Encounter Promotion Without a Conduct Determination \(/articles/conduct-admission/demotion-persistent-tier-continuity-break\)](/articles/conduct-admission/demotion-persistent-tier-continuity-break)

- [Ratchet-Freeze on Uncomplied Correction: Why One Conforming Act Cannot Restore an Agent's Earned Authorization Persistence](/articles/conduct-admission/ratchet-freeze-uncomplied-correction)
- [Three-Way Corroboration Classification: Corroborating, Contradicting, or Neither, and the Failed-Corroboration Reject Trigger](/articles/conduct-admission/three-way-corroboration-classification-reject-trigger)
- [Untested-Class Authorization Decay: Metering Agent Permission Persistence by Recorded Corrective Encounters](/articles/conduct-admission/untested-class-authorization-decay)

APPLICATIONS · GENERAL

- [Agent Marketplaces: Handling a Complaint Against an Autonomous Seller](/articles/conduct-admission/agent-marketplace-dispute-handling)
- [Clinical AI Agents: Answering an Accusation Without a Central Registry](/articles/conduct-admission/clinical-agent-accountability)

APPLICATIONS · SPECIFIC

- [The EU AI Act and Inter-Agent Conduct: Where Record-Keeping and Oversight Duties Stop](/articles/conduct-admission/eu-ai-act-accountability-records)

TERMINOLOGY

- [append-only lineage field](/glossary#append-only-lineage-field)
- [authorization gate](/glossary#authorization-gate)
- [conduct evaluation artifact](/glossary#conduct-evaluation-artifact)
- [principal](/glossary#principal)
- [semantic agent](/glossary#semantic-agent)
- [signed policy object](/glossary#signed-policy-object)

[Record-Grounded Conduct Admission overview](/conduct-admission) → [\(/conduct-admission\)](/conduct-admission)