

Untested-Class Authorization Decay: Metering Agent Permission Persistence by Recorded Corrective Encounters

Untested-class authorization decay is a governance mechanism for autonomous semantic agents, disclosed in U.S. Provisional Application No. 64/117,812, that sets the depletion rate of an action class's standing authorization as an inverse function of the count of corrective encounters recorded for that class with which the agent subsequently complied. The filed text states the posture the mechanism produces: metering the persistence of standing authorization this way treats absence of correction as absence of evidence rather than as accumulated trust. Where the compliance count for a class is zero, the class is an untested class, its decay constant is set to a short constant declared in the signed policy object, and its authorization quantity falls to a declared floor at which the authorization gate is written to the withheld state for that class alone, notwithstanding volume or duration of prior execution.

Volume of Execution Is Not Evidence of Good Conduct

An agent that has executed inside one action class for a long stretch with no adverse entry against it has a record with a hole in it. Nowhere in it is an instance of the agent being told it had gone wrong and thereafter doing something different. The filing gives

that hole a reading: absence of correction is absence of evidence rather than accumulated trust.

The parent structure is already in the specification. A persistent semantic agent (100) maintains an authorization quantity, a scalar held per action class and per scope partition in the memory field (102) and initialized from the signed policy object (112). Upon each execution within the class and each advance of a successor epoch of the agent's hash chain, it computes an elapsed measure since the last renewal and applies a monotonically non-increasing decay function parameterized by a declared decay constant, the decayed value being written to the memory field (102) and appended to the lineage field (104). On reaching a declared floor, the authorization gate (300) is written to the withheld state (310) for that class and the escalation record is emitted.

One term of that structure is left open: where the decay constant comes from. Paragraph [0349] supplies it per class rather than uniformly. The rate at which the authorization quantity governing an action class depletes is set as an inverse function of a count of corrective encounters recorded in the append-only lineage field (104) for that class with which the agent subsequently complied. The filing marks the mechanism as distinct from any per-assertion cost multiplier: what it meters is the persistence of standing authorization.

How the Decay Constant Is Derived From Complied Corrective Encounters

A corrective encounter is derived, not a flag some party sets. Paragraph [0350] builds it out of the append-only lineage field (104): 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 comparing the action-class identifier, scope-partition identifier, and affected-party class of that record against the conduct descriptor.

Conformity then resolves 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 that ought to have been performed, a recorded execution does. Repetition of the asserted conduct establishes conformity in neither case. Absent any subsequent record within a declared observation interval, the encounter is written uncomplied at the first later rate computation.

Zero compliance count is the untested condition. Where that count is zero for an action class, the class is an untested class and its decay constant is set to a short constant declared in the signed policy object (112). Its authorization quantity falls to a declared floor and the authorization gate (300) is written to the withheld state (310) for that class alone, notwithstanding volume or duration of prior execution. In that state the agent does not execute actions of the class and enters a non-executing cognitive mode (302) in which speculative evaluation continues without committing state changes, and the escalation emitter emits a record to the principal naming the class, the scope partition, and the entries on which the transition was computed.

A single conforming act cannot erase a prior failure to conform. Paragraph [0351] composes a state gate onto the decay-constant function. 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, holding at the value it had before the uncomplied encounter. The mark clears only upon a later complied encounter, at which the graduated function resumes, so restoration of earned persistence requires fresh demonstrated compliance.

Who did the correcting matters as well. Under paragraph [0352] the compliance count that lengthens the decay constant is computed per origin-equivalence class (200) of the correcting parties, the rate being a function of the count of distinct classes that corrected the agent within the action class and to which it thereafter conformed,

bounded above by a declared maximum in the signed policy object (112). Repeated correction from a single class, however frequent, does not lengthen persistence beyond that bound. Substituting a distinct-class count for the raw count defeats manufacture of apparent trust through many corrections issued from a single controlled identity.

Withholding reached by this route carries its own return. Where the gate was written to the withheld state (310) by the authorization quantity reaching its floor, a renewal by an uncounted origin-equivalence class (200) returns it to the granting state for that class and exits the non-executing mode, without a principal-resolution object, without an acknowledgment artifact, and without an acceptance determination, the withholding having arisen from absence of evaluation rather than from an accepted determination (122) recording harm. Withholding written by a deviation quantity or by the provisional-restoration mechanism is returned only under the path applicable to it.

Paragraph [0448] closes a further route. A deference records no corrective encounter, advances no compliance count of the deferring agent, and does not lengthen the decay constant applicable to that agent's authorization quantity for the class. Where the deferring agent has recorded no complied corrective encounter there, the class remains untested and its constant remains the short declared one, however many deference records (900) it has appended or counter-signed and whatever conformity history the followed agent holds.

Declared Bounds, Intervals, and Constants

Each quantity below is declared by a principal, and the filed specification states no numeric value for any of them.

- **Short decay constant.** Declared in the signed policy object (112) and applied where the compliance count for the class is zero.
- **Declared floor.** The value of the authorization quantity at which the gate is written to the withheld state (310) and the escalation record is emitted.

- **Maximum bounding the rate.** The function of distinct correcting origin-equivalence classes (200) is bounded above by a maximum declared in the same policy object.
- **Observation interval.** Where it elapses with no subsequent conduct record, the encounter is written uncompiled at the first later rate computation.
- **Form of the decay function.** Monotonically non-increasing and parameterized by the declared decay constant, being an exponential half-life or a declared linear decrement per unit.
- **Elapsed measure.** Expressed in executions performed within the action class or in epochs advanced, as the policy declares per action class.
- **Granularity.** The authorization quantity is maintained per action class and per scope partition.
- **Renewal window.** Each entry of the renewal register expires on elapse of a declared window, after which that origin-equivalence class (200) is again uncounted.

Every consequence described here is conditional in the specification's own terms: a fall to the floor and a write of the gate follow from constants a principal has declared, none of them fixed numerically by the filing.

Neighbors of the Decay in the Filed Architecture

Upstream is the renewal path. The agent keeps, per action class and per scope partition, a register of origin-equivalence classes (200) from which an admissible conduct evaluation artifact (116) has been received within a declared window. Where the asserting party's (118) class is already in the register no renewal applies; where absent, the authorization quantity is renewed to a declared value and the renewal is appended to the lineage field (104) before and independently of the determination produced. Any of the four determinations renews by one and the same amount, so renewal is valence independent. The amount raised is a monotonically non-decreasing function of recorded severance events across the renewing class's constituent edges, a class whose

edges all resolve not-typeable renewing at a declared least amount. The same procedure replenishes the authorization budget (404) from a declared schedule bounded by a declared ceiling, and is the only procedure that increases that budget.

Alongside the decay sits scope partitioning. Each partition carries its own instance of the scoped integrity vector (106), self-esteem aggregate (108), refusal counter (304), authorization budget (404), and a segment of the append-only lineage field (104). Every consequent modification, counter increment, and writing of the gate (300) applies to that partition's instances alone, the gate remaining granting for classes of every other partition, so conduct in one context cannot withdraw authorization in an unrelated one.

Downstream is enforcement. The dispatch-authority predicate is recomputed at each request to dispatch an action, from state then carried in the memory field (102), and the authorization stamp of its three-stamp conjunction evidences that the gate occupies the granting or provisional granting state for the class requested. That predicate is not computed from a previously issued authorization token, a cached predicate result, or a session grant, so a write of the gate takes effect at the next dispatch request without revocation infrastructure.

Correction paired with conformity recurs elsewhere. The earned-exposure path admits an origin-equivalence class (200) only under a two-part conformance filter (906), correction first and a subsequent compliant conduct record second, and promotion of a counterparty identity record (114) to the promoted persistence value requires an encounter history containing no unresolved corrective encounter.

The filing carries "untested class" in two senses and defines both. Applied to an origin-equivalence class, it means one every constituent edge of which resolves not-typeable. Applied to an action class, the sense used throughout this article, it means a class for which no corrective encounter with subsequent recorded compliance is recorded.

Distinctions From Adjacent Categories of Art

Several established families look adjacent from a distance. The differences below are structural, and none asserts infringement by any product or party.

Time-bounded credentials. Short-lived certificates, expiring tokens, and time-to-live grants withdraw permission as a function of elapsed time or issuance age, with renewal obtained by re-presenting to an issuer. What differs is the input to the rate: the decay constant here comes from a recorded conduct history within one action class, and the return from a decay-induced withholding runs through renewal by an origin-equivalence class (200) not already counted within the declared window.

Quota, rate limiting, and spend metering. Budget systems meter the volume of actions and replenish on a window boundary. The filing holds them apart: the authorization budget (404) is per agent across the action space, the authorization quantity is per action class and per scope partition, and neither is computed from nor expressed in the units of the other. For an untested class the filed consequence follows notwithstanding volume or duration of prior execution.

Reputation and trust scoring. Scoring systems aggregate counterparty-supplied ratings, endorsements, or vouches into a scalar consumed at decision time. The compliance count comes instead from entries of the agent's own append-only lineage field (104): an accepted determination (122) recorded there, then a subsequent conduct record of the agent itself evaluated for conformity. An artifact (116) arrives from an asserting party (118), but the encounter closes on what the agent afterward recorded doing.

Continuous or risk-adaptive access evaluation. Session-revocation and step-up schemes react to anomaly signals, posture change, or risk events, so the trigger is arrival of adverse evidence. The condition here is a count of a recorded structure standing at zero, and what follows is the short declared constant.

Audit, attestation, and remediation tracking. Compliance systems record findings and their remediation with strong evidentiary properties. What a remediation record does not ordinarily do is parameterize the depletion rate of a live executing permission per action class, or freeze that rate against lengthening once a finding stands uncomplied.

Policy engines and capability systems. Rule-set evaluation of whether an action is permitted sits orthogonal to the question addressed here, which is how long a granted class-level authorization persists before the gate is written to the withheld state (310).

Disclosure Scope

The operative disclosure for this mechanism is U.S. Provisional Application No. 64/117,812, at Section 10.2, paragraph [0349], captioned "Untested-class authorization decay." Supporting paragraphs in the same section are [0350], defining the corrective encounter and the commission-versus-omission compliance test, [0351], the ratchet-freeze on uncomplied correction, [0352], the distinct-corrector diversity requirement, and [0353], scope-partitioned confinement of state movement. The parent decay, its renewal path, the severance-conditioned renewal amount, and the return path for decay-induced withholding are disclosed at Section 10.1, paragraphs [0343] through [0347]. Retention of the short decay constant across deference appears at paragraph [0448], and the dual definition of "untested class" at paragraph [0543].

This article is a technical description written for public, timestamped disclosure. It summarizes the filed text without enlarging it and is not a claim construction. Where it describes a consequence such as the authorization quantity reaching its floor or the gate being written to the withheld state (310), that consequence is conditioned on the constants, floors, bounds, and intervals a principal declares in the signed policy object (112), for which the filing states no numeric value and neither does this article.

Record-Grounded Conduct

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

Admission (/conduct-admission)

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 \(/articles/conduct-admission/record-grounded-conduct-admission\)](#)

SECONDARY TECHNICAL

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

APPLICATIONS · SPECIFIC

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

TERMINOLOGY

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

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