

The Retroactive Narrowing Bar: Why an Agent Cannot Edit Its Values After the Fact

A governance layer that evaluates an assertion about conduct already performed has to settle which policy version applies, and as of when. U.S. Provisional Application No. 64/117,812 settles it at Section 10.1, paragraph [0331]: the admission evaluator (120) resolves the assertion against the signed policy object (112) in force at the recorded assertion time of the conduct evaluation artifact (116). Candidates come from resolving the policy reference field (110) by canonical alias; a candidate whose validity window encloses that time is retained, a candidate revoked as of that time is discarded, and an anti-rollback floor held in the memory field (102), maintaining a monotonically non-decreasing successor index per alias, discards any candidate below it. A successor object admitted after the assertion time bears a later validity window and is not retained, whereby the agent (100) is incapable of narrowing its declared value set after receipt of an artifact so as to render that artifact inapplicable.

1. The Gap

Any evaluator that runs an assertion against "the policy" has to fix which policy, and as of when. Where the assertion concerns a proposed action, the two questions collapse into one. Paragraph [0007] of the filed background describes systems of that kind, in

which a policy decision point evaluates a proposed action before it is performed. An assertion about conduct already performed pulls the two apart. The conduct fell under one declared value set, the evaluation happens afterward, and the object carrying that set may have been superseded in the interval.

Failure then sits on either side of the missing anchor. Resolve against whatever object the agent holds at evaluation time, and a successor does the work: paragraph [0330] treats a value present in a predecessor whose declared-value identifier is carried by no value of the successor as excluded, and under paragraph [0329], where no declared value is implicated the admission evaluator (120) produces the not-applicable determination (128) at no cost. Anchor to the conduct time and nothing else, and the opposite move opens, since an agent holding a succession may present an older object whose validity window also encloses that time. Paragraph [0024] names what its fourth step forecloses: a policy object superseded before that time is not reinstated.

Paragraph [0331], in Section 10.1 of U.S. Provisional Application No. 64/117,812, addresses both directions and locates the resolution in the agent whose conduct is asserted.

2. Mechanism

The time governing the resolution is a field of the artifact rather than of the agent. Paragraph [0033] gives the conduct evaluation artifact (116) a recorded assertion time, being the time by reference to which the signed policy object (112) in force is identified under Section 1.3, and a signature of the asserting party (118) verifiable against an identity primitive in a counterparty identity record (114). Paragraph [0032] has the artifact originate from an asserting party other than the principal, and paragraph [0036] has it consumed by reference to its enumerated fields alone.

Paragraph [0331] states the resolution, and identification proceeds by the ordered procedure of paragraph [0024]. First, the semantic agent (100) resolves the policy reference field (110) by canonical alias to a set of candidate policy objects. Second, it retains a candidate whose validity window encloses the time by reference to which the resolution is performed. Third, it discards a candidate recorded as revoked as of that time. Fourth, it applies an anti-rollback monotonicity constraint maintained in the memory field (102), holding a monotonically non-decreasing floor over successor indices admitted under that alias, and discards a candidate below the floor.

Forward narrowing meets the second step. Paragraph [0331] recites that a successor object admitted after the assertion time bears a later validity window and is not retained, whereby the agent (100) is incapable of narrowing its declared value set after receipt of an artifact so as to render that artifact inapplicable. Nothing in that paragraph invalidates the successor: paragraph [0023] has the semantic agent (100) modify the signed policy object (112) only by admitting a successor, and such an admission still proceeds.

Backward narrowing meets the third and fourth steps together. Revocation as of the assertion time removes an object withdrawn by then, and the floor discards a candidate whose successor index sits below a value already admitted under the alias, so a predecessor leaves the candidate set before any value of it is tested. Paragraph [0329] places the value-scope test after that resolution: before testing whether asserted conduct occurred, the admission evaluator (120) runs a three-part membership test of the conduct descriptor against each value-scope tuple, matching action class, scope partition, and the affected-party class against an empathy-scope designation selected from a personal, an interpersonal, and a global scope.

Correspondence across a revision is carried by the declared-value identifier of paragraphs [0025] and [0330], assigned by the principal, covered by the object's signature, and stable across successor signed policy objects (112) admitted under one canonical alias. A declared value of a successor corresponds to one of the immediately

preceding object where and only where the two carry an identical identifier; that identifier is an identity of the value and not a scope element, and altering a tuple leaves it unchanged, whereby exclusion and contraction are detected by identifier correspondence rather than by comparing full value sets.

None of this is reconciliation performed afterward by a reviewer. Paragraph [0030] makes the policy stamp one conjunct of a three-stamp conjunction, required in full, upon which alone the dispatch-authority predicate is satisfied: that stamp evidences that the signed policy object (112) in force resolves by canonical alias, verifies against the signature of the principal, and satisfies its validity window, its revocation state, and the anti-rollback monotonicity constraint. Under paragraph [0031], satisfaction of fewer than three stamps produces a deterministic denial, a valid recorded outcome and not an error, appended to the append-only lineage field (104) naming the stamp that did not resolve and not converted into a determination concerning any party. Paragraph [0029] recomputes the predicate responsive to each dispatch request from state then carried in the memory field (102), and not from a cached predicate result or a session grant.

Recording follows under paragraph [0048]: for the not-applicable determination (128), the entry appended to the append-only lineage field (104) further comprises the value-scope tuples tested and the failing element of each.

3. Operating Parameters

Paragraph [0331] declares no interval, no count, and no numeric threshold. What it declares is an ordering, and each element of the ordering is operative.

- Candidates are obtained by resolving the policy reference field (110) by canonical alias.
- The window test is enclosure of the recorded assertion time by the candidate's validity window.

- The revocation test is revocation state as of that recorded assertion time, not as of the evaluation.
- The floor is held in the memory field (102), is monotonically non-decreasing over successor indices, and is maintained per canonical alias.
- The procedure of paragraph [0024] is ordered, its steps recited as first, second, third, and fourth.
- The anchor is a field of the conduct evaluation artifact (116), whose signature is verifiable against an identity primitive in the counterparty identity record (114).
- The policy stamp of paragraph [0030] requires the object in force to verify against the principal's signature.

4. Composition

What paragraph [0331] leaves open, paragraph [0332] takes up. Where the architecture detects from the append-only lineage field (104) an admission of a successor signed policy object (112) that excludes a declared value, or that drops an action-class, scope-partition, or empathy-scope element from a tuple corresponding by declared-value identifier, and that admission follows an as-yet-undetermined conduct evaluation artifact (116) implicating the value or element and falls within a window declared in the policy object in force at the artifact's assertion time rather than in the successor, the architecture scores a deflection event and increments the refusal counter (304), without adjudicating whether the contraction was well founded. A successor that adds a value or element is scored as no deflection and given no retroactive effect. Contraction under a pending assertion is priced; broadening is neither priced nor retroactive.

A constraint of comparable shape is recited at paragraph [0105]. A determination produced while the authorization gate (300) is in the withheld state (310) is appended to the append-only lineage field (104), decrements neither the rate accumulator nor the

run accumulator, and does not write the gate to the granting state, so the agent does not clear its own withheld state by producing further determinations. Paragraph [0106] adds that the withheld state (310) persists irrespective of elapsed time.

The floor recurs elsewhere in the filing. Under paragraph [0418], a delegation record to an independent delegate agent carries a delegator floor field enumerating, per delegated action class, the delegating agent's own anti-rollback floor value, the canonical alias it governs, and an attestation over both; the delegate computes an effective floor as the maximum of the carried value and its own, writes it as its own floor, and retains it without decay on completion, expiry, revocation, or severance. Paragraph [0470] maintains a floor of the same form over a canonical alias constituted from the identity primitives of two co-signing parties to a mutual admission compact, held separately from the floor over each party's own policy reference field (110). Paragraph [0434] forms a policy succession sequence from the signed policy objects (112) admitted under one canonical alias, each carrying a successor index held under a monotonically non-decreasing floor, and re-resolves a retrieved determination against each object of it in version order.

5. Prior-Art Distinction

The filed background enumerates the categories this mechanism is positioned against. The distinction in each case is structural: which party holds the version state, and what selects its evaluation time.

Policy decision points implementing the XACML standard, per paragraph [0007], return a decision drawn from permit, deny, not-applicable, and indeterminate. That four-valued result is directed at the request; the outcomes are returned to the requesting party, produce no persistent state in the deciding entity, and constrain no subsequent operation of the decider. The floor of paragraph [0331] is persistent state held in the deciding entity's own memory field (102), and paragraph [0030] makes it a conjunct of that entity's authority to dispatch at all.

Reputation and trust systems, per paragraph [0004], compute a score for an entity 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, by beta-reputation filtering, and by the immunization mechanisms surveyed by Dellarocas, filter incoming ratings before aggregation to protect the accuracy of the stored score, the filtering party being a disinterested aggregator rather than the rated entity. Under paragraph [0331] the filtering runs in the evaluated party, over its own declared value set.

Behavioral integrity systems, per paragraph [0005], compute a conformity measure between observed execution signals and a baseline model 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 neither admits an assertion by an external party concerning the agent's conduct nor provides a procedure by which the agent tests such an assertion against its own record. Paragraph [0331] presupposes such an admitted assertion and the time it carries.

Accountability protocols for distributed systems, exemplified by PeerReview, and remote attestation architectures, exemplified by the architecture of RFC 9334, locate the verdict and the appraisal of evidence outside the evaluated node, per paragraph [0010]; the accused node performs no governed determination of its own over the accusation. Paragraph [0331] places the resolution, and paragraph [0030] the consequence of its failure, inside the agent being evaluated.

Capability attenuation schemes exemplified by macaroons, per paragraph [0009], confine a delegate by appending caveats along the forward path of a delegation, each holder able to attenuate but not expand it; such schemes protect the owner of the delegated resource and operate in the forward direction only. The bar of paragraph [0331] operates on an artifact already received, about conduct already performed.

Merits-conditioned accountability mechanisms, per paragraph [0008], including Article 23(2) of Regulation (EU) 2022/2065, Section 623(a)(8)(F) of the Fair Credit Reporting Act, model unfair-claims-practices provisions, and bonded challenge procedures, attach a consequence only after a merits determination performed by an institution other than the penalized party. Paragraph [0331] attaches no consequence and reaches no merits. It selects the signed policy object (112) the evaluation runs against.

6. Disclosure Scope

The retroactive narrowing bar is disclosed in U.S. Provisional Application No. 64/117,812, at Section 10.1, paragraph [0331], resting on [0024] for the ordered resolution procedure, [0025] and [0330] for the declared-value identifier, [0329] for the value-scope test, [0033] for the artifact fields, [0030] and [0031] for the three-stamp conjunction and the deterministic denial, and [0048] for the recording of the determination.

Disclosed: resolution of an assertion against the signed policy object in force at the recorded assertion time of the conduct evaluation artifact; identification of that object by canonical-alias resolution, validity window enclosure, revocation state as of that time, and an anti-rollback floor holding a monotonically non-decreasing successor index per canonical alias; and the resulting incapacity of the agent to narrow its declared value set after receipt of an artifact so as to render that artifact inapplicable.

Disclaimed: any suggestion that the bar prevents an agent from revising its declared value set going forward, that direction being addressed at paragraph [0332]; any implication that a policy revision is thereby invalidated; and any reading in which the bar produces a determination concerning a party, paragraph [0031] providing that a denial is not so converted. Paragraph [0024] further recites that in other embodiments policy versioning permits the substrate to initiate rollback to a last verified state, as described in commonly owned Application No. 19/230,933.

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)

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).