

Three-Way Corroboration Classification: Corroborating, Contradicting, or Neither, and the Failed-Corroboration Reject Trigger

Three-way corroboration classification is an admission mechanism disclosed in U.S. Provisional Application No. 64/117,812, in which an admission evaluator sorts each independently received governed observation against the payload of an input into one of three classes: corroborating, affirmatively contradicting, or neither. The test runs over a set of corroboration fields declared for the transaction class, with agreement being equality for a discrete field and a difference within a declared tolerance for a continuous field. An observation corroborates only where it was emitted by a party other than the input's emitter and other than the evaluating party, and only where it agrees on every declared field. An observation affirmatively contradicts only where it agrees on no declared field and disagrees on at least one. Anything between the two is counted in neither class, and the absence of any received observation is not a contradiction and is not converted into one. Where an independently received governed observation affirmatively contradicts the payload, the admission evaluator issues a reject outcome bearing a failed-corroboration classification, one of the named rejection-reason classes of the graduated outcome set.

The Unwitnessed Input and the Disputed Input

A corroboration score assembled as a count of agreeing observations assigns the same low value to an input nobody else reported on and to an input another party disputed. Both fall short of the same target, and the count alone cannot tell them apart.

The filed specification separates them in the closed determination set of Section 1.8, one layer below the settlement machinery. Paragraph [0044] recites that the admission evaluator (120) produces the rejected determination (124) only where a retrieved entry affirmatively contradicts the conduct evaluation artifact (116) on a recorded field of that entry, the contradiction being affirmative where that field holds a value inconsistent with the conduct asserted, and not affirmative where the entry is silent as to it.

Paragraph [0047] closes the alternative: absence of evidence within the append-only lineage field (104) resolves to the not-determinable determination (126) and to no other class, whereby a semantic agent (100) is incapable of refusing an artifact by reason of its own record being silent, incomplete, or unavailable.

Those paragraphs govern what an agent's own record may be made to say. The mechanism here is the corresponding rule for observations arriving from other parties.

Three Classes, One Field Test, and a Named Reject Outcome

Paragraph [0498] states the classification. The admission evaluator (120) classifies each independently received governed observation against the payload of an input, under a set of corroboration fields declared for the transaction class, into one of three classes, and reserves any further corroboration outcome declared in the governing policy object.

Independence is written into the corroborating class, not applied afterward as an adjustment. A received observation corroborates the input where, emitted by a party other than the input's emitter and other than the evaluating party, its payload agrees with the input on each declared corroboration field. Paragraph [0226] states the same

requirement as emission by a third party, and paragraph [0198] closes an adjacent path: where an escalation credential in force upon an observation was self-issued, the receiving party applies a declared self-escalation discount factor to the elevated level and excludes the observation from the corroboration factor and convergence count of Section 6.15, whereby a party is incapable of corroborating or escalating its own input.

Agreement is tested per field, and the declared type of the field selects the test: equality for a discrete field, a difference within a declared tolerance for a continuous one. Corroboration requires agreement on each declared field, without exception among them.

Affirmative contradiction is not the mirror image of that condition. Under [0498], a received observation affirmatively contradicts the input where it agrees on no declared corroboration field and disagrees on at least one within the same test. Both halves are required.

A residual class sits between the two: an observation that neither agrees on all nor disagrees on any is counted in neither. Alongside it, [0498] recites that the absence of any received observation is not a contradiction and is not converted into one, and [0226] states the first of those propositions in the corroboration-factor definition.

Only the contradicting class carries a consequence here. Where an independently received governed observation affirmatively contradicts the input's payload, the admission evaluator (120) issues a reject outcome bearing a failed-corroboration classification, one of five rejection-reason classes recited at paragraph [0225] alongside insufficient authority, failed continuity validation, staleness, and window violation.

The corroborating class feeds a quantity instead. Paragraph [0226] defines the corroboration factor as the quotient of the count of independently received governed observations corroborating the payload field by a declared corroboration target, clamped at unity. It is one of four factors, each in the closed interval from zero to one, and the evidential weight is the sum over those four of the product of each factor and its

declared coefficient, the coefficients being non-negative and summing to unity. That numerator counts corroborating observations, and no term is stated by which a contradicting one enters the sum.

For an input against which no contradicting observation was received, what follows is conditional, and paragraph [0227] sets it by an ordered evaluation in which the first step whose trigger is satisfied issues its outcome. A reject outcome issues there on a credential that does not verify or whose scope excludes the action class, on failure of the acceptance conditions of Section 6.4, on an elapsed interval exceeding the validity duration, on a candidate pair outside an applicable proximity window or declaring non-intersecting partition sets, or on an evidential weight below a declared admission floor. Later steps issue the escalate outcome upon a declared convergence count of independently emitted observations bearing distinct identity primitives; the defer outcome where no second governed observation (610) has arrived within the window or the corroboration factor is below a declared floor, the input being held in the deferral queue; the solicit outcome where that factor stays below the floor after the deferral-expiration parameter has elapsed; the gate outcome where the evidential weight is below a declared admission threshold, the input being admitted at the product of that weight and a declared gate reduction factor; and otherwise the admit outcome. Nothing in that sequence admits an uncontradicted input by itself.

Two outcome vocabularies run in parallel and should not be merged. The reject outcome is a member of the graduated set of Section 6.15. The rejected determination (124) is a member of the closed set of Section 1.8, and paragraph [0050] recites that where the admission evaluator (120) produces that determination, a refusal counter (304) is incremented and writes the authorization gate (300) upon satisfaction of a threshold.

Declared Parameters, and What the Filing Leaves to Policy

Paragraph [0195] fixes what the governing policy object is before fixing what it declares. It is a signed policy object (112) resolved for the settlement context, issued by a deploying authority, bearing a policy identifier and a successor index, and resolved by the policy reference field (110) of the party applying it. Among the settlement parameters it declares are the factor coefficients and thresholds of Section 6.15, the solicitation bound, and the deferral-expiration parameter. Where the parties resolve different policy objects, each applies its own.

Against that, the classification's own parameters divide as follows.

- Corroboration fields are declared for the transaction class. The filing recites the set and its per-class scoping, and states no field list.
- Field tolerance applies to a field declared of a continuous type. The filing fixes the rule, that a difference within the declared tolerance is agreement, and states no tolerance value.
- The corroboration target is declared and forms the denominator of the corroboration factor, the quotient being clamped at unity.
- Factor coefficients are declared as non-negative values summing to unity.
- The admission floor, the floor on the corroboration factor, the admission threshold, the gate reduction factor, the convergence count, the deferral-expiration parameter, and the solicitation bound are each declared. The filing recites the role of each and states no value for any of them.
- The self-escalation discount factor is declared greater than zero and not exceeding unity. That bound is stated; the value is not.
- Further corroboration outcomes beyond the three classes are reserved to the governing policy object.

Paragraph [0228] offers one worked trace by way of illustration. A governing policy object there declares factor coefficients of three tenths, three tenths, two tenths, and two tenths for the authority, continuity, freshness, and corroboration factors, a corroboration target of four, and an admission threshold of seven tenths. An input with an authority level of evidential-weight specification nine tenths, a continuity factor of 0.775, an elapsed interval of twenty units against a validity duration of one hundred, and two corroborating observations obtains an evidential weight of 0.7625, and the admit outcome issues at that weight.

Where the Classifier Sits in the Governed Stack

Section 10.13, which carries the partition-scoped escrow and further settlement rules, holds the classification, and it reaches three ways.

Its input type is the governed observation of Section 6.1, recited at paragraph [0196] as a signed structure comprising an authority credential field, a continuity hash field, a spatial reference field, a temporal reference field, a time-to-live field, a payload field, and an observation lineage field. Paragraph [0197] adds a threshold condition: an observation lacking a verifiable authority credential is not a governed observation within the meaning of the disclosure, and a receiving party verifies the credential and the continuity hash field against its own memory field (102).

The classification runs over the payload field. Sibling fields feed the other three factors of [0226]: the authority factor is the evidential-weight specification of the level borne by the authority credential, the continuity factor is the trust-slope consistency measure computed at paragraph [0206] from the continuity vector field and no other field, and the freshness factor is unity less the quotient of the elapsed interval by the validity duration, clamped above at unity. Corroboration is the one factor of the four computed from observations of parties other than the emitter.

Upward, the classification supplies the composite admissibility evaluation of Section 6.15, where paragraph [0225] recites a settlement admissibility evaluator evaluating each governed observation, each candidate pair, each release request, and each rollback, and producing exactly one outcome from the graduated set. Pairing is governed separately by Section 6.2, where paragraph [0199] binds a first and a second governed observation into a settlement record (604).

Nearby, paragraph [0500] recites a distinct use of counting: a settlement may complete conditionally on an aggregation of a declared count of contributing governed observations, so that a settlement may be made to require a quorum of corroborating tenders rather than a single acceptance. That is a completion condition on a pair, not a classification of an observation against a payload. Paragraph [0229] adds that a settled pair is subject to challenge only through a credentialed dispute procedure.

Sideways, the classification shares its shape with Section 1.8. There, a contradiction is affirmative only where a retrieved entry holds a recorded value inconsistent with the conduct asserted. Here, it requires disagreement on at least one declared corroboration field together with agreement on none. In both, an absent or silent record lands somewhere other than the contradicting class.

Reading This Against Voting, Fusion, and Stance Detection

Several established design families do neighboring work, and the distinctions are structural.

Quorum and Byzantine agreement protocols decide by counting votes toward a threshold over a defined membership, handling a non-response as unavailability against liveness. What is recited here is not an agreement protocol: it carries no membership set and no quorum requirement, it renders a class per received observation rather than a group decision, and an absent observation is neither a contradiction nor convertible into one.

Sensor and data fusion resolves divergence among redundant measurements through outlier rejection, weighted averaging, or covariance-based estimation, absorbing it into a fused estimate. Under this filing, corroboration is quantified into a factor clamped at unity while affirmative contradiction takes a categorical outcome bearing a named reason class, and no term is stated by which a contradicting observation enters the weighted sum.

Reputation and trust-score systems aggregate ratings and non-ratings into a scalar attached to a party. Here the class attaches to a single observation relative to a specific input, under fields declared for that input's transaction class, and not to a party.

Claim verification and stance classification in natural language processing produce a three-way label over textual evidence: support, refute, or not enough information. The three-way shape is close, and the differences lie in what surrounds it. This mechanism operates on signed governed observations bearing verifiable authority credentials, excludes both the input's emitter and the evaluating party from the corroborating class, defines agreement by declared field type, and terminates in an admissibility outcome within a settlement procedure rather than in a label for human reading.

Conflict detection in record linkage and in replicated data types compares field by field, then resolves by merge policy, last-writer-wins, or a commutative merge function, seeking convergence to a single state. The object here is admissibility, and the recited response to an affirmative contradiction is a reject outcome bearing a named classification, not a merge.

None of the foregoing asserts anything about a particular product, protocol, or party; the families are named only to mark what this mechanism recites.

Disclosure Scope

Three-way corroboration classification and its failed-corroboration reject trigger are disclosed in U.S. Provisional Application No. 64/117,812 at Section 10.13, paragraph [0498], with the conditional completion forms of that section at paragraph [0500]. Supporting disclosure appears at Sections 6 and 6.1, paragraphs [0195] through [0198]; Section 6.2, paragraph [0199]; Section 6.4, paragraph [0206]; Section 6.15, paragraphs [0225] through [0228]; Section 6.16, paragraph [0229]; and Section 1.8, paragraphs [0041] through [0047] with paragraph [0050]. This article describes the mechanism as filed. Parameter values, field lists, and thresholds not stated in that specification are declared by the governing policy object of a given deployment and fall outside what is disclosed here. Publication is for defensive and technical disclosure purposes and is not an offer, a license, or a 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 \(/articles/conduct-admission/ratchet-freeze-uncomplied-correction\)](/articles/conduct-admission/ratchet-freeze-uncomplied-correction).
- [Three-Way Corroboration Classification: Corroborating, Contradicting, or Neither, and the Failed-Corroboration Reject Trigger \(/articles/conduct-admission/three-way-corroboration-classification-reject-trigger\)](/articles/conduct-admission/three-way-corroboration-classification-reject-trigger).
- [Untested-Class Authorization Decay: Metering Agent Permission Persistence by Recorded Corrective Encounters \(/articles/conduct-admission/untested-class-authorization-decay\)](/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).