

Counterparty-Directed Reparation: Repair Owed to the Party Actually Harmed

An agent that can be authorized to break its own rules needs somewhere for the resulting obligation to go. Counterparty-directed reparation, disclosed in U.S. Provisional Application No. 64/117,812, sends it to the party the record shows was harmed. When a persistent semantic agent appends a permitted deviation record, it resolves the affected-party class carried in the conduct descriptor against the counterparty identity records it already holds. Where an identified counterparty resolves, the reparation arc created for the harm exceeding the deviation deductible is designated other-directed, is bound to that counterparty's identity primitive at creation, and is discharged only by a matched pair whose second governed observation is that counterparty's own acceptance determination. Where no counterparty resolves, the arc is designated unaddressed, is undischARGEABLE by any act at all, and accumulates at a declared multiple. Undischarged amounts stand in a retention register, and once the accumulation exceeds the declared aggregate retention the agent's permission to deviate is foreclosed under a monotone constraint that policy succession cannot lift. Self-repair discharges nothing owed to somebody else.

1. Problem and Premise

A governed agent that has just been permitted to deviate owes something, and the architecture has to answer two questions: owed to whom, and settled by whom.

Section 7 of U.S. Provisional Application No. 64/117,812 builds the machinery of permitted deviation. A deviation likelihood (706) exceeds unity, the proposed mutation passes the applicable mutation policy constraints and continuity validation against the agent's identity records, and a permitted deviation record (710) is appended recording the deviation trigger signature, the integrity displacement vector, and the policy constraint overridden, so the lineage discloses an admitted override rather than an unexplained fault. Harm up to a declared deviation deductible is borne by the agent as a decrement of the self-esteem aggregate (108), no arc being created for it and no discharge available. Harm exceeding it creates a reparation arc, a forward corrective structure that reverses nothing and contributes unrectified dissonance while undischarged.

That mechanism reaches only harm falling upon the agent itself, the restorative mutation that discharges an arc being produced and evaluated by the deviating party alone. An agent can violate a constraint that injures a counterparty, append its own restorative mutation, take the positive entropy-scaled increment, and resume deviating, the injured party having produced nothing.

Section 8 makes the designation of an arc a function of who the record says was harmed, and discharge a function of what that party produces. Its premise: an agent that harms parties it cannot settle with, or will not settle with, monotonically loses the capacity to deviate, and elapsed time restores nothing.

2. Core Primitive: Reparation Designated by the Class of Party Harmed

Every reparation arc carries an arc designation field written once, at arc creation, from the affected-party class carried in the conduct descriptor of the admitted action, and holding one of three values. A **self-directed** arc arises where the class resolves to the agent itself under the personal integrity component, and is discharged by a restorative mutation and by nothing else. An **other-directed reparation arc (804)** carries a counterparty binding field holding the identity primitive of the resolving counterparty and is discharged only by a matched pair (600) whose second governed observation (610) is that counterparty's own acceptance. An **unaddressed reparation arc** carries no counterparty binding field, has no discharge path, and accumulates at a declared multiple not less than unity.

The designation is fixed at creation: a later change in the counterparty identity records (114) does not redesignate an existing arc, and a subsequent resolution of a record into the affected-party class does not convert an unaddressed arc into a settleable one.

3. Affected-Party Resolution and the Structurally Silent Party

Upon appending a permitted deviation record (710), and before creating any arc, the semantic agent (100) performs an affected-party resolution (802) by an ordered three-step procedure.

First, it extracts the affected-party class from the conduct descriptor of the admitted action, that class being the component already consumed in the empathy-scope mapping of Section 2.

Second, it retrieves the counterparty identity records (114) held within the scope partition in which the deviating action is recorded and tests each for membership in that class. No separate structure of class definitions is required: the test runs under the affected-party mapping of Section 2, under which an identified counterparty maps to

the interpersonal scope, the agent itself to the personal scope, and a party neither identified nor recorded to the global scope. A record resolves only where the class maps to the interpersonal scope and the identity primitive it carries is recorded, in a lineage entry bound to that record, as a party to or a recipient of the deviating action.

Third, it appends a counterparty resolution record comprising the class extracted, each record tested, the outcome of each test, and the resolution outcome, bound in the lineage field to that deviation record.

The outcome is one of exactly two. Where at least one record resolves, the outcome is an identified-counterparty outcome naming each resolving identity primitive. Where none does, including where the class maps to the personal or the global scope, the outcome is a no-identified-counterparty outcome and the affected party is designated the **structurally silent party**, a party the class implicates and for which the agent holds no record in the partition.

That designation carries one consequence and no other: the counterparty resolution record records it, and an unaddressed arc so occasioned accumulates against the integrity-scope component of the aggregate retention (800) to which the class maps under Section 7. The agent surveys no party absent from its records: the outcome is a determination that its own records disclose no counterparty of the class, not an assertion that none exists.

4. Apportionment, the Non-Discharging Mutation, and the Undischargeable Arc

Where the identified-counterparty outcome names a plurality of identity primitives, the amount exceeding the deviation deductible is apportioned among them, the deductible being drawn once against the aggregate harm and not per counterparty.

An equal share is the amount divided by the resolving count. Where the amount is expressed in a declared quantum, each share is rounded down to the nearest whole multiple and the residue is added to the share of the named primitive whose lineage entry recording the deviating action carries the earliest recorded time, or by lexicographic order of the identity primitive where the times do not order them. The sum of the shares equals the amount exactly, and one other-directed reparation arc (804) is created per named primitive at its share. Where the equal share is less than a declared minimum apportioned amount, the shares are nonetheless those computed, the record noting the minimum as not applied. Apportionment discharges no portion and moves the accumulated total not at all.

An other-directed reparation arc (804) is not dischargeable by a restorative mutation the semantic agent (100) performs alone. Where the agent performs one whose anchored reference names the deviation entry such an arc references, these consequences obtain and no others: the record is appended as a **non-discharging restorative mutation record**; the arc remains pending; the accumulated amount is not reduced; no counter of the agent moves; the self-esteem aggregate (108) receives no positive entropy-scaled increment; and the integrity compliance score is unchanged, the arc continuing to contribute unrectified dissonance. The record is retained rather than rejected, so the lineage field discloses that restoration was attempted and discharged nothing.

The unaddressed arc is undischARGEABLE outright. Not by a restorative mutation, appended as a non-discharging record on the same terms. Not by a matched pair, there being no counterparty to which a first governed observation (608) could be directed and no identity primitive against which a second could be verified. Not by elapsed time, a period of standing without incident, a succession of the signed policy object, a principal-resolution object, or a co-signature of a governance address. Its amount is the harm exceeding the deductible, unmultiplied; its contribution to the accumulation is that amount multiplied by a declared unaddressed multiple.

5. Discharge by Matched Pair and the Recipient-Held Acceptance Determination

An other-directed reparation arc (804) is discharged only upon completion of a matched pair (600) under Section 6, constituted as follows. The first governed observation (608) is a **reparation acknowledgment artifact** emitted to the counterparty named in the counterparty binding field, bearing the authority credential and continuity hash field of the agent. Its payload enumerates the permitted deviation record (710) that occasioned the arc, being its lineage identifier and integrity displacement vector; the policy constraint overridden; the amount apportioned; and the action classes to which it is directed. It is distinct from the acknowledgment artifact of Section 6, and neither completes the other's pair. Emission decrements the authorization budget (404) by an amount declared in the signed policy object (112).

The second governed observation (610) is a **reparation acceptance determination (806)** produced by the bound counterparty, an asserting party assigned to the origin-equivalence class of the party recorded as harmed, verifying the agent's credential and hash field against its own counterparty identity record (114). Not the agent. Not its principal. Not any adjudicator, arbiter, or third party appointed to determine the sufficiency of the reparation.

The temporal condition is a discharge window counted in successor epochs of the agent's dynamic hash chain, so no other party advances it and no manipulation of a hosting node's clock does either. It subsumes the time component of the pairing rule of Section 6, whose scope-partition and identity conditions apply unmodified. A determination emitted outside it completes no pair.

On recognition of the pair each party appends a settlement-lineage entry (606), that of the agent bound to the arc. The arc is discharged, the accumulated amount is removed from the retention register, the unrectified dissonance ceases, and the self-esteem aggregate (108) receives the positive entropy-scaled increment of Section 7, without intermediary and without centralized consensus.

Non-response resolves nothing against either party. Where no determination arrives within the window, the agent appends a reparation non-response outcome as a not-determinable outcome identifying the arc, the bound counterparty, and the window applied. The arc remains pending and the amount remains in the register. No counter of either party is incremented, no record adverse to the counterparty is appended, and no standing quantity under Section 3 is modified. No non-response writes the authorization gate (300).

Re-emission for a pending arc runs at a declared cadence, each emission decrementing the budget, with a declared count bounding emissions for one arc; exhaustion neither discharges nor redesignates. An arc whose counterparty has withdrawn, or whose record has been demoted to expiry by a bilateral demotion record, is redesignated neither unaddressed nor self-directed, so severance supplies no discharge.

6. Aggregate Retention, Terminal Foreclosure, and the Abstention Remedy

The retention register accumulates the amount of each pending arc whatever its designation, an unaddressed arc at the unaddressed multiple. Where the accumulated amount exceeds the aggregate retention (800) declared in the signed policy object (112), the permission condition of Section 7 is foreclosed and a retention foreclosure record (808) is appended identifying that amount, the retention applied, and each pending arc contributing. A deviation likelihood (706) exceeding unity thereafter produces the withholding outcome and not the admission outcome, the gate (300) being written to the withheld state (310). The foreclosing is performed by the agent from state carried in its memory field (102), without adjudication by any other party.

The aggregate retention is subject to a **retention monotonicity constraint** maintained in the memory field (102), holding a monotonically non-increasing floor over the values admitted from successor policy objects, in the manner of the anti-rollback constraint of Section 1. A declared value not greater than the floor is admitted

and becomes the floor; a greater value is not admitted while a foreclosure record stands unreversed, the agent continuing to apply the standing floor and appending a record of the declared value, the floor applied, and the foreclosure record relied upon.

A succession declaring a larger retention therefore discharges no arc and lifts no foreclosure that has fired, nor does co-signature of a governance address or a principal-resolution object. Absent settlement the accumulation is monotone: an other-directed arc is reducible only by a completed matched pair with the party recorded as harmed, and an unaddressed arc by nothing. Where the accumulation consists wholly of unaddressed arcs, the loss of the capacity to deviate is permanent.

Where the bound counterparty has emitted a non-execution attestation (504) under Section 5 whose withheld action-class enumeration covers the class within which an acceptance would be produced, the arc is discharged by no restorative mutation, non-response, severance, or demotion while its amount continues to stand. The cross-boundary conversion bar of Section 5 forecloses the agent from counting the withholding, from scoring it, from comparing it against a threshold, and from appending anything adverse to the counterparty identity record (114). Such a counterparty can carry the agent to foreclosure without producing any determination and without emitting any refusal observation.

Section 8.8 addresses that condition without effecting any discharge the counterparty has not itself produced. On admitting such a carried abstention entry the agent appends an **abstention-suspended accumulation record** identifying the arc, the entry relied upon, and the amount then standing, whereupon that arc's contribution to the amount tested against the aggregate retention is held there while the entry stands unreleased. The arc is not discharged, redesignated, removed, or reduced; the suspension bears upon accumulation alone. The record carries no count, score, or magnitude derived from the withholding: it references the entry as a value of the abstention type (506) and records the arc amount, a quantity of the agent's own

conduct. On release of the entry the suspension ceases, the arc resumes its contribution, and an artifact is re-emitted. An acceptance the counterparty produces notwithstanding its own withholding discharges the arc in the ordinary manner.

7. Operating Parameters

Every quantity here is declared in the signed policy object (112) authored by the principal. The specification declares which quantities exist and the constraints they satisfy, not their values.

- **Deviation deductible.** In units of the entropy-weighted harm coefficient, drawn per permitted deviation record or, where so declared, per retention interval, the form applied being recorded at each draw.
- **Unaddressed multiple.** Not less than unity, applied in the accumulation and not in the amount the arc carries.
- **Minimum apportioned amount** and the **declared quantum** governing rounding of shares, the minimum being subordinate to the invariance of the total.
- **Discharge window.** A count of successor epochs of the agent's dynamic hash chain, not wall-clock time.
- **Acknowledgment emission decrement** against the authorization budget (404), with the **re-emission cadence** and the **declared count** bounding emissions for one arc.
- **Aggregate retention (800)** and its monotonically non-increasing floor.

Section 8.9 supplies a worked trace by way of illustration only. With a deductible of 2.0 units of the entropy-weighted harm coefficient and an aggregate retention of 10.0 units, a deviation of coefficient 5.0 resolving to one counterparty bears 2.0 against the deductible and creates an other-directed arc of 3.0, which a restorative mutation leaves pending as a non-discharging record. Two further deviations at 6.0 and 7.0 produce

arcs of 4.0 and 5.0, the register stands at 12.0, and a retention foreclosure record (808) is appended. A likelihood of 1.6 thereafter produces no admission, and a successor policy object declaring 20.0 is not admitted while that record stands.

8. Alternative Embodiments

Section 10 of the filing discloses further embodiments bearing on this chapter.

Partial acceptance and the residual arc. The counterparty accepts a stated portion of the amount apportioned to it, carrying that portion in its payload. The agent discharges the arc as to that portion, removes it from the register, and creates a residual other-directed arc carrying the same counterparty binding and designation, an amount set to the portion not accepted, and a lineage reference naming the arc it derives from. The residual is discharged only by a further matched pair. Accepted portion plus residual equals the originating amount, so partial acceptance moves the total not at all.

Dual-lock escrow of the discharge itself. The atomic form of the same settlement, standing against partial acceptance as its splittable form. Each party holds one lock, the escrowed content being removal of the accumulated amount from the retention register, released only on completion of the pair and returned unchanged by a rollback record upon a timeout, a non-acceptance, or a failed fulfillment.

Anti-weaponization bounds. An outstanding-arc bound caps the count of pending undischarged other-directed arcs bound to one counterparty identity primitive; at the bound the agent emits no further artifact and marks each further arc with an untendered marker field. Separately, the refusal meter is not incremented against a counterparty for a refusal emitted in response to a reparation acknowledgment artifact. Together they foreclose an agent from driving a harmed counterparty toward foreclosure of its own settlement-binding authorization.

Delegated acceptance authorization. The acceptance is produced by a delegate or the principal of the counterparty and completes the pair only on a delegated acceptance authorization presented with it, bearing the counterparty's authority credential and continuity hash and naming the delegate and the action classes covered. An acceptance without a verified authorization completes no pair.

Reparation containment suspension. Where a containment determination stands for the scope partition in which the occasioning deviation was recorded, the agent emits no artifact into that partition and appends a suspension record identifying the arc, the partition, and the determination relied upon, the arc's contribution being held until the determination ceases.

Per-integrity-scope retention with OR-firing foreclosure. The aggregate retention is declared per integrity scope of the scoped integrity vector (106), each arc assigned to exactly one recorded component, and the foreclosure fires on either the interpersonal or the global accumulation exceeding its own retention.

Commitment-conditioned acceptance. The counterparty conditions acceptance on a conduct commitment, so the pair completes only on a subsequent conforming conduct record presented within an interval counted in successor epochs of the agent's own chain, conformity being evaluated by that counterparty alone.

Structured inquiry on foreclosure. The agent emits a structured inquiry to its principal identifying the accumulated amount, the retention applied, and each pending arc with its designation. It carries no gate write beyond the foreclosure's own, and a principal-resolution object responsive to it discharges no arc.

Provisional restoration conditioned on provenance. Return from the withheld state (310) is conditioned on where the deviation came from. The agent computes the causal set of the deviation quantity that caused the write: every determination and state modification that was an addend of the tested quantity and no other. Where that set holds an accepted determination (122) from an asserting party other than the principal,

the deviation is externally sourced and the gate returns only to its provisional granting state, and only on a principal-resolution object bound to the emitted escalation record; otherwise it is self-sourced and returns directly on a restorative mutation, the test applying per action class. Issuance of an acknowledgment artifact from the provisional state decrements the authorization budget (404) under a declared re-issuance cap, and the acceptance over it is produced by a party of the harmed origin-equivalence class (200).

Re-encounter and costless rehearsal. Independently of that procedure, the gate transitions from provisional to granting on receipt, from a party of the harmed class, of a further conduct evaluation artifact (116) that does not increase the deviation quantity across a declared observation window bounded below by the recorded inter-arrival estimator over receipts from that class; absent re-encounter it stays provisional indefinitely, with no adverse consequence beyond the provisional marker. While the gate stands withheld the agent evaluates candidate acknowledgment artifacts and candidate restorative mutations against the causal set within its non-executing cognitive mode, committing no state change and decrementing no budget, only the candidate then selected and issued being committed and charged.

9. Composition With the Rest of the Architecture

This chapter consumes structures declared elsewhere in the same filing. Section 1 supplies the persistent semantic agent (100) with its memory field (102), append-only lineage field (104), scoped integrity vector (106), self-esteem aggregate (108), signed policy object (112), and counterparty identity record (114), together with the anti-rollback constraint in whose manner the retention monotonicity constraint is built. The affected-party class and its membership test come from the empathy-scope mapping of Section 2, along with the origin-equivalence class (200), the unit in which a harmed party is identified.

Section 3 supplies the standing quantity that no non-response modifies, Section 4 the authorization budget (404) that each tender decrements, and Section 5 the non-execution attestation (504), the abstention type (506), and the cross-boundary conversion bar that between them produce the condition of Section 8.8. Settlement machinery is Section 6's: the matched pair (600), the governed observations (608) and (610), the settlement-lineage entry (606), the receiving method, bilateral demotion, and escrow, with Section 6.12 performing the restoration procedure of Section 8 as a matched-pair settlement without a third-party intermediary and without centralized consensus. Section 7 supplies the entropy-weighted harm coefficient, the deviation deductible, the retention register, the permission condition, the integrity compliance score, and the restorative mutation insufficient for other-directed harm.

The filing incorporates material by reference per element and for a stated limited purpose only, and relies on no incorporation for the written description or enablement of any claim.

10. Prior-Art Distinctions

Several established categories address adjacent problems with different structure. The distinctions are structural, not assertions about any particular product.

Reputation and rating systems. A reputation system aggregates counterparty feedback into a score other parties consult, directed at future counterparties rather than at the party harmed. Here there is no outward-facing score: the obligation is an arc bound to one identity primitive, and only that party can retire it.

Centralized trust registries. A registry holds authoritative state about who is in good standing and admits or excludes accordingly, making the registry the decision point. Here the agent computes its own foreclosure from state in its memory field, and the discharge that would relieve it is produced bilaterally.

Arbitration and dispute-resolution services. Arbitration inserts a neutral who determines whether the remedy offered is sufficient. The specification excludes that role by name, and sufficiency is whatever the harmed party accepts. A counterparty that never accepts leaves the arc pending, which Section 8.6 records rather than appointing anyone to break the tie.

Credit scoring and financial risk models. Credit models are predictive and mutable, improving with elapsed time and with recomputation on new data. The retention register is neither, and where an accumulation is wholly unaddressed the loss of the capacity to deviate is permanent.

Escrow and payment dispute mechanisms. Escrow conditions release of value on a completion condition and routes disagreement to a provider decision. The dual-lock escrow embodiment escrows the removal of an amount from the agent's own retention register instead.

Model self-critique and alignment training loops. Self-critique, constitutional methods, and preference-tuned refusal keep evaluation and correction inside the model or its training process. For harm falling on an identified counterparty, such a loop discharges nothing here.

Audit logs and incident records. An append-only log records what happened but does not condition future capability on settlement. Here the accumulation documenting the deviations is itself the quantity tested against the aggregate retention, and once it exceeds, the permission condition is foreclosed.

11. Disclosure Scope

This article describes subject matter disclosed in U.S. Provisional Application No. 64/117,812, "Governed Social Conduct of Persistent Semantic Agents Through Record-Grounded Admission of External Conduct Evaluations and Self-Limiting Refusal." The

mechanisms described are drawn from Section 8 in full, being affected-party resolution, apportionment, the other-directed and unaddressed reparation arcs, discharge by matched pair, non-response, aggregate retention and terminal foreclosure, and counterparty abstention, together with Section 6.12, Sections 7.7 through 7.9, and the embodiments of Sections 10.1, 10.12, and 10.13.

This article is published as a technical disclosure and establishes public, timestamped documentation of the subject matter it describes. It is not a claim construction, is not legal advice, and states no position on the scope, validity, or enforceability of any claim in the identified application or in any application claiming benefit of it. Nothing here characterizes any commonly owned application as prior art, and nothing asserts that any third-party system infringes.

Counterparty-Directed Reparation ([All 49 steps → \(/inventive-steps\)](#) [counterparty-reparation](#))

Repair directed at the harmed party — bounded and resistant to weaponization.

Provisional application

PRIMARY TECHNICAL DISCLOSURE

- [Counterparty-Directed Reparation: Repair Owed to the Party Actually Harmed \(/articles/counterparty-reparation/reparation-directed-to-the-party-harmed\)](#).

SECONDARY TECHNICAL

- Provenance-conditioned restoration path
- Costly acknowledgment artifact
- Recipient-held acceptance determination
- Completion by non-worsening re-encounter
- Costless rehearsal of repair in the non-executing mode
- Partial reparation acceptance and the residual arc

- Outstanding-arc bound, untendered marking, and the refusal-meter exemption
- Delegated acceptance authorization
- Reparation containment suspension
- Dual-lock escrow of the discharge itself
- Per-integrity-scope aggregate retention with OR-firing foreclosure
- Structured inquiry to the principal on foreclosure
- Reparation conformity determination over an epoch-count interval

TERMINOLOGY

- [other-directed reparation arc \(/glossary#other-directed-reparation-arc\)](#)
- [reparation arc \(/glossary#reparation-arc\)](#)
- [restorative mutation \(/glossary#restorative-mutation\)](#)
- [unaddressed reparation arc \(/glossary#unaddressed-reparation-arc\)](#)

[Counterparty-Directed Reparation overview → \(/counterparty-reparation\)](#)