

Narrative Coupling

Interpretive Charge Accumulation in Stratified Agent Architectures

Stephen A. Putman

putmanmodel@pm.me

Version 0.2 — March 2026

Abstract

This paper introduces Narrative Coupling, an engine-agnostic mechanism for modeling how meaning becomes sticky inside an artificial agent. Existing agent architectures often model perception, memory, and action, but leave under-specified the intermediate process by which recurring symbolic overlap acquires interpretive force. Narrative Coupling addresses this gap by defining a minimal accumulation layer between event detection and durable narrative consequence.

In the proposed mechanism, incoming events are represented as symbolic features or tags and compared against an agent’s currently active symbolic field. When overlap is sufficiently high, the event contributes interpretive charge to one or more thematic accumulators. Repeated contributions over time generate thematic pressure, which in turn biases future perception, recall, mood, and response. This process explains how identical event streams can yield divergent internal narratives across agents without requiring hand-authored story branches or metaphysical claims about objective meaning in the world.

The paper formalizes Narrative Coupling in terms of an event stream, an active symbolic field, a coupling gain term, state-sensitive weighting, thematic accumulation, contradiction handling, and saturation dynamics. It further shows how the mechanism can support a range of agent profiles, from low-coupling analytic agents to high-coupling agents prone to fixation under weak contradiction damping. A small deterministic micro-demo illustrates the core claim: the same ambiguous event sequence can produce distinct thematic trajectories depending on coupling parameters and active symbolic state.

Narrative Coupling contributes a portable interpretive layer for stratified agent architectures. It offers a computational account of how symbolic recurrence becomes felt significance, and how that significance begins shaping later interpretation before memory commitment, identity revision, or policy selection occur.

1. Introduction

Artificial agents are often described in terms of what they detect, what they remember, and what they do. Less often formalized is the intermediate process by which repeated symbolic overlap begins to feel significant to the agent itself. An event may be registered correctly without remaining neutral in consequence; once it resonates with an already active interpretive field, it can contribute to a growing thematic pressure that biases later perception, recall, and response. This paper introduces Narrative Coupling as a minimal mechanism for modeling that accumulation process in stratified agent architectures.

The central claim is simple. Agents do not merely detect events and pass them forward unchanged. They accumulate interpretive charge when incoming events overlap with an already active symbolic field, and that accumulated charge alters how subsequent ambiguity is interpreted. Over time, repeated overlap may gather around themes such as exclusion, warning, return, concealment, grief, or care. Once a theme becomes sufficiently active, later events are no longer encountered as isolated inputs. They are filtered through an interpretive pressure that shapes what stands out, what is remembered, and what feels narratively salient.

This paper does not treat narrative as authored story content, nor does it claim that recurring patterns are objectively meaningful in the world. Its concern is narrower and more architectural: how an agent comes to treat patterns as meaningful. Narrative Coupling therefore names neither memory storage nor full identity revision, but an intermediate accumulation layer between perception and durable consequence. It is the process by which meaning begins to stick across time.

That middle layer matters because many existing formulations leave a gap between raw event representation and downstream behavioral consequence. A system may correctly detect a slammed door, a delayed reply, a remembered object, or a repeated phrase without yet explaining why those events begin to cluster into a felt sense of rejection, warning, reconciliation, or return. Narrative Coupling addresses that gap by modeling how symbolic overlap accumulates across events into thematic pressure, and how that pressure biases future perception before it hardens into memory commitment, policy selection, or narrative identity.

In this framing, the same event stream need not produce the same internal narrative across agents. Differences in active symbolic state, affective weighting, contradiction handling, and coupling gain can lead one agent toward suspicion, another toward grief, another toward reconciliation, and another toward near neutrality, even when the raw observations remain unchanged. Narrative Coupling is proposed as the mechanism that explains this divergence without requiring hand-authored story branches or metaphysical assumptions about hidden meaning. It is an engine-agnostic interpretive layer that can be embedded within broader stratified architectures while remaining conceptually distinct from semantic identity, social field propagation, cross-modal binding, and durable memory persistence.

2. Problem Statement

Artificial agents often contain components for perception, memory, and action, yet many architectures remain under-specified in how interpretation gathers force across time. If an event is detected, tagged, and stored, something important may still be missing: the mechanism by which repeated symbolic overlap changes the felt interpretive weight of later events. Without such a mechanism, systems tend to fall back on two unsatisfying extremes. Either each event is treated as largely isolated and narratively inert, or higher-order narrative meaning must be imposed through authored rules, explicit branching, or post hoc explanation.

The missing issue is not whether an event is perceived, but how prior symbolic activation alters the uptake of what comes next. Agents do not confront each new event as blank receivers. They carry forward active thematic residues, partial tensions, unresolved interpretations, and state-

sensitive sensitivities. When a later event overlaps with those active conditions, it may contribute more than information. It may contribute charge. Repeated charge accumulation can create a local interpretive gravity, making some readings more likely and others less available.

This becomes especially important when events are ambiguous or only loosely related at the level of literal causality. A missed greeting, a loud door, a repeated song, a remembered object, and a delayed reply may remain separate at the level of raw observation, yet converge within an agent around a shared thematic cluster such as rejection, warning, or return. Existing systems can often represent each event individually. What is frequently absent is a clean account of how overlap among them becomes thematically adhesive.

Narrative Coupling is proposed to address that architectural gap. It defines a minimal intermediate process in which symbolic overlap between incoming events and an active field contributes to thematic accumulation over time. Once accumulated, that thematic pressure biases later perception, recall, mood, and response. In this sense, Narrative Coupling is not a theory of objective omens, hidden destiny, or story truth. It is a theory of interpretive adhesion inside the agent.

The paper does not argue that patterns are meaningful because they recur. It argues that an agent may increasingly treat them as meaningful when recurrence interacts with active symbolic state under sufficient coupling gain. More specifically, Narrative Coupling is not defined by accumulation alone, but by the accumulation of interpretive frames that alter the meaning-assignment of later events. A system may accumulate salience without producing Narrative Coupling; narrative coupling occurs when accumulated thematic pressure changes how subsequent ambiguity is interpreted rather than merely how strongly prior features remain active. This is the point at which generic carryover becomes narratively relevant.

Narrative Coupling therefore occupies a specific middle layer in stratified agent architectures. It is not semantic identity across environments, not shared representation across agents, not cross-modal binding by itself, not social field contagion, not durable memory storage, and not full narrative identity. It is the process by which meaning begins to adhere before later architectural layers decide what persists, what updates, and what action follows.

3. Core Mechanism

Narrative Coupling begins from a simple observation: events do not enter an agent as isolated facts once an interpretive field is already active. Each incoming event arrives with symbolic features that may overlap, reinforce, contradict, or remain orthogonal to what is already thematically live inside the agent. The proposed mechanism models how such overlap produces interpretive charge, and how repeated charge gradually accumulates into thematic pressure.

At the most general level, the mechanism has six stages:

event → symbol extraction → overlap with active field → interpretive charge → theme accumulation → biased future perception

An event is any incoming unit of significance-bearing input available to the agent. This may be perceptual, linguistic, spatial, social, or memory-evoked. The mechanism does not require that events share a modality. It requires only that they be representable in a symbolic feature space broad enough to support thematic comparison.

Symbol extraction maps the event into a set of features, tags, or weighted symbolic markers. These may be explicit labels such as exclusion, warning, return, or care, or may be lower-level features later grouped into such themes. The important point is not the exact vocabulary, but that the event becomes comparable to an already active interpretive state.

The active symbolic field represents what is currently live inside the agent. It is not the entirety of memory, and it is not identical to a final narrative identity. It is a transient, state-sensitive field of currently salient symbolic tendencies, partial activations, unresolved themes, recent interpretive residues, and contextual weighting conditions. It should be understood as an interpretive surface rather than a stored record. It is active, bounded, and revisable.

A theme, in this paper, is not just a cluster of similar tags. It is an interpretive frame that can bias the uptake of later events. Themes are therefore not merely descriptive categories but active framing states. A feature contributes local symbolic content; a theme organizes repeated symbolic contributions into a developing orientation such as warning, exclusion, return, concealment, grief, or care. That distinction matters because the object of accumulation is not just feature activation, but interpretive framing pressure.

Overlap measures the degree to which the extracted features of the new event resonate with the active symbolic field. This need not be binary. An event may partially support several themes at once, weakly reinforce one and strongly contradict another, or remain largely outside the currently active field. Overlap is therefore best understood not as raw similarity alone, but as context-sensitive compatibility between incoming symbolic content and the interpretive conditions already active inside the agent.

When overlap is sufficiently high, the event contributes interpretive charge. Charge is the incremental increase in thematic weight assigned to one or more candidate themes. Charge is not yet durable memory, belief, or action selection. It is the amount by which the event increases the felt relevance or salience of an interpretive frame.

Across time, repeated charge contributions yield theme accumulation. A theme such as rejection, reconciliation, concealment, grief, or warning becomes stronger as multiple events contribute compatible charge to it. This accumulation may be amplified by affective state, reduced by contradiction, limited by saturation, or allowed to decay. The result is not a single event causing a narrative, but a distributed buildup across a sequence of overlaps.

Once thematic accumulation becomes sufficiently strong, it produces biased future perception. Later events are no longer interpreted as if they arrived in a neutral field. Instead, they are encountered under thematic pressure. Certain readings become easier, faster, or more available; others become less likely or require stronger contradiction to surface. In this way, prior

interpretive accumulation changes the uptake of later ambiguity without requiring explicit story scripting.

This mechanism explains how the same observable sequence can generate different internal trajectories across agents. One agent may treat a sequence of ambiguous cues as weakly related and remain largely neutral. Another may experience the same sequence as increasingly charged toward loss, warning, or return. The difference lies not only in the events themselves, but in how those events are filtered through an already active symbolic field and accumulated over time.

Narrative Coupling therefore models the onset of internal story formation without reducing narrative to authored plot. It formalizes the process by which symbolic recurrence becomes interpretively adhesive. Events do not need to be objectively linked in the world in order to become linked inside the agent. They need only to overlap repeatedly with what is already active enough to receive them.

4. Formal Framing

A lightweight formalism helps clarify the mechanism without overstating mathematical precision. The goal is not to impose unnecessary algebra, but to specify the update structure clearly enough that the mechanism can be implemented, tested, and compared across systems.

Let E_t denote the incoming event at time t .

Let S_t denote the symbolic feature representation extracted from that event.

Let F_t denote the agent's active symbolic field at time t .

Let T_t denote the vector of thematic accumulators active in the agent at time t .

The active field F_t may be understood as a transient interpretive state composed of recent residues, weakly active themes, affective weighting, contextual salience, and other non-durable factors that shape uptake. It is not identical to the full memory store. It is the currently live interpretive surface against which new events are compared.

For each incoming event, the first relevant quantity is an overlap function:

$$O_t = \text{overlap}(S_t, F_t)$$

where O_t may be scalar or theme-distributed depending on implementation. In a simple version, overlap may be computed as weighted symbolic similarity. In a richer version, it may reflect support for one or more interpretive frames under the current field. This distinction matters. If overlap is treated only as similarity, the mechanism collapses toward ordinary activation carryover. If overlap is treated as frame-sensitive compatibility, the mechanism becomes more clearly interpretive.

Interpretive uptake is then modulated by a coupling gain g , along with a state weighting term W_t . The weighting term may incorporate factors such as fear, loneliness, arousal, trust, uncertainty, prior fixation, or recent affective residue. The resulting charge contribution may be written as:

$$\Delta I_t = g \cdot W_t \cdot O_t$$

where ΔI_t is the interpretive charge added at time t . In practice, ΔI_t may distribute across several themes rather than entering a single bucket.

Each thematic accumulator T_t^k for theme k is then updated according to incoming charge, contradiction handling, and saturation or decay:

$$T_{\{t+1\}}^{\{(k)\}} = \text{clip} \left((1 - \lambda) T_t^{\{(k)\}} + \Delta I_t^{\{(k)\}} - D_t^{\{(k)\}}, 0, C^{\{(k)\}} \right)$$

where λ is a decay parameter, $D_t^{\{(k)\}}$ is contradiction damping for theme k , $C^{\{(k)\}}$ is a saturation cap, and clip constrains the accumulator to an allowed range.

This form captures several important behaviors. First, themes can build gradually rather than instantly. Second, they can weaken over time if not reinforced. Third, contradiction can actively reduce them rather than merely fail to increase them. Fourth, saturation prevents unbounded growth and allows thematic pressure to level off rather than spiral indefinitely.

Contradiction should be understood here as active evidence for an incompatible thematic interpretation, not merely the absence of further support. A contradiction term therefore responds to features or events that reduce confidence in the currently dominant frame or positively support competing frames. This point is important because a system that merely stops reinforcing a theme is not yet correcting it. Narrative Coupling requires the possibility of real interpretive revision.

The active symbolic field itself may also be updated as a function of thematic accumulation:

$$F_{\{t+1\}} = \text{updateField} (F_t, S_t, T_{\{t+1\}})$$

This matters because Narrative Coupling is recursive. As themes strengthen, they alter the interpretive field that later events encounter. This means the mechanism is not merely additive. It is path-dependent. Earlier accumulation shapes later overlap, and later overlap further reshapes the field.

A forward bias state may then be derived from the current thematic vector:

$$B_{\{t+1\}} = \text{bias} (T_{\{t+1\}}, F_{\{t+1\}})$$

where $B_{\{t+1\}}$ affects later interpretation, recall weighting, salience ranking, or response preference. To keep the middle-layer boundary operational, this paper treats the output of

Narrative Coupling as bounded thematic-pressure state plus forward-bias weights. These outputs may reweight salience ranking, interpretive scoring, or retrieval preference, but they do not directly authorize durable memory writes, belief commitment, identity revision, or policy finalization. Those decisions belong to downstream persistence and governance layers.

Several implementation choices remain open and should remain engine agnostic. Symbol extraction can be rule-based, embedding-based, hybrid, or hand-authored. Overlap can be cosine-like, graph-based, symbolic, or thresholded. Themes can be predefined bins, learned clusters, or constrained latent channels so long as they function as interpretive frames rather than mere storage slots. The contribution of affect can be explicit or folded into the state weighting term. None of these choices alter the core claim of the paper. The claim concerns the existence of an intermediate accumulation process, not a single mandatory encoding scheme.

The most important architectural consequence of this formal framing is separation. Interpretive charge accumulation should remain distinguishable from durable memory commitment, identity revision, and policy selection. Narrative Coupling describes how meaning gains pressure before later layers determine whether that pressure is remembered, acted on, revised, or allowed to dissipate.

5. Agent Variability

Narrative Coupling is useful in part because it allows the same general mechanism to produce different interpretive styles across agents. Not all agents should accumulate thematic pressure in the same way, and not all agents should carry interpretive residue forward with equal strength. A system that lacks such variability risks either flattening all agents into the same narrative sensitivity or treating thematic accumulation as an exceptional edge case rather than a general architectural layer.

The first major axis of variation is coupling strength. Low-coupling agents exhibit weak thematic carryover. They may register symbolic overlap correctly, but incoming events contribute only modest interpretive charge, and thematic accumulators either grow slowly or decay quickly. Such agents tend to remain relatively literal, detached, or analytic in their treatment of ambiguous sequences. They are less likely to cluster loosely related events into a single developing theme unless the overlap is repeated, strong, and consistent.

Moderate-coupling agents display what might be considered ordinary narrative sensitivity. Symbolic overlap contributes enough charge for themes to form gradually, but contradiction and decay still constrain accumulation. These agents can develop suspicion arcs, recovery arcs, grief loops, or reconciliation trajectories, but such arcs do not harden immediately. Their interpretations remain revisable, and later counter-evidence can still shift the thematic balance. For many practical systems, this regime may offer the most plausible baseline for everyday meaning formation.

High-coupling agents exhibit stronger symbolic adhesion. Overlap with an active field contributes larger increments of interpretive charge, thematic clusters form more quickly, and later ambiguity is drawn more strongly into whatever themes are already active. Such agents may

appear especially sensitive, intuitive, emotionally reactive, or symbolically impressionable. In some settings this may be desirable. A creative companion, a grief-aware therapeutic simulation, or an archetypally tuned narrative agent may benefit from stronger thematic carryover than a literal planning system would.

A second critical axis is contradiction handling. Two agents with similar coupling strength may still diverge sharply depending on whether contradictory evidence actively reduces thematic accumulators or merely fails to reinforce them. Strong contradiction handling keeps thematic buildup revisable. Weak contradiction handling permits interpretive inertia. If contradiction is too weak relative to coupling gain, an agent may begin to force later ambiguity into an already active theme even when fresh evidence no longer supports it.

This creates a failure-prone regime: extreme coupling plus weak contradiction damping. In that regime, thematic accumulation can become self-reinforcing faster than it is corrected. The agent may begin to over-read coincidence, treat weak overlap as strong confirmation, or collapse unrelated events into a single interpretive arc. At mild levels this may look like heightened narrative sensitivity. At stronger levels it can become fixation, projection, or paranoia-like thematic lock. The mechanism described in this paper therefore does not imply that stronger coupling is always better. Coupling must be balanced by contradiction, decay, and saturation if the system is to remain interpretable and stable.

Additional variability can arise through state weighting. A fearful agent may assign more charge to warning-like overlap. A lonely agent may be more sensitive to themes of return, abandonment, or care. A grieving agent may bind neutral recurrences into memory-laden clusters more easily than a stable one. In this way, agent differences need not be encoded solely in static traits. They can also arise from transient conditions that reshape what kinds of overlap are most adhesive at a given time.

This makes Narrative Coupling especially useful for stratified architectures. It allows variability in interpretive style without requiring that every difference be hard-coded at the level of explicit narrative rules. Agent individuality can emerge from parameter differences in gain, damping, saturation, decay, and weighting rather than from authored plot branches. The mechanism remains shared, but the resulting internal narratives differ.

A practical advantage of this framing is that it supports both stability and expressiveness. A low-coupling, contradiction-sensitive configuration may suit a monitoring or compliance-oriented agent. A moderate-coupling configuration may suit a social companion or believable NPC. A higher-coupling but still well-damped configuration may suit an artistic, emotionally responsive, or mythically framed system. By adjusting a small number of parameters, one can produce qualitatively different interpretive behaviors without redefining the mechanism itself.

Narrative Coupling therefore does not describe a single narrative style. It defines a parameterized accumulation process whose behavior depends on how strongly an agent binds overlap into theme, how readily contradiction reduces those themes, and how long thematic residue remains active. These differences help explain why identical event streams need not produce identical internal trajectories across agents, even under a common representational substrate.

6. Cross-Modal Extension

Narrative Coupling becomes more powerful when symbolic overlap is allowed to form across modalities rather than only within a single stream of like-form input. An agent's interpretive field is rarely shaped by language alone, or vision alone, or memory alone. In realistic environments, thematic pressure often emerges when cues from different channels converge into a shared symbolic cluster. A sound, a phrase, a gesture, a color, a remembered object, and a silence may all contribute to the same developing interpretation even though they differ at the level of sensory form.

This matters because thematic accumulation is not fundamentally about modality. It is about symbolic compatibility under an active interpretive frame. A slammed door may overlap with an active field of exclusion or warning. A delayed reply may contribute to the same cluster. A remembered object may reactivate loss or return. A repeated phrase on a radio may reinforce concealment, fate, or reconciliation depending on the agent's current thematic state. These cues do not need to be causally linked in the environment to become linked inside the agent. They need only to land within a sufficiently similar symbolic neighborhood.

Cross-modal operation therefore strengthens the paper's central claim. Narrative Coupling is not merely a tagging system that groups similar perceptual items. It is a mechanism for thematic convergence across heterogeneous inputs. Different kinds of events can produce overlapping charge if they are mapped into a common symbolic space or otherwise routed toward the same thematic accumulators.

In practice, this means the symbolic feature representation S_t need not originate from a single extractor. Visual inputs might contribute tags such as distance, threat, familiarity, or return. Linguistic inputs might contribute rejection, care, warning, or concealment. Spatial inputs might contribute approach, barrier, or absence. Memory-evoked cues might contribute grief, attachment, or unfinishedness. Once represented in a compatible symbolic scheme, these different feature sources can all participate in the same overlap and accumulation process.

The interpretive effect of this convergence can be substantial. A single ambiguous event may carry limited charge on its own. But when multiple cross-modal cues weakly reinforce the same theme, the total thematic pressure may become significant. A muted conversation, an unopened message, a familiar scent, and a returned object may each contribute only partial overlap. Together, however, they may raise a theme such as loss, reconciliation, or warning above threshold. In that sense, cross-modal Narrative Coupling allows meaning to emerge cumulatively without requiring any one event to bear the entire narrative burden.

This section also clarifies what the paper is not claiming. Narrative Coupling does not say that all cross-modal binding is narrative, nor that any two concurrent cues should be fused into meaning. Mere co-occurrence is insufficient. The cues must overlap with an already active symbolic field strongly enough, often enough, or under sufficient state weighting to contribute charge to a shared thematic cluster. Narrative formation remains selective, not automatic.

This selectivity is one reason contradiction handling remains important in the cross-modal case. Multiple weakly aligned cues can amplify a theme quickly, especially in a high-coupling agent. Without damping, the system may over-cluster loosely related signals and mistake ambient noise for thematic coherence. Proper contradiction handling, decay, and saturation help keep cross-modal accumulation from collapsing into indiscriminate symbolic inflation.

Within stratified architectures, the cross-modal extension makes Narrative Coupling especially portable. A host system may differ widely in how it performs perception, retrieval, symbolic tagging, or modality fusion, yet still use the same accumulation layer once features are made comparable. This is one reason the mechanism is best treated as engine agnostic. It sits above raw modality-specific processing and below durable consequence, allowing many different perception stacks to feed a common interpretive process.

The broader implication is straightforward. Narrative pressure often emerges not from one dramatic event, but from repeated, distributed, cross-modal reinforcement. Narrative Coupling offers a computational way to model that process. It explains how heterogeneous cues can gather into a felt theme, and how that theme begins shaping later interpretation before it becomes memory, identity, or action.

7. Deterministic Micro-Demo

To make the mechanism concrete, this paper proposes a small deterministic micro-demo. The purpose of the demo is not to simulate full narrative intelligence, but to isolate the claim that identical event streams do not guarantee identical internal narratives. If Narrative Coupling is real as an architectural layer, then fixed symbolic inputs should produce different thematic trajectories when processed under different active fields and coupling parameters.

The demo should therefore be engine agnostic, minimal, and replayable. A simple Python command-line implementation is sufficient. The core input is a fixed event stream represented as a sequence of symbolic feature sets. Each event is pre-tagged or deterministically mapped into symbolic features so that variation in output cannot be dismissed as stochastic perception noise. This keeps the demonstration focused on interpretive accumulation rather than on upstream extraction variability.

A minimal socially legible stream might contain six to eight ambiguous events such as a greeting not returned, a door shutting loudly nearby, a familiar song playing unexpectedly, a remembered object appearing in view, a stranger glancing over, a delayed reply, and a later supportive message. The events should be chosen so that no single interpretation is forced by the stream itself. Their role is to provide overlapping symbolic material that can plausibly contribute to multiple themes depending on the agent's current field and parameters.

To avoid over-reading the demo as anthropomorphic cherry-picking, the same mechanism can also be illustrated with a more neutral variant: a repeated beacon pulse, an interrupted route marker, an unexpected node reappearance, a missing acknowledgment signal, a confirming ping, and a later contradictory location cue. The point in both cases is the same. Narrative Coupling is

not tied to human social drama. It is a general mechanism for how repeated, frame-relevant overlap can bias the interpretation of what comes next.

Each agent profile in the demo begins with a baseline symbolic field and a small set of accumulation parameters. At minimum these include a coupling gain, contradiction damping, saturation cap, decay rate, and possibly state-sensitive weighting. The same event stream is then replayed across multiple agent profiles, for example a low-coupling contradiction-sensitive agent, a moderate-coupling agent, and a high-coupling weak-damping agent. Because the event stream is fixed, any divergence in outcome is attributable to the accumulation mechanism rather than to changes in input.

For each event, the demo should emit a deterministic trace containing the event identifier, extracted tags, overlap score, interpretive charge contribution, updated theme totals, dominant theme, and forward bias snapshot. These outputs need not be visually elaborate. A structured table or JSON log is enough. The key is that the trace makes the accumulation process inspectable step by step.

The proof case is simple. The same sequence should produce different thematic buildup across the three profiles. One agent may remain near neutral, another may gradually accumulate concern and then relax as supportive cues arrive, and another may lock more quickly into rejection or warning because early overlap is amplified and insufficiently damped. This demonstrates that divergence can arise from parameterized interpretive accumulation even when raw observations are held constant.

A corresponding break case is equally important. The demo should include a contradiction-heavy stream or a later contradiction phase in which new events challenge an emerging theme. Under strong contradiction damping, false thematic lock should weaken or reverse. Under weak damping, it may persist. This break case prevents the mechanism from looking like a mere paranoia engine and shows that thematic pressure can remain revisable rather than automatically self-sealing.

The deterministic micro-demo also helps differentiate this paper from neighboring work. It is not a semantic identity demo across environments, not a shared representation comparison across agents in a single instant, and not a social contagion or field propagation demo. Its focus is temporal accumulation inside an individual interpretive process. The unit of demonstration is not a single object or a crowd field, but a sequence of events acquiring meaning over time.

As a companion to the paper, the micro-demo has practical value beyond illustration. It provides a test harness for reasoning about parameter regimes, failure modes, contradiction handling, and thematic drift. It also offers a portable scaffold for later integration into broader systems, including stratified architectures that distinguish between interpretive accumulation, memory promotion, and action selection.

8. Position Within the Broader Stack

Narrative Coupling is best understood as a middle-layer mechanism within a larger stratified architecture rather than as a replacement for existing modules. Its role is narrow but important: it explains how repeated symbolic overlap acquires interpretive force before later systems decide what becomes durable, governable, or actionable. In that sense, it fills a gap already visible across the surrounding paper set.

Relative to PUTMAN, Narrative Coupling does not redefine interpretation as constrained reconstruction. PUTMAN already frames interpretation shift in terms of relational structure, contextual activation, overlap, recombination, and affect-modulated reconstruction. Narrative Coupling builds on that substrate but narrows its focus to a different question: how repeated event-level overlap contributes to thematic pressure across time. PUTMAN explains how interpretation can shift under changing activation and constraint; Narrative Coupling explains how symbolic recurrence becomes sticky enough to bias what the next event is likely to mean.

Relative to Resonant Field Mapping, Narrative Coupling does not replace runtime modulation, memory gating, governance, or divergence-triggered stabilization. RFM already separates tone modulation from memory persistence and authority, and explicitly insists that tone shifts must not automatically become durable identity-layer updates. What Narrative Coupling contributes here is a more explicit account of what may be happening inside the interpretive interval before persistence decisions are made. In that sense, it is compatible with an active field model: the agent's current field shapes uptake, and accumulated thematic pressure in turn reshapes that field. But Narrative Coupling stops short of governance and persistence. It is a mechanism of interpretive loading, not of memory commitment or lease-controlled action.

Relative to Memory Stratification, the distinction is also clear. Memory Stratification specifies how persistent agency depends on differentiated temporal layers, with short-term activation, compressed regularity, and structural priors operating at different write resistances and timescales. It also defines residual activation and pressure as precursors to assimilation rather than conflating all salience with durable memory. Narrative Coupling belongs upstream of those persistence decisions. It describes how thematic pressure forms while content is still interpretively active but not yet consolidated. This is why it should remain separable from memory promotion logic. A theme may become strong enough to bias later perception without yet meeting the conditions for longer-horizon assimilation.

Relative to Shared Representation, Divergent Interpretation, Narrative Coupling sits nearby but is not identical. That companion note isolates the fact that shared admissible observation and shared representation do not force shared meaning or shared memory consequence. It demonstrates divergence after representation and before persistence. Narrative Coupling takes the next step by specifying one mechanism through which that divergence may unfold over time. The companion note is about the possibility of divergence under shared representation. Narrative Coupling is about the temporal accumulation rule by which such divergence becomes progressively self-reinforcing within an individual agent.

Relative to S.U.N., the separation is simpler. S.U.N. is about preserving semantic identity across heterogeneous perceptual routes and environments. It explicitly separates semantic object identity from perceptual representation so that behavior can remain aligned across engines and

modalities. Narrative Coupling is not about semantic stability across environments. It assumes some event representation is already available and asks what happens when those represented events repeatedly overlap with an active interpretive field. S.U.N. stabilizes the referential layer; Narrative Coupling explains how temporal uptake across represented events acquires thematic force.

Relative to multi-agent field dynamics, the distinction is again one of scope. Field dynamics concerns how activation, coupling, reciprocity, concentration, and dissipation interact across populations, producing local amplification or systemic instability. Narrative Coupling does not begin with social propagation. Its default unit of analysis is the single agent over time. Social or field effects may later amplify, synchronize, or disrupt narrative buildup, but that is downstream from the mechanism introduced here.

Taken together, these relations show why Narrative Coupling is not redundant with the surrounding papers. The broader stack can now be described more cleanly:

PUTMAN provides the reconstruction substrate.

RFM provides active modulation and bounded persistence discipline.

Narrative Coupling provides thematic pressure accumulation across events.

Memory Stratification governs what persists and at what depth.

Shared Representation clarifies that common tags do not force common meaning.

S.U.N. separates semantic identity from perceptual route.

Field dynamics explains multi-agent amplification and dissipation.

Within that arrangement, Narrative Coupling earns its place by naming the intermediate accumulation process through which meaning begins to adhere before it becomes memory, policy, or collective cascade.

9. Limits and Failure Modes

Narrative Coupling is intentionally narrow, and its limits should be stated clearly. It is not a theory of objective significance in the world, not a general account of cognition, and not a claim that all recurrence is narrative. It models one architectural process only: how repeated symbolic overlap inside an active interpretive field can accumulate into thematic pressure that biases later interpretation. That narrowness is a strength, but it also defines what the mechanism cannot by itself guarantee.

A first limitation is false thematic lock. If coupling gain is too high, contradiction handling is too weak, or symbolic features are too loosely mapped, the system may begin to cluster events into a theme that is only weakly supported by the actual sequence. In such cases, ambiguity is not

merely interpreted; it is over-recruited. This can produce fixation, projection, or paranoia-like drift, especially when multiple weak cues are treated as mutually confirmatory without sufficient damping.

A second limitation is symbolic oversensitivity. In cross-modal settings, heterogeneous cues may converge quickly on a shared thematic cluster. That is part of the mechanism's value, but it also creates risk. If every co-occurring or repeated signal is treated as thematically relevant, the system loses selectivity and inflates ambient noise into narrative weight. Mere recurrence is not enough. Overlap must be constrained, contradiction must remain active, and saturation must keep accumulation bounded.

A third limitation is narrative flattening under over-damping. The inverse failure is also possible. If contradiction handling, decay, or saturation are tuned too aggressively, meaningful buildup may never occur. The system then remains too literal or too transient to support believable narrative carryover. Every event becomes locally processed and globally forgotten. This produces stability, but at the cost of expressiveness and temporal coherence.

A fourth limitation is premature entanglement with persistence or policy. Narrative Coupling should not be allowed to directly decide durable memory updates or irreversible action. If thematic pressure is allowed to bypass persistence discipline, then transient interpretive buildup may pollute identity-layer state or drive action escalation without adequate review. This is precisely why the surrounding architecture matters. The middle-layer claim only holds if module boundaries are preserved in practice.

A fifth limitation is opaque tuning and weak auditability. If the mechanism is implemented through hidden or uninterpretable mappings, it becomes difficult to distinguish legitimate thematic buildup from arbitrary drift. This is one reason a deterministic micro-demo matters. Inspectable overlap scores, change deltas, theme totals, contradiction effects, and forward-bias outputs make the mechanism reviewable. More generally, rationale traces and observable precursor signals become valuable whenever interpretive drift can affect later layers.

A sixth limitation is domain misuse. Narrative Coupling can be useful for believable NPC behavior, companion systems, creative agents, symbolic retrieval layers, and interpretive simulations. It is not by itself a sufficient basis for decisions in regulated, safety-critical, or high-authority domains. In those settings, thematic accumulation may still be modeled, but its outputs must remain subordinate to stronger evidentiary, governance, and persistence constraints.

These limitations suggest several design principles. Interpretive accumulation should remain bounded, revisable, and inspectable. Contradiction must reduce themes rather than merely fail to reinforce them. Saturation should prevent runaway growth. Cross-modal convergence should be selective rather than indiscriminate. Most importantly, Narrative Coupling should remain a middle layer. It may bias later perception, but it should not automatically become memory, identity, or action.

10. Conclusion

This paper introduced Narrative Coupling as an engine-agnostic mechanism for modeling how meaning begins to stick inside an artificial agent. The central claim has been narrow throughout: agents do not merely detect events and pass them forward unchanged. When incoming symbolic content repeatedly overlaps with an already active interpretive field, it can accumulate interpretive charge. Over time, that charge gathers into thematic pressure, and that pressure biases how later events are perceived, recalled, and framed.

The contribution is a missing middle layer. Many architectures already represent events, store memory, and generate responses. What often remains under-specified is the intermediate process by which recurring symbolic overlap acquires felt significance before durable consequence is decided. Narrative Coupling names that process. It explains how the same event stream can produce different internal narratives across agents even when raw observations remain unchanged.

The paper has also argued that this mechanism is portable. It does not depend on a single engine, embodiment, or perceptual stack. So long as events can be represented in a comparable symbolic feature space and evaluated against an active field, thematic accumulation can be modeled. This makes Narrative Coupling suitable as a reusable interpretive layer for stratified architectures rather than as a special-purpose narrative trick.

At the same time, the mechanism must remain bounded. Narrative Coupling is not semantic identity, not shared representation, not field contagion, not durable memory, and not action policy. It is the process by which meaning begins to adhere across time. Later layers must still decide what persists, what is revised, what is quarantined, and what may influence behavior.

The practical value of the mechanism is straightforward. It offers a disciplined way to model emergent narrative without hand-authored branching, to represent agent-specific meaning formation, to simulate fixation and recovery under explicit parameters, and to support richer social and emotional behavior in artificial agents. A small deterministic micro-demo is enough to make the core claim testable: identical event streams do not guarantee identical internal narratives when thematic accumulation differs.

In that sense, Narrative Coupling is modest but useful. It does not attempt to explain all interpretation. It explains how repeated symbolic overlap becomes thematically adhesive, and how that adhesion begins shaping later experience before it hardens into memory, identity, or action.