Detecting Structural Coherence Loss in Semantic Systems Under Perspective Shift

Sergi Garcia Mecinas
Independent Researcher, IA_m.ai
sos.prevencio@gmail.com
Preprint — January 2026

Abstract

Modern AI and semantic systems can remain locally functional while silently losing internal coherence when the perspective from which reasoning is initiated changes. Existing evaluation paradigms—similarity-based metrics, robustness analysis, and post-hoc explainability—primarily assess correctness or stability at the level of outputs or representations, often detecting failure only after action has already occurred.

We introduce ***structural coherence*** *as a property of semantic systems defined by the invariance of induced relational dynamics under* ***perspective shift****, operationalized as changes in the restart distribution of a structural random-walk process over a fixed graph. We formalize structural coherence as the complement of the expected distance between perspective-induced stationary distributions, instantiating the distance ￼d as* ***Jensen–Shannon divergence****. We define* ***structural coherence loss*** *as a leading-indicator failure mode in which equilibrium structure fragments and the system's conceptual subspace (axes) becomes misaligned across perspectives, despite unchanged nodes and locally plausible behavior.*

We demonstrate the phenomenon on a minimal experiment comparing a coherent graph G0G_0G0 with a ***minimally contaminated*** *variant G1G_1G1 produced by adding a small number of cross-community "bridge" edges while preserving node set and degree distribution. Using a dual-source structural propagation instrument (semantic diffraction), we show that small structural perturbations can induce substantial coherence degradation without visible system failure. We argue that coherence auditing provides an upstream diagnostic layer complementary to explainability, governance, and AI safety.*

1. Introduction

AI and semantic systems increasingly operate across multiple domains, roles, and viewpoints. In such systems, plausibility or correctness of outputs is often taken as evidence of sound reasoning. However, systems can remain operational while becoming structurally incoherent: their internal organization of meaning no longer remains stable when perspective shifts.

This paper isolates and formalizes a distinct failure mode:

Silent structural failure: the system continues to produce locally plausible outputs, but the internal relational geometry that supports meaning becomes unstable or viewpoint-dependent.

Many current evaluation paradigms are output-centric (accuracy, consistency, calibration) or representation-centric (embedding similarity, representation invariance). Post-hoc explainability (XAI) can provide narratives about why an output occurred, but typically does not measure whether the internal structure remains coherent across perspectives—nor whether the system is entering a fragile state prior to failure.

We introduce **structural coherence** as an observable property of semantic systems, defined in terms of invariance of relational dynamics under perspective shift. We then show, via a minimal controlled experiment, that structural coherence can degrade significantly under small structural perturbations even when no obvious behavioral malfunction is observed. We show that structural coherence loss functions as a **leading indicator**: a condition that increases fragility and risk before downstream failure manifests.

Scope note. We do not claim to model full-scale LLM internals. Our aim is to isolate a *structural* failure mode that can exist in semantic systems, knowledge graphs, and hybrid reasoning architectures, and to make it measurable.

2. Structural Coherence

We treat semantic systems as relational structures. Let $G=(V,E)$ be a directed graph, where nodes represent concepts and edges represent typed structural relations.

2.1 Operationalizing Perspective

A **perspective** is defined as a restart distribution over a set of seed nodes. This provides a computational interpretation of “starting reasoning from a viewpoint” without introducing subjective semantics.

Let $S_\phi \subseteq V$ be the seed set for perspective ϕ. Define the restart vector:

$$r_\phi(v) = \begin{cases} \frac{1}{|S_\phi|} & \text{if } v \in S_\phi \\ 0 & \text{otherwise} \end{cases}$$

A perspective shift is a change $S_\phi \rightarrow S_\phi$ while keeping the graph G unchanged.

2.2 Induced Structural Dynamics

For each perspective ϕ, we define an induced structural dynamic $D_\phi(G)$ via a Personalized PageRank (PPR) stationary distribution:

$$\pi_\phi = \mathrm{PPR}(G, r_\phi; \alpha)$$

where $\alpha \in (0,1)$ is the restart probability. π_ϕ is a probability distribution over nodes, representing steady-state structural influence under perspective ϕ.

2.3 Definition of Structural Coherence

Let $\Phi = \{\phi_1, \ldots, \phi_m\}$ be a set of perspectives. Define the distance between two induced dynamics as a distance over the stationary distributions.

We instantiate d as **Jensen–Shannon divergence** (JSD), which is symmetric and bounded:

$$d\left(D_{\phi_i}(G), D_{\phi_j}(G)\right) := \mathrm{JSD}\left(\pi_{\phi_i} \parallel \pi_{\phi_j}\right)$$

Then structural coherence is defined as:

$$\mathcal{C}(G) = 1 - \mathbb{E}_{\phi_i, \phi_j \in \Phi}\left[\mathrm{JSD}\left(\pi_{\phi_i} \parallel \pi_{\phi_j}\right)\right]$$

Intuition: coherence is high when perspective-induced structural flows remain similar (in distributional geometry), and low when flows become divergent.

2.4 What Structural Coherence Is Not

Structural coherence is not:

- similarity between representations (cosine in embedding space),
- logical consistency of propositions,
- factual correctness,
- output stability.

A system may remain correct and consistent locally while coherence degrades globally under perspective shifts.

3. Structural Coherence Loss as Leading Indicator

We define **structural coherence loss** as a condition where perspective-induced dynamics diverge substantially, causing instability in equilibrium structure and fragmentation of the conceptual subspace, even though local plausibility may remain intact.

3.1 Observable Signatures of Coherence Loss

For a perspective set Φ\Phi$\Phi$, coherence loss is characterized by one or more of:

1. **Distributional divergence increase:** $E[JSD]$ increases, thus $C(G)$ decreases.
2. **Equilibrium fragmentation:** different perspectives yield substantially different equilibrium nodes.
3. **Axis misalignment:** the conceptual subspace supporting equilibria changes across perspectives (low overlap in axis sets).

These signals can appear **before** explicit failure or incorrect outputs, making them useful as upstream diagnostics.

4. Observational Instrument: Dual-Source Structural Propagation ("Semantic Diffraction")

We employ a dual-source propagation mechanism as an observational instrument. The term **semantic diffraction** is used as an analogy; the actual mechanism is explicit and reproducible.

4.1 Dual-Pole Dynamics

Given poles A and B, define two restart distributions r_A and r_B (singletons), and compute:

$$\pi_A = \mathrm{PPR}(G, r_A; \alpha), \quad \pi_B = \mathrm{PPR}(G, r_B; \alpha)$$

4.2 Equilibrium Nodes

Define a balance-sensitive score:

$$s(v) = (\pi_A(v) + \pi_B(v)) - \lambda \, |\pi_A(v) - \pi_B(v)|$$

with $\lambda \geq 0$ controlling asymmetry penalization.

The **equilibrium node** for a pole pair (A,B) is:

$$e(A, B) = \arg \max_{v \in V \setminus \{A,B\}} s(v)$$

4.3 Stability (Operational)

We use two stability checks:

- **Balance:**

$$\mathrm{bal}(e) = \frac{|\pi_A(e) - \pi_B(e)|}{\pi_A(e) + \pi_B(e)} \leq \tau$$

- **Dominance ratio:** Let v_1 and

$$\mathrm{dom} = \frac{s(v_1)}{s(v_2) + \varepsilon} \geq \gamma$$

If either fails, the equilibrium is considered unstable for that perspective.

4.4 Axis Sets (Minimal Definition)

When axis metadata exists, we define the axis set supporting a perspective as the set of axes shared by the poles and the equilibrium candidate. Operationally, we report:

$$X_\phi := \mathbf{axes}(e) \cap (\mathbf{axes}(A) \cup \mathbf{axes}(B))$$

If explicit axes are absent, we treat axis sets as empty and focus on distributional and equilibrium-based measures.

5. Measuring Coherence Loss Across Perspectives

We now connect the formal definition *C(G)* to observable indicators computed from a perspective set.

5.1 Perspective Set Construction

In experiments, we instantiate each perspective ϕ as a pole pair (A_i, B_i), with restart distributions as singleton seeds.

$$\text{Thus, } \Phi = \{(A_1, B_1), \ldots, (A_m, B_m)\}.$$

5.2 Equilibrium Fragmentation

Let $e_i = e(A_i, B_i)$. Define fragmentation:

$$\mathbf{Frag}(\Phi) = \frac{|\{e_i\}_{i=1..m}|}{m}$$

Low fragmentation implies a shared equilibrium across perspectives; high fragmentation implies perspective-dependent centers.

5.3 Axis Alignment

If axis sets X_i are available, define mean pairwise Jaccard overlap:

$$J_X(\Phi) = \frac{2}{m(m-1)} \sum_{i<j} \frac{|X_i \cap X_j|}{|X_i \cup X_j|}$$

Low J_X indicates subspace divergence across perspectives.

5.4 Coherence Loss Condition

We say coherence loss occurs if coherence drops below a threshold:

$$\mathcal{C}(G) < \epsilon$$

and at least one structural signature indicates instability:

- *Frag(Φ)* exceeds τ_{frag}, or
- $J_X(\Phi)$ falls below τ_{axes}, or
- stable equilibria fraction falls below τ_{stable}.

5.5 Equilibrium Dominance (Hub Collapse)

Equilibrium fragmentation alone can be misleading: a system may appear less fragmented while collapsing toward a single high-attraction node. We therefore introduce equilibrium dominance as a complementary indicator.

Let E = {e_i} be the set of equilibrium nodes obtained across perspectives Φ. We define dominance as:

Dom(Φ) = max_v |{i : e_i = v}| / |Φ|

High dominance indicates degenerate convergence toward a single hub, reflecting structural collapse rather than coherent integration.

6. Algorithm 1: Structural Coherence Auditing

Input:

- Graph G = (V, E)

- Perspective set Φ = {(A1,B1), …, (Am,Bm)}

```
  - Restart probability α
  - Balance penalty λ
  - Stability thresholds (ratio_min, balance_max)

Output:
  - Structural coherence score C(G)
  - Equilibrium dominance Dom
  - Stability rate SR

Initialize empty list Π ← []
Initialize empty list E ← []
 for each perspective (Ai, Bi) in Φ do
    Compute πAi ← PPR(G, seed=Ai, α)
    Compute πBi ← PPR(G, seed=Bi, α)
    Compute combined field πφ ← normalize((πAi + πBi) / 2)
    Append πφ to Π
    for each node v in V \ {Ai, Bi} do
        score(v) ← (πAi(v) + πBi(v)) − λ |πAi(v) − πBi(v)|
    end for

    Let eφ ← argmax_v score(v)
    Append eφ to E
 end for

 Compute C(G) ← 1 − mean_{i<j} JSD(Π[i], Π[j])
 Compute Dom ← max_v frequency(v in E) / |E|
 Compute SR ← fraction of perspectives with stable equilibria

 return C(G), Dom, SR
```

The algorithm does not perform inference or generation. It operates purely as an observational instrument, quantifying divergence of structural dynamics under perspective shift.

7. Minimal Experiment

7.1 Why a Minimal Experiment?

The goal is not to claim scalability to full neural models, but to isolate a failure mode and demonstrate it under controlled conditions. The experiment shows that coherence is a global property and can degrade silently.

7.2 Graph Construction

We construct two graphs with the same node set:

- G_0: coherent graph
- G_1: contaminated graph

Both share the same nodes and almost all edges. G_1 is derived from G_0 by adding a small number n of "bridge" edges that alter relational geometry.

7.3 "Minimally Contaminated" Defined Structurally

Contamination is defined structurally, not semantically:

- We identify communities (or high separation regions) under baseline propagation.
- We add n edges that connect nodes across communities with low mutual reachability under π_ϕ.
- We preserve degree distributions approximately (avoid artificially creating hubs).
- We avoid injecting factual contradictions; local plausibility can remain.

Thus, coherence loss in G_1 is attributable to **structural perturbation**, not to "false data".

7.4 Perspectives

We evaluate a fixed set of pole pairs Φ across both graphs. Each perspective is a pole-pair restart distribution (Section 2.1).

7.5 Observed Outcomes

In our experiments, structural coherence decreases under minimal contamination. Specifically, $C(G_0) > C(G_1)$, indicating increased divergence between perspective-induced dynamics.

More importantly, we observe a marked increase in equilibrium dominance. While G_0 exhibits low dominance, with equilibria distributed across nodes, G_1 shows collapse toward a single node ("realidad") across multiple perspectives. This demonstrates that coherence loss may manifest not as fragmentation, but as hub-induced degeneracy.

7.6 Results

For the coherent graph G_0, we obtain $C(G_0)=0.8775$, with equilibrium dominance Dom=0.333. In contrast, the contaminated graph G_1 yields $C(G_1)=0.8474$ and Dom=0.667.

Despite unchanged nodes and locally plausible interpretations, minimal structural contamination induces measurable divergence in perspective-induced dynamics and a twofold increase in equilibrium dominance. Stable equilibria under strict criteria are rare in both graphs, which is expected in small conceptual structures; however, the qualitative difference lies in the pattern of convergence rather than in stability counts.

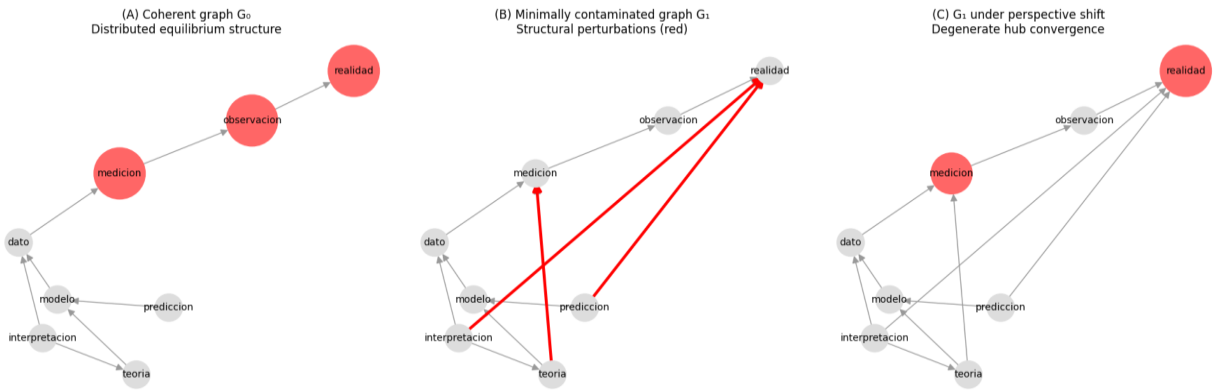

Figure 1 illustrates this phenomenon, contrasting distributed equilibrium structure in G_0 with degenerate hub convergence in G_1 under identical perspective shifts.

8. Discussion

8.1 Distinction from SOTA Robustness and XAI

Robustness analysis typically evaluates output or representation change under input perturbation; gradient-based sensitivity assumes differentiable parametric models. Post-hoc explainability provides narratives that may increase interpretability of outputs but does not quantify whether structural dynamics remain invariant across perspectives.

Structural coherence auditing differs by construction:

- it changes *perspective* (restart distribution),
- not inputs or parameters,
- and evaluates divergence in induced relational dynamics.

Therefore, coherence loss can occur even when:

- outputs are plausible,
- representations remain stable,

- explainability narratives appear consistent.

8.2 "So What?" — Why Coherence Loss Matters

A natural objection is: *if outputs remain correct, why should structural incoherence matter?*

We argue structural incoherence represents a state of **latent fragility**. The system may remain functional under a given perspective, but the loss of a shared structural center implies heightened sensitivity to future perturbations, composition, domain shifts, or multi-agent interactions.

Structural incoherence is analogous to a crack in a building's foundation: the building still stands, but the probability of catastrophic failure under stress increases significantly. Coherence auditing therefore serves as a **pre-failure diagnostic**, enabling earlier intervention than output-based checks.

8.3 Limitations

- The minimal experiment isolates a failure mode but does not claim universal empirical rates in real-world large-scale systems.
- Choice of perspective set Φ influences measured coherence; in operational settings, Φ should be domain-specific.
- JSD provides a stable and interpretable d, but alternative distance functions may be explored (Appendix B).
- When axis metadata is absent, axis misalignment must be approximated by other structural signatures.

9. IA_m Context (Minimal)

This work is developed as part of a broader investigation into **semantic observability** (see Appendix A). The results and definitions in the main paper do not require IA_m to be adopted as a full framework; IA_m merely provides an organizational context in which multiple structural auditors can coexist.

10. Conclusion

We introduced **structural coherence** as an invariance property of perspective-induced structural dynamics, formalized coherence using Jensen–Shannon divergence over stationary distributions, and defined **coherence loss** as a leading-indicator failure mode characterized by divergence, fragmentation, and subspace misalignment under perspective shift.

Through a minimal experiment comparing a coherent graph to a minimally contaminated variant, we showed that small structural perturbations can induce significant coherence degradation without visible system failure. Structural coherence auditing provides a complementary upstream diagnostic layer for semantic systems, AI governance, and safety.

Appendix A — IA_m Framework Context (Non-essential)

IA_m is a meta-framework for **structural observability** of semantic and reasoning systems. It treats meaning as a relational geometry subject to stability, drift, and perspective dependence. IA_m does not train models or generate outputs; it instruments systems to detect coherence loss before action.

Within IA_m, semantic diffraction is one observational instrument among others; the main paper isolates coherence loss independently of IA_m's broader scope.

Appendix B — Alternative Distance Functions ddd (Brief)

While this paper instantiates d as Jensen–Shannon divergence between stationary distributions, other candidates include:

- Total variation distance between π distributions
- Wasserstein distance over graph-embedded node coordinates (if available)
- Earth Mover's distance with topology-based costs
- Distances over derived observables (e.g., top-k overlap)

JSD was chosen for symmetry, boundedness, and interpretability as a divergence between probabilistic structural flows.

Appendix C — Notes on Reproducibility

The proposed definitions require:

- a fixed graph G,
- a defined perspective set Φ as seed distributions,
- a specified PPR implementation with restart probability α,
- and explicit hyperparameters $\lambda, \tau, \gamma, \epsilon$.

A minimal reference implementation can compute:

- $C(G)$ via mean pairwise JSD,
- equilibrium fragmentation,
- axis overlap (if metadata exists),
- and stability rates.