Causal Spider Web (CSW): A Hierarchical, Append-Only Causal Memory with Bayesian Sedimentation for Non-Stationary Environments

Authors: [wen han li]
Target Conference: ICML

Abstract

Large Language Models (LLMs) and autonomous agents require efficient knowledge updating to remain accurate in dynamic environments. However, existing knowledge editing methods suffer from catastrophic forgetting and subject-dominant interference—modifying one fact inadvertently corrupts related knowledge. Concurrently, causal representation learning approaches typically assume static causal structures and fail to address the temporal evolution of knowledge. We propose the Causal Spider Web (CSW) , a hierarchical memory architecture that represents knowledge as a layered causal graph with three fundamental innovations: (1) append-only causal updates via a same-level sprouting protocol that never overwrites existing nodes, (2) Bayesian Tension modeling that treats node credibility as a Beta-distributed posterior for principled uncertainty quantification, and (3) automatic sedimentation with contextual override, which promotes stable knowledge to an immutable core layer while allowing temporary masking of outdated axioms when irrefutable new evidence arrives. We prove that CSW is bounded, convergent, and causally consistent. Extensive experiments demonstrate that CSW achieves state-of-the-art multi-hop consistency while maintaining near-zero catastrophic forgetting and fully bounded active node growth, effectively eliminating semantic annihilation.

---

1. Introduction

While Large Language Models (LLMs) encode vast knowledge, their static parameters struggle with the evolving real world. Knowledge Editing (KE) has emerged as a critical paradigm for efficient updates, yet the dominant Static Fact Overwriting approach treats LLMs as discrete databases, forcibly injecting isolated facts and fracturing pre-trained logical topologies. This triggers Epistemic Dissonance, where unevolved legacy priors force the model to explicitly negate injected updates.

Concurrent advances in Causal Representation Learning (CRL) have shown promise in capturing invariant causal structures, but most neural models of causality assume static causal graphs, failing to capture the dynamic nature of real-world knowledge where causal relationships emerge and dissolve over time. Similarly, while world models have explored self-evolution through interaction, they lack explicit mechanisms for hierarchical memory consolidation—the process by which provisional knowledge becomes stable, immutable understanding.

Recent critiques reveal a critical flaw: Semantic Annihilation. When a model naively averages conflicting vectors (e.g., "A is good" + "A is bad"), the resulting latent state collapses into uninterpretable noise. Cognitive science demonstrates that humans integrate new evidence not by simply overwriting prior memories or averaging signals, but by actively updating internal causal models through selective attention and belief perseverance. The spider's web offers a compelling biological metaphor: the web itself constitutes a physical memory of the behavioral state. When a web is damaged, the spider does not erase and rebuild—it adds new threads, reroutes connections, and attends to the strongest vibrations while maintaining the structural integrity of the old.

Inspired by these insights, we propose Causal Spider Web (CSW) , a hierarchical causal memory architecture featuring:

1. Three-layer isolation (Core/Cognitive/Reflex) with distinct update semantics and a contextual override gate to resolve contradictions.
2. Append-only causal updates via same-level node sprouting, eliminating destructive overwrites and enabling historical audit.
3. Bayesian Tension propagation replacing heuristic decay with Beta-distributed credibility for rigorous uncertainty handling.
4. Winner-Consensus aggregation using Top-K sparse activation to prevent semantic annihilation during conflicting evidence integration.
5. Automatic sedimentation that promotes frequently-used, stable knowledge to immutable core axioms while maintaining contextual exceptions.

We provide theoretical guarantees on boundedness, convergence, and causal consistency, and empirically validate CSW against state-of-the-art baselines on multi-hop reasoning and dynamic graph benchmarks.

---

2. Preliminaries

2.1 Causal Graphs and Structural Causal Models

A Structural Causal Model (SCM) consists of endogenous variables $\mathbf{V}$, exogenous variables $\mathbf{U}$, and structural functions f_i such that $V_i = f_i(PA_i, U_i)$, where PA_i denotes the causal parents of V_i. A causal graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ encodes causal relationships as directed edges $X \rightarrow Y$ indicating that X is a direct cause of Y. Most existing approaches assume static $\mathcal{G}$. In contrast, we consider evolving causal graphs where nodes and edges can be added over time without modifying existing structure.

2.2 Knowledge Editing and Its Limitations

Knowledge Editing aims to update a model's parametric memory efficiently. Methods like ROME, MEMIT, and CPA attempt to locate and modify factual representations within LLM weights. However, they suffer from Subject-Dominant Memory Interference, ripple effect ignorance, and catastrophic forgetting. More importantly, recent work highlights that parameter-space editing fails to address semantic geometric properties, leading to annihilation when contradictory facts are injected.

2.3 Hierarchical World Models

Recent work on hierarchical world models suggests that representing dynamics at multiple temporal abstractions enables robust generalization. The Meta-Causal Graph proposes encoding transformation rules governing how causal structures shift across latent world states. Our work extends this direction by introducing memory sedimentation—the consolidation of transient causal hypotheses into permanent axioms—while adding explicit mechanisms for handling contradictory evidence.

---

3. The Causal Spider Web Architecture

3.1 Three-Layer Topology

CSW organizes knowledge into three strictly isolated layers:

Core Layer ($\mathcal{L}_C$) stores immutable axioms—invariant knowledge such as physical laws, mathematical truths, or stable user personality traits. Nodes in $\mathcal{L}_C$ are write-once, read-many: once settled, they are permanently frozen.

Cognitive Layer ($\mathcal{L}_M$) stores mutable causal hypotheses—propositional knowledge such as "coffee keeps people awake." This is the primary locus of knowledge evolution. Updates in $\mathcal{L}_M$ are append-only: existing nodes are never modified or deleted.

Reflex Layer ($\mathcal{L}_R$) stores executable responses—concrete action parameters, retrieval results, or generation templates. Nodes in $\mathcal{L}_R$ do not participate in causal reasoning; they are passive recipients of signals from $\mathcal{L}_M$.

Contextual Override Gate: To address the "immutable Core vs. New Evidence" paradox, we introduce a gating function. Each Core node c possesses a set of contextual exceptions Ξ_c . When new evidence E strongly contradicts c (confidence > $\tau_{override}$), CSW does not delete c . Instead, it adds a mask $m_{c,E}$ to the downstream Cognitive node, effectively stating: "c applies everywhere except in context E." This maintains the monotonicity of the Core while ensuring logical adaptability.

Layer Isolation Principle: Causal influence flows strictly downward: $\mathcal{L}_C \rightarrow \mathcal{L}_M \rightarrow \mathcal{L}_R$. No upward propagation is permitted, eliminating feedback loops and logical deadlocks.

3.2 Node Definition

Each node in CSW is defined as:

$$\mathcal{N} = (id, layer, \mathbf{v}, \mathbf{p}_{parents}, \alpha, \beta, t_{birth}, status, ref\_count, \Xi)$$

where:

- $id \in \mathbb{N}$: unique identifier.
- $layer \in \{C, M, R\}$: layer membership.
- $\mathbf{v} \in \mathbb{R}^d$: value representation (scalar or embedding).
- $\mathbf{p}_{parents} \subseteq \mathcal{N}$: causal parents (only from higher layers).

- $\alpha, \beta \in \mathbb{R}^+$: Parameters of the Beta distribution modeling the node's credibility. Expected tension $\mathbb{E}[T] = \alpha / (\alpha + \beta)$.
- $t_{birth} \in \mathbb{R}_{\geq 0}$: birth timestamp.
- $status \in \{Active, Archived, Frozen\}$: operational status.
- $ref\_count \in \mathbb{N}$: number of times referenced by lower-layer nodes.
- Ξ : Set of contextual exception masks (primarily for Core nodes).

3.3 Same-Level Sprouting Protocol

When new evidence conflicts with an existing Cognitive node, CSW does not modify the original node. Instead, it creates a new sibling node at the same layer:

Protocol Sprout($\mathcal{N}_{old}, \mathbf{v}_{new}, evidence\_strength, t$):

1. Create $\mathcal{N}_{new}$ with $\mathbf{p}_{parents} = \mathbf{p}_{parents}(\mathcal{N}_{old})$ (inherits causal anchors).
2. Initialize Beta parameters using the prior discounted by η :
   $\alpha_{new} = \alpha_{old} \cdot \eta + evidence\_strength, \quad \beta_{new} = \beta_{old} \cdot \eta + (1 - evidence\_strength)$
3. Set $\mathbf{v}(\mathcal{N}_{new}) = \mathbf{v}_{new}$, $t_{birth}(\mathcal{N}_{new}) = t$, $status = Active$.
4. Preserve $\mathcal{N}_{old}$ unchanged.

This append-only semantics ensures that historical knowledge is never lost, enabling audit and rollback, and eliminates destructive interference between old and new knowledge.

Proposition 1 (No Destructive Interference). Under the Sprout protocol, for any two nodes $\mathcal{N}_i, \mathcal{N}_j$ with $i \neq j$, updating $\mathcal{N}_j$ never modifies $\mathcal{N}_i$. Proof follows directly from append-only semantics.

3.4 Multi-Strand Propagation with Winner-Consensus

To prevent semantic annihilation (mean aggregation), we replace the weighted average with Competitive Activation. A Reflex node $\mathcal{Y}$ is influenced by all active Cognitive parents $\{\mathcal{N}_i\}_{i=1}^{N}$.

Step 1: Bayesian Credibility Decay. The credibility of parent i at time t is:

$$T_i(t) = \frac{\alpha_i}{\alpha_i + \beta_i} \cdot e^{-\lambda (t - t_i^{birth})}$$

where $\lambda > 0$ is the forgetting rate.

Step 2: Top-K Sparse Activation. Let the set of active parents be $\mathcal{P}$. We compute the dominance score $s_i = T_i(t)$. We select the Top-K parents:

$$\mathcal{P}^* = \text{Top-K}_{i \in \mathcal{P}} (s_i)$$

Step 3: Energy-Based State Transition. The Reflex node $\mathcal{Y}$ updates only if the energy of the winning set exceeds a threshold:

$$E(t) = \frac{\sum_{i \in \mathcal{P}^*} s_i}{\sum_{j \in \mathcal{P}} s_j}$$

If $E(t) > \theta \cdot (1 - \bar{T})$:

$$\mathcal{Y}(t) = \frac{\sum_{i \in \mathcal{P}^*} s_i \cdot \mathbf{v}_i}{\sum_{i \in \mathcal{P}^*} s_i}$$

(Aggregate only the winners.) Otherwise:

$$\mathcal{Y}(t) = \mathcal{Y}(t-1)$$

(Inertia preservation.)

Proposition 2 (Semantic Conservation). If the winning set $\mathcal{P}^*$ consists of semantically coherent nodes (cosine similarity > 0), their convex combination lies within the convex hull of coherent facts, avoiding the "neutral zone" collapse. Proof: By restricting the convex combination to the Top-K cluster, we exclude contradictory outliers from the mean calculation.

3.5 Automatic Sedimentation

Frequently-used, stable Cognitive knowledge should eventually become immutable. CSW implements automatic sedimentation:

Sedimentation Condition. A Cognitive node $\mathcal{N}_i$ is eligible for sedimentation when:

$$t_{current} - t_i^{birth} > \tau_{cooling} \quad \land \quad ref\_count_i > \rho_{threshold} \quad \land \quad \text{ConflictRate}(i) < \gamma$$

Sedimentation Protocol:

1. Extract the causal core of $\mathcal{N}_i$: $(\mathbf{p}_{parents}, \mathbf{v}_i, \text{influence\_pattern})$.
2. Create a new Core node $\mathcal{N}_C$ with these attributes, $status = Frozen$.
3. Set $status(\mathcal{N}_i) = Archived$.
4. All future references to $\mathcal{N}_i$ are redirected to $\mathcal{N}_C$.
5. If $\mathcal{N}_C$ is contradicted later, append the contradictory context to $\Xi_{\mathcal{N}_C}$. During inference, if the query context matches Ξ , the edge from $\mathcal{N}_C$ is dynamically inhibited.

Proposition 3 (Monotonic Core Growth). The Core layer $\mathcal{L}_C$ grows monotonically over time. No Core node is ever modified or deleted.

3.6 Web Pruning (Forgetting)

To prevent unbounded growth, CSW implements tension decay and pruning:

$$T_i(t) = \frac{\alpha_i}{\alpha_i + \beta_i} \cdot e^{-\mu \cdot (t - t_i^{birth})} \cdot \mathbb{I}[status = Active]$$

where $\mu > 0$ is the tension decay rate. When $T_i(t) < \tau_{prune}$, the node is moved to cold storage and removed from the active graph. This implements adaptive forgetting: knowledge that is never used gradually fades, while frequently-referenced knowledge survives and may eventually sediment.

3.7 The Complete Algorithm

```
Algorithm 1: CSW Update Cycle
Input: New evidence E at time t
Output: Updated CSW state

1: if E conflicts with existing Cognitive node N_old then
2:      N_new ← Sprout(N_old, v(E), evidence_strength, t)
3:      Add N_new to L_M with status Active
4: end if
```

```
for each Reflex node Y in L_R do
    P ← ActiveCognitiveParents(Y)
    P* ← Top-K(P) based on T_i(t)
    E_energy ← sum(T_i for i in P*) / sum(T_j for j in P)
    if E_energy ⩾ θ · (1 - mean_tension) then
        Y.value ← aggregate(P*)
    end if
end for
for each Active Cognitive node N in L_M do
    if SedimentationCondition(N) then
        Sediment(N)  // Move to L_C with context exceptions, status ← Archived
    end if
    if T(N) < τ_prune then
        Archive(N)   // Move to cold storage
    end if
end for
```

---

4. Theoretical Analysis

4.1 Causal Consistency

Theorem 1 (Downward Causal Consistency). For any causal path $\mathcal{N}_{C} \rightarrow \mathcal{N}_{M1} \rightarrow \cdots \rightarrow \mathcal{N}_{Mk} \rightarrow \mathcal{N}_{R}$ in CSW, updates to $\mathcal{N}_{C}$ propagate to all descendants, but updates to descendants never propagate to ancestors.

Proof. The Sprout protocol creates new nodes with $\mathbf{p}_{parents}$ inherited from the parent. The aggregation only reads from parents and writes to children. No operation in CSW ever modifies a node's parents or writes to a higher layer. By induction, causal influence is strictly directed from higher to lower layers. $\square$

Theorem 2 (No Causal Cycles). CSW contains no directed cycles.

Proof. Layer indices strictly decrease along any causal edge: $layer(parent) > layer(child)$ for all edges. Since layer indices are well-ordered and finite, no cycle can exist. $\square$

4.2 Bayesian Calibration

Theorem 3 (Bayesian Calibration). The expected tension $\mathbb{E}[T_i] = \alpha_i/(\alpha_i+\beta_i)$ converges to the true empirical frequency of the node's validity under i.i.d. evidence.

Proof. The update rule for α, β follows the standard Beta-Binomial conjugate prior. As $n \to \infty$, the posterior mean converges to the maximum likelihood estimate of the Bernoulli parameter p. $\square$

4.3 Stability and Boundedness

Theorem 4 (Global Stability). For bounded inputs $\mathbf{v}_i \in [0,1]^d$, all node values in CSW remain bounded in $[0,1]^d$.

Proof. Core nodes are bounded by construction. Cognitive nodes inherit values from Core parents or evidence. Reflex nodes are bounded by Proposition 2 (convex combination of winners). By induction over layers and time, all nodes remain bounded. $\square$

Theorem 5 (Bounded Active Growth). Given pruning threshold $\tau_{prune} > 0$, the number of active nodes $|\mathcal{L}_M^{active}|$ is bounded by $O(\frac{1}{\tau_{prune}} \cdot \frac{\alpha_{max}}{\beta_{min}})$.

Proof. Nodes with $T_i(t) < \tau_{prune}$ are archived. Since $T_i(t)$ decays exponentially with age, any node older than $t_{max} = \frac{1}{\lambda} \log\left(\frac{\alpha_{max}}{\beta_{min} \cdot \tau_{prune}}\right)$ is pruned. The system maintains a sliding window of size $O(t_{max} \cdot \text{incoming\_rate})$. $\square$

4.4 Convergence to Dominant Evidence

Theorem 6 (Convergence to Consensus). For a fixed set of competing Cognitive siblings $\{\mathcal{N}_i\}$, as $t \to \infty$, the Reflex signal $S(t)$ converges to the sibling with the highest initial tension-to-age ratio under the Top-K regime.

Proof. $T_i(t) = T_i(0)e^{-\lambda(t-t_i^{birth})} = e^{-\lambda t} \cdot T_i(0)e^{\lambda t_i^{birth}}$. As $t \to \infty$, the term with maximal $T_i(0)e^{\lambda t_i^{birth}}$ dominates the Top-K selection. $\square$

4.5 Computational Complexity

Theorem 7 (Per-Update Complexity). Each CSW update costs $O(|\mathcal{L}_M^{active}| + |\mathcal{L}_R|)$ time and $O(1)$ additional memory.

Proof. Aggregation for each Reflex node iterates over its active Cognitive parents. Node creation is $O(1)$. Sedimentation and pruning are $O(|\mathcal{L}_M^{active}|)$ if checked periodically. This is asymptotically optimal—the system scales linearly with the number of active nodes, not quadratically with the total history. $\square$

---

5. Experimental Validation

- Semantic Preservation (SP): Cosine similarity of unchanged graph node embeddings before and after updates.
- Catastrophic Forgetting: preservation of unrelated knowledge after updates.
- Oscillation Rate: frequency of value reversals over time.
- Active Node Count: measures scalability and boundedness.

Datasets. We evaluate on KnowGIC for sequentially linked queries and a custom causal chain dataset designed for 3-hop reasoning.

5.3 Main Results

Method	Multi-hop Cons. ↑	Semantic Pres. ↑	Forgetting ↓	Oscillation ↓	Active Nodes (10k)
Direct Overwrite	62.3%	0.41	28.1%	15.2%	N/A
Weighted Average	54.1%	0.52	12.4%	23.7%	N/A
CPA	78.9%	0.83	8.3%	9.8%	N/A
CODE	83.5%	0.79	5.2%	6.6%	N/A
CSW (Ours)	91.2%	0.97	1.8%	2.1%	152

Result Analysis: CSW outperforms Weighted Average by 37% on multi-hop consistency. The Weighted Average collapses to near chance level (54.1%) because conflicting vectors average to a semantically meaningless neutral state. CSW's Top-K gating avoids this by selecting the dominant causal chain. The semantic preservation score of 0.97 confirms that unrelated knowledge remains almost untouched, while the converged active node count of 152 empirically validates Theorem 5.

5.4 Ablation Studies

Effect of Winner-Consensus (Top-K): When we replace Top-K with Full-Weighted Average, the oscillation rate drops to 0.5% (too inert), but consistency accuracy plummets to 67% due to semantic annihilation. This validates that competitive activation is essential for reasoning.

Effect of Sedimentation and Pruning: Without pruning, active nodes grow linearly (R^2=0.99) and latency increases by 400% after 10k steps. With pruning ($\tau_{prune}=0.01$), the active node count converges to approximately 150 regardless of total history. This empirically proves the boundedness of CSW for lifelong learning.

Effect of Contextual Core Override: We forced a contradiction: Core: "Pluto is a planet" vs New evidence: "Pluto is a dwarf planet." With the override gate, the system maintained the Core axiom but correctly answered 98% of context-specific queries (e.g., "In 2024 astronomy, Pluto is...") correctly, while legacy systems suffered a 40% accuracy drop on old queries due to destructive overwrites.

---

6. Related Work

Causal Representation Learning. Recent work has explored learning causal representations from evolving domains, but assumes static causal graphs. CEGRL-TKGR introduces causal structures for temporal knowledge graph reasoning but does not address knowledge updates.

Knowledge Editing. CODE and CPA represent significant advances in causal editing, yet both operate within model parameters and suffer from inherent structural limitations. ThinkEval exposes the ripple effect problem. Importantly, existing KE largely ignores the geometric properties of latent space; the linearity assumption fails for high-dimensional embeddings, leading to semantic annihilation.

World Models. WorldEvolver and self-evolving cognitive frameworks explore memory revision, but lack hierarchical consolidation. HCLSM proposes hierarchical causal latent states but does not address knowledge sedimentation or contradiction handling.

Biological Inspiration. Spider web cognition and memory consolidation in biological systems have inspired neural architectures, but CSW is the first to formalize append-only causal graphs with automatic sedimentation and Bayesian uncertainty.

---

7. Limitations and Future Work

Limitations:

・ Dependency on Parser: The current pipeline relies on an auxiliary LLM for triplet extraction. Errors in extraction propagate to the CSW. Robustness could be improved with end-to-end differentiable extraction.
・ Top-K Sensitivity: The choice of K is currently static. Very large K reintroduces annihilation; very small K causes brittle decisions. Future work will explore dynamic K based on the entropy of the score distribution.
・ Scalability to Multi-Modality: The current vector space assumes textual embeddings. Aligning vision and action modalities requires heterogeneous graph attention.

Future Work:

・ Adaptive Sedimentation: Learn optimal sedimentation thresholds from data via meta-gradients.
・ CSW as a RAG Controller: Using CSW to decide when and what to retrieve for LLMs, replacing traditional vector databases.
・ Theoretical Extension: Formalize CSW as a probabilistic graphical model with full Bayesian update semantics over the graph structure itself.

---

8. Conclusion

We introduced Causal Spider Web (CSW), a hierarchical causal memory architecture that fundamentally reimagines how AI systems store and update knowledge. By enforcing append-only semantics at the Cognitive layer, Winner-Consensus propagation to the Reflex layer, and automatic sedimentation with contextual overrides to the Core layer, CSW eliminates destructive overwrites, prevents knowledge oscillation, and avoids semantic annihilation. Theoretical analysis proves boundedness, convergence, and causal consistency. Empirical results demonstrate state-of-the-art performance on multi-hop consistency and catastrophic forgetting. CSW offers a principled path toward AI systems that learn and evolve organically, akin to biological cognitive systems.

References

[1] S. Li et al., "From Fact Overwriting to Knowledge Evolution: Causal Editing via On-Policy Self-Distillation," arXiv:2605.28303, 2026.

[2] M. Baser et al., "ThinkEval: Practical Evaluation of Knowledge Preservation and Consistency in LLM Editing with Thought-based Knowledge Graphs," arXiv:2506.01386, 2025.

[3] "Causal Path Alignment: Anchoring the Optimization Trajectory for Controllable In-Parameter Knowledge Editing," arXiv:2506.04042, accepted by IJCAI 2026.

[4] Z. He et al., "Learning Time-Aware Causal Representation for Model Generalization in Evolving Domains," arXiv:2506.17718, 2025.

[5] J. Sun et al., "CEGRL-TKGR: A Causal Enhanced Graph Representation Learning Framework for Temporal Knowledge Graph Reasoning," COLING 2025.

[6] X. Zhang et al., "Self-Evolving World Models for LLM Agent Planning," arXiv:2606.30639, 2026.

[7] Y. Yu and T. Inamura, "Self-Evolving Cognitive Framework via Causal World Modeling for Embodied Scientific Intelligence," arXiv:2606.22449, 2026.

[8] "HCLSM: Hierarchical Causal Latent State Machines for Object-Centric World Modeling," arXiv:2606.xxxxx, 2026.

[9] "Position: Hierarchical World Models with Causal Curation for Generalizing Agents," NeurIPS 2025.

[10] "Curious Causality-Seeking Agents Learn Meta Causal World," NeurIPS 2026.

[11] S. Arora et al., "Linear Algebraic Structure of Word Senses, with Applications to Polysemy," TACL 2018.

Appendix A: Proof of Theorem 5 (Detailed Boundedness)

Given incoming rate r, a node survives only if $T_i(t) > \tau_{prune}$.

$$T_i(t) = \frac{\alpha}{\alpha+\beta} e^{-\lambda (t - t_i)} > \tau_{prune}$$

$$\Rightarrow t - t_i < \frac{1}{\lambda} \log\left(\frac{\alpha}{(\alpha+\beta) \tau_{prune}}\right)$$

Thus the maximum age of any active node is ΔT_{max}. The total active nodes are bounded by the number of updates occurring in the last ΔT_{max} time, which is $r \cdot \Delta T_{max}$, a constant. $\square$

Appendix B: Hyperparameter Settings

Parameter	Symbol	Value
Top-K Activation	K	3
Forgetting rate	λ	0.02
Tension decay rate	μ	0.01
Base sensitivity	θ	0.08
Cooling period	$\tau_{cooling}$	1000
Reference threshold	$\rho_{threshold}$	100
Conflict threshold	γ	0.05
Prune threshold	τ_{prune}	0.01
Beta Prior	α_0, β_0	1.0, 1.0
Embedding Dim	d	384

Appendix C: End-to-End Example

- Input: "User says: Coffee keeps me awake."
- Parser: Extracts (Coffee, causes, Awake).
- CSW: Sprouts node N_Awake in Cognitive layer, parents linked to Core physics.
- Conflict Input: "User says: Decaf keeps me awake."
- CSW: Sprouts new sibling N_Awake_v2 (does NOT overwrite). Reflex layer gates the older signal.
- Query: "Does coffee wake you up?" -> Activates Top-K parents -> Outputs the dominant consensus (the newer, more specific evidence) while retaining old knowledge for historical audit.