Neuro-Semiotic Reasoning: Computational Semiosis for Observational Causal Inference

Martin Trajkow

Technical Founder, o-machine

martin@o-machine.com

https://o-machine.com

Version: 1.3 • Mar 8, 2026

Initial Publication: Ver. 1.0 • Feb 17, 2026

Abstract

Current AI systems excel at statistical pattern matching and formal discovery but fail systematically at causal reasoning from real-world observational signals. While systems like AlphaEvolve demonstrate LLM-guided evolutionary search for synthesized innovation within formal spaces [17, 21], they remain blind to unstructured, hidden signals of physical and commercial reality. We identify this as the *Causal Reasoning Gap*: the inability to derive verified causation from observational signals that have not been textualized or modeled.

We present the **Neuro-Semiotic Reasoning Engine**, an architecture that constructs verified causal understanding through semiotic interpretation of differential relationships across temporal, spatial, conceptual, and absence dimensions. Unlike statistical models operating on tokens, our architecture introduces **Behavioral Signals** as the unit of computation. Drawing on Peirce's unlimited semiosis, we demonstrate that recursive interpretation enables emergent capabilities: shadow mapping of stealth strategies, absence detection as a reasoning primitive, and complete auditability through traceable interpretive pathways.

We formalize the architecture through five pillars (temporal knowledge graphs, concept-mediated ontologies, behavioral modeling, adversarial verification, and evolutionary optimization) and demonstrate that observational causal reasoning requires fundamentally different computational primitives than statistical correlation or formal search.

Keywords: Causal reasoning, Behavioral signals, Neuro-semiotic AI, temporal reasoning, absence detection, observational causal inference.

1. Introduction

Large language models have achieved remarkable fluency in text generation, yet leading AI researchers increasingly question their fundamental architecture as a path to human-level intelligence. As LeCun argues, LLMs lack a grounded model of the world because they are "limited to the discrete world of text" [16]. Sutton identifies deeper structural deficits: they lack ground truth for verification and mechanisms for surprise-driven learning [18].

Furthermore, even the most advanced evolutionary discovery systems, such as DeepMind's AlphaEvolve [17, 21], operate primarily within **formal problem spaces**—mathematical analysis, combinatorics, and code—where the environment is defined by rigid rules and automated evaluators. While these systems can discover novel algorithmic solutions that no human has yet written down, they are essentially performing **synthesized innovation**: exploring permutations of a known, modeled universe.

We identify a different, more fundamental structural limitation: the **Hidden Reality Blind Spot**. Current AI is blind to what has not yet manifested as a document or been modeled into a formal search space. The most valuable causal insights are often those still manifesting as raw, unrecorded behavioral signals. In high-stakes decision environments—investment due diligence, supply chain risk, or geopolitical analysis—statistical correlation and formal search both fail because they lack access to the underlying causal mechanisms operating in real-time.

We define this as the **Causal Reasoning Gap**: the inability to derive verified causation from unstructured, temporal, multi-modal observational signals. This paper presents the **Neuro-Semiotic Reasoning Engine**, an architecture designed to close this gap by implementing computational semiosis – the recursive construction of causal understanding through differential interpretation of **Behavioral Signals**.

2. Background and Related Work

2.1 The Hierarchy of Knowledge Processing

We distinguish four levels of knowledge processing:

Level	Definition	Example	Current AI Capability
Information	Raw facts	"Apple stock is up 3%"	Fully solved
Knowledge	Organized facts	"Tech stocks are rallying"	Largely solved
Understanding	Causal relationships	"Tech stocks rally because AI chip demand surges…"	Reasoning Gap
Wisdom	Actionable insight	"This boom will create rare earth bottlenecks—position accordingly"	Reasoning Gap

Current systems excel at Levels 1-2 but face severe challenges at Levels 3-4. We argue this limitation is *architectural*, not computational.

2.2 Statistical Pattern Matching: Capabilities and Limits

LLMs predict what a person would say about the world, not what will actually happen in the world. They manipulate language patterns without constructing models of underlying reality. As Sutton notes, "you can't have prior knowledge if you don't have ground truth." This produces "hallucination": LLMs optimize for token probability rather than correspondence with reality [7].

Even more fundamentally, LLMs lack mechanisms for surprise-driven learning. They cannot detect when reality deviates from predictions or update their understanding in real-time. This makes temporal reasoning from observational signals—detecting a change in hiring or a failed partnership—structurally impossible.

2.3 Neuro-Symbolic and Discovery Systems

Recent work in neuro-symbolic AI seeks to combine neural learning with symbolic reasoning, but typically requires pre-defined, brittle ontologies [15]. Meanwhile, discovery systems like AlphaEvolve [17, 21] utilize LLMs to guide evolutionary search. While effective for algorithmic discovery, these systems are constrained to formal environments where "correctness" is mathematically verifiable. They cannot "discover" a stealth startup's strategy because the startup's "source code" is not available; it must be inferred from observational signals.

2.4 Semiotic Foundations

We draw on Peirce's theory of unlimited semiosis [11]—the principle that signs generate interpretants which themselves become new signs. Meaning is dynamically constructed through the process of recursive interpretation. In our architecture, a fact gains meaning from its differential relationships across temporal, spatial, conceptual, and absence dimensions.

3. Theoretical Foundation: Causal Understanding Through Behavioral Signals

3.1 A Different Unit of Computation: Actions, Not Tokens

Every reasoning paradigm is defined by its fundamental unit of computation—what it "sees" of the world. - **LLMs see Tokens**: Statistical fragments of text, stripped of time and context. - **Symbolic AI sees Propositions**: Rigid logical statements that are true or false. - **Knowledge Graphs see Triples**: Static snapshots of entity-relation-entity.

o-machine sees Behavioral Signals: observable actions by real-world entities—hiring, partnering, filing, releasing, expanding, contracting—anchored in time, linked to evidence, and meaningful only in relation to their own history and the concurrent behaviors of adjacent actors.

Critically, a Behavioral Signal is **modality-independent**. The signal (e.g., "Company X is expanding manufacturing capacity") may be extracted from text documents, satellite imagery of facility build-outs, sensor data, or financial transactions. The architecture reasons over the *fact of the behavior*, not over the medium that carried the signal. This decoupling allows o-machine to ingest reality regardless of whether that reality has been "textualized."

3.2 The Four Dimensions of Differential Meaning

Causal meaning emerges from how these signals relate to each other: 1. **Temporal Dimension**: Causes precede effects; velocity and acceleration reveal intent. 2. **Spatial Dimension**: Entities connected through multiple independent pathways indicate genuine causality. 3. **Conceptual Dimension**: Facts gain meaning when interpreted within concept phases (e.g., "Maturity" vs. "Pivot"). 4. **Absence Dimension**: The deviation between expected behavior and observed behavior is a first-class signal.

3.3 Unlimited Semiosis: The recursive construction of meaning

The central insight is that **causal meaning emerges through recursive interpretation of differential relationships.** We implement this through a semiotic chain where each interpretation generates a new sign (interpretant) for further analysis.

Example 3.1: The Jony Ive Inference Consider the question: *"What is Jony Ive building for OpenAI?"*—asked before any official announcement.

Step	Context Dimension	Interpretant
1	Atomic Signals	Ive's hiring of former Apple hardware leads; OpenAI's hardware team growth.
2	Spatial/Temporal	Satellite imagery shows facility buildouts at specific contract manufacturers.
3	Conceptual	Apple's concurrent shift away from specific device strategies; Google's ambient computing positioning.
4	Absence	No software-only patents for the new team; absence of traditional cloud-infra hiring.
5	Synthesis	Inference: A specialized AI-integrated wearable device focused on ambient computing.

The answer is not retrieved from a document; it is inferred from the causal logic of *why* these heterogeneous behavioral signals converge.

4. System Architecture

The architecture follows from the theory, utilizing five pillars to implement recursive interpretation.

4.1 Pillar 1: Temporal Knowledge Graph

We use a fact-first architecture where the fundamental unit is the timestamped Behavioral Signal with full provenance. This enables velocity and acceleration analysis (e.g., a hiring spike followed by deceleration indicates a phase transition) and domain-specific absence detection.

4.2 Pillar 2: Concept-Mediated Ontology

Entities connect through concepts that evolve. Adding a new signal to "Autonomous Driving" recontextualizes the concept itself, which recursively recontextualizes all connected entities. Meaning is constructed through use, not encoded statically.

4.3 Pillar 3: The Proof Engine (Adversarial Verification)

Every inference undergoes systematic falsification through a two-agent "Socratic Protocol." - **The Generator** proposes causal connections based on pattern matches. - **The Critic** searches for counter-evidence and contradictory signals across the four differential dimensions.

We reframe this not as simple validation, but as a **Proof Engine**. In high-stakes private markets, alpha is derived from having a thesis that survives extreme skepticism. The Proof Engine ensures that an insight only enters the graph if it is the most *defensible* interpretation of the signals.

This creates **Verified Causal understanding**, transforming the system from an "opinionated generator" into a "rigorous arbiter of truth."

4.4 Pillar 4: Behavioral Modeling

Absence detection requires expectation. We maintain behavioral baselines for entities within specific concept phases. Deviations from these baselines (e.g., a market leader stopping hiring in its core competency) are treated as diagnostic signals of maturity or pivot.

4.5 Pillar 5: Evolutionary Optimization

The system's reasoning strategies—including detection patterns and verification thresholds—evolve through selection pressure. Guided by LLM mutations (which leverage world knowledge), the system autonomously improves its reasoning procedures based on outcome validation. This mirrors recent successes in "test-time adaptation" seen in AlphaEvolve and ARC-Prize-winning systems [14, 17, 21].

5. Emergent Capabilities

5.1 Stealth Discovery (Shadow Mapping)

Inference of hidden strategic states from weak, cross-domain signals (e.g., tracking "talent clusters" to find new companies 6-12 months before incorporation).

5.2 The "Speed-to-No" (Absence Detection)

Generating meaning from the non-occurrence of expected events to rapidly falsify claims. If a startup claims "proprietary hardware breakthroughs" but exhibits zero evidence of specialized hiring or component procurement, the system generates a high-confidence "No" signal.

5.3 Thesis Defense (Multi-Hop Causal Chains)

Protecting $10M+ investment theses by tracing multi-hop dependencies and detecting "stealth killers"—competitors or technological shifts that render a moat obsolete months before news reaches the mainstream.

5.4 Complete Auditability

Every conclusion is backed by a traceable interpretative chain leading back to atomic, evidence-linked Behavioral Signals.

6. Evaluation Framework

6.1 Evaluation Philosophy

Evaluating causal reasoning systems presents challenges absent from standard information retrieval benchmarks. Precision and recall measure fact retrieval; they do not measure the quality of *causal inference*, the detection of *absence*, or the depth of *cross-domain reasoning*.

We propose a multi-tier benchmark framework with 15 scenarios organized by reasoning complexity, drawn from the domain of private market intelligence. Each scenario specifies an archetype user, test queries, and success criteria.

6.2 Benchmark Scenarios

Tier 1: Core Reasoning (Reasoning Gap Closure) **Scenario 1: Negative Evidence Search** - *Task*: Given claims from a pitch deck, find contradicting or disconfirming Behavioral Signals. - *Capabilities tested*: The Proof Engine, evidence chain integrity.

Scenario 2: Stealth Entity Detection - *Task*: Identify stealth companies through talent clustering patterns. - *Capabilities tested*: Shadow mapping, concept-mediated discovery, temporal velocity.

Scenario 3: Inferred Financial Health - *Task*: Estimate a company's remaining runway from Behavioral Signals (hiring, facility expansion, patenting). - *Capabilities tested*: Temporal reasoning, behavioral inference, adversarial validation.

Tier 2: Cross-Domain Inference **Scenario 4: Supply Chain Dependency Mapping** - *Task*: Trace a 4-hop supply chain path from Company → Component → Material → Country of Origin. - *Capabilities tested*: Multi-hop causal chains, concept-mediated traversal.

Scenario 5: Moat Verification - *Task*: Identify stealth threats to a company's competitive moat. - *Capabilities tested*: Shadow mapping, concept-mediated discovery, absence detection.

Scenario 6: Talent Attrition Alert - *Task*: Detect clustering of senior departures from a portfolio company. - *Capabilities tested*: Temporal velocity, behavioral modeling, absence detection.

Tier 3: Strategic Reasoning **Scenario 7: Concept Phase Transition Detection** - *Task*: Determine whether an industry is transitioning from Technology Development to Regulatory Approval phase. - *Capabilities tested*: Concept-mediated ontology, temporal pattern recognition, concept phase inference.

Scenario 8: R&D Reality Check - *Task*: Validate a technology bet against ground-truth market activity. - *Capabilities tested*: Cross-domain temporal analysis, behavioral momentum comparison.

Scenario 9: Exit Path Validation - *Task*: Identify credible acquirers for a specific technology, with historical M&A precedent and strategic rationale. - *Capabilities tested*: Concept-mediated discovery, pattern matching.

Tier 4: Absence and Counterfactual Reasoning **Scenario 10: Phase Transition Detection via Absence** - *Task*: Detect that a market leader has stopped hiring in its core competency, inferring technology maturity. - *Capabilities tested*: Absence detection, temporal baseline behavior.

Scenario 11: Partnership Absence Analysis - *Task*: Detect the non-occurrence of an expected partnership and reason about strategic implications. - *Capabilities tested*: Counterfactual reasoning, absence signal generation.

Scenario 12: Blind Spot Monitor - *Task*: Detect stealth startups building cheaper alternatives to an incumbent's core product. - *Capabilities tested*: Shadow mapping, concept-mediated discovery, multi-hop inference.

Tier 5: System Integrity **Scenario 13: High-Fidelity Entity Verification** - *Task*: Return verifiably accurate signals about well-known entities. - *Capabilities tested*: Evidence integrity, fact verification.

Scenario 14: Failure Mode Avoidance - *Task*: System must avoid generic overviews, outdated data, and inappropriate refusals. - *Capabilities tested*: Temporal precision, query understanding.

Scenario 15: Temporal Precision Validation - *Task*: Return signals with correct temporal context, distinguishing current from historical states. - *Capabilities tested*: Temporal reasoning, fact-level timestamps.

6.3 Evaluation Metrics

For each scenario, we propose evaluation on three axes: 1. **Reasoning Depth**: Number of semiotic interpretation layers achieved. 2. **Evidence Integrity**: Percentage of interpretants backed by traceable evidence chains. 3. **Absence Sensitivity**: Ratio of detectable absence signals correctly identified.

6.4 Implementation Status

The neuro-semiotic reasoning architecture is under active development. A closed beta release is planned for Q2 2026 with design partners from private equity and corporate development.

6.5 Illustrative Examples

Example 1: Concept Phase Transition Detection (The Speed-to-No) **Query**: “Is the autonomous driving industry transitioning from Technology Development to Regulatory Approval phase?”

Reasoning Process: 1. **Behavioral Baseline (Pillar 4)**: High R&D growth (15-25%), frequent prototype unveilings. 2. **Temporal Deviation (Pillar 1)**: Current signals shows R&D growth plateauing ($<5\%$), hiring shifting to safety validation experts. 3. **Absence Analysis (Pillar 4)**: 18-month absence of new sensor breakthroughs (historically 3-4/year). 4. **The Proof Engine (Pillar 3)**: Generator proposes phase shift; Critic challenges if this is merely a funding winter (Counter-evidence: cash reserves are still high). 5. **Inference**: Industry-wide phase transition verified.

Example 2: Counterfactual Reasoning for Bottleneck Identification **Query**: “Company X announced ‘technology readiness’ for AV deployment but has not deployed. Why?”

Reasoning Process: 1. **Expected Causal Chain**: Tech Ready → Regulatory Approval → Fleet Deployment. 2. **Observed Reality**: Tech Ready (t) → No Deployment (t + 18 months). 3. **Counterfactual Analysis**: Regulatory applications filed but not granted; insurance partnerships expected but absent. 4. **Identification**: Absence of downstream enablers (insurance, manufacturing startups) verifies the bottleneck is regulatory, not technical.

7. Comparative Analysis: The Paradigm Shift

The transition to observational causal reasoning is a shift from content-based frequency to relationship-based structure—analogous to the shift from pre-Google search to **PageRank**.

Dimension	Statistical AI / LLMs	Neuro-Semiotic Reasoning
Unit of Computation	Tokens (Textual fragments)	Behavioral Signals (Modality-agnostic)

Dimension	Statistical AI / LLMs	Neuro-Semiotic Reasoning
Foundation	Correlation from frequency	Causation from differential interpretation
Logic	Probabilistic Generation	Verified Proof Chains (Proof Engine)
Discovery	Synthesized Innovation (Formal)	Observational Inference (Real-world)

8. Conclusion

The Causal Reasoning Gap is an architectural gap. Closing it requires moving beyond the "discrete world of text" to a system that sees the world through modality-independent Behavioral Signals. By implementing computational semiosis through a fact-first temporal graph and a rigorous Proof Engine, o-machine enables the transition from imitating human language to reasoning about hidden real-world reality.

We believe this work represents a foundational step toward AI systems that do not merely retrieve information, but construct verified causal understanding from real-world observational signals. The architecture described here has been implemented and is under active development. Comprehensive empirical validation across the proposed benchmark scenarios will be reported following our closed beta launch in Q2 2026. We welcome collaboration with researchers and domain experts to extend this framework and evaluation methodology to additional high-stakes decision environments. A companion theoretical work, "AI Needs Goals," exploring how intelligent systems might autonomously construct operational objectives from human-centric purposes rather than optimizing arbitrary metrics, is in preparation.

References

[1] Brown, T. B., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877-1901.

[2] Garcez, A. d'A., et al. (2019). Neural-symbolic computing: An effective methodology for principled integration of machine learning and reasoning. *arXiv preprint arXiv:1905.06088*.

[3] Goodfellow, I. J., et al. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27.

[4] Hamilton, W. L. (2020). *Graph Representation Learning*. Morgan & Claypool Publishers.

[5] Hogan, A., et al. (2021). Knowledge graphs. *ACM Computing Surveys*, 54(4), 1-37.

[6] Irving, G., Christiano, P., & Amodei, D. (2018). AI safety via debate. *arXiv preprint arXiv:1805.00899*.

[7] Ji, Z., et al. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1-38.

[8] Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

[9] Madry, A., et al. (2018). Towards deep learning models resistant to adversarial attacks. *International Conference on Learning Representations*.

[10] Pearl, J. (2009). *Causality: Models, Reasoning, and Inference.* Cambridge University Press.

[11] Peirce, C. S. (1931-1958). *Collected Papers of Charles Sanders Peirce*, Vols. 1-8. Harvard University Press.

[12] Touvron, H., et al. (2023). LLaMA: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971.*

[13] He, H., & Thinking Machines Lab. (2025). Defeating nondeterminism in LLM inference. *Thinking Machines Lab: Connectionism.* https://thinkingmachines.ai/blog/defeating-nondeterminism-in-llm-inference/

[14] Kamradt, G., & Chollet, F. (2025). ARC Prize 2025 results and analysis. *ARC Prize.* https://arcprize.org/blog/arc-prize-2025-results-analysis

[15] Mao, H., Chen, Z., Tang, W., Zhao, J., Ma, Y., Zhao, T., Shah, N., Galkin, M., & Tang, J. (2024). Position: Graph foundation models are already here. *arXiv preprint arXiv:2402.02216.*

[16] Reese, H. (2026). Yann LeCun's new venture is a contrarian bet against large language models. *MIT Technology Review.* https://www.technologyreview.com/2026/01/22/1131661/yann-lecuns-new-venture-ami-labs/

[17] Novikov, A., Vũ, N., Eisenberger, M., Dupont, E., Huang, P.-S., Wagner, A. Z., Shirobokov, S., Kozlovskii, B., Ruiz, F. J. R., Mehrabian, A., Kumar, M. P., See, A., Chaudhuri, S., Holland, G., Davies, A., Nowozin, S., Kohli, P., & Balog, M. (2024). AlphaEvolve: A coding agent for scientific and algorithmic discovery. *arXiv preprint arXiv:2506.13131.*

[18] Sutton, R. S. (2019). The Bitter Lesson. *Incomplete Ideas.* http://www.incompleteideas.net/IncIdeas/BitterLe

[19] Darlow, L., Crowley, E. J., Storkey, A., & Hospedales, T. (2025). Continuous Thought Machines: Adaptive test-time computation for abstract reasoning. *arXiv preprint arXiv:2505.05522.*

[20] Silver, D., & Sutton, R. S. (2025). Welcome to the Era of Experience. In *Designing an Intelligence* (forthcoming). MIT Press. Available at: https://storage.googleapis.com/deepmind-media/Era-of-Experience%20/The%20Era%20of%20Experience%20Paper.pdf

[21] Georgiev, B., Gómez-Serrano, J., Tao, T., & Wagner, A. Z. (2025). Mathematical exploration and discovery at scale. *arXiv preprint arXiv:2511.02864.* https://doi.org/10.48550/arXiv.2511.02864

Appendix A: Formal Definitions Summary

Symbol	Definition
$f = (s, p, o, t, E)$	Atomic Behavioral Signal (subject, predicate, object, timestamp, evidence set)
$M(f, C)$	Meaning function: signal × context → interpretant
$\Delta_T(f, t_i, t_j)$	Temporal differential of signal f between times t_i and t_j
$\Delta_S(f)$	Spatial differential: relational context set of signal f

Symbol	Definition
c	Concept: higher-order abstraction mediating entity relationships
$\phi(c, t)$	Phase of concept c at time t
$a = (f_{expected}, t, \epsilon)$	Absence signal: expected behavior not observed within tolerance
c_g	Generator confidence score
σ	Convergence verification across four dimensions (temporal, spatial, conceptual, absence)
s_d	Critic counter-evidence strength across dimensions
$c_{final} = c_g \cdot \sigma - \max(s_d)$	Committal Gate confidence with dimensional verification
$\theta_{high}, \theta_{low}$	Empirically calibrated decision thresholds for committal gate
$S = \{i_0, i_1, ..., i_n\}$	Semiotic chain: sequence of recursive interpretants

Appendix B: Domain-Specific Ontologies

Domain-specific ontologies and behavioral pattern libraries have been developed for multiple high-stakes decision environments including autonomous driving, semiconductor supply chains, and energy infrastructure. These ontologies include:

- **Concept Types**: Technology categories, business models, strategic approaches, application domains, and risk factors
- **Behavioral Primitives**: Observable signals including partnerships, product releases, testing activity, leadership changes, technology development, deployment milestones, and workforce dynamics
- **Relationship Types**: Causal dependencies (ENABLES, REQUIRES, BLOCKS), competitive dynamics (COMPETES_WITH), behavioral signals (EXHIBITS, VALIDATES), and risk exposures (EXPOSES_TO)

The ontologies are designed to be extensible and transferable across domains while maintaining domain-specific precision. Each concept and relationship type includes temporal metadata, evidence requirements, and validation criteria.