Neuro-Semiotic Reasoning: Algebraic Semiosis for Observational Causal Inference

Martin Trajkow | Lead Researcher, o-machine

Version 2.0 — May 2026

The Causal Reasoning Gap

Current AI systems—including advanced Large Language Models and formal discovery engines like AlphaEvolve—are structurally blind to reality that hasn't been codified into text or mathematical environments. They excel at statistical pattern matching but fail systematically at deriving causal insight from unstructured, real-world observational signals. We define this as the Causal Reasoning Gap.

The Solution: Algebraic Semiosis

We introduce the Neuro-Semiotic Reasoning Engine, a fundamentally different architecture that reasons over Behavioral Signals—modality-independent facts like a company hiring an engineer, filing a patent, or signing a lease—rather than textual tokens.

Drawing on Peircean semiotic theory, the engine constructs meaning by analyzing how these signals relate across four dimensions: Time, Space, Concept, and Absence.

The core breakthrough of this work is Algebraic Semiosis: the discovery that complex, multi-hop causal reasoning can be computed as constant-time algebraic operations over high-dimensional vectors, rather than through computationally expensive, LLM-based iterative generation—effectively establishing a zero-LLM query understanding and reasoning substrate.

Practical Application: Inferring Unannounced Strategy

If asked, “What kind of hardware is Jony Ive building for OpenAI?”, standard AI searches for text matching those keywords. Our engine computes the “geometric shape” of the unannounced product:

1. Unobserved Inference: The query mathematically resonates with Ive's firm (LoveFrom), pulling it into context without it being explicitly named.
2. Signal Convergence: The system identifies a localized supply chain forming around micro-optics and edge-silicon.
3. Behavioral Absence: The system calculates a mathematical absence of cloud infrastructure, data centers, or traditional screens.

Inferred Hypothesis: The intersection of these behaviors logically forces the conclusion of a screenless, AI-integrated wearable device focused on ambient computing. The engine finds the product by calculating the geometric "hole" it leaves in the supply chain.

Key Empirical Findings

Results validated through systematic ablation across 14 experimental configurations and 109 complex causal queries, with independent evaluation by Gemini 3.1 Flash Lite as judge model:

- 140× Faster: Algebraic query understanding processes in ~50ms compared to ~7 seconds for LLMs.
- Higher Quality: The algebraic system achieved a 4.4/5.0 final answer score on complex causal reasoning tasks, beating the LLM baseline of 4.1 while producing zero context-exhaustion errors.
- The Evidence Text Paradox: Stripping raw text out of the pipeline and providing only structural facts actually improves the AI's causal narration (scoring 4.6/5.0). Verbose text distracts generative models from clean structural logic.
- Adversarial Falsification: The system's ability to mechanically disprove false claims based on behavioral absence proved robust, scoring a perfect 5.0/5.0 across all dimensions on 14 out of 24 adversarial tests.

Conclusion

True causal reasoning requires moving beyond the discrete world of text. By enforcing semantic atomicity and collapsing recursive interpretation into algebraic resonance, this architecture provides a computationally tractable framework for inferring ground truth from the chaos of observational reality.

For formal definitions, mathematical formalisms, and the complete empirical dataset, refer to the full technical paper: Neuro-Semiotic Reasoning: Algebraic Semiosis for Observational Causal Inference (Version 2.0).

Contact: martin@o-machine.com | LinkedIn | https://o-machine.com