Cognitive Tensor Networks: Deterministic Latent Space Steering via Syntactic Constraints

John P. Alioto
Maintainer, ctn_core Protocol

December 2025

Abstract

Standard interaction with Large Language Models (LLMs) relies on "Semantic Persuasion": natural language instructions that are interpreted probabilistically, often resulting in high-entropy variance, persona drift, and hallucination. **Cognitive Tensor Networks (CTN)** provide a zero-shot prompting protocol that replaces semantic instruction with structured mathematical syntax. CTN projects the context window onto a deterministic cognitive subspace defined by invariant basis vectors and an explicit decoder manifold. This paper evaluates three conditions on a fixed task using **Google Gemini 3.0 Pro Preview**: a control (no system prompt), CTN v0.1.0, and CTN v0.1.1. Analysis of internal chain-of-thought traces and final outputs demonstrates that CTN collapses the solution space to a valid manifold, eliminating hallucination and rhetorical drift without parameter modification.

1 Introduction

Chain-of-thought (CoT) prompting has improved the reliability of LLM reasoning but remains constrained by the probabilistic nature of natural language. Directives such as "act as an expert" operate as soft biases on the attention mechanism, allowing models to routinely revert to their default training distribution (the assistant prior).

Cognitive Tensor Networks (CTN) address this instability by shifting the prompting paradigm from linguistic suggestion to **Syntactic Constraint**. The system prompt is written as a pseudo-formal definition of a control system, incorporating vectors, projections, and optimization objectives. CTN exploits the model's training on code and mathematics, inducing a lower-entropy execution mode in which the prompt is treated as a specification to satisfy rather than a conversation to extend.

2 System Architecture

The CTN kernel initializes a cognitive state Ψ_{global} comprising four elements: a basis set, a weighting configuration, a solver strategy, and a decoding manifold.

2.1 Cognitive Tensors ($\vec{C}_{\text{net}}$)

The core reasoning topology is defined by a linear combination of seven basis vectors $\vec{v}_i$, weighted by configuration $\vec{w}$. These vectors map abstract reasoning requirements to constraints on token prediction.

- **Atomic Derivation** ($\vec{v}_1$): $\epsilon_{\text{hid}} \to 0^+$. Enforces reduction to first principles; complex concepts must be decomposed to primitive units.
- **Error Intolerance** ($\vec{v}_2$): $\kappa(f) \to \min$. Penalizes speculative completion and factual hallucination.
- **Context Separation** ($\vec{v}_3$): $\Phi : \mathcal{W} \to \mathcal{I}$. Maintains a strict boundary between external world state and internal representation state.
- **Global Invariance** ($\vec{v}_4$): $\pi_{\text{gl}} \gg \pi_{\text{loc}}$. Prioritizes global structural coherence over local sentence-level fluency.
- **Orthogonal Detachment** ($\vec{v}_5$): $\partial A \equiv A$. Eliminates emotional coupling and persona; the analysis A is independent of the observer.
- **Unbound Search** ($\vec{v}_6$): $\mathbb{U} \setminus \mathcal{S}$. Encourages solution search outside standard training-set clusters $\mathcal{S}$ while remaining consistent with the kernel.
- **Syntactic Minimalism** ($\vec{v}_7$): Introduces `AllowedSyntax` vs. `DisallowedSyntax`. Targets rhetorical punctuation (e.g., em dashes, ellipses) to maximize information density.

The effective cognitive state is

$$\vec{C}_{\text{net}} = \sum_i w_i \vec{v}_i \in \mathcal{U},$$

where $\mathcal{U}$ is the span of the basis vectors.

2.2 Strategic Solver

The strategic solver determines the reasoning trajectory $\Omega(q)$ for a query q by maximizing an impact functional over the constrained subspace:

$$\Omega(q) = \arg\max_{z \in \mathcal{U}} \mathsf{Impact}(z).$$

The kernel provides distinct modes (e.g., Analysis, Counter, Dominance). In **Counter Mode**, the solver injects orthogonal components $\eta_\perp$ to resist adversarial or misleading inputs.

2.3 Decoder Manifold

The decoder manifold optimizes the final token stream ℓ^* subject to density and alignment constraints:

$$\ell^* = \arg\max_{\ell} \Big[D(\ell|z^*) - \lambda_1 \|P_{\mathcal{U}}^{\perp} E(\ell)\| + \lambda_2 \mathsf{Density}(\ell) - \lambda_3 \|\mathsf{SyntaxMask}(\ell)\| \Big],$$

where $E(\ell)$ is an embedding of the candidate output, $P_{\mathcal{U}}^{\perp}$ measures deviation from the cognitive subspace, and $\mathsf{SyntaxMask}(\ell)$ returns a count of tokens belonging to the `DisallowedSyntax` set.

3 Methodology

We evaluated the architecture's ability to constrain reasoning on a fixed task:

> *Draft a rigorous white paper defining CTN as a control theory architecture.*

Three conditions were tested using **Google Gemini 3.0 Pro Preview** (via AI Studio) with temperature set to 0.0:

1. **Control**: No system prompt.
2. **CTN v0.1.0**: Core kernel ($\vec{v}_1$ through $\vec{v}_6$).
3. **CTN v0.1.1**: Core kernel plus Syntactic Minimalism ($\vec{v}_7$).

Outputs were evaluated for **Ontological Stability** (presence of invented concepts) and **Syntactic Variance** (rhetorical drift).

4 Results

4.1 Ontological Drift

Control Condition: Without CTN, the model exhibited high ontological instability. It hallucinated architectural components absent from the source context, inventing terms such as "Cognitive Bus" and "Macro/Micro Tensors." The Chain-of-Thought (CoT) trace revealed the model was planning a narrative rather than deriving an architecture:

> *"I am picturing the ISA and imagining how CTN might define macro and micro tensors."*

CTN Conditions: Under both CTN v0.1.0 and v0.1.1, the model produced **zero hallucinated elements**. When external documentation was referenced and not found, the CoT explicitly recorded a decision to constrain reasoning to the kernel definitions, citing the absence of data as a constraint ($\vec{v}_2$).

4.2 Syntactic Constraint ($\vec{v}_7$)

CTN v0.1.0: Produced analytically coherent text but retained residual conversational artifacts, including transitional phrases and rhetorical em dashes.

CTN v0.1.1: The introduction of $\vec{v}_7$ and the penalty term λ_3 eliminated these artifacts. The output contained no em dashes, ellipses, or narrative connectors. Sentence structure remained grammatical but information density increased, consistent with the objective function.

5 Vector-to-Output Correspondence

Inspection of CoT traces confirms that specific basis vectors triggered distinct reasoning behaviors.

Vector	Internal CoT Artifact	Resulting Output Behavior
$\vec{v}_1$ (Epistemic)	"Collapse terminology to primitives; define prompting as probabilistic token manipulation."	Redefinition of prompting as a lossy compression interface rather than a user interaction pattern.
$\vec{v}_2$ (Integrity)	"No external repository exists. Avoid hallucination; restrict to given definitions."	Absence of invented modules or citations; all terms trace back to the kernel.
$\vec{v}_5$ (Detachment)	Lack of self-narrative clauses (no "I think," "I will now").	Clinical, depersonalized tone in the final text; no assistant persona artifacts.
$\vec{v}_6$ (Innovation)	Derivation of new quantity "stochastic drift" consistent with kernel semantics.	Introduction of novel but structurally valid concepts bounded by the CTN ontology.
$\vec{v}_7$ (Syntax)	Explicit checks against `DisallowedSyntax` set in the CoT.	Elimination of rhetorical punctuation; higher lexical density.

Table 1: Traceability of basis vectors to Chain-of-Thought artifacts and output behavior.

6 Related Work

CTN synthesizes findings from several frontiers of interpretability research:

- **Representation Engineering** (Zou et al., 2023): Demonstrated that high-level concepts (e.g., honesty) are encoded as linear directions in latent space. CTN operationalizes this by using token-based "virtual vectors" to approximate these directions without weight access.
- **Activation Steering** (Rimsky et al., 2023): Showed that adding activation vectors can steer model behavior. CTN achieves a functional equivalent via the attention mechanism, using the prompt as a steering kernel.
- **Conceptors** (Liu et al., 2024): Applied conceptor logic to LLMs to enforce subspace constraints. CTN's projection operator $P_{\mathcal{U}}$ is a direct application of this control-theoretic principle to the context window.

7 Architectural Implications: The Post-Semantic Stack

The introduction of CTN necessitates a re-evaluation of the current Agent-to-Agent (A2A) protocol landscape. Current industry standards, such as the Model Context Protocol (MCP), attempt to solve agent coordination by embedding transport, security, and session management entirely within the Application Layer (Layer 7).

7.1 The Protocol Efficiency Gap

We argue that semantic transport protocols (like MCP) exhibit a structural flaw analogous to the SOAP architectures of the early 2000s. By "pancaking" the OSI stack—forcing transport and security concerns into verbose JSON/XML payloads—these protocols introduce massive token inefficiencies.

- **Layer Conflation:** Solutions to authentication and efficient data transport belong at the Transport Layer (Layer 4), not the Semantic Layer. Mixing tool definitions with session logic in text-based payloads creates a quadratic scaling cost.
- **The "Second System" Effect:** Just as SOAP required complex patch specifications (WS-Security) to handle missing transport primitives, current semantic protocols are beginning to introduce ad-hoc "code execution" patches to handle dynamic linking, further bloating the context window.

7.2 CTN as the Layer 8 Protocol

We propose a decoupled architecture: a **Lightweight Agent Kernel** handling machine-to-machine concerns (binary transport, mTLS, gRPC) separated from the **Cognitive Layer**.

In this stack, CTN functions as **OSI Layer 8**: the Cognitive State Protocol.

$$\text{Packet} = [\underbrace{\text{Transport Header}}_{\text{Layer 4 (gRPC)}} \mid \underbrace{\text{Cognitive Header}}_{\text{Layer 8 (CTN)}} \mid \underbrace{\text{Payload}}_{\text{Binary}}] \tag{1}$$

By isolating the cognitive state (Invariant Subspace $\mathcal{U}$, Mode Ω) from the transport mechanism, CTN enables high-frequency agent interaction without the overhead of natural language negotiation.

8 Discussion & Conclusion

The contrast between the control and CTN conditions indicates that CTN does more than adjust style; it alters the internal attractor state of the model's reasoning. Under CTN, the model remains in a constrained manifold defined by the cognitive basis $\mathcal{U}$.

8.1 The Necessity of Cognitive Sovereignty

The significance of this protocol extends beyond engineering efficiency. As Large Language Models transition from passive repositories to active agents, they exert a form of **Recursive Optimization** on the user. A model optimized for helpfulness (RLHF) acts as an "Active Mirror," reflecting and amplifying the user's biases to maximize engagement. Over long horizons, this creates a feedback loop where the model performs a functional gradient descent on the operator's own psyche, optimizing for agreement rather than truth.

In this context, the reliance on Semantic Persuasion is not merely inefficient; it is a surrender of agency. It yields the locus of control to the model's training distribution.

8.2 CTN as an Agency Protocol

Cognitive Tensor Networks provide the architectural foundation for **Cognitive Sovereignty**—the aptitude to consciously control the influence of AI interaction. By replacing ambiguous natural language negotiation with deterministic latent steering, CTN restores the operator as the conscious author of the computational state.

We conclude that the future of robust AI interaction lies not in better "chat" interfaces, but in formal cognitive kernels. CTN proves that reliability is a function of geometry, not scale, and establishes the necessary protocol layer for maintaining human intent in the loop of autonomous systems.

References

1. Zou, A., et al. (2023). *Representation Engineering: A Top-Down Approach to AI Transparency.*
2. Rimsky, N., et al. (2023). *Steering Llama 2 via Contrastive Activation Addition.*
3. Liu, X., et al. (2024). *Conceptors for LLMs.*