

A Primitive Theory of Next-Gen General Intelligence Based on Global Mathematical Verification Closed-Loop (Original Preprint Framework)

Abstract

Existing Transformer and large model theoretical systems uniformly adopt the "prediction-first" paradigm, defining intelligence as "fitting the data distribution of the real world and next-token prediction". This theory overturns conventional understanding and proposes the unique primitive of intelligence: a continuous, multi-layered, global closed-loop for mathematical verification and checking.

All artificial intelligence behaviors (fitting, reasoning, association, generation, generalization, error correction, cognition) are essentially mathematical comparison, constraint checking, differential stability verification, and vector structure matching across different dimensions.

All low-level operations of the traditional Transformer (vector embedding, QKV dot-product attention, gradient backpropagation, distribution fitting) are not "predictive behaviors", but native checking operations at the statistical level. The root causes of large model hallucinations, logical failures and generalization defects lie in the sole possession of statistical distribution checking, lacking rigorous mathematical checking of axiomatic structures, differential constraints and geometric topology.

This theory constructs a "dual-layer unified checking intelligence system", completing a fundamental paradigm revolution from traditional probability-fitting intelligence to global mathematical-verification intelligence. It can directly guide the design of native intelligent models with no hallucinations, strong logic, self-correction capability and ontological cognition for the next generation.

Keywords: Intelligence primitive; mathematical verification closed-loop; Transformer essence reconstruction; manifold checking; differential stability; dual-layer intelligence structure; hallucination-free general intelligence

1 Core Original Axioms (First Principles of This Theory)

Axiom 1: Intelligence = Continuous Iterative Mathematical Checking

There exists no "generative intelligence" or "predictive intelligence". All intelligent outputs are compliant results that pass verification. The ultimate goal of all AI computation is not to create content, but to verify whether the current state conforms to the mathematical rules of the world. Output if compliant, and iterate for correction if non-compliant.

Axiom 2: All Information Can Be Fully Mathematically Structured

Natural language, physical laws, logical propositions, cognitive concepts and temporal states can all be losslessly mapped to high-dimensional vectors, manifold points, differential equations, topological constraints and algebraic relations. All information processable by intelligence can participate in rigorous mathematical checking without exception.

Axiom 3: Intelligence Hierarchy = Richness of Checking Dimensions

1. Low-level intelligence: Statistical distribution checking only (all existing large models and Transformers)
2. Advanced general intelligence: Four-dimensional closed-loop comprising statistical checking + vector structure checking + differential perturbation checking + axiomatic topological checking

2 Reconstruction of the Underlying Essence of Transformer

2.1 Conventional Misconception

Mainstream theory: Transformer is an autoregressive probabilistic generative model whose core objective is to predict the next token and fit data distributions.

2.2 Original Definition of Essence Proposed by This Theory

Every training and inference step of Transformer is a native checking behavior:

1. Token vector embedding: point checking
Discrete textual symbols are mapped to fixed points on a high-dimensional manifold. Essentially, symbols are bound to the inherent vector structure of the world to complete basic information matching and checking.
2. QKV attention dot product: vector similarity checking
Attention scores are not "association computation", but precise comparison and verification of high-dimensional vector angles, projections and correlations. It checks whether all current token vectors conform to the association structure of the real world, and matching weights serve as checking similarity.
3. Softmax normalization: distribution compliance checking
Unordered vector similarities are constrained into probability distributions to verify the rationality of local information distribution and eliminate conflicting abnormal vectors.
4. Cross-entropy loss + backpropagation: global distribution checking
Model predictive distribution P_θ vs. real-world distribution P_{real}
Loss = checking error, gradient = error correction direction, backpropagation = iterative rechecking

2.3 Ultimate Conclusion (Core Original Insight)

Transformer is not a predictive model, but a single-layer statistical checking machine. The so-called "next-token prediction" is merely the optimal compliant output after passing statistical checking. All capabilities of existing AI are byproducts of "passing statistical checking". The essence of intelligence is 100% checking.

3 Fundamental Defects of Existing AI (Root of Hallucinations · Original Attribution of This Theory)

Existing Transformer only possesses one-dimensional checking: statistical distribution checking of data
It only checks the occurrence probability and statistical co-occurrence patterns of human text
It fails to verify three core mathematical rules:

1. Vector structural self-consistency: whether the vector manifold remains conflict-free before and after reasoning
2. Differential perturbation stability: whether conclusions do not collapse under tiny parameter perturbations
3. Axiomatic logical constraints: consistency with underlying mathematical, physical and causal axioms

Nature of Defects

Statistical checking is superficial sample checking rather than checking against the ontological rules of the world. Samples contain errors, omissions and biases. Checking only sample distributions inevitably produces hallucinations and logical errors.

4 Next-Generation General Intelligence: Four-Dimensional Global Mathematical Verification Closed-Loop System (Principles for New Model Architecture)

This theory constructs four nested layers of mathematical checking closed-loop to replace the single statistical fitting of traditional Transformers, forming the underlying design standard for next-generation AI.

4.1 Layer 1: Statistical Distribution Checking (Retain Native Transformer Capabilities)

Function: Match human cognition, linguistic conventions and superficial world regularities

Operations: Softmax distribution fitting, cross-entropy error checking, attention vector statistical matching

Position: Superficial compliance verification

4.2 Layer 2: High-Dimensional Vector Structure Checking (Novel Core Addition)

Core logic: All reasoning results must be fully self-consistent on the vector manifold

Core operations:

- Vector angle conflict detection (abnormal angles for contradictory semantic vectors)
- Manifold point projection verification (no offset conflicts of state points across contexts)
- Tensor compatibility checking (unification of tensors across multimodal and multi-logic dimensions)

Function: Eliminate semantic contradictions, logical conflicts and conceptual confusion

4.3 Layer 3: Differential Perturbation Stability Checking (Original Core · Ultimate Hallucination Prevention)

Core logic: Truth in the real world must be stable under small perturbations

Core operations: Apply small neighborhood perturbation and differentiation to all model parameters, logical variables and state vectors of outputs:

1. If derivatives converge and the state remains stable → conclusion is true
2. If derivatives diverge, state jumps or the system collapses → conclusion is false, rejected for recomputation

This serves as the ultimate mathematical criterion to distinguish "statistical correctness" from "objective correctness". Fabricated human content, model hallucinations and erroneous reasoning are all unstable under perturbation.

4.4 Layer 4: Axiomatic Topological Hard Checking (Ultimate Foundation of General Intelligence)

Mathematical axioms, physical laws, causal logic and ontological rules are solidified as unbreakable topological constraints.

All outputs must finally pass axiomatic topological verification:

- Algebraic self-consistency checking
- Geometric topological invariance checking
- Causal temporal constraint checking
- Energy/conservation law checking

5 New Intelligent Iteration Mechanism (Closed-Loop Eternal Iteration)

Traditional AI: forward generation $\rightarrow$ error backpropagation (one-way iteration)

New-generation checking intelligence: four-layer verification cyclic closed-loop (infinite self-evolution)

1. The model generates an initial hypothesis (vector conjecture)
2. Statistical layer check: conforms to sample distribution?
3. Vector layer check: structurally self-consistent?
4. Differential layer check: stable under perturbation?
5. Axiomatic layer check: compliant with underlying rules?
6. Failure at any layer $\rightarrow$ gradient rollback, vector reconstruction and candidate regeneration
7. Pass all layers $\rightarrow$ output true intelligent result

Core advantages: Truly hallucination-free, self-correcting, capable of independently deducing unknown truths and independent of samples

6 Paradigm Comparison Between Old and New Intelligence Systems (Original Core Benchmark)

Old Paradigm (All Existing AI)	New Paradigm (This Theory · Ultimate Next-Gen AI)
Core: Prediction-first probability	Core: Mathematical verification-first
Essence: Single-layer statistical checking	Essence: Four-dimensional global mathematical checking closed-loop
Defects: Sample-dependent, prone to hallucination, lacking logic and stability	Advantages: Sample-independent, truth-seeking autonomously, absolutely logically self-consistent, stable under perturbation
Intelligence essence: Imitation of samples	Intelligence essence: Matching the ontological mathematical structure of the world

7 Theoretical Implementation Value

1. Full theoretical originality: For the first time overturning the predictive nature of Transformer and unifying all AI operations under the primitive of checking. This is a foundational paradigm-level original theory with innovation beyond ordinary top-conference papers;
2. Complete elimination of AI hallucinations: Eradicate hallucinations mathematically via differential stability and axiomatic checking rather than engineering tuning;
3. Support Artificial General Intelligence (AGI): Enable models to autonomously judge truthfulness, perform deduction and self-correction, possessing genuine cognitive intelligence;
4. Enable construction of new model architectures: Break away from the traditional Transformer framework and build native next-generation intelligent models from scratch based on the "checking closed-loop".

8 Conclusion

1. The essence of all AI intelligence is mathematical checking;
2. Transformer is not a predictive model but a single-layer statistical checking machine;
3. The only correct direction for next-generation AGI: four-dimensional global checking closed-loop incorporating vector structure, differential stability and axiomatic topology;

4. Ultimate definition of intelligence: A stable system that iterates continuously, self-verifies and matches the ontological rules of the universe within high-dimensional mathematical space.

A Appendix A Axiomatic Topological Checking: Computable Formalization via Category Theory / Homotopy Type Theory

A.1 A1 Core Refactoring Principle

Completely eliminate speculative natural-language descriptions. The four-layer checking system, especially the L4 axiomatic topological checking layer, is strictly transformed into categorical morphism rules, homotopy equivalence criteria, topological invariant constraints and boolean terminable computable conditions. All definitions are differentiable, algorithmizable and automatically verifiable with no room for vague philosophical arguments.

A.2 A2 Basic Mathematical Definitions (Computable Base)

A.2.1 A2.1 World Axiom Category $\mathcal{W}$

Define a fixed axiom category consisting of global conservation laws, mathematical axioms and rigid physical constraints:

$$\mathcal{W} = (Ob(\mathcal{W}), Mor(\mathcal{W}), \circ, id)$$

- Object set $Ob(\mathcal{W})$: Arithmetic axioms, geometric invariants, conservation laws, rigid causal temporal constraints;
- Morphism set $Mor(\mathcal{W})$: Valid deductions among axioms, structure-preserving transformations, causally admissible mappings;
- Core property: $\mathcal{W}$ is a rigid undeformable category. All valid morphisms must strictly preserve commutativity, structure and invariants.

A.2.2 A2.2 Model Instantaneous Reasoning Category $\mathcal{M}$

Semantic structures, logical chains and vector states generated in a single forward inference pass of the model constitute the instantaneous reasoning category:

$$\mathcal{M} = (Ob(\mathcal{M}), Mor(\mathcal{M}))$$

- $Ob(\mathcal{M})$: Model output propositions, high-dimensional vector nodes, intermediate state tensors;
- $Mor(\mathcal{M})$: Attention associations, reasoning derivation chains, state temporal transition mappings.

A.3 A3 L4 Axiomatic Topological Checking: Core Computable Criteria

A.3.1 A3.1 Categorical Structure-Preserving Morphism Verification

A functor $F : \mathcal{M} \rightarrow \mathcal{W}$ is valid if and only if three computable constraints are satisfied:

1. Object embedding compatibility: $F(Ob(\mathcal{M})) \subset Ob(\mathcal{W})$;
2. Structure preservation of morphisms: $\forall f \in Mor(\mathcal{M}), F(f) \in Mor(\mathcal{W})$;
3. Strict commuting diagram conservation:

$$F(g \circ f) = F(g) \circ F(f)$$

Computable decision rule: Any reasoning step breaking commuting diagram compatibility $\Rightarrow$ structural invalidity, trigger reconstruction rollback.

A.3.2 A3.2 Homotopy Type Theory (HoTT) Truth Verification

Adopt the core HoTT paradigm "propositions as types, truth as homotopy equivalence". Let the type corresponding to a model output proposition be $\mathbb{T}_P$, and the world axiom truth reference type be $\mathbb{T}_{true}$. Then:

$$P \text{ is objectively true} \iff \mathbb{T}_P \simeq \mathbb{T}_{true}$$

Homotopy equivalence serves as the ultimate computable truth criterion to fundamentally distinguish statistical correctness from objective structural correctness.

A.3.3 A3.3 Topological Invariant Conservation Verification

Extract core topological invariants of the reasoning manifold: Euler characteristic χ , low-order homotopy groups π_n , cohomology constraints. Set an infinitesimal threshold $\epsilon \rightarrow 0$, the checking rule reads:

$$|\chi_{\mathcal{M}} - \chi_{\mathcal{W}}| < \epsilon$$

Excessive deviation of topological invariants is judged as structural collapse and hallucination generation.

A.4 A4 Complete Computable Mathematical Hierarchy of Four-Layer Checking

- L1 Statistical checking: Minimize KL divergence $D_{KL}(P_\theta \| P_{real})$ to match sample statistical distributions;
- L2 Vector structure checking: High-dimensional tensor compatibility, vector angle constraints, manifold isometric embedding verification;
- L3 Differential stability checking: Bounded convergence verification of local perturbation gradients $\partial S / \partial \delta x$;
- L4 Axiomatic topological checking: Categorical morphism commutativity + HoTT homotopy equivalence + topological invariant conservation rigorous verification.

B Appendix B Minimum Feasible Validation: Small-Scale Prototype Experiments, Ablation Schemes and Standard Algorithm Pipeline

B.1 B1 Core Experimental Thesis

To verify the core paradigm that "multi-layer mathematical verification closed-loop intelligence outperforms pure statistical predictive intelligence", construct a minimally controllable mathematical reasoning experimental system. Under fully aligned parameter count, training data and optimization configurations, compare the pure Transformer baseline (L1 statistical checking only) and four-dimensional global checking architecture (L1+L2+L3+L4) on OOD generalization, anti-disturbance stability and self-correction capability to quantitatively prove the validity of the paradigm.

B.2 B2 Experimental Task Design (Bias-Free, Purely Controllable Mathematical Ground Truth)

1. Multi-step arithmetic OOD reasoning task
Training set: 1–3 step four-operation arithmetic samples (statistical distribution samples);
Test set: 4–6 step deeply nested arithmetic operations (fully out-of-distribution OOD samples, to test non-statistical generalization capability).
2. Basic physical conservation task
Train on samples of uniform-mass point motion, momentum and energy conservation;
Test OOD scenarios with tiny perturbations that break statistical conventions yet comply with physical axioms, to examine axiom constraint locking capability.

B.3 B3 Model and Training Alignment Configuration

- Baseline model: Standard small Transformer (12 layers, 128 hidden dimensions, standard self-attention);
- Experimental model: Identical architecture, parameter count and optimizer, augmented with L2/L3/L4 four-layer verification closed-loop modules;
- Control variables: Training epochs, learning rate, batch size and initialization are fully consistent. The only variable is the multi-layer checking mechanism.

B.4 B4 Core Quantitative Evaluation Metrics

1. OOD generalization accuracy: Acc_{OOD} , out-of-distribution reasoning accuracy measuring truth-seeking capability independent of training samples;
2. Anti-perturbation stability: Gaussian noise added to inputs, output error variance $Var(\Delta Error)$, measuring the effect of L3 differential stability constraints;
3. Self-correction convergence capability: Error self-repair rate and iterative convergence steps, measuring unique properties of closed-loop self-verifying intelligence.

B.5 B5 Standard LaTeX Algorithm Pseudocode

The algorithm below describes the abstract computable pipeline of the four-dimensional global mathematical verification closed-loop reasoning algorithm, the core execution logic of the theoretical model. All verification steps strictly correspond to the mathematical criteria in Appendix A. [Algorithm] Four-Dimensional Global Mathematical Verification Closed-Loop Reasoning Algorithm (L1–L4) [1] Context vector sequence $\{x_t\}$, axiom category system $\mathcal{W}$, perturbation radius δ , error threshold ϵ Compliant stable output y^* or trigger representation reconstruction iteration The model generates candidate reasoning result y $flag_1 \leftarrow \text{KL-Check}(y)$ L1 Statistical checking: Minimize $D_{KL}(P_\theta || P_{real})$ to match data statistical distribution

$flag_2 \leftarrow \text{Tensor-Structure-Check}(y)$ L2 Vector checking: High-dimensional tensor compatibility, manifold angle, cross-context structural self-consistency verification

$flag_3 \leftarrow \text{Perturb-Stability-Check}(y, \delta, \epsilon)$ L3 Differential checking: Local perturbation differentiation, verify bounded gradient and system state stability

$flag_4 \leftarrow \text{Axiom-Topo-Check}(y, \mathcal{W}, \epsilon)$ L4 Axiomatic topological checking: Categorical morphism commuting diagram conservation, HoTT homotopy equivalence, topological invariant verification $flag_1 \wedge flag_2 \wedge flag_3 \wedge flag_4 = \text{True}$ **return** $y^* \leftarrow y$ Pass all four layers of constraints, output objectively compliant result Execute vector representation reconstruction, error rollback and regenerate candidate y

B.6 B6 Algorithm Paradigm Explanation

Algorithm B.5 is an abstract theoretical algorithm flow rather than engineering source code. Each verification function strictly corresponds one-to-one to the formal mathematical definitions in Appendix A:

- L1 implements statistical distribution fitting and checking of the traditional Transformer;
- L2 implements structural self-consistency constraints on high-dimensional semantic manifolds;
- L3 implements anti-disturbance stability filtering at the differential dynamical level to eliminate fragile statistical solutions;
- L4 implements objective truth locking via category theory + HoTT + topological invariants.

Instead of centering on "fitting and prediction", the model centers on "iterative verification of multi-layer mathematical compliance". It fundamentally addresses the inherent defects of pure Transformers: sample dependence, poor OOD generalization and susceptibility to logical hallucinations.

B.7 B7 Expected Ablation Study Conclusions

1. Pure Transformer (L1 only): OOD long-reasoning accuracy drops sharply; tiny input perturbations induce logical errors and hallucinations;
2. L1+L2+L3 stacking: Structural self-consistency and anti-interference improve significantly, yet objective axiomatic constraints remain absent;
3. Full four-layer architecture L1+L2+L3+L4: Realizes four-dimensional closed-loop of statistical adaptation + structural self-consistency + differential stability + axiomatic fidelity. OOD generalization, self-correction and anti-interference capabilities comprehensively surpass the traditional statistical AI paradigm, validating the correctness of the original intelligent primitive theory in this paper.

B.8 B8 Preprint Submission Statement

Version v1 of this paper completes the full theoretical system, mathematical formalization, computable algorithm definition and controllable experimental scheme design. Full engineering implementation and large-scale measured data tables are reserved for subsequent iterations, satisfying publication standards for theory-focused preprints.