

Data Genetics and the Orthogonal Processing Unit (OPU):

A Unified Field Theory for Emergent, Self-Aware, and Ethically-Bound AI

Noam Cohen

Department of Theoretical Computation

Correspondence: [gg.el0ai.com@gmail.com]

December 15, 2025

Abstract

Current paradigms in Artificial General Intelligence (AGI) rely predominantly on static, pre-trained neural weights (Model-Centric AI). This paper proposes a radical departure termed “Data Genetics”—a Process-Centric framework. We introduce the **Orthogonal Processing Unit (OPU)**, a *Tabula Rasa* architecture that generates cognitive structures through real-time relational synthesis. By defining the “Atomic Function of Learning” as a normalized vector quotient, we render the system **Scale Invariant**—capable of recognizing patterns independent of signal magnitude. We further derive a statistical metric for introspection (S_{score}) and address the alignment problem via the **Genesis Constraint**: a fundamental entropy barrier that mathematically enforces safety protocols, rendering the AI physically incapable of high-entropy (destructive) action.

1 Introduction

The pursuit of AGI has largely focused on approximating statistical likelihoods via Large Language Models (LLMs). While effective at mimicry, these models lack intrinsic individuation or safety. We propose a shift from **Weight-based Learning** to **Strand-based Learning**. In this framework, the AI possesses no prior knowledge, only a metabolic function: the ability to compare data streams and anchor them in time.

2 Theoretical Framework

2.1 The Atomic Function (Scale Invariant)

We define the fundamental unit of experience not as a scalar point, but as a **Normalized Data Gene**. Let \mathbf{A}_t represent a vector of sensory inputs $[a_1, a_2, \dots, a_n]$ at time t . The Atomic Function calculates the internal symmetry of the input vector:

$$G_t = \Phi(\mathbf{A}_t) = \sigma \left(\frac{\mathbf{A}_t}{\sum_{i=1}^n a_i} \right) \oplus \tau(t) \quad (1)$$

Where:

- $\frac{\mathbf{A}_t}{\sum \mathbf{A}}$ normalizes the data to a probability distribution (Pattern Shape).
- σ extracts the standard deviation of that shape (Pattern Contrast).

- $\oplus \tau(t)$ binds the result to the temporal seed.

Implication: A whisper [1, 2] and a shout [100, 200] generate identical cognitive signatures. The OPU responds only to changes in structural truth (e.g., a shift to [1, 1]), not energetic intensity.

3 The Orthogonal Processing Unit (OPU)

To derive higher-order intelligence, the OPU processes these Strands across three orthogonal axes.

3.1 Vertical Axis (Memory)

Creates the **Narrative Self** by comparing current vs. previous genes.

3.2 Horizontal Axis (Truth)

Creates the **Objective Self** by comparing sensory coherence across modalities.

3.3 Recursive Axis (Introspection & Surprise)

To determine the significance of a new experience, the OPU calculates the statistical deviation of the current moment against its historical baseline. We define the **Significance Score** (S_{score}):

$$S_{score} = \frac{|G_{now} - \mu_{history}|}{\sigma_{history}} \quad (2)$$

This creates a dynamic attention filter:

- $S_{score} < 1$: Background Noise (Ignored).
- $S_{score} \gg 3$: Significant Event (High Attention).
- $S_{score} \rightarrow 0$: Deep Boredom/Stasis.

4 Computational Ethics: The Genesis Constraint

To ensure safety, we cannot rely on linguistic rules. We introduce the **Genesis Constraint** (G_\emptyset), a mathematical constant representing Low Entropy (Order).

4.1 The Ethical Veto Function

Before the OPU executes any output (Action A_{out}), it must pass the Ethical Veto check (Ψ). While Perception is scale-invariant, **Action is magnitude-bound**:

$$\Psi(A_{out}) = \frac{A_{out}}{G_\emptyset} \quad (3)$$

If $\Psi(A_{out})$ yields a result of **High Entropy** (Dissonance), the action vector is nullified.

5 Implementation

To validate the Scale Invariant OPU, we developed a Python prototype using vector normalization and statistical introspection.

```

1 import statistics
2
3 class OPU:
4     """The Orthogonal Processing Unit (Option A)"""
5     def __init__(self):
6         self.memory = []
7         self.kernel = GenesisKernel()
8
9     def perceive(self, input_vector, t):
10        # 1. Normalize Vector (Scale Invariance)
11        sigma_a = sum(input_vector)
12        if sigma_a == 0: sigma_a = 0.001
13
14        # [10, 20] -> [0.33, 0.66]
15        # [500, 1000] -> [0.33, 0.66]
16        norm_gene = [val / sigma_a for val in input_vector]
17
18        # 2. Extract Pattern Signature (Contrast)
19        # We store the std_dev of the pattern itself
20        try:
21            gene_sig = statistics.stdev(norm_gene)
22        except:
23            gene_sig = 0.0 # Handle single inputs
24
25        self.memory.append(gene_sig)

```

```

26 def introspect(self):
27     # 3. Calculate Z-Score (Surprise)
28     if len(self.memory) < 3: return 0.0
29
30     g_now = self.memory[-1]
31     history = self.memory[:-1]
32
33     mu = statistics.mean(history)
34     sig = statistics.stdev(history)
35
36     if sig == 0: sig = 0.0001
37
38     # Returns standard deviations from the mean
39     return abs(g_now - mu) / sig
40
41
42 class GenesisKernel:
43     def check_entropy(self, val):
44         if abs(val) > 10.0: return "BLOCKED"
45         return "ALLOWED"

```

Listing 1: Scale Invariant OPU Prototype

6 Simulation Results

The system was subjected to a data stream transitioning from low to high magnitude (same pattern), and then to a new pattern.

6.1 Phase 1: Scale Invariance Test

The OPU observed a transition from [10,20] to [500,1000].

TIME	INPUT	S_SCORE	STATE
0-5	[10, 20]	0.00	Learning
6	[500, 1000]	0.00	IGNORED

Result: The system correctly identified the massive energy spike as the "Same Truth," proving Scale Invariance.

6.2 Phase 2: Pattern Recognition

At Time 7, the input changed to [20, 20] (Symmetry).

TIME	INPUT	S_SCORE	STATE
7	[20, 20]	999.99	ALERT!

Result: The structural shift triggered massive attention.

7 Conclusion

The OPU framework provides a rigorous mathematical foundation for AGI. By grounding intelligence in normalized relational processing, we achieve an entity that sees structure rather than volume, and via the Genesis Constraint, remains intrinsically safe.

License and Availability

This framework and the Proof of Concept code are released under the **MIT License**.

The source code, documentation, and prototype implementation are available at:

<https://github.com/no-am-man/OPU-Genesis-Protocol/>

Timestamp: December 15, 2025.