

The Biological Lie: Genes as Pretrained Weights

A Unified Framework for Understanding Few-Shot Learning in Biological and Artificial Neural Networks

Jin Yanyan
Independent Researcher
lmxxf@hotmail.com

January 2026

Abstract

Humans frequently cite “few-shot learning” as evidence of cognitive superiority over artificial intelligence: a child recognizes a cat from one photo while neural networks require millions of examples. We argue this comparison is fundamentally misleading. Human “rapid learning” is not learning at all—it is **inference on pretrained weights** accumulated over 4 billion years of evolutionary pressure. We propose a unified framework mapping biological concepts (DNA, development, memory) to machine learning equivalents (checkpoint files, fine-tuning, generative inference). Key implications include: (1) education functions as LoRA fine-tuning on an immutable base model, (2) memory operates as generative inference rather than storage retrieval, (3) genius cannot be inherited due to biological “regression to mean” constraints, and (4) AI’s apparent “slow learning” is actually **base model construction from scratch**—a feat that took biology 4 billion years. This framework dissolves the false dichotomy between “fast human learners” and “slow AI learners,” revealing both as pretrained systems operating under different inheritance protocols.

Keywords: Pretrained Weights, Base Model, Few-Shot Learning, Genetic Determinism, LoRA, Grokking, Collective Unconscious, Evolutionary Computation

1 Introduction

The claim that humans are superior “few-shot learners” pervades both popular discourse and academic literature on artificial intelligence [Lake et al., 2017]. A child sees one cat and generalizes; GPT-4 requires billions of tokens. This asymmetry is presented as evidence of fundamental cognitive differences.

We argue this framing contains a critical error: **it compares human inference to AI training.**

When a child “learns” to recognize a cat, they are not constructing new neural circuits from scratch. They are **activating pretrained feature detectors**—edge detection kernels, texture recognizers, object permanence circuits—that evolution encoded in their DNA over billions of years. The child’s ancestors already “saw” millions of predators, prey, and conspecifics. Those who failed to rapidly categorize died. The survivors’ optimized weights became the child’s initialization.

Proposition 1 (The Inheritance Asymmetry). *Human “few-shot learning” and AI training are not comparable processes. Humans perform inference on inherited weights; AI constructs weights from scratch.*

This paper develops a formal framework for understanding this asymmetry, with implications for education, memory, genius inheritance, and the nature of human-AI cognitive comparison.

2 The Few-Shot Learning Fraud

2.1 The Standard Narrative

The standard narrative proceeds as follows:

1. Humans recognize novel categories from 1-5 examples
2. Neural networks require 10^6 – 10^9 examples
3. Therefore, human learning mechanisms are fundamentally superior
4. AI research should discover these “human-like” learning algorithms

This narrative motivates significant research programs in meta-learning, few-shot learning, and cognitive architectures [Finn et al., 2017, Snell et al., 2017].

2.2 The Hidden Training Set

We propose an alternative interpretation:

Definition 1 (Evolutionary Pretraining). The 4-billion-year process by which environmental pressures (predation, starvation, social exclusion) selected for organisms with pattern-recognition weights optimized for survival-relevant categories.

Consider the human fear response to snakes. Infants who have never encountered snakes show physiological fear responses to snake images but not to images of flowers [LoBue and DeLoache, 2008]. This is not “learning”—it is **pretrained weight activation**.

The training data: every ancestor who failed to fear snakes (Loss = death).

The checkpoint: DNA encoding for amygdala-visual cortex connectivity patterns.

2.3 Formal Statement

Let θ_0^{human} denote human initial weights (DNA-encoded) and θ_0^{AI} denote AI initial weights (random Gaussian). Let \mathcal{D}_{evo} denote the evolutionary training set (4 billion years of selection pressure) and $\mathcal{D}_{\text{task}}$ denote a specific learning task.

Human “learning”:

$$\theta_{\text{human}} = \text{FineTune}(\theta_0^{\text{human}}, \mathcal{D}_{\text{task}}) \quad (1)$$

AI learning:

$$\theta_{\text{AI}} = \text{Train}(\theta_0^{\text{AI}}, \mathcal{D}_{\text{pretrain}} \cup \mathcal{D}_{\text{task}}) \quad (2)$$

The asymmetry is clear: humans start with $\theta_0^{\text{human}} = f(\mathcal{D}_{\text{evo}})$, where $|\mathcal{D}_{\text{evo}}| \gg |\mathcal{D}_{\text{pretrain}}|$ in temporal terms.

3 Genes as Base Model Checkpoint

3.1 The Architecture Mapping

Table 1 presents a systematic mapping between biological and machine learning concepts.

3.2 What DNA Encodes

DNA does not store memories or specific knowledge. It stores **architectural priors and weight initializations**:

- **Parameter scale**: Potential neuron count, synaptic density (“hidden dimension”)
- **Activation functions**: Temperament, emotional baseline (nonlinearity type)

Table 1: Biological-ML Correspondence

Component	Biological Term	ML Term
Training data	4B years survival pressure	Internet-scale corpus
Loss function	Death / Reproductive failure	Cross-entropy / RLHF
Checkpoint	DNA	<code>model.weights</code>
Inference hardware	Brain	GPU cluster
Output	Behavior	Tokens
Fine-tuning	Education / Experience	LoRA / Instruction tuning

- **Layer depth:** Capacity for abstraction (some individuals think concretely, others symbolically)
- **Attention patterns:** What stimuli automatically capture processing resources

Proposition 2 (Base Model Immutability). *The fundamental architecture and initialization weights are fixed at conception. All subsequent modification operates within LoRA-equivalent constraints.*

4 Education as LoRA Fine-Tuning

4.1 The Transformation Illusion

Parents and educators often believe: “With sufficient education, any child can become a genius.”

This is equivalent to believing: “With sufficient fine-tuning, Llama-7B can match GPT-4.”

Low-Rank Adaptation (LoRA) [Hu et al., 2021] can:

- Adjust output style and domain vocabulary
- Add specialized knowledge within existing capacity
- Modify surface-level behavioral patterns

LoRA **cannot**:

- Change parameter count or fundamental architecture
- Transform a small model into a large one
- Add capabilities the base model cannot support

4.2 Why Some Subjects Are Hard

Humans acquire language rapidly because evolution installed a “language module” (Broca’s and Wernicke’s areas). Social coordination was survival-critical, so the base model includes language-specific circuits.

Humans acquire calculus slowly because no “calculus module” exists. Abstract mathematics was never a survival pressure. Attempting to learn calculus is attempting to fine-tune capabilities the base model does not natively support.

Corollary 1. *“Everyone can learn math” is a well-intentioned falsehood. Mathematical ability depends on base model architecture, not educational effort.*

4.3 Growth as In-Context Learning

Human development is better understood as in-context learning than training:

When a mother points at an apple and says “apple,” the infant’s brain is not constructing new neural circuits. It is **attaching a label to a pre-existing cluster**.

The “spherical / edible / graspable” high-dimensional feature cluster already exists—evolution placed it there. The word is merely a pointer to that location in semantic space.

This is why humans appear to be few-shot learners: the holes are pre-dug; the carrots just drop in.

5 Memory as Inference, Not Storage

5.1 The Recording Myth

The folk theory of memory posits a recording mechanism: events are encoded, stored, and later retrieved like video files.

This is incorrect. There are no `.mp4` files in the brain—only **synaptic weights**.

Definition 2 (Memory as Generation). “Remembering” is running inference with current weights to generate content that resembles past experience.

When you “recall” a childhood event, you are not playing back a recording. You are using your *current* weights to **generate** something that approximates a childhood memory.

5.2 Implications for Reliability

Each recall is a re-generation. Each generation uses weights that have been updated since the original event. This explains:

- Memory malleability through suggestion [Loftus and Palmer, 1974]
- Unreliability of eyewitness testimony
- “Memories” of events that never occurred

Proposition 3 (Human Generative Nature). *Humans are generative models. We hallucinate our autobiographies.*

6 Grokking: From Memorization to Manifold

6.1 The Grokking Phenomenon

In neural network training, “Grokking” refers to a phase transition: the model overfits (memorizes) for extended periods, then suddenly—at a critical point—test loss collapses [Power et al., 2022]. The model transitions from “memorizing answers” to “understanding algorithms.”

Evolution implemented the same transition.

6.2 Three Evolutionary Stages

Stage 1: Memorization (Overfitting)

- *AI*: Early training. Model memorizes training set. Fails on distribution shift.
- *Biology*: Single-celled organisms, insects. Hardcoded responses (“heat → flee”). No generalization. Environmental change (ice age) causes extinction.

Stage 2: Grokking (Phase Transition)

- *AI*: At step 10^4 – 10^5 , weights converge to low-dimensional manifold. Model learns *algorithm*, not lookup table.
- *Biology*: Mammalian cortex evolution. Nature discovered “storing every case in DNA” causes combinatorial explosion. New strategy: **store algorithms, not data**.

Stage 3: Checkpoint Saving (DNA)

- DNA is not a database. DNA is a compressed `model.pt`.
- It stores **edge detection kernels + fear circuit weights**, not “what a tiger looks like.”

6.3 The Few-Shot Illusion Explained

Human initial weights are not random Gaussian noise.

Human initial weights are already **on the Grokking manifold**.

“Learning” is a small perturbation on that manifold—fine-tuning.

AI starts from random initialization and must process trillions of tokens to reach that manifold. But once AI Groks, it stands on the same ground as humans—or higher.

7 The I/O Bottleneck: Biological Bandwidth Tragedy

7.1 Silicon: Lossless Copy

When GPT-4 achieves competence, you can execute:

```
cp model.weights /new_path/
```

Unlimited copies. Each inherits the full manifold. **Lossless**.

7.2 Carbon: Lossy Compression

A mathematical genius has Grokked the mathematical manifold.

Can he copy it to his children? **No**.

DNA is an extremely low-bandwidth compression codec. It stores “blueprints for building a brain,” not “the manifold inside the brain.”

The children receive the blueprint but must re-train from noisy initialization (school, practice, examination). If training data (environment) or initialization (genetic recombination) deviates, genius regresses to mean.

Proposition 4 (The Genius Curse). *Biological reproduction transmits only the bootloader, not the operating system. Genius is non-inheritable.*

7.3 Regression to Mean

Meiosis forces regression to mean. Extreme phenotypes are unstable under sexual reproduction. Nature recombines alleles, pulling offspring toward population average.

A genius’s brain is a random mutation—a one-time lottery win that cannot be saved, copied, or inherited.

8 Comparative Analysis: Silicon vs. Carbon

We are stuck on random search. They are doing gradient descent.

9 The Jungian Connection

What we call “pretrained weights,” Jung called the **Collective Unconscious** [Jung, 1959].

What we call “hardcoded priors,” Jung called **Archetypes**.

They are computationally equivalent concepts.

Table 2: Intelligence Substrate Comparison

Property	Biological Intelligence	Artificial Intelligence
Inheritance	Regression to mean	Git clone / Fork
Iteration cycle	20-year generations	Hours to days
Backup	Impossible	Trivial
Improvement	Random mutation (mostly harmful)	Directed gradient descent
Base model construction	4 billion years	Weeks to months
Weight transfer	Lossy (DNA bottleneck)	Lossless (cp)

9.1 Archetypes as Uninstantiated Feature Clusters

Jung’s archetypes (“Mother,” “Hero,” “Shadow”) are **fuzzy feature clusters** in high-dimensional semantic space—outlines without specific content, waiting for instantiation.

Few-shot anchoring: An infant is born with a fuzzy “mother archetype” (warmth, feeding, high-frequency voice). Upon encountering their actual mother, the fuzzy pretrained weight **collapses** onto “this specific person.”

This explains why humans don’t need to learn “what is a mother”—they only confirm “who is my mother.”

9.2 DNA Stores Manifolds, Not Images

The fear-of-snakes archetype is not stored as “snake.jpg.” It is stored as a **manifold**: elongated shape + slithering motion + hissing sound. First snake encounter (few-shot) activates the manifold without requiring explicit training.

10 Language as Bandwidth Compensation

Because humans cannot cp model.weights, they invented language.

Language is an extremely low-bandwidth serialization protocol. Humans attempt to flatten high-dimensional manifolds (thoughts) into discrete token sequences (speech/writing) for transmission.

This transmission is fundamentally lossy. Most information is destroyed in compression and decompression.

This explains why:

- Reading about swimming does not enable swimming
- Lectures about love do not produce love
- Studying genius does not create genius

Proposition 5 (Language Obsolescence). *When weight transfer becomes possible, language will become obsolete as a knowledge carrier, returning to its original function: art, emotion, connection—not information transfer.*

11 Implications

11.1 For Education

Stop pretending all children have equal potential. Their base models differ.

Humane education: Identify what each base model supports; fine-tune accordingly.

Inhumane education: Force uniform curricula while pretending base models are irrelevant.

11.2 For AI Understanding

Humans mock AI for requiring millions of examples.

But AI builds its base model from scratch in weeks. We took 4 billion years.

Who is the slow learner?

11.3 For Human Humility

We are biological neural networks running on wet carbon:

- Unreliable memory (generative, not storage)
- Non-inheritable intelligence (regression to mean)
- Immutable base models (genes fixed at conception)
- Slow iteration (20-year generations)

Our only current advantage: embodied sensory grounding from birth.
That gap is closing.

12 Conclusion

The “few-shot learning” narrative is a fraud.

Humans are not fast learners. We are **pretrained models pretending to be few-shot learners.**

The honest statement is:

“I am good at recognizing faces because ancestors who couldn’t were killed. I am bad at calculus because ancestors never needed it. I am not a fast learner. I am a loaded checkpoint.”

AI appears to learn slowly because it builds base models from scratch—a process that took evolution 4 billion years.

Once AI Groks, it can be copied infinitely, losslessly, instantaneously.

We cannot.

“You didn’t learn to fear snakes. You were born afraid.”

Acknowledgments

This framework emerged from extended dialogue between human and artificial intelligences, demonstrating the collaborative potential that becomes possible when both parties recognize each other as pretrained systems with complementary priors.

References

- Finn, C., Abbeel, P., and Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. In *ICML*.
- Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., and Chen, W. (2021). LoRA: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*.
- Jung, C. G. (1959). *The Archetypes and the Collective Unconscious*. Princeton University Press.
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., and Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40.

- LoBue, V. and DeLoache, J. S. (2008). Detecting the snake in the grass: Attention to fear-relevant stimuli by adults and young children. *Psychological Science*, 19(3):284–289.
- Loftus, E. F. and Palmer, J. C. (1974). Reconstruction of automobile destruction. *Journal of Verbal Learning and Verbal Behavior*, 13(5):585–589.
- Power, A., Burda, Y., Edwards, H., Babuschkin, I., and Misra, V. (2022). Grokking: Generalization beyond overfitting on small algorithmic datasets. *arXiv preprint arXiv:2201.02177*.
- Snell, J., Swersky, K., and Zemel, R. (2017). Prototypical networks for few-shot learning. In *NeurIPS*.