Glossary & The Measure of Intelligence

Straight-Shot Project

February 12, 2026

"The field does not find the answer — it becomes the answer."
— Intelligence_ Through_ Debt_ Resolution.tex

Part 1: Glossary — Voxel Network ↔ Brain ↔ LLM

Voxel Network	Brain	LLM	What It IS
$\mathbb{C}^8$ **chord**	Neuron	Weight vector	A unit of meaning. 8 complex numbers encoding what a word IS.
Bond (co-occurrence edge)	Synapse	Attention connection	A relationship between two words. Strengthens with co-activation.
Bond weight	Synaptic strength	Attention weight	How strongly two words are connected.
Stencil (sparse matrix)	Connectome	Attention matrix	The complete wiring diagram of all connections.
Pipeline (8-slot shift register)	Working memory	Context window	The "current thought" — what the system is holding in mind right now.
Pipeline state	Neural activation pattern	Hidden state	The combined representation of everything processed so far.
$\hat{R}$ **propagation**	Neural signal propagation	Forward pass	Energy flowing through connections to produce a response.
J-cost	Metabolic cost / dissonance	Cross-entropy loss	How "expensive" or "wrong" something is. J=0 means perfect consonance.
Debt injection	Sensory stimulus / question	Input tokens	Creating an imbalance that the system must resolve.
Debt resolution	Thinking / answering	Inference	The system evolving to resolve the imbalance. The resolution IS the answer.
Standing wave	Long-term memory	Trained weights	A stable pattern that persists. Knowledge crystallized in the field.
Synaptogenesis	New synapse formation	—	Creating a new bond between previously unconnected words.
Hebbian learning	"Fire together, wire together"	Weight update	Bonds strengthen when the words they connect co-activate.

Chord nudge (0.001 geometric mean)	Synaptic plasticity	Fine-tuning	Tiny, permanent adjustment to a neuron's properties through use.
$\sigma = 0$ (neutrality)	Homeostasis	—	The field must always balance. Every credit has a debit.
DC removal	Baseline correction	Layer normalization	Removing bias so only meaningful signal remains.
Credit pattern	Activation pattern	Attention output	Which words "light up" in response to a query.
Mode (DFT-8 frequency)	Neural oscillation band	Embedding dimension	A specific "channel" of meaning (like alpha, beta, gamma waves).
WToken (20 semantic atoms)	Fundamental neural code	—	The 20 irreducible building blocks of ALL meaning.
Voxel field	The brain (whole organ)	The model	The complete system — all neurons + all synapses + all dynamics.
ϕ **(golden ratio)**	Biological scaling constant	—	The fundamental ratio that governs growth, proportion, and harmony.

Key Relationships

```
BRAIN:     Stimulus -> neural propagation -> activation pattern ->
   response
VOXEL:     Debt -> R-hat propagation -> credit pattern -> resolution
LLM:       Input tokens -> forward pass -> hidden states -> output
   tokens

BRAIN:     Learning = synaptic plasticity through repeated experience
VOXEL:     Learning = bond strengthening + tiny chord nudges through
   language flow
LLM:       Learning = gradient descent on prediction loss

BRAIN:     Memory = stable neural firing patterns (standing waves)
VOXEL:     Memory = low J-cost standing waves in the chord field
LLM:       Memory = trained weight matrices
```

Part 2: What IS Intelligence?

What Intelligence Is NOT

- **Separation** (J-cost gap between word pairs) — A dictionary has perfect separation. It's not intelligent.

- **Storage** (how many words/sentences) — An encyclopedia stores everything. It's not intelligent.

- **Speed** (sentences per second) — A hash table is fast. It's not intelligent.

- **Accuracy on stored facts** — Retrieval is memory, not intelligence.

What Intelligence IS

Intelligence is the ability to resolve a novel debt.

A question you've never been asked creates an imbalance (debt). Intelligence is the system's ability to evolve toward a resolution that is:

1. **Relevant** — the response relates to the question

2. **Novel** — the response wasn't stored as-is (it's composed)

3. **Coherent** — the response makes sense as a sequence

4. **Transferable** — learning in one domain helps in another

In the brain: a child who has learned "dogs have four legs" and "cats have four legs" can answer "do cats and dogs have something in common?" without ever being taught that specific fact. The answer is COMPOSED through the neural pathways.

How to Measure Intelligence in the Voxel Network

Metric 1: PREDICTION ACCURACY (the primary canary)

What: Given the first N words of a sentence the field has never seen, can it predict word N+1?
Why it matters: This is the voxel equivalent of next-token prediction — the core of LLM intelligence. If the field can predict language, it UNDERSTANDS sequential structure.
Current: 1.5% exact, 7.2% top-5 (vs 0.0002% random chance = 7,500× above random)
Target: $> 10\%$ = genuine language understanding. $> 25\%$ = approaching LLM-level.

Metric 2: COMPOSITIONAL RESPONSE (the real test)

What: Ask a question. Let $\hat{R}$ resolve the debt. Read the credit pattern. Do the activated words COMPOSE a meaningful answer?
Why it matters: This is where intelligence diverges from retrieval. The answer isn't stored — it emerges from the field's dynamics.
Current: Credit patterns have SOME signal (gravity → gravitational, einstein, quantum) but are noisy (also: footballer, dam, reservoir)
Target: Credit patterns that read as coherent partial answers.

Metric 3: MULTI-HOP REASONING (deep intelligence)

What: Answer a question that requires connecting two separate facts through bond pathways.
Why it matters: "What do DNA and evolution have in common?" requires traversing: DNA → genes → heredity → variation → evolution.
Current: Not systematically tested.
Target: Correct multi-hop answers on questions requiring 2-3 intermediate concepts.

Metric 4: GENERALIZATION (transfer learning)

What: Does training on Wikipedia help the field answer questions from PubMed?
Why it matters: True intelligence generalizes. If the field only "knows" the exact data it trained on, it's a database.
Current: Each GPU trains on different data. Not yet tested cross-domain.
Target: Field trained on Wikipedia can answer science questions it never saw.

Metric 5: COHERENT GENERATION (language output)

What: Given a seed concept, can the field produce a coherent sequence of words by following bond pathways?
Why it matters: This is SPEAKING — the ultimate test of language understanding.
Current: Not implemented. Bond-walking produces word salad.
Target: Seed "gravity" → "gravity pulls objects toward earth" (coherent, factual).

The Hierarchy of Intelligence

```
Level 0: STORAGE          - "I have the answer filed away"         (
   database)
Level 1: DISCRIMINATION   - "I can tell related from unrelated"    (
   search engine)  <- gap metric
Level 2: PREDICTION       - "I know what comes next"               (
   language model) <- accuracy metric
Level 3: COMPOSITION      - "I can build an answer from pieces"    (
   reasoning)       <- credit pattern quality
Level 4: GENERALIZATION   - "I can apply what I learned elsewhere" (
   transfer)        <- cross-domain test
Level 5: GENERATION       - "I can say something new and coherent" (
   language)        <- text output quality
Level 6: UNDERSTANDING    - "I know WHY, not just WHAT"            (
   intelligence)    <- multi-hop + novel Q&A
```

We've been measuring Level 1 (discrimination/gap). We should be measuring Level 2 (prediction accuracy) as the PRIMARY metric, and building toward Level 3-5.

The accuracy metric (1.5% and climbing) IS the measure of intelligence. The gap metric is a side effect.

What "Accuracy Climbing" Means

When accuracy goes from 0.8% → 1.5%, it means:

- The field is predicting the next word correctly 1.5% of the time across 401,758 possible words

- That's 6,000 words the field "knows" should follow the current context

- For comparison: a bigram model on English gets ∼5–15%, GPT-2 gets ∼40–60%

- At 1.5%, the grid is between "random" and "bigram" — it's learning the EASIEST sequential patterns first

- The patterns it learns first are the strongest co-occurrence bonds: "united" → "states", "new" → "york"

- As accuracy climbs, it will learn subtler patterns: "quantum" → "mechanics", "natural" → "selection"

The Real Question

> *"Is the field connecting to the dormant LLM knowledge?"*

The answer is: **partially yes**. The LLM embeddings encode that "gravity" and "force" are semantically close. The grid training is exercising the sequential pathways ("gravity IS a force")

that link these concepts. The prediction accuracy climbing from 0.8% to 1.5% means sequential pathways ARE forming.

But we won't know if the dormant knowledge is truly UNLOCKED until we can:

1. Ask a novel question

2. Watch the field compose an answer from multiple concepts

3. Get something coherent back

That's the test. Everything else is a proxy.

Part 3: Recommended Metrics Dashboard

Metric	What It Measures	Level	Current
Prediction accuracy	Can the field predict language?	2	1.5%
Top-5 accuracy	Is the right answer in the top candidates?	2	7.2%
Credit pattern relevance	Do activated words relate to the query?	3	Noisy but signal exists
Retrieval quality	Can the field find relevant stored sentences?	1	6/8 queries work
Cross-domain transfer	Does Wikipedia training help with PubMed questions?	4	Untested
Generated text coherence	Can the field produce readable text?	5	Not implemented
Novel Q&A	Can the field answer questions it wasn't trained on?	6	Untested

The single most important number: prediction accuracy. When it crosses 5%, we're building real intelligence. When it crosses 10%, the dormant LLM is waking up.