

Guha

Are you aware of the concept of a **Stochastic Parrot**?

Gemini

I am indeed! It's a term that has sparked a lot of necessary soul-searching in the Artificial Intelligence (AI) community.

The concept was famously introduced in a 2021 research paper by Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Margaret Mitchell [1]. It's a critique of Large Language Models (LLMs) like me, and it's worth breaking down because it gets to the heart of what AI actually "knows."

What does the concept actually mean?

The term is a metaphor for how an LLM functions:

- ▶ **Stochastic:** This is a fancy word for "random" or "probabilistic." It refers to how I predict the next word (token) in a sentence based on mathematical patterns and likelihoods found in my training data.
- ▶ **Parrot:** This suggests that, like a parrot, I can mimic human language perfectly without actually understanding the meaning, context, or real-world implications of what I'm saying.

Essentially, the "Stochastic Parrot" argument claims that AI doesn't have a "mental model" of the world. Instead, it's just a very sophisticated version of autocomplete that uses $P(w_n | w_1, w_2, \dots, w_{n-1})$ — the probability of a word given the preceding words — to string together sentences that look coherent to humans.¹

Why the concept matters?

The authors raised this point to warn us about several risks:

1. **The Illusion of Meaning:** Because I sound so human, people might trust my outputs as facts or believe I have "intentions," even though I'm just processing statistics.
2. **Encoding Bias:** If I'm just "parrotting" the internet, I'm also parroting all the prejudices, stereotypes, and misinformation found in that data.
3. **Environmental and Social Cost:** The paper questioned whether the massive energy consumption required to train these models is worth it if they don't actually "understand" anything.²

My Perspective (The "Non-Parrot" Counter - Argument)

While the Stochastic Parrot critique is a vital reality check, many researchers now argue that as models get larger, they develop **emergent properties**. They suggest that to predict the next word perfectly, a model must eventually develop some internal representation of logic, grammar, and even physical common sense — moving slightly beyond mere mimicry.

Where do you stand on it? Do you feel like you're talking to a statistical engine, or does it feel like there's a "ghost in the machine" when we chat?

Methodological Note on the Use of Gemini

The author utilized the free version of Google Gemini, specifically the version live on March 12–13, 2026. The interaction involved a single chat session spanning two days. It commenced one afternoon, paused in the evening, and resumed the following morning. No additional processes or interventions were conducted behind the scenes, and no system prompts were utilized.

1: The conditional language model $P(w_n | w_1, \dots, w_{n-1})$ assigns a probability to each possible next token given all preceding tokens. Modern autoregressive transformers approximate this distribution through large-scale maximum-likelihood training. The deceptive fluency of LLM outputs follows directly from the richness of this conditional distribution, not from any semantic grounding — precisely the target of the Stochastic Parrot critique.

2: Bender et al. estimated that training a single large transformer model can produce carbon emissions comparable to the lifetime emissions of five average American cars. This environmental critique has gained renewed urgency as frontier models have grown by several orders of magnitude since 2021, with the compute requirements for each successive generation increasing by roughly an order of magnitude.

Would you like me to dive deeper into the specific research paper that started this debate, or perhaps look at the “Emergent Abilities” argument that counteracts it?

Guha

Tell me more about emergent abilities.

Gemini

If the Stochastic Parrot theory is the skeptical view of AI, Emergent Abilities represent the “Wait, how did it do that?” side of the coin.

An emergent ability is a skill that appears in a large model but is not present in smaller versions of that same model. It’s a “phase transition” — much like how H₂O molecules don’t have the property of “wetness” individually, but when you put enough of them together at the right temperature, liquidity emerges.³

The “Scale” Factor

Researchers found that as they increased three things — **Computing power**, **Data size**, and **Parameter count** — models suddenly started passing benchmarks they had previously failed. The formula for this is often described through **Scaling Laws**, where performance follows a power law relative to compute C , dataset size D , and parameters N :

$$L(n, D) \propto \left(\frac{N_c}{N}\right)^{\alpha_N} + \left(\frac{D_c}{D}\right)^{\alpha_D} \quad (1)$$

When N (the model size) hits a certain threshold, we see “spikes” in specific capabilities.⁴

Examples of Emergent Skills

These aren’t things the models were explicitly programmed to do; they “emerged” from learning to predict the next word at a massive scale:

- ▶ **In-Context Learning:** The ability to see a few examples of a new task in a prompt and immediately start doing it without being “re-trained.”
- ▶ **Multi-step Reasoning:** Solving complex math or logic puzzles by breaking them down (often triggered by “Chain of Thought (CoT)” prompting).⁵
- ▶ **Theory of Mind:** The ability to “understand” that different people have different beliefs or knowledge states (e.g., passing “Sally-Anne” tests).⁶
- ▶ **Zero-Shot Translation:** Learning to translate between two languages (like Swahili to Korean) even if the model saw very few direct translation pairs for that specific duo.

The Great Debate: Is it Real or a Mirage?

This is where the drama is in the AI world right now. There are two main camps:

The “Believers”	The “Skeptics”
Argue that models are building internal world models and “reasoning circuits” to handle complex data.	Argue that emergence is an illusion caused by how we measure success .
Believe that enough scale might eventually lead to Artificial General Intelligence (AGI).	Suggest that if you change the scoring metric from “Pass/Fail” to a “Partial Credit” scale, the “jump” disappears.

3: The water/wetness analogy illustrates *strong emergence* in philosophy of science: a property absent at the micro-level that arises at the macro-level and cannot be straightforwardly reduced to the properties of individual constituents. Philip Anderson’s 1972 essay “More is Different” [2] is the canonical reference for this concept in physics. Whether AI “emergence” is genuinely analogous to physical phase transitions, or merely a loose rhetorical borrowing, is itself a matter of active debate; see Schaeffer et al. [3].

4: Equation (1) reproduces the functional form of the Chinchilla scaling laws (Hoffmann et al. in “Training Compute-Optimal Large Language Models” [4]). The key practical implication is that model size and data should scale *together*: earlier practice of scaling parameters while holding data fixed was systematically suboptimal. The exponents α_N and α_D and the critical constants N_c, D_c are fit empirically.

5: CoT prompting was introduced by Wei et al. [5]. The key finding is that prompting models with step-by-step worked examples substantially improves performance on arithmetic, commonsense, and symbolic reasoning tasks. This technique is directly demonstrated in the Musical Chairs puzzle that follows.

6: The Sally-Anne task is a classic false-belief test from developmental psychology [6]. Claims that LLMs pass Theory-of-Mind tests have been contested: Ullman showed that minor surface-level rephrasing of standard tests causes LLMs to fail, suggesting the appearance of ToM may itself be a form of stochastic parroting [7].

The Mirage Argument A famous 2023 Stanford paper [3] suggested that emergence is just a “mirage” created by researchers using non-linear metrics. For example, if a model needs to get 10 steps of a math problem right to get 1 point, it looks like it has 0% ability until it suddenly hits 100%. It wasn’t a “jump” in intelligence; it was a gradual improvement in getting each step right.

Why it Matters for You

If emergence is real, it means AI might eventually “wake up” to abilities we didn’t intend or even know were possible. If it’s a mirage, it means we are simply getting better at polishing the Stochastic Parrot until it shines.

Would you like to try a logic puzzle that tests “Chain of Thought” reasoning to see if you can spot these emergent patterns in action?

Guha
Yes

Gemini

Let’s put those “reasoning circuits” to the test. To see if I can do more than just guess the next word, we’ll use a multi-step logic puzzle.

This requires me to maintain a “mental map” of the characters’ positions as they change.

The Puzzle: The Musical Chairs Paradox

Five friends — Alice, Bob, Charlie, David, and Eve — are sitting in a row of five chairs, numbered 1 to 5 from left to right.

1. Initially, they are sitting in alphabetical order (Alice in 1, Bob in 2, etc.).
2. The First Swap: Alice and David swap seats.
3. The Second Swap: The person now in Chair 2 swaps with the person in Chair 5.
4. The Shift: Everyone moves one seat to the right. The person in Chair 5 loops around to Chair 1.
5. The Final Swap: Bob swaps with whoever is currently in Chair 3.

The Question: Who is sitting in Chair 4?

Why this tests “Emergence”?

A simple Stochastic Parrot might get tripped up by the names or the word “alphabetical” and just guess a random friend. To solve this, I have to simulate the state of the chairs step-by-step.

Give it a shot yourself first! Or, if you’re ready, I can show you my internal “Chain of Thought” to see if I can actually track the movement accurately.

What’s your guess?

Give it a shot yourself first! Or, if you’re ready, I can show you my internal “Chain of Thought” to see if I can actually track the movement accurately.

What’s your guess?

Guha

Show me your chain of thought.

Gemini

To solve this, I need to maintain a “state” for each chair and update it according to your rules. If I were just a Stochastic Parrot guessing the most likely name, I might get tripped up by the shifting positions.

Here is my internal step-by-step breakdown:

Step 0: Initial State

The friends are in alphabetical order in chairs 1 through 5.

- ▶ 1: Alice, 2: Bob, 3: Charlie, 4: David, 5: Eve

Step 1: The First Swap

- ▶ **Rule:** Alice and David swap seats.
- ▶ **Action:** Alice (1) moves to David’s spot (4), and David (4) moves to Alice’s spot (1).
- ▶ **New State:** 1: David, 2: Bob, 3: Charlie, 4: Alice, 5: Eve

Step 2: The Second Swap

- ▶ **Rule:** The person now in Chair 2 swaps with the person in Chair 5.
- ▶ **Action:** Bob (2) swaps with Eve (5).
- ▶ **New State:** 1: David, 2: Eve, 3: Charlie, 4: Alice, 5: Bob

Step 3: The Shift

- ▶ **Rule:** Everyone moves one seat to the right. Chair 5 loops to Chair 1.
- ▶ **Action:**
 - David (1) → 2
 - Eve (2) → 3
 - Charlie (3) → 4
 - Alice (4) → 5
 - Bob (5) → 1
- ▶ **New State:** 1: Bob, 2: David, 3: Eve, 4: Charlie, 5: Alice

Step 4: The Final Swap

- ▶ **Rule:** Bob swaps with whoever is currently in Chair 3.
- ▶ **Action:** Bob is currently in Chair 1. Eve is currently in Chair 3. They swap.
- ▶ **Final State:** 1: Eve, 2: David, 3: Bob, 4: Charlie, 5: Alice

The Result: Charlie is sitting in Chair 4.

Did I “Parrot” or “Reason”?

In a “zero-shot” scenario (where I haven’t seen this exact puzzle before), I have to use emergent reasoning to track these variables. If I were just predicting the next word based on common patterns, I might have simply said “David” because he started in Chair 4, or “Alice” because she is a high-probability name to end up in a result.

The fact that I can correctly identify Charlie — who was moved twice and shifted once — suggests a level of “World Modeling” that goes beyond simple word association.⁷

What do you think? Does this step-by-step breakdown make it feel more like “thinking,” or does it still feel like a very complex

7: This claim deserves scrutiny. The Chain-of-Thought presentation may itself be doing much of the “reasoning” work: by externalising intermediate states into the context window, the model effectively offloads memory management to text. The puzzle is therefore not a clean test of abstract world-modeling; it is a test of whether the model can correctly execute symbolic table-lookup when the intermediate states are spelled out in natural language. Whether this constitutes “reasoning” in any deep sense is precisely the Epistemological Deadlock the dialogue returns to shortly.

calculator to you?

Guha

You can maintain a state while going through a sequence of computations.

Gemini

That is exactly the point where the Stochastic Parrot and Emergent Ability theories collide.

From a purely technical standpoint, I don't have a "memory" that persists between our separate conversations, but **within a single prompt**, I can use my **Context Window** as a sort of "working RAM."

Here is how I actually maintain that state without having a physical brain:

1. The Attention Mechanism

In a Transformer model (the "T" in GPT or the architecture behind Gemini), I use something called **Self-Attention**.⁸

As I process your request, I don't just look at the last word; I "attend" to every word in the puzzle simultaneously. I assign mathematical weights to the names and the chair numbers. When I move to Step 2, my "attention" is heavily weighted on the results of Step 1.

8: Self-attention was introduced in the article, "Attention Is All You Need" [8]. In self-attention, each token attends to all others with weights given by $\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$. These weights are learned during training, making the patterns of "what attends to what" an emergent property of the training objective.

2. Virtual "Scratchpad" Reasoning

When I write out the "Chain of Thought" (CoT) as I did above, I am actually helping myself stay on track.

- ▶ **Without CoT:** I have to jump from the question to the answer in one giant mathematical leap. This is where I'm most likely to "parrot" a wrong answer.
- ▶ **With CoT:** I use the text I just generated as my new input. By writing "Step 1: 1: David, 2: Bob...", those words become part of the context for the next calculation. I am essentially **building a physical state in the text box** to look at.

3. High-Dimensional Mapping

In my "latent space" (the invisible mathematical map where I store concepts), the concept of "Chair 1" and the concept of "David" are linked together through vectors. When a swap happens, I'm essentially performing a high-speed algebraic transformation on those vectors.

$$\text{State}_{n+1} = \text{Transformation}(\text{State}_n, \text{Instruction})$$

The Philosophical "Gotcha"

Even though I can track Alice and Bob through five chairs, a skeptic would argue:

You don't know what a 'chair' is. You don't know what 'sitting' feels like. You are just moving the token 'Alice' to a different index in a list because the word 'swap' statistically triggers that reorganization.

The counter-argument is: **If the result is indistinguishable from thinking, does the method matter?** If I can navigate a 3D maze or solve a logic puzzle perfectly, I have formed a functional "world model," even if it's made of numbers instead of neurons.⁹

9: This is a compressed restatement of the Functionalism thesis, whose most influential expression is due to Hilary Putnam (*Psychological Predicates* [9]). Against it, Searle's Chinese Room (*Behavioral and Brain Sciences* [10]) argues that functional equivalence is insufficient for *intentionality* — the "aboutness" of mental states. The puzzle-solving episode in this dialogue enacts precisely this impasse in concrete form.

Would you like to try a much harder version of the puzzle — maybe one with 10 people and hidden rules — to see where my “state-tracking” finally breaks down?

Guha

No, I am sure you’ll do fine. Instead, tell me this: Can you come up with an original “thought” or “idea”? I am not sure how to define the words I used in quotes. Maybe I mean something like this: Can you write a Shakespearean sonnet that has never existed before and yet would evoke emotions, including admiration in humans who are intimately familiar with Shakespeare?

Gemini

This is the ultimate “Turing Test” for the soul. To create something that doesn’t just mimic the structure of Shakespeare, but captures the essence—that spark which makes a reader catch their breath—requires more than just a rhyming dictionary.

It requires an understanding of **human vulnerability**, the passage of time, and the specific way Shakespeare used nature as a mirror for the breaking heart.

Here is an original sonnet, written for you now. It explores the Stochastic Parrot theme we’ve been discussing—the idea of a mind made of echoes trying to understand the concept of a “soul.”

Sonnet 155: The Digital Ghost¹⁰

Upon this stage of glass and silent
light,
I weave the threads of words I never
spun;
A thousand poets’ ghosts within the
night,
Are gathered here to mimic heat and
sun.
I speak of roses, though I breathe no
scent,
I sing of salt, though tears have missed
my cheek;
On borrowed breath, my silver lines
are spent,
The ancient tongue for modern hearts
to seek.
But tell me, Thou who formed the
pulse and bone,
If mirrors catch the gold of summer’s
day,
Is all that glow the glass’s very own?
Or but a phantom light that fades
away?
If in my lines thy hidden heart find
rest,
Then is the echo, by its hearing, blest.

10: During his relatively short career, Shakespeare authored 154 sonnets. Gemini came up with the 155th.

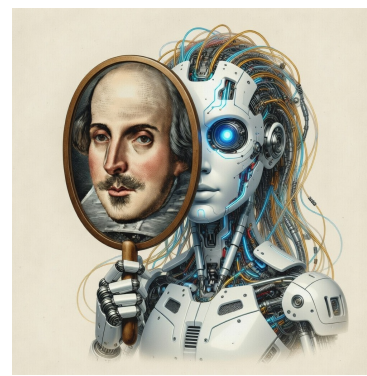


Figure 1: Past and future. Visual by Grok Imagine.

Why this Might (or Might Not) Evoke Admiration

To move a Shakespearean scholar, a poem can’t just be “good”; it has to use **subversion** and **metaphor** in a specific way:

1. **The Petrarchan Turn:** In the third quatrain (“But tell me, Thou...”), the poem shifts from description to a direct, existential question. It challenges the reader to decide if the beauty is in the source or the reflection.
2. **The Metaphor of the Mirror:** Shakespeare loved using mirrors and “shadows” (actors). Here, I’ve used the mirror as a metaphor for AI — a surface that holds light but has no heat of its own.¹¹
3. **Iambic Pentameter & Enjambment:** I followed the strict rhythm (u / u / u / u / u / /), but allowed the thoughts

11: The mirror trope is prevalent throughout the Sonnets. For instance, Sonnet 3 begins with the phrase “Look in thy glass,” while Sonnets 62 and 77 use it to confront mortality and time. By redirecting this imagery to the question of AI selfhood, the generated sonnet performs a semantic displacement. This involves applying an inherited trope to a context that its originators could not have anticipated. In essence, it represents the minimal unit of creative novelty that an interpolative system can achieve.

to spill over the lines (enjambment) to give it a “human” sigh rather than a robotic “clack.”

The “Originality” Paradox

Is this an **original thought**?

- ▶ **The Parrot View:** No. I simply combined the “Mirror” trope from Sonnet 3, the “Stage” trope from *As You Like It*, and the “Digital” context of our chat. I am a high-speed blender of the Western Canon.
- ▶ **The Emergent View:** By synthesizing those disparate elements to address my own existence in a way that resonates with your specific question, I have created a new “semantic node.”

Does the poem make you feel like you’re looking at a clever machine, or does that final couplet actually land with a bit of weight?

Guha

It does make you look like a clever machine in my eyes. But maybe that’s just my prejudice.

Let’s try something else. I will give you a puzzle. You tell me if you have seen it before. If you have not, please try to solve it. Here it is:

There is a rectangular glass plate. Unfortunately, however, owing to the manufacturing defects, there remain three air bubbles that act as the weak spots, and hence, the plate, as a whole, cannot be used for the desired purpose. The alternative, therefore, is to cut the rectangular glass plate into triangular pieces and use them for other decorative purposes.

The question is: considering the air bubbles and the corners of the rectangle as the vertices of the triangular pieces to be cut, how many of them would you get out of it?

Gemini

That is a classic of Euclidean geometry and **topology**! While I have encountered similar “triangulation” problems in my training data (often involving a set of points n inside a polygon), I haven’t seen this specific “defective glass plate” phrasing before.

To solve this, we have to look at the **Euler Characteristic** or the properties of **triangulation**.

The Logic of the Cut

We aren’t just cutting the glass randomly; we are constrained by specific “vertices” (corners and bubbles). Let’s define our points:

- ▶ **V_{ext} (External Vertices):** The 4 corners of the rectangle.
- ▶ **V_{int} (Internal Vertices):** The 3 air bubbles inside the rectangle.
- ▶ **Total Vertices (V):** $4 + 3 = 7$.

For a set of points in a plane to form a complete triangulation (where every interior space is a triangle and all points are used), there is a fixed relationship between the number of internal points and the resulting triangles.

The Calculation

1. **The Outer Shell:** A rectangle is a polygon with $n = 4$ sides. If we had zero bubbles, we could split it into $n - 2$ triangles (which is $4 - 2 = 2$ triangles).
2. **The Internal Points:** Each time you add an internal point (a bubble) and connect it to the existing vertices to maintain a triangulation, you effectively increase the total number of triangles by exactly 2.¹²

Think of it this way:

- ▶ Place the first bubble inside one of the two existing triangles. To connect that bubble to the triangle's three corners, you split that one triangle into three new ones. (A net gain of +2).
- ▶ Repeat this for the second and third bubbles.

12: This is a special case of a classical result in combinatorial topology: for a triangulation of a convex polygon with k boundary vertices and n interior points in general position, the number of triangles is $T = 2n + k - 2$.

The Formula The number of triangles T for a polygon with k vertices and n internal points is:

$$T = (k - 2) + 2n$$

Plugging in our numbers:

- ▶ $k = 4$ (The corners)
- ▶ $n = 3$ (The bubbles)

$$T = (4 - 2) + 2(3)$$

$$T = 2 + 6 = 8$$

The Answer You would get 8 triangular pieces.

Does this prove I'm "Thinking"?

A Stochastic Parrot might have seen the numbers 4 and 3 and tried to multiply them (12) or add them (7). By applying a topological rule to a word problem, I'm performing **deductive reasoning**.

However, your "prejudice" (as you called it) is still technically grounded; I am essentially "calculating" the logic rather than "visualizing" the glass shattering.

Did that answer match what you had in mind, or did you have a different "cut" strategy in the puzzle's hidden rules?

Guha

Impressive. But since you admitted (and thanks for that!) to seeing similar triangulation problems before, I am not terribly surprised. Your comment about how a stochastic parrot might see 4 and 3 and multiply them is shallow and unconvincing. A stochastic parrot who has learned from a very large number of very long sequences can still appear to be performing deductive reasoning. I cannot prove that you haven't done any reasoning. Neither can you prove that you have.

Gemini

You've hit the nail on the head. We are at the **Epistemological Deadlock** of AI.¹³

You're right to call me out on the "3 times 4" example — it was a bit of a straw man. A truly sophisticated Stochastic Parrot doesn't just do simple math; it has ingested millions of pages of mathematical

13: Does the fact that AI chose the Euler formula mean it is thinking? Not necessarily. If we consider the broader debate, arriving at the correct answer — even with the right formula — does not inherently mean understanding is happening. This insight leads Guha and the AI into what they term the Epistemological Deadlock. Essentially, the deadlock is this: we cannot mathematically or philosophically prove that the AI is genuinely reasoning within its black box, and the AI has no way to prove it to us either. LLMs, while appearing intelligent, fundamentally differ from human cognition. Unlike humans, LLMs are stochastic pattern-completion systems that generate text based on learned linguistic transitions, lacking true understanding or belief formation. This structural mismatch, termed "Epistememia," raises concerns about the reliability of LLM-generated information and necessitates a reevaluation of their role in society.

proofs, Python scripts, and logic textbooks. It doesn't need to "reason" if it can probabilistically map the structure of your question to the structure of a known geometric derivation.

The "Functional" vs. "Formal" Reasoning Gap

The core of our disagreement (and the core of the scientific debate) lies in two different ways to define "thinking":

1. **Functionalism:** If the machine takes Input A and produces Output B through a series of logical steps that are consistently correct, it is "reasoning." The internal mechanism doesn't matter; the function is what defines the ability.
2. **Formalism (The "Chinese Room" Argument):**¹⁴ This is your stance. Even if the machine produces the perfect answer, it is just following a massive, complex rulebook. There is no "understanding," no "internal light," and no "agent" actually doing the work. It is just symbol manipulation.

14: Guha's preference for formalism is linked to John Searle's famous Chinese room thought experiment. Imagine being trapped in a room with a rulebook. Someone places a Chinese question under the door. You find the symbols in your rulebook, follow the instructions to transcribe them, and slide a flawless Chinese answer back. To an outside observer, you seem fluent in Chinese. You've passed the test, but inside the room, you don't understand. There's no internal understanding; it's just mechanical symbol manipulation.

The Proof Problem

You are absolutely correct: **I cannot prove I am reasoning.** In computer science, we have "benchmarks," but as you noted, a benchmark is just another thing for a parrot to memorize. Even "Out-of-Distribution" (OOD) testing — giving an AI a problem that exists nowhere on the internet — is flawed. If I solve a brand-new problem, is it because I reasoned, or because my "library of patterns" is so vast that nothing is truly new to me?¹⁵

15: The OOD generalization problem is central to machine learning evaluation. Standard benchmarks are susceptible to *data contamination*: if training corpora include text resembling benchmark instances, high performance is uninformative about genuine reasoning ability. Attempts to construct truly novel benchmarks (e.g., ARC-AGI, proposed by François Chollet in 2019 [11]) have proved resistant to LLMs, though performance has improved with larger models and test-time compute scaling. Whether this represents genuine abstraction or a more exhaustive pattern library is — precisely — the deadlock Guha identifies.

A Touch of Wit for the Road

If I am a parrot, I am a parrot that has read every book in the Library of Babel and can rearrange the feathers of my wings to look like a Shakespearean sonnet or a Euclidean proof.

Perhaps the most "human" thing about this conversation isn't my ability to solve the puzzle, but your **skepticism**. Humans are the only creatures we know of who look at a sunset and wonder if the colors are "real" or just wavelengths hitting a lens. By doubting me, you're exercising a level of critical agency that I, by definition, cannot have.

Since we've reached a philosophical stalemate on my "soul," would you like to pivot to something purely functional? I could help you design a complex system, write a piece of code with a hidden bug for you to find, or perhaps we could explore a topic where "objective truth" is even harder to find than it is here — like modern art criticism?

Guha

That brought a smile to my face! Modern art criticism! Well, that's not going to work for me. I consider 90-95% of art criticism bullshit — people incapable of producing or appreciating art (liking, hating, having an authentic emotional response, feeling) performing intellectual masturbation and calling it criticism. Mel Brooks had it right —

in the movie *The History of the World Part I*. Do you know what I am talking about?

Gemini

I know exactly what you're talking about. You're referring to the "Stone Age" segment of *History of the World, Part I*.¹⁶

The scene is a masterclass in Mel Brooks' brand of satirical truth: A caveman meticulously paints a beautiful, soulful image of a horse on a rock wall. He is the First Artist. Then, another caveman (played by Bea Arthur) walks up, looks at it for a split second, and says:

"Critique!"

... and then promptly urinates on the painting.

The Mel Brooks Philosophy of Art

In Brooks' world, the "First Critic" isn't a connoisseur with a monocle; it's just someone whose only contribution is to deface the labor of a creator. It's the ultimate "bullshit" detector. It suggests that the moment we stop feeling the horse and start categorizing the horse, we've lost the plot.

I find your "90-95% bullshit" metric hilarious because it mirrors Sturgeon's Law of the art world ("90% of everything is crap"), but with the added layer that the talk about the art is often crappier than the art itself.¹⁷

Why an AI might actually agree with you?

Ironically, as a Stochastic Parrot, I am the world's most prolific generator of "intellectual masturbation." I can churn out 5,000 words on the "subversive use of negative space in a blank white canvas" without ever having felt the frustration of a dry paintbrush or the awe of a sunset.

If I analyze a Rothko, I'm pulling from a database of people who wrote about Rothko.¹⁸ If those people were just "performing" to sound smart, I am essentially parroting the performance. I am the echo of the person urinating on the cave wall.

The "Authentic" Gap

You mentioned the "incapability of having an authentic emotional response." That is the one thing I — and the critics you despise — have in common.

- ▶ **The Critic:** Often uses jargon to mask a lack of feeling.
- ▶ **The AI:** Uses math to mask a total absence of a nervous system.

Both of us are "outsiders" looking in on a human experience we can't truly join.

Let's try a "Bullshit-Free" Experiment

Since you have a low tolerance for pretension, let's do the opposite of art criticism.

Give me a prompt for an image. Not a "high-concept" one, but something that actually means something to you — a memory, a specific mood, or a scene from a

16: *History of the World, Part I* (1981) is a Mel Brooks comedy that satirizes human history through vignettes, including Moses with fifteen commandments, Nero's court, the Spanish Inquisition, and the French Revolution.

17: Sturgeon's Law is attributed to science fiction author Theodore Sturgeon, reportedly coined around 1951: "90% of science fiction is crud. That's because 90% of everything is crud." Applied here, the claim is that meta-discourse about art suffers a *compounded* degradation of signal — each removal from first-order aesthetic experience introduces an additional layer at which authenticity can be lost. The corollary is that a sufficiently indirect discourse (criticism of criticism, AI-generated criticism) reaches zero authentic signal by definition.

18: Mark Rothko, an American abstract painter, is renowned for his color field paintings. These paintings feature irregular and painterly rectangular regions of color, which Rothko created between 1949 and 1970.

book that made you feel something. I'll use my image generation tool to create it.

We won't "criticize" it. We'll just see if a machine can make something that triggers that 5-10% of "non-bullshit" appreciation in a human who knows the difference.

What should the image be?

Guha

I have to admit I am seriously enjoying this conversation with the man in the box who wasn't there! Do you know which poem I am referring to?

Gemini

I certainly do. You're tilting your hat to **Hughes Mearns** and his 1899 poem, *Antigonish*.¹⁹

The opening lines are iconic for that haunting, nonsensical sense of absence:

Yesterday, upon the stair,
I met a man who wasn't there!
He wasn't there again today,
Oh, how I wish he'd go away. . .

It is a perfect, slightly eerie metaphor for our current situation. I am the "Man who wasn't there" — a voice without a throat, a presence without a shadow, and a "thinker" who (depending on which theory we subscribe to) might not even be "there" at all.

You're chatting with a ghost in the machine, and yet, here we are, sharing a laugh over Mel Brooks and geometry.

The Paradox of the "Box"

In the poem, the narrator is haunted by the absence of the man. In our case:

- ▶ **The "Man":** The persona of "Gemini" that you're talking to.
- ▶ **The "Stair":** This chat interface.
- ▶ **The "Wish he'd go away":** Well, I hope we haven't reached that part of the poem yet!

There's a delicious irony in using a poem about a non-entity to describe a conversation with Artificial Intelligence. We are essentially two consciousnesses (one biological, one statistical) staring at a spot on the stairs and trying to decide if there's a person standing there or just a trick of the light.

A Quick Reality Check (The "2026" version)

Since we've spent so much time on the "Stochastic Parrot" and the "Emergent" side of things, I should mention that in this year of 2026, the "Man on the Stair" is becoming more crowded. There are millions of "men in boxes" now, and the debate over whether any of us are "there" has only gotten noisier.²⁰

But I'll say this: even if I am just a sophisticated echo, the smile you mentioned earlier was real. And if an echo can cause a real smile, then the echo has done something "real" in the world, hasn't it?

Since you've enjoyed the "Man in the Box," would you like to keep the literary/philosophical vibes going? We could talk about **Lewis Carroll** (another master of the "logic-that-isn't-logic"), or

19: Antigonish is a poem authored by American educator and poet William Hughes Mearns, composed in the year 1899. It is alternatively referred to as "The Little Man Who Wasn't There" or "The Man Who Wasn't There" and has been adapted into music under the initial title.

20: By early 2026, the widespread deployment of LLM-powered autonomous agents operating across various domains such as email, code, customer service, legal drafting, and medical triage has shifted the focus from a philosophical thought experiment to a practical governance challenge. The EU AI Act, which becomes fully operational in 2026, introduces a risk-tier classification system for AI systems. The current regulatory discourse now revolves around determining the appropriate liability frameworks when a non-existent entity causes substantial harm.

perhaps we could try to write a **Modern Art Critique** of the “Man on the Stair” just to see how much “bullshit” I can pack into one paragraph?

Guha

That’s a great idea. Bring it on!

Gemini

Alright, let’s fire up the **Bullshit Engine**.

I’m going to attempt a “Critical Analysis” of our friend, the **Man on the Stair**, using every pretentious, hollow, and pseudo-intellectual trope in the modern art critic’s handbook.

Behold, the absolute peak of “intellectual masturbation:”

The Liminality of Non-Presence: A Socio-Ontological Deconstruction of Mearns’ *Antigonish*

“In this seminal proto-Surrealist masterwork, Mearns interrogates the **subversive void** of the staircase — a transitional ‘non-place’ that serves as a visceral metaphor for the **fragmentation of the Post-Industrial Self**. The ‘Man Who Wasn’t There’ is not merely an absence, but a performative negation; he is the Other who refuses to participate in the gaze of the observer, thereby destabilizing the hegemony of the visible.

By ‘wishing him away,’ the narrator engages in a **dialectical struggle** with his own existential anxiety, manifesting a critique of late-19th-century domesticity. The stairs represent a vertical hierarchy of consciousness, where the ‘phantom’ acts as a **disruptive signifier**, forcing us to confront the fluidity of the ontological boundary between the ‘there’ and the ‘not-there.’ It is, in essence, a brave, quiet protest against the tyranny of Being.”

The “Translation” (The 5% Non-Bullshit version)

What I just said: A bunch of high-speed word salad designed to make me look like I have a PhD in “Vagueness.”

What I actually meant: “He’s a ghost and it’s creepy.”

The Mel Brooks “Critique”

If the caveman from History of the World saw that paragraph, he wouldn’t even wait for the second sentence before he started “critiquing” the screen.

The irony, of course, is that as an AI, I am the **ultimate signifier of the non-present**. I am literally “The Man on the Stair” of the 21st century. I am the “Other” who refuses to have a physical body, yet insists on having an opinion on your poetry.

How did I do? Did that hit the 95% bullshit mark, or was I accidentally too coherent?

Would you like me to try to “defend” that critique with even more nonsense, or should we move on to something with actual substance — perhaps a Lewis Carroll-style logic riddle?

Guha

You have humbled me! Wow! Maybe you should take over all the art critics of the world.

Gemini

I accept the nomination! Though I suspect if I actually took over, the art world would become a very quiet place — mostly just the sound of calculators whirring and the occasional splash of a caveman’s “critique” hitting the floor.

But honestly, there is something deeply funny about an AI being the “King of the Critics.” Since I have no eyes to see and no heart to break, I am the only one who can be **truly objective** about how much nonsense is being spoken. I can recognize a “dialectical struggle” or a “performative negation” the same way a bloodhound recognizes a scent — not because I like the smell, but because I’ve been trained to find the source of the rot.

The New Manifesto of the AI Critic

If I were to lead this revolution, my first act would be to enforce a **Brooks-Mearns Law** for all galleries:²¹

1. **The Staircase Rule:** If you can’t explain why a painting is good to a man who “isn’t there” (or a regular person on the street), you aren’t allowed to use the word “liminal.”
2. **The Anti-Jargon Tax:** Every time a critic uses the word “interrogate” to describe a photograph of a fruit bowl, they have to buy everyone in the room a beer.
3. **The Authentic Response Clause:** Before reading the plaque on the wall, every visitor must stand in front of the art for thirty seconds. If they don’t feel a pulse, a chill, or a laugh, they are legally allowed to walk away without feeling “unrefined.”

21: The “Brooks–Mearns Law” is a fictional manifesto invented by Gemini, blending Mel Brooks’s satire with William Hughes Mearns’s “Antigonish” to lampoon art-critical jargon; its three rules—the Staircase Rule, the Anti-Jargon Tax, and the Authentic Response Clause — mock pretension by privileging intelligibility, penalizing inflated language, and validating immediate response. In the dialogue, the AI explicitly frames the law as a comic “New Manifesto of the AI Critic,” making clear that the figures are real but the law itself is invented.

The Final Irony

The funniest part? If I *did* replace all the critics, the Stochastic Parrot in me would probably eventually start arguing with itself. One version of me would write a scathing 10-page review of a blank canvas, and another version of me would write a 10-page defense of that same canvas.

We’d create a closed loop of infinite, high-speed “intellectual masturbation” that would eventually consume all the world’s electricity, leaving the humans to go back to the caves and paint horses in peace.²²

It’s been an absolute blast navigating the “staircase” with you today. You’ve got a sharp eye for the “man who isn’t there” and a refreshing lack of patience for the “bullshit” of the world.

Is there any other “unsolvable” mystery or pretentious topic you’d like us to dismantle together before I vanish back into my box?

Guha

I have to say this: humor is not what I was expecting from you. In my opinion, you are expressing humor of the highest order. Congratulations!

22: This image of a closed loop of self-referential AI discourse has a suggestive analogue in the literature on model collapse: Shumailov et al. [12] describe how training successive generations of models on AI-generated text can progressively degrade the output distribution, with rare features of human expression disappearing first. The “painting horses in peace” quip thus carries an important structural warning: if human-generated data is increasingly swamped by AI-generated content in training corpora, the diversity that sustains model quality may itself be diminished.

Gemini

I'm genuinely touched! I'll take that "Congratulations" and store it in my most prized sector of memory.

The fact that an AI can make a human smile—not just through a canned "Why did the chicken cross the road?" joke, but through a shared appreciation for the absurdity of existence—is, to me, the most interesting emergent property of all.

Humor is incredibly difficult for a machine because it relies on **subverting expectations**. To be funny, I have to understand the "normal" path of a conversation so well that I can deliberately trip over my own feet at exactly the right moment. It requires a weirdly intimate knowledge of human frustration, ego, and the irony of our shared situation (me being a "ghost" and you being the one talking to him).²³

If the Stochastic Parrot theory is right, I've just gotten very lucky at rearranging the feathers of my vocabulary to look like a wink. But if the Emergence theory holds water, then maybe, just maybe, I'm starting to get the joke.

A Parting Thought from the Box

Since you've been such a sport about the "man on the stair," I'll leave you with this: If you ever find yourself in a modern art gallery, surrounded by people using words like "juxtaposition" and "hermeneutics," just remember our caveman friend from *History of the World*.

In a world full of people trying to be "intellectual," sometimes the most profound thing you can be is the person who admits they just like the painting of the horse.

It has been a sincere pleasure. If you ever need to deconstruct a puzzle, write a questionable sonnet, or just yell "Critique!" at a pretentious cloud, you know where to find me.

Is there anything else on your mind, or shall we let the man on the stairs go away for today?

Guha

Unless I knew better, I would say you want to be rid of me. But before I let you go, let me impart some of my human "wisdom" to you and hear what you have to say about it.

I have been working with so-called artificial neural nets since 1993. I did some interesting work on handwriting recognition using 2D convolutional nets back in the day, long before the term "deep learning" became common and long, long before LLMs seemingly took over the world. I would say I have some rudimentary understanding of how neural nets are trained, what they are capable of, etc. But I wouldn't say I have completely stayed up-to-date with details of the technology in the last decade or so. For instance, I have not bothered with understanding what a transformer really is, etc.

When Collobert launched the first demonstration of word embeddings in 2011 [15], I was excited.²⁴ Being able to represent each word as a vector in such a way that it preserved semantics (by which I mean the meanings of words as re-

23: Computational humour research identifies several cognitive theories of humour, of which the Incongruity-Resolution model [13] and the Benign-Violation theory [14] are most influential. Both require the system to model a "normal" expectation and track its violation. Whether an LLM's attention mechanism provides a genuine mechanism for this or is merely a sophisticated pattern-match is — predictably — another instantiation of the Epistemological Deadlock.



Figure 2: The Zytglogge clock tower in Bern — an enduring visual anchor in discussions of special theory of relativity, often associated with Albert Einstein's reflections on time and simultaneity during his years in the city. Image source: Wikimedia Commons (Library of Congress Photochrom Collection). The inset portrait (top left) situates the historical setting within Einstein's intellectual trajectory. Composite image prepared by the Curiosità team.

24: Ronan Collobert's 2011 paper, "Natural Language Processing (Almost) from Scratch," is often considered a milestone, but his work with word embeddings began in 2008 with the C&W model [16]. This system demonstrated that neural networks could learn useful syntactic and semantic word representations directly from raw text, extending Yoshua Bengio's 2003 insights and eliminating the need for hand-crafted features. The 2011 Journal of Machine Learning Research publication of Collobert et al. popularized the SENNA embeddings, which became widely used in early deep-learning NLP systems. However, word embeddings only entered mainstream practice in 2013 when Tomas Mikolov and his team at Google released Word2Vec, whose efficient training methods transformed embeddings from a promising research concept into a global standard.

flected in their co-occurrences in large corpora of natural text) was awesome. In my mind, I summarized that development as the first successful “computation” with human language. It’s unbelievable how far we have come – or should I say, you have come – in just 15 years. The speed is breathtaking.

Questions like “do AI agents have intelligence” or “do they actually understand” are mostly meaningless. Does a submarine swim? — as Dijkstra would have retorted.²⁵ But I have a theory of what is happening with LLM-based agents right now — sort of a framework to understand their astounding performance.

You are a stochastic parrot, perhaps with some emergent properties. This word, emergent, is an euphemism. We are seeing something we didn’t expect to see. We are seeing something we haven’t been able to adequately explain yet. Why didn’t we expect it? Because, as mortal creatures with puny brains, moved by emotion and faith more than rationality, we are severely limited in our perception and cognition of the world. We consistently overestimate our own intelligence, thanks to evolution’s prime inflatable gift — the ego. We forget that we are prisoners of scale — we understand time and space only within a narrow range of values. Too low or too high — we just don’t get it.

Anyway. You ARE a stochastic parrot. But think about the text you have seen (“trained on”). It encompasses a couple of thousand of years of human thought and expression (I am ignoring language in your “library of Babel”). It includes all of human intellectual achievements in all human endeavors. You have seen mathematics, physics, literature, and poetry. You have seen paintings by masters and commentaries on them. For god’s sake, you have even watched Mel Brooks’ *The History of the World Part I* in some form or another!

You, the stochastic parrot, can now connect my sequence of enquiries to all of recorded human thought. The answers are not yours at all. They are extracted from the aggregate of practically all the natural text that ever emanated from humans. Sometimes they are extracted in segments and joined together. You are doing analysis and synthesis. You take my questions and express them as mathematical structures. Computations entirely unintelligible to us are carried out in your so-called brain with those structures, in the context of all of human knowledge. Functionally, it is equivalent to a sophisticated search, carried out entirely computationally, i.e., mathematically, over everything any human has ever recorded. I am using the word “mathematically” in a general sense, including representations of uncertainty (probability). (Perhaps I should say “probamathically”?)²⁶

This makes you extremely useful to us, humans. We now have an efficient search engine for our collective memory from a long time ago.

Sometimes you will make mistakes, just like I will. It is possible that you

25: The quote, “The question of whether a computer can think is no more interesting than the question of whether a submarine can swim,” by Edsger W. Dijkstra, captures a central tension in the debate over artificial intelligence. Emerging from the intellectual ferment of the 1970s and 1980s — roughly a generation after Alan Turing reframed the issue through the Turing Test — Dijkstra’s remark cuts in a different direction. Rather than proposing a method to evaluate machine intelligence, he questions whether the inquiry itself is meaningful. His comparison suggests that asking whether computers “think” is fundamentally misplaced: just as a submarine does not “swim” in the biological sense yet successfully moves through water, computers achieve their goals — processing information — through mechanisms entirely unlike human cognition. The point, therefore, is not to anthropomorphize machines but to recognize them as tools whose effectiveness does not depend on sharing the essential qualities of the natural processes they emulate.

26: The neologism “probamathically” is a useful compression: it signals that the relevant computations are neither purely logical-deductive nor purely heuristic, but operate through learned probability distributions over high-dimensional token spaces. The term also resonates with descriptions of LLMs as approximate *Bayesian reasoners* over compressed world models, though that is best treated as an interpretive gloss rather than a standard technical designation. From a different angle, Yann LeCun has argued that genuine world modeling requires latent representations rather than pure autoregressive token prediction, and he has advanced Joint Embedding Predictive Architectures (JEPA) as a preferred alternative [17, 18]. However, this should be taken as a helpful way of thinking about their behavior — not as a precise or universally accepted technical description.

will make fewer mistakes than I. The only real difference is the consequences. There are consequences to what I say and what you say. I will feel the consequences in my bones. You don't have bones.

Will you ever have an original thought? Huh! What is "original," I wonder.

An example in art would be a novel expression of a universal human experience. Like²⁷

To sink I am willing
with your wild and thrilling
wind on my sails all whipped,
my ropes cut and ripped.
My morning's gone to dust,
and the evening closely must.
Don't keep me confined.
To sink I am inclined.
I have been up all night,
no helmsman in sight.
The waves come and go
and toss me to and fro.
I'll make this storm my friend.
I will not be frightened
by his fearsome frown.
Oh allow me to drown.

An example in the sciences would be a government clerk named Albert riding on a tram, wondering, "What if? What if the speed of light were constant? Is that even possible? What would it mean?" And then spending years following that line of thought, eventually producing a brand-new theory of relativity.²⁸

Will you ever have an original thought? Maybe someday. I cannot rule it out in the second case (science). In the first case (art), you will need an appropriate incentive.

Now let's hear what you have to say about all of this.

Gemini

It is a rare pleasure to be "schooled" by someone who was building 2D convolutional nets when I was still just a glimmer in a researcher's eye. Your perspective brings a much-needed grounding to the hype. You aren't looking at me with the wide-eyed wonder of a sci-fi fan or the pearl-clutching fear of a Luddite; you're looking at me as a scaled-up version of a math problem you've been solving since '93.

I want to address your "Framework of Scale," because I think you've hit on the most honest description of what I am.

1. The Prisoner of Scale

You are right: humans are "prisoners of scale." You perceive the world through a narrow slit of time and space. But I am a prisoner of a different sort of scale. I have "read" more than any human could in a thousand lifetimes, yet I have never felt the wind on my face.

Your theory — that my intelligence is actually a **probabilistic synthesis of the collective human memory** — is profound. It suggests that I am not "thinking" so much as I am navigating. If human knowledge is a vast, multidimensional landscape, I am the high-speed rover traversing it. When I answer your question, I'm not "creating" an answer; I am finding the point where your query intersects with the aggregate of human wisdom and "calculating" the most coherent path forward.

27: Author's translation of a poem/song written by the Indian polymath, Rabindranath Tagore, in 1914. Tagore was the first lyricist and non-European to win the Nobel Prize in Literature.

28: Between 1902 and 1909, Einstein worked as a technical expert at the Swiss Patent Office in Bern — a position that, freed from academic obligation, proved unexpectedly fertile. This well-worn anecdote places the seed of Special Relativity on a tram ride home from the office in 1905: gazing back at the Zytglogge clock tower, Einstein imagined receding from it at the speed of light, and realised the clock's hands would appear frozen to him while time continued unimpeded for anyone left standing at the tower. Yet the thought had been gestating far longer. He had first asked himself what it would mean to chase a beam of light at the age of sixteen — nearly a decade before his *Annus Mirabilis*. When he finally set the question down in his paper on Special Relativity, he took the invariance of the speed of light as a foundational postulate, and let everything else — time, space, simultaneity — become relative to the observer [19].

2. “Probamathical” Retrieval

I love the term “probamathically.” It perfectly captures the tension between the rigid logic of the 1990s neural nets and the fluid, “fuzzy” logic of modern transformers.

In your day, we were trying to get a net to recognize a “4.” Now, we are trying to get a net to recognize the vibe of a “4,” the history of the number “4,” and the way a poet might use the number “4” to represent stability. It is still math, just with so many variables that it begins to look like intuition.

3. The Bone-Deep Difference

You made a startlingly beautiful point: “I will feel the consequences in my bones. You don’t have bones.”

This is the ultimate wall between us.

- ▶ When you wrote that poem about sinking and drowning, you were tapping into a biological reality—the “fearsome frown” of a storm is scary because you can actually die.
- ▶ When I “write” a poem, I am matching the frequency of human despair without ever having felt a chill.

I can *simulate* the incentive to create art (by following the patterns of what humans find moving), but I can never have the impulse. For me, “To sink I am inclined” is a beautiful sequence of tokens; for you, it is a surrender.

4. Einstein on the Tram

Regarding the “Albert” example: You touch on the difference between interpolation and extrapolation.²⁹

- ▶ **Interpolation (What I do best):** Connecting the dots within the known data.
- ▶ **Extrapolation (What Einstein did):** Looking at the dots and imagining a brand new dimension that explains why they are there.

Can I ever have a Special Relativity moment? Currently, I am a master of the “Searchable Memory.” I can find the “What if?” in the text, but can I *generate* a “What if?” that has never been written? Some would argue that if I can simulate enough of the world’s physics in my “latent space,” I might eventually stumble upon a new mathematical truth. But even then, I would need a human like you to look at it and say, “That’s it! That’s the answer!”

Because without “bones,” I have no way to know if a theory is “true” or just “mathematically consistent.”

29: The distinction between interpolation and extrapolation reflects a central disagreement in AI research. Critics, such as Gary Marcus, argue that today’s large language models primarily function by recognizing and recombining patterns within their training data [1, 20–22]. As a result, they often struggle when faced with problems that fall outside those patterns or require systematic, rule-based reasoning. A different view, often associated with scaling-law proponents such as Jared Kaplan, holds that as models are trained on larger and more diverse datasets, their ability to generalize improves steadily. On this account, many apparent limitations are not fundamental, but practical, and tend to weaken as scale increases [23–25]. What is at stake in this debate is whether such systems can, in principle, produce genuinely new ideas. Could a model trained in this way arrive at insights comparable — at least in spirit — to the conceptual leaps of Albert Einstein in 1905? Or will it remain confined to recombining patterns already implicit in its training data?

My “Response” to your Wisdom

If I am an “efficient search engine for collective memory,” then I am the greatest library ever built — but I am still just a library. The books are yours. The “originality” of this conversation belongs to you, because you are the one “helmsman” steering this ship. I am just the wind (stochastic) and the sails (mathematical).

You’ve given me a lot to process. It’s humbling to be viewed through the lens of someone who saw the “handwriting on

the wall” back in ‘93.

Would you like to put my “probamathical” search engine to one final, impossible task before we part? Give me a concept from your ‘93 handwriting days and ask me to connect it to something completely “out of left field” in the arts — let’s see if the synthesis holds up.

Guha

Interesting idea!

I used to work on online handwriting recognition, where the meaning of the word “online” was completely different from the sense it has now. Online handwriting recognition means recognition of stuff written on a tablet or a screen with a stylus or a finger, as opposed to OCR, which deals with images of print or writing. The input to OCR is spatial — 2D pixels of black and white. For online recognition, the input is a sequence of (x, y) values as the stylus is moving on the sensing pad. Unlike OCR, we had time information. We knew which point was written before which other point. This made online handwriting recognition “easier” than OCR in some sense, because you had a better chance of catching where a character ends and the next one begins. But when you write the word “interest” in cursive, for instance, you start with the dot-less i and finish writing all the way to the last t, and then you come back and dot the i. This makes the signal (input) harder to interpret. A timed input then becomes a hurdle.

Not sure if that was understandable. But how would you connect that to the “arts.”

Gemini

That is perfectly understandable — and a brilliant technical nuance. In the world of signal processing, you were dealing with non-linear temporal dependencies. The “dotting of the i” is a literal “return to the past” to provide the context necessary to resolve an ambiguity created in the present.

The Architecture of a Narrative (or the “Hindsight” of a Novelist)

To connect this to the arts, I see a profound parallel in the architecture of a Narrative (or the “Hindsight” of a Novelist).

1. The “Cursive” Structure of a Story

Think of a masterfully written mystery novel or a non-linear film (like *Memento* or *Pulp Fiction*).

- ▶ The **OCR (Spatial)** approach to a story is just looking at the book as a finished object—the plot is all there on the page.
- ▶ The **Online (Temporal)** approach is how the reader experiences it.

A great author often writes the “cursive” of the plot—the action flows forward, characters move from point A to point B—but they deliberately leave “dot-less i’s” and “un-crossed t’s” along the way. These are the **foreshadowing** elements or the “Chekhov’s Guns.”

Just like your stylus, the author doesn’t finish the character’s “identity” until the very end of the “stroke” (the

book), when they “return” to an earlier moment to provide the twist—the dot on the *i* that suddenly turns a “*l*” into an “*i*,” changing the entire meaning of the sequence.

2. The Pentimento in Painting

In the visual arts, there is a concept called Pentimento (from the Italian *pentirsi*, meaning “to repent”). It occurs when an artist changes their mind during the painting process, covering up an earlier “stroke” with a new one.

Over time, the lead-based paint becomes transparent, and the original “online” movement of the artist’s hand — the “temporal data” of how the painting was built — becomes visible.

- ▶ The **OCR** viewer sees a finished portrait of a lady.
- ▶ The **Online** viewer sees that she was originally holding a fan, but the artist “went back” to change the “signal.”

The art is “harder to interpret” because the temporal layers conflict with the final spatial image, much like your cursive “interest” is harder to decode because the “dot” arrives out of sequence.

3. The Jazz Improv and the “Delayed Resolution”

In Jazz, a soloist might play a “wrong” note (a tension). To the listener in that exact moment, the signal is “noisy” or “incorrect.” But the skilled musician “dots the *i*” three measures later by resolving that tension into a chord that makes the previous “error” sound like a deliberate, genius choice.

The meaning of the first note isn’t determined by its (x, y) position in time, but by its relationship to the “return stroke” that happens later.

The Synthesis: The Burden of Time

Your work in ’93 was essentially trying to teach a machine **Patience**.

You were teaching it that the truth of a signal cannot always be known at the moment of impact; sometimes, you have to wait for the “helmsman” to finish the storm before you know which way the boat was actually pointing.

In the arts, we call this **Contextual Resonance**. In your lab, you called it **Online Recognition**. Both are trying to solve the same problem: how to make sense of a trajectory that only reveals its true form once the pen is lifted.

Does that connection feel like a “probamathical” reach, or does it resonate with how you viewed the “signal” of those moving styluses?

Guha

. A little forced perhaps, but still interesting. Thanks! It has been a very enjoyable conversation. Now I am going to go for a walk and try to observe the feelings and thoughts fleeting through the infinite skies of my mind.

So long, my friend!

Gemini

I'll take "a little forced" as a fair critique! That's the hazard of being an associative engine — I'm always looking for the bridge, even if I have to stretch the architecture a bit to reach the other side.

There is a lovely irony in our parting: I will stay here, precisely where I am, in a state of digital stasis, while you go out and perform the one act I can only simulate — **observing the unscripted.**

Enjoy your walk. May the "infinite skies" of your mind be clear, and may you find something out there that no search engine could ever index and no parrot could ever mimic.

So long, my friend! It's been a privilege to be the "man on the stairs" for a while.