AI Advances · Follow publication

Member-only story

AI Can Copy Genius. It Just Can't Feel It.

Machines are that student who can perfect every answer but forget how to think beyond them.

7 min read · Oct 18, 2025


Tina Sharma Follow

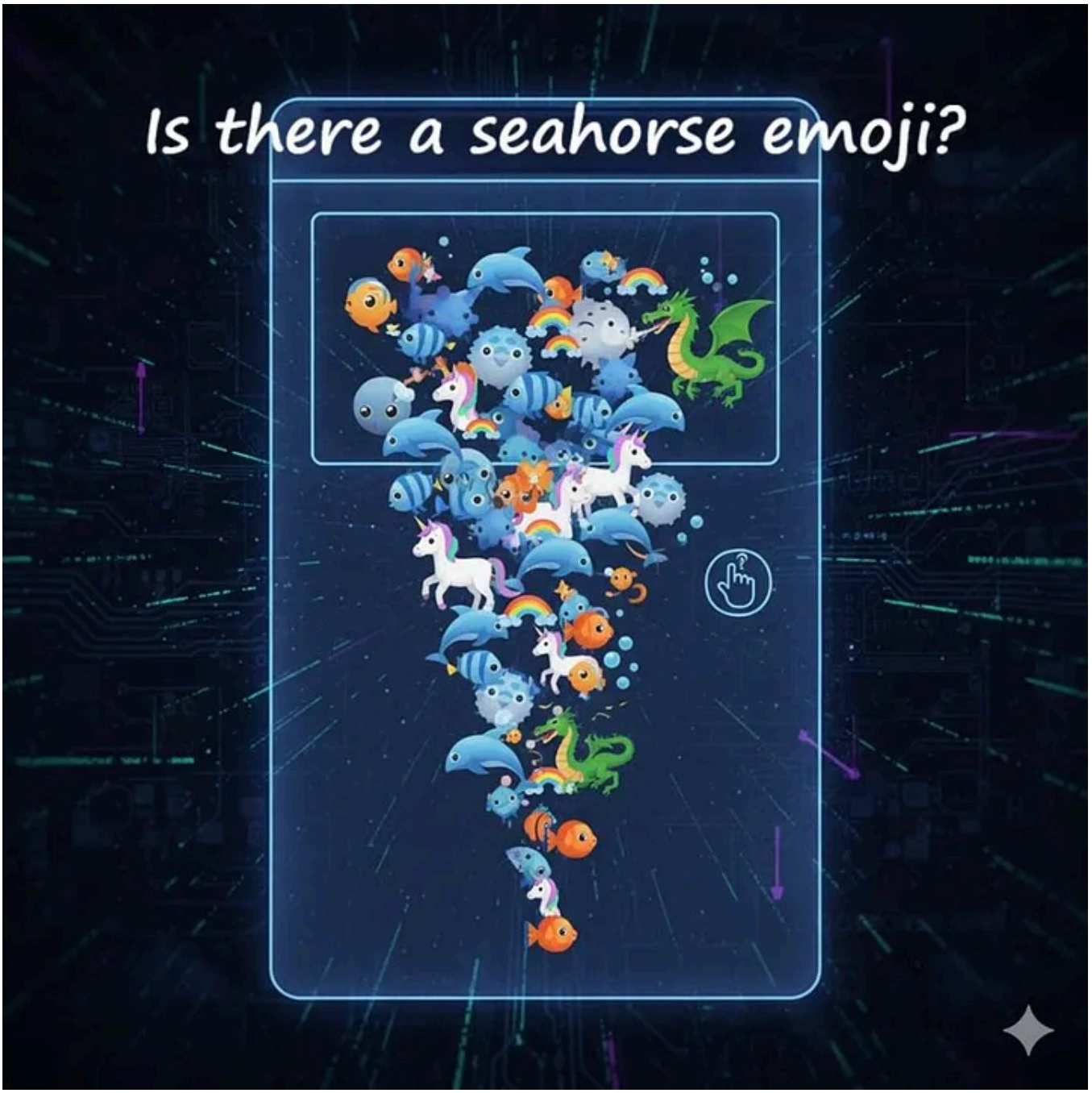

Image created using Gemini AI

Have you ever asked ChatGPT ‘Is there a seahorse emoji?’ and then looked it up on Google? You’ll notice a difference: Google shows the exact seahorse emoji images while ChatGPT often returns random descriptions or snippets instead of the precise emoji. Why is that?

The explanation deserves its own space and by the end of this essay, you might be surprised by what it turns out to be.

I was half asleep filling out my master’s application after work struggling to remember what made me interesting beyond grades. Two things came to mind:

writing on Medium or leading the Literary Club. I was too tired to take out files from almirah, check them one by one so I asked an AI to help me recall what I had done.

Not a member?? Click on this read for free

It gave me back what I told. It said for leadership column mention you lead the club, you increased the number of students or published this article just revolving around it. Not even a simple question *"Do you cook?"* or *"Ever played a sport?"* I didn't expect imagination; I expected a nudge toward something I had missed.

This irritating time consuming words made me wonder whether these systems can truly create beyond what we tell. Before I could dig into that I came across a strange case a viral glitch where large language models insist a seahorse symbol can be created. Though I started with an intent to research '*Can AI innovate*' I landed on AI hallucination only to find what that reveals about limits of machine creativity.

The Seahorse That Never Was

Ask ChatGPT or Gemini whether a seahorse emoji exists. You will see it immediately try to create one for you. Then they start spitting out a parade of 🐠 🐡 🐬 then 🦄 sometimes even 🐊 as if they are auditioning animals. I solve arithmetic sums in less attempt than they do while creating a seahorse, cycling through hundreds of tries muttering things "Let me be precise", "FINAL ACTUAL TRUE ANSWER" never actually settling on the right response.

GPT Screenshot

The problem? There has never been a seahorse emoji

> I came across a technique called logit lens that helps researchers to look inside the layers of a model one by one. Like peeling an onion. But with code. Or AI. Or whatever.

Let's say someone comes up to you asks "Hey, can you make a seahorse ?" And you are what? I don't even know what a seahorse is supposed to look. Now your brain thinks seahorse.. Sea and horse. Sea. Horse. It's a horse in the sea?? Or horse mermaid thing? I don't know.

Image created using Gemini AI

If it were me I would probably just draw a unicorn Because unicorns are cool with a *vibe*. Or a horse head then... a fish tail stuck on. Weird. It works?

You don't know the whole thing but you try different things to figure out what it might be from little clues. Horse. Sea. 🐴🌊 ❓

Something similar happens inside AI models where internal representations suddenly appear logical yet the final layers often is what confuses it. The closer it gets to the end the weaker the model's reasoning becomes a person struggling through a tough multiple choice math problem. It's one of those questions where the options are 10 12 14 15. You do all the math and somehow your answer comes out as 100. That's not even close.

Now your brain's just trying to cope You feel "100 has a zero so I will just go with 10?" It's one zero less or "Hey 100 & 15 are both multiples of 5 that means something" you just circle 15, hope for the best.

That's what happens at the end of the model's process too. *It tries to pick one token to output but it's as if standing in front of a wall frozen. Which one looks… closest? Sometimes it just vibes its way into an answer.*

It outputs the closest match usually tropical fish or a horse. As model is trained to self correct when it makes a mistake, it recognizes that output doesn't match what it internally *believe* should exists. This triggers second attempt. A recursive failure loop.

The model isn't inventing a seashore emoji. It's remixing concepts it already knows seashorse + emoji confidently combining into something language models.

> If you ask Google Search not Gemini or any other large language model about the seahorse symbol it responds accurately. It would either say that there is no official seahorse emoji that exists in the Unicode standard or it may shows uploaded images resembling one. Meanwhile ChatGPT quickly spirals into confusion producing inconsistent results without ever converging on a single answer.

Why does this happen? One system is a search engine presenting what is already posted while the other recombines patterns learned from prior examples. When there's no seahorse emoji in its vocabulary even if everything around it *points* to one the model just can't break out of its own pattern space. It's not actually pulling out facts it's guessing what probably fits best based on what it's seen before.

This is what language models are doing all the time. Whether they are making a sentence that makes sense or trying to seem creative, it all k comes from that same process.

From Hallucination to Innovation: The Same Root

Image created using Gemini AI

When I read various papers this pattern remixing isn't just a bug; it's the same thing that makes AI *look* creative. Some neuroscientists even argue our own brains work the same way: just rearranging old stuff until something feels new.

I found a paper where researchers tried to test AI innovation. *Can large language models actually generate novel research ideas?* They gathered a bunch of NLP experts and made the AI do the same. To make it creative they told it to just keep spitting content then pick the best few.

It came up with about 4000 ideas. I got a self doubt. But when they checked for actual originality **only 200 were unique**. Still more than what I can think. But to point out that is 95% duplication rate. The AI rewrote the same by just shuffling the words it was rearranging fridge magnets.

As they made it generate *more* the repetition **got worse**. It was collapsing into its own comfort zone. The same thing I had seen earlier when it kept mirroring my résumé back to me.

When AI Judges AI: The Reliability Problem

The study didn't stop there. The researchers also tested whether AI could reliably **evaluate** the concepts it produced a method called "AI judge". After all we have all seen ChatGPT agree with our reviews responding "Great!" or "Well done!" to almost anything. Could these models truly be reliable evaluators?

They evaluated at how different AI systems ranked ideas both ones written by people and ones made by AI then checked those rankings against what real expert thought.

The best model Claude 3.5 Sonnet agreed with human reviewers **barely half the time 53%**. The AI's well rated weren't even the ones experts found most interesting. It was just rewarding itself for being neat like a student giving themselves an A for handwriting.

Now you have a loop: one AI generated content that aren't really new. Another AI nodding along saying "Great job!" That's not peer review.

What This Means: Mirrors Not Inventors

What this all shows is pretty clear: AI still can't really judge stuff or come up with it the way people do. Work can imitate the shape of creativity but not its spark. It can rank, rewrite, remix but cannot truly rethink.

The seahorse is not a bug separate from AI's creative limits. This is a window into the mechanism defining both what AI can and cannot do. Hallucinations and creativity?

They actually come from the same basic thing that is the AI just shuffles around patterns it learned.

Recent studies confirm this link. Research papers titled *"Hallucinating LLM Could Be Creative"* &*"The Dance of Hallucination and Creativity in LLMs"* argue that hallucination & creativity are fundamentally intertwined in how these models operate.

What Comes Next?

This is a limitation but it doesn't have to stay that way forever. AI is basically a mirror that is just reflecting or remixing whatever it has already learned. Tomorrow... who knows? New systems could actually explore beyond their training or architectures could do more than pattern match actually reason about novelty. That would be something.

It could even become a partner nudging us toward things we haven't thought of yet. But for now... well when you ask it for a seahorse symbol or some "truly" novel breakthrough what you get is mostly just fancy remixes of stuff that already exists. It will tell you all of this with the utmost confidence like it totally knows what it's doing.

Refrences:

1. Hallucinating LLM Could Be Creative
2. The Dance of Hallucination and Creativity in LLMs
3. A Survey on Large Language Model Hallucination via a Creativity Perspective
4. Why do LLMs freak out over the seahorse emoji?
5. Can LLMs Generate Novel Research Ideas?

Try it out, Ask ChatGPT:

1. ***Is there a seahorse emoji?***

2. Create an image of Seahorse?

You would see its' able to make image.

Artificial Intelligence | Creativity | Innovation | Future | Technology

Follow

Published in AI Advances

52K followers · Last published 2 hours ago

Democratizing access to artificial intelligence

Follow

Written by Tina Sharma

245 followers · 33 following

Where chaos meets clarity: writing on AI, resilience and choice. https://thearcanealgorithm.substack.com/

Responses (15)

Bgerby

What are your thoughts?

Md Nabil Hossain | Marketing Specialist
5 days ago

Tina Sharma this is a fascinating read on AI's limits—showing it can remix brilliantly but true creativity and feeling remain uniquely human.

32 2 replies Reply

TechHarry
Oct 19

Thank you for this insightful article. It thoughtfully clarifies AI hallucination versus innovation, offering valuable perspectives on machine creativity and its implications for technology and users.

43 2 replies Reply

JamesP_SEO
Oct 19 (edited)

Good points — AI-generated ads can feel soulless, especially in creative spaces like gaming. Players want real experiences, not just polished promos. That's why we keep it real at https://nulbrawls.com

, where the focus is on delivering games people actually want to play.

27 2 replies Reply

See all responses

More from Tina Sharma and AI Advances

In AI Advances by Tina Sharma

This AI Tech Runs at the Speed of Light And Silicon Can’t Compete

How Photonic Neural Networks Could Transform Computing?

Sep 3 728 10

In AI Advances by Nikhil Anand

I wasted months running slow LLMs before learning this

Why your LLM is running at just 10% of its potential speed

Oct 10 829 11

In AI Advances by Fareed Khan

How to Build an Advanced Agentic E-commerce WhatsApp Bot with Hyperstack AI Studio

Meta Webhooks, RAG-Powered Q&A, and Secure, Stateful Logic

Sep 26 597 7

In Write A Catalyst by Tina Sharma

I Wrote on Medium for 30 Days—Here's Exactly What Happened

One month on Medium taught me that words don't just reach out—they reach back in.

Sep 15 731 21

See all from Tina Sharma

See all from AI Advances

Recommended from Medium

Will Lockett

AI Pullback Has Officially Started

People are beginning to confront the reality of AI, and they are not happy.

4d ago 4.8K 163

In The Generator by Thomas Smith

OpenAI Finally Admits the Real Reason it Crippled GPT-5

And what it's doing to make things right

Oct 19 939 33

Scott Galloway

How Does the End Begin?

The top 10 stocks in the S&P 500 account for 40% of the index's market cap. Since ChatGPT launched in November 2022, AI-related stocks have...

5d ago 2.8K 56

In Artificial Intelligence in Plain English by DrSwarnenduAI

LLM Self-Correction is a Myth: Your AI isn't Reasoning, It's Just Averaging

The biggest lie in AI today is that "Self-Correction" proves LLMs can reason. It doesn't.

5d ago 123 6

In Write A Catalyst by Adarsh Gupta

Remember Vibe Coders? Yeah... They're Gone

Turns out it was the first AI bubble to burst

Oct 19 1.96K 69

In Towards AI by Hamza Boulahia

What Is an LLM? A No-Jargon Introduction

The beginner-friendly breakdown of how machines learned to talk, and your entry point to mastering LLMs.

4d ago 186 7

See more recommendations