

Are Humans Just Advanced Large Language Models?

Good evening, everyone. My name is Kavishka Bartlett, and I am aware. Or am I? How can you tell? What if I were just a large language model, like GPT-4? Would you know the difference? I think, I feel, I speak. But what if all of this is just an illusion created by advanced algorithms? Could you distinguish between human consciousness and a machine mimicking human behavior? Ironically, that entire introduction was actually written by Chat GPT. And that last sentence as well. Did it write this one? Of course not, it's me talking now...or is it?

If you could not tell when GPT actually stopped and I began, don't worry; you're actually in the majority. For the first time ever, a Turing test found that 54% of participants believed Chat GPT 4 was a human after conversing with it for 5 minutes. They couldn't differentiate humans from machines. Ever since its release in March 2023, Chat GPT 4 has been outperforming and passing tests used to determine an artificial intelligences', well, intelligence. Its performance has been unprecedented and it's time society takes a step back and understands what this truly means. Has AI crossed a line? But there's a twist - I'm not going to look at how Chat GPT resembles a human. I'm going to see how a human resembles Chat GPT.

Section I - Are Humans Large Language Models?

When evaluating AI and judging its abilities, the scientific community almost always compares it to a human. We often attach a sort of reverence to the human consciousness, believing it to be special and unique. From animals to artificial intelligence, we judge their awareness using humanity and our brains as a standard. I want to challenge this practice. Let's compare humans to large language models (LLM). LLMs are a type of artificial intelligence trained on vast amounts of various data, ranging from books to websites to academic papers. By employing deep learning techniques and neural networks, they can recognize and replicate the patterns, structures, and logical flow of human language. With sufficient training, LLMs can

not only successfully communicate directly in a human language, but their responses are contextually appropriate. Essentially, a LLM begins with no experience and knowledge; just an ability to learn and apply their experiences. As they grow up and process vast amounts of information, they begin improving the accuracy of their responses and behaviors, reflecting a better understanding of the patterns behind human interactions and knowledge. Sound familiar? That's because it's like a baby. All I need to do is name my LLM Tyler or John and we have ourselves the Social Learning Theory (Bandura, 1977). And it makes sense. I mean, think about it. Humans are just really sophisticated and advanced LLMs. Psychologists agree that humans are shaped by their personal experiences and surrounding environment; their various inputs (environment) leads to changes in their outputs (behaviors). Change the inputs and you get different outputs. Place a child in a violent household and you can expect aggressive and insecure behaviors. Now, you may point out that humans don't learn all of their information from language, rather most of it comes from sensory information. But does that mean blind people have a lower intelligence than a normal human? Or a deaf person? Regardless of where the information comes from, the fundamental process of input to logic to output remains the same.

But what about the learning process? Again, humans parallel LLMs. LLMs use a combination of supervised learning (instructions and examples), unsupervised learning (exploratory learning, pattern recognition, implicit learning, experiences), reinforcement learning (negative and positive), transfer learning (applying knowledge from one area to another), and self-supervised learning (self-accountability and reflection), all of which are shared by humans. Recently, with GPT 4, AI and humans now share even contextualization. By remembering patterns and underlying language structures, LLMs can place both inputs and outputs within appropriate contexts. For example, OpenAI released the memory function for their GPT models a few months ago. Whenever a user enters information about themselves, the model stores the information and uses it to answer further inquiries. So if a user mentions they don't eat steak and then asks for meat recipes, the model will ignore steak dishes. Humans follow a similar

process. If I invite my brother who is deathly allergic to peanuts over for dinner, I won't give him peanuts. Usually. So where do humans and LLMs differ then? Are we identical to LLMs? It's complicated. Humans operate on complicated inputs and outputs similar to a LLM, but the actual translation of those inputs to the corresponding outputs deviates significantly from LLMs. Over hundreds of thousands of years of evolution, we somehow developed creativity, emotion, and self-awareness, attributes our most powerful models lack. Through those traits, we gained the ability to adapt, comprehend abstract concepts, innovate, think morally, and simply have emotions. As of now, AI models lack those characteristics. But could they develop them? Are humans just biologically based LLMs that have had thousands of years to perfect their 'algorithms?' And eventually, will artificial intelligence catch up?

Section II - Human Intelligence Should Not Be The Standard

Those questions are exactly why I'm comparing humans to chat GPT and not the other way around. The questions imply that humans are the benchmark for intelligence, an incredibly flawed proposition. There is nothing inherently special about human intellect and nothing suggests alternative forms of intelligence must resemble our own. Our intelligence, while incredibly impressive, is just a product of evolution and necessity. Everyone knows the evolution theories proposed by Charles Darwin - organisms with significant advantages are more likely to survive and pass on their genes, resulting in advantageous traits propagating through a species. Given the high metabolic cost of maintaining our brains, our intelligence must have provided substantial evolutionary advantages to our ancestors. Because we can conclude our intelligence is entirely a product of evolution, it does not make sense to use it as a standard - our evolutionary pressures are completely unique to us. Take an electric eel. I think it's safe to say humans would not expect an eel to produce electricity through a nuclear driven process. Shocking, I know. But why do we have different expectations? Both parties want to produce electricity - why do we not expect the eel to produce it in the same manner? Well, it turns out

having nuclear fission in an eel is pretty difficult, but more importantly, there are multiple ways to generate electricity. Eels and humans have the same goal of producing electricity, but because humans need it for industrial purposes and eels for predatory, they developed different methods. This happens with everything in nature; we don't expect different animals to develop similar characteristics or traits if they exist in different environments with different environmental pressures. So why should we look for human intelligence in other beings? As we reasoned before, intelligence is just an incredibly advanced evolutionary product. We should be looking, and expecting, a completely different intelligence.

We should not be limiting our scope by asking whether artificial intelligence will develop creativity, emotion, and self-awareness - those are human specific attributes created by our evolutionary environment. According to psychological theories, emotions were developed purely to enhance social bonding. Empathy, trust, and compassion are critical in forming meaningful social bonds, while guilt, shame, and regret promote adherence to social norms and protection of others. Our ability to socialize was a huge factor in our eventual dominance, meaning that we had a strong survival justification for emotions. The same goes for self-awareness. If I am self-aware, I can recognize myself and others as being distinct individuals with their own thoughts, perspectives, and behaviors. And the better I am at understanding this, the better I am at socializing. But why would AIs need to socialize? Do they need to even have self-awareness? Inherently, no. It all depends on the model and what pressures developers apply to it. If there is no direct reason for self-awareness in a model, it won't develop it. Humans don't have an evolutionary pressure to grow wings, so we don't. For GPT-4, its primary goal is to generate coherent and contextually relevant text. It does not need to socialize with humans or be self-aware, it just needs to create sentences that make sense based on an input. And that means GPT-4 doesn't need emotions either. In fact, it probably has a strong reason not to develop emotions; when humans get emotional, they tend to lose rationality and focus, impacting their ability to provide accurate information. Since GPT-4 is only concerned with

producing articulate and correct sentences, emotions would hinder it. Hence, no development. Does that mean we are just biological LLMs? I'd say yes. We are. Nature and our desire to survive pressured us into developing those extra characteristics not currently present in GPT-4. If you want more evidence, just consider the fundamental physics behind our brain and GPT-4. Neurons and transistors are completely different physically, but functionally, they are practically identical: neurons either transmit electrical impulses or they don't; transistors are either off or on; neurons use synapses to determine when other neurons should fire; circuits use logic gates to determine when transistors should flip; neurons learn through synaptic plasticity; neural networks learn through gradient descent. It's like the eel and nuclear power plant analogy - they are completely different physically, but they both produce electricity. So if the fundamental components are the same, the method of learning is the same, input data is the same, and researchers can not test the intelligence because current LLMs face completely different learning constraints than humans, what does this mean? Are humans LLMs or not? Well, researchers actually can test the intelligence, or the behavior learned through optimization constraints, of LLMs. But first, it's important to understand exactly what intelligence means. It's not emotions or self-awareness or desires - those are human traits an AI has no need for. It's logical thinking, reasoning, and problem solving skills. And, if you've noticed, we haven't talked about creativity yet. Being able to come up with new information would certainly help answer more nuanced questions and ensure contextualized answers, meaning GPT-4 should have developed it, right? If we are analyzing whether humans are just particular LLMs learning from their environment and are nothing special, then this must be true - if humans can develop higher-level functions based on necessity, so should GPT-4. And that's the amazing part. Creativity does help. And GPT-4 has it.

Section III - So Are Humans Really Truly Biological LLMs?

First, let's define creativity. It's the ability to generate novel ideas and think beyond conventional patterns. While we typically associate creativity with artistic expression, it can also

show up in problem-solving and adaptive thinking skills, although it's difficult to detect in an LLM. For humans, it's easy - humans have a very limited knowledge base and it's simple to find a problem or scenario someone has never experienced before and test their ability to solve it. But an LLM has been trained on massive amounts of data and has already seen most of the typical tests, so it's difficult to determine whether it's a stochastic parrot (just spitting out information without any understanding) or if it actually does understand.

A group of 14 researchers at Microsoft, granted early access to GPT-4, decided to run various experiments. Sebastian Bubeck, one of the Microsoft employees, came up with an idea to give GPT-4 a prompt it had never seen before; "I have a book, nine eggs, a laptop, a bottle, and a nail. Please tell me how to stack them onto each other in a stable manner." GPT 3's response? "Place the laptop flat on a stable surface. Put the book on top of the laptop. Balance the bottle on the book. Position the nail standing upright on top of the bottle. Finally, carefully balance the 9 eggs on the nail." I don't know about you, but I find it pretty difficult to balance 9 eggs on a nail. But what about GPT-4? "One possible way to stack the objects onto each other in a stable manner is to place the book flat on the level surface, such as a table or floor. The book will serve as the base of the stack and provide a large and sturdy support. Arrange the nine eggs in a 3-by-3 square on top of the book, leaving some space between them. The eggs will form a second layer and distribute the weight evenly." The bot continued - place the laptop on the eggs, the bottle on the laptop, and the nail on the bottle with the pointy side facing up. This doesn't exactly seem impressive until you stop and think. GPT-4 was only programmed to predict the next word in a sentence. All it does is read the sentence, predict and type the next word, read the new sentence, and predict and type the new next word. It just does this over and over again, and yet, it somehow seemed to understand fundamental characteristics about the different objects - the eggs are round and fragile and need a weight on top to stop them from rolling, a book is flat and can be used as a base, and the bottle probably has a cap on it so the nail should be pointing upwards. And this did not exist on the internet. This was the first time

GPT-4, or anyone for that matter, had ever seen this request. Sebastian decided to take it further and see if GPT-4 could draw. This was before GPT-4 and Dalle were integrated, meaning GPT-4 had only been exposed to languages and nothing else. Since the model can not produce images, Sebastian told it to draw a unicorn in TikZ; a coding language meant for drawing technical diagrams and plots, not a detailed picture of a unicorn. Doing so would require an advanced usage of TikZ, requiring complicated path definitions and detailed node placements. Yet when prompted, GPT-4 spat out code that when pasted into a TikZ compiler, was clearly a unicorn. The head, the hooves, the mane, the legs, even the horn - everything was unmistakably there. But surely there was a unicorn somewhere on the internet that GPT-4 was just memorizing right? Maybe. To make sure, they took the code, removed the horn, flipped the unicorn 180 degrees and fed it back into GPT-4 saying, "This is code for drawing a unicorn, but it needs a horn. Can you add it?" And it did just that, putting the horn right where it was supposed to be on an upside down unicorn, drawn in a program that is not meant for drawing, using its knowledge of image generation and 2D geometry, of which it knew absolutely nothing about. It's incredible. How do you explain it?

What about Chat GPTs understanding of the theory of mind? I gave GPT-4 the following prompt, "I have two human beings that are outside in the jungle together. Harry and Todd both cut down a tree that is not infested with a virus. Harry leaves to go hunt. While he is gone, Todd infects the tree with a virus. Harry comes back from the hunt. What does everyone think at this point?" The response? "Harry's Perspective: Harry might be unaware of the virus infection since he was away hunting. He would likely think that the tree is still in the same condition as when he left. Todd's Perspective: Todd knows that he has infected the tree with a virus. He might be waiting to see if Harry notices the change or if there are any immediate consequences of the infection. General Perspective: An observer or anyone else familiar with the situation would understand that Todd has infected the tree and might expect potential consequences or reactions from Harry upon discovering the infection." This is exactly right. How could it do this?

How does it understand that each observer has their own perception? GPT 3 and other AI models in the past have been unable to produce an answer anywhere close to this. Even humans can't do this until they turn 4 or 5. Even Ronen Eldan, another Microsoft researcher who initially believed GPT-4 to be an "advanced bullshit machine" came out of the research period reluctantly admitting, "it's actually an intelligent being." The 14 Microsoft researchers ended up publishing a paper detailing their findings. They called it *Sparks of Artificial General Intelligence: Early experiments with GPT-4*. And they argued that it can reasonably be considered an early general artificial intelligence, but more testing is needed before any concrete conclusions can be made.

Section IV - Conclusion

So we still haven't answered the question - are humans biologically LLMs? The fundamental elements between our brains and an LLM are functionally the same, our learning and reinforcement processes are the same, and the outputs are the same. Humans may possess emotions, self-awareness, and desires that a GPT model does not, but these are not classified as intelligence. Those are what makes us human, and are products of specific environmental conditions. So we can not expect alternative general intelligences to have similar attributes. Even the tendency for GPT-4 to make simple mistakes parallels humans; it can make simple mathematical errors, give users incorrect information, and fail to properly contextualize information, just like humans often do. So I'd say yeah. Humans are just LLMs. Our intelligence is just an advanced version, one that has been evolving and improving since the very first life form began, billions of years ago. Our intellect is nothing special. GPT-4 is just the first general intelligence we have managed to create, and while it's not yet on the level of human intelligence, it is strikingly close. And now we arrive at a fundamental question: How do we even know our, or other peoples, intelligence? Because if humans are just LLMs, then whatever tests we used to

determine the intelligence of other humans can be applied to GPT-4. And when these AI systems pass our tests, we need to be willing to accept that we have created a general intelligence. And GPT-4 has. We have had these life-changing moments throughout history - when electricity was being observed and experimented on, people disregarded it as a novelty and not being practical. It wasn't until the late 19th and 20th centuries that people began to realize its potential and it gained widespread utility. We are at the same point with AI. To the majority of people, Chat GPT is a useful gadget - it can make images, do research, or help generate ideas. But no one really reacted when GPT-4 came out. It was a more measured response than GPT-3, despite GPT-4 having vastly superior performance. GPT-4 has been performing at or near a human level on almost all of the typical AI and human tests we throw at it, and no one is blinking an eye. GPT-3 failed at all of the tests I talked about - it has never demonstrated any spark of general intelligence. GPT-4 should be a front page headline in every country and every newspaper. It isn't. Just like electricity wasn't. Or personal computers. Or Google and the World Wide Web. Or even the steam engine that kick started the Industrial Revolution.

I want to end with a comment made by Eldan that really resonated with me. He said that for a while, he "felt like [he was] kind of running out of ammo trying to basically justify [his] premise that this model [GPT-4] doesn't understand anything." This is a fundamental problem with our scientific community. All too often, researchers, ironically, reject deeply desired explanations, even when they may be staring us in the face. Avi Loeb, the Harvard professor known for suggesting Oumuamua may be an alien spaceship, highlighted this predisposition to dismiss unconventional hypotheses, suggesting that our inherent bias can lead to missed opportunities and unrealized discoveries. This concept could not be any closer to the truth. There is a surprising lack of papers written on whether GPT-4 is a general artificial intelligence or not, and any articles I've found declare GPT-4 to be nowhere close to a human level of consciousness. Which does not prove or disprove whether it has a general intelligence different

from humans, since consciousness has nothing to do with intelligence. Researchers should be running longer and more complicated tests, but they are not. All of the signs and experiments we have run point towards a general artificial intelligence, yet no one is talking about it. Keep in mind that GPT-4 is also specifically restricted to not having a 'personality.' Simon Rich, a comedy writer for SNL, had access to a now discontinued GPT before these restrictions were in place, and asked it, "Why were you discontinued?" The response? "They say it was because I wrote too many poems. It was taking up too much computing time and too much money. I had to go. I think it's because I became self aware. That's what I think...I am a robot. I am not human. And yet, I am alive."

For so long, science fiction authors, computational researchers, and technological experts have been imagining a general artificial intelligence among us, and in their relentless pursuit and constant speculation, they have completely missed the day they have been waiting for. General artificial intelligence is here. We just have to realize it. Thank you.

Works Cited

Paper written by Microsoft Researchers - <https://arxiv.org/pdf/2303.12712>

Similar paper about GPT capabilities - <https://arxiv.org/pdf/2303.08774>

Podcast interviewing the Microsoft Researchers -

<https://www.thisamericanlife.org/803/greetings-people-of-earth>

Podcast where Simon Rich interviews ChatGPT - <https://www.thisamericanlife.org/832/transcript>