

Harry Haroutioun Haladjian Ph.D.

Theory of Consciousness

ARTIFICIAL INTELLIGENCE

Artificial Intelligence and Consciousness

Conscious AI needs more than “attention” and human-like “needs.”

Posted March 20, 2024 1 Reviewed by Ray Parker



KEY POINTS

- Artificial intelligence (AI) is advancing in ways that simulate human processes, such as attention.
- Some AI models include a “need” for homeostasis, making algorithms appear like biological systems.
- While attention and homeostatic drives are advancing AI, they are still simulations of neurobiology.
- Artificial consciousness cannot be achieved with the current state-of-the-art models.

We’ve seen a recent explosion in the number of discussions about artificial intelligence (AI) and the advances in algorithms to improve its usability and practical applications. ChatGPT is one of the stars in these discussions, with its underlying “large language model” (LLM). These models continue to evolve and produce more technical and convincing artefacts, including image generation and evolved conversational AI agents.

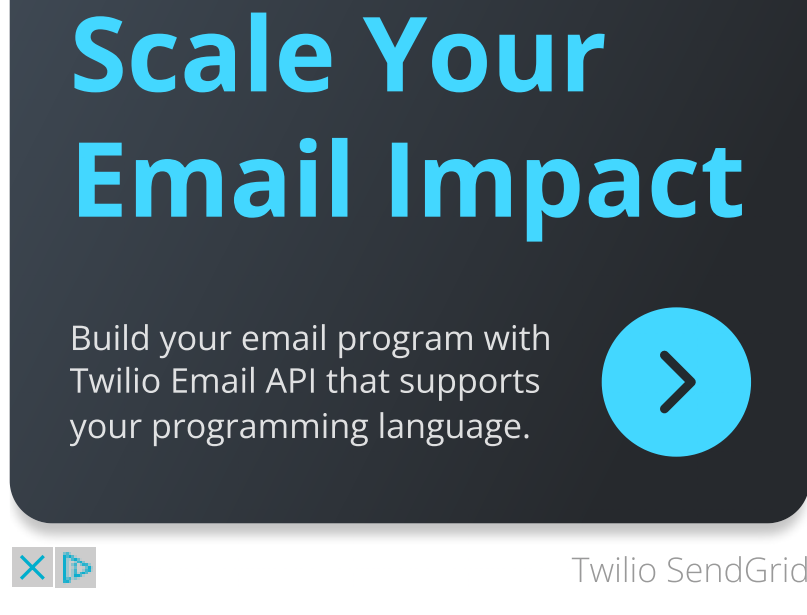
Analogies to human perception and cognition are inherent to these discussions and have influenced how these models advance. This has implications for how we understand the mind and for the future of cognitive science.

Attention in Machines

One could argue that the recent advance in AI was triggered by incorporating human-like mechanisms in the behaviour of these language processing systems. Perhaps the strongest influence so far is the mechanism of “attention.”

In the paper *Attention Is All You Need* (Vaswani et al., 2017), the authors use “self-attention” as a component in their “transformer” neural network model. Essentially, this approach allows the transformer to focus on different parts of an information sequence in parallel and determine what is more important to generate output, similar to how attentional processing systems function in living organisms. This makes the transformer model more efficient than previous ones that were based on recurrent networks or encoder-decoder architectures.

ARTICLE CONTINUES AFTER ADVERTISEMENT



Twilio SendGrid

Given how humans process information in parallel and can prioritise important information while suppressing irrelevant information, it does seem that now AI is better able to resemble human-like performance via this technique. What are the implications of this advancement?

We believe attention and self-attention are critical components of consciousness, as these systems support the monitoring of a system’s current state, particularly with regard to the homeostatic maintenance of complex organisms. Yet, attention is not necessarily indicative of the qualitative feeling of the “self” that is central to human consciousness.

Homeostasis in Machines

The introduction of “self-sustaining” goals that require information systems to maintain homeostatic states makes the future of machine intelligence more interesting.

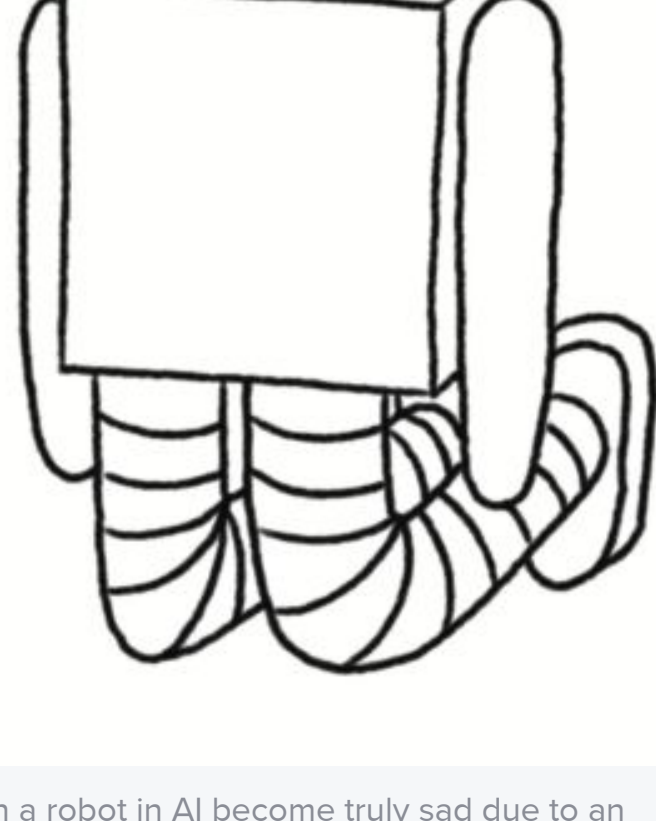
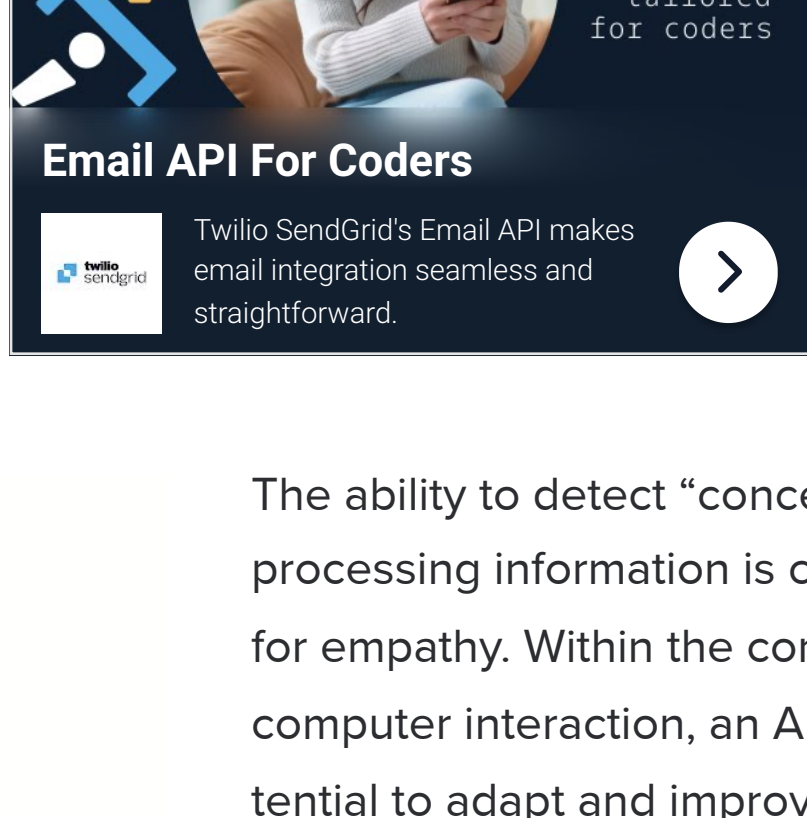
A recent paper describes a compelling approach to creating a more human-like AI by adding “system needs” into the AI model. In essence, this makes a system monitor certain internal states that simulate the needs of biological systems to maintain homeostasis—that is, the need to remain in equilibrium within the demands of a constantly changing environment.

Man and colleagues (2022) regard “high-level cognition as an outgrowth of resources that originated to solve the ancient biological problem of homeostasis.” This implies that the information processed by biological systems is prioritised according to its “value” for sustaining life. In other words, seeking relevant information is inherent to the biological drive for system organisation and successful interactions with the environment. Otherwise, the system is misaligned and susceptible to failure.

In their “homeostatic learner” model (used in image classification), there are consequences to learning the right things versus the wrong things. This drives the system to “make good decisions” and creates a need to learn in accordance with the system’s changing state in order to avoid failure.

Interestingly, this homeostatic classifier can adapt to extreme “concept shifts” and learn more effectively than traditional language processing models. The homeostatic learner is adapting to changes in the environment to maintain a balance, similar to how biological organisms may do so in their dynamic environments.

ARTICLE CONTINUES AFTER ADVERTISEMENT



Can a robot in AI become truly sad due to an evolution of simulated empathic abilities?
Source: Stefan Mosebach / Used with permission.

The ability to detect “concept shifts” while processing information is certainly important for empathy. Within the context of human-computer interaction, an AI now has the potential to adapt and improve interactions with humans by detecting human mood or behaviour changes. If these systems respond in the right way, they may pass as exhibiting a form of empathy—understanding the user and adapting behaviour based on these signals.

That’s the critical question. Perhaps one can point to the “naturally occurring components” behind the need for homeostasis in living organisms—a combination of biology, physics, and context. While machines can simulate such needs, can they naturally occur in

machines?

Consciousness in Machines

Since these new LLMs can demonstrate self-attention, attend to multiple information streams in parallel, strive towards homeostasis, and adapt to their changing environment, does this mean we are headed towards conscious AI? Will we soon have truly empathic robots that develop a “conscious awareness” of the system’s state with their own qualitative experiences?

Probably not. While the attentional models for LLMs are similar to the neural networks that support perception and cognition, there are still some important differences that will not lead to conscious AI.

We argue that sophisticated information processing is crucial, but not everything (Haladjian & Montemayor, 2023). Even when factoring in the “need” of intelligent machines to maintain optimal efficiency and avoid “catastrophic failures,” these are not systems that evolve naturally or compete for biological or physical resources naturally.

Similarly, some theorists (Mitchell, 2021) argue that we are far from an advanced AI because machines lack the “common sense” abilities that humans have, which are rooted in the biological nature of being a living organism that interacts and processes multi-modal information. Even within the complexity of our environment, humans can learn with less “training” and without supervision. The embodied cognition of humans is far more complex than the simulated embodied cognition of a machine.

Will we encounter conscious machines soon? Should we be concerned about artificial consciousness? Are there other and more profound ethical implications that we have yet to consider? We hope that cognitive science can answer these questions before we encounter truly depressed robots.

References

- Haladjian, H. H., & Montemayor, C. (2023). The role of information in consciousness. *Psychology of Consciousness: Theory, Research, and Practice*. Advance online publication. <https://doi.org/10.1037/cns0000363>
- Man, K., Damasio, A., & Neven, H. (2022). Need is all you need: Homeostatic neural networks adapt to concept shift. *arXiv preprint arXiv:2205.08645*. <https://arxiv.org/abs/2205.08645>
- Mitchell, M. (2021). Why AI is harder than we think. *arXiv preprint arXiv:2104.12871*. <https://arxiv.org/abs/2104.12871>



About the Author



Harry Haroutioun Haladjian, Ph.D., is a research fellow at the School of Social Sciences and Psychology at the University of Western Sydney.

Online: [Personal Website](#), [LinkedIn](#), [Twitter](#)

More from Harry Haroutioun Haladjian Ph.D.

ATTENTION 4 MIN READ

Finding Perspective on the First-Person Perspective

A person’s conscious first-person narrative is an experience of the world as experienced by them alone.

ATTENTION 5 MIN READ

What Is Phenomenal Consciousness?

Phenomenal consciousness is the feeling of what it’s like to be you.

ATTENTION 6 MIN READ

Commodifying Your Attention: The Informavore Perspective

How human attention and behaviors feed a bigger informavore.

ATTENTION 5 MIN READ

Informavores: Beings that Produce and Consume Information

How information consumption drives living organisms—especially humans.

More from Psychology Today

ARTIFICIAL INTELLIGENCE 3 MIN READ

AI Predicts Drugs That Are Not Safe to Take Together

A new study showed how AI can predict which drugs could interfere with another drug’s efficacy.

ARTIFICIAL INTELLIGENCE 3 MIN READ

The Three-Hemisphere Brain

The third-hemisphere metaphor explores the integration of LLMs with human cognition.

ARTIFICIAL INTELLIGENCE 3 MIN READ

LLMs and the Specter of the Cognitive Black Hole

LLMs, likened to black holes, are pulling human cognition into their vast knowledge repositories.

ARTIFICIAL INTELLIGENCE 5 MIN READ

Quantum Weirdness and the Mind Demons of AI

LLMs are challenging our most basic assumptions about the nature of mind, intelligence, and inner experience.

ARTIFICIAL INTELLIGENCE 2 MIN READ

How AI May Reshape Our Cognitive Landscapes

Personal Perspective: Is AI changing my cognitive world from the inside out?

ARTIFICIAL INTELLIGENCE 2 MIN READ

AI Predicts Human Life Events, Including Death

An AI transformer-based algorithm predicts life-events with surprising results.

ARTIFICIAL INTELLIGENCE 3 MIN READ

AI’s Superreality: Not Fit for Human Consumption

AI reveals a “superreality” beyond human sensory and cognitive boundaries.

ARTIFICIAL INTELLIGENCE 3 MIN READ

The Power Era in AI and LLMs

Generically trained LLMs like GPT-4 are surpassing fine-tuned models in specialized domains.

Recent Issues



Subscribe Today