

Artificial intelligence called GPT-3 can write like a human but don't mistake that for thinking

September 18 2020, by Guillaume Thierry



Credit: Pixabay from Pexels

Since it was unveiled earlier this year, the new AI-based language generating software GPT-3 has [attracted much attention](#) for its ability to

produce passages of writing that are convincingly human-like. Some have even suggested that the program, created by Elon Musk's OpenAI, [may be considered](#) or [appears to exhibit](#), something like artificial general intelligence (AGI), the ability to understand or perform any task a human can. This breathless coverage reveals a natural yet aberrant collusion in people's minds between the appearance of language and the capacity to think.

Language and thought, though obviously not the same, are strongly and intimately related. And some people tend to assume that language is the ultimate sign of thought. But language can be easily generated without a living soul. All it takes is the digestion of a database of human-produced language by a [computer program](#), AI-based [or not](#).

Based on the relatively few samples of text available for examination, GPT-3 is capable of producing excellent syntax. It boasts a wide range of vocabulary, owing to an unprecedentedly large knowledge base from which it can generate thematically relevant, highly coherent new statements. Yet, it is profoundly unable to reason or show any sign of "thinking".

For instance, one passage written by GPT-3 predicts you could [suddenly die](#) after drinking cranberry juice with a teaspoon of grape juice in it. This is despite the system having access to information on the web that grape juice is edible.

Another passage suggests that to bring a table through a doorway that is too small you should cut the door in half. A system that could understand what it was writing or had any sense of what the world was like would not generate such aberrant "solutions" to a problem.

If the goal is to create a system that can chat, fair enough. GPT-3 shows AI will certainly lead to better experiences than what has been available

until now. And it certainly allows for a good laugh.

But if the goal is to get some thinking into the system, then we are nowhere near. That's because AI such as GPT-3 works by "digesting" colossal databases of language content to produce, "new", synthesized language content.

The source is language; the product is language. In the middle stands a mysterious black box a thousand times smaller than the [human brain](#) in capacity and nothing like it in the way it works.

Reconstructing the thinking that is at the origin of the language content we observe is an impossible task, unless you study [the thinking](#) itself. As philosopher [John Searle](#) put it, only "machines with internal causal powers equivalent to those of brains" could think.

And for all our advances in cognitive neuroscience, we know deceptively little about human thinking. So how could we hope to program it into a machine?

What mesmerizes me is that people go to the trouble of suggesting what kind of reasoning that AI like GTP-3 should be able to [engage with](#). This is really strange, and in some ways amusing—if not worrying.

Why would a computer program, based on AI or not, and designed to digest hundreds of gigabytes of text on many different topics, *know* anything about biology or social reasoning? It has no actual experience of the world. It cannot have any human-like internal representation.

It appears that many of us fall victim of a mind-language causation fallacy. Supposedly there is no smoke without fire, no language without mind. But GPT-3 is a language smoke machine, entirely hollow of any actual human trait or psyche. It is just an algorithm, and there is no

reason to expect that it could ever deliver any kind of reasoning. Because it cannot.

Filling in the gaps

Part of the problem is the strong illusion of coherence we get from reading a passage produced by AI such as GPT-3 because of our own abilities. Our brains were created by hundreds of thousands of years of evolution and tens of thousands of hours of biological development to extract meaning from the world and construct a coherent account of any situation.

When we read a GPT-3 output, our brain is doing most of the work. We make sense that was never intended, simply because the language looks and feels coherent and thematically sound, and so we connect the dots. We are so used to doing this, in every moment of our lives, that we don't even realize it is happening.

We relate the points made to one another and we may even be tempted to think that a phrase is cleverly worded simply because the style may be a little odd or surprising. And if the language is particularly clear, direct and well constructed (which is what AI generators are optimized to deliver), we are strongly tempted to infer sentience, where there is no such thing.

When GPT-3's predecessor GPT-2 [wrote](#), "I am interested in understanding the origins of language," who was doing the talking? The AI just spat out an ultra-shrunk summary of our ultimate quest as humans, picked up from an ocean of stored human language productions—our endless trying to understand what is [language](#) and where we come from. But there is no ghost in the shell, whether we "converse" with GPT-2, GPT-3, or GPT-9000.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Artificial intelligence called GPT-3 can write like a human but don't mistake that for thinking (2020, September 18) retrieved 19 September 2024 from <https://techxplore.com/news/2020-09-artificial-intelligence-gpt-human-dont.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.