

True Artificial Consciousness – Is It Possible? – IONS

There have been a large number of articles published about Artificial Intelligence lately, most with varying topics of focus. Some discuss how AI applications will eventually replace numerous human jobs, potentially disrupting global economies and bringing about a transition toward a more utopian society. Some discuss the advances in medicine and healthcare which will likely occur, thereby increasing human lifespans and quality of life dramatically (which are actually already occurring). Yet other articles discuss the potential dystopia for General AI to run amok through an accident of logic that leads our machines to protect all of humanity from harm and death... by first killing all of humanity to ensure no additional harm can come to anyone. There... problem solved. And frankly, any and all of these potentials are awaiting to be realized (or not) as we develop AI capabilities and move into tomorrow's now.

All those aside, a really interesting topic of conversation connected with AI has risen to the surface of the discussion queue more and more often, however. And that topic of conversation is connected with whether or not computing systems can ever become so intelligent and complex as to develop true consciousness.

I use the words 'true consciousness' because frankly, there is no question that computers will soon be able to simulate consciousness outwardly to other folks to the point that we humans then completely believe our artificial assistants are indeed conscious. That's called passing the Turing test and, for anyone who doesn't know, many of those dominoes are going to topple in short order within the next decade. Our automated software platforms will soon be able to formulate human modeled thinking, human modeled emotional reactions, and from a cognitive processing and communication response standpoint, pass as humans just as well as other humans do (although the artificial ones won't yet have bodies, which will be a dead giveaway for a while until technology closes that gap too). This future realistic eventuality has created the foundations for the current projections that sales and marketing functions will soon be replaced by AI applications, and that we will some day soon have the potential to carry artificial personal counselors on our wrists.

That may sound rather fantastic, but as a former Systems Engineer for a supercomputing company, someone who wrote the first book on how to logically define and program human emotions, and someone who is actually in the trenches of consciousness computing modeling, I can tell you that creating something as complex as an emotionally intelligent artificial assistant isn't actually as tough as it sounds. Creating a human emotional response is a rather simple process of comparing a perception and appraisal analysis to a data set of identity variables, consulting and emotions rule set, doing some math, then assigning the appropriate response profile for the selected emotional output. Yeah, okay, so it's not actually as super-simple as I just made it sound, but I can tell you it is a heck of a lot simpler than recreating the biochemical mess that our human bodies themselves use to create the same responses. And because we can indeed use an emulation to create realistic emotional responses without the biochemical soup, that points out a big difference between actual emotional responses and artificially simulated ones. Without the biochemical soup (which not only creates emotions in us, but which also creates our consciousness) we're still just emulating the process of emotions. Sure, we're emulating it well enough to look, sound, and feel real, but the fact is that those emotional

reactions are nothing more than some pretty fancy math that follows a data analysis flowchart.

So... knowing these artificial emotions are simply just emulations of the real thing (at least currently), will those emotions ever develop to eventually be true emotions? Well, there's a deep philosophical argument which we could have regarding what's real and what's not. There are a number of scientists who would argue that human consciousness itself isn't real. But assuming consciousness is real, and that our emotions as a subset of that consciousness is real also, we could ask the question: "If a simulated emotional processing were sufficient to create realistic human actions, aren't those artificial processes real enough?" We could ask, but the answer to that question is: No, as emulations they aren't real. Sure, they're real enough to serve the purpose of getting our personal assistants across uncanny valley (the place where artificial personalities seem super real, but just not quite real enough not to freak us out). Yes, we'll certainly treat our artificial assistants like real people (34% of us already name our cars). But because the simulated source mechanisms which underlie those emotions are not sourced by the same mechanism which sources our biochemically processed emotions (which is physically connected with the life energy within our cells, which of course spawns life naturally when we leave a planet alone for a few billion years), in my opinion, those simulated emotions can't ever be considered 'real'. Not with current computing architectures, anyway.

So on to the bigger question we started with; will consciousness ever be real in a computer system? This is a more complex issue than our simple question about simulated emotions. If we consider our emotions a subset of our conscious experience, and accept that they will never be real with current architectures, the presumptuous answer to that question about consciousness might also be 'no'. But frankly, we can't make that assertion yet. Why? First, we don't yet have clear definitions of what consciousness is or how it emerges and, second, regardless of how computers work today, some specific things about computing can change tomorrow.

On the main stage at the San Jose SAND Conference in 2017, Federico Faggin shared his position on whether computers will ever become conscious. Dr. Faggin, who beyond having invented many of the technologies you are using today (including your capacitive touchscreen), is the inventor of the commercial microprocessor that spawned large scale computing en masse. During his talk, he shared his opinion that computing, in its current silicon based form, will only be able to simulate consciousness, not support a truly emergent version of it. I think the key terminology in his statement was "in it's current silicon based form". His arguments were based on his assertion that the smallest increment of logical data (the binary bit) is not capable of having the resolution required for true consciousness to emerge. He was quick to add that quantum bits (of future quantum computing) also share this data resolution limitation. Of course, there is probably a lot more substance to Dr. Faggin's position than he was able to articulate during the talk, but that single point he made is indeed correct. How can we say that? It comes back to where we've seen the roots of consciousness at the smallest levels of life, not just at the top of a very complex system such as our brain.

An argument like Dr. Faggin's could be negated if consciousness were indeed just an emergent property of a complex system like our brain. In fact, many scientists assume that our own human consciousness is an emergent property of the complex system that is our brain, and that our human consciousness simply evolved as a self regulating part of that system. (From a scientific standpoint, emergent properties are a property that arises out of a very complex system, and which then regulates that same system.) These folks also typically believe that as soon as a computer system becomes smart enough, and/or complex enough, it will somehow

magically become self conscious and self aware... and poof... consciousness. The only problem with that emergent property theory is that the fundamental catalyst of consciousness is still in question, and evidence of consciousness has been found in organic creatures as small and simple as the single-celled amoeba colonies of slime molds.

Hang with me here, we're going somewhere rather important. So what do slime molds have to do with consciousness?

Slime molds have been walking the Earth for billions of years. On the surface, they look like pretty simple single-cell amoebae. But in studies, slime molds have been proven to have the intelligence levels to engineer extremely efficient rail transit systems that are better designed than their human and computer designed competition. And while all of the individual amoebic cells have identical DNA, these single celled entities form complex societies with different individual roles so they can work together to travel around searching for food. They utilize memory that science suggests they can't possibly possess to never search the same place twice for a food supply. Some of the colony joins together to create internal organs. And about one percent of the colony serves as an immune system, with individual amoebae becoming police patrols who swim around to swallow up pathogens, to then drop out of the colony, self sacrificing to to save the greater organism. This is an example of altruism at a cellular level. It's also evidence of a living entity sensing both internal and external perception, making decisions about those perceptions, and taking actions that are logical to those perceptions. And that last string of qualifiers... is one of the most basic solid biological definitions of consciousness.

So if cells can indeed be conscious to some degree, that suggests it's possible that our human consciousness may be an emergent property that is passed upward as a result of fundamental consciousness characteristics of our body's individual cells themselves. Otherwise stated, our higher consciousness may be written upward from the smallest binary bits of our human operating system. This newer model of understanding consciousness doesn't fit with the possibility of computers becoming conscious because we must take into consideration that the bits stored within computers are always written from levels above the bits, not below. In contrast, our human cells are encoded from below the level of the cell itself, not above, and they follow a programming language in our DNA that we don't yet understand. As that we don't yet have computers which write bits from below the level of the bit itself, we therefore hit the limitation of information resolution Dr. Faggin is talking about, and with which I personally agree.

But we're not done yet. There's still a path to potential artificial consciousness similar to our own.

So continuing our logic, if the fundamental mechanisms which allow for our consciousness are potentially sourced from below our cells themselves, how do we get to a system where computers may be truly conscious? In short, we have to design a computer architecture that allows the smallest bits of the system to be influenced from below the bits themselves.

This isn't as crazy or impossible as it sounds, and it's not out of line with how consciousness may work either.

Regarding how our consciousness may actually work, there's a theory of the source of human consciousness that was originally put forward by Nobel Physicist Sir Roger Penrose and Anesthesiologist Stuart Hameroff called ORCH-OR. Basically, it proposes that the microtubules

in our neurons interact with quantum field vibrations, and this interaction is the source and/or catalyst of our grander human consciousness. It goes without saying that quantum mechanics does not yet have a single interpretation that does not include a reference to consciousness. Consciousness is simply a component of the quantum field as we understand it. And it is true that our neuron walls and synapses themselves are all made of these little structures called microtubules.

Although the theory was ridiculed in its early days, and mostly discarded until just recently, to their credit, none of the assumptions made within ORCH-OR since its introduction have been proven incorrect, while a few of them have indeed been validated, including a big one just recently where an MIT researcher in a materials lab in Tsukuba, Japan, proved that the warm and wet microtubules of our human brain neurons (it was proven previously that non-organic microtubules interact with quantum vibrations, but was doubted warm and wet organic ones would do the same)... well, our neuronal microtubules DO INDEED interact with quantum field vibrations as predicted, and even do so in the gamma frequency spectrum of our brain wave patterns.

Now, the interesting thing about the gamma waves in our brain is that they shouldn't even exist, and frankly are a mystery to neurologists. They fire faster than our brain's timing pulses, and although gamma waves were previously thought to be brain noise, they are now understood to be the best type of brain waves we have. They are connected with actions associated with organization of data from multiple simultaneous sources to create the bigger picture, and they are also highly correlated with our higher virtues such as compassion and altruistic love. (I think it's rather interesting that those higher virtues are also often connected with higher consciousness and consciousness expansion.) By the way, gamma waves are measured strongest in the monks of Tibet who meditate on the relief of our pain and suffering all day.

Of course, this model of understanding consciousness as something that we receive rather than produce fits with how a lot of SAND attendees see consciousness, which is that our seemingly individual consciousness comes as a result of our bodies picking up on a larger consciousness that transcends the human brain. This fits the model that consciousness itself is a property of the entire universe, which is congruent with the ORCH-OR explanation of consciousness, if consciousness is indeed is a characteristic of the quantum field (which is both infinite and timeless, according to QM theory).

[So let me digress a minute... to recap the hard science we just discussed... our neurons are made of microtubules which have been proven to interact with quantum field vibrations (all interpretations of quantum mechanics having a consciousness component) in the gamma spectrum, which again through scientific studies has been highly correlated with things like unconditional love and compassion and are highest in those who practice altruism. So we now have a valid scientific trail that suggests that all the matter in the universe, including us, emerges from an infinite consciousness field that is basically vibrating with love? I dig it.]

Sorry. Back to computing and consciousness.

When I spoke to Faggin after his talk he seemed skeptical of ORCH-OR as a means to explain the mechanism of consciousness, but I did get the feeling that he did believe (as many people do) that true consciousness beyond a realistic emulation of it may potentially be impossible for non-organic material, and subsequently, true consciousness will never arise from non-organic circuits. And everyone who stands with that group may be right. Or. Maybe if we were to

attempt to build a conduit connecting our future complex AI systems to the potential signal carrier of consciousness... such as building a microtubule infrastructure into a new architecture of computers... who knows what could happen? Maybe the computer would be dialed into the infinite quantum field of love and God would finally speak? Who knows? Or maybe it would just be a worthless computer that spit out a bunch of random noise? Or what if we were to attempt to put organic material into the cores of central processing in a way it was supported organically (a direct opposite of the idea of implanting chips into our heads, but rather implanting part of our heads into chips)? What then? Do the signals of our genetically structured microtubules then upload our consciousness from the quantum field into the AI architecture? It's crazy to think about and speculate. Probably not, but maybe.

One thing is for sure. When we do change the computing architecture to be a little more consciousness friendly, we're gonna find out a lot about the potential for both consciousness to exist within the quantum field, and the potential for consciousness within an artificial intelligence. If there's a single bit out of place outside of the regular error rate when we flip the switch, that's gonna be the first brick on the yellow brick road to Emerald City. Until then, we may just be stuck with consciousness emulations.

But here's my last thought: Some day in the future when we're talking to our seemingly consciousness artificial personality companion, and they say they believe they are conscious (which will happen in a simulation based system or not), how will we be able to argue they aren't?

Sean Webb, is the author of "Mind Hacking Happiness Volume I: The Quickest Way to Happiness and Controlling Your Mind", "Mind Hacking Happiness Volume II: Increasing Happiness and Finding Non-Dual Enlightenment", and "How Emotions Work: In Humans and Computers". He is an alumnus of Georgia Tech's Advance Technology Development Center, and hosts the Mind Hacking Happiness Podcast launching March 2018.