**Research Essay**

# Why AI Will Never Surpass Human Intelligence

Bradley Y. Bartholomew*

## Abstract

A paper entitled "The unreasonable effectiveness of deep learning in artificial intelligence" argues that the way forward towards achieving general AI, that is to say a human level intelligence, is to copy how an organic brain does if for humans. The paper argues that AI has to move from a very limited 2D-space which is referred to as Flatland to a tera-dimensional space that represents the million billion synapses between the neurons in the cortex of the human brain. It is pointed out that the move from 2-D AI to tera-dimensional AI is actually a move in the wrong direction if they ever hope to achieve general AI. The fact is that although there are a million billion synapses between the neurons in the cortex, human consciousness is one dimensional or holistic. In order to achieve general AI the machine will have to do everything a human can do where there are no gaps or seams in the output. A model of the human brain is offered where different sections of the cortex are specialized for different functions and these disparate regions communicate with each other electronically at the speed of light via brainwaves and this is how the brain generates a global holistic 1-dimensional consciousness in us. Also as numbers don't exist in Nature an organic brain, unlike deep learning, generates intelligent output without the aid of numerical programs or statistics.

**Keywords**: Deep learning, AI, statistics, neural networks, mathematics, Singularity, transhumanism.

This paper is a reply to a recent paper entitled "The unreasonable effectiveness of deep learning in artificial intelligence" which recalls the celebrated paper by Eugene Wigner "The unreasonable effectiveness of mathematics in the natural sciences". [1,2] The author says:

> In his essay "The Unreasonable Effectiveness of Mathematics in the Natural Sciences," Eugene Wigner marveled that the mathematical structure of a physical theory often reveals deep insights into that theory that lead to empirical predictions. Also remarkable is that there are so few parameters in the equations, called physical constants. The title of this article mirrors Wigner's. However, unlike the laws of physics, there is an abundance of parameters in deep learning networks and they are variable. We are just beginning to explore representation and optimization in very-high-dimensional spaces

* Correspondence: Bradley Y. Bartholomew, Independent Researcher, France. Email: brad.bartholomew2@gmail.com

The remarkable thing is that although the title to the paper under reply mirrors Eugene Wigner's paper, the theme or tone of the two papers are strikingly different. Whereas Wigner was marveling at how effective mathematics has been to describe physical processes and predict physical outcomes, in this paper the author seems to be arguing that deep learning in artificial intelligence systems has hit some sort of ceiling in the pursuit of making a machine that can match human intelligence. His argument is that further advancement towards general artificial intelligence can only be achieved by studying and mimicking actual brain processes. If AI is ever to become truly conscious and have a general intelligence to equal or surpass humans then it will have to mimic the way the brain generates consciousness and intelligence in us.

On its face this of course is an eminently reasonable and valid argument, but it has a hidden irony much like the 'hidden' layers in the deep learning networks themselves, that deep learning and AI in general relies heavily, one might say almost exclusively, on statistical mathematics in order to generate output, whether this be recognizing a physical object in an image, or a word or a sound, or answering any question that may be put to it. Essentially it makes five guesses and gives you an answer that is 'probably' the best answer. If that output happens to be wrong then the 'hidden layers' in the system will update the statistical analysis accordingly, and the next time around it will be less likely to make that mistake. This is where the 'deep learning' comes in. The statistical analysis of the data is gradually improving with every iteration of the system. This has proved to be remarkably effective in certain limited areas that involve human intelligence, which gets us back to this word 'unreasonable' in the title of Eugene Wigner's article. Essentially we have to consider the true relationship between human mathematics and natural processes.

## AI and the Age of Spiritual Machines

The ideas that led to the first programmable computers came out of mathematicians' attempts to understand human thought—particularly logic—as a mechanical process of "symbol manipulation." Digital computers are essentially symbol manipulators, pushing around combinations of the symbols 0 and 1. To pioneers of computing like Alan Turing and John von Neumann, there were strong analogies between computers and the human brain, and it seemed obvious to them that human intelligence could be replicated in computer programs.[3]

This notion that computers and human intelligence are in some way related or involve the same operations has permeated the information technology community ever since its inception and has led to the wildest and most extravagant, dare one say absurd, claims by leading lights in the tech industry as to what computers will ultimately be capable of. Notwithstanding that computer output is to a 2D screen it has been seriously predicted that eventually computers will be simulating whole universes, will be bringing cryonically suspended cadavers (and severed heads) back to life whereupon they will live forever as cyborgs, terminally sick and aged people will have eternal life by having their brain uploaded with their personal consciousness of self remaining intact, machines will not only surpass human intelligence in specific domains and generally, but will have humanlike consciousness, aspirations and emotions. Picture if you will an evil paperclip machine that 'decides' to turn the whole universe into paperclips, or a pious machine who stops into a church to pray.

The hype for AI has become a religion known as transhumanism. Notwithstanding the fact that these machines have no mental life apart from crunching numbers, they will look upon us as lesser beings. They will be 'conscious' of being superior to us, just as we are conscious of being superior to an intelligent animal such as a dog. God is not dead; in the religion of Transhumanism, god is a Machine.

## What is mathematics

All this hype that has turned AI into a religion is largely due to the fact that the speed of computers for many years now has been increasing exponentially, as has memory storage capacity. For the AI community the fact that computers can process exponentially growing data at exponentially increasing speeds means that computers are becoming exponentially more intelligent. Surely this is consistent with the initial inspiration of Turing and other mathematicians that because computational 'symbol manipulation' is exponentially improving then the computer's ability to 'replicate' human thought is exponentially increasing as well.

The sticking point is this word 'replicate'. Can machines crunching numbers replicate or simulate living conscious human beings in such a way that they will actually surpass the human beings intellectually, creatively and spiritually. Can enhanced number crunching create beings with a 'higher' consciousness? This certainly resonates with the enigmatic assertion of Pythagoras that "all things are number" which surely embraces our soul, our intellect, our psyche, our mind and our consciousness. A subtitle to this paper could be: Solving the 'hard problem': Consciousness is a numerical phenomenon.

The only problem with all this euphoria surrounding numbers which has evidently been going on for two and half millennia is that numbers don't actually exist in Nature. In earlier papers I quoted the philosopher Oswald Spengler who to my mind stated this fact most emphatically.[4] He also points out that the crown jewel of mathematics in the physical sciences, differential calculus introduced independently by Leibniz and Newton, is mere statistics. As Spengler points out even the fundamental premise in Physics that space consists of three dimensions, the xyz axes in Cartesian coordinates, is a mathematical fiction.

> We feel — and the feeling is what constitutes the state of all-round awareness in us — that we are in an extension that encircles us; and it is only necessary to follow out this original impression that we have of the worldly to see that in reality there is only one true "dimension " of space, which is direction from one's self outwards into the distance, the "there" and the future, and that the abstract system of three dimensions is a mechanical representation and not a fact of life.[4]

It is not necessary to name any particular mathematical functions as an example of something that cannot exist in Nature. The fact is mathematics is about the properties of idealized concepts such as numbers, right angle triangles and circles that don't actually exist in Nature. If you take all these idealized concepts out of mathematics then there is nothing left. All of the mathematics in the physical sciences is a fiction invented by humans which enables them to observe, explain and predict natural processes on a statistical basis. Take the march of time for example. Mathematicians count the passage of time in seconds, minutes etc and put a t for time into their equations but obviously that is not what is actually occurring in Nature. The much touted Laws of Physics are merely descriptive; not dynamic. So deep learning in AI is not being intelligent in its own right, it is not even artificial intelligence, it is merely statistical manipulation of symbols that simulates processes in a human brain.

## The unreasonable effectiveness of AI

The paper under reply presents a very reasonable and balanced argument as to what AI has to do to move to the next level. The author is essentially saying that AI has already achieved astounding successes but the way forward is to go back and study how the brain actually does it. In fact the astounding successes the author is referring to are still examples of what is called "narrow" or "weak" AI. The terms narrow and weak are used to contrast with strong, human-level, general, or full-blown AI (sometimes called AGI, or artificial general intelligence)—that is, the AI that we see in movies, that can do most everything we humans can do, and possibly much more. General AI might have been the original goal of the field, but achieving it has turned out to be much harder than expected.[3]

Ever since its inception the development of AI has seen several seen several cycles where further advancement seemed bleak. These are the AI winters. And then a new idea pops up, or in the case of deep learning an old idea that was dead and buried comes to life again, and there is an AI spring. Like every AI spring before it, our current one features experts predicting that "general AI"—AI that equals or surpasses humans in most ways—will be here soon. "Human level AI will be passed in the mid-2020s," predicted Shane Legg, cofounder of Google DeepMind, in 2016. A year earlier, Facebook's CEO, Mark Zuckerberg, declared, "One of our goals for the next five to 10 years is to basically get better than human level at all of the primary human senses: vision, hearing, language, general cognition." The AI philosophers Vincent Müller and Nick Bostrom published a 2013 poll of AI researchers in which many assigned a 50 percent chance of human-level AI by the year 2040. Much of this optimism is based on the recent successes of deep learning programs.[3]

Interestingly the high priest of the Transhumanism religion, Ray Kurzweil, who wrote the Transhumanism bible "The Age of Spiritual Machines" made the following prediction in 2002: "A careful analysis of the requisite trends shows that we will understand the principles of operation of the human brain and be in a position to recreate its powers in synthetic substances well within thirty years." Kurzweil is vague on how all this will happen but assures us that to achieve human-level AI, "we will not program human intelligence link by link as in some massive expert system. Rather, we will set up an intricate hierarchy of self-organizing systems, based largely on the reverse engineering of the human brain, and then provide for its education … hundreds if not thousands of times faster than the comparable process for humans." Reverse engineering the brain means understanding enough about its workings in order to duplicate it, or at least to use the brain's underlying principles to replicate its intelligence in a computer.[3]

The paper under reply commences with the following statement:

> In 1884, Edwin Abbott wrote Flatland: A Romance of Many Dimensions. This book was written as a satire on Victorian society, but it has endured because of its exploration of how dimensionality can change our intuitions about space. Flatland was a 2-dimensional (2D) world inhabited by geometrical creatures. The mathematics of 2 dimensions was fully understood by these creatures, with circles being more perfect than triangles. In it a gentleman square has a dream about a sphere and wakes up to the possibility that his universe might be much larger than he or anyone in Flatland could imagine. He was not able to convince anyone that this was possible and in the end he was imprisoned. We can easily imagine adding another spatial dimension when going from a 1-dimensional to a 2D world and from a 2D to a 3-dimensional (3D) world. Lines can intersect themselves in 2 dimensions and sheets can fold back onto themselves in 3 dimensions, but imagining how a 3D object can fold back on itself in a 4-dimensional space is a stretch that was achieved by Charles Howard Hinton in the 19th century. What are the properties of spaces having even higher dimensions? What is it like to live in a space with 100 dimensions, or a million dimensions, or a space like our brain that has a million billion dimensions (the number of synapses between neurons)?

The author doesn't enlighten us as to why he thinks the solution for the advancement of AI is to increase the number of dimensions that AI is capable of processing. It is an interesting analogy between AI and Flatland because the output of AI is in fact 2-dimensional. The output invariably appears on a 2-D computer screen. So AI literally does exist in a 2-D space that could be aptly named Flatland. The author then suggests that in order to upgrade from narrow AI to general AI it will have to 'learn' to do processing in a space that has a million billion dimensions (the number of synapses between neurons). The author makes this suggestion in the form of a question where he doesn't actually know the answer which can be found by simply 'reverse engineering' his question. Instead of increasing the number of dimensions in the AI space, the secret is to decrease the number of dimensions to a 1-D space. The output of a human brain is not 2-dimensional or 3-dimensional or tera-dimensional. As Oswald Spengler observes above, our consciousness is one dimensional, that is to say it is holistic. The issue is how does a million billion neurons combine to generate in us a perfectly unified and seamless consciousness.

This question has been around for a long time. Erwin Schrödinger alluded to "The Arithmetical Paradox: The Oneness of Mind" in a series of lectures on *Mind and Matter* at Cambridge University in 1956. Schrödinger ponders the question why every living cell in an organism (say a human being) contains a complete copy of the DNA for that organism. He quotes Sir Charles Sherrington:

> The cell as a component of the body is not only a visibly demarcated unit but a unit-life centered on itself. It leads its own life … The cell is a unit-life, and our life which in its turn is a unitary life consists utterly of the cell-lives.

This paradox is particularly striking in relation to the brain, where the cortex is made up of a sheet of trillions of individual cells, each containing a complete copy of the human DNA in the nucleus, so each cell appears to be an autonomous unit; yet somehow this "commonwealth of cells" produces in us the impression of having a unified mind. "Matter and energy seem granular in structure, and so does 'life', but not so mind." Here we have one mind based ostensibly on many cell-lives; the only explanation Schrödinger could offer is that there must be a "sub-mind" associated with the individual cells that enables them to act perfectly in concert to produce a unified effect. He immediately dismisses this notion of a sub-mind in every living cell as an "absurd monstrosity." However, it is for every living cell, not just those of the brain, that this paradox arises. All living creatures consist of individual autonomous cells, whether one or millions or trillions, which act in concert to produce a unified effect.[5]

This is the essential problem for the AI community if they ever want to 'supersize' from narrow to general AI. The intelligence or consciousness that they generate has to be 1-dimensional. The irony of course being in order to 'supersize' they will have to decrease the number of dimensions of the AI generated space, not increase them. The author seems to suggest that with the 'unreasonable effectiveness' of mathematics they should be able to increase the number of dimensions of the AI space to an infinitely large number, and that is probably the case. 3-D printing is an example of moving from 2-D Flatland to 3-D at least. The problem is the more dimensions they introduce mathematically into the AI space the further away they are actually becoming to achieve general AI.

It's probably bad form to introduce anything so frivolous into an academic discussion about consciousness, but I would like to refer you to a video on YouTube that has the world champion Boogie-Woogie dancers for the years 1991-2011.[6] Consider, if you will, what the brain is actually doing to generate a 1-dimensional and seamless consciousness in these dancers of dancing the boogie-woogie. In addition to heightened general input-output for their autonomic nervous system, their brain is simultaneously merging streaming input from all five of their senses, and merging that with the parts of their brain responsible for proprioception, balance, body schema, intelligence, emotion, language, memory and musical appreciation and producing a unified consciousness on the cortex of their brain as well as simultaneously sending instructions via motor regions to literally every muscle in their body. Their 1-dimensional consciousness of dancing the boogie-woogie is simultaneously infinitely dimensional. And it is for this reason I contend that AI will never achieve general AI, let alone actually surpass human intelligence. There will be those who might deny that a champion boogie-woogie dancer is displaying intelligence. To them I say that the essential talent of a good dancer is their ability to 'interpret' the music.

These are world champion boogie-woogie dancers which of course AI will never ever come close to matching let alone surpass, for if it were to do so, deep learning would display 'unreasonable effectiveness' indeed. The examples of 'unreasonable effectiveness' of deep learning that the author speaks of are considerably more modest although they do indeed involve beating human champions at their particular game of expertise. AI has beaten a world champion chess player at the game of chess, a world champion Go player at the game of Go, and there are several other specific games that AI can play at somewhere near human levels of skill. These are all classic examples of narrow AI. There is an AI that can play a specific game, and none other.

The other area where narrow AI has been 'unreasonably effective' is in the area of language recognition. The author says:

> Natural language applications often start not with symbols but with word embeddings in deep learning networks trained to predict the next word in a sentence, which are semantically deep and represent relationships between words as well as associations. Once regarded as "just statistics," deep recurrent networks are high-dimensional dynamical systems through which information flows much as electrical activity flows through brains.

One might question whether statistically predicting what word is most likely to follow a particular word in a sentence actually represents a flow of 'information' where any 'intelligence' is involved, but it is true that narrow AI has achieved somewhere near 90% effectiveness in transcribing spoken language into text. If the spoken language is a question where a human is seeking 'information' then this request is submitted to a massive database that virtually encompasses the 'information highway' – the internet. In this way AI has managed to achieve a summit in narrow AI – it was able to beat a champion at the popular television game of Jeopardy. Here I must take issue with the author. The effectiveness of AI at playing the game of Jeopardy is not 'unreasonable' considering that the AI had in its database all the questions that have ever been asked in the game, and in fact the text of the question was submitted to the AI although it was made to appear as if the AI was recognizing the spoken question. Also the fact that the AI was simply quicker than the human at pressing the buzzer would have given the AI a 'superhuman' ability but I doubt if that really counts as 'intelligence'. Having said that, any output of an AI or an organic brain may be called 'intelligence' in the widest sense of the word.

## Deep learning vs neural networks

The paper under reply sets out what AI needs to learn from the neural networks in the brain in order to achieve general AI. You will see that he is stating the problem correctly if AI is ever to match the brain, however he doesn't seem to realize the enormity of the problem. Imagine if they were to develop narrow AI for every single human skill, talent, emotion, creative activity, area of knowledge and expertise and then have that network of narrow AI modules simultaneously communicate with each other to create in the machine a holistic intellectual consciousness. (Bearing in mind that for each of those broad categories of human endeavor there will be hundreds of specific processes that the narrow AI would have to individually master; for instance a human can learn to play hundreds of different games within the 'skills' category). In order to have general AI they would first have to have narrow AI at advanced level for *everything* that a human can do, and then have those thousands of modules of narrow AI instantaneously communicate with each other to produce a seamless holistic output. This is what our brain does and the author clearly recognizes this, but what he doesn't seem to recognize is that such a deep learning neural network would be so 'unreasonably effective' as to be impossible. He writes:

> Features of neurons are likely to be important for their computational function, some of which have not yet been exploited in model networks. These features include a diversity of cell types, optimized for specific functions; short-term synaptic plasticity, which can be either facilitating or depressing on a time scales of seconds; a cascade of biochemical reactions underlying plasticity inside synapses controlled by the history of inputs that extends from seconds to hours; sleep states during which a brain goes offline to restructure itself; and communication networks that control traffic between brain areas. Synergies between brains and AI may now be possible that could benefit both biology and engineering. The neocortex appeared in mammals 200 million years ago. It is a folded sheet of neurons on the outer surface of the brain, called the gray matter, which in humans is about 30 cm in diameter and 5 mm thick when flattened. There are about 30 billion cortical neurons forming 6 layers that are highly interconnected with each other in a local stereotyped pattern. The cortex greatly expanded in size relative the central core of the brain during evolution, especially in humans, where it constitutes 80% of the brain volume. This expansion suggests that the cortical architecture is scalable— more is better—unlike most brain areas, which have not expanded relative to body size. Interestingly, there are many fewer long-range connections than local connections, which form the white matter of the cortex, but its volume scales as the 5/4power of the gray matter volume and becomes larger than the volume of the gray matter in large brains. Scaling laws for brain structures can provide insights into important computational principles. Cortical architecture including cell types and their connectivity is similar throughout the cortex, with specialized regions for different cognitive systems. For example, the visual cortex has evolved specialized circuits for vision, which have been exploited in convolutional neural networks, the most successful deep learning architecture. Having evolved a general purpose learning architecture, the neocortex greatly enhances the performance of many special-purpose subcortical structures. Brains have 11 orders of magnitude of spatially structured computing components. At the level of synapses, each cubic millimeter of the cerebral cortex, about the size of a rice grain, contains a billion synapses. The largest deep learning networks today are reaching a billion weights. The cortex has the equivalent power of hundreds of thousands of deep learning networks, each specialized for solving specific problems. How are all these expert networks organized? The levels of investigation above the network level organize the flow of information between different cortical areas, a system-level communications problem. There is much to be learned about how to organize thousands of

> specialized networks by studying how the global flow of information in the cortex is managed. Long-range connections within the cortex are sparse because they are expensive, both because of the energy demand needed to send information over a long distance and also because they occupy a large volume of space. A switching network routes information between sensory and motor areas that can be rapidly reconfigured to meet on going cognitive demands.

As it happens there is a model of the brain that does indeed purport to supply all the information the author is seeking which will enable him to build his general AI. It is the model of the brain as an electronic device.[7,8] According to this model consciousness is generated by electronics, and the brains of all living creatures are connected electronic devices. It has now been found that the action potentials of the neurons specifically in the cortex of the brain are not an 'all-or-nothing' event as was previously thought, but in fact they generate a waveform that is capable of communicating information to other neurons in the cortex via brainwaves (ELF radio waves). That is to say that neurons in disparate parts of the cortex that are 'hard wired' to perform specific functions are able to communicate with each other at the speed of light. This means that input from an infinite number of terminals in the brain and body (including the senses) are able to merge into a single holistic experience of consciousness at the level of the cortex at the speed of light. In neuroscience this is referred to as the 'binding problem' and this is exactly what the author refers to in the passage above. "*The cortex has the equivalent power of hundreds of thousands of deep learning networks, each specialized for solving specific problems. How are all these expert networks organized?*"

## Conclusion

The author has written a paper arguing that if AI is to ever achieve general AI they must pay more attention to the way the brain does it. While his argument is perfectly sound, unfortunately it demonstrates at the same time the impossibility of this goal of general AI that the AI community has set itself. Essentially they will have to develop hundreds of thousands of narrow AI modules and have them all communicate with each other at the speed of light in order to simulate a global or holistic consciousness; no perceptible gaps or seams anywhere. But although it has demonstrated the impossibility of ever achieving general AI, his article has at least given the AI community a roadmap for the next five millennia at least. It's conceivable in the very distant future that conventional computing will take place at the speed of light, and so it's conceivable in the very far distant future that there will be something very closely approximating human intelligence. The only thing that will remain impossible is for general AI to surpass human intelligence. That would involve a programming that did not involve numbers, and an AI where 'intelligent' decisions are made that didn't involve mathematics and statistics. That will only occur when and if humans can actually create living human beings by means other than natural procreation. But the intelligence of these creatures would no longer be 'artificial'. Behold the 'singularity'.

*Received March 12, 2020; Accepted March 28, 2020*

## References

[1] Sejnowski, T.J. “The unreasonable effectiveness of deep learning in artificial intelligence”. *Proceedings of the National Academy of Sciences,* (2020) https://doi.org/10.1073/pnas.1907373117

[2] Wigner, E., “The Unreasonable Effectiveness of Mathematics in the Natural Sciences”. Communications in Pure and Applied Mathematics 13.1 (1960) New York: John Wiley & Sons, Inc.

[3] Mitchell, Melanie. “Artificial Intelligence”. Farrar, Straus and Giroux. Kindle Edition.

[4] Spengler, Oswald. “The decline of the West.” London Allen & Unwin. Kindle Edition.

[5] Schrodinger, E., “What Is Life?” (1945) Cambridge: Cambridge University Press.

[6] World Champion Boogie-Woogie 1991-2011 https://youtu.be/Svcs5jTsTaA

[7] Bartholomew, B.Y. “Solving the ‘Hard Problem’: Consciousness is an Electronic Phenomenon.’ *Journal of Consciousness Exploration and Research.* (2020) 11:46-60 https://jcer.com/index.php/jcj/article/view/862

[8] Bartholomew, B.Y. “The Electronic Waveform of Action Potentials in the Brain.” *Journal of Consciousness Exploration and Research*. (2020) 11:185-197 https://jcer.com/index.php/jcj/article/view/871