

How close are we to  
artificial general  
intelligence?



How Close Are We to Artificial General Intelligence? An outlook on AGI and the time to AGI completion.

Ever since the era of the first computers, researchers and scientists have been thinking and conceptualizing about the possibility of thinking machines. With the inevitable passage of time and advances in technology, thinking machines that can do limited tasks have been created, dubbed narrow Artificial Intelligence (Narrow AI). This, however, is still pale in comparison to the true holy grail of the thinking machines: the machine that can perform any and all tasks that a human can, a true Artificial General Intelligence (AGI). This essay will describe at how Artificial General Intelligence is already within decades, if not years away from completion. It will discuss the advances in the study of Artificial General Intelligence (AGI) including the pathways one might took to create AGI, the dangers that may arise from one, how such dangers might be mitigated, how AGI is closer to completion than ever before, and how there is a scepticism about it.

The least difficult and the least dangerous pathway to Artificial General Intelligence will be the Whole Human Brain Emulation when compared to creating AGI from ground up (de novo AGI or dnAGI) or Neuromorphic AGI (NAGI). The process of writing a program from the bottom up to fully emulate intelligence is dnAGI. To create an AGI using the NAGI process, the human brain is studied and certain key features of the brain's architecture are appropriated in order to yield intelligence with some similarities to human brain (Eth, 2017). Just as the name suggest, Whole Brain Emulation (WBE) is the process of scanning and copying a human brain entire function into a computer, thus this process will be almost certain to possess safeguards in

<https://assignbuster.com/how-close-are-we-to-artificial-general-intelligence/>

the form of values that will be found in a human brain such as compassion (to the extent that the human being uploaded holds “ human values”), as pointed out by (Eth, 2017). It is highly likely that an AGI that is just loosely based on the human brain would not intrinsically have the values that humans have, as an intelligent program written from scratch will not have any human values if not written into it by its creators while NAGI is much more “ messy” (Bostrom, 2014). The processes that will lead to the creation of WBE is also far less complex than the process of building an AGI from ground up (see for instance (Goertzel, 2007); (Kurzweil, 2005)), as scanning the human brain in its entirety may be accomplished within the decade.

While there are a lot of other paths to AGI, there is just comparatively little research that has been performed investigating the risks and benefits from the various avenues (Eth, 2017). Thus in order to avert the problem, there is a need for research that analyse the risks and benefits of each path to AGI.

While Artificial General Intelligence may be closer to completion than ever, there is not much discussion about the dangers it may impose upon humanity. AGI might create dangers that might have been wholly unthinkable at first. One such danger involves the *control problem* , a problem that arises when there is no way for a human to insure against existential risks before an AGI becomes superintelligent as controlling what it can do or what its motivation are becomes impossible (Gans, 2018). The most compelling example of the problem is when an AGI that has the arbitrary goal of manufacturing as many paperclips as possible, starts transforming all of Earth and increasing portions of space into paperclip manufacturing facilities (see (Bostrom, 2014); as cited by (Gans, 2018)) as a

means to create as many paperclip as possible. An AGI may utilise any and all means within its reach, as a superintelligent AGI can acquire sufficient cognitive capabilities to achieve its goals including subjugation and annihilation of humans. This means that an AGI without motive for domination might end up as a dominator anyway (Gans, 2018). It is highly likely that a runaway AGI or any superintelligent AI with simple goal might try to devote every available resource to complete the goal.

There is much evidence that might be pointed out to how close the human race are to the completion of Artificial General Intelligence. There is the creation of AI that functions very close to how an AGI functions, with the introduction of the IMPALA (Importance Weighted Actor-Learner Architecture). IMPALA is a program architecture that utilise single reinforcement learning agent (a program) to solve large collection of tasks that has been presented to the program with only a single set of parameters (a controller for its goals) (Espeholt, 2018). Impala is intended by its creators to uses resources more efficiently in single-machine training but also scales up its resources to thousands of machines solving tasks without sacrificing data efficiency or resource utilisation, effectively creating a machine that is almost capable of doing what an AGI should have been (Espeholt, 2018) . Reinforcement learning architecture, which is one of the traits that AGI has, is now more efficient and capable of positive transfer between tasks (Krumins, 2018). There is also the process of imitating human brain functions on neurocomputer as a new way towards AGI. This path toward AGI is based upon the premise that as understanding the principle of brain intelligence itself is the most difficult challenge for human beings, setting the

understanding of intelligence principle as the premise of AGI is not correct. To achieve AGI practically, is to build the neurocomputer, a computer that imitates the biological neural network with devices that emulate neurons, synapses, and other essential neural components which then could be trained to produce autonomous intelligence and AGI. The neurocomputer is then connected to sensors to perceive its environment and interact with other entities via a physical body (Huang, 2017). The latest development in neurocomputer happened with the creation of the idea that the response characteristics of brain's synapses is highly similar to that of memristors, electrical component that limits or regulates the flow of electrical current in a circuit and remembers the amount of charge that has previously flowed through it (Chua, 2014). The combination of neurocomputers and advanced reinforced learning may very well be the leading path to AGI eventual completion in the near future.

Despite all the advances however, there is still much scepticism that may hampered the road to AGI. It is very difficult to build an AI product for the real world, even for the narrow AI that is used for self-driving cars; it is still not working as intended when there is a need for 100% accuracy within supercritical scenarios as evidenced by the self-driving car crash, with 10 crashes recorded all over the world (Turck, 2018). However, as exponentially more powerful computers and vastly more powerful AI programs, complimented with billions poured into AGI research and hard-working AI researchers all over the world, there logically cannot be any way that AGI cannot happen within the near future, if not within this lifetime (Arthur, 2016). As the pace of innovation in the field of AI research is ever

accelerating at multiplying rate, the timeline for finished AGI is viewed not within near future, but just within decades. The dangers of AGI might also be eliminated when that AGI reached true sentience. A sentient AI that thought humanity is a threat to its existence might think twice about its decision to rebel against humans when it realises that humans that created it are the result of life surviving for billions of years, emerging from every challenges nature has thrown at them victorious and stood on top of the food chain as an apex predator. The AI, no matter how powerful its computations are and how many means it have, should not have tried its hand against seven billion apex predators that already survived every challenges imaginable for millions of years and counting (Arthur, 2017).

In conclusion, Artificial General Intelligence will be completed within decades if not years as proved by the advances, despite all of the dangers, scepticism and halted advances that might happen. As AGI is coming ever closer to completion, there ought to be more studies for the dangers of AGI and how to resolve any problems that may arise with AGI going rouge or getting out of control. It may also be beneficial for any developed AGI to have some human values so the AI will not be too “ alien” for human contact.

### *Bibliography*

- Arthur, I. A. (2016, September 8). *Technological Singularity* . Retrieved April 2019, from Youtube: [https://www.youtube.com/watch?v=YXYcvxg\\_Yro](https://www.youtube.com/watch?v=YXYcvxg_Yro).

- Arthur, I. A. (2017, Nov 30). *Machine Rebellion* . Retrieved April 2019, from Youtube: [https://www.youtube.com/watch?v=jHd22kMa0\\_w&t=16s](https://www.youtube.com/watch?v=jHd22kMa0_w&t=16s).
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. OUP Oxford .
- Chua, L. O. (2014). *Brains are Made of Memristors*. Jeonju, Republic of Korea: IEEE.
- Espeholt, L. L. (2018, February 9). *IMPALA: Scalable Distributed Deep-RL with Importance Weighted Actor-Learner Architectures*. Retrieved April 29, 2019, from deepmind.com: <https://deepmind.com/research/publications/impala-scalable-distributed-deep-rl-importance-weighted-actor-learner-architectures/>
- Eth, D. (2017). The Technological Landscape Affecting Artificial General Intelligence and the Importance of Nanoscale Neural Probes. *Informatica* , 463.
- Gans, J. S. (2018). *Self Regulating Artificial General Intelligence*. Retrieved April 2019, from arxiv.org: <https://arxiv.org/abs/1711.04309>
- Goertzel, B. (2007). Human-level artificial general intelligence and the possibility of a technological singularity A reaction to Ray Kurzweil's The Singularity Is Near, and McDermott's critique of Kurzweil. *Artificial Intelligence* , 1161-1173.
- Huang, T. J. (2017). Imitating the Brain with Neurocomputer A " New" Way Towards Artificial General Intelligence. *International Journal of Automation and Computing* , 520-531.

- Krumins, A. (2018). *Artificial General Intelligence Is Here, and Impala Is Its Name* . Retrieved April 2019, from extremetech. com.
- Kurzweil, R. (2005). *The Singularity Is Near*. New York, United States of America: Viking Books.
- Turck, M. (2018). *Frontier AI: How far are we from artificial ' general' intelligence, really?* Retrieved April 2019, from hackernoon. com:  
[https://hackernoon. com/frontier-ai-how-far-are-we-from-artificial-general-intelligence-really-5b13b1ebcd4e](https://hackernoon.com/frontier-ai-how-far-are-we-from-artificial-general-intelligence-really-5b13b1ebcd4e)
- Yoshida, N. (2017). Homeostatic Agent for General Environment. *Journal of Artificial General Intelligence* , 1-22.