FROM HUMAN-LEVEL ARTIFICIAL INTELLIGENCE TO SUPERINTELLIGENT ARTIFICIAL INTELLIGENCE: SCENARIOS, PATHS, AND COUNTER ARGUMENTS

Wesley Y. Tian
University of Michigan
12/21/2017

1. Introduction

In a recent talk at the 2017 Beneficial AI Conference, panelists Bart Selman, David Chalmers, Elon Musk, Jaan Tallinn, Nick Bostrom, Ray Kurzweil, Stuart Russell, Sam Harris, and Demis Hassabis were asked to answer the question: *"Once we get to human-level AI, what do you think is the sort of timescale involved in reaching superintelligence?"* ("Superintelligence: Science or Fiction?", 2017). David Chalmers and Elon Musk answered that it would likely happen within days of achieving human-level AI, while Jaan Tallinn, Sam Harris, and Nick Bostrom said it would take years. On the other hand, Toby Walsh and Theodore Modis, not part of the panelists, have stated that superintelligence is unlikely to ever happen (Walsh, 2016), (Modis, 2012). However, Demis Hassabis offers an alternative–it would depend on the type of cognitive architecture that the human-level AI is based on. This was my thinking exactly and provides the basis to my paper. I will discuss the different kinds of human-level AI, how each kind may affect the time it takes to create superintelligent AI and rebuttals against the naysayers.

2. Background

Superintelligent AI will most likely be achieved through human-level AI systems. We will assume the premise that human-level AI will exist one day as there are no physical laws currently imaginable that will prevent us from creating it besides lack of research.

Since the birth of human civilization about 200,000 years ago, the majority of mankind's greatest technological advances have been achieved within the past 50 years (Susan, 2004). At this current velocity of growth, many experts and business leaders in AI predict that it will not be long until a human-level AI is created. Once that happens, it won't be long until superintelligence is developed, and then the technological singularity.

2.1 The Technological Singularity

The technological singularity is named after the singularity in a black hole due to their similarities in nature. When each of these respective points in spacetime is crossed, there is no turning back. In a black hole, it is due to the crossing of the event horizon. In technology, it is because it will trigger runaway growth that will lead to irreversible and unforeseeable technological advances that could drastically change the fate of humanity. The creation of human-level AI is commonly regarded as the first phase of the technological singularity. Although there have been predictions on when this event will occur, such as 2045 by Kurzweil (2006), the time it may take to create this technology is highly variable. The second phase of the technological singularity is the step from human-level AI to superhuman-level AI or superintelligence.

Besides human-level AI, there are many other possible paths to the technological singularity such as eugenics, nanotechnology, and a new form of human-computer interface (HCI) that would allow humans to access information near the speed of light (Kurzweil, 2016). These technologies are surely interesting and may also be possible, but by creating a human-level AI, we would solve human intelligence and have the extra brain power to research other technologies at an increased rate. Once human-level AI is created, we would be able to speed up the research performed by these AIs and produce research results in all other fields. Currently, almost all of the leading researchers in AI and believe superintelligence is the most likely path to the technological singularity.

For the purposes of this paper, I will treat superintelligence and the technological singularity as one because superintelligence will trigger runaway technological growth that will result in inexplicable and irreversible changes. It is also one of the most discussed scenarios for achieving the technological singularity.

2.2. Artificial Intelligence

As the field of AI continues to develop, one transition that we are currently seeing is from application-specific AI to general AI. Currently, applications of narrow AI are responsible for producing your Facebook news feed, personalized advertisements, and the computer vision software in self-driving cars. Other applications of narrow application-specific AI include IBM's Deep Blue computer program which beat Garry Kasparov, the former world champion in chess in May 1997 (McPhee et. al., 2015). Recently, Alphabet Inc.'s Google DeepMind AlphaGo computer program also beat the world champion Lee Sedol in the ancient game of Go ("Google's AlphaGo beats Go master," 2016). So far, we have been incredibly successful in developing this type of AI. In contrast, we have not been as successful in developing general AI.

General AI, artificial general intelligence (AGI), or human-level AI is defined by Murray Shanahan (2015) as an AI that "*can match the performance of an average human in all, or nearly all, spheres of intellectual activity*". Both computer scientists and cognitive scientists find this type of AI to be incredibly difficult to create. Instead of discussing the paths to human-level AI–for the sake of this paper–we will assume the premise of human-level AI technology and focus on what could happen thereafter.

3. Human-Level AI Scenarios

Before we get ahead of ourselves, it is natural to assume that the paths to superintelligent AI will depend on the availability of human-level AI. Without human-level AI, superintelligent AI may still be possible, but highly unlikely. A similar analogy would be the unlikely creation of a MacBook Pro before the invention of the ENIAC vacuum computer. Although some countries in Africa can leapfrog the innovation gap and jump straight to mobile phones without needing to invent landline phones, ("What technology can do for Africa", 2017) it is uncertain and impossible to predict whether that will happen to the human race.

Human-level AI may be created in various ways and each realization of human-level AI may inherently lead to a different path to superintelligent AI. For example, a primarily algorithmic human-level AI may be easily duplicated leading to collective superintelligence, while this may not be possible for a primarily hardware-based human-level AI that requires many more resources to duplicate. Biological human-level AIs have the same biological limitations that humans do and so this type of human-level AI developing into superintelligence is unlikely (Bostrom 2014). Software-based human-level AIs may differ vastly from one another, and each specific detail in its design may lead to different scenarios. We will explore different types of these artificial human-level AI in detail, see how each type may lead to superintelligent AI, and their challenges.

3.1 Software-Based

A piece of software cannot exist without its hardware counterpart. At its core, software is just the manipulation of bits through transistors and circuits. However, a software-dominant human-level AI may require less physical resources than a hardware-dominant one. In current AI research, we have developed machine learning techniques such as deep learning and neural networks that can help us classify objects and generate novel pictures that are almost indistinguishable from real pictures. Deep learning can be used to identify the objects in a picture. However, in order to do this accurately, deep learning models require billions of training examples (Shazeer et. al., 2017) and immense parallelized processing power to train the models and help them make inferences. Even if we have the storage capacity for the data, it still requires multiple GPUs and a couple of hours to train such a model. A software-based human-level AI system could use more efficient algorithms and take advantage of parallel processing capabilities to reduce the time needed to train massive neural networks. (Chung et. al., 2016) It may be possible for such algorithm to reduce the runtime and processing power of current programs by multiple orders of magnitude.

If we take a look at our only known example of intelligence, the human brain is much more algorithmically efficient than our current machine learning techniques. A toddler only needs a to have seen a few dogs and cats to be able to distinguish a cat from a dog with high accuracy. If a machine learning algorithm only required as many examples as a toddler, then we could forgo the large storage capacity required to store all the training data.

Certain programming languages can also improve the efficiency of a system. Compiled languages such as Fortran, C, and C++ are able to make faster calculations than interpreted languages such as Python. (Bright, 2014) To turn the code into instructions in the registers at the hardware level, a computer must have programs that interpret the code. For a compiled language such as C++, the source code is first parsed to detect typos and semantic mistakes such as calling a function that doesn't exist. Next, a code generator is used to produce executable code that is then executed by the machine. An interpreted language such as Python, on the other hand, must go through additional steps such as looking up functions already contained in the language before being executed by the program. Each additional layer will reduce performance. In general, Python code is less verbose and simpler to learn and write for the typical software engineer and offers more support via libraries. C++ requires more boilerplate code but will run faster. Programming languages' seem to have a tradeoff between ease of use and performance. Modern approaches to programming have evolved from punchers to high-level languages such as Octave in order to reduce the learning curve for programming. Once programs have been algorithmically verified and written, it is not unusual to see a program translated into more efficient languages. Advances in compilers and programming languages may be able to further speed up the runtime at the software level.

Another factor related to increased computational efficiency is an algorithm's complexity. In computer science terms, the worst time complexity of an algorithm is described used what is called *big O notation*. Intuitively, the big O notation of an algorithm will describe the runtime growth rate as the number of data increases. For example, an algorithm that sorts a hand of cards that iterates through the cards and handpicks the card with the largest value and places it on the far right has a big O notation of $O(n^2)$. Conceptually, this means that each additional card in the hand will increase the runtime by a factor of n^2. If sorting 2 cards using this algorithm used 4 operations, sorting 100 cards would use 10,000 operations. A more efficient algorithm such as quicksort, which I will not go into detail here, has a more efficient runtime of $O(n \log_2 n)$. Sorting 2 cards with this algorithm would use 0.6 operations while sorting 100 cards using quicksort would use only 200 operations. As the number of cards increase, it is obvious how increases in algorithmic complexity could affect the runtime. While multiplying matrices has a complexity bounded by $O(n^3)$, using Virginia Williams' algorithm based on the Coppersmith-Winograd construction (2014) can result in an bound of $O(n^{2.373})$. It should be noted that boosted algorithms such as Williams' has a large constant and will only be effective under specific circumstances. Improvements in algorithmic complexity can further speed up the efficiency of a software-based human-level AI.

If the software for this type of AI requires less data and is more efficient to run, then such a human-level AI will not require nearly as much computing power. The extent to which our algorithms will be able to reduce the amount of hardware necessary for creating human-level AI remains unknown, but if both the software and hardware capabilities for a human-level AI system become mature enough, then AI becomes just another program like Microsoft Word. We could then we can imagine a situation where we easily duplicate these programs. Even if the application were very large and needed a few exabytes of storage, with our current growth in storage capacity, it wouldn't be long before we will be able to have a computing device that could easily store the AI system. If we think back to floppy disks, a large flat rectangular storage device could only contain up to a few hundred kilobytes of data. It would have been unthinkable to use those storage devices and transfer even a high-resolution picture which may take up to ten megabytes of storage.

A software-based AI that is extremely efficient can come in various forms. A human-level AI can be imagined to require varying orders of magnitudes of processing power and storage. It can be computationally efficient to run on portable computing, specialized hardware such as a home desktop computer, a supercomputer such as China's Sunway TaihuLight, IBM Sequoia, a megastructure such as the Large Hadron Collider or the Dyson Sphere (Stapledon , 1937). Each of these varying levels of computing power are necessary for human-level AI and would greatly alter the path leading up to superintelligent AI. For the sake of simplicity, we will look at two cases. The first is a human-level AI that is accessible to the general public. Although it may require special hardware, anybody who is decently well-off would be able to purchase such a system. Next, we will imagine the case where a human-level AI requires the resources that only the largest corporations or countries could afford. For future reference, let us denote the first type of AI as Type S1 for the smaller software-based human-level AI and Type S2 for the larger software-based human-level AI.

There are plenty of ways the software could become efficient enough to reduce the load on the hardware, but more importantly, software-based human-level AI requires a programmed cognitive architecture that may utilize current and new AI and machine learning techniques. The blueprint to such a cognitive architecture that has human-level intelligence is still in the workings, but as long as it is like other programs that we have created thus far then such an AI program should also be feasible.

The difference between software-based AI and hardware-based AI is that software-based AI is not reliant on a special type of hardware. As long as our current machines continue advancing, the software is a possible solution to creating a human-level AI.

3.2 Hardware-Based

The alternate realization of a human-level AI may not result from advances in software but enhancements in hardware. The most straightforward hardware-based approach to human-level AI is whole brain emulation.

If we obtain ample processing power to fully simulate the brain and create hardware components that mimic each component in the human brain, then it seems reasonable to assume that we would then have a machine that acts just like the brain.

This hardware-based approach of human-level AI could be accomplished through Copy-and-Transfer, which involves mind uploading through scanning and mapping the salient features of a biological brain and then copying, transferring and storing that information state into a computer system. The simulated mind could then reside within a virtual reality or simulated world.

When asked whether simulating the entire human brain was possible, Henry Markam (2007), lead researcher of the "Blue Brain Project" replied:

> *"It will be very difficult because, in the brain, every molecule is a powerful computer and we would need to simulate the structure and function of trillions upon trillions of molecules as well as all the rules that govern how they interact. You would literally need computers that are trillions of times bigger and faster than anything existing today."*

An approach like this would be different from the Types S1 human-level AI as it would obviously require specialized and intensive hardware that most people would not have access to. Once successfully developed, many of these whole brain emulation supercomputers could be owned by select companies and countries. They could be distributed to communities and act as a community librarian as imagined in the future in The Time Machine (Wells 1995). For future reference, let us denote this type of human-level AI as Type H.

3.3 Software and Hardware-Based

The approaches discussed previously are either software focused or hardware focused. However, there is another approach that contains both specialized software and hardware that would need tremendous research efforts to make a reality. This differs from the previous approaches because we cannot just continue advancing in the same areas that we are currently developing in order to achieve this approach. We cannot just innovate horizontally by improving what we already know. We must innovate vertically and come up with entirely novel ideas and technologies (Thiel and Masters, 2014). According to Thiel,

> "*Horizontal or extensive progress means copying things that work -- going from one to n. Horizontal progress is easy to imagine because we already know what it looks like. Vertical or intensive progress means doing new things, going from zero to one."*

This software and hardware approach has software that is able to mimic the human brain but has a robotic body in order to make it a complete human-level AI. This approach comes from the embodied cognition theory which states that there are many features of cognition that are entirely dependent on the body of the organism. Thus, in order to create human-level AI, it may be necessary to create both the mind and the body. Similar to the case with the software-based approaches, not only would such a realization of human-level AI require research in cognitive architectures but could also require anthropomorphic features such as senses and actions to become a human-level AI.

Why is embodiment necessary? For starters, we have the computation required to simulate the processing power of the brain of a bee, yet we still have not been able to create a simulation of a bee. Besides the lack of the bee's embodiment, other reasons for this lack of progress could be due to an oversimplified model of the neural cells and a poor understanding of higher cognitive processes. (Clark and Chalmers, 1998)

This kind of human-level AI could fall somewhere between Type S1 and Type S2 in terms of accessibility and affordability after some amount of time. Initially, it may seem plausible that there would only be prototypes developed by a large research company and would be extremely expensive to manufacture. This kind of AI would be least likely to advance into superintelligent AI quickly. If the cost of such a robot decreased over time as the economies of scale came into play; it may even be possible to have these human-level AIs working day and night to create more and more of themselves. Duplication of such a human-level AI would not lead to superintelligent AI, but is much more likely to result in recursive self-improvement.

4. Paths to Superintelligence

4.1 Duplication

As with any other piece of software, a human-level AI system such as Type S1 could be easily copied. Since Type S1 requires inexpensive hardware, a wealthy person would be able to own many of these systems. A large corporation or country could fund thousands to millions of these systems.

Similar to a human, a human-level AI may be an expert in some fields and be fairly average in other fields. As AIs are not biological, they could be programmed and trained to suit the needs of the user. Additionally, an AI would have capabilities beyond a normal human being such as being able to access all of the data on the world wide web, having a direct interface with computers, and being able to work in a different timescale compared to regular humans. These human-level AIs would not need basic human needs such as sleeping, eating, and other bodily functions to survive, thus giving them additional time to work and learn. These AI systems could also be sped up in a virtual environment. Doing so could increase their learning rate or productivity to many times faster than a normal human being.

The following scenario is meant to show how a collective group of human-level AI can display superintelligence. It takes inspiration from Murray Shanahan (2015), but is more realistically detailed and brings together Bostrom's knowledge on collective superintelligence (2014). Imagine two companies have one year to bring a new laptop computer to market. One of the companies is a large multinational organization with over fifty thousand of employees. Let's call this company ECorp. ECorp has many products, patents, and previous generations of laptops. The other company is a small start-up with human-level AI technology. Let's call this company AICorp. ECorp decides to split up and delegate different tasks involved in creating a new laptop to different sectors of the company. AICorp decides to generate ten off-the-shelf human-level AI systems that start with the knowledge and skills of an average college graduate. AICorp informs the AI systems about the task at hand and enrolls them in their relevant virtual environment graduate schools. Able to be sped up inside the virtual environment, AICorp's employees are able to obtain the same education as the employees of the larger company in only a few weeks of real time. The human level AI systems do not require food and rest compared to the human employees and are capable of working and learning 24 hours a day in their virtual environment. Over just a short amount of time, this leads to a dramatic increase in engineering and research outcomes. Though ECorp had many more resources to begin with, and was able to draw on the knowledge of its past products, AICorp is not hindered, but instead has an advantage of having an entirely open playing-field to stretch their expertise and innovation.

When a year has passed, and it is finally time for both companies to release their respective products, ECorp releases a new generation of laptops, improving from its previous iteration. It has a better processor, the option to upgrade to 32 cores, a 2TB SSD, a thinner and lighter body, a screen with a better resolution and brighter colors, and other standard incremental improvements commonly seen in a new generation of laptops. For a regular company, a year's time isn't much. ECorp's new laptop may have some new features such as FaceID or TouchID as featured in the iPhone, or a change to an OLED display from an IPS display, or include wireless charging. However, these new features are still incremental and are seen as linear improvements.

Now, let's see what AICorp is up to. Using their newly educated human-level AIs and much more simulated time than the ECorp, they are able to reinvent the laptop. They developed new technologies in almost every domain involved in the development of a new laptop. They made breakthroughs in material design, hardware, and software. The laptop looks like something from the future and looks nothing like the laptops we've used to seeing. The body is made up of a material with varying plasticity and can be rolled up like a poster, yet still has a vibrant screen. The laptop is built so thin that it seems like the hardware has nowhere to go, yet it still has at least an order of magnitude more processing power and battery capacity than current generation laptops. It can access information much faster than normal computers through a new type of RAM that is non-volatile. Additionally, it is able to read the user's brainwaves as a new form of brain-computer interface. A person can simply look at this magical paper-like device and it will do whatever they want it to do. This technology seems to work seamlessly with modern users without needing a learning curve because of how well it was designed.

AICorp didn't just incrementally improve a computer, they reinvented it using technologies that ECorp wasn't even close to developing. By having virtually an infinite amount of simulated time, AICorp was able to solve research problems in multiple domains and use all of the solutions in their new product.

What this example shows is that even though each human-level AI individually is no different from a normal human being, collectively, they display superhuman-like intelligence. To us, the new technology looks like magic. It would be as if Thomas Edison were shown the modern day laptop. To him, it would look like something that only a man with superhuman intelligence could have created. To us, many human-level AI systems working together collectively show superhuman level intelligence.

4.2 Recursive Self-Improvement

Another path to superintelligent AI is through recursive self-improvement.

One example of recursive self-improvement is in humans. Treff, an in vitro fertilization specialist, has recently helped advance eugenics enough to be able to predict the IQs, genetic defects, and diseases of embryos (Regalado, 2017). Over time, this artificial selection of infants will lead to smarter humans who then grow up and help advance related fields. Each generation of humans will improve their ability to enhance further generations. Humans are currently experiencing recursive self-improvement in a lengthened timespan because we have no control over our biological growth factors and it takes time for humans to mature to a stage where we can contribute to improving ourselves. However, AI systems would not have this limitation and could potentially go through rapid iterations of recursive self-improvement within a very short amount of time.

AI systems who are as intelligent as the humans who developed the program could build on AI, then the AI itself would be able to build on itself. An AI that is intelligent enough to redesign and update itself is known as a *Seed AI* (Yampolskiy, 2015). A *Seed AI* that has the engineering capability that matches or surpasses its creators would then have the potential to upgrade its own hardware or software. This more capable machine could then go on to develop a machine that has an even greater capability. These iterations of recursive self-improvement could accelerate, potentially allowing enormous quantitative changes before any upper limits imposed by the law of physics set in.

Even if this process is slow at first, because non-biological AI systems have no biological needs such as sleeping and eating and can be easily replicated; recursive self-improvement is another possible path for superintelligence.

5. Counter Arguments

5.1 The "Meta-intelligence" argument

According to Toby Walsh (2016), a strong critic against the technological singularity and superintelligent AI, the intelligence needed to perform a task is confused with the capability to improve the intelligence to do a task. Walsh claims that this is one of the strongest arguments against the idea of a technological singularity. He argues that since machine learning is likely to be a part of a human level AI system and frequently tops out at particular tasks, no amount of tweaking, be it feature engineering or parameter tuning, appears able to enhance their learning ability. Therefore, human-level AI systems will also be unable to enhance their learning ability. Using deep learning techniques to recognize speech or identify objects has not lead to an improvement in deep learning itself. Walsh also uses humans as an example. Our IQ has only slowly increased over the last century, so he ponders that perhaps electronic brains will also struggle to boost their performance quickly and never get beyond a fraction of their fundamental capabilities (Walsh, 2016).

There are a few rebuttals worth mentioning here. First, machine learning algorithms may only be an insignificant part of a human level AI system. He bases his claim that it is "indeed likely" for a human level AI system to contain machine learning algorithms. However, it is just as likely that a human level AI system does not contain machine learning algorithms. If we create a human level AI system using Kurzweil's proposed method of whole brain emulation, then there would be no need for neural networks or large datasets (Kurzweil, 2006). However, we can grant Walsh the premise of machine learning being part of a human level AI system, as it does not matter to the main content of my rebuttal.

Walsh seems to think that because we are experiencing diminishing returns for scoring well on speech recognition or object classification tasks, that it will not be possible to advance intelligence in a human level AI system dramatically. Additionally, improving scores on these tasks through deep learning have not improved themselves. First, a human level AI system does not need to achieve higher scores in specific tasks such as object recognition and speed recognition to transition to show superintelligence. As explained in 4.1, human level AI systems may display superintelligent like abilities when working collectively and/or at an increased pace.

Second, AutoML, a machine learning effort made by Google, has proved that a neural network can tune other neural networks better than humans can (Zoph, et. al, 2017). So, although Walsh is right in saying that deep learning techniques for object classification have not advanced deep learning, advances in one field in machine learning can help improve performance in other fields. Therefore, it is not necessary for a deep learning algorithm to improve its own ability to learn. It is only necessary for an algorithm within a program to advance another algorithm which can then advance the initial algorithm. This is another form of recursive self-improvement.

5.2 The "Fast Thinking Dog" Argument

According to Walsh, the argument put forward by proponents of the technological singularity is that silicon has a significant speed advantage over our brain's wetware, and this advantage doubles every two years or so according to Moore's Law (Walsh, 2016). However, speed does not bring increased intelligence.

Pinker (2008) sums it up well:

> *"There is not the slightest reason to believe in a coming singularity. The fact that you can visualize a future in your imagination is not evidence that it is likely or even possible. Look at domed cities, jet-pack commuting, underwater cities, mile-high buildings, and nuclear-powered automobiles all staples of futuristic fantasies when I was a child that has never arrived. Sheer processing power is not a pixie dust that magically solves all your problems."*

There is a significant difference between superintelligence and the fantasies (domed cities and jet-pack commuting) that Pinker came up with. Just because some fantasies have never arrived, does not mean that all other fantasies will not arrive. Some "fantasies" dreamt up by people in the 19th century have become a reality. Personal drones, self-driving cars, hoverboards, and biometric devices are just a few examples of "fantasies" that were actualized. What determines a fantasy's actualization is its current demand and forecasted value to society. Pinker's examples of fantasies such as domed cities and jet-pack commuting are not inventions that bring value to society. Humans can barely navigate in 2D space. Jet-packing commuting would involve navigating in 3D space and would be a hazard to both humans wearing the jet-packs and pedestrians. However, self-driving cars and personal drones have many practical uses such as autonomous fleets and drone delivery services. There's no practical reason for domed cities either. Only fantasies that are predicted to be valuable to society call for research and development.

Human-level AI is a different *fantasy* than the ones Pinker mentioned because it is forecasted to have a monumental impact to society. Creating human-level AI is solving intelligence and doing so will essentially enable us to speed up research and development for all other technologies. It would be able to automate millions of low-level jobs and create more advanced jobs such as those in the mobile app industry created from the automation of farming. Therefore, because human-level AI is predicted to be extremely valuable society, it is much more likely to become a reality.

Additionally, supporters of the technological singularity don't just believe that sheer processing power will bring about superintelligence. Superintelligence is created from a crucial software component and possibly a hardware component which I went into detail previously. It's accepted that more than just computing power is necessary to create human-level AI or superintelligence. Nobody is saying that sheer processing power is all you need. I would agree with Pinker and Walsh that processing power is part of the equation, but it is not the entire equation. In addition to other technological advances in domains such as cognitive architecture and algorithms, I see superintelligence and the technological singularity as certainly possible.

6. Conclusion

There are many different paths that can lead to the technological singularity. One possible path that involves artificial intelligence could happen in two stages. First, technology for human-level AI needs to be created. Next, human-level AI may evolve into superintelligence through ways such as collective superintelligence or recursive self-improvement.

In this paper, I devise different possible categories of human-level AI in order to create more plausible scenarios for the evolution of superintelligence. I make distinctions amongst four different kinds of AI systems–two software-based systems, a hardware-based system, and a system that is both software and hardware-based. Because the topic of this paper is speculative in nature, devising these types allows me to build up more accurate scenarios depending on the type of human-level AI that will one day be developed. For example, human-level AI systems that are software-based such as Type S1 may lead the way to collective superintelligence more easily than Type S2 human-level AI systems.

After detailing these possible types of human-level AI systems, I introduce two different ways in which a society with human-level AI technology could develop superintelligent AI. If we assume Type S1 AI systems, then a collective group of them could display superintelligent behavior while the Type H system could more easily develop superintelligence through rapid iterations of recursive self-improvement.

Lastly, I tackle some of the existing counter arguments presented by naysayers of superintelligent AI technology and the technological singularity. This paper does not encompass all of the possible realizations of human-level AI or the paths that can lead up to the technological singularity as there are infinitely many possible scenarios. However, I try to generalize by creating four primary categories of human-level AI systems to build off of.

Due to a lack of a clear path towards human-level AI, there is a low probability of predicting which type of AI system will be engineered. Since the timescale to develop superintelligence is conditional on the type of AI system developed, there is also a low probability of predicting if and when the technological singularity will occur. Therefore, this paper isn't meant to convince you whether the technological will or will not happen or when it will happen but rather offer insights on how the timescale to reach superintelligence from human-level AI changes from scenario to scenario.

7. References

[Allen and Greaves 2011] Allen, P., and Greaves, M. 2011. The singularity isn't near. *MIT Technology Review* 7-65.

[Bostrom 2014] Bostrom, N. 2014. *Superintelligence: Paths, Dangers, Strategies.* Oxford, UK: Oxford University Press.

[Bright 2014] Bright, P. 2014 *Ask Ars: Why are some programming languages faster than others?* Ars Technica. Retrieved from https://arstechnica.com/information-technology/2014/05/ask-ars-why-are-some-programming-languages-faster-than-others/

[Chung et. al. 2016] Chung, I., Sainath, T. N., Ramabhadran, B., Picheny, M., Gunnels, J., Austel, V., Chauhari, U., and Kingsbury, B. *Parallel Deep Neural Network Training for Big Data on Blue Gene/Q*. 2016. IEEE Vol. 28: Issue 6: *pp. 1703-1714.* Retrieved from http://ieeexplore.ieee.org/document/7738586/

[Clark and Chalmers 1998] Clark, A., and Chalmers, D. J. 1998. *"The extended mind".* Analysis. 58 (1) 7-19. MIT Press. *pp. 27-42.*

["Google's AlphaGo beats Go master" 2016] *"Artificial Intelligence: Google's AlphaGo beats Go master Lee Se-dol".* 2016. BBC News. Retrieved from http://www.bbc.com/news/technology-35785875

[Markram 2007] Markram, H. Blue Brain Project FAQ Archived. 2007. Retrieved from https://web.archive.org/web/20070127195551/http://bluebrain.epfl.ch/page18924.html

[Kurzweil 2006] Kurzweil, R. 2006. *The Singularity Is Near: When Humans Transcend Biology.* Penguin (Non-Classics).

[McPhee et. al. 2015] McPhee, M., Baker, K. C., and Siemaszko, C. *"Deep Blue, IBM's supercomputer, defeats chess champion Garry Kasparov in 1997"*. 2015. New York Daily News. Retrieved from http://www.nydailynews.com/news/world/kasparov-deep-blues-losingchess-champ-rooke-article-1.762264

[Modis 2012] Modis, T. "*Why the Singularity Cannot Happen*". 2012. *Singularity Hypothesis: A Scientific and Philosophical Assessment. pp. 311-346.* Berlin, DE: Springer Verlag.

[Pinker 2008] Pinker, S. 2008. Tech luminaries address singularity. *IEEE Spectrum.*

[Regalado 2017] Regalado, A. 2017. *"Eugenics 2.0: We're at the Dawn of Choosing Embryos by Health, Height and More"*. MIT Technology Review. Retrieved from https://www.technologyreview.com/s/609204/eugenics-20-were-at-the-dawn-of-choosing-embryos-by-health-height-and-more/

[Shanahan 2015] Shanahan, M. 2015 *The Technological Singularity*. MIT Press.

[Shazeer et. al. 2017] Shazeer, N., Mirhoseini, A., Maziarz, K., Davis, A., Le, Q., Hinton, G., and Dean, J. "*Outrageously large neural networks: The sparsely-gated mixture-of-experts layer".* 2017. Submitted to ICLR 2017.

[Stapledon 1937] Stapledon, O., *Star Maker*. 1937. London, UK: Methuen.

["Superintelligence: Science or Fiction?" 2017] "Superintelligence: Science or Fiction?" [Video file] Retrieved from https://www.youtube.com/watch?v=h0962biiZa4

[Susan 2004] Anton, S. C., "*Early Dispersals of homo from Africa".* 2004. Annual Review of Anthropology. *33: pp. 271-96*.

[Thiel and Masters 2014] Thiel, P., and Masters, B. *Zero to One: Notes on Startups, or How to Build the Future*. 2014. New York, US: Penguin Random House Company.

[Wells 1995] Wells, H. G. *The time machine.* 1995. Dover thrift editions.

["What technology can do for Africa" 2017] *"What technology can do for Africa".* 2017. The Economist. Retrieved from https://www.economist.com/news/special-report/21731038-technology-africa-making-huge-advances-says-jonathan-rosenthal-its-full

[Williams 2014] Williams, V. V., *Multiplying matrices in o(n2.373) time*. 2014. Stanford University.

[Yampolskiy 2015] Yampolskiy, R. V. *From Seed AI to Technological Singularity via Recursively Self-Improving Software.* 2015. arXiv:1502.06512 [cs.AI]. Retrieved from https://arxiv.org/abs/1502.06512

[Zoph, et. al 2017] Zoph, B., Vasudevan V., Shlens, J., and Le, Q. 2017, "*AutoML for large scale image classification and object detection".* Google Research Blog. Retrieved from https://research.googleblog.com/2017/11/automl-for-large-scale-image.html