Daniel Yount

Mr. Farmer

ENG 111

November 19, 2014

A.I.

Singularity. Derived from late Latin “singularis”, old French “singularite” and modern English “singular”, the word is undeniably well traveled. To mathematicians, a singularity is defined, naturally, from maths, as a point at which a function takes an infinite value, as at the center of a black hole; the point at which all known laws of physics break down into the void darkness of infinite eternity. To researchers in the field of Artificial Intelligence (A.I.), however, it breathes life into another meaning entirely. Singularity. A hypothetical moment in time when Ai and other technologies have become so advanced that humanity undergoes a dramatic and irreversible change. This is the ultimate culmination of the seemingly endless journey that life, as it exists, has been on. The parallels between A.I. and the mathematical studies of black holes do not end at similar word choices though. The singularity is, more universally, an expression used to describe an event horizon at which all logical preconceived notions and every ounce of knowledge that is of this world dissipates into impossible uncertainty. In the same way astrophysicists are utterly, mentally disarmed at physics pertaining to black holes, the field of A.I. musters a corresponding bewilderment from anyone and everyone who attempts to study it. However, unlike A.I., humanity is not racing towards the center of a black hole anytime soon. The pace at which technology is recursively improving is nigh

impractical, yet extraordinarily predicable. Researchers, such as Ray Kurzweil of Google, have consistently indicated the world is a mere fourteen years from the birth of a conscious computational being from silicon substrate (Kurzweil 435). In his most recent publication *How to Create a Mind,* he states the following:

> There is now a grand project under way involving many thousands of scientists and engineers working to understand the best example we have of an intelligent process: the human brain. It is arguably the most important effort in the history of the human-machine civilization. In *The Singularity Is Near* I made the case that one corollary of the law of accelerating returns is that other intelligent species are likely not to exist. To summarize the argument, if they existed we would have noticed them, given the relatively brief time that elapses between a civilization's possessing crude technology (consider that in 1850 the fastest way to send nationwide information was the Pony Express) to its possessing technology that can transcend its own planet. From this perspective, reverse-engineering the human brain may be regarded as the most important project in the universe. The goal of the project is to understand precisely how the human brain works, and then to use these revealed methods to better understand ourselves, to fix the brain when needed, and—most relevant to the subject of this book—to create even more intelligent machine (Kurzweil 26).

Other estimates, such as a short anecdote in the form of a tweet, a few days ago from serial entrepreneur Elon Musk (founder of PayPal, SpaceX, Solar City, and Tesla Motors), would, instead, indicate a mere five years until said birth. The consensus

among academia and Wall Street is clear. This is not a matter of if, but when. What is more, an A.I. system will be born much sooner than humanity will be ready to face the consequences that pertain to it. As was mentioned previously, it is almost impossible to look beyond a singularity, however, there is merit in imagining and preparing as best as possible for the most conceivably likely scenarios.

When discussing advancements in technology that effect the world at such a large scale, it is important to analyze the potential reaction modern economics and systems will have, if only to better understand how to contextualize the future moving forward. A.I. will certainly have an unprecedented effect on the economy. The moment humans create machines smarter than themselves, is the moment humans release control of their evolutionary destiny. Introducing a machine, which can think, operate, and sense at levels equal to or far beyond human limits completely obliterates the entire structure of modern economics. Why should a corporation continue to pay people for work? The people complain, go on strikes, have “needs” like food, water, and housing, and can be hard to work with. A computer, on the other hand, will work for as long as it needs to and offers seemingly endless advantages over humans. Ultimately, this type of systematic replacement of humans in the economy is to be encouraged. The long term benefits of a system where humans are not forced to work for a living are paramount to solving many social and individual physiological problems however, the short term implications of the quickly approaching change of systems will almost certainly crush people, unless properly planned for. As humans leave the work force, the cycle of capitalism is broken. Employers, by not employing people, yet saving money, paradoxically eliminate the source of income their business relies on: the people. If the people do not

work, thus make no money, who buys the products and services that the employers produce? This is a categorically systemic problem and as the world moves into a post capitalistic technological era, the social systems and economic systems must overhaul completely to meet this change. No amount of political band aid fixes are enough to force the system of capitalism to be compatible with a world in which A.I.s are a part of.

On a personal level, the fallout from the quick advancements that eliminate human workers is already being sensed in society today. Examples of automation exist everywhere, and are only becoming more prevalent as time marches onward. The primary concern, therefore, is to allow people the ability to transition to a new system in a way that causes minimal financial, physical and mental discomfort. Sweden's government is already in the process of tackling this problem before it fractures too deep into society and devastates families. In 2013, a grassroots campaign in Sweden brought the idea of a universal basic income to the halls of government. A universal basic income is a program that exists to primarily help people make the transition from our modern capitalistic economy to one where humans are not needed. The basic premise is, through taxation, a small amount of money which is deemed a livable amount, is given to every citizen regardless of the employment status, thus ensuring a safety net while mitigating the danger of incentivizing laziness and unemployment to reap benefits. The critics of this system often do not believe in the long term effects of A.I. (an extreme position considering the evidence to the contrary) and therefore have trouble seeing universal basic income as little more than government handouts, thus advocating for laziness and unemployment. Advocates, nonetheless, push onwards claiming that by constructing a system in which people are given the safety net amount regardless of their employment

status, people are not incentivized to stay unemployed and will search for work if what they seek is an increase in income however, it also allows the unemployable people (i.e. the people who want work but can't find any) the peace of mind that, under no circumstance, will they starve or fall under a certain threshold of livability. The UBI campaign is still ongoing, but it offers a look at some real world solutions to the problems in which humanity is about to come face to face with in the coming years.

The extreme job automation and systematic replacement of humans in the process of evolution may be a seemingly disparaging concept to mentally handle, but A.I. also brings outrageously miraculous possibilities that unarguably work to improve human life far beyond anything that has come before. One aspect, which is worth touching on, is the conceivability of digital immortally. Being that an A.I. is essentially an emulation of a human neuron network, there is an extremely likely scenario in which humans are able to upload their minds to a surrogate A.I. core and never fear a death of a biological origin ever again. Again, this is not simply unfounded predictions based on dreams of grandeur, but real science providing the basis of physics at which this idea operates.

The concept of mind uploading enabling digital immorality is so complex and so amazingly multifaceted, it would take, at minimum, an entire book to do the theories, technical research, physiological studies, and questions justice. For example, assuming that digital immorality is not simply a goose chase into frivolous territories, how can someone be sure the mind they upload into a silicon substrate is them and not simply a copy? Put simply, what makes you, *you*? Such a simple question is surprisingly hard to answer. A classical thought experiment known as the ship of Theseus thought experi-

ment, illustrates the quandaries associated with personal identity and parallels the struggles of certainty in mind uploading. The details of this thought experiment cannot be referenced without first referencing the whole of the thought process:

> Imagine you have a wooden ship in your backyard and decide that one day you want to take it apart piece by piece. You go about taking it apart, delicately removing each part as if you were going to use those same parts to rebuild it later (should you decide to do so) and storing it in your garage. Now, here's part of the problem. When you remove each piece, it still seems to be (intuitively) the same Ship. Removing one plank from the floor, for example, doesn't seem to make it a different ship. However, if you keep up with the process, you will end up with all of the pieces of the Ship in your garage, but it will no longer be the Ship of Theseus, since, by hypothesization, you won't have a ship – you'll have a pile of wood that used to be a ship. But at some point in your deconstruction, the Ship had to move from existence to non-existence, unless you want to say all the pieces in the garage is the Ship. At what discrete point did the Ship cease being the Ship? A further wrinkle: suppose that, rather than just tearing the Ship apart, you decide to replace every wooden piece you removed with an aluminum piece of the exact same dimensions. So, when you start, you have a completely wooden Ship, but at the end, you have a completely aluminum Ship. But, at each discrete stage of time, you only have a ship that is one piece different than it was in the previous moment. An even further problem: suppose that you decide to use the wooden planks you re-

> moved in the case above to build another Ship which is materially identical to the original ship. At the end of that project, you'll have two Ships, one aluminum and one wooden, that each have a claim to being the Ship of Theseus. They can't both be the Ship of Theseus, but it could be true that they both could not be the Ship of Theseus, but the problem becomes, when was the Ship of Theseus destroyed? (Productive Flourishing).

Mind uploading follows a similar paradox of thought as the "Ship of Theseus" thought experiment and could very well be the most important long term effect of an A.I.'s birth into the world. Humans are still trying to grasp the concept of consciousness form the clutches of scientific uncertainty. For some, consciousness is what it is by the individual parts that make it up. The first time a plank of wood is removed from the ship, it ceased to be. For others, consciousness is what it is because of its structure, like a river that flows: ever changing, but still in essence a river. Finally, some believe, consciousness is what it is because of its history. Past experiences and memories make personal identity. The discussion and debate could go on forever, yet it is important to not overlook the key consequence: For the first time in human history, a conceivable form of immortality presents itself in the form of technology and will unequivocally change the entire definition of what it means to be human. But what if A.I.s have their own minds made up about the fate of the human race? How would humans react to an apocalypse at the hands of intelligences billions and trillions of times removed from human intelligence?

Science fiction, movies, comic books, and practically any other form of pop culture have tended to paint A.I.s in a negative light, most often detailing a rise to power

and eventual mutiny against humans, its creators. But is this fear rationally founded? Or simply the product of a hyperactive imagination? The bottom line? This is a question without an answer as of now. It is impossible to predict the personality of an A.I. when researchers are just now contemplating the physics behind whether it is at all possible to build an intelligence. That being said, the likelihood of human success at quelling a robotic A.I. uprising is so small it is beyond laughable. An A.I. by its very definition is so far removed from anything humans have ever experienced on earth before that should a war ever start between humans and just a single A.I., the battle would be over in hours, more likely in minutes. Such is the nature of the beast humans are destined to release, however, we cannot make sweeping generalizations about such a yet unknown piece of the puzzle. Some questions are only answered in practice, not in theory.

In the wake of an A.I. birthed into the world, humanity will be thrown into a singularity and be changed forever. The consequences of this, while yet to materialize, are nevertheless important to postulate and prepare for. A.I. could very well be the last great act of humanity as the torch of progress is based from biology to technology. It is the ultimate effect of humanity.

Works Cited

Kurzweil, Ray. How to Create a Mind: The Secret of Human Thought Revealed. New York: Viking, 2012. Print.

"The Ship of Theseus and Personal Identity." *Productive Flourishing*. Web. 20 Nov. 2014.