Madison Fichtner
Future AI Philosophy
Annotated Bibliography
Writ 221 - Section 1

[1]C. Veliz, "The Challenge of Determining Whether an A.I. Is Sentient", Slate Magazine, 2016. [Online]. Available: http://www.slate.com/articles/technology/future_tense/2016/04/the_challenge_of_determining_whether_an_a_i_is_sentient.html. [Accessed: 26- Oct- 2016].

The article, "The challenge of determining whether an AI is sentient" begins by giving a situation where you may own a robot that does work around the house, and over time, this robot develops feelings of being sad and unuseful. The author proposes that a part of sentience is that it warrants moral consideration, and whether we should owe moral consideration to something that is only a thing, not a living human creature or person. The author paraphrases philosopher Thomas Nagel who considers that a thing is only a thing and nothing more. The history of determining sentience in animals, about how they were only things that acted out of mechanical reflexes, then about how they were incapable of feeling pain, and then finally about how in 2012 it was established that humans weren't the only organisms capable of experiencing subjectively. It's difficult to agree upon consciousness of something other than human. Since we do not fully understand what consciousness may mean in a thing, or even animals, it may be possible that we will create a self-learning sentient machine without even knowing it. The author relates this idea by saying that a self-learning machine is a black box, even to it's creator. Finally, should we owe sentient creatures moral consideration if we do in fact create them?

[2]S. T. H. Limited, "Artificial intelligence software," in Sentient Technologies, 2016. [Online]. Available: http://www.sentient.ai/. Accessed: Oct. 27, 2016.

Sentient Technologies has created the largest, and most powerful AI platform to date. The platform is built on deep learning algorithms to develop and scale neural network technologies, evolutionary computation that gives the AI the ability to learn, adapt, and react to different situations to provide the best solutions possible, and finally the ability to scale to millions of cores. Sentient is designed to help with ecommerce, digital marketing, finance, and research. Sentient is developing a product coined as "Sentient Aware" which enables businesses to staff an AI sales representative that deals with customers and sells products efficiently, it interacts with customers on a human scale. This is a prime example of AI being used sentiently in a business background to help humans.

[3]"Ethics of artificial intelligence," in Wikipedia, Wikimedia Foundation, 2016. [Online]. Available: https://en.wikipedia.org/wiki/Ethics_of_artificial_intelligence. Accessed: Oct. 27, 2016.

The above article on Wikipedia discusses the ethics being building artificial intelligence. Important aspects are covered well in the article including the possible weaponization of AI, the translation from AI in fiction to reality, and the unintended consequences of building a human-like computer.

The information I am focusing mostly on in this article is the unintended consequences and the threats to human dignity that an AI might pose.

[4]B. Azarian and R. Story, "A neuroscientist explains why artificially intelligent robots will never have consciousness like humans," Google+, 2016. [Online]. Available: http://www.rawstory.com/2016/03/a-neuroscientist-explains-why-artificially-intelligent-robots-will-never-have-consciousness-like-humans/. Accessed: Oct. 30, 2016.

Sentient AI will in all likelihood not happen; at least not in the manner that humans have sentience. Despite top scientists and technological experts such as Stephen Hawking, Bill Gates, and Elon Musk thinking that advances in the field of machine learning will soon yield self-aware AIs that will seek out the destruction of humans; there isn't much chance of machines becoming 100% sentient. There is a definite possibility of machines becoming extremely dangerous autonomous weapons of mass destruction, but not on their own. The article states that no matter how fast or how little memory the computer uses, it will never be able to to truly be like a human in its ability to experience subjectively simply because it is and always will be a computer that can only process information one bit at a time that is either in a state of being a 1 or a 0.

[5]L. Bell, "IBM is one step closer to mimicking the human brain," WIRED UK, 2016. [Online]. Available: http://www.wired.co.uk/article/scientists-mimicking-human-brain-computation. Accessed: Oct. 30, 2016.

Scientists at IBM claim they have made a computation breakthrough by imitating large numbers of neurons. IBM was inspired by the way the human brain transmits information to develop the first "randomly spiking" artificial neurons that can store and process data in a way similar to humans. It is hopeful that this breakthrough will be a major step forward in the creation of energy-efficient, and ultra-dense technology that will be used in applications of cognitive computing and possibly sentience. The neurons consist in two states, a "amorphous" and a "crystalline"; the amorphous is its form of not consisting of a defined structure where the crystalline form has a rigid structure, and both of them are analogue instead of digital, just at the human brain is. The importance of this breakthrough is astronomical in terms of future AI philosophy and the possible development of sentience, and i will use this breakthrough as a possible example and step towards the advancement of AI.

[6]S. Anthony, "Google AI invents its own cryptographic algorithm; no one knows how it works," Ars Technica UK, 2016. [Online]. Available: http://arstechnica.co.uk/information-technology/2016/10/google-ai-neural-network-cryptography/. Accessed: Oct. 30, 2016.

Google's artificial intelligence team, Google Brain, have recently developed two artificial intelligences that have evolved by use of machine learning to develop their own cryptographic

algorithm that protects messages they send to each other from a third artificial intelligence who was trying to evolve to crack the cryptographic algorithm. Google has two major teams working on future AI and sentience in a sense; the Google Brain team, and the Deep Mind team. What is incredible is the fact that the AI's were not told how to encrypt stuff, or even what crypto techniques they could use. The systems are not fail-proof in developing a message algorithm but the fact that they can do it at all is incredible, let alone at a rate where most of the time it works. The machines are not connected in any way other than the two communicating AI's having the same starting "key" to work with. Other than the key they are given when they are booted up, no human instruction is given to the computers to develop this neural network. This development of machine learning shows exactly how advanced we are in the current age; these computers are actually capable of communicating with each other in a cryptographic manner than the human developers of the AIs can't even figure out how they are doing it. This is one of my major examples of the current state of AI's.

[7]I. E. of Philosophy, "Internet encyclopedia of philosophy,". [Online]. Available: http://www.iep.utm.edu/art-inte/. Accessed: Nov. 2, 2016.

The above encyclopedia entry has contents starting from the origin of AI, the proofs of existence of AI, how AI looks, and different arguments about the possibility of AI. I plan to use this encyclopedia entry as a base, a place where i can find a multitude of information on the topic of AI and machine learning. The entry is very broad, and it is not narrowed on anyone aspect of AI, but I mostly plan to use the information from the appearances of AI section, and the proofs of existence of AI.

[8]N. V. Patel, "Marvin Minsky: Read his 1961 paper on artificial intelligence," Inverse, 2015. [Online]. Available: https://www.inverse.com/article/10641-marvin-minsky-read-his-1961-paper-on-artificial-intelligence. Accessed: Nov. 2, 2016.

Marvin Minsky was a genius of the late 1900's who developed some of the earliest ideas of contemporary artificial intelligence. He bridged the gap between thinking in humans and thinking in computers. He helped found MIT's AI lab which is still in existence today. Minsky, along with John McCarthy, is credited for actually coining the term "artificial intelligence," in his 1961 paper, "Steps Toward Artificial Intelligence." In this paper, Minsky outlines the five essential skills for AI; search, pattern-recognition, learning, planning, and induction. The thing about these 5 skills is that they are definitely capable of being mathematically constructed into a computer through a programming language, but the ultimate goal of AI is for the computer to actually understand. I will use this article for it's information about one of the first AI philosophers and the man who actually coined the term itself. Minsky philosophized that it is more important for us as humans to solve problems not as a set of mechanical procedures, but as a system of creativity and improvisation that can be applied to reach a desired destination, and that is important to keep in mind.

[9]M. Minsky. [Online]. Available: http://www.cs.utexas.edu/~jsinapov/teaching/cs378/readings/W2/Minsky60steps.pdf. Accessed: Nov. 2, 2016.

This is Marvin Minsky's landmark paper, "Steps Toward Artificial Intelligence." He coins the term "artificial intelligence" in this paper. The ideas included in the paper are extremely important to artificial intelligence, and programming in general. The ideas may be from 1961, but they are ever so pertinent. I want to use the information of what the problems we face as humans when it comes to trying to create artificial intelligence. Can we solve the problem tf allowing the machine to search? Can we teach a machine how to recognize complex patterns in real time? Minsky thought about these ideas in the mid 1900's, and now that we have the technology to create AI, we can actually implement and use his ideas. His ideas are still very relevant today, and I would like to compare the thoughts of the past on AI with the thoughts in the present on AI.

[10]. [Online]. Available: http://www.slate.com/articles/technology/future_tense/2016/04/the_philosophical_argument_against_artificial_intelligence_killing_us_all.html. Accessed: Nov. 2, 2016.

The topic of this article is about letting artificial intelligence evolve on its own. The article gives multiple cases of how AI might evolve into something that will kill all of humanity, but there is also the argument that until AI actually has feelings and morals of its own, it will never have true desire to do anything. The current software we have, is incapable of achieving this consciousness, for example, Google's AlphaGo recently beat the Go champion, but it was done with machine deep learning, and not through actual intent by the machine. Our machines right now have no sensation of doing anything, they just do what they are designed to do and are incapable of actually thinking about it like a human. There are many hypothetical situations mentioned in the article about the possible outcomes of artificial intelligence. The author gives many different perspectives, and I need different perspectives and arguments for my research.

Outline

a. Hopes:
    i. Service-based industries or processes
        1. Fear of job loss for humans?

b. Fears:
    i. Abuse of intelligence
    ii. Militaristic usage
    iii. Sentience that does not coincide with human beliefs or morals
    iv. Surpasses the general intelligence of humans
c. Arguments for and against the possibility of developing human-mimicking, advanced AI

2. The challenge of determining sentience
    a. The struggle of determining sentience in foreign things or beings
    b. What is classified as sentience in machines or robots
    c. Should we give sentient/advanced AI moral respect and consideration?
    d. The possibility of creating thinking AI without us even knowing
    e. The problem with the Turing test in the modern era and the actuality of sentient AI