Madison Fichtner
Writ 221
Literature Review
madisontfichtner@msn.com

Artificial intelligence is a growing trend in computer science, and with it comes the need for new information about it. We need to philosophize the possible outcomes of advanced AI that artificially mimic humans. We need to have expectations and we need to have an idea of what we are actually working to build in AI. We need to explore the possibility of creating advanced, and possibly sentient AI, the ethics of creating such AI, the possible things that the systems will be used for, such as warfare, and finally we need to decide whether it is actually helpful to us as humans to develop something artificial that has the same abilities as we do to process information in real time and experiencing subjectively. These things are what we should focus on most in artificial intelligence, and these are the things that the used sources review.

As outlined in the wikipedia resource, an important aspect of the possibility of developing intelligent or sentient machines is the ethics of creating them, and the possible threats these machines may pose. If a machine is given the capability to understand natural language such as English, there is possibility for it to be used for malicious intent to read every email, text, and listen to any phone call it pleases; with the fear of being spied on already being in the minds of people, especially citizens of the United States (NSA), the thought and fear of machines being capable of fully understanding humans is warranted. Computers are already being used to replace humans in doing specific jobs and tasks, but some argue that computers should never be used in jobs that require respect and care for others. Joseph Weizenbaum was a computer scientist who believed that there were jobs that should never be replaced by machines. He believed the most crucial of these jobs were; soldiers, judges, police officers, therapists, elderly care workers, and customer service representatives. An obvious fear of advanced AI is the use of them in warfare, but there is also a hope among many that autonomous AI will be in fact used in warfare. A fully autonomous weaponized robot could make decisions more effectively than a human, there would be no grey area of morality or decision making as there is with human soldiers. Some fear that AI weaponry will be extremely detrimental to society and humans, and possibly even cause the downfall of mankind; an outspoken believer of such a possibility is the great Stephen Hawking, who believes that computers will indeed cause mankind to stumble. Along with the weaponization of AI comes the area of fictional, and science-fiction based AI hypotheses. Movies such as The Matrix or Bicentennial Man both deal with the idea of sentient artificial life.

As discussed in the article about Google's recent advancement in AI and the mimicking of human behavior; there have been recent leaps in the development of thinking AI, and Google has developed intelligence that is capable of creating a cryptographic language, and neural network. This project by Google consists of three computers, two of these computers goal is to create a neural network between each other, while the other's goal is to "break" their code, intercept and understand their language and neural network. The outcome is generally that the two computers are capable of establishing some sort of communication between each other and transfer and analyse data, while the third computer is not as successful in accomplishing its goal. The only tools that the computers are given are given to the two computers who are trying to establish a connection, and they are only given a single encryption for establishing the same

language, but the entirety of the neural network is created without developer or coding help by the Google team. Since there are mixed results and it is not completely consistent, the trial still shows that area of machine learning and problem solving AI is growing quickly.

In the final listed resource, the author considers the importance of recognizing sentience, when it is important to be recognized, and whether it is possible for computers to be sentient, and have feelings and emotions in the same sense humans do. Should we offer “things” moral consideration? Things cannot be hurt, they have no personal interests or concerns, and they have no preferences. But if we give a “thing” the ability to feel and perceive for itself, an important aspect we need to approach is how we should go about offering moral consideration towards these things. Humans are not naturally good at recognizing sentience, we have made considerable changes in our beliefs about sentient life forms throughout our recent history (17th century - present.) Descartes, a famous philosopher claimed that animals were not sentient, but were merely biological machines that acted by reflex, and after Descartes claim, a mechanist named Malebranche claimed that animals were incapable of feeling pain, but finally in 2012, a group of scientists at Cambridge declared that humans are not the only living creatures that feel pain and experience subjectively. An argument as to why we have trouble detecting sentience in other creatures is because we as humans are so different from the things, or creatures we are trying to determine sentience in. So how can we truly determine if a machine is sentient if it is so different from a human?

Sentience is a complicated topic, and it is harder to justify sentience in a thing or creature, than it is to simply understand sentience; and the following sources help give a little insight on the possibility of advanced AI. Before we as a society try to create artificial sentience, we need to understand exactly what we are trying to classify as sentient; as in, what should we actually classify as sentience in terms of a machine or AI. We have to worry about the morality issues AI may pose. Is Google creating software that discretely makes neural networks moral? Should there be a true fear of communicating without the knowledge of developers and humans in general? Understanding the possibility of sentient computers requires the understanding of many different aspects. The articles outlined in this review are just three of the possible criterias that are involved with the future philosophy of artificial intelligence, and the possibility of total sentience.

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

[2]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.

[3]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].