Marcus Domingo

ENGH 302-N12

Literature Review

**Introduction**

Artificial Intelligence (AI) is rapidly growing throughout human society which shows promise, but could also cause issues in the future. AI has many concepts behind it that help revolutionize the world around us. A rough explanation of AI is the ability for a computer system to learn and react to its environment. Dietterich and Horvitz (2015) described the common AI systems that we encounter every day such as GPS, Siri, Cortana, Google Now, and Facebook (para 1.). These AI systems have been implemented to help us with random questions we may have or to record the events for the day. Other AI implementations include robotics, machine learning, and something Jones (2014) described as deep learning (para. 3). AI is being used in one form or another in almost every aspect of life today and although the intentions are innocent, there is a risk factor when relying on AI too much. This review will not focus on the history of artificial intelligence. The key takeaways from this review of artificial intelligence are some of the implementations of artificial intelligence, successes in artificial intelligence applications, what are the risks that are involved, and the direction AI is headed which should give better perspective on the implementation and design of AI.

**Definitions**

AI is best described as the ability for computer systems to perform human tasks. The range of abilities of a computer system is based on the design and its use. The abilities that will be focused on throughout will be decision making, machine learning, and deep learning. Machine learning is an ability that is designed with special algorithms that allow a computer system to learn without being coded word for word. The computer system learns from its actions and acts accordingly in the future for the data it has gathered. Jones (2014) introduced deep learning as being the same as the older concept of neural networks (para. 3). He describes deep learning as a method to process large amounts of data to solve difficult problems (para. 2). Deep learning is the representation of human learning through connections between neurons in a simulated brain. Deep learning can be perceived as a map of multiple machine learning systems that are connected to represent a bigger system. While only Jones outlined the meaning of deep learning, other authors described machines that implemented this concept.

**Successes and Applications**

AI has had many successes in human society, although only a few will be highlighted. AI has made significant progress in energy consumption, audio processing, and leukemia detection (Crawford and Whittaker, 2016, para. 1). Throughout scientific discovery AI has showed promise with search engines. Scientists are using these search engines to find related scientific articles along with figures, videos, blogs, data sets, and computational services (Gil, Greaves, Hendler, & Hirsh, 2014, para. 4). This is quite exceptional because before search engines would look for articles of the topic that is being searched, but now when a topic is searched other forms of related media are also presented. Cognitive systems, which process the information gathered, are becoming rapidly popular in scientific discovery. Systems like IBM’s Watson are being used to update medical literature to allow doctors to stay current with diagnoses and treatments (Gil et al., 2014, para. 6). Gil et al. stated that neurological inspired computation, as Jones (2014) described it as deep learning, shows advances in the processing of online images and video that pertain to a wide range of scientific models and layouts (para. 6).

Commonsense in AI has also had success that has affected the world. Davis and Marcus (2015) described one of the successes in commonsense as action and change (p. 4). Action and change is performed just as it sounds. The AI performs an action, based on that action it waits for an event to occur, which then follows another action that the AI performs based on the previous event. This has come to be of prevalent use throughout the construction of many AI systems. Schölkopf (2015) described a system that Mnih et al. created with the use of Q-learning, a type of machine learning based on action and change that has a reward system where the goal is to maximize the reward (para. 4,5). This system can teach itself to play 49 different vintage video games that uses the game score as the reward (para. 5). Designing a system to teach itself to play video games may seem of no use to human society, but it does open doorways to many more possibilities to come. With the use of systems based on concepts of Q-learning, we will have opportunities to solve real world problems in the future.

**Risk Factors and Actions**

Reliability and dependability is important when it comes to AI. The more AI is integrated into the world, there are more and more risks involved. Dietterich and Horvitz (2015) categorized five sets of risk factors: 1) errors in programming, 2) vulnerability to cyberattacks, 3) perception of words, 4) real-time switching between human and AI, and 5) the influences on socioeconomics and wealth (p. 39-40). These set of risks they have provided highlight the possibilities of where AI can go wrong. The first set of risks is based solely on the designer and programmer of the system. Errors in programs can lead to fatal situations. An AI system that is designed to expose patients to certain levels of radiation, with the most minor miscalculations, could lead to radiation overdose and even the death of patients. Therefore, when an AI system is produced it undergoes various tests and procedures. AI systems are also designed to try and handle situations outside the spectrum of the system. Vulnerability to cyberattacks is a constantly concerning risk. Recently a team of Chinese hackers could hack into a Tesla and control the car from a computer. This shows that although certain AI software may have been worked on for years, there are still holes and needs for improvement. The next set of risks pertains to an AI system’s perception of words. Systems may not perceive words or sayings as humans intend them to be. Humans use commonsense reasoning to understand what someone is saying but an AI system without substantial knowledge may take the saying literal. Davis and Marcus (2015) described real-world knowledge being key to commonsense and allowing an AI to pick up on sentence structure, grammar, and intentions (para. 4). Thus, allowing the AI to learn real-world knowledge would help with language perception. Switching between human and AI control shows risk factors in many aspects. Dietterich and Horvitz described airline crashes being linked to pilots not knowing the state of the plane when switching from autopilot to manual control (p. 3). This risk is dependent on the user’s understanding of the current situation. The last set of risks is the influence of AI in economy. Parkes and Wellman (2015) described these risks as AI over or under pricing merchandise because of the lack of commonsense (para. 4). They also described that multi-agent AI systems must be designed to rationally respond and not react upon actions (para. 11).

**Looking Forward**

AI is around us more than we think. Systems are integrated into our phones, social media, search engines, and even appliances. All authors have expressed that the road AI is headed down is a bright one with many options and opportunities to come. They each have described different methods and approaches to help work around the challenges and risk factors involved with AI. Ghahramani (2015) described that probabilistic machine-learning is a great method to use for parts of future systems when there is uncertainty involved. This would show use when an AI system is entering boundaries that it wasn’t programmed for so it then uses probabilistic reasoning to analyze what to do next. The review and reflection produced by Dietterich and Horvitz (2015) along with the popular piece by Crawford and Whittaker (2016) make a good outline for risks that are involved with AI and where it can go wrong. Both express the need for safety with AI and better means in which to monitor issues with AI. Davis and Marcus (2015) provide many forms of methods to teach AI commonsense but also describe that there isn’t one form that covers all aspects of commonsense. They highlight that integration is key when trying to produce a system based on commonsense (p. 103). AI has come to show great success in the world around us, but it still has a far way to go. AI is merely an aid to our everyday lives and should be kept that way because, as Crawford and Whittaker (2016) emphasized, the more we rely on AI the more the possibilities of severe consequences (para. 8).

**Conclusion**

AI is very helpful in the world around us. AI allows us to refine our searches, collect our data, or even talk to it. The research and articles that were reviewed reflect different methods, risks, and futures for the growing field of AI. Most researchers used machine learning as a crucial point to the successes of AI. Many of the solutions and methods were based on machine learning and showed great use throughout the different forms of AI. With machine learning as the basis of a system, the system can practically be designed for anything as shown in the different articles. The different risks and consequences involved with AI impact the design and implementation of systems. The deeper AI research goes, the better the solutions and methods. AI will stay in our society for a while, which allows the AI community to aim for perfection in the design of Artificial Intelligence systems.

References

Crawford, K. & Whittaker, M. (2016). Have we given artificial intelligence too much power too

soon? Retrieved from http://qz.com/787302/artificial-intelligence-holds-growing-power-over-our-everyday-lives-but-we-have-no-idea-how-well-it-works/

Davis, E., & Marcus, G. (2015). Commonsense reasoning and commonsense knowledge in

artificial intelligence. Communications of the ACM, 58(9), 92–103. https://doi.org/10.1145/2701413

Dietterich, T. G., & Horvitz, E. J. (2015). Rise of concerns about AI: reflections and directions.

Communications of the ACM, 58(10), 38–40. https://doi.org/10.1145/2770869

Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. Nature,

521(7553), 452–459.

Gil, Y., Greaves, M., Hendler, J., & Hirsh, H. (2014). Amplify scientific discovery with artificial

intelligence. Science, 346(6206), 171–172. https://doi.org/10.1126/science.1259439

Jones, N. (2014). The Learning Machines. Nature, 505(7482), 146–148.

Parkes, D. C., & Wellman, M. P. (2015). Economic reasoning and artificial intelligence. Science,

349(6245), 267–272. https://doi.org/10.1126/science.aaa8403

Schölkopf, B. (2015). Learning to see and act. Nature, 518(7540), 486–487.