

Vidar Minkovsky

Professor Stacey Suver

Technical Writing CSE300

31 March 2016

What is it learning, and should we be afraid?

From science fiction to computer science, the field of artificial intelligence has made leaps and bounds in progress, and so has the literature pertaining to it. What is artificial intelligence? Is it conscious? Is it dangerous, and on what level? When will marriage between person and robot be accepted in society, let alone legalized? These are the questions I hope the literature written on the subject will be able to answer. But before we delve into AI, let's talk about I. I am a person, a self conscious observer of mathematically defined phenomena hitting my phenomena receptors. In the case of vision, light hits my eyes and I make up a world in my head and am able to navigate it. Intelligence is something I think I have, and in the act of thinking that I have it, I prove to myself that I do, in fact, have it. Intelligence is consciousness - the ability to think is directly attached to the ability to perceive. In the present time, spring of 2016, artificial intelligence has not, as far as we know, gained self consciousness, but it certainly has gained some consciousness. Artificial intelligence is getting smarter every day and even learning how to see and act.

The article, "Artificial Intelligence: Learning To See And Act", discusses some of the underlying principles of how artificial intelligence works. It talks about how AI understands patterns from observational data and tweaks its algorithms to attain a higher reward in things like video games. Coverage of these topics in this paper only scratch the surface, but do so in an easy

to understand way. Within the topic of how artificial intelligence works in the application of video games, some of the things that need answering are: how the computer knows what to try next, is the computer able to find the best possible solution, and how can the computer learn to do things where there is no reward. This paper explains briefly how computers use Q* learning to decide what to try next, but does not go in depth as to what Q* is or how it works. The paper does not directly address whether the computer is able to find the best possible solution, but it does talk about how AIs have been able to consistently beat humans in an ever growing list of games, which at least means it's finding better solutions than we can. The paper glances over the problem of being able to learn where no reward is present, but it does not provide a solution because this is still an area of research that needs to be done.

Moving on from video games to waking reality, Savage writes about AI seeing and understanding real life visual situations. Savage's article explains how the computer sees through looking at small overlapping portions of the image and figuring out if something is in them. It checks to see how confident it is that an object is within a few pixels, then moves to the next few pixels and does the same thing. When these confidence matrices are layered together, the computer is able to determine where something is and if it is there at all. Part of the process initially in having the computers understand what's going on is having a lot of people write descriptions of images, and then having the computers analyze the similarities and find patterns between pixels and descriptions. Computers were scored against humans in a turing test of writing captions for images, and the computers did around half as well as the humans, an impressive feat. Some of the questions that need answering regarding how computers can see real life images are how they're able to do it, when they'll be able to do it as well as humans, and

what the implications are. This article explains how computer can see, and that seeing as well as humans is still a long ways away. This makes sense because an adult human, who is capable of recognizing things as well as an adult human, took 20ish years of experience to get to that point. The article goes somewhat into the implications of computers seeing, such as being able to recognize what images should go where when someone post them to a social network or bringing up the right results when someone does a google image search,

Uses of AI are discussed in “Grounding AI” from determining which houses have potential for lead poisoning, to picking schedules for students. The article discusses some of the ways in which government could benefit from AI because AI has the capability of spending much more time on a problem than humans would be willing, or have the resources, to. Some of these problems include determining which houses should be searched for lead, or which domestic abuse cases should be responded to first. The article places AI into four categories: prevention, resource prioritization, policy formation and benchmarking, but according to the other papers written on AI, there are many other more important possible uses like understanding consciousness and figuring out what the future could hold. The article also discusses the importance of the interaction between humans and computers and how the most effective use of AI is to augment our own minds and ability to think. One more potential use of AI discussed is the ability of it to determine what the best course schedule would be for a given person, but not just out of classes that the person wants to take - the AI would determine what the best path for the person in life would be and determine which courses and in what order would have them achieve that end. The article didn’t discuss how the AI would be able to do any of these things, which it would have benefitted from doing.

Moving on from the uses and basic workings of AI, there are also articles written on the potential dangers or overhype of the potential dangers. Boden's paper "Artificial Intelligence: Cannibal Or Missionary?", talks about the possible dangers including taking jobs from people, and outright annihilation of the human race. The paper talks about the non objectivity of AI programs and how they would have the ability to pervert their instructions and do potential evil. It also talks about how AI could be potentially dehumanizing since computers would be able to be just as conscious as we are, making us not so special anymore.

"Don't Worry About Superintelligence", in response to Boden, says that there is a currently unsolved problem of how we will control AI, but claims that a solution to that problem will come alongside advances in AI. The paper talks about AI becoming super intelligent and wanting to kill off the human race because, like ants to us, we would be annoying to superintelligent AI and a waste of space as well. The paper gives a partial solution to this problem by saying that we should program AI to have human friendly goals, but also shows the negative effects of that - programming the AI to make us happy or smile could lead to the AI paralyzing our faces into a smile or planting electrodes in the pleasure centers of our brain. The paper claims that, following the current trend in AI progress, the first superintelligent AI are likely to be human friendly because we have been making AI to help with our needs rather than to have its own.

There is a lot to still learn about AI, particularly in how to make it reality but also in how to use and control it once it is real. The papers cited talk about some of the possible uses, dangers, and existential disasters associated with AI, but there is still much left to research in the

field. AI could give us a completely new perspective on our own consciousness, and I think that is the most important thing that has not particularly been gone over in the current research.

Works Cited

Agar, Nicholas. "Don't Worry About Superintelligence." *Journal Of Evolution & Technology*

26.1 (2016): 73-82. *Academic Search Complete*. Web. 29 Mar. 2016.

Boden, Margaret. "Artificial Intelligence: Cannibal Or Missionary?." *AI & Society* 21.4 (2007):

651-657. *Applied Science & Technology Source*. Web. 29 Mar. 2016.

Savage, Neil. "Seeing More Clearly." *Communications Of The ACM* 59.1 (2016): 20-22.

Business Source Complete. Web. 29 Mar. 2016.

Schölkopf, Bernhard. "Artificial Intelligence: Learning To See And Act." *Nature* 518.7540

(2015): 486-487. *MEDLINE with Full Text*. Web. 29 Mar. 2016.

Wood, Colin. "Grounding Ai." *Government Technology* 29.1 (2016): 16-21. *Applied Science &*

Technology Source. Web. 29 Mar. 2016.