Philip Dumaresq

Values in the workplace

Leslie Elliot

March 17, 2017

Artificial intelligence

1. When talking about artificial intelligence, we’re really talking about different parts of AI. Sometimes it’s ANI, artificial narrow intelligence. These are systems like the AI in video games – Their sole purpose is to try and kill you without getting killed. This mean repeating basic tasks over and over, like crouching and standing to fire or reload in a shooter game. This is the kind of AI that exists today, AI that can only perform one certain task. When we think of the future however, we think of robots that are able to think for themselves and who are as intelligent, or more intelligent than we are. We think of AI ranging from home assistance systems, like Jarvis in the Iron Man and Avenger’s movies to the supercomputer Deep Thought in hitchhiker’s guide to the galaxy. AI is certainly heading in the direction that these sci-fi movies think they’ll end up in the future. Google’s DeepMind team is working to build AI that can teach itself based on observing the result of the same task repeated over and over. Eventually, this will be able to teach itself anything and learn anything in a matter of minutes or hours.

To get to the point of having this super intelligence (ASI), there’s several different paths that researchers think that we can take. One way to get there is through plagiarizing our own brains. We could reverse engineer it and build an artificial neural network out of it to create AI that’s as smart as us. Another possibility that’s been looked into is to emulate evolution inside the computer. We program algorithms that determine which way of completing tasks is the best, based on the input of any given task and get computers to merge their software together based on the highest performing amongst them. Terminate the less successful AI and merge the capabilities of the most successful until we have one AI that’s capable of doing everything the weak ones can. While this is a good possibility, it’s always worse to take a bunch of odd pieces and chunk them together, so the last, and possibly the best, is to have computers figure it out. We could build an ANI computer that’s sole purpose is to research AI and learn how to code. Then as it discovers different things about AI, it could code changes into itself and constantly self improve. It would be the fastest way to do it, since an AI could probably learn everything that we know about AI incredibly quickly, but this would require years of work on our end before we could ever build and AI and just set it off to do it on it’s own.

1. One of the big issues with developing AI is that a lot of people believe that it could spell the end of humanity. ASI poses a big threat, and our fear of it is often shown in different movies, like Terminator, The Matrix or Tron. These movies all have these ASI super villains that attempt to destroy humanity, and AI could head that way if we aren’t careful. Multiple people in the scientific community, such as Stephen Hawking and Elon Musk have warned that we have to be incredibly cautious with AI. Building something that learns aggression, yet has more knowledge than every human on earth who’s ever lived could easily prove to be an existential threat to humanity. It could wipe out the entire world as we know it, or drive us to extinction. This is the biggest worry among AI community. While we’re still probably a couple decades down the road from getting to a point where we can develop and AI that powerful, we have to be careful now and learn to build AI responsibly and determine what we should and shouldn’t do.
2. I’m cautiously optimistic about the future of AI. I believe that getting the point of having ASI could do amazing things for the human race. We could develop cures to every disease known to man, we could invent energy sources that are 100% green, we could discover life on other planets and truly travel space. As humans, we strive to progress, and we will do that at any price, which is where the problem comes in. People can make mistakes, much more than machines do, and since we’re the ones creating them, they too will be flawed. And if computers get to the point where they’re re-writing their own code, they will be flawed beyond our comprehension. Disasters that have occurred by problems with ANI, like failed power grids or nuclear plant malfunction could become things that are far more catastrophic, beyond what we could even image today. I don’t really think that there’s much we can do to mitigate the risk of something happening, because one way or another, we’re going to get to the point where ASI exists, and there will be problems we’ll have to work out. The only thing we can really do is build one, see what the outcome is, and fix it built off what we see as a result. I don’t think that we’re going to do something like build some evil overlord capable of enslaving the human race, but if we develop colonies on other planets and AI fail for some reason, it could kill thousands or millions of people. AI developed to monitor threats of volcanic eruptions could do the same thing. Threats on the scale are very real, and probably very close. So while I’m optimistic that AI will drastically better our lives in the future, I do believe we have to be very cautious and only implement things that will work and take no shortcuts in research and development, which is something that people have a tendency to do.