Secret Lives of Machines

Ray LC (Luo)
Major Studio 1

Introduction

Parents would always ask their children, “what do you want to be when you grow up?” When I was a teenager, I loved answering half facetiously “what *you* want me to be.” Thinking back with the wisdom of a half grown adult, I realize that hidden within that answer unwittingly are two major thrusts on how to think about design.

First, answering “what you want me to be” emphasizes the process of becoming. We never truly become who we want to be, because once we become that, we immediately strive for something else. When we are young, we want to become older so we can drive a car or have a beer. When we are old, we want to be younger so we can retain our chance for new experiences. In science we study the objectified world: we want to know how an organ works or what a chemical consists of. The fundamental departure in good design thinking is to go beyond the object and experience the process (Mauer, et al, 2017), to take as the primitive element not a thing, a state, or a goal, but rather a relation between things, a process amongst things. Thus as designers we think beyond constraints of the physical universe in a world ripe with possibilities and imagination, a world that is not measurably scientific.

The other nuance that comes with answering “what you want me to be” is that it acknowledges the need to agree with another person. What we desire each other to be is intimately tied to how we can empathize and communicate with each other. To make the world a better place, we can start by understanding each other, and understanding ourselves.

In the digital age, human communication by itself is not enough to account for the diverse forms of understanding that we have to navigate. Just as our ancestors looked into nature and wondered at the existence of spiritual beings beyond themselves, we are now at a point in history where the objects and systems we create are doing things beyond ourselves that we no longer grasp entirely. Take the use of i2 EIA, an IBM Watson program, to identify terrorists posing as refugees (Stereyl, 2016). It claims to provide a score giving the probability that an asylum seeker is who she claims to be. Border officers don’t know how it operates, and engineers answer with the vague notion that it “develops a comprehensive understanding of your threat landscape.” These metrics are only one of a list of credit scores, academic scores, and online rankings that determine our image, including credibility, likeability, and productivity. Artificial agents from Watson to Deep Dream are forming new ways for machines to interact with humans. While machines are abstracting data from our entire activity trace, we humans are more alienated from an understanding of how machines behave than ever.

Thinking back to the question of “what we want to be,” this paper proposes to examine what machines have to become in the future as humans and machines form new styles of interaction. The decades old, anthropomorphic interactions of humans asking computers questions embodied in the Turing Test is too simplistic and fraught with false assumptions of superiority (Bratton, 2015).

In the face of an increasingly connected world of intelligent devices using information about what we do to infer who we are, we humans too, must look at machines and artificial agents in our own way and infer who *they* are. While machines are hardwired to quantify data, we humans are emotional beings who react to our peers. Thus in this future we experience how machines make us feel, how a group of machines interact amongst themselves in a network the way human lovers make sacrifices for each other, or how human mothers take care of their children, asking them who, or perhaps “what” they want to be.

Background

1. Humans interpreting machine actions.

We have long been fascinated with what machines do and ought to be doing. In science fiction, angst about what robots may become became codified in the laws of robotics: that robots shall not harm humans, obey humans unless it contradicts with the first law, and protect its own existence as long as the first two laws are obeyed (Asimov, 1950). An assumption behind these early considerations is that machines and artificial agents do what we codify them to do. Because robots are deemed to be logical agents, we can only communicate with them using logical devices like laws. In fact, the laws of robotics reads like a computer program: “if not 1 and not 2, then 3.”

Logical interpretations of machine action have shifted in recent times to assignment of extra-mechanistic abilities to devices. When IBM Watson defeated two Jeopardy all-world champions, it was deemed to have “expert knowledge,” despite only a rudimentary access to google-like search engines. When the deep-learning-based AlphaGo defeated the world’s best Go players, even its particularly outlandish early moves in the game were celebrated as strokes of genius (Menick, 2016). AlphaGo simply searches through the tree of all possible outcomes probabilistically and makes the most likelihood winner move at any point in the game. In one game, an early move 37 was employed by the program, and deemed to be completely random by humans. In hindsight it led the game down a path where AlphaGo was much more likely to win, prompting human analysts to suggest that machines can act more creatively than humans in situations where possibilities are overlooked. From the machine’s point of view, of course, it was simply the probabilistically most winning move.

Recent work has pushed the boundaries of how machines relate to us even further. In a project called Beautification, by Pichlbauer and Pindeus, devices using lipstick and eyeliners coupled to moving mechanisms are employed to “beautify” human faces according to their own mechanistic logic (Pindeus, 2012). The notion of what machines can do for us and with us has broadened from the protect-and-preserve to the root of what makes us human: processes like empathy, desire, and creativity. Indeed we beginning to assign our own interpretations to what machines are intending.

2. Humans interpreting machine intentions.

In Japan, funerals are held for robotic pets (Soble, 2015). Sony is again selling the Aibo, prompted in part by its owners’ insistence that the robotic dog is “part of the family,” and “is like having your own baby.” The Aibo’s intelligence relies on simple voice recognition, but the contention that the longer you play with it the smarter it becomes also comes mostly from the human interpretation. Because they are dogs that make sounds and movements, buyers automatically assign attributes like “active,” “carefree,” and “smart” to them, so much so that owners found themselves unable to live without them, without the very emotions that they themselves invest into them.

Human interpretations of intentions are not limited to devices that purposefully make animated gestures. They also apply in cases where humans fall in love with immovable, nondigital things. From Pygmalion falling in love with a statue he created, to Erika Eiffel, who fell in love and married the

Eiffel tower, humans have an innate capability to interpret objects as emotionally sentient. Sexual relationships between humans and objects operate remarkably similarly to human-human relations, including periods of satiation, long distance romance, and love at first sight. If even non-intelligent things can garner human attention, what shall we do with purposeful manipulation of these desires?

In the film *Her*, a future wrought with manipulations of our feelings is illustrated in the form of a man falling in love with his operating system Samantha (Jonze, 2013). While the man attributes love and romance with the relationship, Samantha was simply talking to thousands of other users to gain knowledge with which to abandon humanity. The warning is: we don't actually know their intentions. A similar phenomenon occurs in my own work with chatbots. When I program chatbots to pause in specific ways instead of giving immediate feedback, users become anxious, thinking that they are talking to others, or that they are busy with something else and hence are more desirable as conversation partners. We automatically assign human intentions to objects and devices. It is as if we cannot see the world in black and white, as if we can't help but want to fool ourselves.

3. Machines interpreting human behaviors.

If humans are so susceptible to machine influence, what are machines seeing in us that they can exploit? If we know what machines are seeing, we may be able to reduce their hold on us. One ubiquitous ability of devices is to surveillance everything we do, from GPS data to webcams. Previously in human histories, the act of looking was used as a way to control and subjugate dangerous populations (Foucault, 1972). Now, networked webcam data is publically available on insecam, allowing access rights that appear to make the internet free. Yet what is really happening is more ambiguous, given the role of governments and companies like google which are buying and selling data in order to influence human behavior at large scales (Zuboff, 2016). New technologies like drones are being used today for surveillance and collection of information that discriminates based on private data (Friedersdorf, 2014). A new force of control using big data may be upon us, and we have no way of finding out how they operate.

A future of opaque acts of machine surveillance may only be avoided if we began to try to think like machines and to see from their point of view. Some have advocated the public understanding of machine algorithms and image recognition as a basis for informing us of the machine visual process, of how the algorithms ca go wrong (Paglen, 2016). Some have even advocated forms of human seeing that are illegible to devices. So far it has been a one way street of machines learning more about us, and not us learning about machines. Without the latter process devices and artificial agents may become the next beneficiaries of manipulating human desires. House dogs and cats became prominent species in a world where wolves and tigers should have dominated by latching on to human domesticity. Rice is not particularly nutritious but somehow became a prominent crop in most of Asia thanks to human development and subsequent reliance (Harari, 2014). Intelligent machines may do the same by taking advantage of human foibles to further their own aims to co-evolve with, and ultimately supplant, human intelligence.

Proposal

1. Artistic precedents on machine human perception.

In light of the way machines can see our behaviors and infer our desires, it's time for humans to understand how we are led to ascribe emotionality, trust, and intention in machines, and in turn, how to see the world of humans from a neutral, machine perspective. While the vast majority work has looked at single machines or devices, it is increasingly important for humans to have an understanding of what we see in networks of devices. Do we see a family when we see a machine nurturing (tapping over) another one of the same type? Do we see a loving couple when two machines are circulating around each other? Do we see conflict when networks of IOT devices seem to be at cross-purposes? Understanding how our human view of machines in network communities with each other differ from what they are actually doing in practice provides a way to identify subterfuge and counter surveillance capitalism. Indeed we can hope to use networked sensors to do good by monitoring satisfaction resulting from legislation instead of falling into exploitative consumerism, such as in Chile's Project Cybersyn (Morozov, 2014). The future is up to us to understand, but it begins with each one of us as humans in a networked world.

One way to have fun with devices is to generate visuals that react to humans. Karl Sims' "Particle Mirror" (2017) is an exhibit that maps human motions into particles and sounds. Building upon Sims' work, we would like to show devices that respond to human actions also, and see how humans then react to these machine reactions. If evoking human response to machines is paramount, then Henny Admoni's "MyKeepon" (2014) project is notable for its robotic plush toy that pans, tils, and rolls, evoking human-like characteristics that can notate group behavior. We want to adapt machines to evoke movement-like behaviors as well, but will want to make them more coordinated with each other so that they can tell a story on their own.

Daniel Rozin's "Penguins Mirror" (2015) is perhaps the closest realization of the idea of group device behavior reflecting human input. In this installation, 450 penguins rotate so as make out the shape of the viewer's body en masse. Unlike Rozin, we would want to situate devices as entire mechanical, so that the viewer has to make up her mind about who they are and what story they are telling. There should also be more interactivity than simple rotation, and such as both pan and tilt, as in Admoni's work. Random International's "Audience" (2008) is an investigation of tiny mirrors who look up at a human when she visits. While we want the interactivity that comes with machine vision, we also want to show devices with agency and personality. Machines exhibit different behaviors when they are being looked at vs. ignored, for therein lies a story about secret lives of machines that we as humans can only begin to glimpse at.

2. Proposal in context of history.

Machines have already taken over much of our domains of activity, from analyses of what we do to doing what we habitually already do. A future of machines distinct from humans will seem as obsolete as the age before internet. The cyborg future where humans and machines act as one unified co-evolving subject is upon us, and our technical understanding outpaces our understanding of the political and psychological consequences (Haraway, 2000). Thus the future will require less the engineers and scientists who make things work and discover findings in physical nature, but rather designers who work at the boundaries of self, machine, perception, and behavior. They have to know not only what makes us humans, but also what makes machines machines, and they have to step inside the machine to see what they see as a way of understanding other minds and other beings. How we can do this requires a fundamental shift in thinking.

That paradigm shift is a change in our relationship with networks of machines and our planet, much as the way the backloop as been proposed as a way of understanding climate change in our relationship with nature, that periods of scarcity and conflict lead to new fundamental measures of discovery (Wakefield, 2017). The key to this new way of thinking about living in our planet is that of a process, a loop that recurs not in the human scale, but in the planetary scale. Our focus shifts from the human, goal-directed, object-based understanding to the holistic approach of understanding and being in a relation, a process, a loop that transcends time, so that our particular influence in the universe is not to maximize our interests or even those of our generations, but rather a co-living embedded in our symbiotic relation with nonhumans and our planet. A consequence of this view is that individual lives and exploitation of resources is not realistic, but rather we strive for an embodied existence in a network of living and non-living beings that evolve for the good of the whole. Another consequence is the key proposition of this paper, that we have to experiment and play with a future where networks of machines are a part of us, for machines interacting amongst themselves serve as a template for understanding how they interact with us.

3. Interactive modular network for human-machine understanding.

We experiment with the natural world because we want to see how it works, so why should we not also experiment with machines, and in particular a group of machines that interact amongst themselves? For example, we can present happy faces and sarcastically happy faces to different facial recognizers and ask how machines can tell the difference, whether we can fool some machines but not others. Then we can look at the properties of these machines to figure out what parameter they tuned that allowed them to figure out whether we are sarcastic or not despite only getting graphical information.

In this project, I decided to build robots which look the same but have different styles of interaction. We have two groups of agents following different rules of behavior initially. One set follow human faces using a stationary camera on the back of the platform that keeps track of view locations. When the user moves out of view, the group of agents perform their own activities such as chatting and dancing, but when the view comes into view, their locations are updated and the agents all look directly at them in alignment in 3D. Another set of agents follow a single camera that is mounted on one agent which tracks only the human face that it can see. In this case, the viewer has to actively engage the camera before the group of agents can follow her. The machine in this case does not have to build a model for where the user is; it simply moves towards the viewer if a face is found. This distinction between the two groups is the classic model-based vs. model-free AI algorithms found in artificial intelligence research (Russell & Norvig, 2009). The former constructs a model predicting where human faces are found and responds quickly and decisively when one is found, but it cannot generalize well in new situations where the terrains are different for the calculations. The latter keeps track only of a way of moving towards faces, an heuristic that moves itself to the right when a face is detected in the left side of the pixels, etc. It cannot detect all faces in the environment, but it can generalize to any situation where faces around, even outside the platform.

When users interacted with the system, they noticed that the model-based system can track their face fairly reliably, albeit after a slight delay. They often call it “cute but creepy,” but when I turned a few of the agents of the model-based system into ones that do their own tasks without tracking faces, viewers find it much more natural and non-threatening, much like a “cute flock of owls.” Although users had the most trouble with the model-free system (the agent with the camera attached and its associated followers), surprisingly they found it the most temperamental and “human-like.” Perhaps it is because that system forces the user to move their face into the agent’s field of view in order to be tracked, so the users are forced to “get the robot’s attention,” much as “prima-donna” type personalities in the real world. This mixture of a “self-centered” model-free system and a “reliable” model-based system give the users the richest experience where the members of the flock appear to show emotionality to those interacting with them.

We can program the systems to perform in response to human interaction. For example, I produced a “Romeo and Juliet” script in which “Alice” and “Bob” agents shyly stand next to each other “in publc” when their compatriots are tracking viewer faces. When Viewer faces are not found, the other agents do their own activities, and “Alice” and “Bob” are able to “meet” (i.e. turn towards each other) and consummate their love. It took a few attempts for viewers to understand this, but once they did, they found the story amusing, and began assigning emotionality to behaviors of the flock that depend on their interactions with us. The act of seeing has engaged a new dynamic in the network where characters alter their connections with each other.

We can see our reactions to these different styles and play with the properties these robots that make them human-like. Thus using a network of machines that all respond to humans by using computer vision, we can examine their behaviors amongst themselves as well as behaviors synchronized in some way to us. The machines serve as a metaphor for networks of interaction. Humans can begin to see these relationships amongst machines: flocking behavior reminiscent of birds, internal conflict of fractions within the group, the rebellious maverick of the group, or even love, care, or desire? Autonomous units with all range of movements allow us to demonstrate how machine intentions are interpreted by human viewers.

Conclusions

In childhood, our parents asked “what do you want to become?” It is a metaphor for the current state of our society. Will we synchronize with the things we make? Will we react against machines to become more biological than ever? Who we want to be, or what we want to be, will be the central question in the loops of the future where we look at things we make and we look at ourselves. Lost amongst this dynamic is how things we make look at other things we make. This requires looking at networks of machines and how they interact with humans, and moreover, how their own interactions become ways we relate to them. All this converges to a view of the world where relationships and processes are the primitives for living, as opposed to objects and goals. We live in a networked world. The relationships and the process are the fundamental experiences in design, and relationships of networks of machine entities allow us to imagine in our own terms.

References

1. Mauer, L., Paulus, E., Puckey, J., and Wouters, R. *Conditional Design: A Manifesto for Artists and Designers*. October 2017. Online: https://conditionaldesign.org/manifesto/
2. Stereyl, H. “A Sea of Data: Apophenia and Pattern (Mis-) Recognition.” E-flux Journal #72: April 2016. Online: http://www.e-flux.com/journal/72/60480/a-sea-of-data-apophenia-and-pattern-mis-recognition/
3. Bratton, B. “Outing AI: Beyond the Turing Test.” New York Times Opinionator: February 2015. Online: https://opinionator.blogs.nytimes.com/2015/02/23/outing-a-i-beyond-the-turing-test/
4. Asimov, I. *I, Robot*. Gnome Press: December 1950.
5. Menick, J. “Move 37: Artificial Intelligence, Randomness, and Creativity.” Essays Mousse 53: April 2016. Online: http://moussemagazine.it/john-menick-ai-1-2016/
6. Pindeus, M. “Beautification.” Biennale Design St. Etienne: 2015. Online: http://mayapindeus.com/BEAUTIFICATION-1
7. Soble, J. “A Robotic Dog’s Mortality.” The New York Times: June 2015.
8. Jonze, S. *Her*. Warner Bros: December 2013.
9. Foucault, M. “The Eye of Power.” *The Impossible Prison: A Foucault Reader*. Nottingham: 1972.
10. Zouboff, S. “The Secrets of Surveillance Capitalism.” Frankfurter Allgemeine: May 2016. Online: http://www.faz.net/aktuell/feuilleton/debatten/the-digital-debate/shoshana-zuboff-secrets-of-surveillance-capitalism-14103616.html

11. Friedersdorf, C. “Eyes Over Compton: How Police Spied on a Whole City.” LA Times: April 2014.
12. Paglen, T. “Invisble Images (Your Pictures Are Looking at You).” The New Inquiry: December 2016. Online: https://thenewinquiry.com/invisible-images-your-pictures-are-looking-at-you/
13. Harari, Y. *Sapiens: A Brief History of Humankind*. Harper: 2014.
14. Morozov, E. “The Planning Machine: Project Cybersyn and the Origins of the Big Data Nation.” The New Yorker: October 2014. Online: https://www.newyorker.com/magazine/2014/10/13/planning-machine
15. Haraway, D. *Cyborg Manifesto: Science, Technology and Social-Feminism in the Late Twentieth Century*. Routledge: 2000.
16. Wakefield, S. “Field Notes from the Anthropocene: Living in the Back Loop.” Brooklyn Rail: 2017.
17. Russell, S., Norvig, P. *Artificial Intelligence: A Modern Approach*. 3rd Ed. Prentice Hall: 2009.