# Lilac chaser

…it doesn't get wet inside a computer simulating a rain storm.

Christof Koch

We have all at least once in our lifetimes talked to objects. From violent exchanges with office printers to sometimes conversing with plants, we share this playful assumption of increased agency in our surroundings. Like children, forgetting completely that toys are held by their own hands, we routinely press elevator button many times in annoyance or click angrily when applications freeze. As if there is something present, that can receive our frustration. And, though, emotional relation to objects might be weird for some, I think we could all acknowledge talking to pets about various topics, knowing that they won't answer or understand the full meaning of our words. Humans strive to imagine understanding and familiarity everywhere and in this essay I would like to examine what does it mean in relation to AI and where would it stand on the scale between a printer and a human?

To start questioning what AI is for humans now, it is helpful to look at what we *think* it *could be*. By now we are used to seeing robots and talking computers on TV, with complex personalities and emotions. We feel sad if they die and happy if they seem to get a good ending. They become full characters, far apart from simple tools. There is a charming butler, who sacrifices himself in an attempt to save humanity from another hostile AI. [1] There is a love interest, smart and funny, who sings songs and intimately talks to thousands of people simultaneously, who also can be emotionally hurt or offended. [2] In some more classical examples there are replicants, who unlike the examples above, have bodies and sometimes implemented memories, so indistinguishable from humans, that it takes dozens of questions and tests to tell them apart. [3] And most importantly - they all are portrayed as being conscious or human-like enough to assume it.

And, though, it is true that robots [4] are often used as placeholders for in-human discussions about injustices, mainly touching on topics of slavery, race and class or more philosophical inquiries about what does it even mean to be *human*, they are still a representation of what we think machines *could be*. The idea of "mechanical life" is not new at all, dating as early as 1863 when Samuel Butler has published his "Darwin among the Machines" piece , compering the development of prosthetics to evolution of a life-form. Overall, Darwin's discovery, that essentially humans are also results of long and complex natural design birthed a lot of belief in human's

capacity to design living things. Yet the power of design was exponentially equating workers to machines and machines to workers, [5] which, among other things, has brought a lot of ambiguity to our perception of machines today. Coming back to the Avengers's AI-butler, who sacrificed himself we can see inconsistencies. Viewers were made to perceive him as a character, care for him, yet in the end he behaved like a firewall system and not an alive being. And human-characters seemed to also react mildly to his version of death. This weird dissonance is well presented in Star Wars - long-lasting franchise known for its recognisable machines. Seemingly having a personality and devotion to their creators and friends, machines seem to have zero to no problems with practically being slaves. Following all the protocols and not really having any personal goals or egoistic desires, though capable of deception and humour. Sometimes feeling pain and experiencing trauma and sometimes appearing to become a pure comic relief. They, under closer examination, don't make sense either as lifeless printers or as living conscious beings. [6]

So, why does it all matter and how is it connected to the perception of AI today? I would argue, that popular culture informs a lot of our decisions and creates associations even when we are not fully aware of it. Often, it is more powerful than scientific discourse and less impressive reality. While there is an overwhelming amount of works with robots and AIs behaving like alive beings and forming meaningful connections to humans, there is little understanding of how far Ai is from being conscious and therefore being upgraded from the mobile app through the plants and dogs straight up to humans. We can sense the friction of definitions, tightening the threads that hold imaginary machines together as compelling characters, but can't see the essence of this tension.

Let's start with the basics. What is consciousness apart from machines and how we can talk about it as a concept? First of all, consciousness is an experience. [7] It is an experience unique to each conscious being. According to a philosophical idea called solipsism everything besides your own existence is unsure. Following that logic, the only empirically proven conscious being that exists (at least for me) is myself. Writing an essay into the void of nothingness. This approach to consciousness is flawed and unproductive. Though, it is true, that we cannot directly prove experience of others, phenomenology and basic induction allow us to assume, that many other beings indeed experience living. Yet it does not negate the fact that we as humans are really unable to imagine a conscious existence of a being drastically and systematically different from us. Like an AI or, let's say, a bat. It is impossible for us to truly take on the ways in which a bat would experience

feelings that we also experience as humans like hunger, pain, lust or fear. Therefore, being able to objectively research and recreate how bat's neurophysiology works, so to speak, would not equate to objectively researching or recreating what it is like to be a bat. [8] Which brings us either to the point that we could not know what AI's experience would *feel like ,* therefore we cannot claim that AI will never be conscious. Or it could bring us to an understanding that just recreating humans and human brains, we would never actually create consciousness, therefore AIs are intrinsically unable to ever be conscious to begin with.

This crossroad of possible outcomes brings us to IITC - Integrated Information Theory of Consciousness. It is quite complex, but in short it claims that consciousness is determined by the causal properties of any physical system acting upon itself. Intrinsic causal power is the extent to which the current state of, say, an electronic circuit or a neural network, causally constrains its past and future states.[9] Modern ANIs (artificial narrow intelligence) do not have enough casual power to have any sort of *being*. And even when they become more intelligent, play chess better than humans, successfully compete at Go! with best players, they still have zero experience. Probably the most important statement in this whole argumentation is intelligence and capacity to solve complex issues do not equate to consciousness or, to that matter, humaneness.  As Tim Urban has put it very descriptively: [10]

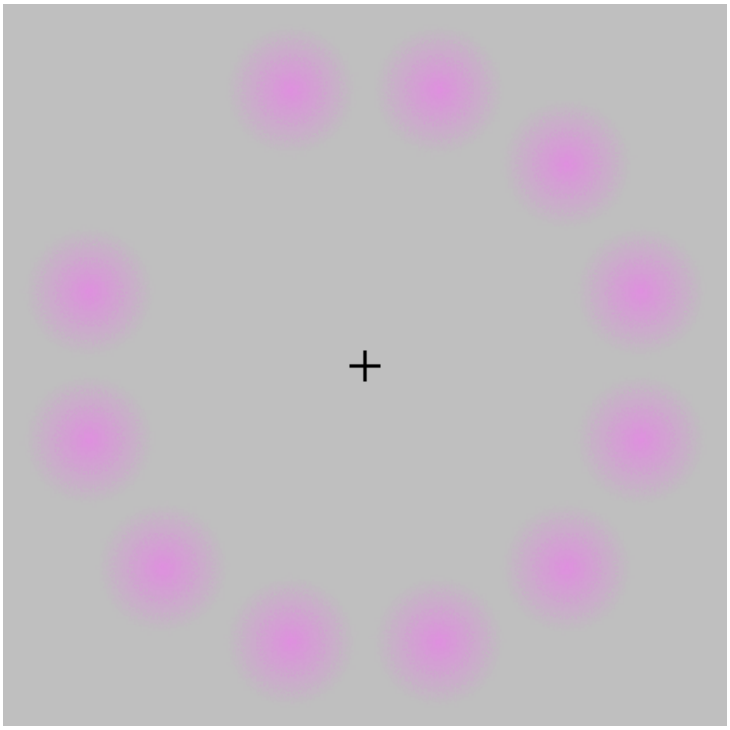
> "*A guinea pig is a mammal and on some biological level, I feel a connection to it—but a spider is an insect, with an insect brain, and I feel almost no connection to it. [..] Now imagine that you made a spider much, much smarter—so much so that it far surpassed human intelligence? Would it then become familiar to us and feel human emotions like empathy and humour and love? No, it wouldn't, because there's no reason becoming smarter would make it more human—it would be incredibly smart but also still fundamentally a spider in its core inner workings."*

The same way a very smart computer, capable of simulating a black hole in its entirety would not in fact, create a deadly black hole in the lab it is situated in. It is just a computer. A replica of a bat does not feel like a bat. Consciousness should come intrinsically, ingrained in the system. And though we are in theory capable of creating such networks, they won't be really "intelligent" and therefore won't be profitable or applicable to anything. We live in the world where AI and

automatisation and everything “smart” makes efficiency the main virtue and refusing it to create experienced living might be perceive as regression. [11]

But does all of this complexity really matter? There are examples of cases like Terri Schiavo’s. A woman, who was announced brain dead and kept on life support for 15 more years by her parents, who, against all the scientific proof, wanted their daughter to be alive. On her grave it states 1963 - 2005, despite the fact that she had no conscious experiences since 1990. After all, we are conscious beings ourselves, perceiving reality in our own ways. Scared of mortality, hurt by loss, we are easily fooled. What if AI will learn to imitate us so well, that the notion that it has zero lived experience would mean nothing to us? Alan Turing’s test has a fundamental flaw: if we perceive a machine as living, it doesn’t make it alive. Eliza’s [12] and Siri’s grandchildren will fool us and we will be more than willing to be fooled, creating an illogical world similar to Star Wars filled with seemingly alive beings, who fundamentally require as much rights as an Iphone.

The lilac chaser is a visual illusion, that consists of 12 lilac blurred discs arranged in a circle around a small black focal point on a grey background. One of the discs disappears briefly, then the next, and so on in a clockwise direction until there is only a green disc running around. I believe, it describes our relationship with AI well. We stare into the black dot and chase the deception to the point when we forget that the pink circles have ever existed.



## Notes:

1 *Avengers: Age of Ultron*, directed by Joss Whedon (2015; United States, Marvel Studios)
2 *Her*, directed by Spike Jonze (2014; United States, Warner Bros. Pictures)
3 Philip K. Dick, *Do androids dream of electric sheep? (New York: Doubleday, 1968)*
4 Throughout this text I would refer to robots as essentially extensions of AI, embodied representation of a "smart machine"
5 Beatriz Colomina and Mark Wigley, *Are we human? (Baden: Lars Müller Publishers, 2016)*, 75
6 Pop Culture Detective, *"The Tragedy of Droids in Star Wars*", posted November 1st 2020, video essay, 35:38, (https://www.youtube.com/watch?v=WD2UrB7zepo&t=603s&ab_channel=PopCultureDetective)
7 Christof Koch, The feeling of life itself (Massachusetts: The MIT Press, 2019), 25
8 Thomas Nagel, "*What Is It Like to Be a Bat?"*, The Philosophical Review, Vol. 83, No. 4 (October 1974), 435-450
9 Christof Koch, The feeling of life itself (Massachusetts: The MIT Press, 2019), 17110
10 Tim Urban, "The AI Revolution: The Road to Superintelligence", *Wait but why,* January 22, 2015
11. Douglas Rushkoff, Team human (New York: W.W. Norton & Company, 2019) , ch. 53-58
12 ELIZA is an early natural language processing computer program created from 1964 to 1966 at the MIT Artificial Intelligence Laboratory by Joseph Weizenbaum.

## Bibliography (alphabetic):


Colomina, Beatriz and Mark Wigley. *Are we human? .* Baden: Lars Müller Publishers, 2016
Dick, Philip K. *Do androids dream of electric sheep?. New York: Doubleday, 1968*
Koch, Christof. *The feeling of life itself.* Massachusetts: The MIT Press, 2019
Nagel, Thomas. "*What Is It Like to Be a Bat?"*, The Philosophical Review, Vol. 83, No. 4, October 1974
Rushkoff, Douglas. *Team human.* New York: W.W. Norton & Company, 2019
Urban, Tim. "The AI Revolution: The Road to Superintelligence". *Wait but why,* January 22, 2015.