The Total Turing Test. Robotics from Japanese and European perspectives

Jos de Mul (Erasmus School of Philosophy. Erasmus University Rotterdam)

It's a great honour and pleasure to be here again at the Graduate School of Sociology of Ritsumeikan University. I have wonderful memories of my 2016 stay in Kyoto as a guest professor.[1] It was a privilege to work with my Japanese colleagues – especially Yuko Nakama, with whom I have collaborated in the past decade in different projects about landscape and space – and to discuss with the students who attended my course at Ritsumeikan, on the relevance of Greek tragedy for understanding the human condition in our present, high-technological world. During my stay in 2016 I also spent quite some time researching android robotics in the Kansai region. Especially interesting were the visits to the Hiroshi Ishiguro Laboratories in the Advanced Telecommunications Research Institute International (ATR) in Kansai Science City.

Although the subjects mentioned – landscape, tragedy, and robotics – seem to be quite diverse, my research in these fields share a comparative approach, bringing in dialogue Eastern and Western, more particularly Japanese and European perspectives in these three domains. In each of these domains we find striking similarities as well as fundamental differences. In my lecture today I hope to demonstrate this, taking the so-called Turing Test as starting point for a reflection on the similarities and differences of Asian and Western perspectives on and attitudes towards robotics.

In the first part I will analyze three recent Western science fiction movies in which the Turing test plays a prominent role. Although all three movies are fiction films, they reveal some important characteristics of the Western view on robotics. In the second part I will contrast the Western approach with the way the Turing Test is approached in Japanese robotics, more particular in social android robotics. Hiroshi Ishigiro's ERICA (ERato Intelligent Conversational Android) will be my main example. In the third and final part I will I will argue that, in the final analysis, the difference in approaches in Western and Eastern robotics is closely connected with different religious worldviews, which even in a secularized world still inform robotics and AI research at a fundamental level.

Western fear about robots

Now that robots no longer can only be found in factories, but have started to enter the social world (e.g. as care robots) and are being used as military weapons (e.g flying robots known as drones), the latent fear for robots in the Western world increasingly invades the headlines of the newspapers and magazines. In 2016, University of Oxford associate professor Michael Osborne predicted that robots will pick up 50 percent of current jobs within the next twenty years.[2] Scientists warn about the development of fully 'autonomous drones', which select and destroy targets without interference by humans.[3] And in January 2015 dozens of leading experts – including the famous physicist Stephen Hawking, entrepreneur Elon Musk, co-founder of DeepMind Demis Hassabis,, Director of Future of Humanity Institute Nick Bostrom, Google's director of research Peter Norvig, and Harvard

professor of computer science David Parkes - have signed an open letter calling for researchers to take care to avoid potential "pitfalls" of Artificial Intelligence.[4] In a BBC interview Stephan Hawking even warned that superior artificial intelligences could end mankind.[5]

Whoever believed that these developments will take their time, was startled by the news that in 2014, for the first time in history, an artificial intelligence passed the Turing test.[6] This test should determine whether an artificial intelligence can be distinguished from natural (human) intelligence. The test was designed in 1950 by Alan Turing (1912-1954), the brilliant mathematician who invented the programmable computer.[7] He was also one of the pioneers in the domain of Artificial Intelligence (AI)[8], and the one who cracked the secret Enigma code of the Nazis.[9]

In 2014, the last mentioned achievement was the subject of the movie *The Imitation Game* (2014). However, this movie was not the only recent movie in which Turing's "imitation game", also known as the Turing Test, appears. The test plays a major role in three science fiction films that were released around the same time: *Her* (2013), *Ex Machina* (2015) and *Uncanny* (2015). Each of these films deals with a love affair between a human being and a more or less human-looking (that is: android) robot. In each movie, the main philosophical theme is Turing's question whether it is possible to distinguish the intelligence of an android robot from a human being, and each time it runs poorly if not tragically for the human protagonist.

The Turing Test

Before analyzing the three movies, let me first say a few words about the Turing Test. In his 1950 paper 'Computing Machinery and Intelligence', Turing starts by remarking that the question 'Can machines think?' is very difficult to answer, because there is not much agreement among philosophers about what the word 'thinking' exactly means.[10] Moreover, we have no access to the 'inner ' of the computer. That applies, according to Turing, to human beings as well. After all, we can neither 'look in the head' of our fellow humans to determine whether there is some intelligent spirit or consciousness active. According Turing, for that reason it is better to look at the behavior of the machine in order to determinate if it is intelligent.

Turing describes an ' imitation Game ' for three players, where a man and a woman are in one room, and an interrogator in another room. They communicate via notes. The interrogator has formulate questions that will help him (or her) to find out who of the two persons is the woman. The woman must answer all questions correctly, the man must try to deceive the questioner. In the 1950 version of the imitation game Turing replaces the man by an intelligent machine, pretending that it is the woman. In a later version, often referred to as the standard version, the man and woman are replaced by a human and a computer. The task now is to determine who is the human being and 'who' the machine. The human has to assist the interrogator, the 'intelligent machine' (actually a chatbot, a software program) has to try to deceive the interrogator. In this case the communication takes place through a teleprinter. In order to pass the test, the intelligent machine – that is the computer program it runs – has to be able to convince at least 30% of the interrogators for at least 5 minutes that it is the human being. Turing predicted that it would take 50 years

before a machine would be able to pass the test, that is: in the year 2000. As we saw, the prediction was not that bad: in 2014 the first computer program passed the test.

The Turing test is not uncontroversial. Is being able to conduct a conversation not a far too narrow conception of human intelligence? Turing also observes that the best way to make an intelligent machine would be to build a complete human being, with senses and limbs, who could explore the world, and learn from his experiences. But even is such an artificial human being would be realized – which, according to Turing in his 1948 paper 'Intelligent Machinery' was not very realistic, because of the sheer size of the 'artificial brains' - "the creature would still have no contact with food, sex, sport and many other things of interest to the human being".[11]

The most fundamental criticism one can bring forth against the Turing test is that in the test it's not so much about testing intelligence, but rather about generating the *illusion* of it, that is, arousing in the interrogator the feeling that he or she is dealing with a human being. Is it therefore that Turing calls the idea of intelligence "emotional rather than mathematical"?[12] After all, deception is already ingrained in The Imitation Game, as it is the task of the computer to *deceive* the human interrogator.

Seen in this light, the title of the movie about Turing's life - *The Imitation Game* – is, in spite of the fact that the Turing test plays no role in the movie, well chosen. After the 'intelligent machine' designed by Turing had cracked the Nazis' Enigma code, the British had to deceive the Germans, preventing that they would realize that the enemy knows the code, because in that case they would immediately change it. Turing therefore refused to warn the ship on which the brother of one of his employees was located, for an imminent attack. And this was not the only deadly game in which Turing became trapped. Because was prohibited by law in those years, he was forced to play that he was heterosexual. After being caught, Turing was forced to go to jail or to be chemically castrated. He choose chemical castration, causing the development of female breasts. It was his last 'imitation play'. Only 42 years old, he committed suicide by taking a bite of a poisoned apple in the footsteps of his favorite film – Disney's *Snow White* (1937). Unfortunately for Turing, there was no prince to awake him from death with a life-giving kiss.

Perhaps the Turing test still fascinates us so much because it is based on ruse and deception. Since 1990 Hugh Loebner organizes an annual competition in which chatbots try to convince a jury that they are human beings. Despite the limited nature of this intelligence test, up to 2014 no chatbot succeeded to deceive the jury. But in that year the Russian chatbot 'Eugene Goosman' succeeded in convincing the jury in more than 30 percent of the tests at least for five minutes – Turing's criterion – that it was a human being. According to critics, that was mainly due to a trick: the chatbot presented itself as a 13-year-old Ukrainian boy, who had learned English only as a second language.[13] The interrogators were in fact deceived by a 'digital dummy'. Some critics were of the opinion that this was unfair and argued that Eugene Goosman did not pass the Turing test.

The Turing test in recent science fiction movies

But maybe the Russians did exactly what the Turing test is all about. Turing might be right in presupposing that the ability to deceive is an important quality of intelligence. At least, it is difficult to escape this impression when one watches the three recent science fiction movies

I've already mentioned. In *Her*, the main theme seems to be self-deception. The film deals with the lonely Theodore, who falls in love with the new talking operating system of his computer (with the voice of Scarlett Johansson), named Samantha. She appears to be a future version of virtual assistants like Siri (Apple) and Cortana (Windows 10). That Theodore falls in love with his operating system seems naive, as he knows that Samantha is just a computer program. However, by now it is a well-established fact that we emotionally attach ourselves to interactive devices, especially when they ask our attention (the Tamagotchi) or give attention to us, think about care robots like Zora, to which the elderly pour out their hearts). That tendency only becomes stronger if we do not know to have an artificial of doing. That happened for example the American psychology professor and Turing test specialist Robert Epstein. In an amusing article in the *Scientific American* (2007) he describes his amorous correspondence with a Russian lady through a dating site. Only after four months he got suspicious, and he emailed the handsome Ivana: "asdf; kj as; kj I; jkj; j; kasdkljk; klkj KLASDFK; Asjdfkj. With Love, Robert." When Ivana cheerfully sent him another mail about her mother in reply, Epstein finally realized that 'she' actually was a chatbot, and that he was the victim of self-deception and "darned clever programming."[14]

In *Ex Machina*, the central theme seems to be self-deception, too. In this film, the software programmer Caleb, who works with the search engine Bluebook, is invited by the eccentric founder of the company for a variant on the Turing test with Ava, a beautiful female robot. Ava may possess particularly attractive body shapes (the film is a long masculine fantasy), a transparent part of it clearly shows its mechanical inner. Can Ava convince Caleb that she has real emotions, although he sees that she is a robot? The loving Caleb fails his Turing test: he is seeing blind, and just like Theodore in *Her,* it runs badly with him. However, the fact that Ava turns out to be an android femme fatale and also kills her creator Nathan, who reigns as a modern Bluebeard over his collection of robotic sex slaves, raises the uncanny question of whether Ava's emotions might be real, and the even uncannier question whether it will be still possible in the future to find out.

The idea that in the future we might no longer be able to distinguish people and robots is the central theme of *Uncanny*. Just like in *Ex Machina* in Uncanny one of the characters – in this case the female science journalist Joy (a former artificial intelligence student who dropped out) – is subjected to an emotional Turing test by an evil genius, called David, for a week. Joy is confronted with the male android robot Adam, who in appearance and behavior – he even has a digestive organ – is not distinguishable from a human being. Like in *Ex machina,* sexual attraction plays a role. Yet *Uncanny* is also the counterpoint of *Ex Machina*, because we are dealing here with a male android robot, and the love initiative this time emanates from the robot. When Joy engages in a love affair with David, Adam begins to exhibit obsessive jealous behavior and the Turing test takes a bizarre twist. I don't want to spoil the surprising plot for those who did not see the movie yet, but I can assure you that in this case not only Joy is being fooled during this 90 minutes Turing test, but the spectator as well.

Origins of the Turing test

Although the Turing test is closely connected to the computer age, several centuries earlier mechanical dolls already raised similar questions. In the seventeenth century, for example, Descartes, confronted with mechanical ducks and other automatons, wondered how we can

distinguish a perfectly moving mechanical doll from a human being.[15] Descartes could still reassure himself with the idea that such dolls are unable to carry out a meaningful conversation with us. However, such an reassurance no longer seems possible in the epoch of computers, artificial intelligence and speaking robots.

Maybe the most fundamental reason artificial intelligences evoke our deepest fears, is that they affront the narcissism of the human species. Darwin had already convincingly argued that we are not exceptional, God-created beings, but an ape-related primate species. And now robots are even blurring the distinction between human and lifeless matter. Luckily, so far all this is mainly happening in science fiction movies In the real world, digital dummies like Eugene Goosman still dominate the overtone. And if we think of the perplexing complexity of human intelligence, that will remain the same for the time being.

However, the Goosmans of the digital world already have begun to 'test ' us. For example, think of Captcha's (Completely Automated Public Turing test to tell Computers and Humans Apart), the programs on websites that ask you to enter a set of letters or numbers to check if you are a human or a robot.

It is all the more apparent that our lives are increasingly in the grasp of such algorithms. I am, unlike Hawking, not so much afraid that in the near future we will be outstripped by superior intelligences. My fear is rather that we will be dominated by inferior artificial intelligences. And that, after the human race have long been extinct, they will continue to reproduce themselves thoughtlessly until the end of time.

Or is this a Western, too Western anxiety and preoccupation?

Asian alternative?

During my guest professorship at Ritsumeikan in 2016, I spent part of my time studying the Japanese approach to robotics. Asin Europe, in Japan there was an early interest in automatons, such as mechanic tea-serving dolls. But whereas in Europe such android – humanlike - robots did not become very popular, in Japan, especially in the Kansai region, there is a strong interest in android robotics. During the writing of my book *Artificial by Nature. On the way to Homo sapiens 3.0*, I became fascinated by this Japanese preference for androids. And during my stay in Kyoto I've visited the Ishiguro Laboraties, which are part of the Advanced Telecommunications Research Institute International (ATR) in Kansai Science City.

Hiroshi Ishiguro, professor of Robotics at the Graduate School of Engineering Science at Osaka University, is one of the most famous researchers in this field. He acquired international fame in 2006 with his robot Doppelgänger, Geminoid HI-1 (from Latin *Gemini*, twins, see www.geminoid.jp/). My first visit to his laboratory was to attend a demonstration session of this geminoid robot and of his latest model, the communication robot Erica.

Geminoid HI-01 actually is a so-called telerobot. He does not act independently, but is controlled *in real time*   by a human operator, in this case Ishiguro himself. The robot has been made in collaboration with artists. The robots have a metal skeleton and a plastic skull, and especially their silicone skin and human hair (from Ishiguro itself) makes HI-01 eerily realistic. The movements – Ishiguro's Doppelgänger has fifty different movements and facial

expressions – are produced by means of an air compressor and pneumatic actuators. They run synchronously with the movements Ishiguro, located in another room, is making. His movements are translated into robot movements via tracking devices and teleoperation software. With the help of a microphone and speakers in the geminoid, Ishiguro can also speak through the robot. Ishigoru himself undergoes impressions of his environment thanks to sensors and cameras in the robot.

Ishiguro uses his mechanical Doppelgänger, among other things, to hold lectures from his home at his university or elsewhere in the world. However, such practical applications – which also successfully serve his public relations - are only one side of Ishiguro's *android science*. He also uses the geminoids for (neuro)psychological experiments in the field of human-robot interaction and telepresence. One of the notable findings is that even without tactile feedback the user will undergo a physical sensation when his twin robot is touched. Probably, mirror neurons appear to play a crucial role in this phenomenon. When watching an action by another person, the same neural patterns are triggered as those which are active during the performance of the action. This not only plays a role in evoking empathic feelings for other people, but also in identifying with a robotic body and literally including them in their own body schedule.

Ishiguro's telerobots realize materially what the philosophical anthropologist Helmuth Plessner calls the excentric positionality of man.[16] According to Plessner, we have a threefold relationship to our body. Like plants, we *are* our body, and like other animals, we *have* our body in the sense that we can control it thanks to our nervous system and brain which constitute the center of our experience. However, we distinguish ourselves as human beings form other animals because we are also positioned *outside* our bodies in the sense that we can reflect on ourselves from an ex-centric position. Whereas in ordinary life, sitting on the sofa in our house, we can only imagine that we walk through Manhattan in New York, rescue a victim from a burning house, or walk on Mars, thanks to telerobot – which we can equip with cameras and microphones and artificial senses like infrared eyes and echo location and artificial limbs, we can actually do these things now.

Interacting with the robot Doppelgänger is a remarkable experience. Because of the strong resemblance to the real Ishiguro and the fact that we speak with 'the man behind the robot', we are inclined to approach the robot as a human being. However, as the mechanical and stereotypes movements and facial expressions break through that familiar pattern, cognitive dissonance occurs.

The Japanese robot designer Masahiro Mori argues in an article from 1970 that things make a more familiar impression as they appear more like a human being, but that in cases where the similarity becomes very close, but not complete – as we experience with bodies of deceased people or zombies in a horror movie, and android robots, the familiarity turns into disgust. [17] In a chart in an article Mori discusses this this phenomenon he calls it *bukimi no tani genshō*. In English it became *Uncanny valley*, with which a link was made to Sigmund Freuds essay on the uncanny (in German: 'Das Unheimliche'). Freud connects the uncanny among other things with mechanical Doppelgängers, and associates it with the fear of death, suppressed sexual feelings and narcissistic feelings of omnipotence.

However, during my visit to Ishiguro's laboratory, my curiosity was especially focused on Erica. Erica is also an android robot, but she's not a geminoid in the sense that she's not a copy of a specific human individual. Instead, she has been designed by Ishiguro on the computer by combining thirty pictures of women he found particularly beautiful. (Having Freud in my mind, involuntary I had to think of the aforementioned movie *Ex Machina*). Erica is an autonomous robot, which interacts with its surroundings without immediate control. However, 'her' freedom of movement is limited. Like Ishigiro's Doppelgänger she can't walk, but is attached to her seat. This has a reason. Erica, an acronym for ERato Intelligent Conversational Android, is an Android version of a chatbot that provides additional problems. To be able to communicate, Erica must be able to hear the voice of her interlocutor. However, auditory *speech recognition* in noisy rooms with a group of people – visitors who attend the demonstration of Erica - is a particularly difficult task. For this reason, Erica is invisibly connected to a number of microphones and sensors, so that she can pinpoint her interlocutor even when he or she is moving.

Erica, whose voice is synthesized in real time, is not only able to respond to questions and remarks, but she also continues to follow her interlocutors with her eyes and her face shows the emotions appropriate to the conversation. At least, that was what Dr. Takashi Minato, who accompanied Erica's demonstration, told me. As the day I attended the demonstration the audience was mainly consisted of Japanese visitors, the question and answer session was entirely in Japanese and I do not master Japanese. However, the fact that Dr. Minato explained the demonstration to me in English, confused Erica a couple of times. As she can handle only one language at a time, she turned several times to our English conversation without being able to change to English and to reply to us.[18]

Afterwards I discussed with Dr. Minato the many obstacles on the way to the perfect conversational robot. The focus of Erica on her interlocutor is still child's play compared to the task of letting her make an everyday conversation about every possible subject. Although her skills, thanks to the use of an *open domain conversational system* (which searches the web for usable, similar dialogues, roughly like Google translate does in translating) and *deep learning* (a contemporary version of neural networks) are quite impressive, the number of subjects that can be discussed with Erica is currently still very limited. So you can talk to her about her hobbies, but in many other subjects she keeps the boat off (as she also does - apparently not just a human habit - if you ask her for her age).

The Total Turing Test

It is Ishiguro's ambition to develop Erica to the level that will enable her to pass the famous Turing test. That means, as you will remember, that Erica should be able to fool at least 30% of her interlocutors for 5 minutes that she is a human being, regardless of the subject being addressed. As in 2014 the chatbot Eugene Goosman has passed the Turing test, this seems to be a realistic goal.
       However, Ishiguro's ambitions reaches much further, his aim is a robot capable of passing the Total Turing Test. Unlike the traditional Turing test, where the call takes place via keyboard and screen, this requires that the future version of Erica – like the robot Ava in the *Ex Machina* – a questioner in the same room must be able to convince that she is a real

human being. The experiments conducted by Ishiguro so far show that 20% of subjects record within one second that they deal with a robot. And in my estimation the remaining 80% will follow in the next four seconds.

The hype is no less so. The day after I attended the demonstration, I came across a page big picture of Erica on the back page of the New York Times with the text "Meet Erica. She didn't go to school. She doesn't have DNA. Soon She will be smarter than you. " However much I was impressed by Erica, the demonstration suggests that this ' soon ' strikes an evolutionary time scale. When I hold a lecture on the Turing test a few weeks later for Ishiguro and his team, I notice how optimistic Ishiguro and his team are about the possibility of traversing the *uncanny valley* and passing the Total Turing test. Ishiguro's optimism is not shared by everyone. The critics not only doubt the feasibility of the project, but also its usefulness. Planes can fly excellently without doing it like a bird. And a conveyor belt is a lot more efficient than android robots carrying boxes.

In the discussion that followed my lecture, Ishiguro frankly admits that android robots are not in all cases the best solution. However, in the case robots enter the social world, they might be the best option. Human beings have been evolutionary adapted and from childhood on trained in dealing with other humans. According to Ishiguro, robots that resemble people will not only facilitate a smooth interaction between robots and human beings, but also enhance the acceptance of robots in the human society. Something that looks like a human being is more likely to be also treated as a human being. In *Human-Robot Interaction in Social Robotics* (2013)[19], Ishiguro discusses a series of field experiments with android guide robots in a museum, a shopping mall and a train station. These experiments not only confirm the effectiveness of the services, but also show that most adults and children have a lot of fun in interacting with the robots.

Because the robots remember the questions and preferences of their customers with the help of ID chips, there is even something like a ' personal bond ' with the robots, and even emotional attachment takes place. At least, as long as the uncanny valley has not been crossed. That's why cartoonized robots often work better than 'photorealistic' androids do, as they don't cause the kind of disgust the imperfection of the latter do. This explains the success of the catoonized care and communication robots designed by Ishiguro and other robot engineers, which are applied in the care of the elderly and the education and the guidance of autists. In the Netherlands and Belgium cartoonized care robots like Alice and the aforementioned Zora made their appearance, too. The difference with Japan, however, is rmarkable. Japan is the most strongly robotized society in the world. Not only do more than 250,000 industrial robots do their daily work, the government and the business world have invested billions in the development of social and affective robotics over the past decades. Robots are immensely popular. They dance and sing at electronics fairs, Flasher bridal clothing on the catwalk and perform in television programs. The annual ROBO ONE robot competitions are attended by whole families. And from the introduction of the Robot Dog Aibo in 1999, consumer robots eagerly find deductions. The cartoonized android Pepper, marketed as the first emotional robot, was sold out in 2015 within one minute of its introduction.

Since 2015 in the theme park Huis ten Bosch in Nagasaki, enthusiasts can also stay in the (practically) completely robots-run Hennna Hotel, and in the film *Sayonara* (2015) Ishigoru's female Geminoid-F plays one of the leading roles. At the Tokyo International Film Festival she even was nominated for the best female leading role. The film, based on the Oriza Hiratan's Theatre production *Sayonara II*, deals with the friendship between a human and an android, which develops in the aftermath of a Fukushima-modeled nuclear disaster. The central theme is their vulnerability. The android even was nominated for the best female actress at the Tokyo Film Festival. With the stay of the companion robot Kirobo in the International Space Station, this human-robotic friendship model has now passed the cinematic imagination.

It is often claimed that the Japanese love for the robot is a practical necessity. Not only economically, to be able to compete with the rapidly developing economies of other Asian countries like China, but also socially. Japan has the strongest ageing population in the world and has a relatively low number of immigrants that could lower the average age. As a result, the country suffers from an ascending shortage of personnel in care, education, and other service professions.

However, the practical necessity seems insufficient to explain the emotional bond that Japanese maintain with their robots and which encourages them to bless their robotic dog or android by letting Shinto priests and with a Buddhist ritual to their last resting place to accompany. Such uses suggest that the unique other relationship that Japanese maintain with their Android robots cannot be detached from their worldview, which in some respects differs greatly from the western.

In the Western, Christianity-shaped culture, a taboo rests on 'playing God'. Man has been appointed by God as a steward of nature and may even experiment with it, but the creation of life, especially human life, is the privilege of God. Geneticists and robotics who violate this taboo, even in the most secular western societies, are quickly dealing with the reproach to act from *hubris*, in the Christian tradition the mother of all mortal sins. And as also the Greek tragedies learn, *hubris* leads to disaster and catastrophes.

It is for that reason that the vast majority of Western science fiction that deals with robots is apocalyptic in nature. Earlier in my talk I already referred to the bad ending of the movies *Her, Ex Machina* and Uncanny, but these are no exceptions, they stand in a long tradition that goes back at least to *Frankenstein* (1817) of Mary Shelley and also characterizes iconic science fiction movies like Cyberspace Odyssey (1968), *Blade Runner* (1982) and *Terminator* (1984) and , androids are invariably focused on the destruction of humans. Even in the famous robot stories of Asimov, which revolve around the three robotic laws that must prevent robots from ever doing harm to humans, the plot almost always revolves around its circumvention. And that apocalyptic view is not confined to fiction. And as we have seen, Western scientists also warn against robots taking away people's jobs, drones causing death and destruction, or even starting to dominate the human race and eventually will replace them. Even in the most optimistic versions – like the paradisiacal end-time fantasies of Hans Moravec and Ray Kurzweil about robots and singularities which transcend humans – there is no longer a place for *human* being.

In comparison with this, the robotic image in Japanese science fiction is usually much more positive. Robots like Astroboy, the main character in a manga comic that appeared between 1952 and 1968 and has been repeatedly edited into animated film, are not enemies, but helpers of humanity. They do not form the evil negative of man, such as Freud's uncanny Doppelgänger. Human beings and robots are rather – as in the movie *Sayonara* - of the same nature and relying on each other.

This cannot be dissociated from the worldview of Shintoism and Buddhism, which does not know the separative cosmology, which dominates the modern western worldview. Whereas in Western thinking the separation between life and death, body and mind, man and animal, man and woman is often absolutized, in Asian thinking the boundaries between the opposing parts are much smoother and blurred, as the paradigmatic yin-yang symbol shows. This is reflected in many aspects of everyday life, but especially in the religious traditions.

This seems to color the Japanese attitude towards robots in a fundamental way. In Shintoism, for example, everything is attributed a spiritual dimension, a *kami*. 'Kami' is a concept which is difficult to translate in Western languages. It refers to 'holy powers', so sometimes it is translated as 'gods' or 'spirits' (for example those of venerated dead persons), but it can also refer to living human beings, to other animals, trees, plants, and even to stones, mountains, oceans - all may be kami. According to Edo scholar Motoori Norinaga "...any being whatsoever which possesses some eminent quality out of the ordinary, and is awe-inspiring, is called *kami*." [20] Within this worldview it is not strange that *kami* is also attributed to robots. And from a Buddhist background, the robotic engineer Mori suggests that robots also strive to realize their Buddha nature.[21]

Like in the West, Japanese people compare with animals, but that is not only to appoint the differences, but also, precisely, the *similarities*. Traditionally mostly were apes that fulfilled this mirror function, in modern Japan that role is played ever more frequently by robots. Within the reflexive anthropomorphism that characterizes the Japanese worldview, robots are not opposed to humans, but they share a common nature.[22]

The Turing-Wittgenstein debate

Although relativizing the difference between human beings and other animals is gradually entering the Western worldview (think about the work of the popular primatologist Frans de Waal[23]), attributing a spiritual dimension to robots seems to be still a bridge too far. Certainly within a mechanistic worldview, as we find it paradigmatically expressed in the title of Julien De la Mettrie's book *Machine Man* (*'l Homme machine*[24]) In this view, to quote the contemporary successor of De La Mettries's materialism, the famous American philosopher Daniel Dennett, who claims that human beings actually are nothing more than 'moist robots'.[25]

However, in my view in Western thinking we also see a gradual turn to overcome the unfruitful opposition between materialism and spiritualism. Let me try to explain this by returning for a moment to the movie *Ex machina.* Not only Turing's famous test plays an important role in this movie, but also the spirit of Turing's intellectual rival Wittgenstein is remarkably present. The name of Nathan's company, Bluebook, refers to one of

Wittgenstein's writings[26], and that is just one of the many references to Wittgenstein in the film. For example, in one of the scenes, we see Gustav Klimt's famous portrait of Wittgenstein's sister hanging on the wall of Nathan's house.

In the years that Turing worked on his 'thinking machines ' he attended the lectures of Wittgenstein on the foundations of mathematics.[27] Like Turing, Wittgenstein was a philosophical behaviorist, who believed that you cannot separate the mental life from the bodily behavior. Both agree that intelligence is not a mysterious inner power, but shows itself in the behavior of human beings. On one point, however, they were passionately disagreeing. In his *Bluebook*, Wittgenstein argues that the question of whether a machine can think is just as nonsensical as asking for the color of the number three.[28] The meaning of words lies in their use, and in the ' language game ' around machines words like ' thinking ' and ' emotions ' do not belong. According to Wittgenstein, attributing an autumn depression to a vacuum cleaner is a blunt category mistake or a poetic metaphor at best.

However, in this particular debate Turing seems to be more consistent in his behaviorism than Wittgenstein. He follows the motto that something that looks, swims and croaks like a duck, is also a duck. A perfect robot that in all its behaviors is not distinguishable from a human, the predicate thinking cannot be denied. Turing, more than Wittgenstein, seemed to realize that language use changes with our practices. Where it is indeed rather absurd to attribute intentions or emotions to a traditional vacuum cleaner, we are inclined to do so in the case of artificial intelligences. So my grandchildren believe that the vacuum cleaner robot turns around before the stairs because he is afraid to fall down. However, they perfectly know the robot is programmed that way. But aren't we programmed as well, in this case not by a human programmer, but by natural selection in the course of human evolution?

The grammar and vocabulary of Western philosophy still needs finetuning if it tries to develop a common future for humans and robots that is mutually advantageous. Here, the Japanese worldview, imbued by Shintoism and Buddhism, seems to be conceptually far better equipped to think and shape the common future of humans and robots and can be a source of fruitful inspiration. Of course, we have to prevent digital orientalism, an uncritical idealization of oriental robotic wisdom. Not only because the opposites between East and West are not absolute (who is of this opinion, is still victim of a problematic separative cosmology), but also because also in the East there is often a clash between the high ideals and expectations on the one hand, and the often crude reality on the other. Think for example of the Buddhist violence against the Rohingya minority.

And even in robophile Japan the introduction of robots is not without problems. This is, for example, already being read out to the user agreement that SoftBank allows the buyers of Emorobot Pepper to sign, and which includes the provision that the user promises to refrain from sexual and otherwise indecent acts with the robot. Contrary to the *Campaign against Sex Robots* initiated in England in 2015[29], that provision seems not so much to be motivated by the fear that such robots dehumanize sexuality by transforming sex into a commodity, but mainly by concern for the fragile robot soul.[30] The Buddha nature has not yet been realized in Japan. Maybe we should be grateful for the fact that that there is not yet a consumer version of Erica for sale.

Endnotes

[1] See http://www.ritsumei.ac.jp/ss/education/professional/mul.html/

[2] https://www.techtimes.com/articles/41932/20150324/robots-replace-half-jobs-20-years.htm

[3] https://www.dailymail.co.uk/sciencetech/article-3099703/Humans-left-defenceless-killer-drones-Flying-AI-robots-pose-threat-lives-expert-warns.html

[4] https://futureoflife.org/ai-open-letter

[5] https://www.bbc.com/news/technology-30290540

[6] http://www.reading.ac.uk/news-archive/press-releases/pr583836.html

[7] Turing, A. (1937). On computable numbers with an application to the Entscheidungs problem. *Proceedings London Mathemetical Society*, 42 (July), 230-265. In this article, Turing presents three proofs of the thesis that that some decision problems are "undecidable". In these proofs Turing uses an imaginary like "computing machine". This type-writer like device is able to process a simple set of rules in a mechanistic way. This leads Turing to the idea of a "universal computing machine", "a single machine which can be used to compute any computable sequence" (p.241).

[8] Turing, A. (1950). Computing Machinery and Intelligence. *Mind, LIX*(2236), 433-460.

[9] https://www.iwm.org.uk/history/how-alan-turing-cracked-the-enigma-code

[10] Turing A. Computing machinery and intelligence. *Mind* 1950; LIX: 433-60.

[11] Idem, 420.

[12] Idem, 431.

[13] http://www.reading.ac.uk/news-archive/press-releases/pr583836.html

[14] Epstein, R. (2007). From Russia, with Love. How I got fooled (and somewhat humiliated) by a computer. *Scientific American Mind* (October/November), 16-17.

[15] Kang, M. (2017). The mechanical daughter of René Descartes: the origin and history of an intellectual fable. the mechanical daughter of rene descartes: the origin and history of an intellectual fable, 14(3), 633-660.

[16] Plessner, H. (2019). *Levels of Organic Life and the Human* (M. Hyatt, Trans.). New York: Fordham University Press.

[17] Mori M. The uncanny valley. *IEEE Robotics & Automation Magazine* 1970; 19: 98-100.

[18] See for an English speaking conversation: https://www.youtube.com/watch?v=_NTj88EdPtM.

[19] Ishiguro, H., & Kanda, T. (2013). *Human-Robot Interaction in Social Robotics*. Boca Raton: CRC Press. Taylor & Francis Group.

[20] Gall, Robert S. (1999). "Kami and Daimon: A Cross-Cultural Reflection on What Is Divine". *Philosophy East and West*. **49** (1): 63–74.

[21] Mori, M. (1982). *The Buddha in the robot: a robot engineer's thoughts on science and religion*. Tokyo: Tuttle.

[22] Is it therefore that the Japanese behavior in Western eyes often appear to be somewhat robotic? In the local train that took me from Ryoan-ji station towards center, I look at every ride fascinated to the machinist-conductor, who runs a script at lightning speed. Before every single act – looking in the mirror, on the clock, on the time schedule; turning the handle that starts the train - ,he points with his stretched finger to the object.

[23] "We start out postulating sharp boundaries, such as between humans and apes, or between apes and monkeys, but are in fact dealing with sand castles that lose much of their structure when the sea of knowledge washes over them. They turn into hills, leveled ever more, until we are back to where evolutionary theory always leads us: a gently sloping beach.", ??.

[24] La Mettrie, J. O. de (1981 [1748]). *L'homme-machine*. Paris: Denoèel/Gonthier.

[25] Statement in an interview with Jennifer Schuessler: Schuessler, J. (2013, April 29). Philosophy That Stirs the Waters. *The New York Times*.

[26] Wittgenstein, L. (1958). *Preliminary studies for the "Philosophical investigations," generally known as the Blue and Brown books.* Oxford,: B. Blackwell.

[27] Wittgenstein, L., Bosanquet, R. G., & Diamond, C. (1976). *Wittgenstein's Lectures on the foundations of mathematics*, Cambridge, 1939 : from the notes of R. G. Bosanquet, Norman Malcolm, Rush Rhees, and Yorick Smythies. Ithaca, N.Y.: Cornell University Press.

[28] Wittgenstein, L. (1958). *Preliminary studies for the "Philosophical investigations," generally known as the Blue and Brown books*. Oxford,: B. Blackwell, 47.

[29] See https://campaignagainstsexrobots.org/

[30] See https://www.wired.co.uk/article/pepper-robot-sex-banned