



Dutch-Japanese Workshop in Philosophy of Technology

Saturday July 7, 2018

*Open session:
Robotics and Philosophy*

*Vulnerability, Risk,
and Humanity*

SHOJI NAGATAKI

CHUKYO UNIVERSITY

shojinagataki@gmail.com

- Human ancestors have survived a narrow pathway of evolution. **Our ancestors were vulnerable from the very first.**
- In the following, we consider the history of humans in terms of struggles for overcoming **vulnerability**, and situate contemporary **technologies** advanced through scientific and industrial revolutions on the background of those struggles.
- Then, we analyze what possible products of robotics mean to us from a viewpoint of **vulnerability**, with a special attention to the problem of **robot-human coexistence**.

i. Facing New Type of Risks

- Humans have devised a variety of tools and changed their environmental world in order to compensate for their vulnerabilities. As a result, while old types of risk, such as shortage of food, menace of predators and so on, have been substantially reduced, new ones have emerged.
- For example, various machines developed after the Industrial Revolution helped to relieve us, even though partially, of harsh manual labor.

- However, they have transformed the form of labor and work environment, which is conducive to the mismatch diseases "such as type 2 diabetes and heart disease" (Lieberman 2013, 269) caused by a poor balance between the amount of eating and that of energy consumption.
- There are new forms of vulnerability which have never been before.

[N]ew technologies ... always create new risks and vulnerabilities, thus transforming human vulnerability rather than substantially reducing it. (Coeckelbergh 2013, 12:5)

- We can view the current development of AI and robotics from a historical perspective of coping with human vulnerability.
- It has reduced the burden of vast calculation and precise work that are troublesome for humans.
- Just as considerable part of physical labor has been taken over by machines since the Industrial Revolution, more and more part of intellectual labor is being transferred to them since the latter half of the 20th century.

- Some people respond as modern Luddites, having fear that human labor would be negatively affected by the newcomers.
- This situation is also, though not very serious, a manifestation of human vulnerability.
- According to Cartesian ontology, products of scientific technology, no matter how excellent they are, belong to *res extensa*, thus essentially different from humans in a metaphysical sense.
- In the near future, robotics and AI research may succeed in creating very humanlike beings which can exceed human intelligence in many senses.

- Human desire to replicate themselves may make such beings more than just industrial products, just *res extensa*.
- Those beings, which are a kind of externalization of human intelligence, could become *res cogitans* like us in the sense of beings with mind and consciousness.
- We might even notice within us "the archaic remnants of emotions which may linger in our revulsion" (Habermas 2003, 25) against such beings.
- In that case, we will have a much more crucial problem of whether we should accept such intelligent and humanlike robots as our partners,

- What is necessary for such robots to be accepted as social members for us, or to coexist with us?
- How can they be not just mere objects, but intersubjective beings which can share with us “processes of reaching understanding and self-understanding” (Habermas 2003, 10)?
- My thesis is that they have to be a *moral agent* with a kind of humanity.
- Otherwise, such robots can be a new type of *significant risk* for us.

2. Robot As Moral Agent

- What kind of beings do humans accept as moral agents? Analyzing situations in which someone is deemed a moral agent, there are, among others, two conditions to be met.
- First, it can be seen as being basically similar with each other in terms of bodily structure, cognitive ability, and so on.
- Second, despite those similarities, there is a variety of differences in each individual, some of which are inscrutable from the first-person perspective.
- In the following, I would like to clarify what similarity and difference are.

i. Embodiment and Psychological Abilities

- Psychological abilities specific to humans are bodily restricted. Our cognitive style is largely determined by physical features that we have.
- For example, the perceptual world appears in a perspectival way due to the bodily constraints. In spite of, or rather because of this perspectivality, our cognitive ability or intelligence functions in such a way as to extend a limited range of information.

- In addition, human cognition is not based on symbol processing separated from the environmental world, but on bodily interacting with and cognitively adapting to it.

- Various individuals are also included in our perceptual world. When communicating and interacting with them, the condition that a physical isomorphism holds between us is of great importance.
- Whether the other is a human or a human-like robot, clues to properly capture its intention are provided by our having similar bodies.
- Such similarity helps us to predict how it perceives the outer world and what intent it has.

ii. Bodily Similarity and Ontological Homogeneity

- As developmental psychologists have shown, even a newborn child can understand the basic meanings of others' expressions through a primitive body scheme. Understanding facial expressions is one type of imitative behavior with an implicit intention, and as such very primitive.

- From a phenomenological point of view, this kind of psychological ability is based on bodily similarity and ontological homogeneity among us.
- The similarity makes it possible for us to intuitively comprehend, for example, the ways how organs such as eyes and a mouth function.
- It also enables us to recognize our relationship with things in general.

- "[T]he primordial intersubjectivity" (Zahavi, 1999, 171) has always already been established between myself and others by our participating in the world.
- Self and others have reached a fundamental mutual understanding tacitly and preconsciously on the basis of the ontological homogeneity --- that is, what "makes us simultaneous with others ... in the most private realm of our life" (Merleau-Ponty, 1960, 24).

iii. Skin, Vulnerability, and Humanity

- Humans can be subject to pain, illness, injury, disability, and death. They can feel pleasure as well.
- Injuries and diseases are caused when the skin and inner tissues are physically damaged, when ultraviolet rays erode the skin, or when bacteria, viruses, or toxic substances affect the body through the skin-boundary.
- No matter how science and technology progress, our vulnerability comes down to the fact that the skin is thin and susceptible to damage.

- The direct and mutual relationship between humans and the world and things in it is sometime expressed by *a metaphor of touching*.
- If robots and AIs produced by scientific technologies in 21st century do not share the common ground with us and the world in some way, they cannot be moral agents.

iv. Alterity, Irreplaceability, and Machine as a Moral Agent

- The discussions so far suggest the importance of bodily similarity and ontological homogeneity with us.
- However, these do not suffice for something's being a moral agent. There is another element to be considered: alterity, or otherness, against such affinities.

- Alterity means an irreplaceability in some essential respect. To be a moral agent is to bear its own responsibility which others cannot take for it.
- My thesis is that such an irreplaceability consists in its having a rich inner world. The personhood of a moral agent, which is irreducible to a mere difference of trait or feature of individuals, is firmly rooted in such an inner world.

- The irreplaceability can be viewed along another dimension; it is related to the problem of whether a first-person perspective can be attributed to the other in question.
- This kind of perspective involves a private realm to which other people cannot have direct access, and which provides one reason for us to treat something as the other and to accept it as a moral agent.

- Such private realm is where our personality and irreplaceability, including that of moral responsibility, lie in. On Merleau-Pontian conception, such a realm is based on the ontological common ground.
- Habermas makes much the same point when he talks of “the morally relevant limit to instrumentalization” of other people. He discusses genetic intervention in humans, arguing that the limit “is set by what, in the second person, will be out of my reach” (Habermas 2003, 55) .
- This out-of-reach-ness, which resonates with Levinasian thought, is an essential element that constitutes rich inner-world of others.

- Thus, for example, a machine which functions in a predictable or required way does not have its “alterity,” even if it is as good an industrial product as can be.
- When coordinating ourselves in order to engage in a cooperative activity, we will feel an affinity between us, while when failing in it, a sense of alterity, impenetrability, or inscrutability will be imposed upon us.
- In fact, such alterity is very familiar. It is a common experience that we find similarities as well as differences between us.

3. Concluding Remarks

- As I discussed in 2-v, human beings have their own inner states which are inscrutable each other. This aspect is, in relatively large part, realized by human specific psychological abilities.
- Alterity and morality are based on the irreplaceability which is closely related with such inscrutable inner affluence.
- If we can implement this inner affluence in robots, they are probably become moral agents for human beings in some way.

- As I argue in 2-iii and 2-iv, it is necessary for something's being a moral agent (or a moral subject) us to share the ontological homogeneity with a human being. This homogeneity can be grasped by the concept of, for example, common vulnerability with human. It is impossible to realize such human vulnerability in machines like robots at the current moment.
- Is it reduced to the technological problem of (bio-)engineering realization of a humanlike skin? Or is there a deeper metaphysical problem?