

Reply to the questions from Ms. Karin Wegsjö, \*k\*wp in Sweden

Nov. 1<sup>st</sup>, 2016  
Masayoshi Shibata  
Vice President of Kanazawa University

1. What would it mean to make robots with personality?

Basically it means to create robots with mind. “Mind without personality” is a conceptually confused idea. We are now making robots without mind all over the world, but, at next the stage, we can’t help but making ones with mind. Because making robots with personality originates from our innate desire and curiosity to make whatever is technologically possible. So we can’t stop it.

On the other hand, we increasingly need to have our partners or friends when we get older and closer to death, though we couldn’t always get them. And it becomes harder and harder for elderly husband and wife to care each other. So many elderly couples would want to have nursing-care robots, and a human touch relation with them. Therefore making robots with personality is what meets our demand in the aging society in the future.

2. How to develop a moral society with robots?

It is very difficult to describe a concrete procedure to build and reach a moral society with robots. That society presupposes the existence of robots with autonomy and mind. So it is a little far away from now, and I have to leave the work of telling the story of how to develop such a society to the imagination of novel writers or film producers, rather than to the analytic thinking of philosophers like me. Unfortunately it is beyond my power.

3. How do robots develop morality? Can they?

I believe robots could have mind and morality. In my view, morality is ultimately the adjustment among many agents’ interests, but there is no absolutely correct way of such adjustment. Even we, humans have to make a judgement about what is good and bad, not by the single authorized moral theory but by rules of thumb gained through our experiences. Essentially the same with robots.

In order to acquire moral sense, robots have to have the equipment of self-learning system, because they have to learn how to morally react to various real situations from their experiences. Not by pre-programed rules for actions, but by renewed sense of morality gained through interacting with humans and other robots, they would come to be moral agents.

Of course all of these robots must have the capacity of autonomous judgments and actions as prerequisites.

4. Who decides over this development? Which morality?

No one could decide whether some robots are autonomous moral agents or not, because it is a fact anyone cannot deny arbitrarily. But we could decide whether they should be admitted as a member of our moral community, as in the case of women or black people in the old days. And I think we should include robots into our moral community if they have mind and personality. It is us as a member of moral community who decide this matter.

I'm afraid I don't see through the point of your question, "Which morality?" Although I may misunderstand you, I will say what I take to be the related issue. I think it is our duty to create a new moral system, a new idea of morality for us living in the highly technological society in and after 21st century. But I can't figure out now what type of moral rules is appropriate for us. The only thing I can say now as a starting point is that the new moral system has to permit other moral agents to have "the greatest possible freedom", as J. S. Mill---Libertarianism in political philosophical sense--- has insisted.

5. What are the social consequences of unbridled technical development?

(a) I think everything has both good and bad aspects. If the political and economic situation is stable and the difficult global problems--- such as a deep gap between the rich and the poor people, and the refugee issues---are solved, we could have so much time to enjoy our personal interests and concerns. In my opinion as a philosopher, humans have, as their essence, 3 innate inclinations: Truth, Freedom, and Expression. (1) We are starving to know the truth at any risk. (2) We can't stand if we are not free. (3) We feel an irresistible desire to express ourselves. If things go well, we would flourish in holding this human identity. We could do well without any resort to religions.

(b) But if we employ new technologies, e.g. bio- or gene-technology to ourselves "beyond therapy", we might feel disturbed by the question of what the human identity is and who we are, and lose it at the end of the day. Losing the identity means the end of the human kind, although the other kind of existence might emerge and continue to live.

(c) But far before we reach that point I suggested in my presentation slide, we may destroy ourselves due to formidable weapons our new technology can make possible.

6. Who decides how this should be managed?

The answer: the human kind as a whole. Of course the government, the parliament, academic organizations, mass media, all kind of NGO, NPO, and so on, depending on each situation.

But in the end, our rational thinking and imagination would decide it.

7. In Japan, technical development has progressed even further; what consequences can we already see?

I could say that, now in Japan, we are at the halfway point between the good 7-(a) and bad 7-(c) results caused by the advanced technology. For example, we can see in recent news people's mixed feelings of hopes and fears about regenerative medicine involving IPS cells or reproduction technology such as artificial insemination.

8. How will we, as humans, relate to robots? About your study:

I think we need to divide 2 cases: ( $\alpha$ ) robots without mind or personality, and ( $\beta$ ) robots with it. In the case of ( $\alpha$ ), we will be basically treating robots no more or less than useful tools, however deeply we feel attachment with them like dolls in childhood. But in the case of ( $\beta$ ), as I suggested at my answer to 4, we will be living with them as our partners and giving them a membership of our moral community. It wouldn't be, in essence, different from the cases of encountering strangers from yet-unknown areas in the Earth or aliens from other planets.