I. Introduction

What is necessary for robots to coexist with human beings? In order to do so, robots must be moral agents. To be a moral agent is to bear its own responsibility which others cannot take for it. We argue that such an irreplaceability consists in its having an inner world. The personality of a moral agent is firmly rooted in such an inner world.

II. Purpose of This Presentation

It is necessary to implement similar bodily and psychological abilities in someone, or something to be accepted as a moral agent, or another person in a human society. Then the irreplaceability, which we mentioned, 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. This is where our personality and irreplaceability, including that of moral responsibility, lie in.

In fact, such an otherness is familiar. It is a common experience that we find similarities as well as differences between us. Suppose that you and I agree to have lunch together, but you force me to eat something I have never expected in a restaurant. In that situation, I would feel I’ve lost my initiative. This happens in our everyday life. We have a sense of otherness in unexpected transfers of initiative. In order to make explicit such an aspect of our daily experience, we design an experiment of Bodily Coordinated Motion Task (BCM Task). A bodily coordinability is a social art and one of the key elements which enables us to have a social relationship with others (cf., William H. McNeill (1997) Keeping Together in Time: Dance and Drill in Human History, Harvard University Press). When coordinating ourselves well and getting along with each other, we feel an affinity between us, while when failing in it, a sense of otherness or impenetrability is imposed upon us.

The purpose of this presentation is to explicate a condition in which humans attribute the status of moral agency to a robot. For that purpose, we are planning to set up an experiment of interaction between a robot and a human. Before that, we develop some hypotheses about the anticipated results of the experiment.

IV. Anticipated Results and Two Hypotheses

One hypothesis is that a subject will attribute a certain moral agency to the other (even to a robot) with whom the subject can bodily coordinate in a better manner. This is because the coordination involves the process of mutual understanding in some respects. Generally speaking, even with a new acquaintance of others, we naturally develop a concern for them. In parallel with that, we come to think that others should have a similar concern for us in turn. One can recognize a primitive basis for ethics in this situation.

Another hypothesis is that the richer world we recognize within others, the more demand for morality we make. On a simple setting of BCM Task, however, what is it like to recognize a richness -- or an inescapable realm which underlies personality -- within others? In our experiment, subjects may succeed in bodily coordination or fail. There are also conflicts as to which subject possesses the initiative. In the case of a conflict where the coordination once fails, a transition of the initiative eventually will take place, and a new coordination will hold, we presume. In the bodily coordination process with the other, you may have not only a sense of the opponent’s being in tune with yourself. You may also feel her/his resistance or the shift of the initiative to the other side. The experience of coordination can be a complicated one full of twists and turns.

We think that a subject will find a richness within his opponent through a complicated process of co-ordinations, divergences and transitions of the initiative. This process occurs when, for instance, the subject feels its opponent’s purposely making an unexpected move. In such a situation, it seems natural for us to attribute intention, desire, responsibility, and so on to others. This is when we recognize others as moral agents and accept them into the intersubjective world of morality.

III. Experimental Design

Participants

Students of university in Japan : 40 right-handed people (all male).

Based on a between-subjects design, they were allocated into one of four conditions.

Four conditions

<table>
<thead>
<tr>
<th>Human-Human</th>
<th>Human-Robot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face-to-face</td>
<td>Back-to-back</td>
</tr>
<tr>
<td>Rich mechanism</td>
<td>Poor mechanism</td>
</tr>
<tr>
<td>1. HH.EF</td>
<td>2. HH.BB</td>
</tr>
</tbody>
</table>

Tasks

1. Bodily Coordinated Motion Task

2. Moral Judgment Task

IV. Anticipated Results and Two Hypotheses

One hypothesis is that a subject will attribute a certain moral agency to the other (even to a robot) with whom the subject can bodily coordinate in a better manner. This is because the coordination involves the process of mutual understanding in some respects. Generally speaking, even with a new acquaintance of others, we naturally develop a concern for them. In parallel with that, we come to think that others should have a similar concern for us in turn. One can recognize a primitive basis for ethics in this situation.

Another hypothesis is that the richer world we recognize within others, the more demand for morality we make. On a simple setting of BCM Task, however, what is it like to recognize a richness -- or an inescapable realm which underlies personality -- within others? In our experiment, subjects may succeed in bodily coordination or fail. There are also conflicts as to which subject possesses the initiative. In the case of a conflict where the coordination once fails, a transition of the initiative eventually will take place, and a new coordination will hold, we presume. In the bodily coordination process with the other, you may have not only a sense of the opponent’s being in tune with yourself. You may also feel her/his resistance or the shift of the initiative to the other side. The experience of coordination can be a complicated one full of twists and turns.

We think that a subject will find a richness within his opponent through a complicated process of co-ordinations, divergences and transitions of the initiative. This process occurs when, for instance, the subject feels its opponent’s purposely making an unexpected move. In such a situation, it seems natural for us to attribute intention, desire, responsibility, and so on to others. This is when we recognize others as moral agents and accept them into the intersubjective world of morality.