

# Joint Attention Realized in a Robot with Intentional Agency

S. Nagataki<sup>1</sup>, M. Shibata<sup>2</sup>, T. Konno<sup>3</sup>, T. Hashimoto<sup>3</sup>, and H. Hattori<sup>4</sup>

<sup>1</sup> Chukyo University, 101-2, Yagoto Honmachi, Showa-ku, Nagoya, Aichi, Japan

<sup>2</sup> Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa, Japan

<sup>3</sup> JAIST, 1-1 Asahidai Nomi Ishikawa, Japan

<sup>4</sup> Nanzan University, 18 Yamazato-cho, Showa-ku, Nagoya, Aichi, Japan

## 1 Pre-linguistic communication and mind-reading system

Human beings have a distinctive ability to understand what is going on in others' minds. This ability, usually called "folk psychology," comprises a system of understanding others' intentions of varied levels. But how have we developed such an ability or mind-reading system? The most popular, though rough, answer is that it has been developed through evolution. Language, obviously one of our most useful tools to communicate with each other, is also acquired through this process. However, both we and our ancestors presumably depend on the common basic mechanism for communication underlying the linguistic interaction. That is, we deploy what might be called "primitive" parts of folk psychology for the purpose of communication.

The aim of our presentation is to show the basic functional structure of primitive folk psychology as a pre-linguistic tool. It is certainly not easy to explicate this dimension of mental function in an objective way because of its peculiar elusiveness. In order to cope with this inherent difficulty, we adopt a "constructive approach" to human communication, which attempts to construct a computational model of the phenomena in question, and to implement it into artificial systems, such as computers and robots, and then, by examining the behaviors of those systems, to grasp the mechanism in the part of the human communication system. Assuming that there are relevant similarities between natural and artificial systems, experimentally studying the behaviors of the latter can much contribute to the understanding of the former.

## 2 Making a robot with intentional agency

As an example of mind-reading function, we have selected "joint attention", one of the communicative behaviors already found in infants, in which they follow their parents' eyes and try to look at things in the same direction as their parents. It is known that this kind of behavior shows up as a reflective one, but then develops gradually into the cognitive states where infants understand and share others' attention. To share eyes direction with others seems to require a kind of "nested structure" of intentions and understandings among us through

which I understand that he/she understands that I intend to look at some object. In this regard, Tomasello suggests that infants become intentional agents before they can understand and share others' attention. As a result of our research, we could find how to construct a computational model of this process of becoming intentional agents and to implement it into a robot. Through this model, we can see how that robot interacts with humans to share their intentions to see something.

We believe that our computational model makes primitive intentional agency of infants possible by some association mechanism comprising the following two functions: the one is to form a memory of connections between directions of others' eyes and gazed objects, and the other is to look for a target object according to the memory.

If we discriminate 3 stages in this developmental process, our robot corresponds to the 3rd stage. This stage is an immature type of joint attention in that subjects intend to look at an object they recall through some association mechanism. If we briefly characterize the 1st and the 2nd stages here, the first is one where subjects look at objects through some reflexive mechanism, merely reacting to such external stimuli as sound or light, and the second where subjects look at objects in the same direction as their parents through another reflexive mechanism. We can verify that our robot attained an immature type of joint attention by confirming that it had a chance of looking for an associated object in many directions in case of failure.

### **3 From primitive intentional agency to understanding others' intentions**

Our robot in the 3rd stage does not intend to understand humans' intentions to see something. We may call this type of joint attention "joint attention with primitive intentional agency". Through making our robot, we have found that infants have to pass through two more stages, the 4th and the 5th, in order to share others' intention in a full-fledged sense. But, regrettably enough, we have not succeeded in constructing robots in those stages yet, so we could only show a conceptual sketch of the mechanism of our future robots at this time.

In the 4th stage, our robot can infer humans' intended objects (of course, not without any failure), using its inference mechanism which allows it to reach those objects from the past experiences. In other words, our robot will intend to understand humans' intentions and to make their intended objects its own desired objects. This may be called "joint attention with understanding of others' intentions". In the final stage of joint attention, i.e. the 5th stage, our robot will want to know that humans understand that it intends to share their intentions to look at some object. This means that when it knows by some assessment mechanism that its desired object does not coincide with humans' intended object, it will correct its inference and change its desire in order to make the two objects the same. This type of "joint attention with shared intentions" could realize a kind of nested structure of intentions and understandings. From making robots

in the 4th, and 5th stages, we would know what kind of further functions are required to construct a full-fledged mind-reading system, as the “constructive approach” tells us.