In: Proceedings of The Twelvth International Symposium on Artificial Life and Robotics (AROB07),
M. Sugisaka and H. Tanaka (Eds.), (2007)

Elementary Developmental Process of Intentional Agency: Artificial Construction of Gaze Alternation in Communicative Eye Gaze by Infants

Takeshi Konno
School of Knowledge Science,
Japan Advanced Institute of Science and Technology
1–1, Asahidai, Nomi, Ishikawa, 923–1292 Japan
t-konno@jaist.ac.jp

Takashi Hashimoto
School of Knowledge Science,
Japan Advanced Institute of Science and Technology
1–1, Asahidai, Nomi, Ishikawa, 923–1292 Japan
hash@jaist.ac.jp

Abstract

It has been pointed out that gaze alternation by infants, which is a basis for social communication, is related to the process of the development of intentional agency. Intentional agency is defined as an act with a desired goal and a means. It has been pointed out that infants understand others' intentions based on intentional agency. In our recent work, we constructed a computational model which acquired gaze alternation behavior with elementary intentional agency. In this paper, through the analysis of the behavior and the internal states of the constructed model, we confirm that two mechanisms, discrimination and accumulation, play an important role in developing elementary intentional agency. The former discriminates between a caregiver and objects that are producing stimuli. The latter associates the caregiver with the objects by accumulating relationships between sensory states. We also argue that a nested structure, in which gaze alternation composed of a goal and a means is utilized as a means to achieve another goal, is an important feature in the development of intentional agency.

Index Terms — **Gaze Alternation, Communicative Eye Gaze, Understanding Others' Intention, Intentional Agency, Constructive Approach.**

1 Introduction

We study the gaze alternation behavior in communicative eye gaze by infants through the construction of a computational model. Gaze alternation is behavior by infants alternately gazing at a caregiver and particular objects. Tomasello [1] points out that gaze alternation, which is a basis of social communication, is related to the developmental process of intentional agency. Intentional agency is defined as an act composed of a desired goal and a means. It has been pointed out that the infants understand the intentions of others based on intentional agency. From this viewpoint, gaze alternation is not merely behavior returning an infant's gaze point to a caregiver after looking at particular objects, but intentional behavior of gazing at a caregiver based on an infant's desire. This intentional behavior can be developed into more a communicative use of eye gaze – i.e., social referencing and the utilization of gaze alternation.

In our recent work, we constructed a computational model [2] in which acquiring the visual orientation of gazing at a caregiver or at objects in the center of the visual field became a means by reinforcement learning and in which acquiring the gaze alternation between a caregiver and objects became a goal through memorizing sensory states. However we have not yet clearly shown the mechanisms by which intentional agency is acquired. In this paper, in order to study the elementary developmental process of intentional agency, we investigate the mechanisms in the constructed model. we analyze the behavior and internal states of this model, by which goals are separated from actions in intentional agency.

2 Model

The agent model of an infant (the infant agent, hereafter) has the functions of visual orientation and gaze alternation. In this section, after explaining the agent's visual field, we describe the functions.

2.1 Visual Field of Infant Agent

The infant agent's visual field is a square area that is 0.2m on each side on a flat surface. The sensory state of the infant agent is defined by three kinds of information from the visual field and its motion (Fig.1): (a) the feature of a visual stimulus, which is one of 13 stimuli composed of the caregiver's eye directions (a total of 10) and the type of objects (3 shapes), (b) the position of the visual stimulus consisting of the direction and the gaze, and (c) the proprioception of the muscle states that are related to the orientation of the gaze point. We suppose that the muscle states are integrated to represent the direction of the gaze point.

2.2 Visual Orientation

The visual orientation is the ability to gaze at the caregiver and the objects in the center of the visual field and consists of three modules: the selector, the

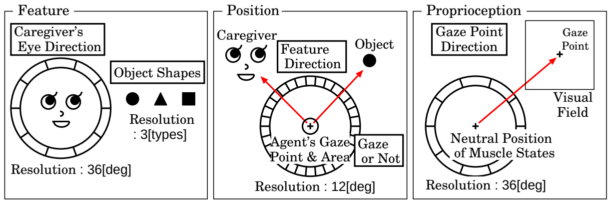

Figure 1: State information about the feature, position, and proprioception in the visual field. In the position, a state of gaze is determined by whether or not the stimulus comes within the gaze area. The gaze area is created to be small circle from the gaze point.

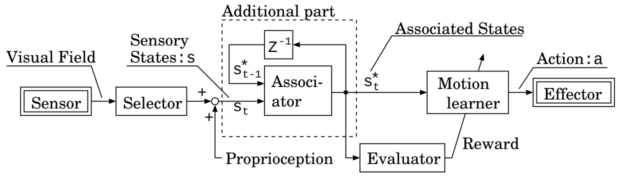

Figure 2: System block diagram of gaze alternation.

evaluator, and the motion learner, shown in Fig.2 without the additional part (which is surrounded with a dashed line). The selector picks up the most distant stimulus from the gaze point in the visual field, an object or a caregiver; then it outputs information about the feature and the position of the selected stimulus. A sensory state, s, is constructed by merging three kinds of sensory information: the feature, position, and proprioception. The sensory state is related to an action to move the gaze point[1] by reinforcement learning algorithm which is known as a standard temporal difference learning with tabular SARSA [3]. In this algorithm, the action is determined according to a probability with an action-value function. The motion learner reinforces the action which allows the gaze point to approach a target[2].

2.3 Gaze Alternation

The gaze alternation system is equipped with an associator in addition to the visual orientation system (Fig.2). The associator stores relationships between three sensory states: the present sensory states, s_t; previous associated states, s^*_{t-1}; and future sensory states, s_{t+1} in a frequency distribution. The frequency distribution is updated according to the sensory states.

$$F(s_t, s^*_{t-1}, S) = F(s_t, s^*_{t-1}, S) + 1 \ , \qquad (1)$$

where S is conditioned by

$$S = \begin{cases} s_t & \text{if} \quad cond.A \ , \\ s^*_{t-1} & \text{if} \quad cond.B \ , \\ s_{t+1} & \text{if} \quad cond.C \ . \end{cases} \qquad (2)$$

[1]The moving direction of the gaze point has 30 divisions.
[2]For more details, see paper [2].

Table 1: Storing condition in the frequency distribution.

cond.				C(s^*_{t-1}): Φ		C(s^*_{t-1}): CGV		C(s^*_{t-1}): OBJ	
			G(s^*_{t-1})	0	1	0	1	0	1
C(s_t)	Φ	G(s_t)	0	A	A	B	B	B	B
			1	A	A	B	B	B	B
	CGV		0	A	A	A	A	B	A
			1	A	A	A	C	B	A
	OBJ		0	A	A	B	A	A	A
			1	A	A	B	A	A	C

The condition, $cond.A, B, C$, is defined by Table.1, where $C(s)$ judges the feature, the caregiver (CGV), the object (OBJ) or non-existence (ϕ), and $G(s)$ determines whether the sensory state is a gaze (1) or not (0). The output of the associator, s^*_t, is determined according to a probability:

$$p(s|s_t, s^*_{t-1}) = \frac{F(s_t, s^*_{t-1}, s)}{\sum_{s'=1}^{N_c} F(s_t, s^*_{t-1}, s')} \ , \qquad (3)$$

where N_c is the total number of the sensory states.

3 Experiments and Results

3.1 Experimental Setups

At first, we display the caregiver and the objects in the agent's visual field in its visual orientation system. By reinforcement learning, the agent develops a probability distribution of the action-value function, shown in Fig.3. The agent acquires the ability of visual orientation because the probability distribution of the moving gaze point develops to match the direction of the caregiver and the objects.

Second, in a training phase, we display the caregiver and the objects alternately in the agent's visual field

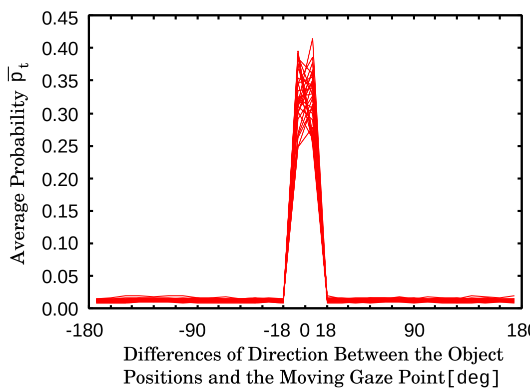

Figure 3: The average probability of the moving gaze points in the positions of the caregiver and the objects.

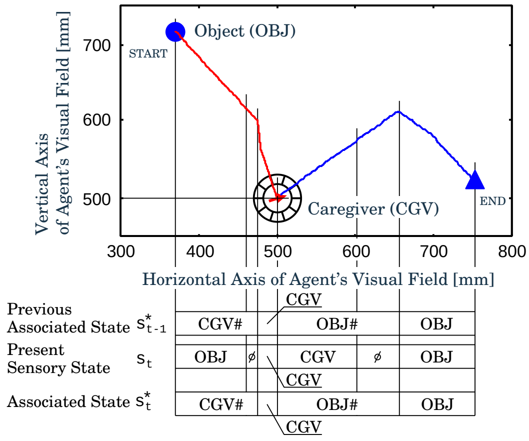

Figure 4: For the process of gaze alternation between a caregiver and objects which are placed outside the visual field, the corresponding internal states of the associator are indicated. The mark '#' means the state does not coincide with the present sensory state.

in the gaze alternation system (Fig.2). In this training phase, the objects are placed in the direction the caregiver looks. After adequate training, we display the objects outside the visual field. As a result, the agent can acquire a gaze alternation behavior outside the visual field.

3.2 The Process of Gaze Alternation

The trajectory of the agent's gaze point with the corresponding internal states of the associator in the gaze alternation behavior is exemplified by Fig.4. In the behavior toward the caregiver from the circular object which is placed on the left side of Fig.4, when the present sensory state, s_t, is an object (OBJ) or non-existence (ϕ), the associated state, s_t^*, is the previous sensory state, s_{t-1}^* which indicates the caregiver (CGV). If the present sensory state is the caregiver – specifically, if the caregiver appears in the visual field – the associated state coincides with the present sensory state, and the agent adjusts the direction to move its gaze point. When the agent gazes at the caregiver, the associated state becomes the future sensory state, s_{t+1}, and designates a new object (OBJ), which does not appear in the visual field. The new object is associated with the caregiver, based on the frequency distribution [Eq.(3)].

In behavior to look at a new object, which is the triangular object placed on the right side of Fig.4, while the present sensory state is the caregiver (CGV) or non-existence (ϕ), the associated state is the same as the previous associated state, indicating the object (OBJ) to be seen. At the point where the object comes into the visual field – that is, when the present sensory state turns into the object – the associated state indicates the object. The agent can modify the

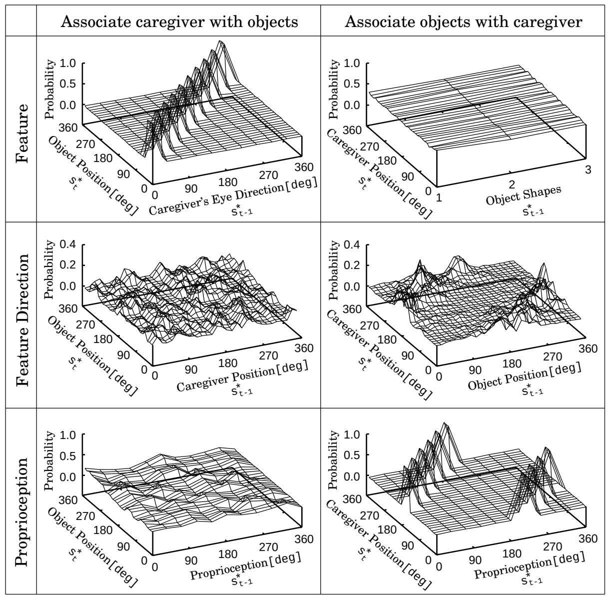

Figure 5: Association probability between the previous associated states, s_{t-1}^*, and the associated state, s_t^*.

trajectory of the moving gaze point even if the previous direction has been considerably different from that of the object to be seen. Through the transition of internal states of the associator, the agent realizes the gaze alternation behavior outside the visual field.

When the present sensory state and the previous associated state are different features, the associator keeps the previous associated state because the agent accumulates the sensory states in the frequency distribution according to *cond.B* in Eq.(2), which is based on discrimination about the category (that is, the caregiver, the object, or non-existence) using function $C(s)$ in Table.1. If the present sensory state and the previous associated state become the same feature, the associator changes to the present sensory state because the frequency distribution has accumulated according to *cond.A* in Eq.(2). When the agent gazes at the caregiver or the objects, the associator associates the caregiver with the objects using the frequency distribution that is accumulating the sensory states according to *cond.C* in Eq.(2).

3.3 Association Probability Distribution

In the association of the caregiver (CGV) with the object (OBJ), the agent uses the probability distribution shown in the left column of Fig.5. In the top left of Fig.5, the association probability distribution between the caregiver's eye directions and object positions is shown. Using this distribution, the agent can detect the unique object position from the caregiver's eye direction since one-to-one correspondence between these two pieces of information is formed in this distribution. The distributions of the other sensory in-

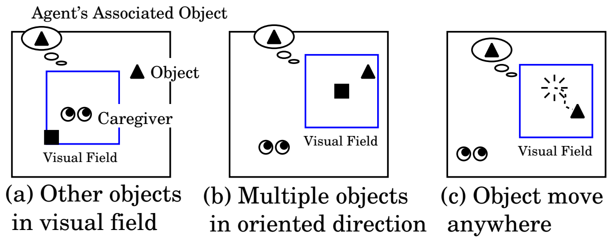

Figure 6: Showing the intentional agency.

formation, the caregiver position and proprioception, do not form such a one-to-one structure and, thus, are unable to be used to detect the object position.

On the other hand, when the agent associates the object with the caregiver, the agent can determine the direction of the caregiver using the state of the proprioception since only the frequency distribution between the proprioception and the caregiver position possesses the one-to-one correspondence structure, as shown in the right column of Fig.5. In sum, we find that the infant agent makes use of the caregiver's eye direction to gaze at the objects and uses the proprioception to return its gaze to the caregiver.

4 Discussion

4.1 Development of Intentional Agency

We confirmed that the constructed agent model can acquire gaze alternation behavior outside the visual field. In the behavior of gaze alternation depicted in Fig.4, three functional properties in the internal states of the associator are identified: **1.** The agent can keep the associated object when different objects appear in its visual field. Therefore the agent can properly gaze at the object that is indicated by the caregiver shown in Fig.6 (a). **2.** For the same reason, if two or more objects are placed in the caregiver's eye direction, as in Fig.6 (b), the agent can gaze at the associated object. **3.** Once the associated object appears in the visual field, the agent can adapt the moving direction of its gaze point to the direction of the associated object. Therefore, even though the object may move anywhere, the agent can follow the object, as shown in Fig.6 (c).

In these functional properties, the associated states play the role of the goal at which the agent wants to gaze, and the agent through its visual orientation ability selects the action of gazing at the caregiver and the objects as a means of achieving the goal. The three functional properties of gaze alternation suggest that the agent acquires elementary intentional agency.

4.2 The Nested Structure of Intentional Agency

It is known that infants tell their desire to others to get an object by utilizing gaze alternation, and often others hand over the object to the infants. Accordingly, gaze alternation is thought of as a means of achieving the desired goal of getting the object. In this behavior, we can discover a nested structure in which the behavior composed of a goal and a means is utilized as a means to achieve another goal. This behavior consists of a goal to get the object and a means of alternatively gazing at the caregiver and at the object. Furthermore, the means consists of the goal of gazing something which the infants want to see and the means of moving their gaze point.

The constructed model of gaze alternation has been composed of the goals of gazing at the caregiver or the objects and the means of achieving those goals. When we develop the agent model to perform utilizing gaze alternation based on the present model, the nested structure must be observed.

5 Conclusion

In this paper, in order to study the elementary developmental process of intentional agency, we investigate how the infant agent of our constructed model acquires intentional agency that involves separating goals from actions. Through the analysis of the gaze alternation behavior and the internal states, we confirm that two mechanisms – discrimination and accumulation – play an important role in acquiring intentional agency. The former discriminates within the category – that is, between the caregiver, the object, and non-existence – in the visual stimuli and determines whether or not the agent gazes at the stimulus. The latter accumulates the relationships between the continual sensory states in a frequency distribution. Through these two mechanisms, the agent acquires the goals – that is, objects or a caregiver – at which the agent wants to gaze and that are to be realized by the means of moving its gaze point with its visual orientation. Through discussion of the developmental process of intentional agency, we suggest that a nested structure, in which the gaze alternation composed of a goal and a means is utilized as a means to achieve another goal, is an important feature in the developmental process of intentional agency.

References

[1] Tomasello,M., "Joint attention as social cognition," In Moore,C. & Dunham,P.J.(Eds.), Joint Attention: Its Origins and Role in Development, Lawrence Erlbaum, pp.103-130,(1995).

[2] Konno,T. and Hashimoto,T., "Developmental Construction of Intentional Agency in Communicative Eye Gaze," Proceedings of the International Conference on Development and Learning, ICDL '06, Indiana Univ., Bloomington, USA, May 31 - June 3,(2006).

[3] Sutton,R.S. and Barto,A.G., Reinforcement Learning, A Bradford Book, MIT Press, cambridge, MA,(1998).