Proceedings of the 14th Biennial
International Conference on Infant Studies
May 2004

How does an infant acquire the joint attention ability?: A Constructivist Approach

Yukie Nagai*, Koh Hosoda*,†, and Minoru Asada*,†
*Osaka University, †Handai Frontier Research Center

This study explores how a human infant can acquire the ability of joint attention through interactions with its caregiver, from a constructive robotics approach. In our constructive model, a robot learns a sensorimotor coordination for joint attention based on visual attention social input, and "self-evaluative" feedback. Because visual attention does not always correspond to joint attention, the robot has incorrect learning situations as well as correct ones. However, the robot is expected to statistically "unlearn" the data of incorrect situations as outliers through the learning algorithm, and consequently acquire the ability of joint attention, even if the environment is not controlled nor the caregiver provides any task evaluation.

The learning module was implemented in a robot with two degrees of freedom in its eye cameras. It consists of a three-layered neural network. In forward processing, this module receives the image of a human "caregiver's" face and the angle of its camera head as inputs, and outputs a motor command vector. The caregiver's face image is used to estimate the motor command to follow the caregiver's gaze direction. The generated motor command is sent as output to the learning module. In the learning process, this module learns sensorimotor coordination by back propagation when triggered by the internal evaluator. To train the robot, in learning trials, multiple salient objects were presented at random positions, and the human caregiver attended to one of the objects. This served as input to the robot.

The experimental results suggest that the proposed model could explain the developmental progression of the infant's joint attention. For example, in the beginning of learning, the robot had a tendency to attend to an interesting object in its field of vision. By the final stage, the robot acquired a "mature" ability of joint attention. It could identify the object that the caregiver attends to, and produce a motor response to look at the object, whether or not that object was first in the robot's visual field. This progression is analogous to data on human infants described by Butterworth and Jarrett (1991). Thus, in some regards the robot's joint attention behaviors were similar to those seen in infant development. Ongoing work is seeking to refine how the learning model combines information from different sources, to better understand the mechanism of infant development.