

A developmental model of initiating joint attention through constructing state space

Tsukasa Nakano*, Shinya Fujiki*, Yuichiro Yoshikawa[†] and Minoru Asada*

*Dept. of adaptive Machine Systems, Graduate School of Engineering, Osaka University

2-1 Ymadaoka, Suita, Osaka, Japan 565-0871

Email: {tsukasa.nakano, shinya.fujiki, asada}@ams.eng.osaka-u.ac.jp

[†]Dept. of Systems Innovation, Graduate School of Engineering, Osaka University

1-3 Machikaneyama, Toyonaka, Osaka, Japan 560-8531

Email: {yoshikawa}@sys.es.osaka-u.ac.jp

I. INTRODUCTION

Inferring other's intention is an important factor that can accelerate the development of joint attention (JA). In this paper, we aim to reveal the developmental mechanism of JA based on such inferring capability by a synthetic approach.

Human infants have gradually come to follow the gaze direction of caregivers from 6-months of age[1]. Until 12-month, they have accurately been able to find an object that her gaze and/or pointing are referring to[2], [3]. These behaviors are called responding to JA (RJA). Meanwhile, they have developed the ability to lead her attentions to an object by using their pointings, utterances, etc[2]. These actions are called initiative behaviors of JA (IJA). However, the internal mechanism which enables such a development of both RJA and IJA has not been well understood.

On the other hand, some studies have synthetically suggested the developmental models of JA in human infants as advocated in cognitive developmental robotics[4]. Those showed that a robot acquires gaze-following only through statically mapping gaze and locations when it finds something salient[5], [6], and that a robot acquires gaze-alternation in JA through discovering and reproducing the contingency of caregiver-infant interactions based on information theory[7]. However, those have not dealt with the development of inferring caregiver's internal states, nor explain about the relationship between it and JA.

This work presents our working model for a learning mechanism of optimal actions for JA in a scheme of reinforcement learning (RL). It integrates methods for the autonomous state space construction and the hidden state estimation. It seems an important but formidable step for infants to estimate the caregiver's hidden state, whether she is interactive for JA or not, since it can not be directly observable from them but largely affects their future rewards. To overcome this, the proposed method has three steps. The first step is to divide state space, with consideration of the differences between the caregiver's hidden states, based on the extended method of ishiguro et al.[8]. The second step is to learn the probabilistic model of state transition based on HMM. Through above two steps, the learner can acquire the ability to estimate the hidden states according to not only the directed observation at the time but also the history of their interactions. The final step is to learn the optimal action on each state based on RL. We applied the proposed model to the computer simulation, and try to reproduce the developmental processes of both RJA and IJA in human infants.

II. LEARNING MODEL OF JA

Here, a certain type of caregiver-infant interaction was supposed. The learner first observes the caregiver and the objects in their environment, estimates her attentions and the hidden states. Then it selects one from possible actions: lookings at the caregiver, looking at an object, and utterance. If the JA between them is realized, the caregiver gives a reward. After experiencing those interactions, the learner constructs the state space, and learn the state transition model and optimal actions on each state. By repeating the learnings and the interactions, the it can gradually understand her hidden states, and acquire the behaviors of RJA and IJA.

In this work, we assume that the caregiver's hidden states are three kinds of the modes: *initiative mode*, *responsive mode*, and *non-interactive mode*. These modes affect her policies of giving rewards and state transition. When she is the initiative mode or the responsive mode, the learner can get rewards by JA. On the other hand, it can't get rewards when she is the non-interactive mode, even if it can follow her gaze directions. However, the learner can let her mode change to the responsive one by its particular behaviors: gaze-alternations and speakings to her. Therefore, it should induce her responsive mode in order to get rewards when she is non-interactive mode.

A. System for JA

The system has three modules: *state estimation module*, *action selection module*, and *memory module*. The learner infers the caregiver's modes by state estimation module, selects an action to get rewards by action selection module, and stores the interaction data for learnings in memory module.

The state estimation module is composed of two parts that are *sensor classifier* and *state transition model*. The sensor classifier calculates the current state probabilities from observable sensor data: gaze directions of the caregiver's and the learner's, and locations of two objects. Meanwhile, the state transition model also calculates those probabilities from the prior data: Those are the state probabilities, The learner's selected action and the caregiver's reward value at the last interaction. The learner estimates the current state by integrating the state probabilities based on two different way. Note that the way to give rewards is different depending on the caregiver's mode, the learner's action, and the current state. Therefore, the state transition model has very important role for estimating her current mode. By this method, the learner identifies the current state according to her modes.

In the action selection module, an action is stochastically selected from the identified state and the state-action value. The state-action value based on RL means the probabilities that the action enables the learner to get rewards on the states. By this method, the learner selects an action that is seemed to be optimal.

B. Learning algorithm

The learner needs to recognize a variety of states to achieve JA depending on the caregiver's mode which change following to the infant's behavior. In order to discriminate these modes which are not observable at the moment, the system has two stages: State space construction for RL to find adequate features based on an extension method of ishiguro et al. [8] and the parameter estimation of the state transition model based on HMM.

At the state space construction stage, the system first divides the continuous sensor space into two kinds of areas after each action, according to whether the learner is given a reward in the area or not. Moreover, the divided area may have more division if it has multiple peaks in the histogram of reward values depending on the different modes. As a result, the learner constructs the state space which is expected to reflect the differences among the modes.

At the parameter estimation stage of the state transition model, the system re-estimates the states from the past interactions. Based on the HMM, it calculates the probabilities of the state transition in terms of the learner's prior action, the reward given, and the previous state. The learner consequently understand how the states and modes change depending on its own behavior.

Moreover, the learner needs to learn adequate behaviors to achieve JA in each the caregiver's mode. In action learning, the system calculates state-action values by Q-learning of RL. The merit of RL can be to consider the near future rewards. Therefore, the learner can select behaviors not only directly but also indirectly linked to rewards.

III. SIMULATION OF DEVELOPMENT OF IJA

As the results of computer simulations, there were the differences of the selection probabilities of the learner's action on each modes, after that the learner acquired the states with consideration of the caregiver's modes. Before learning her modes, the learner tended to select gaze-following in any caregiver's mode. This was because it was learned that the learner has possibilities to get rewards by gaze-following although it could not always get rewards by the behavior. However, after learning her modes, it came to select gaze-alternation and speaking to her more than gaze-following when she is the non-interactive mode. This was because it was learned that the learner should induce her mode to the responsive one in order to get rewards.

IV. CONCLUSION

In this work, we proposed the method that can acquire optimal action based on reinforcement learning, which integrates methods for state space construction and hidden state estimation. Results of computer simulation showed that the proposed method can learn not only gaze-following but also initiative behavior of joint attention by estimating the caregiver's hidden states. Those results indicated the possibility that the learning method reproduced one of the developmental processes of joint attention affected by understanding other's intention in human infants.

The caregiver's model in this work was a specific setting. Therefore, the future work will investigate what kinds of parameters in caregiver models are the important factors for the behavior acquisition of joint attention in human infant.

ACKNOWLEDGMENT

This study is partially supported by Grants-in-Aid for Scientific Research (Research Project Number: 22220002) and Grant-in-Aid for JSPS Fellows(Research Project Number:21.57061).

REFERENCES

- [1] Butterworth and Jarrett. What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Dev. Psy.*, 9:55-72, 1991.
- [2] Carpenter et al. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the society for research in child Dev.*, 63:, 1998.
- [3] Amanda L. Woodward. Infants' developing understanding of the link between looker and object. *Developmental Science*, 6:297-311, 2003.
- [4] Asada et al. Cognitive developmental robotics: a survey. *IEEE Transactions on Autonomous Mental Development*, 1:12-34, 2009.
- [5] Nagai et al. A constructive model for the development of joint attention. *Connection Science*, 15:211-229 2003.
- [6] Triesch et al. Gaze following : why (not) learn it. *Develpmental Science*, 9:125-147, 2006.
- [7] Sumioka et al. Development of Joint Attention Related Actions Based on Reproducing Interaction Contingency. *Proc. of the 7th ICDL*, 2007.
- [8] Ishiguro et al. Robot Oriented State Space Construction. *In Proc. of the IEEE/RSJ Int'l Conf. on Intelligent Robots and System*, 3:1496-1501, 1996.