The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Co-creation of Human-Robot Interaction Rules through Response Prediction and Habituation/Dishabituation

Takatsugu Kuriyama and Yasuo Kuniyoshi

***Abstract*— A joint learning approach is described that meets a major challenge with social robots — developing a methodology for learning communicative behaviors. We focus on interaction rule that is relationship between a robot's action and a partner's response. In this approach a robot is simultaneously a learner and proposer of interaction rules. The human partner and robot continuously search for and co-create new rules as inspired by the social games played between an infant and a caregiver. A simple and universal scheme with response prediction and habituation/dishabituation was developed, and a robot model was built using the scheme. The robot generates actions, observes the partner's response, and get to predict them. It identifies relationships between its actions and the responses, and generates actions designed to elicit particular responses from the partner. After it is habituated to the responses, it generates other actions to search for other rules. In experiments of human-robot interaction based on this model and using a ball, different patterns of interaction emerged, such as passing the ball back and forth, rolling and catching, and feint passing. Response prediction and appropriate habituation supported the emergence of interactions, indicating that the scheme and the model are effective. This joint learning should lead to natural communication between human partners and social robots beyond teach/taught relationship.**

I. Introduction

A. Background

Social robots, such as pet, elderly-care, child-care, museum tour-guide, and alternative remote, are now being introduced into daily life environments. For example, AIBO is a dog-shaped commercial robot for pet [1]. It has abilities of image and voice recognition, and can chase after a colored ball and respond to voice commands. Paro is a baby-seal-shaped robot for mental commitment, and is used at elderly day homes and hospitals [2]. It has touch sensors and ability of voice recognition, and can feel being held and recognize greeting voice.

If action and response of social robots are all preprogrammed, we cannot expand the relationship and get bored as a long-term companion. AIBO has parameters of playing preference and degree of growth. These parameters have been adjusted and the behavior has changed through interaction. The scenarios of the changes are, however, defined in advance. The process of growth is a trace and choices of the scenarios. Paro identifies its name as the name is repeatedly called. The response to the calling is, however, hard-coded.

T. Kuriyama and Y. Kuniyoshi are with Department of Mechano-Informatics, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, JAPAN 113-8656. They are also with JST ERATO Asada Synergistic Intelligence Project. {kuriyama, kuniyosh}@isi.imi.i.u-tokyo.ac.jp

In human-robot interaction, it is a major challenge with social robots to develop a methodology for learning communicative behaviors for sustainable relationship.

B. Related Work

Several researches introduce further learning ability into robots to adjust and acquire social relationship. Mitsunaga et al. utilized reinforcement learning for adjusting interaction distance, gaze meeting, motion speed, and timing in interaction, through watching minute body signals of a partner [3]. Breazeal et al. made a robot that recognizes and identifies correspondence between perceived feature and its own, and realize facial imitation [4]. Ishihara et al. focused on sensorimotor magnetic effect in which a partner unconsciously acts in one's familiar way when the partner imitates a robot [5]. The robot finally acquired vowel through vocal mutual imitation. Taniguchi et al. focused on scheme of role reversal imitation [6]. At the first learning phase, the robot observes partner's demonstration. At the second learning phase, the robot demonstrates actions, and the partner gives a corresponding word to the robot, then the robot finds the correspondences between the actions and the words. At the recognition phase, the robot presents a word corresponding to the partner's demonstration.

As stated above, conventional approaches are basically based on teaching, and also suppose convergence of the relationship. In the real relationship between humans, however, social relationship is between stable and fluctuating, and also not one-directional but co-creative. Both are simultaneously a learner and proposer, and search for a good relationship for each other. Conventional ways, such as teach/taught paradigm and optimization for a single state, have limitation to meet human.

C. Goals and Approach

We aim to realize a robot which continues to search for and co-create relationships with a human partner. Human can build social relationships with human, so the process where human gets to communicate is considered to serve as a useful reference for building a social robot. We focus on *social games* such as peek-a-boo, ball game, give and take, gonna get you, point and name and so on [7]. They are considered to be proto-communication, and are played between a caregiver and an infant in the first two years [8].

Bruner et al. pointed that social games are consisted of rules [8][9]. Stern pointed that *Infant-elicited social behaviors* plays a significant role in social games [10]. We focus on interaction rule that is relationship between a robot's action

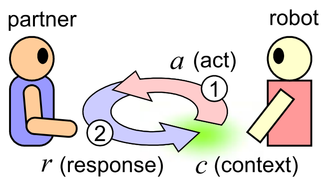

Fig. 1. Interaction rule: Correspondence between a robot's action in a context and a partner's response

in a context and a partner's response (**Fig.1**). A partner's social cue to a robot is included as a context of the rule.

We previously proposed a scheme and a robot model where a robot learns interaction rule through imitation [11]. While a partner demonstrates, a robot imitates the partner. While a partner responds to a robot, the robot is motivated to elicit the partner's response. The robot adaptively identifies communicative actions in imitated numbers of actions and responses. It is, however, teach/taught relationship.

In this paper, we propose a scheme and a robot model where a partner and a robot are simultaneously a learner and proposer, and continue to search for and co-create interaction rules.

II. Scheme for Co-creation of Rules

A. Development of Social Games

Development of social games progresses through the four stages described below [9][12]. An infant gradually moves from a passive role to an active one over course of playing.

1) observing passively: The infant merely observes the caregiver passively; the caregiver may physically assist the infant to play.
2) taking part in one of the game's elements: The infant takes part in one of game's elements and eventually grows to initiate more of elements.
3) sharing of the game's activities: Each player takes a turn in a well-organized fashion based on the convention of the game.
4) generating modifications: The infant generates variations within the rules of the game. The infant has a sufficient understanding of the game's rule structure to be able to add new rules.

B. Proposed Scheme

We reinterpret the developmental stages of social games for co-creation of rules. In our proposed scheme, a partner and a robot co-create rules within an interactional sequence. These phases are not clearly separated or repeated in the sequence.

1) The robot simply tries various actions. The partner waits for or assists the robot's interpretable action. The interaction is not incorporated into the rules.
2) The robot performs an interpretable action by chance, and the partner readily responds to the action. The robot has thus got to predict the response and found a rule. This is "making response prediction", and the way a rule is co-created.

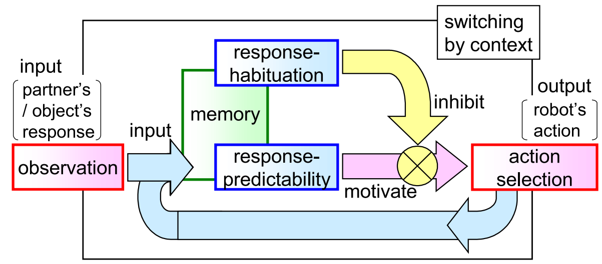

Fig. 2. The proposed robot model with the two indices, response-predictability $ruled(c, a)$ and response-habituation $habit(c, a)$

3) The robot repeats the action and, if the partner's response is in accordance with the rule, the robot understands that the rule is reliable (i.e., the partner's response is not by chance). This is "confirming response prediction". The partner's response is then a social cue for the robot. Turn-taking occurs in accordance with the rule, and the interaction is reciprocated as a result.
4) The robot inhibits and avoids the confirmation of well-confirmed rules, and instead it performs actions for which the corresponding response is not well known. As a result, the robot generates modifications of the interaction, and creates room for new rules to be found. The result is co-creation of new rules. This is "habituation/dishabituation of predicted response"

III. Robot Model for Co-creation of Rules

A. Model Constitution

We propose a robot model for co-creation of rules, based on the scheme through "making response prediction", "confirming response prediction", and "habituation/dishabituation of predicted response".

We describe a as a robot's action, r as a corresponding partner's response, c as a context which is situation when the robot begins to act. We define two indices described below.

- response-predictability $ruled(c, a)$: degree of how much partner's responses are ruled (predictable) with a robot's action a under a context c.
- response-habituation $habit(c, a)$: degree of how often partner's responses were correctly predicted with a robot's action a under a context c.

The model is illustrated in **Fig.2**. Robot observes at an appropriate timing after action. The robot acts and observes in the following way.

- acting: The robot calculates $ruled(c, a)$ according to past observed partner's responses, and acts a in which $ruled(c, a)$ is large and $habit(c, a)$ is small.
- observing: The robot memories (c, a, r) and updates $habit(c, a)$.

B. Expected Behavior of the Model

We describe how the proposed model works. In **Fig.3**, the left is the phases of the scheme, and the right is corresponding states of the two indices on the act space. The red frame and the blue frame are the same act space. In

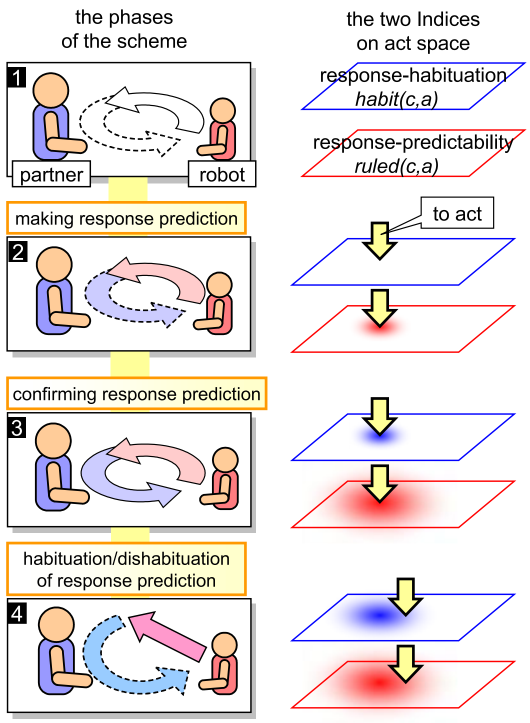

Fig. 3. The phases and the scheme of co-creation of rules, and corresponding states of the two indices on the act space.

the red frame, the red spread is the area where $ruled(c, a)$ is large on the act space. In the blue frame, the blue spread is the area where $habit(c, a)$ is large on the act space. The yellow arrow points to the action a to be selected. On a, $ruled(c, a)$ is large and $habit(c, a)$ is small. We describe how the two indices correspond to the phases.

1) At beginning of interaction, there is no action to which partner's response is predicted or habituated. The robot just tries some actions.
2) The robot performs an interpretable action by chance, and the partner readily responds to the action. Then corresponding $ruled(c, a)$ increases on the act space. This is "making response prediction".
3) The action a in the rule is selected to act because $ruled(c, a)$ is large. As a result, the robot acts to elicit the partner's predicted response. Through this repetition, the robot understands the rule is reliable, not by chance. This is "confirming response prediction". Then $ruled(c, a)$ increases and $habit(c, a)$ does due to response prediction.
4) $habit(c, a)$ of the confirming action is now large enough, then the action is inhibited and the robot gets to perform an action that is a different from the confirming one. As a result, the robot seems to search for another rule. This transition of focused response from well-known to unknown is "habituation/dishabituation of response prediction."

In this way, with $ruled(c, a)$ and $habit(c, a)$, the robot follows "making response prediction", "confirming response prediction", and "habituation/dishabituation of response prediction", and the partner-robot dyad co-creates rules.

C. Mathematical Formulation of the Model

We mathematically formulate the conceptual model. The robot selects an action of which response-predictability $ruled(c, a)$ is large and response-habituation $habit(c, a)$ is small. We define degree, that is how large $ruled(c, a)$ is and how small $habit(c, a)$ is, as motivation $motiv(c, a)$ in (1). The robot selects an action that maximizes $motiv(c, a)$.

$$motiv(c, a) = ruled(c, a) \times e^{-habit(c,a)} \ . \qquad (1)$$

$ruled(c, a)$ is calculated as (2).

$$ruled(c, a) = I(c, a; R)\ trust(c, a) \ . \qquad (2)$$

where $I(c, a; R)$ is correlation between a and all the partner's responses R. If $I(c, a; R)$ is large, partner's response is ruled by and predicted from a. It is calculated as (3).

$$I(c, a; R) = \int_{r \in R} \frac{p(c, a, r)}{p(c, a)} log \left(\frac{p(c, a, r)}{p(c, a)p(r)} \right) dr \ . \qquad (3)$$

This is a deformation of the Mutual Information [13]. In (3), there is time gap between the time the robot acts and the time the robot observes a partner's response, thus this is equivalent to simplified the Transfer Entropy which is a causality measure [14]. $I(c, a; R)$ of unexperienced (c, a) is over-estimated. $trust(c, a)$ is degree of how often the robot experienced (c, a), and it is a constraint function for domain of $I(c, a; R)$. $trust(c, a)$ is calculated as (4).

$$trust(c, a) = (p(c, a))^{\beta} \ , \qquad 0 < \beta < 1 \ . \qquad (4)$$

When the robot observes a partner's response, $habit(c, a)$ is updated as (5).

$$\Delta habit(c, a) = h(H(R \,|\, c, a) - I(r \,|\, c, a)) \ . \qquad (5)$$

Initial value of $habit(c, a)$ is 0, and constrained to be greater or equal to 0, not to be over dishabituated. (5) judges whether an observed response is predicted or not. $I(r|c, a)$ is information of a present observed response, and $H(R|c, a)$ is average of information (entropy) of a response. (5) compares these. If $I(r|c, a)$ is smaller, the response is same as prediction. If $I(r|c, a)$ is larger, the response is out of prediction. $H(R|c, a)$ is the baseline for evaluation of $I(r|c, a)$. $h(x)$ is a translation function from information to habituation. $h(0) = 0$ and $h(x)$ should be monotone increasing.

IV. Experiment

We test that, in an interactional sequence, the partner and the robot can co-create and reciprocate interaction rules which are not defined in advance.

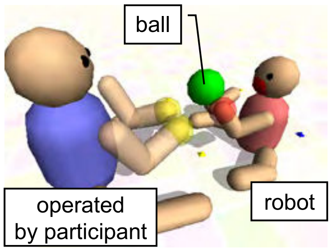

Fig. 4. Simulation environment

Fig. 5. Game pad

Fig. 6. Grasping ball

Fig. 7. Grasping

A. Setup

The task is a free play with a ball because it is a simple case of triadic interaction and the partner can easily interpret robot's actions in the play. Time when the robot acts and observes is discrete. The robot acts at every second. The robot observes when a target (the partner or the ball) turns or stops. The robot also observes when one second elapsed after a previous observation.

We experiment with a simulation environment (**Fig.4**). The left is the agent that a participant operates, and the right is the robot. A participant operates the agent with DUALSHOCK3 (**Fig.5**) which is the game pad of SONY PlayStation3 and expected to be familiar with a participant. In the environment, the physical simulation (Open Dynamics Engine) is introduced so that a participant and the robot can utilize physical phenomena (such as rolling ball) as communicative actions.

For simplicity, position of their hand and the ball are constrained to the sagittal plane. Posture of the two is consisted of hand position (2 dimensions) and degree of hand grasping (1 dimension), and each dimension is normalized within 0 to 1. Robot's action is defined as posture at one second after the robot begins to act. Context c and response r are continuous value. Action a is discrete value as (0.1, 0.3, 0.5, 0.7, 0.9) due to computational cost.

Hand position of the agent is controlled with the 2-axes analog joystick, indicated as "for hand reaching" at **Fig.5**. Hand grasping of the agent is controlled with the button, indicated as "for grasping" at **Fig.5**. While the button is pressed, degree of grasping is increasing to 0.8. While it is not, the degree is decreasing to 0.2. Hand of the two is colored yellow if degree of grasping is less than 0.5. When the degree crosses over 0.5, hand is colored red and the ball is grasped as **Fig.6** if the ball is touched. Hand is colored orange as **Fig.7** if not. With this indication, a participant can see whether the agent and the robot grasp the ball or not.

We utilize the Kernel Density Estimation to estimate $p(c, a, r)$. The robot memories experience (c, a, r), putting

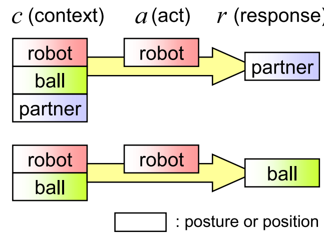

Fig. 8. 2 types of rule

Gaussian kernels at experienced (c, a, r) in $p(c, a, r)$ space. $p(c, a, r)$ is calculated as summation of the kernels.

The robot calculates each $motiv(c, a)$ of 2 types of rules in (c, a, r) (**Fig.8**). One is for interaction with a partner using the ball, and the other is for ball manipulation. The robot selects the action that maximizes summation of each $motiv(c, a)$. In other words, the robot performs the action $act(c)$ in (6).

$$act(c) = \underset{a}{\operatorname{argmax}} \left(motiv_{partner}(c, a) + motiv_{ball}(c, a) \right) . \tag{6}$$

where $motiv_{partner}$ is motivation to elicit partner's responses, and $motiv_{ball}$ is motivation to elicit ball's responses. With this summation, the robot can explore the ball especially when partner's responses are unknown or habituated.

We experiment in the following conditions.

1) The author and the robot in faster habituation interact for 1 minute.
2) Each participant A, B, C, D, E and the robot in faster habituation interact for 3 minutes.
3) Each participant F, G, H, I, J and the robot in slower habituation interact for 3 minutes.
4) Participant K and the robot in faster habituation interact for 14 minutes.

$h(x)$ is defined as $h(x) = \gamma \quad if\ x > 0, \quad 0 \quad else.$ $\gamma = 3$ in faster habituation. $\gamma = 1$ in slower habituation.

B. Result

A variety of interactions were observed in our experiment. First, we describe interactions with the robot in faster habituation. With participant K and the author, the partner and the robot interacted by passing a ball back and forth two consecutive times (**Fig.9**). With participants A and C and the author, the partner and the robot interacted by the robot dropping the ball, the partner picking it up, the robot taking it, and dropping it again (**Fig.10**). With participant D, the partner and the robot interacted by rolling a ball back and forth three consecutive times (**Fig.11**). With participant B, the partner and the robot interacted by the robot extending its arms with the ball, the partner extending his arms to take it, the robot pulling its arms back without handing the ball to the partner, and the partner pulling back his arms. This was done three times and was also done once more six seconds

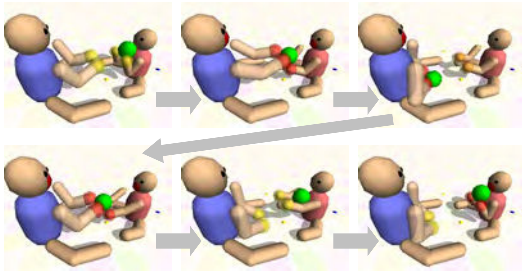

Fig. 9. Passing the ball back and forth

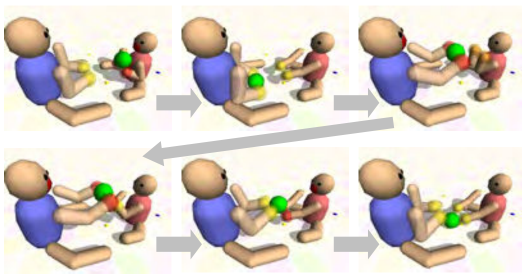

Fig. 10. Dropping the ball and taking it

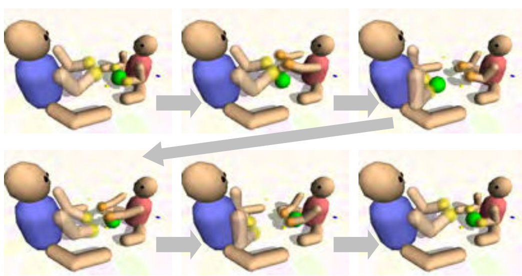

Fig. 11. Rolling and catching

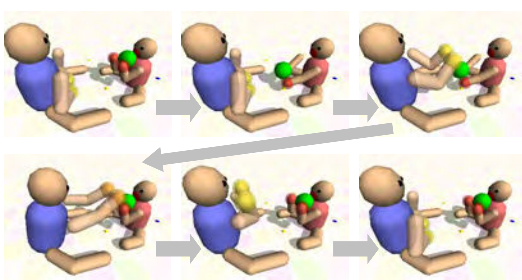

Fig. 12. Feint passing

later (**Fig.12**). This is considered to be feint passing. With participant K, a slightly different interaction was performed six times and then performed once more one minute later. With participant E, no obvious interaction was observed.

Next, we describe interactions with the robot in slower habituation. With participant H, the partner and the robot interacted by the robot extending its arms, the partner pulling back her arms, the robot pulling back its arms, and the partner extending her arms. This was done two consecutive times. This is considered to be complementary interaction. With participants F, G, I, and J, no obvious interactions were observed. In the case of slower habituation mode, the robot tended to repeat the same actions.

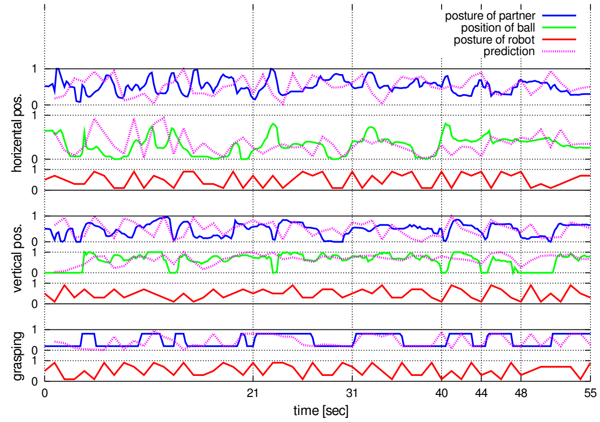

Fig. 13. Transition of interaction with the author and the robot's response prediction. 0 is the robot side, and 1 is the partner side.

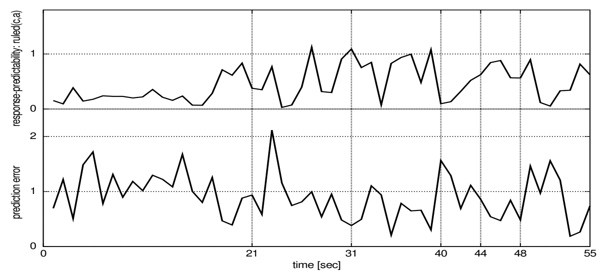

Fig. 14. Transition of response-predictability $ruled(c, a)$ and prediction error of response.

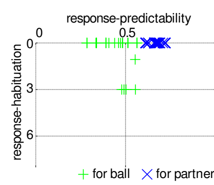

Fig. 15. Action distribution at 21 sec

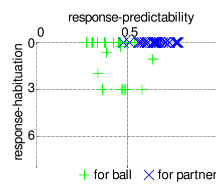

Fig. 16. Action distribution at 31 sec

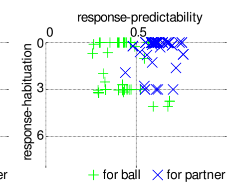

Fig. 17. Action distribution at 40 sec

We show transition of interaction with the author and the robot's response prediction (**Fig.13**). The interaction, passing the ball back and forth, is reciprocated in 21-31 sec and 31-40 sec. The interaction, dropping the ball and taking it, is reciprocated in 40-44 sec and 44-48 sec. We show transition of response-predictability $ruled(c, a)$ and prediction error of response in the interaction (**Fig.14**). $ruled(c, a)$ is larger in 31-40 sec than 21-31 sec, and in 44-48 sec than 40-44 sec. Prediction error is smaller in 31-40 sec than 21-31 sec, and in 44-48 sec than 40-44 sec. We also show distribution of actions for ball/partner in response-prediction/response-habituation space at 21, 31, 40 sec in the same interaction (**Fig.15**, **Fig.16**, **Fig.17**). The robot tried various actions and their response-predictability and response-habituation were dispersed. Especially, repeated actions' response-habituation increased. These included game-relevant rules such as the partner grasps the ball and pulls back his arms after the robot extends its arm and releases the ball, but many were not interpretable. These phenomena were also seen in other interactions.

V. Discussion

Several interesting conclusions can be drawn from the results for the reciprocated interactions. Our finding that response-predictability was higher at the second time than at the first time, indicates that the second interaction is predictable. That is, the first interaction is by chance, and the second one is intentional.

The results for the effect of habituation also lead to some interesting conclusions. When the habituation was faster, the robot tended to vary its actions more easily. This enabled the robot to basically perform game-relevant actions within the duration of the experiment. When it was slower, the robot tended to repeat the same actions. Practically reciprocal interactions were less likely to emerge because the robot tended to stick to non game-relevant actions and had difficulty performing game-relevant actions within the duration of the experiment.

The result of action distribution gives a suggestion. There were many actions with high response-habituation. They included the game-relevant rules, but many were not interpretable. This indicates that the rule definition is insufficient because what the rule realizes is action coordination between the two. The definition should be revised to support intention.

Various interactions emerged in the experiment. Those co-created rules were not determined at the beginning of the interactions but emerged through mutual involvement. The feint passing interaction in particular went beyond the expectations of the authors and the participant. A conventional social robot is basically supervised. Therefore, its only interactions are those that are taught. The proposed robot model, however, went beyond this limitation. Its interactions were not only those that were taught but also those that developed through interactions with the participants.

VI. Concluding Remarks

A. Summary

We focus on interaction rule that is relationship between a robot's action and a partner's response. We propose a perspective where a robot is simultaneously a learner and proposer of interaction rules. The human partner and robot continuously search for and co-create new rules. A simple and universal scheme with response prediction and habituation/dishabituation was developed, and a robot model was built using the scheme. The robot generates actions, observes the partner's response, and get to predict them. It identifies relationships between its actions and the responses, and generates actions designed to elicit particular responses from the partner. After it is habituated to the responses, it generates other actions to search for other rules. In experiments of human-robot interaction based on this model and using a ball, different patterns of interaction emerged, such as passing the ball back and forth, rolling and catching, and feint passing. The feint passing interaction in particular went beyond the expectations of the authors and the participant. This is significant as a result of the co-creation model.

Response prediction and appropriate habituation supported the emergence of interactions, indicating that the scheme and the model are effective. This joint learning should lead to natural communication between human partners and social robots beyond teach/taught relationship.

B. Future Work

According to the conclusions about habituation, we believe that regulation of habituation should be done appropriately. We point that emotion such as excitement is a key for the regulation. The robot watches excitement of a partner. If a partner is little excited, the robot accelerate habituation and search for rules. If a partner is excited enough, the robot decelerate habituation and stay with current rules. The robot should also express excitement, based on knowledge of infant's emotion [15]. In this way, a partner and the robot mutually regulate habituation and then progress of interaction. We think that this regulation of novel stimuli leads to sharing enjoyable mood.

VII. Acknowledgments

We acknowledge JST ERATO Asada Synergistic Intelligence Project for support.

References

[1] M. Fujita, On activating human communications with pet-type robot AIBO, *Proceedings of the IEEE*, vol. 92, no. 11, pp. 1804-1813, 2004.

[2] K. Wada, T. Shibata, T. Saito and K. Tanie, Effects of robot-assisted activity for elderly people and nurses at a day service center, *Proceedings of the IEEE*, vol. 92, no. 11, pp. 1780-1788, 2004.

[3] N. Mitsunaga, C. Smith, T. Kanda, H. Ishiguro and N. Hagita, Robot Behavior Adaptation for Human-Robot Interaction based on Policy Gradient Reinforcement Learning, Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1594-1601, 2005.

[4] C. Breazeal, D. Buchsbaum, J. Gray, D. Gatenby and B. Blumberg, Learning From and About Others: Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots, *Artificial Life*, vol. 11, no. 1-2, pp. 31-62, 2006.

[5] H. Ishihara, Y. Yoshikawa, K. Miura and M. Asada, Caregiver's sensorimotor magnets guide infant's vowels through auto mirroring, 7th IEEE International Conference on Development and Learning, pp. 49-54, 2008.

[6] T. Taniguchi and N. Iwahashi, Computational model of role reversal imitation through continuous human-robot interaction, Proceedings of the 2007 workshop on Multimodal Interfaces in Semantic Interaction, pp.25-31, 2007.

[7] G. E. Gustafson, J. A. Green and M. J. West, The infants' changing roles in mother-infants games: The growth of social skills, *Infant Behavior and Development*, vol. 2, pp. 301-302, 1979.

[8] N. Ratner and J. Bruner, Games, social exchange and the acquisition of language, *Journal of Child Language*, vol. 5, pp. 391-401, 1978.

[9] J. S. Bruner and V. Sherwood, Peekaboo and the learning of rule structures, *Play: Its role in development and evolution*, pp. 277-285, 1975.

[10] D. Stern, The First Relationship: Infant and Mother (Developing Child), *Harvard University Press*, 1977.

[11] , T. Kuriyama and Y. Kuniyoshi, Acquisition of Human-Robot Interaction Rules via Imitation and Response Observation, *Proceedings of the 10th International Conference of the Simulation of Adaptive Behavior*, pp. 467-476, 2008.

[12] T. Rome-Flanders and L. Cossette, Comprehension of Rules and Structures in Mother-Infant Games: A Longitudinal Study of the Early Two Years of Life, *International Journal of Behavioral Development*, vol. 18, no. 1, pp. 83-103, 1995.

[13] C. E. Shannon and W. Weaver, The Mathematical Theory of Information, *University of Illinois Press*, 1949.

[14] T. Schreiber, Measuring Information Transfer, *Physical Review Letters*, vol. 85, no. 2, pp. 461-464, 2000.

[15] P. Rochat, *The Infant's World*, Harvard University Press, 2001.