

Sociable Dining Table: Incremental Meaning Acquisition Based on Mutual Adaptation Process

Khaoula Youssef, P. Ravindra S. De Silva, and Michio Okada

Interaction and Communication Design Lab, Toyohashi University of Technology
1-1 Hibarigaoka, Tempaku, Toyohashi, Aichi
youssef@icd.cs.tut.ac.jp, {ravi,okada}@tut.jp
www.icd.cs.tut.ac.jp

Abstract. Our main goal is to explore how social interaction can evolve incrementally and be materialized in a protocol of communication. We intend to study how the human establishes a protocol of communication in a context that requires mutual adaptation. Sociable Dining Table (SDT) integrates a dish robot put on the table and behaves according to the knocks that the human emits. To achieve our goal, we conducted two experiments: a human-controller experiment (Wizard-of-Oz) and a human-robot interaction (HRI) experiment. The aim of the first experiment is to understand how people are building a protocol of communication. We suggest an actor-critic architecture that simulates in an open ended way the adaptive behavior that we have seen in the first experiment. We show in a human-robot interaction (HRI) experiment that our method enables the adaptation to the individual preferences in order to get a personalized protocol of communication.

Keywords: Mutual Adaptation, Communication Protocol, Actor-Critic.

1 Introduction

Developing robots with mutual adaptation skills and understanding the meaning acquisition process in the human-human interaction is a cornerstone to build robots that can work alongside humans and learn swiftly from intuitive interaction. By using the natural ability of humans to adapt to other artifacts, the robots can be capable of adapting to humans. Such an adaptation process would commonly be observed in a pair who can communicate smoothly, such as a child and a caregiver. Understanding how the caregiver behaves with the child affords many ideas to design intuitive robots facilitating the communication with people [1]. In fact, the caregiver's voice and physical contact lead to a mutual interest in communication. As a response the child generates some movements and utterances transferring his own assumptions to the caregiver. Incrementally, mutual adaptation evolved since both parties are trying to find the common successful patterns of communication which we name a communication protocol [2]. Our main goal is to explore how a communication protocol is established during the

mutual adaptation process in a human-human context and a human-robot context. We intend to develop a computational architecture that helps to simulate the human’s adaptive capability using the Sociable Dining Table (SDT). SDT affords the possibility to interact with the humans by displaying its behaviors while the human can interact through a knocking sound with the robot (Fig.1). Knocking is the only channel of communication used in our study that helps to draw a minimalistic scenario similar to the child-caregiver interaction’s scenario. It requires mutual adaptation from both parties in order to master and mirror the different most successful knocking and robot’s behaviors combinations [3].



Fig. 1. A participant interacts with the sociable dining table

2 Background

To enable the robot to learn flexible mapping relations when interacting with humans in daily life, many studies point out to the mutual adaptation as a very promising solution [4][5]. Mutual adaptation guarantees that if the human proposes new behaviors during the HRI, the robot will try to adapt and acquire the meaning of these new behaviors. Meanwhile, humans also will try to adapt to the robot if it proposed new behaviors [5]. The concept of adaptation was explored in many HRI studies [6][7]. Thomaz et al [8] used the active learning to adapt the robot’s knowledge. The robot addresses multiple types of explicit queries to learn the new concepts. Subramanian et al [9] used the explicit answer of the Pacman’s users concerning the best interactive options to propose a convenient adaptive Pacman agent that can learn from users. These studies explore the one-sided explicit adaptation (the artifact’s adaptation) while a mutual adaptive behavior has to exploit two levels of adaptation to evolve a flexible communication protocol. They also depend on explicit meaning affordance to teach the robot while the meaning can be inferred implicitly in the behavioral interaction between the human and the others. As an example, one can refer to the implicit communication between the caregiver and the child when they autonomously create their own meaning structure through a series of implicit interaction. Our work focuses on the implicit meaning’s acquisition and the incremental communication protocol formation through mirroring the patterns of each others’ behaviors to guarantee that double sided adaptation emerges.

3 Experiment 1: Human-Controller Experiment

3.1 Experimental Protocol

We conducted a Wizard-Of-Oz experiment that aims to ground the interaction between the human and the controller. 32 participants were grouped into 16 pairs (controller that controls the robot and a user that emits the knocking patterns) in order to lead the robot to the different checkpoints (Fig.2). To avoid the distraction by other sensory channels, the controller is located in another room, ignores the goal, checkpoints and refers only to the knocks. The user knows about the different checkpoints and has to lead the robot through knocking to the final goal after passing by the different checkpoints. The robot uses 5 reflectors [10] to avoid falling from the table. There are 3 trials where in the 1st and the last one we have chosen several configurations by proposing different checkpoints coordinates to guarantee the diversity of the patterns suggested by the participants (Fig.2). Both parties were informed that during the 1st trial the robot can operate only two behaviors (*right, forward*). Since we hypothesized that the pairs will try to build together a communication protocol, we chose 2 behaviors for the 1st trial in order to facilitate finding the successful patterns of communication. In the trial 2 and 3, we increase the degree of difficulty. We told the pairs that the robot can execute 4 behaviors (*right, left, back, forward*). Trial2 is a transitional stress-free session without any checkpoints which we believe that it can enhance the mutual understandability between the two parties. We informed the knocker and the controller during the trial2 that there were no specific trajectories nor checkpoints that the robot has to land on. By changing the configurations and the sessions' conditions, we aim at verifying whether the pairs human-controller can always mutually adapt to each others' behaviors.

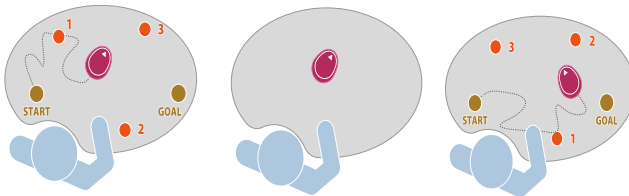


Fig. 2. In the 1st trial (left), each participant has to move the creature into 5 places (start, 1, 2, 3, goal) by knocking using 2 behaviors (right, forward). The 2nd trial (center), is a stress-free session where we do not assign any configuration. In the 3rd trial (right), we changed the place of the former points, and then the user has to guide the robot into the new points using 4 behaviors (right, forward, left, back).

4 Experiment 1: Results and Discussion

4.1 Behavior Adaptation Process

Although, we set up 20 minutes as a time limitation to achieve the task, all the participants reached the different checkpoints in less than 15 minutes. Thus,

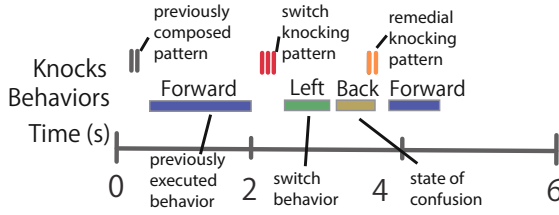


Fig. 3. A scenario showing examples of switch knocking pattern, switch behavior, state of confusion and remedial knocking pattern

to study the incremental adaptation to each others' behaviors, we calculated the number of switch knocking patterns, switch behaviors, states of confusion and the remedial knocking. Figure 3 helps to understand the meaning of these four concepts. As you may see in the Figure 3, the robot executed initially the *forward* behavior and when the controller detected that he received the switch knocking pattern (3 knocks in red), he picked *left* as a new behavior which we call according to this scenario a switch behavior. Thus, we call a switch knocking pattern a new received pattern that is different from the previous received one and a switch behavior the controller's picked behavior as a response for the received switch knocking pattern. Within few milliseconds, we can see that again the controller changes the behavior to *back*. We call such situation a state of confusion since the controller changes the behavior without being prompt by any knocking. As a response the knocker, composed 2 knocks (in orange) as a remedial knocking pattern for the controller's state of confusion. If for each switch knocking pattern, we have systematically a switch behavior then we may conclude that the controller is trying to adapt to the knocker's patterns of knocking. The presence of states of confusion indicate that the controller is trying to establish the rules of communication but may go through some confusing states. Consequently, the knocker also tries to adapt to the controller's state of confusion by composing a remedial knocking pattern and thus the existence of mutual adaptation can be proved. We calculated the test of independence between the switch knocking patterns and the switch behaviors. Table 1 exhibits the Chi-square test results and Cramer V values. A Cramer V value ranging from 0,15 to 0,20 showed that a minimally acceptable dependence exists between the two measured variables while a value ranging from 0,20 and 0,25 showed that we have a moderate dependence and finally a value ranging between 0,35 and 0,41 showed that a very strong relationship exists between the two variables.

Table 1 revealed that during the trial 1, there was no statistically significant relationship between the knocker's switch knocking and the controller's switch behaviors. However, during the trial 2 and 3 we had significant values with p-values respectively equal to 0,036 and 0,0001. By comparing the two Cramer's V values of trial 2 and trial 3, we have $Cramer's V_{trial2} = 0,170 \leq Cramer's V_{trial3} = 0,245$ showing that the dependency between the two variables is becoming gradually larger. This proves that there was incrementally an attempt to combine each pattern to a robot's behavior.

Table 2 revealed that during the trial 1, there was no statistically significant relationship while during the trial 2 and 3 the p-values were respectively equal to 0,019 and 0,004 were significant. By comparing the two Cramer’s V values of trial 2 and trial 3, we have $Cramer's V_{trial2} = 0,260 \leq Cramer's V_{trial3} = 0,279$ showing that the dependency between the two variables is becoming gradually larger. This proves that the controller was trying to adapt himself and thinking about the best behavior that may correspond to the heard patterns.

We calculated the test of independence between the states of confusion and the remedial knocking. Table 3 exhibits the Chi-square test results and Cramer’s V values.

Finally, the Table 3 revealed that during the trial 1, there was no statistically significant relationship. However, during the trial 2 and 3 the p-values were significant with values respectively equal to 0,043 and 0,001. By comparing the two Cramer’s V values of trial 2 and trial 3, we have $Cramer's V_{trial2} = 0,316 \leq Cramer's V_{trial3} = 0,410$ showing that the dependency between the two variables is becoming gradually larger. This proves that the knocker was adapting himself in order to afford for the controller the suitable pattern so he can find his way to the correct behavior. Consequently, based on the 3 tables we can confirm that a double sided adaptation emerges.

4.2 Interaction Smoothness

It is generally assumed that almost any human behavior that involves information processing and decision-making tends to increase the reaction time. We

Table 1. The test of independence between the switch knocking patterns and the switch behaviors as well as the Cramer’s V values by means of trial

Trial	χ^2 value	P-value and significancy	Cramer’s V (CV)
Trial1	$\chi^2 = 1,112;df=4$	P-value=0,892 at $\alpha = 0.05$ not significant	no significance
Trial2	$\chi^2 = 22,104;df=12$	P-value=0,036 at $\alpha = 0.05$ significant	CV=0,170
Trial3	$\chi^2 = 42,987; df=12$	P-value=0,0001 at $\alpha = 0.05$ significant	CV=0,245

Table 2. The test of independence between the switch knocking patterns and the states of confusion as well as the Cramer’s V values by means of trial

Trial	χ^2 value	P-value and significancy	Cramer’s V (CV)
Trial1	$\chi^2 = 2,334;df=4$	P-value=0,675 at $\alpha = 0.05$ not significant	no significance
Trial2	$\chi^2 = 24,16;df=12$	P-value=0,019 at $\alpha = 0.05$ significant	CV=0,260
Trial3	$\chi^2 = 28,787;df=12$	P-value=0,004 at $\alpha = 0.05$ significant	CV=0,279

Table 3. The test of independence between the states of confusion and the remedial knocking by means of trial as well as the Cramer’s V values

Trial	χ^2 value	P-value and significancy	Cramer’s V (CV)
Trial1	$\chi^2 = 2,635;df=4$	P-value=0,621 at $\alpha = 0.05$ not significant	not significance
Trial2	$\chi^2 = 4,505;df=12$	P-value=0,043 at $\alpha = 0.05$ significant	CV=0,316
Trial3	$\chi^2 = 33,227;df=12$	P-value=0,001 at $\alpha = 0.05$ significant	CV=0,410

wanted to verify whether the controller’s response time¹ changes by means of trial (Fig.4). If the response time becomes shorter, we conclude that an adaptation process has facilitated the decision making. The results showed that 75% of the reaction time is in the range of [2-4] seconds . Kruskal-Wallis test proved that there were statistical differences concerning the controller’s reaction time during the different 3 trials with(K (Observed value)=13.835; df=2; p-value (Two-tailed)=0.001; alpha=0.1). The multiple pair wise comparisons using the Steel-Dwass-Critchlow-Fligner test showed that there were significant differences between the trial 1 and 2, the trial 3 and 1 but there was no significant differences between the trial 3 and 2. Figure 4 depicts the average reaction time by means of trial for each one of the 16 pairs (knocker-controller) where blue color corresponds to trial 1, red to trial 2 and green to trial 3. During the trial 2 and 3 that involves a higher degree of difficulty, the reaction time decreases slightly in comparison to the trial 1 when the pairs were trying to adapt with a lower task difficulty (2 behaviors). Consequently, even if the complexity of the task increases, the pairs were more engaged during the 2 last trials to acquire incrementally the communication protocol and the decision making becomes easier.

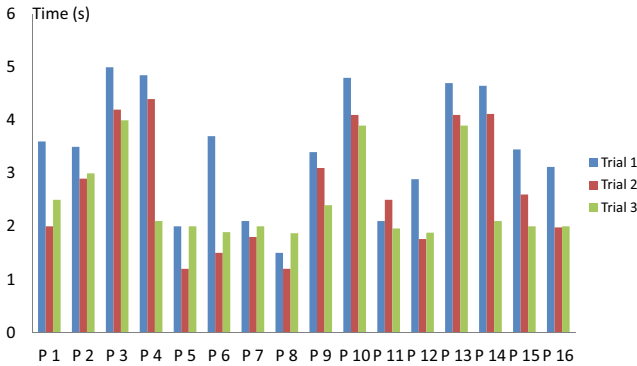


Fig. 4. The response time during the three trials

4.3 Visualization of the Incremental Acquisition of the Protocol of Communication

Using a visual approach which is the correspondence analysis, we succeed in representing the protocol of communication that can be defined as a map which represents the different pairs’ knocking patterns and the robot’s behaviors. The frequency for each behavior (*forward*, *right*, *left*, *back*) and for each knocking pattern (e.g: 2 knocks, 3 knocks) is considered in order to expose the Euclidean distance in two dimensions. Figure 5 depicts the correspondence analysis for the pair 15 during 3 trials. The red triangles represent the robot’s behaviors and

¹ It is the time between the onset of the knocking pattern and the time of the 1st response of the controller regardless of whether it was correct or not.

the blue circles represent the knocking patterns. During the trial 1 (Fig.5 (left)), the *right* behavior was associated with 4 and 2 Knocks and *forward* with 3 and 6 knocks. During the trial 2 (Fig.5 (center)), the behavior *back* was associated with 3 knocks, *left* and *forward* with 1 knock while *right* with 2 and 4 knocks. Finally, during the trial 3 (Fig.5 (right)), the pair successfully distinguished the different combinations where 4 knocks was associated with *back*, 2 knocks with *right*, 1 knock with *left* and 3 knocks with *forward*. The different correspondence analysis results proved that the pairs try to establish a communication protocol incrementally.

4.4 The Convergence to a Protocol of Communication

We wanted to explore statistically the differences between the participants' communication protocols during the 3 trials. For this issue, based on the correspondence analysis results, we calculated the euclidean distance between each of the robot's behaviors (red triangles as presented in the Fig.5) and the different patterns (blue circles as presented in the Fig.5). After, we picked for each behavior the minimum distance. We sum up the 4 minimum distances² and the resultant value which we call convergence metric, affords an information about the minimum distance that the pair knocker-controller reached to form stable rules. We repeated the same procedure for the 16 pairs and for the three trials.

To verify whether there was statistically convergence differences during the three trials, we used the Kruskal-Wallis test. As the computed p-value=0,01 is lower than the significance level alpha=0,1, we accept the alternative hypothesis confirming that there was a clear statistical difference concerning the convergence to a protocol between the different trials. We applied the multiple pair wise comparisons using the Steel-Dwass-Critchlow-Fligner test to verify the significant differences between the different trials. The statistical results showed that there were differences between the trial 2 and 3 and between the trial 1 and 3. Combining the statistical tests and the different correspondence analysis, we conclude that there was a tendency to associate for each behavior a knocking pattern especially during the trial 3.

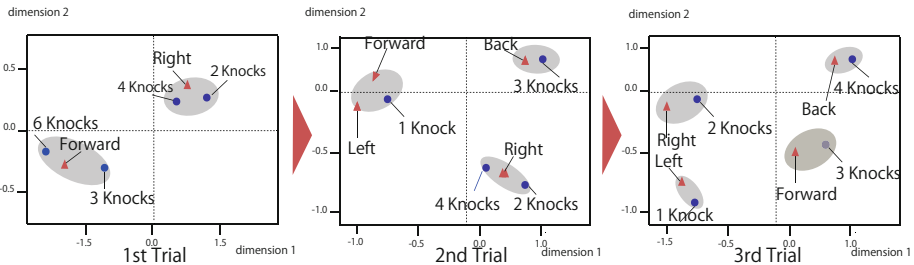


Fig. 5. The correspondence analysis representing the communication protocol during the trial 1 (left), the trial 2 (center) and the trial 3 (right)

² Each minimum distance is associated with one behavior.

5 Actor-Critic Architecture

Through the 1st experiment, we noticed that incrementally people use in a trial-error process the different successful combinations of (knocking pattern/ robot's behavior) to establish the rules of communication. We proposed a similar trial and error method that is based on the reinforcement learning. Our solution consists on an actor-critic architecture which we expected that it will help to establish a communication protocol.

5.1 Actor Learning

Each knocking pattern has its own distribution $X(S_t) = N(\mu_{X(S_t)}, \sigma_{X(S_t)})$ where $X(S_t)$ is defined as the knocking pattern, $\mu_{X(S_t)}$ and $\sigma_{X(S_t)}$ are the mean value and the variance. We chose 2 s as a threshold for the user's reaction time based on the human-controller experiment. In fact, the results showed that the reaction time is in the range of [2-4] seconds (s) and thus we assumed that we assumed that we have a disagreement state if the human interrupted the robot when it is executing the chosen behavior within 2s. When the robot observes the state S_t that is materialized by a knocking pattern, the behavior is picked according to the probabilistic policy $\Pi(s_t)_{nbknocks}$. If within 2s there was no knocking pattern, we suppose that the robot has succeeded by choosing the right behavior and the critic reinforces the value of the executed behavior in the state S_t to increase its chances to be picked in the future if the robot receives the same knocking pattern. Finally the system will switch to the state S_{t+1} . But if a new knocking pattern is composed before that 2s elapsed, the state of the interaction changes to the state S_{t+1} indicating that the knocker disagrees about the behavior that was executed, the probabilistic policy failed to propose the correct behavior. The critic updates thus the value function before choosing any new behavior. As long as the knocker is interrupting the robot's behavior before that 2 seconds elapsed, the actor chooses the action henceforth by pure exploration (until we meet an agreement state: no knocking during 2 seconds) based on (1). The random values vary between $0 \leq rnd1$, and $3 \leq rnd2$ the above range was decided to bring the values of the action (1) between 0 and 3 (corresponding to the behaviors' (forward, right, back, left) numerical codes). We assume in such case that the knocker will randomly compose the patterns just to switch desperately the robot's behavior.

$$A(S_t) = \mu_{X(S_t)} + \sigma_{X(S_t)} * \sqrt{-2 * \log(rnd_1)} * \text{Sin}(2\Pi * rnd_2) \quad (1)$$

5.2 Critic Learning

The critic calculates the TD error δ_t as the reinforcement signal for the critic and the actor according to Equ.2

$$\delta_t = r_t + \gamma V(s_{t+1}) - V(s_t) \quad (2)$$

with γ is the discount rate and $0 \leq \gamma \leq 1$. According to the TD error, the critic updates the state value function $V(s_t)$ based on (3).

$$V(S_t) = V(S_t) + \alpha * \delta_t \quad (3)$$

where $0 \leq \alpha \leq 1$ is the learning rate. As long as the knocker disagrees about the executed behavior before 2 s elapsed, we refine the distribution $N(\mu_{X(S_t)}, \sigma_{X(S_t)})$ which helps us to choose the action according to (1). The distribution update consists on computing (4) and (5).

$$\mu_{X(S_t)} = \frac{\mu_{X(S_t)} + A_{S_t}}{2} \quad (4)$$

$$\sigma_{X(S_t)} = \frac{\sigma_{X(S_t)} + |A_{S_t} - \mu_{X(S_t)}|}{2} \quad (5)$$

6 Experiment 2: the Human-Robot Interaction

6.1 Experimental Setup

A second experiment HRI was conducted to show that our architecture learns in real time how to establish the protocol of communication based on the knocking patterns. In this experiment, 10 participants accomplish the same task as in the 1st experiment with two different configurations for the two trials that are also different from those used in the trial 1 and 3 of the experiment 1 (Fig.2) .

6.2 Visualization of the Incremental Acquisition of the Protocol of Communication

We remarked that the human-robot pairs were able to establish communication protocols that allowed the robot to reach the different checkpoints. As in the first experiment, we applied the correspondence analysis for all the participants' interaction data to visualize the communication protocol. Figure 6 exhibits respectively the results of the 1st (left) and the 2nd (right) trial. Figure 6 (left) shows that there was some tendency to attribute for the behaviors different patterns. *Right* was combined with 1 knock, *forward* with 2 knocks with some confusion for the *left* behavior (1 and 4 knocks). During the 2nd trial (Fig.6(right)), the Euclidean distance between *forward* and the pattern 2 knocks decreases, *right* was combined with 1 knock and *left* with 3 knocks.

6.3 The Convergence to a Protocol of Communication

As in the 1st experiment, we calculated for the two trials, the convergence metric values of the 10 participants based on the correspondence analysis results. To verify whether there was statistically some convergence differences during the 2 trials, we used the Mann-Whitney two-tailed test. As the computed p-value=0.027 is lower than the significance level alpha=0.05, we accept the alternative hypothesis confirming that there were a clear differences concerning the

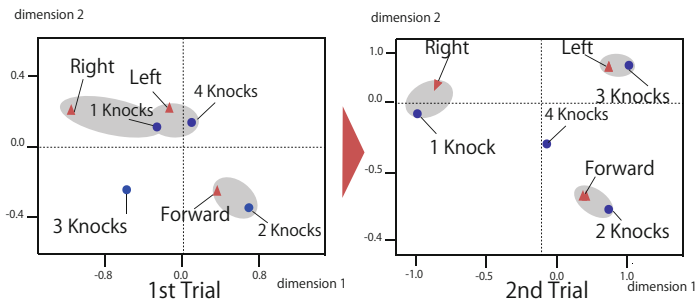


Fig. 6. The correspondence analysis displaying the communication protocol during the trial 1 (left) and the trial 2 (right)

convergence between the trial 1 and 2. As a conclusion, we acknowledge that each participant is collaborating with the robot in order to find out the common best practices associating each behavior with the most convenient generated knocking pattern exactly as in the human-controller experiment.

7 Conclusion

The results showed that the WOZ experiment helps to explore how mutual adaptation evolves between the controller and the knocker and how a protocol of communication can emerge incrementally. The 2nd experiment indicates that there was an incremental formation of a protocol of communication as in the 1st experiment. Although the promising results that we gathered, we have seen that in some cases there are some participants that have slowed adaptation in comparison to others which can be justified by the fact that there are some people that gets along with a different kind of learning. In our future work, we intend to elaborate a learning method that helps to boost the convergence to a communication protocol using inarticulate sounds.

Acknowledgments. This research is supported by Grant-in-Aid for scientific research of KIBAN-B (26280102) from the Japan Society for the Promotion of science (JSPS).

References

1. Michaud, F., Laplante, J., Larouche, H., Duquette, A., Caron, S., Letourneau, D., Masson, P.: Autonomous spherical mobile robotic to study child development. In: IEEE International Conference on Systems, Man, and Cybernetics, vol. 4, pp. 1–10 (2005)
2. Condon, W.S., Sander, L.W.: Neonate movement is synchronized with adult speech: interactional participation and language acquisition. *Science* 183, 99–101 (1974)

3. Matsumoto, N., Fujii, H., Okada, M.: Minimal design for human agent communication. In: *Artificial Life and Robotics*, pp. 49–54 (2006)
4. Okada, Y., Ueda, S., Komatsu, K., Takeshi, O., Kamei, K., Yasuyuki, S., Nishida, T.: Formation conditions of mutual adaptation in human-agent collaborative interaction. *Applied Intelligence*, 208–228 (2012)
5. Xu, Y., Ueda, K., Komatsu, T., Okadome, T., Hattori, T., Sumi, Y., Nishida, T.: Woz experiments for understanding mutual adaptation. *AI Society*, 201–212 (2008)
6. Thomaz, A.L., Breazeal, C.: Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence*, 716–737 (2000)
7. Mitsunaga, N., Smith, C., Kanda, T., Ishiguro, H., Hagita, N.: Robot behavior adaptation for human-robot interaction based on policy gradient reinforcement learning. In: *Intelligent Robots and Systems*, pp. 218–225 (2005)
8. Crystal, C., Cakmak, M., Thomaz, A.L.: Transparent active learning for robots. In: *Human-Robot Interaction*, pp. 317–324 (2010)
9. Subramanian, A., Charles, K., Isbell, L., Thomaz, A.L.: Learning options through human interaction. In: *Agents Learning Interactively from Human Teachers*, pp. 208–228 (2011)
10. Kado, Y., Kamoda, T., Yoshiike, Y., De Silva, P.R.S., Okada, M.: Reciprocal-adaptation in a creature-based futuristic sociable dining table. In: *2010 IEEE RO-MAN*, pp. 803–808 (2010)