

# Sociable Dining Table: Meaning Acquisition Exploration in Knock-Based Proto-Communication

Khaoula Youssef<sup>1</sup> · Takahiro Asano<sup>1</sup> · P. Ravindra S. De Silva<sup>1</sup> · Michio Okada<sup>1</sup>

Accepted: 23 July 2015 / Published online: 30 August 2015  
© Springer Science+Business Media Dordrecht 2015

**Abstract** In order to build social robots that can coexist with human beings, it is necessary to understand the mechanisms of how communication protocols are developed in human–robot interactions. Our main goal is to explore how a communication protocol can be established incrementally between a human and our minimally designed robot which is called sociable dining table (SDT). SDT integrates a dish robot put on the table and behaves according to the knocks that a human emits. To achieve our goal, we conducted two experiments: a human–human experiment (Wizard-of-Oz) and a human–robot interaction (HRI) experiment. The aim of the first experiment was to explore how people build a protocol of communication. Based on the first experiment, we suggested an actor-critic architecture that simulated in an open-ended way the adaptive behavior which we determine in the first experiment. After that, we demonstrated in the HRI experiment how our actor-critic architecture enabled the adaptation to individual preferences in order to obtain a personalized protocol of communication.

**Keywords** Actor-critic architecture · Sociable dining table · Protocol of communication · Usercentered design

---

✉ Khaoula Youssef  
khaoula.youssef10@gmail.com

Takahiro Asano  
asano@icd.cs.tut.ac.jp

P. Ravindra S. De Silva  
ravi@tut.jp

Michio Okada  
okada@tut.jp

<sup>1</sup> Interaction and Communication Design Lab., Toyohashi University of Technology, 1-1 Hibarigaoka, Tempaku, Toyohashi, Aichi-ken, Japan

## 1 Introduction

As robots move from the research lab to the real world, it is interesting that users, including those without programming skills, can teach robots customized behaviors [1,2]. If sophisticated methods were developed in order to allow users to transfer their knowledge, we may be able to guarantee long-term communication and mutual understanding. Developing robots with mutual understanding skills and exploring the meaning acquisition process in the human–human interaction is a cornerstone to build robots which can work alongside humans. By using human adaptation capability adequately, robots are capable of adapting to humans and will be easily adaptable as well. Such a process can commonly be observed in a pair who can communicate smoothly, such as a child and a caregiver.

Understanding how a caregiver behaves with a child is required to achieve key ideas about the behaviors, that can be used to design intuitive robots [3–6]. Many issues have been of interest to the HRI community, such as how children learn to talk [7], grasp an object [8], and navigate [9], etc. Understanding how such issues occur helps roboticists building intuitive robots. During a child-caregiver communication scenario, the child and the caregiver try to adapt to each other using a limited number of communication channels which they initially do not master in the same way. Incrementally, they become familiar to each other's patterns of communication. The meaning decoding of each other's behavior is no more difficult for both parties. In fact, each party implicitly infers the meanings of the other party's most commonly used patterns and links the most often used patterns to the context of the interaction. Such linking leads to an implicit formation of (patterns-meanings) cartography, which in our study is called a “communication protocol”. A non-expert user and a minimally designed robot also try to customize a commu-

nication protocol which depends on the patterns emerging from limited communication channels that are not mastered in the same way during the initial communication stages.

In this vein, the purpose of this study is to explore how non-expert users can cooperate with a minimally designed robot in order to acquire a communication protocol. The challenge is to investigate how people aggregate communication patterns. We want also to investigate how to adequately take advantage from the adaptation ability of humans in order to enable our minimally designed robot SDT to adapt to new situations during a novel interaction scenario that integrates minimal communication channels. Understanding how to take advantage from the human's adaptation strategy helps us to tailor a control model for minimally designed robots that have a minimal number of communication channels. The final designed control model has to guarantee the establishment of flexible communication protocols just as in the child-caregiver interaction context.

Therefore, we draw a scenario inspired from the child-caregiver interaction and opt for knocking as the only one communication channel used by humans. Knocking is a novel communication channel that had not been used in a similar task. This guarantees that the user and the robot have the same amount of knowledge about the communication scenario. Thus, to have a successful interaction both parties need to adapt to each other. To explore how the adaptation occurs, we conduct our first experiment. It is a human–human (H–H) experiment. For each instance of interaction during the H–H experiment, we engage two participants. The first participant is the one that knocks on the table while watching the robot moving on the table [room (A)]. The second participant is the one remotely controlling the robot according to the knocking sounds. Thereby, the robot is controlled via an interface. The second participant is located in another room [room (B)].

Both parties have to cooperate in order to make the robot visit different checkpoints marked on the table. We informed each new pair (knocker-controller) that the robot can use 4 behaviors (going forward, going back, going left, going right). Based on this experiment, we want to investigate whether the task can be achieved using our only communication channel. In the case of a successful interaction, we want to explore what are the stages that the communication went through and what are the best adopted practices that led to the emergence of a communication protocol? After that, we want to implement in the robot the components and the functionalities that may guarantee to make our robot adaptive. Finally, we conduct another experiment (HRI experiment) to verify whether our robot was adaptive like in the H–H experiment. Also, we compare the H–H and the HRI experiments in terms of performance, emergent communication protocols, and the way the task is solved in each experiment.

## 2 Background

Adaptation is a term referring to the ability to adjust to new information and experiences, track the new facets of the environment and adopt the most convenient strategies based on the sequentially gathered information. Many studies point out the robot and human's adaptation to each other as being a very attractive and promising solution for the HRI [10, 11]. Robot and human's adaptation to each other consists of the fact that if the human changes his behavior, the robot must adapt to this new behavior. Humans also have to change their behavior patterns to adapt to the robot's new proposed behaviors during an instance of an HRI [10]. Yamada et al. [12] investigate the capability of the human and the agent to detect each others' state of mind based on few social cues such as facial expressions [13]. The concept of adaptation is explored in many other HRI studies [13–15].

Some studies use many modalities integrated into the robot [16–18] in order to design an adaptive artifact. Other studies [19, 20] examine how a speaking robot can infer the adequate speech by combining words to particular contexts through observing different situations. Kanda et al. use the robot Robovie in HRI studies to investigate children's interaction in a museum [21] and a school [22]. Thomaz and co-workers [23] investigate the active learning to refine the robot's knowledge where multiple types of queries are used by the robot to demand an explicit spoken answer facilitating the robot's concept learning process. Subramanian et al. [24] use the explicit answer of Pacman game users concerning the best interactive options that they imagine are effective for the agent teaching. These interactive options are learned in an offline mode and introduced later into the robot. These studies [20–22, 24] explore the explicit verbal communication to implement adaptive systems while the meaning can be inferred in real time implicitly based on the behavioral interaction. We do not address the general problem of multimodal communication channels and instead we focus on a minimal communication channels concept which we expect can guarantee the emergence of simple communication patterns and is suitable for minimally designed robots.

Minimal Design Policy is first proposed by Matsumoto et al., who conclude that the robot's appearance should be minimized in its use of anthropomorphic features so that the humans do not overestimate or underestimate the robot's skills [25]. By minimal design, we mean eliminating the non-essential components and keep only the most fundamental functions. We expect that in the future minimally designed robots will be affordable. People will use such minimally-designed robots for many tasks such as cleaning, and here we may mention the Roomba robot [26] or to engage more with autistic children through therapeutic sessions of interaction while cooperating with Keepon the robot [27], etc.

Minimal design policy is applied to develop many other robots such as Muu [28], ROBOMO [29], CULOT [30], etc. The simple nature of minimally designed robots allows humans to interact easily with such robots on a daily basis. On the other hand, we must pay attention to sociability and adaptation factors. In fact, interacting with an affordable minimally designed robot may represent the first experience of a human interacting with a robot. This, lead us to assume that people will possibly have high expectations about the robot's adaptive capabilities.

In addition to humans having a natural tendency to forget quickly, there are not exact details of how an interaction occurs and what are the instructions used. For this, a human attempts to come up with any similar instructions to solve the problem. A similar phenomenon occurs in the human-pet interaction when the human forgets the exact instruction taught to the pet [31]. Interestingly, the human in that case does not recognize the difference and the pet tries to grasp the meaning incrementally in order to satisfy the human's request. In this context, we believe that robots need an extra capability which enables them to grasp the meaning of the newly introduced instructions and satisfy the human's new request. Kiesler [32] concurs with our point of view while he confirming in his studies that a minimally designed robot has to integrate a process which makes it adaptive [33]. Thus, one contribution of this paper is to determine how a minimally designed robot can incorporate an adaptive process that helps establishing a communication protocol with non-expert users and adapt to their different communication patterns.

To achieve the above goal, we chose to conduct a WOZ experiment to explore how a communication protocol can be established between the users and a minimally designed robot. It is a well-known principle in robot design, that the roboticist should involve humans early in the design process, rather than in the final evaluation phase [34]. Many HRI studies [35–37] use the WOZ experiment in order to test early aspects of the robot's design. We agree with the fact that WOZ can help in exploring the best features which can be later incorporated in the robot's design. Also, we believe that robots are not sufficiently advanced to interact autonomously with people in a socially appropriate way. Therefore, we started our study by conducting a WOZ experiment that helped exploring the best practices humans adopt in order to establish a communication protocol. Based on the first experiment, we gained some insights in order to incorporate in our robot's architecture the best adopted practices that can get along with people's communication patterns in the context of the SDT interaction. Finally, we attempted to validate our robot's architecture through an HRI experiment in order to compare the HRI performance to the WOZ experiment performance.

We start by exposing the architecture of the SDT in Sect. 3. In Sect. 4, we explain our H–H experiment. In Sect. 5, we



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

explain our proposed architecture. Finally, in Sect. 6 we validate our minimal architecture based on an HRI experiment (Figs. 1, 2).

### 3 Architecture of the SDT

Our system consists of a webcam to compute the robot's positions and its angle of orientation. The robot's coordinates are used only for further analysis purposes (Fig. 3). The robot uses four microphones to localize the knock's source based on the weighted regression algorithm [38]. It communicates with the host computer through Wi-Fi using a control unit [a macro computer chip (AVR ATMEGA128)] and employs a servomotor that helps to exhibit the different behaviors: right, forward, left and back. Finally, five photo reflectors are utilized to automatically detect the boundaries of the table and avoid falling (Fig. 4).

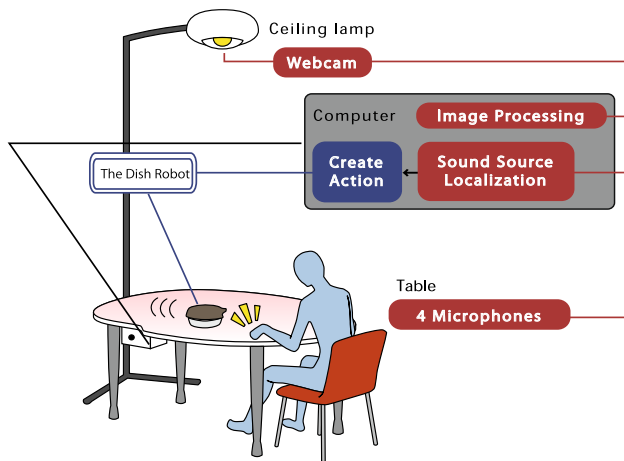
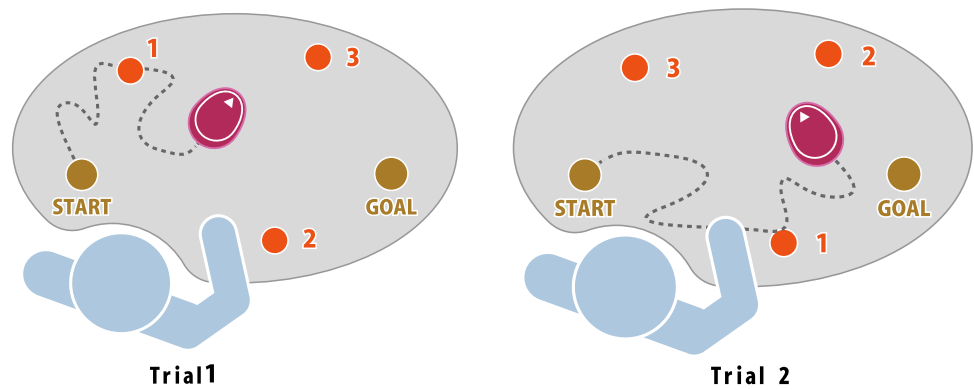
### 4 Experiment 1: Human–Human Interaction

We expect that H–H experiment allows the envisioning of future useful features that can be integrated into the robot's architecture in order to make our minimally designed robot SDT adaptive.

#### 4.1 Experimental Setup

Each time we conducted an instance of the H–H experiment, we gathered a new pair of participants and assigned the first one to the role of a knocker while the other to the role of a controller. The knocker was the one that has to knock on the table in order to help the robot visit different points marked on the table. The controller was the one that has to remotely control the robot based on the knocking.

**Fig. 2** In the first trial (*left*), the controller tries to understand the knocker's patterns of knocking in order to move the robot into five decided places on the table (start, 1, 2, 3, and goal) by means of knocking patterns. In the second trial (*right*), we change the place of the former points on the table, and then the knocker and the controller have to exploit the emerged rules of communication of the first experiment to guide the robot into the newly defined points



**Fig. 3** The overall architecture of the SDT: the human's knock is detected by four microphones while the robot executes the different behaviors using the servomotor



**Fig. 4** A close-up picture showing the inside of the SDT robot

Before a knocker enters the experimental room (A), the instructor told him the purpose of the experiment is to help the robot to land on different checkpoints marked on the table. The knocker did not know that a human controlled the robot when he knocked, while the controller did not know that another person emitted the knocking. This helped us to simulate convenient conditions guaranteeing that any possible emerging communication protocol would emerge if we

were in a real HRI. Also, by exploring how gradually a communication protocol emerged we may find out the key ideas that we needed to integrate in order to elaborate a convenient adaptive architecture for our robot. The knocker was located in a first room (A) and can visualize the robot as well as all the checkpoints on the table. In another room (B), the controller remotely controlled the robot while listening to the knocking without seeing the predefined checkpoints. The controller could only visualize an interface showing the robot moving since he was in another room. We isolated each party in a different room in order to make sure that no eye contact or facial expressions could be exchanged between both parties. The instructor told the controller that he needed to listen to the knocking, guess the meaning and then choose the convenient direction based on his own opinion. Finally, after the experiment ended we interviewed both participants (knocker and controller). Importantly, we asked them to describe their experience with the robot through simple phrases.

In the first trial, the pair (knocker-controller) had to cooperate in order to lead the robot to different sub-goals (Fig. 2). In the second trial, we changed the coordinates of the former points and the pair (knocker-controller) had to cooperate to reach the new check points. We chose several different configurations. At each time the goal position and the intermediate check points were changed. This may guarantee that the participants were not accustomed to the configuration. Also, it helped us confirming the pairs (knocker-controller) used their adaptation abilities and the emerging communication patterns rather than memorizing the different transitions that helped to achieve the task in the previous trial. There are two trials, each lasting 20 minutes<sup>1</sup> and video-recorded. During each new trial, the new controller and the new knocker try to cooperate in order to achieve the task. We did not indicate for the pairs that they must follow a special knocking strategy so that they interact in a natural way with the robot and we can also see whether they aggregated some redundant patterns to form a communication protocol with the robot.

<sup>1</sup> We estimated this period based on a previous pilot study.





**Fig. 5** The first (*left*), second (*center*) and final (*right*) time segmentations of an extract of the interaction from the first experiment where in the *first line* we have action executed by the robot: *F*, *R*, *L* and *B*

stands for forward, right, left and back behaviors; in a *second line*, the corresponding knocking patterns such as 2 or 3 knocks, etc.; and in a *third line* the time progress in seconds

## 4.2 Subjects

We hired thirty Japanese students (ages: Mean ( $M$ ) = 20.2, Standard Deviation ( $SD$ ) = 2.0 [years]) from different universities. Sessions 1 and 2 were performed with thirty subjects (eighteen males and twelve females). A written informed consent was obtained from all the subjects.

## 4.3 Results

After the experiment was finished, we attempted to analyze the interaction scenarios in order to verify whether a communication protocol was established between the knockers' knocking patterns and the chosen actions. We also attempted to detect the components that led to the possibly emergent communication protocols.

We analyzed the video data by annotating with a video annotation tool called ELAN. Two coders, one of the authors and one volunteer, analyzed the behavioral data captured in the video camera using the same coding rules for the first and the second trials. We picked ten data sets arbitrarily from our entire data set which were coded based on rules. We calculated the average of Cohen's kappa to investigate the reliability. As a result, we confirmed that there was a reliability with  $\kappa = 0.98$

### 4.3.1 Evaluation of the Command-Like and the Continuous-Knocking Patterns Based on the Videos

We remarked that there are 2 types of patterns: continuous - knocking patterns and command-like patterns. Command-like pattern consisted of combining each behavior with a different combination of knocks (e.g., 2 knocks for *Forward*). Continuous-knocking was used when there was contiguous interruptions in the robot's behavior.<sup>2</sup> We counted the number of both types of patterns based on the coded data for each participant and for the two trials. We noticed that there was a significant usage of the command-like patterns (90.26 % of the patterns were command-like during trial 1 compared with 89.47 % of the patterns during trial 2).

To verify whether the usage of command-like was statistically significant, we conducted a  $t$  test between the number

of command-like patterns and the number of continuous-knocking patterns used by the participants during the trial 1: ( $t = 6.973$ ,  $d.f. = 14$ ,  $p \text{ value} < 0.01$ ) and trial 2: ( $t = 4.750$ ,  $d.f. = 14$ ,  $p \text{ value} < 0.01$ ). For both  $t$  tests, we found that there was a significant difference between both types of patterns usage during trials 1 and 2, highlighting that participants were trying to simplify the input in each interaction cycle for the robot.

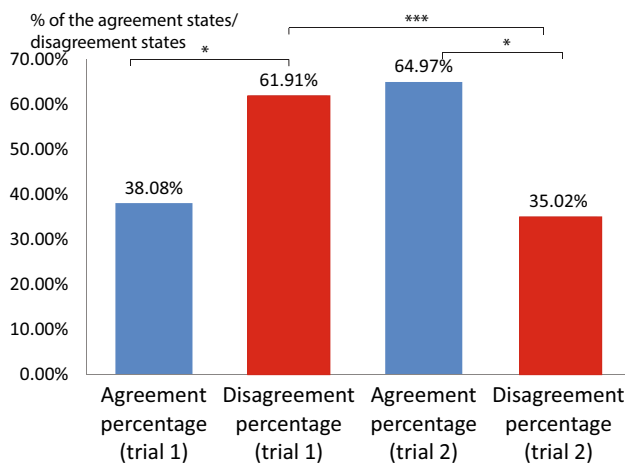
Participants confirmed through most of their answers that they wanted to simplify the input for the robot. One of the participants indicated: "...I was confused initially but as time goes by I start to compose simple redundant input to get the regular intended output...", another participant confirmed that: "...The robot is smart, while there are some repetitive combinations between my knocking and the chosen actions and thus I started to track the best knocking that led to the convergence to stable combinations. It has to be slow modulated knocking..."

### 4.3.2 Evaluation of an Interaction's Scenario

To investigate the different stages of pattern emergence, we tried to explore the flow of the interactions. A sample flow of pair 15 is depicted in Fig. 5 where in grey we have the knocking while the corresponding action is represented by the colorful line.

Figure 5 shows that most of the time when the controller received a knocking pattern, the latter waited a small period of time in order to choose the behavior that he thought the most appropriate for the received knocking pattern. As an example, we could see that when the knocker emitted a new knocking pattern, the controller stopped for a while to think before attributing the behavior according to his own assumptions (all red circles). Consequently, if the knocker was satisfied with the controller's choice he would not knock, otherwise the knocker would knock again before 2 s (based on the knocker's reaction time (KRT) distribution: [mean: 1.93; sd: 0.12] seconds) elapsed in order to implicitly indicate to the controller that he must change direction again. Some exploration was adopted [55–57 s] when encountering a new pattern. In fact, the controller chose the correct behavior for the new pattern (1 knock) even if the pattern was encountered for the first time. Interestingly, if we track the mapping of the knocking patterns and the robot's behavior, we find that in some occasions the rule was maintained for several times such as for the

<sup>2</sup> Continuous-knocking was related to the presence of contiguous disagreements about the shared rules, and we defined a disagreement state in the Sect. 4.3.2.



**Fig. 6** The percentage of agreement and disagreement states during the experiment 1

pattern (2 knocks) when it was associated with the *left* behavior ([15–16 s], [45–47 s] and [79–81 s]), and the (3 knocks) pattern when it was associated with *right* behavior ([30–32 s] and [102–104 s]). However, at other times there was a change in the rule combination such as when (1 knock) was initially associated with the *forward* behavior ([55–57 s]) and later with the *back* behavior ([114–116 s]).

When the controller and the knocker shared the same assumption about one of the knocking pattern-robot's behavior combinations that was maintained over time we call that state an "agreement state". If the combination knocking pattern-robot's behavior changed over time we call that state a "state of disagreement". The participants were then blending incrementally in a trial-and-error process the agreement and disagreement states in order to establish shared rules organizing the communication.

#### 4.3.3 Adaptation's Evaluation Based on the Agreement and Disagreement States Comparison

To evaluate the different pair interactions' convergence toward a stable protocol, we counted the number of the agreement and the disagreement states based on the coded data for both trials and all the pairs. We computed the  $t$  test between the agreement and the disagreement states of the trial 1. The results were significant with  $t = 2.242$ ,  $d.f. = 14$ ,  $pvalue = 0.033 < 0.05$ ). Figure 6 shows the percentage of the agreement states (blue color) as well as the percentage of the disagreement states (red color) during the trials 1 and 2.<sup>3</sup> By examining the percentage of the agreement and disagreement states of the trial 1, we deduced that during the trial

1, disagreements (61.91 %) were more significantly frequent than agreements (38.08 %) (Fig. 6).

We computed the  $t$  test between the agreement and the disagreement states of the trial 2. The results were also significant with ( $t = 2.067$ ,  $d.f. = 14$ ,  $pvalue = 0.048 < 0.05$ ). By displaying the percentage of the agreement and disagreement states of the trial 2, we deduced that during trial 2, agreement states (64.97 %) were more significantly frequent than disagreement states (35.02 %) (Fig. 6). Finally, we calculated the  $t$  test between the trial 1 and 2 disagreement states. The results were statistically significant with ( $t = 2.948$ ,  $d.f. = 14$ ,  $pvalue = 0.006 < 0.01$ ). By displaying the percentage of the trial 1 disagreement states (61.91 %) and the percentage of the trial 2 disagreement states (35.02 %), we deduced that during the trial 1, disagreement states (61.91 %) were significantly more frequent than disagreement states of the trial 2 (35.02 %) (Fig. 6).

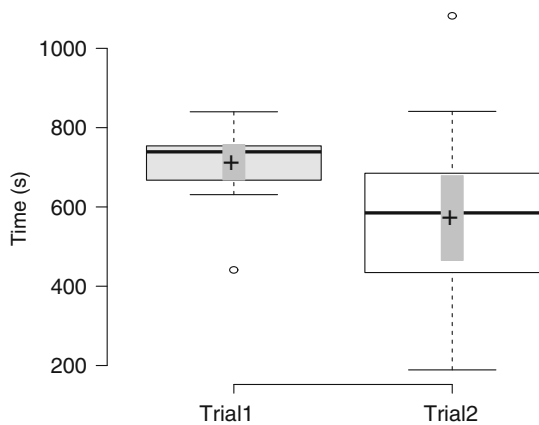
#### 4.3.4 Comparison of the Task Completion Time in Trial 1 and Trial 2

The time to reach the different sub-goals was estimated based on the videos. The distribution of the task completion time datasets of the trial 1 (first boxplot in grey) and 2 (second boxplot in white) are represented in Fig. 7. Results showed that there is a decrease on the task completion time during the trial 2 (Fig. 7). A  $t$  test showed that there was a statistically significant difference between the task completion time of the trial 1 and 2 with ( $t = 2.143$ ,  $d.f. = 14$ ,  $pvalue = 0.041 < 0.05$ ). This highlighted that although during the second trial we changed the configuration by changing the point coordinates (which may imply that the pairs would have to adapt to each other again in a new context), the pairs succeeded on achieving the task more quickly during the trial 2.

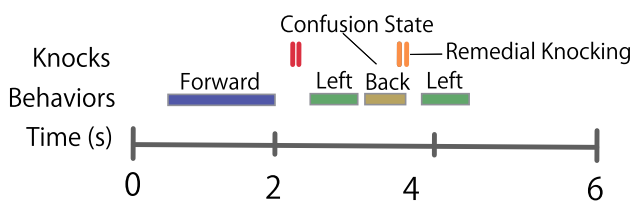
#### 4.3.5 Cooperative Communication for the Task Achievement

To study the incremental adaptation to each others' behaviors, we calculated the number of confusion states and the remedial knocking states. Figure 8 helps to understand the meaning of these two practices. As you may see in the Fig. 8, the robot executed initially the forward behavior, and when the controller detected that he received a knocking pattern (2 knocks in red), he picked left as a new behavior. Within a few milliseconds, we can see that the controller changed the behavior to back. We called such situation a state of confusion since the controller changed the behavior after recently choosing an action and without being prompt by any knocking. As a response the knocker, composed of a remedial knocking pattern (2 knocks in orange: the same previous knocking pattern) so as to help the controller overcome the situation by resuming with the previous executed behavior.

<sup>3</sup> As an example, the percentage of agreement states = number of agreement states/(number of agreement states + number of disagreement states).



**Fig. 7** Task completion time distributions during trials 1 and 2 (experiment 1)



**Fig. 8** A scenario showing an example of a state of confusion and a remedial knocking pattern

The presence of states of confusion indicated that the controller tried to establish the rules of communication but may go through some confusing states. Consequently, the knocker also tried to adapt to the controller's state of confusion by composing a remedial knocking pattern.

We calculated the Pearson correlation between the confusion states and the remedial knocking of the first and second trials. The value of  $R$  during the trial 1 is 0.6149 with ( $p$  Value from Pearson ( $R$ ) = 0.014; d.f. = 13; The result was significant at  $p < 0.05$ ) and during the trial 2 with  $R$  value ( $p$  value from Pearson ( $R$ ) = 0.00019. d.f. = 13; The result was significant at  $p < 0.01$ ). This meant that there was a tendency for high confusion states values went with high remedial knocking values (and vice versa). Consequently, if the confusion states occurred more frequently, the knocker would try to cooperate most of the time with the controller in order to maintain the rules which he thought they were shared between him and the controller.

#### 4.3.6 Communication Protocol Analysis

The subjective results and the previously discussed objective analysis showed that there was a cooperation between the knockers and the controllers in order to adapt to each other and establish communication protocols. To visualize the emergent communication protocols, we used the correspondence analysis. Correspondence analysis is an exploratory

technique that helps analyzing the two-way frequency cross-tabulation tables containing measures of correspondence between the knocking patterns and controllers' interpretations of these patterns. The results provide information which is similar in nature to those produced by Factor Analysis techniques, and they allow us to explore the structure of our two variables (knocking patterns and controllers' interpretations to these patterns) by means of derived dimensions  $F_1, F_2, \dots, F_n$ .

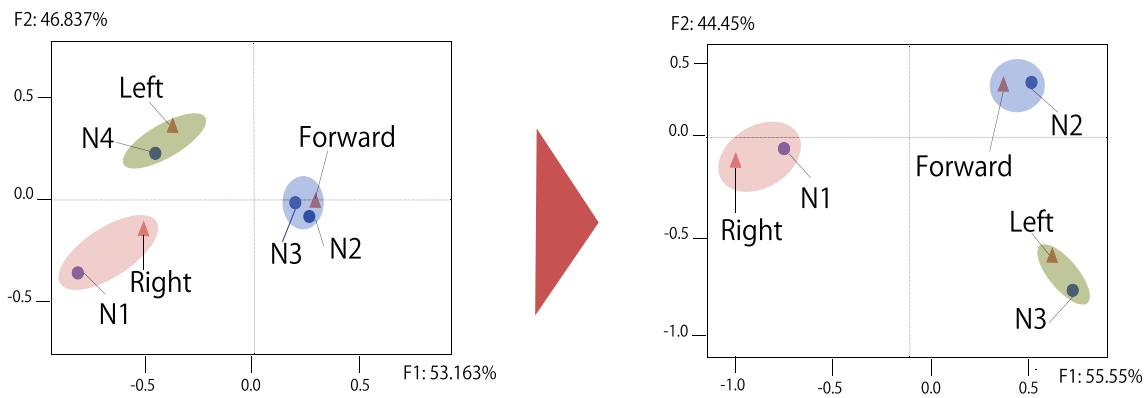
To understand how the dimensions are derived, we need to consider the Chi-square statistic for two-way tables like in our example (knocking patterns and the related controllers' interpretations of these behaviors). Any deviations from the expected values (expected under the hypothesis of complete independence of the knocking patterns and the controllers' interpretations) would contribute to the overall Chi-square. Thus, another way of looking at correspondence analysis is to consider it a method for decomposing the overall Chi-square statistic (or Inertia = Chi-square/Total N) by identifying a small number of dimensions in which the deviations from the expected values can be represented. This is similar to the goal of Factor Analysis, where the total variance is decomposed, so as to arrive to a lower-dimensional representation of the variables that allow us to reconstruct most of the variance matrix of variables.

For a matter of illustration, we chose to depict the associations between knocking patterns and controllers' interpretations of pair 9 (Fig. 9). It appeared that based on the two-way frequency table associating the pair 9's knocking patterns to the controllers' interpretations, we had two derived dimensions. With a single dimension  $F_1$  (trial 1:  $F_1 = 53.163\%$  and trial 2:  $F_1 = 55.550\%$ ) as we represented in Fig. 9 53.163% in trial 1 and 55.550% in trial 2 of the inertia can be "explained," that is, the relative frequency values can be reconstructed from a single dimension and reproduced 53.163% of the total chi-square value (and, thus, of the inertia) for the case of our two-way table. Two dimensions allowed us to explain 100% of the data with  $F_2$  (trial 1:  $F_2 = 46.837\%$  and trial 2:  $F_2 = 44.450\%$  (Fig. 9).

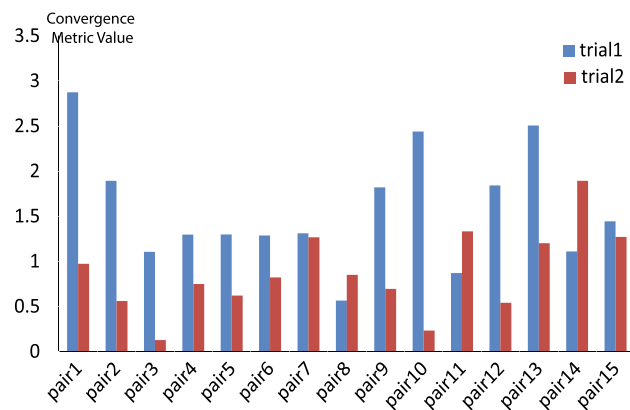
Based on the (Fig. 9 (right)), we remarked that *right* behavior is materialized by 1 knock, *forward* was represented by 2 and 3 knocks, and *left* by 4 knocks. In the second trial (Fig. 9 (left)), the protocol was slightly ameliorated where we could see a clear categorization of *forward* that was represented by only 2 knocks while *left* was represented by 3 knocks and *right* was always represented by 1 knock.

#### 4.3.7 Performance Evaluation Based on the Convergence Metric Values

We wanted to explore whether there was a statistically significant difference between the convergence level to a stable communication protocol during trials 1 and 2. For this pur-



**Fig. 9** Correspondence analysis for both trials for the pair 9 (*Left* first trial, *Right* second trial) in the first experiment where  $N_i$  represents the knocking patterns; e.g.,  $N_2$  represents 2 knocks



**Fig. 10** The convergence metric values of the first and second trial (experiment 1)

pose and 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. 9) and the different patterns (blue circles as presented in the Fig. 9). Thus, for each behavior we calculated the  $n$  possible Euclidean distances (assuming that we have  $n$  possible patterns). After that, we picked for each behavior the most minimal distance. We summed up the 4 most minimal distances and the resultant value afforded information about the most minimal distance that the pair knocker-controller achieved to form stable rules. We called this value the convergence metric which evaluated the system's performance. We repeated the same procedure for the 15 pairs and for the two trials.

We computed the  $t$  test between the convergence metric values of the trial 1 and 2 which revealed significant differences:  $t = 2.503$ ,  $d.f. = 14$ ,  $p \text{ value} = 0.018 < 0.05$ . We displayed the results of the trial 1 and 2 convergence metric values where in blue we had the convergence metric values of the first trial and in red the convergence metric values of the second trial (Fig. 10). Figure 10 showed that 12 out of the 15 pairs (80 %) succeeded in reducing the convergence metric

values during the second trial, indicating that the pairs were closer to the convergence to stable protocols' formation.

#### 4.3.8 Consistent Protocol Formation Evaluation

To statistically measure the relationship between the knocking patterns and the different behaviors, we computed the test of independence (Chi-square) between the knocking patterns and different behaviors as well as the Cramer's V-values. Tables 1 and 2 exhibited the results of the first and second trials for the different participants. Based on the Table 1 we deduce that 7 out of 15 pairs (46.66 % of the pairs) succeeded in establishing a stable communication protocol during trial 1, where the chi-square values were significant for 7 pairs, with a Cramer's V-values ranging from 0.331 to 0.823, indicating a strong relationship between the knocking patterns and the controller's interpretations of these patterns. We noticed that during the trial 2 (Table 2), the number of pairs that succeeded in establishing a communication protocol increased to 11 out of 15 pairs (73.3 % of the pairs) with high Cramer V-values, indicating that there was also a strong relationship between the knocking patterns and the controller's interpretations of these patterns. Consequently, we deduced that gradually there was a strong relationship between the knocking patterns and the controller's interpretations of these patterns.

#### 4.4 Discussion

We started with a H–H experiment to evaluate the knockers' and controllers' adopted behaviors that led to the emergence of communication protocols. Understanding both parties' strategies facilitated for us the tailoring of a control model that could be integrated into the robot and may lead to a similar flexible communication protocol formation.



**Table 1** The test of independence (Chi-Square) between the knocking patterns and the robot's behaviors as well as the Cramer's V (CV) values of the trial 1 (experiment 1)

Pairs	Chi-square	CV
Pair1	$\chi^2 = 55.515$ , d.f. = 18, ***p value < 0.01	0.446
Pair2	$\chi^2 = 21.978$ , d.f. = 18, p value = 0.233	–
Pair3	$\chi^2 = 12.394$ , d.f. = 9, p value = 0.192	–
Pair4	$\chi^2 = 18.6565$ , d.f. = 12, p value = 0.097	–
Pair5	$\chi^2 = 9.828$ , d.f. = 12, p value = 0.631	–
Pair6	$\chi^2 = 26.345$ , d.f. = 15, **p value = 0.035	0.331
Pair7	$\chi^2 = 10.222$ , d.f. = 3, **p value = 0.017	0.698
Pair8	$\chi^2 = 2.475$ , d.f. = 6, p value = 0.871	–
Pair9	$\chi^2 = 12.634$ , d.f. = 8, p value = 0.125	–
Pair10	$\chi^2 = 50.068$ , d.f. = 18, ***p value < 0.01	0.590
Pair11	$\chi^2 = 5.528$ , d.f. = 9, p value = 0.786	0.166
Pair12	$\chi^2 = 19.307$ , d.f. = 9, **p value = 0.023	0.529
Pair13	$\chi^2 = 9.828$ , d.f. = 12, p value = 0.631	–
Pair14	$\chi^2 = 14.215$ , d.f. = 2, ***p value < 0.01	0.823
Pair15	$\chi^2 = 17.071$ , d.f. = 18, p value = 0.518	–

\* confidence interval is 90%

\*\* confidence interval is 95%

\*\*\* confidence interval is 99%

**Table 2** The test of independence (Chi-Square) between the knocking patterns and the robot's behaviors as well as the Cramer's V (CV) values of the trial 2 (experiment 1)

Pairs	Chi-square	CV
Pair1	$\chi^2 = 31.640$ , d.f. = 12, ***p value = 0.002	0.568
Pair2	$\chi^2 = 28.119$ , d.f. = 8, ***p value < 0.01	0.404
Pair3	$\chi^2 = 0.877$ , d.f. = 2, p value = 0.645	–
Pair4	$\chi^2 = 10.297$ , d.f. = 12, p value = 0.590	–
Pair5	$\chi^2 = 4.422$ , d.f. = 4, p value = 0.352	–
Pair6	$\chi^2 = 18.033$ , d.f. = 8, **p value = 0.021	0.308
Pair7	$\chi^2 = 4.6$ , d.f. = 4, p value = 0.331	–
Pair8	$\chi^2 = 8$ , d.f. = 2, **p value = 0.018	0.9
Pair9	$\chi^2 = 12.036$ , d.f. = 4, **p value = 0.017	0.501
Pair10	$\chi^2 = 26.813$ , d.f. = 12, ***p value = 0.008	0.829
Pair11	$\chi^2 = 22.610$ , d.f. = 6, ***p value < 0.01	0.408
Pair12	$\chi^2 = 17.714$ , d.f. = 4, ***p value < 0.01	0.859
Pair13	$\chi^2 = 23.517$ , d.f. = 6, ***p value < 0.01	0.637
Pair14	$\chi^2 = 34.476$ , d.f. = 15, ***p value = 0.003	0.384
Pair15	$\chi^2 = 32.799$ , d.f. = 9, ***p value < 0.01	0.594

\* confidence interval is 90%

\*\* confidence interval is 95%

\*\*\* confidence interval is 99%

#### 4.4.1 Evaluation of the Command-Like and the Continuous-Knocking Patterns Based on the Videos

Based on the coded videos, we remarked that the communication was patterned. It was crucial for the pairs to scaling

the problematic to a small number of entry states (1 knock, 2 knocks, etc.). The use of continuous-knocking was a way to overcome the contiguous disagreements. By examining the percentages and the  $t$  test results, we remarked that there were potential trend to use the command-like more frequently during the trials 1 and 2. The objective of the pairs was to minimize the expected infinite horizon of states to a small number of states in order to easily track each of the states successful combinations with the controller's interpretations of these patterns. Thus, during the communication protocol establishment, users restricted the number of states to facilitate inferring the communication rules (even if we do not impose for the human a way of an interaction with the minimally designed robot).

#### 4.4.2 Evaluation of Interaction Scenarios

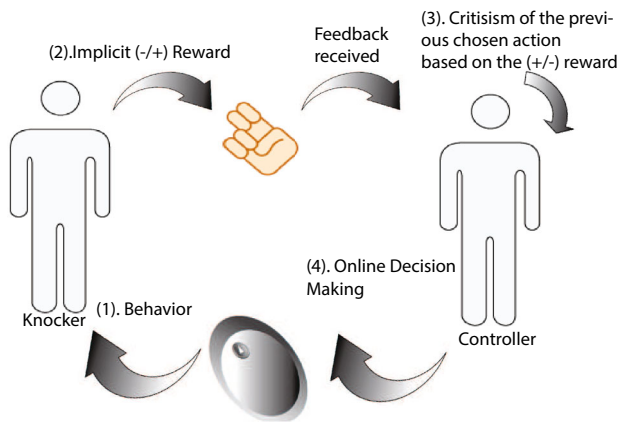
Interrupting the controller's executed action was associated with the presence of knocks (negative reward for the controller), while no knocks implied the controller was doing well (positive reward). Based on this trial-and error process, the pairs were incrementally establishing communication protocols by mainly going through multiple agreements and disagreements about the shared rules as the Fig. 5 showed.

#### 4.4.3 Adaptation Evaluation Based on a Comparison of Agreement and Disagreement States

Based on the  $t$  test results and Fig. 6 we concluded that disagreement states decreased significantly from trial 1 to trial 2. We deduce also that the agreement states were significantly inferior than the disagreement states during the trial 1 in addition to the fact that the same thing occurred during the trial 2. These results suggested that even though the pairs normally had to adapt again to each other during trial 2 in order to share the communication rules (since we had a new configuration with different checkpoint coordinates), there was a better convergence during the trial 2. We deduced implicitly that there were some first trial rules which facilitated the convergence during trial 1 and that were transferred to trial 2. As an example, we saw in Fig. 9 that the rule combining the behavior *right* with the pattern 1 knock was maintained during the trial 2.

#### 4.4.4 Cooperative Communication for the Task Achievement

By examining the data and Pearson correlation test values, we maintained that there was a significant correlation between the confusion states and the remedial knocking. On the one hand, this indicated that the controller was trying to maintain stable rules that he thought organizing the interaction. On the other hand, this indicated that the knocker cooperated with



**Fig. 11** Cooperative behavior between the controller and the knocker during the communication protocol formation

the controller in order to altogether shape a stable protocol of communication (Fig. 11).

During the interaction, the controller tried to establish the communication rules by choosing the behavior that was previously more frequently (greedy policy<sup>4</sup>) associated with the received knocking pattern. He also auto-criticized his strategy based on his own assumptions and this was proved by the presence of some confusion states. He refined his assumptions according to the new rules that he imagined shared with the knocker. Finally, he chose a new behavior. His choice might lead to an agreement or a disagreement state. These insights led us to think about a model which integrated two components during the communication protocol formation, one related to the action choice and the other to the criticism of the executed action.

#### 4.4.5 Performance Evaluation

Shared rules formation led to a significant decrease (as the  $t$  tests and Fig. 7 shows) of the task completion time during trial 2. We also noticed that there was a decrease in the convergence metric values during trial 2 (Fig. 10). We deduced then that the pairs were growing closer to the stable communication protocol formation. This decrease was revealed by the elaboration of clear rules. As an example, pair 9 in Fig. 9 succeeded on associating for the *forward* behavior 2 knocks during trial 2 after being confused during trial 1 between two patterns (2 knocks, 3 knocks). By applying the chi-square and Cramer's V (Tables 1 and 2) tests, which evaluated the relationship between the knocking patterns and the controller's

<sup>4</sup> It consists of choosing the most frequent behavior that was previously associated to the same number of knocks previously received and led to an agreement state; e.g.,: choosing the left behavior when we have 3 knocks led most probably to an agreement state while choosing back may have led to a disagreement because it has led less frequently used for an agreement state based on the previous interactions.

interpretations of these patterns, we found that the number of pairs showed a statistically significant relationship between the patterns, and that the behaviors increased from 7 out of 15 pairs (46.6 %) to 11 out of 15 pairs (73.3 %), indicating our scenario helped the users to acquire the meaning of the different emergent patterns and form communication protocols incrementally based on the previous interactions.

## 5 Modeling the Architecture of the Robot

### 5.1 Insights from the Human–Human Experiment

We seek to enable non-expert users to shape a communication protocol with a minimally designed robot. The fact that the robot used a novel minimal communication channel caused some confusion for the human. It required adaptation from him in order to understand how to provide the most convenient input for the robot while guaranteeing the intended output. In this vein, we noticed that people aggregated a small number of redundant patterns (such as 1 knock, 2 knocks, etc.) in order to guarantee a systematized output (e.g.: 1 knock for the left direction, 3 knocks for the back direction, etc.). For each instance of interaction, the controller chose an action based on the received knocking while he tried to affect for the gathered pattern the most frequently successful action that was tested previously. Afterward, the knocker would judge the controller's choice. If the chosen action did not converge with the knocker's desired direction, the knocker would compose another knocking pattern in 2 s (approximated value based on the KRT distribution) indicating that the controller's choice was incorrect. Since the controller tried to track the best combinations between the knocking and the robot's action, any new knocking that disrupted the execution of the newly chosen action (action interrupted before 2 s elapse) would lead to a disagreement with the controller's assumptions about the knocking pattern-action combinations. However, if no knocking was received the action is correct and consolidated the controller's assumptions about the knocking pattern- action combinations. We also found there were times that when the controller chose the action, he got confused and changed the action without being prompt by any knocking. This indicated that the controller chose the action but also criticized his choices. The knocker sometimes detected the controller's confusion which confirmed again that there were rules shared between both parties. The knocker then tried to cooperate by composing the same previous knocking pattern, indicating that the controller (or the robot here since the knocker did not know that a controller wizarded the robot) had to return to the other recently executed action.

In parallel to our insights, Reinforcement Learning (RL) is “learning through a *trial-and-error process* how to *associate*

states to actions in order to maximize a numerical reward. The learner has to discover which actions yield the most rewarding state using the greedy policy and finally reach a meaningful state-action combinations” [39]. Therefore, if we suppose that:

- Command-like patterns referred to the states in the RL while we had different states such as 1 knock state, 2 knocks state, etc.
- The different robot’s behaviors were the actions for the RL (4 actions: right, left, back, forward).
- The controller’s choice that consisted of choosing the most frequently used action previously tested corresponded to the greedy action chosen based on the greedy policy.
- The presence of knocking after the robot started the execution of the chosen action and before 2 s elapsed is the negative reward.
- The absence of knocking (for 2 s) after the robot started the execution of the chosen action was the positive reward.
- The fact that the interaction went through agreement and disagreement states indicated that the adaptation corresponded to a sequential trial-and-error process just like in the RL.
- Both parties established different combinations of (knocking pattern - controller’s interpretations) corresponded to the (state - action) cartography that emerged during a RL process.

We may deduce then that RL algorithms fitted to our problematic adequately. In addition, the decision making should be in a *real time*<sup>5</sup> because we obtained different communication protocols for the different pairs, indicating that any hand-programming of a possibly supposed same protocol adopted by all the pairs would fail. We should therefore reduce the scope of useful RL algorithms to only the online RL algorithms. Finally, and based on the first experiment’s insights, we found that the controllers at times were auto-criticizing their strategies. This made us think about the actor-critic as an online RL algorithm that fitted to our problematic. An actor-critic algorithm integrates a critic and an actor. The critic uses a temporal difference learning (TD) to criticize the action that has been chosen, and the actor is updated based on the information provided by the critic [40]. Incrementally, the actor chooses the greedy action while the critic observes the relevance of the actor’s choice after receiving the feedback. The relevance of an executed action

<sup>5</sup> Real time: Because the communication patterns emerge in a sequential fashion and we remarked that communication protocols were personalized to the pairs, any attempt to integrate a batch learning method to the robot’s architecture could not succeed in establishing the same customized protocols that we had seen in the first experiment, and that it is why we needed an online machine learning method. An online machine learning method gathers the data and learns incrementally.

is materialized in our case, by the presence (negative reward) or the absence (positive reward) of the knocking and leads to an agreement or a disagreement state. The proposed actor-critic model should lead to similar performance (decrease in the disagreement states, the task completion time and the convergence metric values) as in the H–H experiment. It should also guarantee transfer learning of the shared rules during trial 2 (while some combinations knocking-action of the first trial’s communication protocol should be used during trial 2) so that stable communication protocols emerge.

## 5.2 Actor-Critic Algorithm

### 5.2.1 Actor Learning

Each knocking pattern (state) has its own distribution.  $X(s_t) \approx N(\mu_{X(s_t)}, \sigma_{X(s_t)})$  where  $X(s_t)$  is defined as the number of knocks,  $\mu_{X(s_t)}$  and  $\sigma_{X(s_t)}$  are the mean value and the variance while  $\Pi(s_t)$  is the corresponding probabilistic policy associated to  $X(s_t)$ . We also assigned a distribution for the continuous-knocking pattern<sup>6</sup> that also helps in learning what behavior should be chosen once a continuous knocking is received by the robot. Initially, the action is chosen according to the probabilistic policy  $\Pi(s_t)$ . The state of the interaction changes to the state  $s_{t+1}$  according to the user’s knocking presence (disagreement)/absence (agreement). If the human interrupts the robot’s behavior execution before 2 s<sup>7</sup> by composing a new knocking pattern, we have a disagreement state about the previous pattern’s meaning (which was received from about 2 s). Consequently, the action that is chosen based on the probabilistic distribution in an attempt of exploiting the emerged knowledge failed. The actor updates the probabilistic policy  $\Pi(s_t)_{nbknocks}$  and chooses the action henceforth (until we meet an agreement state as a closure for the current pattern meaning’s decoding process) by a pure exploration based on the equation

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

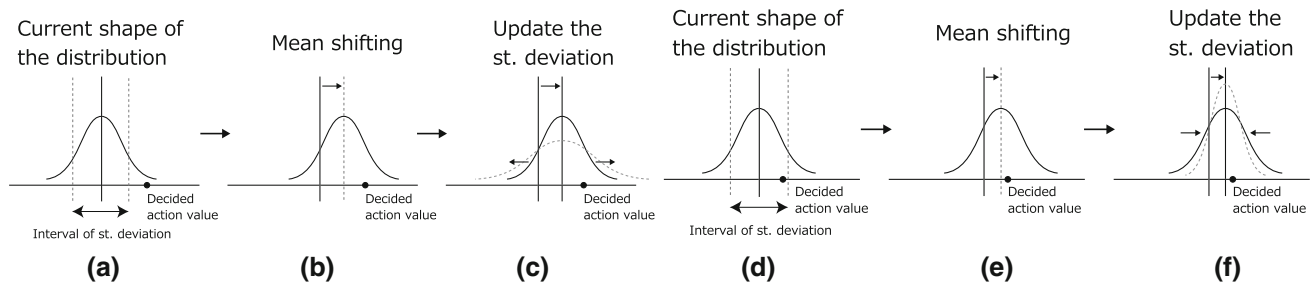
where *rnd1* and *rnd2* are random equations that are designed to bring the values of the action between 0 and 3.

### 5.2.2 Critic Learning

After each action selection, the critic evaluates the new state to determine whether things has gone better or worse than expected. The action is evaluated based on the presence or absence of knocking (positive or negative reward). This evaluation process is called the temporal difference (TD) error.

<sup>6</sup> We suppose that a knocking pattern that involves a number of knocks superior than 4 knocks.

<sup>7</sup> We calculated approximately the value based on a pilot study.



**Fig. 12** Figure shows the re-adjustment procedure of state parameters; (1) the decided action value is outside of the standard deviation interval (**a, b, c**): **a** current shape of the state distribution and decided action value, **b** mean shifting has started, and **c** the state parameters are updated and a new shape of the distribution is established; (2) Re-adjustment

procedure of state parameters when the decided action value is inside of the standard deviation interval (**d, e, f**): **d** the current shape of the state distribution and decided action value, **e** indication the shifting has started, and **f** the state parameters are updated and a new shape of the distribution is established

The critic calculates the TD error ( $\delta_t$ ) as the reinforcement signal for the critic and the actor where

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

with  $\gamma$  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 the equation:

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

where  $0 \leq \alpha \leq 1$  is the learning rate. A positive TD error indicates that the tendency to select  $a_t$  when receiving the  $i$ -th current pattern should be strengthened for the future. A negative TD error indicates that the tendency to use that action with the gathered current pattern should be weakened, and in our case we weaken the possibility to choose the action  $a_t$  for the  $i$ -th current received pattern. As long as the current pattern meaning's decoding is not achieved (exploration phase), (exploration phase), the critic will each time it encounters a disagreement state updates  $\delta_t$ ,  $V(s_t)$  and the distribution  $N(\mu_{X(s_t)}, \sigma_{X(s_t)})$ :

$$\mu_{X(s_t)} = \frac{\mu_{X(s_t)} + A(s_t)}{2} \quad (4)$$

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

The modification during the update process helps to readjust the shared rules according to the previous interactions and assigns the most frequently correct behavior for the  $i$ th current pattern received.

The idea here is to attempt to obtain the correct action inside the interval that represents the possible actions which should be executed when gathering the  $i$ th pattern. The chosen behavior can be inside (when the action is chosen based on the probabilistic policy) or outside of the distribution (when the previously chosen action fails). If the behavior was outside of the distribution of the pattern, this means that the

human has changed the rule concerning the  $i$ th pattern. We operate in this case the mean shifting and the variance enlarging to recuperate the value inside the distribution (Fig. 12c). As the decided action value is already inside the standard deviation interval and the TD was positive (Fig. 12d), then our approach attempts to shift the mean value (Fig. 12e) toward the action value while minimizing the standard deviation (Fig. 8f). Shifting occurs when TD is positive by choosing the correct behavior as a part or the center of the distribution. In fact, if the action was outside the distribution then we assume that we are not sure that it is the new sustained rule (we only know that it was correct for one time) so we recuperate it inside. If that same action was combined with the same knocking pattern to which it was previously associated ( $i$ th pattern), it becomes the mean because the robot is more certain it is the new rule of the  $i$ th pattern.

## 6 Experiment 2: Human–Robot Interaction

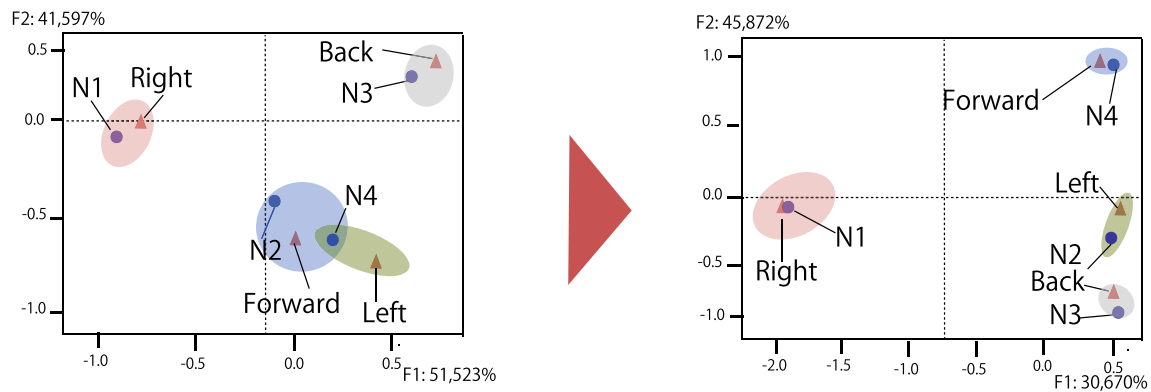
Through this experiment, we tried to validate the robot's implemented architecture and verify whether the human and the robot can establish stable communication protocol.

### 6.1 Experimental Protocol

Each time we had a new participant, the instructor told him that he had to lead the robot to different checkpoints marked on the table before reaching the final goal point using knocking (Fig. 2). We had two different configurations for the two trials of the experiment 2. We asked the participants to describe their experience when they finished the task.

In the first trial (Fig. 2 (left)), we expected the knocker to cooperate with the robot to invent his own protocol of communication by focusing on the most successful patterns that led mostly to agreement states just like in the first experiment. Meanwhile, we expected that the robot would focus on the rules' acquisition. The robot has to keep on guessing the





**Fig. 13** Correspondence analysis for both trials for the participant 3 (*Left* first trial, *Right* second trial) in the first experiment where  $N_i$  represents the knocking patterns, e.g.: N2 represents 2 knocks

most possibly correct behavior that must be combined with the right knocking pattern. It has also to refresh its assumptions in real time so that a stable communication protocol could be finally established. In the second trial, we assumed that the communication would become smoother as in the second trial of the first experiment. In this experiment, we had 10 participants (6 male, 4 female) ranging in age from 20 to 24 years old.

## 6.2 Results

After the experiment was finished, we tried to analyze the interaction scenarios in order to verify whether a communication protocol was established between the knockers' knocking patterns and the chosen actions.

We analyzed the video data by annotating with a video annotation tool called ELAN. Two coders, one of the authors and one other volunteer analyzed the behavioral data using the same coding rules for the first and the second trials. We calculated the average of Cohen's kappa from six arbitrarily selected videos in order to investigate the reliability. As a result, we confirmed that there was a reliability with  $\kappa = 0.819$ .

### 6.2.1 Evaluation of the Command-Like and the Continuous-Knocking Patterns Based on the Videos

Based on the coded data, we counted the number of continuous-knocking pattern and the number of command-like pattern for all the participants and for the two trials to see whether participants had tendencies to use the command-like mode just like in the experiment 1. We discovered the participants were mainly using the command-like patterns with percentages (trial 1: 91.14 % of the patterns were command-like) and (trial 2: 95.46 % of the patterns were command-like). We conducted 2  $t$  tests to verify whether there was a significant difference between the 2 patterns

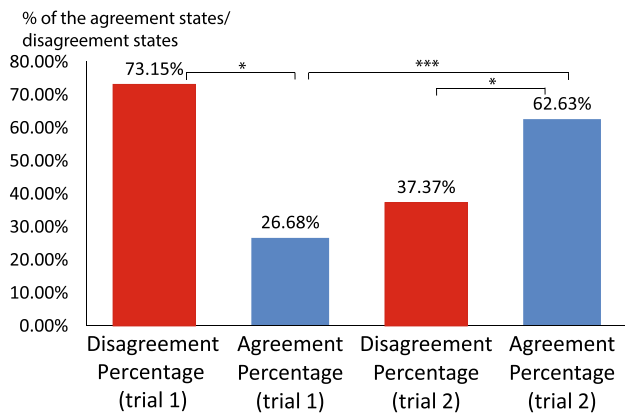
usage: trial 1: ( $t = 4.596$ ,  $d.f. = 9$ ,  $p \text{ value} < 0.01$ ), and trial 2: ( $t = 7.486$ ,  $d.f. = 9$ ,  $p \text{ value} < 0.01$ ). As a result, we found a significant effect for usage of the command-like patterns during both trials, while a new state in the interaction cycle corresponded most of the time to a command-like pattern just as in the first experiment.

Participants confirmed also the fact that they need to use the simple command-like mode while one of the participants said: "...I tried to knock slowly, to focus on the most useful knocking that will lead the robot to execute the right direction...", another one said: "...It is clear that I have to pay attention to the knocking and then I tried to affect 1, 2 knocks, etc. to facilitate remembering of the most convenient knocks..."

### 6.2.2 Communication Protocol Analysis

For a matter of illustration, we had chosen to depict the associations between knocking patterns and robot's chosen behaviors of the participant 3 based on 2 dimensions for the trial 1: (F1 = 51.523 %–F2 = 41.597 %) and trial 2: (F1 = 45.872 %–F2 = 30.670 %) <sup>8</sup>, just as in the first experiment (Fig. 13). Based on the (Fig. 13 (right)), we maintained that *right* behavior was materialized by 1 knock, *forward* and represented by 2 and 4 knocks, *left* by 4 knocks and *back* by 3 knocks. In the second trial (Fig. 13 (left)), the protocol is slightly ameliorated where we can see a clear categorization of *forward* that is represented by only 4 knocks, while *left* is represented by 2 knocks, *right* is always represented by 1 knock, and *back* by 3 knocks.

<sup>8</sup> Here we had actually 3 dimensions for each of the trials F1, F2 and F3 to reach 100 %, but the highest possible representation in 2 dimensions consisted of choosing the F1 and F2 more so than either F1 and F3 or F2 and F3.



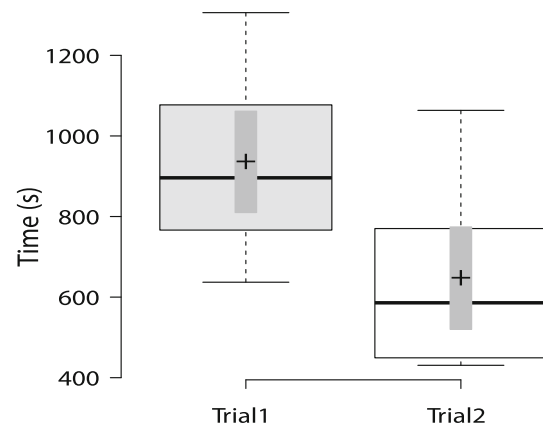
**Fig. 14** The agreement and disagreement percentage during trials 1 and 2 (experiment 2)

### 6.2.3 Adaptation Evaluation Based on the Agreement and Disagreement States Comparison

We counted the number of agreements and disagreements during trials 1 and 2 and for all the participants. A  $t$  test showed that there were significant differences between the number of agreements and the number of disagreements usage during the trial 1 with a value:  $t = 2.37$ ,  $d.f. = 9$ ,  $p \text{ value} = 0.028 < 0.05$ . We displayed the percentage of the first trial's agreements and disagreements in the Fig. 14, where in blue we have the percentage of the agreements and in red we have the disagreements during the trial 1 and 2.<sup>9</sup> Based on the Fig. 14, we noticed also that the number of disagreement states (73.15 %) was higher than the number of agreement states (26.68 %) during the first trial. A  $t$  test showed that there were statistically significant differences between the number of agreements and disagreements during the trial 1 with a value  $t = 2.37$ ,  $d.f. = 9$ ,  $p \text{ value} = 0.028 < 0.05$ .

Based on Fig. 14, we also noticed that the number of agreements exceeded the number of disagreements with a percentage value respectively 62.63 and 37.37 % during the trial 2. A  $t$  test between the agreement and disagreement states during trial 2 showed that this excess was statistically significant with ( $t$  test:  $t = 2.108$ ,  $d.f. = 9$ ,  $p \text{ value} = 0.049 < 0.05$ ). Finally, by calculating the  $t$  test between the number of agreements of the first trial and the second trial, we obtained the above value ( $t = 5.359$ ,  $d.f. = 9$ ,  $p \text{ value} < 0.01$ ). We can therefore conclude then that even though the second trial involved a configuration with new checkpoints, there were a higher number of agreements during trial 2. This implies that a transfer of learning occurred and facilitated the formation of a communication protocol during the trial 2 just like in the second trial of the first experiment.

<sup>9</sup> The percentages were calculated based on the same formula used during the experiment 1.



**Fig. 15** Task completion time distributions during trial 1 and 2 (experiment 2)

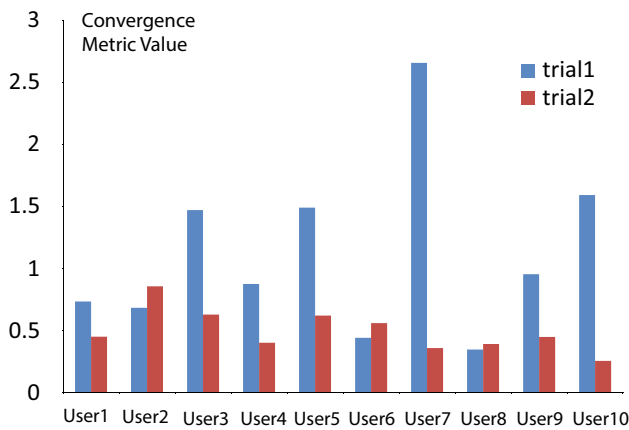
### 6.2.4 Comparison of the Task Completion Time of the Trial 1 and 2

The distribution of the task completion time datasets during the trial 1 (first boxplot in grey) and 2 (second white boxplot) were represented in Fig. 15. Figure 15 shows that there was a decrease in the task completion time during trial 2. We applied a two-tailed  $t$  test to verify whether there were statistically significant differences between the task completion time of the first and second trial. The results were significant with  $t$  test value: ( $t = 2.959$ ,  $d.f. = 9$ ,  $p \text{ value} = 0.008 < 0.01$ ).

### 6.2.5 Performance Evaluation Based on the Convergence Metric Values

We wanted to explore whether there was a statistically significant difference between the system's performance during trials 1 and 2. For this purpose and 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. 13) and the different patterns (blue circles as presented in the Fig. 13). Thus, for each behavior we calculated the  $n$  possible Euclidean distances (assuming that we have  $n$  possible patterns). After that, we picked for each behavior the most minimal distance. We summed up the 4 most minimal distances and the resultant value afforded information about the most minimal distance that the pair knocker-controller achieved to form stable rules. We called this value the convergence metric which evaluated the system's performance. We repeated the same procedure for the 10 participants and for the two trials.

As in the first experiment, we display the results of trials 1 and 2 convergence metric values, where the convergence metric values of the first trial are shown in blue and the convergence metric values of the second trial are shown in red (Fig. 16). Figure 16 shows that 70 % of the pairs (7 out of 10 pairs) succeeded in reducing the convergence metric val-



**Fig. 16** The convergence metric values during trials 1 and 2 (experiment 2)

ues during the second trial, which indicated the pairs where closer from the convergence to stable communication protocols formation during the trial 2.

We computed the *t* test between the convergence metric values of the trial 1 and 2 to verify whether there were statistically significant differences. We found then significant differences with *t* test result as follows ( $t = 2.776, d.f. = 9, p \text{ value} = 0.012 < 0.05$ ), indicating that users attempts to converge to stable protocols were more significant during the trial 2.

### 6.2.6 Communication Protocol Evaluation Based on the Independence Test Results

To statistically measure the dependency between the knocking patterns and the different robot’s behaviors, we computed the test of independence (Chi-Square) between the knocking patterns and the different behaviors as well as the Cramer’s V values. Tables 3 and 4 exhibited the results of the first and second trials for the 10 participants. Based on the Table 3, 7 out of the 10 participants (70 %) succeeded in establishing a communication protocol with a Cramer’s V-values ranging from 0.206 to 0.525 and thus ranging from a moderate to very strong relationship. During trial 2 (Table 4), the number of pairs that succeeded in establishing a communication protocol was almost the same despite the new configuration (the point coordinates of the checkpoints have been changed) which required adaptation for the human and the robot. Cramer’s V-Values ranged from 0.283 to 0.387, which meant the relationship between the behaviors and the knocking patterns was moderately strong.

## 6.3 Discussion

### 6.3.1 Command-Like and Continuous Knocking Usage Evaluation

We remarked that command-like was more frequently used in comparison to the continuous—knocking mode. We

**Table 3** The test of independence (chi-square) between the knocking patterns and behaviors, as well as the Cramer’s V (CV) values of trial 1 (experiment2)

Users	Chi-square	CV
User1	$\chi^2 = 51.977, d.f. = 10, \text{***}p \text{ value} < 0.001$	0.425
User2	$\chi^2 = 9.747, d.f. = 10, p \text{ value} = 0.463$	–
User3	$\chi^2 = 20.531, d.f. = 8, \text{***}p \text{ value} = 0.009$	0.206
User4	$\chi^2 = 8.613, d.f. = 6, p \text{ value} = 0.197$	0.4194
User5	$\chi^2 = 12.727, d.f. = 6, \text{**}p \text{ value} = 0.048$	0.477
User6	$\chi^2 = 13.847, d.f. = 6, \text{**}p \text{ value} = 0.031$	0.397
User7	$\chi^2 = 73.605, d.f. = 10, \text{***}p \text{ value} < 0.001$	0.511
User8	$\chi^2 = 11.563, d.f. = 3, \text{***}p \text{ value} = 0.009$	0.525
User9	$\chi^2 = 28.119, d.f. = 8, \text{***}p \text{ value} < 0.001$	0.404
User10	$\chi^2 = 6.111, d.f. = 6, p \text{ value} = 0.411$	–

\* confidence interval is 90%  
 \*\* confidence interval is 95%  
 \*\*\* confidence interval is 99%

**Table 4** The test of independence (chi-square) between the knocking patterns and behaviors as well as the Cramer’s V (CV) values of trial 2 (experiment2)

Users	Chi-square	CV
User1	$\chi^2 = 14.772, d.f. = 10, p \text{ value} = 0.141$	–
User2	$\chi^2 = 31.977, d.f. = 10, \text{***}p \text{ value} < 0.001$	0.381
User3	$\chi^2 = 22.419, d.f. = 6, \text{**}p \text{ value} = 0.001$	0.387
User4	$\chi^2 = 14.625, d.f. = 6, \text{**}p \text{ value} = 0.023$	0.355
User5	$\chi^2 = 26.883, d.f. = 8, \text{***}p \text{ value} = 0.001$	0.291
User6	$\chi^2 = 13.073, d.f. = 6, \text{**}p \text{ value} = 0.042$	0.291
User7	$\chi^2 = 17.885, d.f. = 8, \text{**}p \text{ value} = 0.022$	0.283
User8	$\chi^2 = 3.583, d.f. = 8, p \text{ value} = 0.893$	–
User9	$\chi^2 = 10.044, d.f. = 10, p \text{ value} = 0.437$	–
User10	$\chi^2 = 15.714, d.f. = 6, \text{**}p \text{ value} = 0.015$	0.304

\* confidence interval is 90%  
 \*\* confidence interval is 95%  
 \*\*\* confidence interval is 99%

concluded that the command-like mode was chosen spontaneously so that the problem can be decomposed into static number of states without telling the participants that they needed to modulate their knocking just like in the first experiment.

### 6.3.2 Interaction’s Evaluation Based on the Agreement and Disagreement

Based on the Fig. 14, we found that the percentage of disagreement states exceeded the percentage of agreement states during the trial 1 and that the percentage of the agreement states exceeded the percentage of the disagreement states during trial 2 as well just as in the first trial. The *t* test between

**Table 5** A comparison between the first and second experiments in terms of states of aggregation and performance

	H–H Exp, trial1	HRI Exp trial 1	H–H Exp, trial2	HRI Exp trial 2
%State Based Input CL-CK	90.26–9.73	91.14–8.85	89.47–10.52	95.46–4.53
%Agreement–%disagreement	38.08–61.91	26.68–73.15	64.97–35.02	62.63–37.7
%users with stable protocols	40	70	73	70
%users convergence metric < 0.25	20	50	40	90

agreement and disagreement states was significant during the trial 1 and 2. This indicated that, even though the second trial evolved a new configuration (former checkpoints coordinates changed), the participants were able to achieve significantly more agreement states during the second trial. This paved the way to conclude that during the second trial the pairs did not start from scratch again to establish the communication protocol, although there were some previously shared practices which helped to facilitate the communication protocol formation (transfer learning) just like in the first experiment.

### 6.3.3 Performance Evaluation

The rules sharing led to the significant decrease of the task completion time (Fig. 15) with a significant  $t$  test between the task completion time of the trial 1 and 2 where  $p$  value =  $0.008 < 0.01$ . We also remarked that the interaction led to better performance during trial 2 (Fig. 16). The  $t$  test showed that there were significant differences between the trial 1 and trial 2 convergence metric values. These results indicated that the participants were growing closer to the stable communication protocol formation. By applying the chi-square and Cramer's  $V$  tests, which evaluated the relationship between the patterns and the behaviors, we found out that the number of pairs showed a statistically significant relationship between the patterns and the behaviors did not decrease. This indicated that gradually there was a strong relationship between the knocking patterns and the robot's chosen behaviors.

## 7 Summary of the H–H and the HRI Experiments Results

We may conclude based on the previous results of the HRI experiment that most of the participants succeeded in establishing personalized communication protocols. In the Table 5, we attempted to compare the human–human experiment (H–H Exp) and the human–robot experiment (HRI Exp) results, while CL and CK correspond respectively to command-like and continuous-knocking patterns. Based on

the Table 5, we can see that the number of disagreements of the experiment 2 and during the two trials 1 and 2 (trial 1: 73.15 %–trial 2: 37.7 %), exceeded the number of disagreements of the experiment 1 (trial 1: 61.91 %–trial 2: 35.02 %). We may explain this by the absence of an implemented strategy in the robot that can decode the continuous-knocking patterns which occurred less during the HRI experiment and dropped during the trial 2 (trial 1: 8.85 %, trial 2: 4.53 %) versus a higher value during the H–H experiment which increased during trial 2 (trial 1: 9.73 %, trial 2: 10.52 %). This increase during the H–H experiment can be explained by the fact that the controller could detect the hazardous continuous-knocking patterns and decode them, while the knockers detected in the first trial that the continuous-knocking was handled by the wizarded robot. If we compare the percentage of participants that reached a convergence metric value under 0.25 during experiments 1 and 2, we found that only 40 % of participants finally reached 0.25 as a convergence metric value versus 90 % of participants who finally reached 0.25 as a convergence metric value during the HRI highlighting. Even though we did not implement a strategy that handled the continuous-knocking patterns that emerge during the interaction, we still had better results in terms of convergence to stable protocols formation. We may explain this by the fact the participants during the HRI might have detected that command-like was the best strategy to guarantee a systematized output and that continuous-knocking led to a hazardous output, so they adapted themselves and implicitly avoided that strategy.

## 8 Conclusion and Future Work

In this paper, we presented a human–human WOZ experiment, an actor-critic architecture and an HRI experiment. The WOZ experiment aimed at tracking down the interaction between the knocker and the controller to identify the best practices that may lead to the mutual sharing of the communication rules and facilitated the tailoring of a flexible control model which can be integrated in a minimally



designed robot. We extrapolated these emerging patterns and the pairs (knocker—controller) succeed by shaping their adaptive strategies. In a second step, we implemented the robot's control model. Finally, we conducted the HRI in order to validate our architecture. Our work afforded a methodology that helped bootstrapping how an adaptive model can be tailored and integrated in a minimally designed robot as we expect that it is a persuasive way of guaranteeing long-term use and high sociability factor for such kinds of robots.

In our future work, we intend to integrate inarticulate sounds to the robot's feedback modalities to further investigate whether a simple feedback channel, such as inarticulate sounds, can communicate back to the user, increase expressiveness and boost the convergence toward a more stable communication protocol on a long term basis.

**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

- Thomaz AL, Hoffman G, Breazeal C (2006) Experiments in socially guided machine learning: understanding how humans teach. In: ACM SIGCHI/SIGART conference on human–robot interaction, pp 359–360
- Breazeal C, Hoffman G, Lockerd A (2004) Teaching and working with robots as a collaboration. In: Joint conference on autonomous agents and multiagent systems, pp 1030–1037
- Michaud F, Duquette A, Nadeau I (2003) Realistic child robot Affetto for understanding the caregiver-child attachment relationship that guides the child development. *Syst Man Cybern* 3:2938–2943
- Michaud F, Laplante JF, Larouche H, Duquette A, Caron S, Letourneau D, Masson P (2005) Autonomous spherical mobile robot for child-development studies. *Trans Syst Man Cybern* 35(4):471–480
- Breazeal C, Scassellati B (2000) Infant-like social interactions between a robot and a human caregiver. *Adapt Behav* 8(1):49–74
- Ishihara H, Yoshikawa Y, Asada M (2011) Realistic child robot Affetto for understanding the caregiver-child attachment relationship that guides the child development, *Development and Learning (ICDL)*, pp 1–5
- Levinson SE, Niehaus L, Majure L, Silver A, Wendt L (2012) Can a robot learn language as a child does?. In: AAI spring symposium: designing intelligent robots
- Ugur E, Nagai Y, Celikkanat H, Oztop E (2014) Parental scaffolding as a bootstrapping mechanism for learning grasp affordances and imitation skills. *Robotica*, pp 1–18
- Squire KM, Levinson SE (2007) HMM-based concept learning for a mobile robot. *Evol Comput* 11(2):199–212
- Xu Y, Ueda K, Komatsu T, Okadome T, Hattori T, Sumi Y, Nishida T (2009) WOZ experiments for understanding mutual adaptation. *AI Soc* 23(2):201–212
- Xu Y, Ohmoto Y, Okada S, Ueda K, Komatsu T, Okadome T, Kamei K, Sumi Y, Nishida T (2012) Formation conditions of mutual adaptation in human-agent collaborative interaction. *Appl Intell* 36(2):208–228
- Yamada S, Idea K (2003) interaction design for interaction. *Jpn Soc Fuzzy Theory Intel Inform* 15:185–189
- Mohammad Y, Nishida T (2008) Human adaptation to a miniature robot: precursors of mutual adaptation. In: *Robot and human interactive communication*, pp 124–129
- Mitsunaga N, Smith C, Kanda T, Ishiguro H, Hagita N (2008) Teachable robots: understanding human teaching behavior to build more effective robot learners. *Artif Intell* 172(6):716–737
- Mitsunaga N, Smith C, Kanda T, Ishiguro H, Hagita N (2005) Robot behavior adaptation for human-robot interaction based on policy gradient reinforcement learning. *Intell Robots Syst*, pp 218–225
- Castellano G, McOwan PW (2009) Analysis of affective cues in human-robot interaction: a multi-level approach. In: *WIAMIS*, pp 258–261
- Zeng Z, Pantic M, Roisman GI, Huang TS (2009) A survey of affect recognition methods: audio, visual, and spontaneous expressions. *Pattern Anal Mach Intell* 31:39–58
- Vinciarelli A, Pantic M, Bourlard H, Pentland A (2008) Social signal processing: state-of-the-art and future perspectives of an emerging domain. In: *ACM international conference on multimedia*, pp 1061–1070
- Salvi G, Montesano L, Bernardino A, Santos J (2012) Language bootstrapping: learning word meanings from perception-action association transactions on systems. *Man Cybern* 42(3):660–671
- Odahara Y, Ohshima N, De Silva PRS, Okada M (2013) Talking ally: toward persuasive communication in everyday life, human computer interace, pp 394–403
- Shiomi M, Kanda T, Ishiguro H, Hagita N (2006) Interactive humanoid robots for a science museum, human-robot interaction, pp 305–312
- Kanda T, Hirano T, Eaton D, Ishiguro H (2004) Interactive robots as social partners and peer tutors for children: a field trial. *Hum-Comput Interact* 19(1):61–84
- Chao C, Cakmak M, Thomaz AL (2010) Transparent active learning for robots, human-Robot Interaction, pp 317–324
- Subramanian K, Charles L, Andrea T (2011) Learning options through human interaction. *Workshop on agents learning interactively from human teachers*, pp 208–228
- Matsumoto N, Fujii H, Goan M, Okada M (2005) Minimal design strategy for embodied communication agents. In: *Robot and human interactive communication (RO-MAN)*, pp 335–340
- Forlizzi J, DiSalvo C (2006) Service robots in the domestic environment: a study of the roomba vacuum in the home. In: *ACM SIGCHI/SIGART human–robot interaction*, pp 258–265
- Kozima H, Yasuda Y, Nakagawa C (2007) Social interaction facilitated by a minimally-designed robot: findings from longitudinal therapeutic practices for autistic children. In: *Robot and human interactive communication*, pp 303–310
- Okada M, Sakamoto S, Suzuki N (2004) Muu: artificial creatures as an embodied interface. In: *Computer graphics and interactive techniques*, pp 91–92
- Youssef K, Yamagiwa K, De Silva PRS, Okada M (2014) ROBOMO: towards an accompanying mobile robot. In: *International conference of social robotics*, pp 196–205
- Kina N, Tanaka D, Ohshima N, De Silva PRS, Okada M (2013) CULOT: sociable creature for child's playground. In: *Human-robot interaction*, pp 407–408
- Seiji Y, Tomohiro Y (2004) Training AIBO like a dog—preliminary results. In: *Robot and human interactive communication*, pp 431–436
- Kiesler S (2005) Fostering common ground in human-robot interaction. In: *Robot and human interactive communication*, pp 729–734
- Clark HH (1992) *Areas of language use*. University of Chicago Press, Chicago
- Green A, Htenrauch H, Severinson EK (2004) Applying the wizard of Oz framework to cooperative service discovery and configuration. In: *Proceedings of IEEE RO-MAN*, pp 575–580

35. Shiomi M, Kanda T, Koizumi S, Ishiguro H, Hagita N (2007) Group attention control for communication robots with wizard of OZ approach. In: ACM Human-robot interaction, pp 121–128
36. Kahn PH, Freier NG, Kanda T, Ishiguro H, Ruckert JH, Severson RL, Kane SK (2008) Design patterns for sociality in human-robot interaction. In: ACM human robot interaction, pp 97–104
37. Shiomi M, Sakamoto D, Takayuki K, Ishi CT, Ishiguro H, Hagita N (2008) A semi-autonomous communication robot: a field trial at a train station. In: Joint conference on autonomous agents and multiagent systems, pp 303–310
38. Vijayakumar L, Sethu D, Schaal S (2005) Incremental online learning in high dimensions. *Neural Comput* 17(12):2602–2634
39. Sutton R, Barto AG (1998) Introduction to reinforcement learning. MIT Press, Cambridge
40. Grondman L, Busoniu L, Lopes GA, Babuska R (2012) A survey of actor-critic reinforcement learning: standard and natural policy gradients. *Trans Syst Man Cybern* 42(6):1291–1307

**Khaoula Youssef** is a Ph.D. student in Department of Computer Science & Engineering at Toyohashi University of Technology (TUT), specializing in Human–Robot Interaction. Her research interest include mutual adaptation, social robotics, minimal design, and communication protocols.

**Takahiro Asano** is a graduated student from Department of Computer Science & Engineering at Toyohashi University of Technology (TUT), specializing in Human–Robot Interaction. His research interest include humanrobot interaction, social robotics, persuasive technology and ecological psychology.

**P. Ravindra S. De Silva** is an assistant professor in the Department of Computer Science and Engineering at the Toyohashi University of Technology (TUT). He received PhD in Computer Science and Engineering at University of Aizu in 2006. He was a researcher at Toyota Technological Institute (TTI), from 2006 to 2009. His research interest

include social robotics, proxemics on dynamic interactions, human-robot interactions, user modeling based on statistical approach, and exploring robotic systems based on models of human intelligence. He also received various award include best paper finalist from computer animation and virtual world (non-verbal communication), best paper award from international conference on social robotics ICSR (Springer), best-late breaking report from ACM/IEEE conference on humanrobot interaction, two awards from inaugural Design Competition (international conference on social robotics (Springer)) and Ravindra is the author or coauthor of book chapters, journal, conferences articles and member of IEICE and IEEE.

**Michio Okada** is a professor of Interactions and Communication Design (ICD) Lab at Toyohashi University of Technology (TUT), Japan. He has a Ph.D. in Computer Science at Tohoku University in 1987. From 1987 to 1995, he was a researcher at NTT Basic Research Laboratories, Japan. From 1995 to 2005, he has been with ATR International, where he was a senior researcher and department head in Department of Ecological Communications at ATR Network Informatics Laboratories. From 1998 to 2005, he is a visiting associate professor in Graduate School of Informatics at Kyoto University. His research interests include social robotics, humanrobot interaction and cognitive science, especially ecological socio-cultural approach to human cognition. He has received several awards, including the best paper finalist from ICSR 2010 (Springer), award for best late-breaking report from HRI2010 (ACM/IEEE), and award for best article of the year 2010 from Human Interface Society, Japan.