Autonomous Evaluation and Representation of Social Experiences for Long–term Human–Robot Interaction

Hamed Mahzoon

March 2019

Autonomous Evaluation and Representation of Social Experiences for Long–term Human–Robot Interaction

A dissertation submitted to
THE GRADUATE SCHOOL OF ENGINEERING SCIENCE
OSAKA UNIVERSITY
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY IN ENGINEERING

BY

Hamed Mahzoon

March 2019

Abstract

Recently, social robots have received lots of attention due to their potential in contributing to various aspects of human society. With the reduction of the production cost and consequently facilitation of implementing robots by different companies, organizations, and universities, the robots have applied not only in the factories and industries but also have studied to be adopted in the daily life of human and human society; such as guiding people in shopping malls, playing with children, talking with elderly people, caring at hospitals and clinics, or even teaching English in a classroom. For such robots, establishing a long–term interaction with the human is one of the most important factors to successfully accomplish its duty. However, previous studies mentioned that having a long–term interaction requires lots of interaction patterns programmed in the robot which also have to be able to adapt to different situations. Providing a huge variety of preprogrammed scenarios for such adaptive purpose would not be feasible by a programmer, while on the other hand there are several reports discussing that the most of the participants in a long–term robotic field experiment have got bored after a while. In this study, the open–ended development of the robot was focused as a way to enable the long–term interaction of the robot with the human, and also the way to increase the long–term motivation of the human to continue interacting with the robot was investigated.

To realize the open–ended development of the robot, learning the causality of the events by evaluating the contingency among the observations and the actions of the robot, or in other words the experience of the robot, was proposed. While there are several studies focusing on such contingencies to enable the open–ended learning and development of the robot in a computer simulation, however, they are too much time–consuming and not feasible to adopt to a real–world interaction of a real–world robot with a human; and/or the accuracy of the estimation in the contingency evaluation was low so that the simulated robot could learn only a limited number of the causalities. To treat with this issue, we proposed the local evaluation of the contingencies and showed how the performance of the learning could be increased in terms of speed and accuracy, in the computer simulation. Also, the techniques required for implementing the mentioned mechanism to the real–world robot was proposed. While the mentioned mechanism reduced the time required for the learning of the causalities four to eight times faster than before and the accuracy of the learning was increased around 60%, however in the real world, the learning performance of the robot in terms of speed and accuracy was

still low, especially for the complex causalities, i.e. the causalities regarding to the sequence of the several observations and actions, which are essential for the open–ended development of the robot. As a solution for this issue, two algorithms were proposed for the contingency evaluation unit of the real–world robot: weighted learning of the experiences including an ostensive cue from the interacting human (namely OsL algorithm), such as mutual eye contact with the robot, and distinguishing the evaluation of the experiences concerning the complex causalities from the simple causalities (namely XEP algorithm). The former was expected to increase the learning speed while the latter was expected to improve the accuracy of the learning. By conducting a human–robot interaction experiment, it was shown that the proposed algorithms were effective on the learning performance, and consequently, the robot became able to learn the complex causalities through a feasible interaction time.

On the other hand, the motivation of the human to continue the interaction with the robot is an important factor for the learning of the robot. To increase such a long–term motivation, we focused on the previous reports about the effect of robot's mind and interactability as perceived by the human on the long–term motivation, i.e. increasing the perception of mind and interactability increases the motivation of human to interact again with the robot. However, the ways to increase such perceptions were not investigated. As a way to increase them, in this study we proposed that the sociability of the robot as perceived by the human has an effect on the perceived mind and interactability of the robot. A human–robot interaction experiment supported our hypothesis and showed that sociability of the robot increases the perception of agency and positive experience of the robot as the factors of mind, as well as the perception of likeability and enjoyment as the factors of interactability.

Dedication

To my late grandfather... who encouraged all of us to learn...

'*True knowledge exists in knowing that you know nothing.*'

Socrates

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and Objectives

Is it possible to build a robot which is able to learn autonomously how to interact with the human? Can the robot decide what to learn and how to treat with the people for a long–term interaction? Primitive social skills such as joint attention, pointing, nodding and social referencing were mentioned as the most important abilities for establishing successful communication and continuing interaction with the others [1–3]. These skills were also proposes to be adotped by the interactional robots in order to fasilitate the social interaction of the robot with the human [4–7]. They are expected to help the human more easily understand the intention of the robot through the interaction.

As a way to equip the interactional robots with the primitive social skills, the implementation of the sensorimotor mapping regarding the sensor data as the input of the mapping and the action of the robot as the output of the mapping has proposed in several studies [8–10]. Adopting such mapping enables the robot to produce an intentional reaction to its observation based on the strategies designed for the mapping. In these studies, the strategy utilized for the sensorimotor mapping plays an important role in the expression of the intention of the root during the interaction.

For the implementation of sensorimotor mapping, several studies proposed an autonomous approach by the robot to acquire the mapping. Since there are various types of considerable reactions against the observation of the robot based on the different interaction scenes and the intention of the robot, programming all of the possible mappings considering the different conditions for using each of them seems not to be a feasible way to implement the sensorimotor mapping on the robot. Instead, providing a general mechanism which enable the robot to explore the proper mapping for the current environment have proposed [11], which consequently makes the robot become able to adapt to different environments and situations.

In order to enable the autonomous acquisition of the sensorimotor mapping by the robot, the contingency among the events has focused in several works [12–14]. In these studies, the estimation of the contingency among the events by the robot was utilized to evaluate the relation of the variables, i.e. the robot's motion and the change of the environment as the robot's observation. Based on such evaluation, the causality of the events observed by the robot could be detected and consequently could be adopted to plan a suitable reaction for the next observations in order to meet the predicted result.

Implementation of such robots and robotic systems was, on the other hand, proposed to contribute to clarifying the mechanism of the developmental learning, both for robots and even for the human infants. Several constructive approaches were proposed for this aim and it was discussed how these research contribute to the field of the cognitive developmental robotics [15]. Some of these studies tried to reproduce the developmental process of the human infant by a robot inspired from the reports about the infant's developmental behaviors, and some of them proposed their original developmental model and compared the result with those of the infant (see the survey [15]).

Since the only source of the social developmental robot for the learning is the behavior of the interacting human, considering the tendency and the property of the natural teaching behavior of the human could be a way to enable an efficient learning of the robot. Inducing the human to show more teaching behaviors by choosing the intentional reactions by the robot would be another strategy to increase such efficiency. Since the studies about the natural pedagogy of the human for teaching infants was reported in several previous works (see a review [16]), designing such reactions inspired from these studies could be proposed as the implementation strategy for the robot's behavior. Also, the tendency of the perceptions of the human about the interactional robots was widely studied in several reserch [17–19]. Considering these perceptions, such as the psychological aspect of the robot as perceived by the human, it's agency and expected intelligence, as well as the theory of mind about the robot, seems to be the effective way to increase the efficiency of the learning.

On the other hand, the long–term motivation of the human in order to continue interacting with the robot has been discussed as another important factor for realizing the long–term interaction of the human with the robot [20]. Here, the essential claim is even if the robot was equipped with a powerful social skill learning algorithms, a useful applications or even an attractive shape/face, however if we couldn't evoke the motivation of the human to continue the interaction with the robot and/or interact again with the previously interacted robot, then the robot would be turned into a statue and not be used by the human. For the social robot's, it would not have enough opportunity to collect interaction data and consequently the learning of the robot would not be successful. Several studies investigated the ways to increase such long–term motivation and conducted research on the factors that affect the motivation of the human [21–24].

In this dissertation, we address the ways which enable the interactional robot to learn social skills and have long–term interaction with the human in a real–world environment. In the next section, the problems which required to be tackled to achieve such robot was mentioned and the solution proposed by our study for each of them was discussed. Finally, the contribution of each chapters was explained and the idea proposed in each chapter to solve the mentioned problem was briefly expressed.

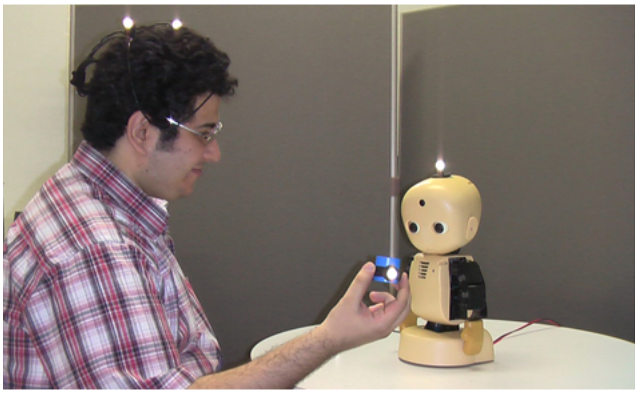


Figure 1.1: A real–world interaction environment of the robot with the human.

1.2 Problems to Solve

In order to enable the interactional robot to learn social skills through the interaction with the human in a real–world environment such as the one shown in Fig.1.1, realizing the evaluation of the contingencies during a feasible period of time would be one of the most important factors. However, most of the contingency evaluation methods which work based on the interaction experience of the robot require huge number of interaction experience with the human. This problem is more essential for the evaluation of the contingencies concerning a longer sequence of experiences. Based on the previous reports, the fastest social skill regarding a contingency with two sequences of events required 40,000 steps of turn-taking interaction of the robot with the human. Obviously, this amount will not be feasible for the implementation of a real–word robot. Decreasing the time required for the evaluation of the contingencies, especially the ones with the longer sequence of events, is the first problem to solve in order to realize the implementation of the contingency evaluation mechanism on a real–world robot.

Even if the time required for the evaluation of the contingencies was reduced enough so that the implementation on a real-word robot would become feasible, solving the problem of the dependency of the contingencies is essential. Since the contingency among different events were evaluated simultaneously, the independence of events from each other should be guaranteed to have an accurate contingency evaluation for them. However, detecting whether two events are completely independent from each other seems to be difficult because the real causality among the events is not clear, even for a human. Considering this constraint and designing a system which gets less effect from this problem would be the other essential issue for realizing the implementation of the learning method on a real–world robot.

For the problem concerning the long–term motivation of the human, different and various factors were reported as the effective ones on increasing the motivation of human in order to interact again with the robot, such as perceived agency, enjoyment, safety, anthropomorphism, likeability and so on. However, providing a robot with the acceptable degree for all of these perceptions seems to be difficult. Investigating more general and easier way to increase such perceptions about the robot, and consequently increasing the long–term motivation of the human to continue interacting with the robot would be the other challenge of this study.

In **chapter 3**, the problem of the huge time consumed for the evaluation of the contingencies was discussed. In this chapter, the global evaluation of the contingencies as well as the synergistic contribution of the contingencies were mentioned as the main problems which cause the essential delay of the contingency evaluation. Also, these problems was mentioned as the reason of decreasing the accuracy of the evaluation. In order to deal with these problems, a method utilizing the local evaluation of the contingencies was proposed. Also, a way to correctly consider the synergistic effect of the contingencies was explained. By using a comparative computer simulation experiment, the contribution of the proposed mechanism was verified. The result showed that the proposed method improved the performance of the learning in terms of learning speed, accuracy and resistance against noise. Also, it was shown that the proposed method enables the acquisition of more complex skills in more complex environment.

In **chapter 4**, the application of the system proposed in chapter 3 to a real–world environment by utilizing a real-world robot was proposed. As mentioned above, in this chapter the problem of the dependency of the contingencies was discussed as one of the most essential problem for the implementation, and a method to prevent the overestimation of the contingencies regarding this dependency was proposed. Also, a way to synchronize the teaching phase of the human with the learning phase of the robot was proposed, which was inspired from the natural pedagogy of human, in order to increase the speed of the learning. By using a human–robot interaction experiment in a real–world environment, it was shown that the proposed method significantly improved the speed of the learning, and consequently increased the number of the learned social skills. Also, the predictability of the robot was improved and as a result, a longer turn–taking interaction of the human with the robot was realized.

In **chapter 5**, the long-term motivation of the human to interact again with the robot was proposed as an important factor for the learning process of the robot, and the way to increase such motivation was investigated. The mind and interactability of the robot as perceived by the human was proposed as the effective factors to increase such motivation, and a hypothesis about the effect of the perceived sociability of the robot by the human on these factors was constructed. By conducting a human-robot interaction experiment comparing the social and non-social conditions, it was shown that the result of the experiment supports the hypothesis: the perception of mind and the interactability of the robot by the interacting human was affected by the perceived sociability of the robot, and could be increased by the self-representation of sociability of the robot.

Finally, in **chapter 6**, the conclusion of the dissertation as well as the future works were proposed.

Chapter 2

Previous Works

In this chapter, the previous works concerning the acquisition of social skills by the robot as well as the long–term motivation of the human to interact with the robot were mentioned. Several research studied the ways to equip the robot with the primitive social skills in order to facilitate more natural interaction of the robot with the human. Also, understanding the mechanism of the learning of the social skills by the human infant utilizing the constructive approach has became the other goal in some of these studies. On the other hand, the long–term motivation of the human to interact with the robot has became one of the most important bottlenecks for the field of human–robot interaction. Several studies tried to understand the feature of the human during interacting with the robot and investigate how to increase such motivation. In the following, the previous studies concerning each of these concepts are briefly introduced.

2.1 Learning of Primitive Social Skills

2.1.1 Acquisition of Joint Attention Skill

Joint attention is one of the most important primitive social skills for establishing and continuing a successful communication with the others [1–3]. It was defined as looking at the same place with the others and widely used through the interaction in order to share intention, attract attention, and/or even teach a concept to the others. Due to the important role of this ability in the acquisition of various social skills, several research have been conducted regarding it in the field of cognitive science and developmental psychology.

In the field of social robotics, several research focused on the capabilities of this ability through the interaction with the human, and tried to adopt it to a robot in order to realize more natural interaction [4–7]. Also, by the implementation of the behaviors concerning the joint attention on an infant robot, some of these studies tried to clarify the mechanism of the acquisition of joint attention by the human infant. For example, some previous works [8,9] focused on the causality between the gaze behavior of the human infant and the caregiver, and proposed a learning mechanism based on the acquisition of the sensorimotor mapping from the patterns of the face of the caregiver to the robot's motor command. They showed that the system could learn autonomously and adapt the learning to the probable different behaviors of the caregiver based on the inherent causality. However, the robot could not distinguish the modalities and the observations that are related to each other by itself, but the information about such relation was required to be programmed before. The other work [11] proposed that evaluating the transfer entropy [25] among the variables could be utilized to detect the relation among the events, and consequently the robot becomes able to autonomously decide about the variables of the sensorimotor mapping. They showed that the basic social skill such as gaze following could be acquired by the robot through a simplified face–to–face interaction of human with the robot. However, the environment of the experiment was very simple and the doubt about the capability of the proposed method to be applied to more general environment was still remained.

2.1.2 Developmental Process of Learning Social Skills

To enable the learning of more complex social skills by the robot, the developmental approach was utilized in several works. In these works, the robot uses the learned concepts to examine the world and find the other concepts to learn. Sumioka et al. [11] focused on the chain of the contingency among the events through the face–to–face interaction of the infant with the caregiver, and proposed a mechanism to evaluate the contingency of such events, sequentially. In there mechanism, the robot first evaluates the contingency among the robot's observation, robot's action and the consequent observation. If any contingency was detected, the robot acquires the sensorimotor mapping concerning the found contingency, i.e. the observation as the input and the action as the output of the mapping. After that, the robot starts to evaluate the relation among the found contingency and the event following that. Since it was evaluating a contingency of the occurrence of an event after another contingency, the sequence of the contingency was defined as the contingency chain. They showed that by evaluating the contingency chains, the robot become able to enhance its learned social skill and develop it to more complex one. In practice, their infant robot acquired the gaze following behavior and developed it to the social referencing behavior, which is the other primitive and important social skills during the interaction, especially through the developmental process of the human infant.

Interestingly, comparing the achieved development by the proposed robotic system with the developmental process of the human infant has became one of the constructive approaches to understand human. The concept was declared as the field of cognitive developmental robotics [15] and several research tried to understand the mystery of infant's learning and development by such constructive approach (see the review [15]). For example in the previous work [11], the developmental process of the implemented robotic system was compared with that of 6 to 18 months old infant, and it was discussed that by using their method, not only the robot could enhance the learned social skill such as gaze following, but they could reproduce the developmental process of the human infant. This could be expressed that the contingency evaluation method utilize by the system potentially can describe the essence which the infant consider for enhancing its learning and realize the development. However the work was proposed in a computer simulation environment, and the capability of the proposed method for the application in a real–world environment in order to reproduce such development process in a real–world robot was not investigated.

The other work [13] applied the reinforcement learning to a robotic system in order to incrementally adjust the dynamic Bayesian network for modeling the occurrence of the predictable events in a continuous environment. In this study, the hierarchy of the primitive motions of the robot was sequentially combined to accomplish the tasks which were asked from the robot. By enhancing the model of the planner of the system from a simple planning unit which were prepared for a simple tasks such as pushing an object on the table to a complex planing unit which were constructed for executing more complex tasks such as garbing an object on the table, by combining the learned simple models. However, adopting this model to the interaction scene of the human with the robot seems not to be feasible, because the proposed prediction system had no mechanism to distinguish the contingencies that were not related to the state of the agents, i.e. the human and the robot. It is understandable, because the aim of the study was to become able to execute as many as probable tasks provided from the human, therefore storing as many as possible contingencies in the system as a planner increases the power of the system to treat with more commands and accomplish more tasks. This approach would not work with the environment of the face–to–face interaction of the robot with the human, in which the detection of the irrelevant variables become one of the challenging problem to solve in order to learn suitable social skills. In chapter 3 of the dissertation, we address the methods to treat with these issues.

2.1.3 Open–ended Development

The concept of the open–ended development for the robot refers to the learning framework in which the robot can continue the learning of the new concepts in an open–ended manner. It have been proposed in various studies by equipping the robot with different properties such as curiosity, intrinsic and extrinsic motivations, desire or even developing the learned simple concepts to more complex ones [26–28]. With this method, it is expected that the robot become able to autonomously continue its learning without any specific reward, guide of correction from the external environment. Consequently, the system would *potentially* be able to investigate and learn the concepts that difficult to find/teach and hard to investigate even for the human. For example, Oudeyor et al. [27] proposed a learning system with the intrinsic motivation, in which the robot was able to change its learning goal base on its internal curiosity. They showed that by setting the predictability of the observations as the negative rewards for the learning of the robot, how the curiosity could be expressed and the robot could set a new goal for its learning. This approach was discussed as a way to implement the open–ended learning and development of the agent, and could be applied to a real–world robot. However, this approach did not consider the sequential properties of the human's teaching/learning behaviors, and adopting it to the interaction scene of a learning robot with a teaching human will loose a huge information included in the interaction manner of the human. Since through such interaction, the human first tries to teach a simple concept to the learner, and then tries to enhance it to more complex concepts, therefore not considering this tendency could causes to examine the huge number of contingencies to find the complex concepts. this problem is verified in chapter 3 of this dissertation, and a solution for this problem is proposed as well. Also, as the best of our knowledge, most of the studies concerning the open–ended development of the learning robot were designed for a very limited conceptual frameworks and specialized to a very specific environment. Additionally, these studies were conducted for a simulation environment and the methods to enable the open–ended development of a real–world robot in a real–world environment have not been investigated enough. For example, the previous work [11] showed that the evaluation of the contingency chain enables the learning of the primitive social skills in an open–ended manner; however the system required a huge number of the turn–taking interaction with the caregiver and consequently the learning was feasible only in a simulation environment. In chapter 4, the methods which enables the implementation of such learning and development on a real–world robot was discussed as well.

2.2 Motivation of Human to Continue Using Robot

With regard to the decrease of the production costs of the robots and development of different types of robot by various companies, organizations and universities, the number of the robots with applications in social environments have been increased. These robots usually have been designed considering the probable interaction scene of the robot with the human. While the first interaction of the human with such robots would be interesting for them and the designed interaction strategies, physical shape, reactions and behavior of the robot would be successful to construct an active interaction between the human and the robot, however several studies reported that the motivation of the human to interact again with the robot decreases with the pass of time [20]. Therefore, keeping the motivation of the human to interact again with the robot and/or continue using the robot have became one of the most important topics for the implementation of the social robot.

Several research studied the behavior of the human against the robot during a relatively long time to understand how their treatment with the robot change and when the human looses his/her motivation to interact with the robot (see survey [20]). Most of them reported that even if the implemented robot was equipped with some interesting application, had very cute face and shape, spoke various sentences and even perceived very attractive for children, however after a relatively short period of time (depends on the type of the robot as well as the field experiment) they stopped using the robot and the robot became just a statue than an interactional robot. The studies analyzed that the human got bored due to several factors, such as the non-natural behavior of the robot, limited variation of the reaction of the robot, robot's predictable actions as well as the break through the content of the interaction. Therefore, implementing a robot which the human could honestly feel that he/she wants to continue using it for a long-term seems to have faced with several essential problems.

On the other hand, in order to increase the long–term motivation of the human to interact with the robot, several factors have been proposed as the effective ones. The agency, intelligence, autonomy, experience, independency, animacy, anthropomorphism and life–like behavior of the robot [29–36] have been proposed as the factors that express the mind of the robot, which affects the motivation of human. The likeability, safety, enjoyment and physical attractiveness of the robot [4,37–42] have been proposed as the factors that express the interactability of the robot, which are also effective on the human's motivation. However, considering all of these factors in the design and the implementation of the robot seems not to be feasible. Also, the way to increase the perception of the interacting human with the robot about the effective factors, i.e. the perceived mind and the interactability of the robot has not been investigated enough. Therefore, conducting more research on the ways that enable the robot to make the human to be interested in continuing the interaction and using again the robot seems to be essential for the field of the social robots, especially the implementation of them. This issue is tackled in chapter 5 of this dissertation.

Chapter 3

Evaluation of Local Contingencies for Open-Ended Development of Infant Robot

The developmental algorithm proposed in the previous work [11] was too much time consuming for the implementation to the real world robot. In this chapter, the reason as well as the solution for this problem is proposed. The proposed solution was compared with the previous work and the other famous ones [13,27] in terms of different aspects of the system performance by using a computer simulation. Moreover, the potential of the system utilizing the proposed method for applying in more complex environment was evaluated. This issue showed the potential of the proposed method to be applied in the real–world human–robot interaction.

3.1 Introduction

Joint attention is one of the most basic cognitive functions in human communication. It is simply defined as looking where someone else is looking [1], and extensive research has been

This work was published in: Mahzoon H, Yoshikawa Y, Ishiguro H. Social skill acquisition model through face-to-face interaction: Local contingency for open-ended development. Frontiers in Robotics and AI. 2016; 3:10.

conducted by those investigating the developmental process of following the gaze of others [1–3], including a report that a human infant shows this capability before birth [44]. These results have garnered research interest in cognitive science and developmental psychology owing to its important role in the acquisition of social capabilities such as language communication and mind reading [45]. Recently, a number of research efforts in the field of robotics have focused on the issue of joint attention in human-robot interaction [7] as it also appears to be a necessary building block in this type of interaction. The development of robotics technologies and a consequent possible future for human society with interactive robots adds to the importance of studies on joint attention, which has implications for creating communicative robots [5,6].

Understanding the developmental process of the human infant by producing an infant model based on the utilization of synthetic approaches [15] has also garnered increasing research interest. The role of joint attention-related behavior is also being explored in this research area in an effort to understand the mystery surrounding the development of a human being [8,9]. These synthetic studies focused on the significant role of joint attention behavior in the acquisition of social skills through the interaction of a human infant with its caregiver. They considered the causality between gaze behavior of the infant and the caregiver, and proposed that a robotic model of an infant could acquire social behavior, such as gaze following, by acquiring sensorimotor mapping from the face pattern of the caregiver to its own motor command by estimating the causality among them. In these works, the programmer had to specify the set of variables on which the robot should focus to learn the sensorimotor mapping. However, to acquire various types of social behavior, the programmer needed to redefine the set of variables for each of the behavioral types. Therefore, a mechanism to automatically find the appropriate sets of variables seems to be necessary to apply the learning robot in different interactive scenes in different environments.

Oudeyer et al. addressed intrinsic motivation for the learning robot, which enables open-ended development [27]. They showed how an internal reward system that equates a lower prediction error with a larger reward enables the learning robot to automatically acquire various types of behavior in an incremental manner. However, they did not consider the sequential properties of human behavior in the interaction in depth, which seems to be essential for continuing the interaction. For example, when a robot becomes more social by learning to respond to a human in an appropriate way, the human will try to continue the interaction with it by, for example, talking to it or touching it in response to its reaction. Therefore, by learning actions that result in contingent sequences, a robot would be able to continue such interactions with the human.

Mugan and Kuipers extended the reinforcement learning method to incrementally acquire dynamic Bayesian networks for modeling predictable events in a continuous environment [13]. Because a longer sequence of events gradually becomes predictable by sequentially combining the models found, it can be used to plan the hierarchy of the actions to accomplish the given tasks, ranging from simple tasks such as hitting an object, to complex ones such as grabbing the object. However, it is difficult to apply this method to learning social skills through social interaction, because the tasks that have to be accomplished in the interaction are not always (or even rarely) explicitly given. Furthermore, the pure predictability of the events is considered to evaluate the causality of the events, which is implemented by dynamic Bayesian networks. Therefore, causalities that are independent of the state of others are included in the found causalities. In other words, there is no mechanism to detect and avoid these causalities, which leads the robot to be incapable of choosing a suitable action for interacting with the human through social interaction.

In recent studies, Sumioka et al. [10] derived a contingency evaluation measure focusing on the effect of a robot's own action only in relation to specific conditions, including those reflecting the state of others. An informational theoretical measure based on transfer entropy [25] was utilized for the evaluation, and applied to the model of an infant robot interacting with a human caregiver model in a computer simulation. The result shows that evaluating contingency among variables leads the robot to find a combination of the variables that should be focused on to acquire a sensorimotor mapping, which enables the robot to behave in a social way, such as performing gaze following and social referencing. Regarding the gradual changes of the response of the caregiver model with the emergence of the infant-robot's communicative abilities, which were designed based on those reported for the interaction of a human caregiver with its infant [46,47], the robot obtained a meaningful response from the caregiver model when it used an acquired social skill. For example, the caregiver model had a strict rule such as looking at the robot when the robot successfully performed the acquired gaze following skill. Therefore, the robot could find further contingency between "using the gaze following skill" and "consequent face direction of the caregiver" (which is looking at the robot due to the strict rule). This contingency was expressed as a chain of contingencies, which consists of a sequence of two consequences: 1) finding the object (by following the gaze of the caregiver); and 2) finding the caregiver's frontal face (by looking at the caregiver after gaze following).

In [11], they focused on the importance of the sequence of contingent sub-actions in several social behaviors of human infant, where lots of them such as social referencing behavior consist of these sequences. Therefore, they extended the mechanism to find the chain of contingencies by evaluating contingency among the skill that was used, the action that was taken, and the consequent observation of the robot. However, in their method, the acquisition of complex skills, i.e., skills consisting of a contingency chain, is too time consuming; in addition, they did not sufficiently check whether complex skills with longer sequences could be acquired. Accordingly, only the acquisition of the simple skills is reported in the implementation of this system in a real-world robot [48]. Moreover, the performance of this mechanism is not compared with the results obtained by others, such as [27] and [13], in which simple concepts, such as the pure predictability of the events for the evaluation of the causalities, are used.

In this chapter, we propose a mechanism to overcome these two significant problems, namely poor skill acquisition and the large number of time-consuming steps. Our proposed mechanism is based on two main ideas. First, we introduce a new informational measure, named transfer information, which evaluates contingency among specific values of variables. Previously, [10,11, 48], the expectation of contingency among whole values of the variables was utilized to evaluate contingencies. Therefore, gathering a sufficient number of samples for all values of the variables was highly time consuming. Instead, in the proposed method, it is sufficient to gather samples of specific values of the variables for the evaluation. In this way, fast contingency evaluation is realized in the proposed method. Second, we utilize transfer information to produce a measure which evaluates the synergistic contribution of values in contingencies. It enables the robot to distinguish the contingencies, which consist of the synergistic effect of taking a specific action in a specific state to the environment, from those composed of a single effect. This approach leads the robot to acquire more complex skills, i.e., skills with a longer contingency chain and interaction sequences, compared with work by other researchers.

In the results section of this chapter, we compare the performance of our proposed mechanism with those of others using computer simulation: work which utilizes simple concepts, such as predictability for the skill acquisition process of the robot, such as [27] and [13]; and work which uses more complex concepts, such as finding saliency of contingencies among variables to prevent the acquisition of a huge number of behavioral rules, as well as those unrelated to the state of the interacting person, such as [10], [11], and [48]. In this chapter, the former is implemented using transfer information to evaluate events at the value level (locally), whereas the latter utilizes transfer entropy which evaluates events at the variable level (globally). We refer to these methods as "local pure predictability method" (l.p. method) and "global contingency method" (g.c. method), respectively. For comparison, we consider the interactive environment used in the previous work [11], but the measure used in the skill acquisition process of the robot differs for each method.

The remainder of this chapter is organized as follows: first, we describe the assumed interactive scene of the robot with its caregiver in our experiment. Then, we explain the system schema of the proposed mechanism and its components. After that, we analyse the results of the two experiments, i.e., the computer simulation performed to validate our system. In the first simulation, we compare the performance of the proposed method with the other mechanisms in a simplistic interactive world model. In this comparison, accuracy and speed of skill acquisition as well as robustness against uncertainty is examined. In the second simulation, we examine whether our system remains feasible when the robot has additional sensing and action modalities and the caregiver behaves accordingly based on more complex rules. It shows the capacity of the proposed mechanism for application in a more complex environment such as in real-world interaction between robots and humans.

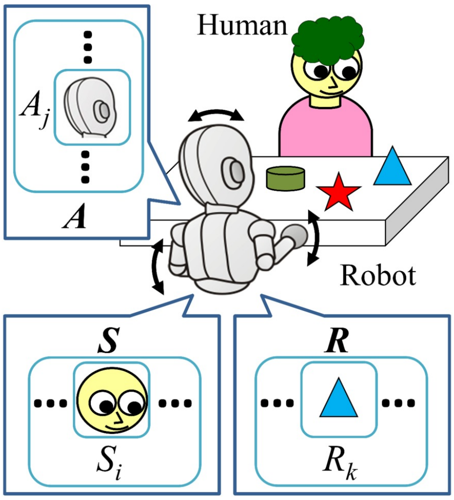


Figure 3.1: Interaction Environment: A human caregiver interacting with a robot. They sit across a table, and there are some objects. During interaction with the human, the sensory data, taken actions, and resultant observations are stored in the sensory, action, and resultant variables, respectively ($\boldsymbol{S}$, $\boldsymbol{A}$ and $\boldsymbol{R}$).

3.2 Mechanism

We assume a face-to-face interaction between a human caregiver and an (infant) robot (Fig.3.1). In each time step, the robot observes its environment, and sends action commands to its joints. The robot obtains the observation as sensory variable $\boldsymbol{S}$, and taken action as action variable $\boldsymbol{A}$. In addition, the robot retains the resultant observation after taking the action as resultant variable $\boldsymbol{R}$. These variables are consist of some elements, S_i ($i = 1, 2, ..., N^s$; N^s denotes the number of types of sensory data), A_j ($j = 1, 2, ..., N^a$; N^a denotes the number of different kinds of actions) and R_k ($k = 1, 2, ..., N^k$; N^k denotes the number of types of resultant sensory data), respectively.

Fig.3.2 shows the structure of the proposed mechanism. The system consists of two main components: a Contingency Detecting Unit (CDU) and an Action Producing Unit (APU). After the observation of the current state at time t, i.e., updating the variables $\boldsymbol{S}^t$ and $\boldsymbol{R}^t$, the APU produces an action for each joint of the robot ($\boldsymbol{A}^t$) based on the current state. The robot observes the consequent result of taking action $\boldsymbol{A}^t$, and saves it as $\boldsymbol{R}^{t+1}$. The CDU evaluates the contingency based on the variables $\boldsymbol{S}^t$, $\boldsymbol{R}^t$, $\boldsymbol{A}^t$, and $\boldsymbol{R}^{t+1}$. If the CDU detects contingency, it adds a new Contingency Reproducer (CR) to the APU, which enables the robot

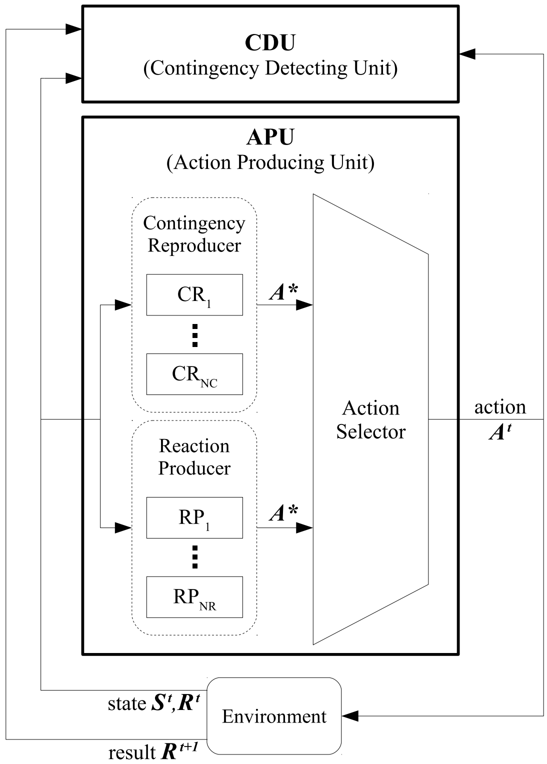


Figure 3.2: System schema.

to reproduce the found contingency by taking suitable action when it is observing a specific state. The action is mentioned as $\boldsymbol{A}^*$ in Fig.3.2. Therefore, at the beginning, there is no CR in the APU; the APU produces action based on the output of another component, named the Reaction Producer (RP). The RP is designed to enable the robot to take pre-programmed reactive action to a specific observation. The RP outputs suggested action $\boldsymbol{A}^*$ in each time step. To produce many reactions, the system can have several RPs, from RP_1 to $\mathrm{RP}_{\mathrm{NR}}$. A component named Action Selector, finally, selects the outputting action to each joint of the robot, among the suggested actions $\boldsymbol{A}^*$ from the RP and the CR (if any). Therefore, at the beginning of the interaction of the robot with the caregiver, the APU outputs the action based on the $\boldsymbol{A}^*$ produced by RPs. Continuing the interaction leads the CDU to find contingency and add CRs to the APU. After that, the Action Selector chooses the outputting action $\boldsymbol{A}^t$ from the $\boldsymbol{A}^*$ of the CRs and the RPs. The robot continues the interaction and continues acquiring additional CRs if there are further contingencies during the interaction. In this section, we explain each of the components in detail.

Note that the global structure of the g.c. and l.p. method which we implemented for the comparison are same with the proposed mechanism, which have/will be described in this section/following subsections. In the last part of the following subsections, we will mention about them if there is any difference among the proposed method and the g.c./l.p. methods. For the detail of the g.c. method, see [11].

3.2.1 Contingency Detection Unit (CDU)

The CDU has two roles: detecting contingent experiences and generating new CRs based on them. It tries to find the contingencies using a histogram of experiences obtained through the interaction. The CDU is equipped with an informational measure to evaluate the contingency of the experiences. Once an experience is judged to be contingent, a CR is added to the APU to enable the robot to "reproduce" the contingency, i.e., take a suitable action in a specific state to be able to repeat the experience. We define the quaternion $\boldsymbol{e} = (r_k^{t+1}, s_i^t, a_j^t, r_k^t)$, where $r_k^{t+1}, r_k^t \in R_k$, $a_j^t \in A_j$ and $s_i^t \in S_i$ as an *experience*. It indicates that in state s_i^t taking the action a_j^t made the transition of r_k^t to r_k^{t+1}.

Using this definition, the system tries to learn the knowledge "when, what to do, for which transition". It is expressed as acquiring social skills in our study.

Evaluating Contingency

Assume X and Y denote two discrete-time stochastic processes which could be approximated by a stationary Markov process. When X takes the value x^t at time t, the evolution of the process is described by the transition probability $p(x^{t+1}|x^t)$. Using transfer entropy [25] the dependency of the process X on the process Y can be quantified as:

$$T_{Y \to X} = \sum_{x^{t+1}, x^t, y^t} p(x^{t+1}, x^t, y^t) \log \frac{p(x^{t+1}|x^t, y^t)}{p(x^{t+1}|x^t)}. \tag{3.1}$$

In other words, transfer entropy evaluates the effect of process Y on the transition of process X. We introduce "transfer information" which estimates the effect of a specific value of process

Y, i.e., y^t, on the specific transition of process X, i.e., x^t to x^{t+1} as follows:

$$I_{y \to x} = \log \frac{p(x^{t+1}|x^t, y^t)}{p(x^{t+1}|x^t)}. \tag{3.2}$$

If the value of the transfer information is high, it shows that the specific transition of x^t to x^{t+1} has high dependency on the specific value y^t. We refer to this dependency as *local contingency*, or simply "*contingency*". It is named local, because it does not evaluate the (averaged) dependency among all values of the processes (such as transfer entropy); instead, it performs the evaluation among the specific values of these processes.

Applying this to our environment, we can evaluate the effect of the specific values of the sensory variable S_i, i.e., s_i^t, and the action variable A_j, i.e., a_j^t, on the specific transition of the resultant variable R_k, i.e., r_k^t to r_k^{t+1}, by the following equations, respectively:

$$I_{s_i \to r_k} = \log \frac{p(r_k^{t+1}|s_i^t, r_k^t)}{p(r_k^{t+1}|r_k^t)} \tag{3.3}$$

$$I_{a_j \to r_k} = \log \frac{p(r_k^{t+1}|a_j^t, r_k^t)}{p(r_k^{t+1}|r_k^t)} \tag{3.4}$$

In other words, they evaluate the contingency of a specific transition on a specific state and action, respectively. Moreover, the joint effect of the specific values of the sensory variable and the action variable can be evaluated with:

$$I_{(s_i,a_j) \to r_k} = \log \frac{p(r_k^{t+1}|s_i^t, a_j^t, r_k^t)}{p(r_k^{t+1}|r_k^t)}. \tag{3.5}$$

In other words, it evaluates the contingency of a specific transition on a specific action in a specific state. Considering an experience $\boldsymbol{e} = (r_k^{t+1}, s_i^t, a_j^t, r_k^t)$, we can evaluate the local contingency of the experience on the state s_i^t, on the action a_j^t, or on the action a_j^t in the state s_i^t, using equations (3.3) to (3.5), respectively:

$$L_s(\boldsymbol{e}) = I_{s_i \to r_k} \tag{3.6}$$

$$L_a(\boldsymbol{e}) = I_{a_j \to r_k} \tag{3.7}$$

$$L_{sa}(\boldsymbol{e}) = I_{(s_i, a_j) \to r_k} \tag{3.8}$$

We term them "single" contingencies (of experience $\boldsymbol{e}$) on s_i^t, on a_j^t, and "joint" contingencies on (s_i^t, a_j^t), respectively. However, evaluation of the joint contingency may reflect the contingency only on either s_i^t or a_j^t, but not on both of them. To evaluate the synergistic contribution of both of the values (s_i^t, a_j^t) on the joint contingency, we need to eliminate the contribution of single values, s_i^t and a_j^t. The following equations eliminate the single contribution of values s_i^t and a_j^t, respectively:

$$^{S}E(\boldsymbol{e}) = \log \frac{p(r_k^{t+1} | s_i^t, a_j^t, r_k^t)}{p(r_k^{t+1} | s_i^t, r_k^t)} = L_{sa}(\boldsymbol{e}) - L_s(\boldsymbol{e}) \tag{3.9}$$

$$^{A}E(\boldsymbol{e}) = \log \frac{p(r_k^{t+1} | s_i^t, a_j^t, r_k^t)}{p(r_k^{t+1} | a_j^t, r_k^t)} = L_{sa}(\boldsymbol{e}) - L_a(\boldsymbol{e}) \tag{3.10}$$

In the first equation, the transition probability $p(r_k^{t+1} | s_i^t, a_j^t, r_k^t)$ in the numerator is compared with $p(r_k^{t+1} | s_i^t, r_k^t)$ in the denominator. This compares the contribution of (s_i^t, a_j^t) on the transition of r_k^t to r_k^{t+1}, with the single contribution of s_i^t. In other words, it compares the dependency of the transition of r_k^t on (s_i^t, a_j^t) with the dependency on s_i^t. Therefore, $^{S}E(\boldsymbol{e})$ shows the difference of contributions on joint contingency between those of (s_i^t, a_j^t) and s_i^t. In other words, it eliminates the single contribution of s_i^t on the joint contingency. According to equations (3.3),(3.6), and (3.5),(3.8), $^{S}E(\boldsymbol{e})$ could be written as the subtraction of $L_s(\boldsymbol{e})$ from $L_{sa}(\boldsymbol{e})$. For the same reason, $^{A}E(\boldsymbol{e})$ eliminates the single contribution of a_j^t on the joint contingency. To eliminate both of the single contributions and achieve the synergistic contribution of the values on the joint contingency, the following measure could be used:

$$E(\boldsymbol{e}) = \min\{{}^{S}E(\boldsymbol{e}), {}^{A}E(\boldsymbol{e})\}. \tag{3.11}$$

Using this measure, the robot would be able to distinguish the experiences which are dependent on both of s_i and a_j, i.e., reflecting the knowledge "when (s_i^t), what to do (a_j^t), for which transition (r_k^t to r_k^{t+1})". The robot uses this measure to evaluate experiences in the proposed mechanism.

Note that the measure used in the g.c. method for the evaluation is designed based on transfer entropy, which evaluates contingency among the variables S_i, A_j, and R_k [11]:

$$\begin{aligned} C_{i,k}^{j} &= T_{(S_i,A_j)\to R_k} - (T_{S_i\to R_k} + T_{A_j\to R_k}) \\ &= \sum_{\substack{r_k^{t+1}, r_k^t \in R_k \\ s_i^t \in S_i, a_j^t \in A_j}} p(r_k^{t+1}, s_i^t, a_j^t, r_k^t) \log \frac{p(r_k^{t+1}|s_i^t, a_j^t, r_k^t)\, p(r_k^{t+1}|r_k^t)}{p(r_k^{t+1}|s_i^t, r_k^t)\, p(r_k^{t+1}|a_j^t, r_k^t)} \end{aligned} \tag{3.12}$$

Furthermore, the l.p. method uses $L_{sa}(\boldsymbol{e})$ for the evaluation, which reflects the (pure) predictability of the experience $\boldsymbol{e}$, without elimination of the single contributions of the values on the joint contingency, such as those done in the equation (4.5) or (4.6).

Adding new CR to the APU

In each time step of the interaction, the robot calculates $E(\boldsymbol{e})$ for all experiences utilizing their histograms. If the robot experiences specific $\boldsymbol{e}$ more than θ times and $E(\boldsymbol{e})$ is higher than the acquisition threshold C_T, the robot accepts it as a *contingent experience* and adds a new CR to the APU based on it. Using the CR, the robot tries to reproduce the contingent experience in the next steps of the interaction. Each CR is mentioned as $\pi_{\boldsymbol{e}}$ based on the experience $\boldsymbol{e}$ that the CR is generated.

When a CR $\pi_{\boldsymbol{e}}$ is added to the APU, the CDU adds a new binary sensory variable S^{π} to the set of the sensory variable $\boldsymbol{S}$. The new variable indicates whether $\pi_{\boldsymbol{e}}$ has been used in the previous time step; it takes the value "1" if it has, and "0" otherwise. The CDU then continues evaluating the $E(\boldsymbol{e})$ of the experiences, including the new variable S^{π}. As a result, the CDU can evaluate a chain of contingencies stemming from the use of the found contingency, i.e., the contingencies relying on more than one time step related to the generated CR. Note that, the CDU does not count experiences which do not contain S^{π}, when the robot has used an acquired $\pi_{\boldsymbol{e}}$. We expect this trick to lead the CDU to focus on the effect of using acquired $\pi_{\boldsymbol{e}}$ on the environment, and consequently to enable the robot to evaluate contingency chains in shorter time steps.

Note that g.c. and l.p. method uses $C^{j}_{i,k}$ (equation (3.12)) and $L_{sa}(\boldsymbol{e})$ (equation (3.7)) instead of $E(\boldsymbol{e})$ which described above, respectively.

3.2.2 Action Producing Unit (APU)

The APU obtains the current state of the robot, and outputs an action to each joint of the robot (see Fig.3.2). The unit consists of three components: 1) Contingency Reproducer(CR), which suggests an action that leads to reproduce found contingency, 2) Reaction Producer(RP), which suggests an action designed to produce a specific reaction to a specific state, and 3) Action Selector, which chooses the outputting actions to each joint of the robot from those suggested.

Contingency Reproducer (CR)

The CR obtains the current state of the robot, and outputs a suggested action to reproduce the found contingency. It is generated by the CDU and added to the APU as mentioned in section 3.2.1. Assume a CR $\pi_{\boldsymbol{e}}$ is created with $\boldsymbol{e} = (r^{t+1}_{k}, s^{t}_{i}, a^{t}_{j}, r^{t}_{k})$ where $r^{t+1}_{k}, r^{t}_{k} \in R_k$, $a^{t}_{j} \in A_j$ and $s^{t}_{i} \in S_i$. Each CR is a sensorimotor mapping, which maps the robot's current state (s^{t}_{i}, r^{t}_{k}) to a suggested action a^{*}; therefore, it is expressed as follows:

$$a^{*} = f(s^{t}_{i}, r^{t}_{k}), \tag{3.13}$$

where $f(s_i^t, r_k^t)$ indicates the sensorimotor mapping of $\pi_{\boldsymbol{e}}$, which outputs a_j^t as a^*. The CR sends the a^* to the Actions Selector together with its predictability Z. We use ${}^{A}E(\boldsymbol{e})$ as the measure reflecting the predictability of CR (see equation (4.6)), because it considers the cases in which the robot has/has not taken the action a^* in the state (s_i^t, r_k^t), and compares their transition probability to the desired observation (i.e., comparing $p(r_k^{t+1}|s_i^t, a^*, r_k^t)$ with $p(r_k^{t+1}|s_i^t, r_k^t)$). If the ${}^{A}E(\boldsymbol{e})$ is high, it means that taking action a^* in the state (s_i^t, r_k^t) increases the probability of achieving desired observation r^{t+1}, or in other words, the predictability of $\pi_{\boldsymbol{e}}$ is high.

To enable comparability with previous works [10,11,48], we mention the CR using another notation $\boldsymbol{\Pi}(R_k^{t+1} = r_k^{t+1}|S_i^t = s_i^t, A_j^t = a_j^t, R_k^t = r_k^t)$, which means that in state (r_k^t, s_i^t) the CR suggests the action a_j^t to reproduce the found contingency and expects the resultant observation r_k^{t+1}. We denote the expected resultant observation of CR with r^*. In addition, to indicate the variables and the values separately, we may use the notifications $\boldsymbol{\Pi}(R_k|S_i, A_j)$ and $\pi(r_k^{t+1}|s_i^t, a_j^t, r_k^t)$ for the CR, respectively. Both of them are labels we use to indicate the contingency reproducer, the former shows the policy on the specific variables, whereas the latter shows it on the specific values.

Note that in l.p. method, $L_{sa}(\boldsymbol{e})$ (equation (3.7)) is used as te predictability Z, because it reflects the (pure) predictability of the experience $\boldsymbol{e}$. For the g.c. method, refer to [11].

Reaction Producer (RP)

The RP obtains the current state of the robot and outputs a suggested action to produce a preprogrammed reaction to a specific state. It sends the suggested action to the Action Selector as well as its predictability Z. An RP sends a constant value α as the predictability. To simplify the quantitative analysis of the research result, in this work we assumed one RP for the system which outputs a random action for any inputting state.

Action Selector

The Action Selector selects one action among candidates from the RP and the CR for each joint of the robot, regarding their predictability Z. Assume that the number of candidate actions generated by the RPs and CRs for a specific joint j are N_j^R and N_j^C, respectively. We use a soft-max action selection method to calculate the probability of selecting the action suggested from component i for joint j:

$$P_i^j = \frac{\exp\left(Z_i/\tau\right)}{\sum_{k \in N_j^R + N_j^C} \exp(Z_k/\tau)}, \tag{3.14}$$

where Z_i is the predictability of component i, and τ is a temperature constant. Note that if Z_i is reduced to less than the omission threshold C_O, the action selector does not consider that component as an input. With this mechanism, the system can refrain from using the CRs that have been incorrectly acquired. The calculated contingency of a CR usually decreases after the acquisition due to the probabilistic feature of the environment such as unpredictability. We set $C_O = C_T - \varepsilon$ to enable the system to tolerate such a feature, where C_T is the acquisition threshold (see section 3.2.1) and ε is a constant value. Furthermore, to avoid the acquisition of contingency chains which consist of a chain of the same actions for the same results, such as executing the same behavior for several time steps in a static environment and having consistent observation, the Action Selector does not use those CRs which do not change the observation, i.e., $r^* = r^t$.

3.3 Computer Simulation

In this section, we discuss the two experiments we conducted to evaluate our proposed mechanism using computer simulation. In the first experiment, we compared the performance of our system with the g.c. and l.p. methods (see section **??**) in terms of precision and recall, i.e., F-measure, learning speed, length of the acquired sequence of skills, and noise tolerance, to determine whether our system is more suitable for real-world implementation. In the second experiment, we implemented our system in a more complex environment to verify its ability to find (more complex) rules produced by a more complex caregiver model, and acquire (more complex) social skills related to these rules.

The basic assumptions used in the experiments were the same as those adopted by Sumioka et al. [11]. The environment used in the experiments is illustrated in Fig.3.1. The environment comprises a robot and a human caregiver sitting across a table, and an object on the table. There are n possible positions on the table for the object, and the caregiver moves it to a random spot every m time steps. In the experiments, we set $(n, m) = (3, 10)$. We set the other simulation parameters as $(\theta, \alpha, \varepsilon) = (8, 0, 0.1)$ based on our experiences (see section 3.2.1, 3.2.2, and 3.2.2 for the parameters, respectively).

3.3.1 The First Experiment

We utilized a simplistic environment for the first experiment because comparison of the performance of the systems would be difficult in a complex environment. We considered a small number of sensory/action/resultant variables for the robot, and a corresponding small number of behavioral patterns for the caregiver model. Consequently, we were able to determine the combinations of variables and values that should be evaluated as a contingent experience. In other words, we were able to compile a list of CRs the robot should acquire in the experiment before performing the computer simulation. This made it possible to evaluate and compare the performance of the systems using F-measure, i.e., the harmonic mean of precision and recall of the skill acquisition algorithm. To enable a fair comparison, we first determined the best threshold parameter C_T for each system. This parameter is denoted as C_T^* and produces the best performance, i.e., the highest F-measure, of the system. Then, we analyzed the learning speed, length of the acquired sequence of CRs, and noise tolerance to demonstrate the level by which our proposed method improved on that achieved previously by others. In this experiment, we set the constant parameter as $\tau = 0.3$ (see section 3.2.2).

Table 3.1: Variable setup for the experiment

Type	Variable Name	Symbol	Elements
$\boldsymbol{S}$	Caregiver's Face	C	$S_1 = \{f_1, f_2, f_3, f_r, f_\phi\}$
	Object	O	$S_2 = \{o, o_\phi\}$
$\boldsymbol{A}$	Gaze shifting	G	$A_1 = \{g_1, g_2, g_3, g_c\}$
	Hand Gesture	H	$A_2 = \{h_1, h_2, h_3, h_4\}$
$\boldsymbol{R}$	Frontal face of caregiver	F	$R_1 = \{0, 1\}$
	Profile of caregiver	P	$R_2 = \{0, 1\}$
	Object	O	$R_3 = \{0, 1\}$

Experimental Setup

The same variables as those used by Sumioka et al. [11] were utilized in this experiment (see Table 3.1). The sensory variable, $\boldsymbol{S}$, consisted of two elements, the visual pattern of the caregiver's face (S_1) and existence of the object (S_2). The action variable, $\boldsymbol{A}$, consisted of two elements, gaze direction (A_1) and hand gesture (A_2). Lastly, the resultant sensory variable, $\boldsymbol{R}$, consisted of three elements, the frontal face of the caregiver (R_1), the profile of the caregiver (R_2), and existence of object (R_3). S_1 could be one of five values: f_1, f_2, and f_3, which indicate the visual pattern of the caregiver's face when the caregiver looks at either of the positions on the table; f_r, which indicates when the caregiver looks at the robot; and f_ϕ, which indicates that the robot was not able to detect the caregiver's face. S_2 could have one of two values: o, which indicates that the robot has detected an object; and o_ϕ, which indicates the opposite situation. A_1 could be one of four values: g_1, g_2, and g_3, which indicate whether the robot is looking at the corresponding spot on the table; and g_c, which indicates that it is looking at the caregiver. A_2 could also have one of four values: h_1 to h_4, which indicate that the robot has made the corresponding hand gestures. If the robot was able to sense the frontal face of the caregiver, profile of the caregiver, and existence of the object, the value of the resultant variables R_1, R_2, and R_3 became one; otherwise, it was set to zero.

Caregiver Model

We adopted simpler rules for the behavior of the caregiver by simplifying the rules used by Sumioka et al. [11] to facilitate easier quantitative analysis of the performance of the system. In each time step, the caregiver either looked at the object or the robot in a random manner (OR-behavior). The only exception was when the robot followed the caregiver's gaze direction. In such a case, the caregiver would look back at the robot with probability P_{LB} (LB-behavior). In this experiment, we set $P_{LB} = 1.0$. This setting enabled us to guess which experience would be acquired as the CR, and consequently, which contingency chains would be detected.

Expected Contingencies

The simplistic implementation of the caregiver model in a simplistic environment allowed us to infer the contingencies that would be found and acquired as CRs by the robot. By analysing all combinations of the variables and their values in terms of whether any dependencies exist among their current and future values, we can find the combinations that should be treated as contingent experiences. Namely, through these analyses, the proposed method is expected to find the following combinations as CRs:

- First, the gaze following skill, GF, is expected to be acquired. Since the caregiver always looks at either the object or the robot, the robot should be able to find an object if it follows the gaze of the caregiver when the caregiver is looking at the object. The CR is $\mathbf{\Pi}_{\mathrm{GF}}(R_3 = 1 | A_1 = g_j, R_3 = 0, S_1 = f_i)$ where $i = j$ and $i = \{1, 2, 3\}$. This contingency is related to the transition from time t to $t+1$; hence, it expresses transition in one time step. Therefore, we define the "level" of this contingency as one. Let S^{Π}_{GF} be a sensory variable which represents whether the robot used $\mathbf{\Pi}_{\mathrm{GF}}$ in the previous time step.

- Second, we expect the gaze returning skill, GR, to be acquired. This is a complex skill in which the robot returns its gaze to the caregiver after using the acquired skill of gaze following, i.e., GF, which leads it to perceive the face of the caregiver from the front. Since the caregiver looks at the robot with high probability ($P_{LB} = 1$) after the GF behavior,

$\mathbf{\Pi}_{\mathrm{GR}}(R_1 = 1|A_1 = g_c, R_1 = 0,\ S^{\Pi}_{\mathrm{GF}} = 1)$ should have a high value for $E(\boldsymbol{e})$ and be found by the CDU. This contingency expresses a transition in two time steps, from $t-1$ to t, and then from t to $t+1$; therefore, the level of this contingency is two. Let S^{Π}_{GR} be a sensory variable that signifies whether the robot used $\mathbf{\Pi}_{\mathrm{GR}}$ in the previous time step.

- The next expected skill, which is named "object looking after gaze returning" (OL), is a complex skill with a level of three. Acquisition of this skill enables the robot to look back at the previous place where the object had been found after using GR and lead it to find the same object again. The CR is $\mathbf{\Pi}_{\mathrm{OL}}(R_3 = 1|A_1 = g_i, R_3 = 0,\ S^{\Pi}_{\mathrm{GR}} = 1)$, where $i = \{1, 2, 3\}$. The g_i is the same as that used in the two preceding steps, i.e., used in GF. This contingency is at level three because it is based on three transitions: from $t-2$ to $t-1$ by GF, from $t-1$ to t by GR, and from t to $t+1$ by itself. Let S^{Π}_{OL} be a sensory variable that signifies whether the robot used $\mathbf{\Pi}_{\mathrm{OL}}$ in the previous time step.

- Finally, a complex skill at level four was also expected. After using OL, keeping the gaze direction at the current place should result in reconfirmation of the current object. This skill is denoted as OL2, with $\mathbf{\Pi}_{\mathrm{OL2}}(R_3 = 1|A_1 = g_i, R_3 = 1,\ S^{\Pi}_{\mathrm{OL}} = 1)$, where $i = \{1, 2, 3\}$. Its contingency belongs to four transition steps: from $t-3$ to $t-2$ by GF, from $t-2$ to $t-1$ by GR, from $t-1$ to t by OL, and from t to $t+1$ by itself, therefore its level is four.

These contingencies are referred to in the evaluation of the system performance, in the next section, as those which should be acquired by the robot. As mentioned above, a consistent and simple behavior rule of the caregiver model in combination with a simplistic environment ensures that the robot is not disturbed in its attempts to detect and acquire the skills. However, there is another contingency which is not counted in the evaluation of the system performance in the next section, because it could be acquired or not. It is named "Object Permanency" (OP) and the CR is mentioned as $\mathbf{\Pi}_{\mathrm{OP}}(R_3 = 1|A_1 = g_i, R_3 = 1,\ S^{\Pi}_{\mathrm{GF}} = 1)$, where $i = \{1, 2, 3\}$. The g_i is the same as that used in a step before, i.e., used in GF. It means that after using GF, keeping the gaze in the same direction leads to reconfirmation of the same object. However, the acquisition of GR disturbs the acquisition of OP, because in the state $S_i = S^{\pi}_{\mathrm{GF}}$, GR outputs $A_1 = a^*_{GR} = g_c$, since OP needs the experience of $A_1 = g_1/g_2/g_3$ in that state, to be experienced and acquired by the robot. Conversely, the acquisition of OP does not disturb the acquisition of GR, because the Action Selector will not choose the output of acquired OP while $r^*_{OP} = r^t_{OP}$ (see section 3.2.2 for the selection prevention algorithm). Therefore, OP is not counted in the evaluation of the next section.

Results and Discussion

In this section, we compare the performance of three different methods: the proposed, g.c., and l.p. methods. The performance is expressed as the accuracy of each method in terms of the acquisition of CRs, and defined as their F-measure:

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}. \tag{3.15}$$

For the calculation of precision and recall, the CRs listed in 3.3.1 are counted as the relevant elements. An acquired CR is regarded as true positive if it is listed in 3.3.1, and false positive otherwise.

However, the number of true positives and false positives are strongly affected by the value of acquisition threshold C_T. A large value of C_T leads to the acquisition of only those experiences with very high contingency as CR (or even not to acquire any CR), whereas setting C_T to a very small value leads to the acceptance of many experiences as contingent ones and the acquisition of all of them as CR. The former decreases the number of true positives, whereas the latter increases the number of false positives; hence, both of them decrease the F-measure of the system. Therefore, we need to determine the best value of C_T for each method, i.e., the value which leads to the highest performance. We mention it with C^*_T, and denote it for each method as C^*_{T1}, C^*_{T2}, and C^*_{T3} for the proposed, g.c., and l.p. methods, respectively. Moreover, the highest performance is denoted with F^*: F^*_1, F^*_2, and F^*_3 for the three methods, respectively.

To find the value of C^*_T, we ran the simulation with different C_T values, from 0 to a very large value (which results in $F = 0$). We ran 30 simulations for each value of C_T, where each run comprised 100,000 time steps. The average F-measure in the 30 runs for each value of C_T is plotted in Fig.3.3(top row). The F-measures of the methods are compared in this graph. C_T^* and F^* are denoted with vertical and horizontal dotted lines, respectively. According to the graphs, we have $F_1^* = 0.85$ with $C_{T1}^* = 0.75$, $F_2^* = 0.51$ with $C_{T2}^* = 0.01$, and $F_3^* = 0.33$ with $C_{T3}^* = 1.7$. Therefore, it could be concluded that the proposed method delivers the highest performance compared with the other methods. The middle row of Fig.3.3 shows the average number of acquired CRs in different C_T. It is categorized to the omitted CRs (see section 3.2.2 about the omission), true positives (T.P.), and false positives (F.P.). Note that the maximum capacity for the number of CRs is set to 100 in the simulation. In the proposed and g.c. methods, increasing C_T from 0 to C_T^* leads the number of F.P. to be decreased, and T.P. to be increased. In addition, the total number of acquired CRs are reduced, which means that with a suitable C_T, i.e., C_T^*, the system could avoid acquiring a huge number of CRs. However, in the l.p. method, the system continues to acquire many CRs ($N_C = 100$) even with C_T^*. According to the graph, most of them are omitted or are F.P., and the ratio of T.P. seems to be small. Consequently, the F^* of the l.p. method is smaller than the corresponding value of the other methods. The bottom row of Fig.3.3 shows the ratio of T.P., F.P. and false negatives (F.N.) of each method with $C_T = C_T^*$, in terms of precision and recall. It is obvious that for the proposed method they are both high (more than 75%); for the g.c. method they are around 50%; and for the l.p. method the number of T.P. is very small, which makes both of them small (less than 25%) and consequently the F-measure of the system is also reduced.

Checking the acquired CRs for the g.c. method indicates that it could usually acquire GF, but difficult to acquire GR, OL, and OL2. According to equation(3.12), the contingency is evaluated among the variables in the g.c. method. Therefore, the averaged dependency of different values of R_k on each value of S_i and A_j is evaluated. However, in GR, OL, and OL2, the averaged dependency would be small, because the dependency only exists among a specific value of R_k on a specific value of S_i and A_j. For example, in OL, the dependency exists only among $R_3^{t+1} = 1$, $R_3^t = 0$, $S_{GR}^t = 1$, and $A_1^t = g_1/g_2/g_3$. Therefore, the g.c. method could not acquire them easily, and consequently its F-measure was unable to attain a high value. In the case of the l.p. method, although it acquires many predictable experiences as contingent ones, as mentioned in section **??** and according to equation(3.8), it could not detect which one

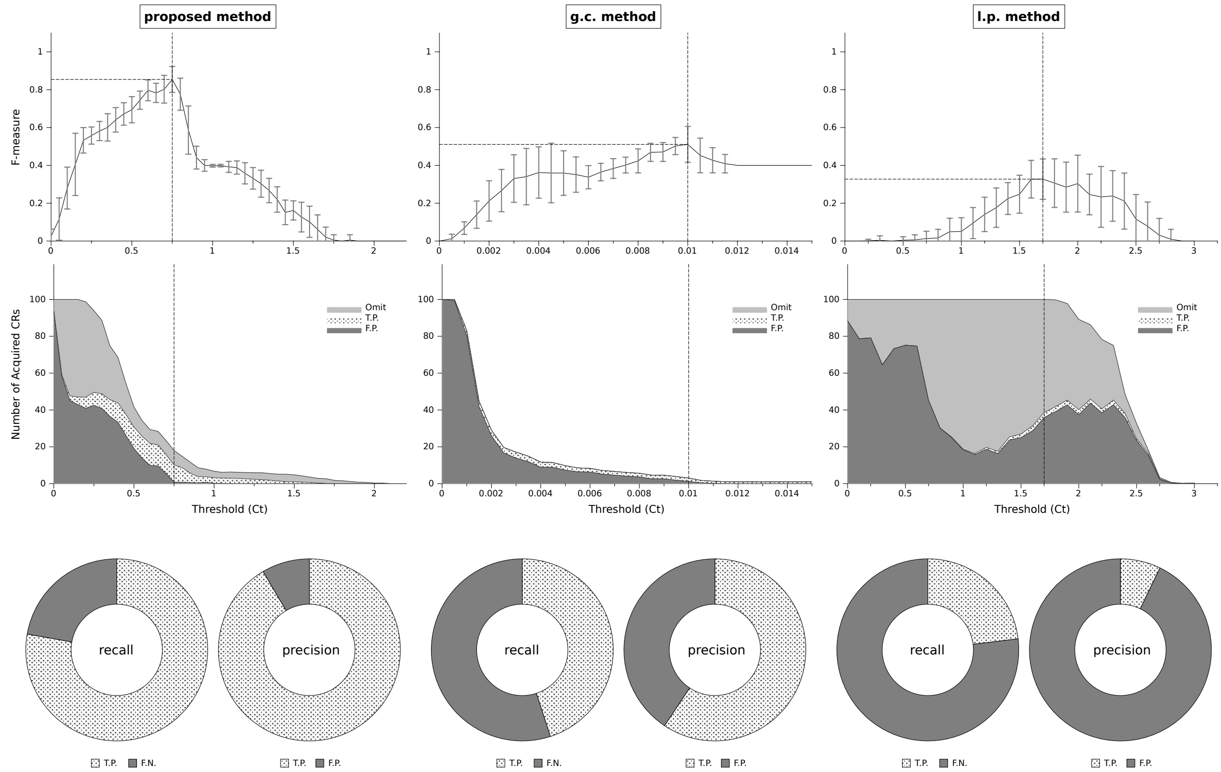


Figure 3.3: *Top*: F-measure of the algorithms. Changing acquisition threshold C_T to find C_T^*. It and F^* are denoted with vertical and horizontal dotted lines, respectively. *Middle*: number of acquired CRs, categorized to omitted ones, true positives (T.P.) and false positives (F.P.) *Bottom*: precision and recall of the methods when $C_T = C_T^*$, mentioned with the ratio of acquired CRs: true positives and false negatives (F.N.) for recall; true positives and false positives for precision.

is dependent on the state of the caregiver. Therefore, the desired CRs, such as GF and GR, could not be distinguished from the others, and as a result the F-measure of the system is very small.

After detecting C_T^* and comparing the performance of the methods, we are now able to compare another important factor for the implementation of the system in real-world robots: the speed of the algorithms. Fig.3.4 shows the speed of acquiring CRs at different levels. We ran the simulation consisting of 30,000 time steps 30 times for each method with $C_T = C_T^*$, and plotted the average remaining time steps to the end of the simulation, when the CR is acquired. For the g.c. method, which uses the global contingency measure, we judged the CR of a level to be acquired when the set of variables (S_i, A_j, R_k) is determined to be contingent using equation (3.12). To ensure a fair comparison, in the case of the proposed and l.p. methods, which use a local contingency measure, the level is judged to have been acquired when all of

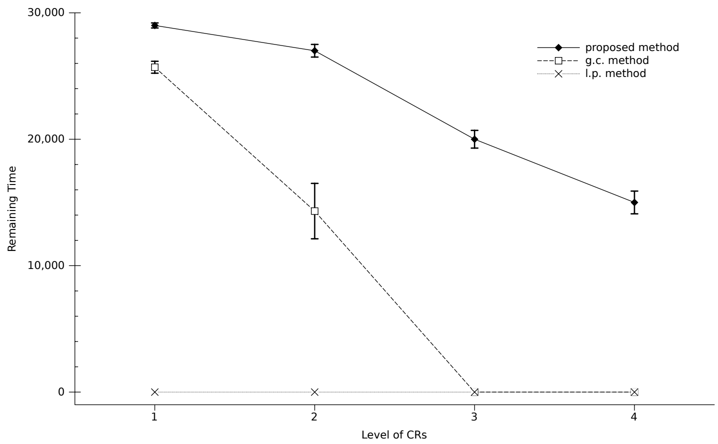


Figure 3.4: Required time steps for the acquisition of skills in each method. The proposed method is four to eight times faster than the g.c. method.

the CRs denoted in the list in section 3.3.1 for each of the levels are obtained by the robot. For example, GF was judged as acquired when the gaze following to the left (g_1), right (g_2), and central (g_3) directions were obtained.

According to the graph, the proposed mechanism acquired CR with level one (i.e., GF) four times faster than the g.c. method, and CR with level two (i.e., GR) eight times faster. Further, the proposed method was able to acquire CRs with levels three (OL) and four (OL2) even in shorter time steps than the g.c. method for CRs with level two, whereas the g.c. method could not acquire CRs with level three and four. The result of the l.p. method is zero for all levels of CRs, because it could not acquire all possible CRs denoted in the list in section 3.3.1, for any of the levels. The reason seems to be shown in the middle graph of Fig.3.3. According to the graph, it acquired many CRs of which most are not true positive, having used full acquisition capacity $N_C = 100$. Therefore, even for the level of one it could not acquire all of the true positives, i.e., the CRs denoted in the list of GF in section 3.3.1. The reason for the late contingency detection of the g.c. method is the same as that described for Fig.3.3: evaluating contingency among the variables. Since there is contingency among specific values of the variables, evaluating the average of the contingency among all of the values of the variables leads to the underestimation of the contingency. Therefore, the system needs to gather more experiences until the average exceeds the threshold value C_T, which requires a larger number of time steps to acquire a CR.

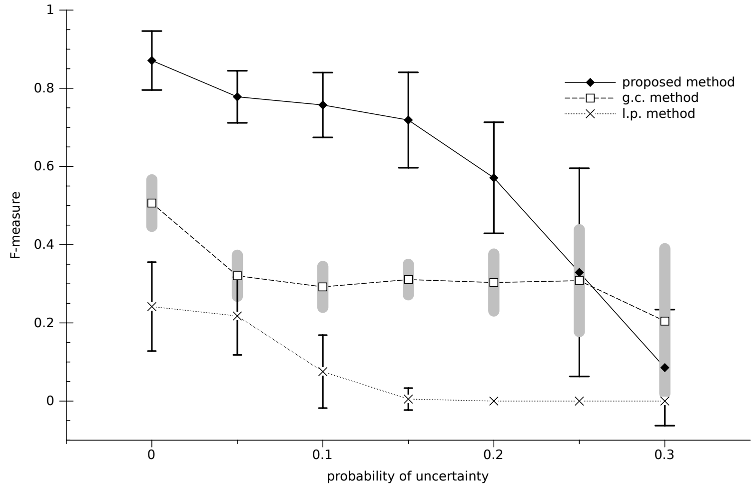


Figure 3.5: Performance of the systems with uncertainty.

Fig.3.5 shows the effect of uncertainty on the performance of the system. It is implemented by considering wrong data/action for both the perception and motor commands of the robot, because we cannot ignore any of these in a real-world robot. The probability of the uncertainty is defined by the variable η. We ran 100,000 time-step simulations 30 times for different values of η, with $C_T = C_T^*$. The average of the F-measure over the 30 runs is plotted in Fig.3.5. As expected, increasing η causes a reduction in the F-measure of the system, for all the methods. Since the contingency is evaluated based on the histogram of the experiences, having wrong data disturbs the histogram and increases the calculation error of the contingencies, which leads the F-measure to be decreased. However, as Fig.3.5 shows, the F-measure of the proposed method is more than the twice as large as those of the other methods when $\eta < 0.25$. Therefore, in real-world implementations, small mistakes in the behavior of the human or the sensors of the robot are expected to be tolerated when our proposed mechanism is used in the implementation.

3.3.2 The Second Experiment

In the previous section, we showed that our system operates faster, finds more complex CRs, and displays a higher resistance against uncertainty compared with the other methods. However, the environment was simplistic: the number of modalities of the robot were small and the contingencies stemmed from a single rule of the caregiver's response to the robot. As a result, the question as to whether the system would be capable of acquiring social behavior in a more

Table 3.2: Additional Variables for the experiment

Type	Variable Name	Symbol	Elements
A	Utterance	U	$A_3 = \{u, u_\phi\}$
R	Utterance of the caregiver	T	$R_4 = \{0, 1\}$
	Smile of the caregiver	M	$R_5 = \{0, 1\}$

complex environment which more closely resembles the interactive environment of humans, arose. In this section, we use a more complex environment to examine our proposed mechanism. The number of modalities are increased and the response of the caregiver is designed with multiple rules relying on the different modalities. In this experiment, we set the constant parameters as $(C_T, \tau) = (0.75, 1.0)$ base on our experiences.

Experimental Setup

We add a new modality and variables to the robot for this experiment (see Table 3.2). Specifically, we enable the robot to utter a sound, and to hear the voice of the caregiver. In addition, we enable the robot to recognize the emotion on the face of the caregiver. Thus, A_3 represents the utterance of the robot, u indicates that the robot utters, whereas u_ϕ indicates that it does not. R_4 represents whether the caregiver uttered, whereas R_5 indicates whether the caregiver is smiling.

Caregiver Model

We increase the behavioral complexity of the caregiver model according to the changes in the variables. In principle, the caregiver behaves in the same way as described in section 3.3.1, but also executes the following additional actions:

1. If the robot uttered to the caregiver after following the gaze of the caregiver, the caregiver responds to the robot by uttering (UU·GF-behavior).

2. After the gaze following behavior of the robot, if the robot returned to the caregiver and kept looking at him/her for a while (here it is one time step), the caregiver utters to the robot (UK·GR-behavior).

Note that in this experiment, the caregiver smiles when the robot looks at the caregiver after the GF behavior, in addition to LB-behavior (see section 3.3.1). We denote this with SLB-behavior.

Expected Contingencies

Since new response rules are added to the behavior of the caregiver model (section 3.3.2), the following behavior is expected to be acquired by the robot in addition to that described in section 3.3.1:

- uttering after GF behavior, which would lead the caregiver to respond to the utterance of the robot (GU behavior)
- keeping the gaze on the caregiver after GR behavior, which would lead the caregiver to utter (KT behavior)

However, a quantitative analysis of the performance of the system, such as in section 3.3.1 is not feasible. This is because there were many parallel causal behavioral actions from the caregiver model in this experiment, and acquisition of one contingency may disturb the acquisition of another one. For example, acquiring the KT behavior may disturb the acquisition of OL behavior, because KT suggests $A_1 = g_c$ as the output a^*, whereas OL suggests $A_1 = g_1/g_2/g_3$. Therefore, instead of performing a quantitative analysis, we inspected the acquired CRs one by one after the simulation to determine what kind of behavior the robot could represent by each of them.

Results and Discussion

We ran the simulation consisting of 10,000 time steps to determine the interaction of the robot with the caregiver model. Table 3.3 shows all 28 of the CRs acquired by the robot in this

Table 3.3: Social skills acquired by the robot.

CR Level	Variables	Symbol	CR Number	Input (r^t, s^t)	Output (r^*, a^*)	$E(\mathbf{e})$	Z	Interpreted Function
1	$\Pi(O\|C, G)$	**GF**	π_1	$(0, f_2)$	$(1, g_2)$	1.50	1.12	Gaze Following
			π_2	$(0, f_1)$	$(1, g_1)$	1.46	1.06	
			π_3	$(0, f_3)$	$(1, g_3)$	1.53	1.10	
2	$\Pi(O\|\mathbf{GF}, G)$	**OP**	π_4	$(1, \pi_2)$	$(1, g_1)$	1.25	1.68	Object Permanency
			π_5	$(1, \pi_3)$	$(1, g_3)$	1.39	1.43	
			π_6	$(1, \pi_1)$	$(1, g_2)$	1.36	1.32	
2	$\Pi(M\|\mathbf{GF}, G)$	**GR**	π_7	$(0, \pi_2)$	$(1, g_c)$	1.58	1.83	Gaze Return
			π_8	$(0, \pi_1)$		2.00		
			π_9	$(0, \pi_3)$		2.30		
2	$\Pi(T\|\mathbf{GF}, U)$	**VR**	π_{10}	$(0, \pi_2)$	$(1, u)$	0.86	2.40	Vocal Response
			π_{11}	$(0, \pi_1)$		0.76		
			π_{12}	$(0, \pi_3)$		0.83		
3	$\Pi(O\|\mathbf{VR}, G)$	**CA**	π_{13}	$(0, \pi_{11})$	$(1, g_2)$	0.75	0.98	Check Again
			π_{14}	$(0, \pi_{10})$	$(1, g_1)$	0.97	0.90	
			π_{15}	$(0, \pi_{12})$	$(1, g_3)$	1.26	1.02	
3	$\Pi(T\|\mathbf{GR}, G)$	**GU**	π_{16}	$(1, \pi_7)$	$(1, g_c)$	1.20	1.21	Get Utterance
			π_{17}	$(1, \pi_8)$		1.22		
			π_{18}	$(1, \pi_9)$		1.21		
3	$\Pi(T\|\mathbf{VR}, G)$	**KT**	π_{19}	$(1, \pi_{10})$	$(1, g_c)$	0.76	0.87	Keep Talking
			π_{20}	$(1, \pi_{11})$		1.06	1.08	
4	$\Pi(T\|\mathbf{CA}, U)$	**VR4**	π_{21}	$(0, \pi_{13})$	$(1, u)$	0.77	0.67	Vocal Response (Lv4)
			π_{22}	$(0, \pi_{14})$		1.10	0.59	
			π_{23}	$(0, \pi_{15})$		0.75	0.79	
4	$\Pi(O\|\mathbf{CA}, G)$	**OP4**	π_{24}	$(1, \pi_{13})$	$(1, g_2)$	1.51	1.48	Object Permanency (Lv4)
			π_{25}	$(1, \pi_{14})$	$(1, g_1)$	1.57	1.68	
			π_{26}	$(1, \pi_{15})$	$(1, g_3)$	1.33	1.53	
4	$\Pi(M\|\mathbf{CA}, G)$	**GR4**	π_{27}	$(0, \pi_{14})$	$(1, g_c)$	1.18	0.47	Gaze Return (Lv4)
5	$\Pi(O\|\mathbf{VR4}, G)$	**CA5**	π_{28}	$(0, \pi_{21})$	$(1, g_2)$	0.85	0.94	Check Again (Lv5)

experiment, which are classified in 11 behavioral categories (note that in this example, the system acquired 34 experiences as CR, but during the interaction 6 of them were omitted which none of them represented any contingency, and 28 skills were remained). The column labeled "CR Level" indicates the length of the contingency chain of each skill, which is described in section 3.2.1. The column "Variables" determines the variables of the CRs. In the column "Symbol" we assigned a symbol to CRs based on the behavior of the robot when it uses the CRs. We use the symbols to indicate the S^{π} of added CR, in the column "Variables" of Table 3.3. In "CR Number" we allocated an ID to each CR. This would enable the value to be used if the CR is the input of another CR. The columns "Input" and "Output" show the input and the output of the sensory-motor mapping of each CR. Furthermore, the value of $E(\boldsymbol{e})$ when the robot acquired the CR and the predictability Z at the end of the simulation is shown for each CR. Finally, an interpretation of the CR is given based on the functionality in the last column. Below, we explain each behavioral type briefly:

- **GF**: This behavior, named Gaze Following, enables the robot to follow the gaze of the caregiver when it detects that the caregiver is looking at a point of the table (when $C = f_1/f_2/f_3$, outputs $G = a^* = g_1/g_2/g_3$). Due to the OR-behavior of the caregiver and infrequent movement of the object ($m = 10$), using GF leads the robot to (usually) find the object($O^{t+1} = r^* = 1$). Therefore, GF appears to be a social skill for finding the object.

- **OP**: This behavior, named Object Permanency, enables the robot to keep its gaze along the same direction, when it used GF behavior and detected an object (when $\mathbf{GF} = 1$ and $O = 1$, outputs $G = a^* = g_1/g_2/g_3$). Due to the infrequent movement of the object ($m = 10$), using OP leads the robot to (usually) see the object again ($O^{t+1} = r^* = 1$). Therefore, OP appears to be a social skill to keep looking at the found object.

- **GR**: This behavior, named Gaze Return, enables the robot to look at the caregiver when the robot used GF behavior (when $\mathbf{GF} = 1$, outputs $G = a^* = g_c$). Due to the SLB-behavior of the caregiver, using GR leads the robot to detect the smiling face of the caregiver ($M^{t+1} = r^* = 1$). Therefore, GR appears to be a social skill for looking back at the caregiver to obtain a prize by smiling, when it succeeded in finding the object (by GF).

- **VR**: This behavior, named Voice Response, enables the robot to utter a sound when it used GF behavior (when **GF** $= 1$, outputs $U = a^* = u$). Due to the UU·GF-behavior of the caregiver, using VR leads the robot to detect the vocal response of the caregiver ($T^{t+1} = r^* = 1$). Therefore, VR appears to be a social skill for uttering a sound to elicit a vocal response from the caregiver, after the robot succeeded in finding the object (using GF).

- **CA**: This behavior, named Check Again, enables the robot to look at the previous location of the object when it used VR behavior (when **VR** $= 1$, outputs $G = a^* = g_1/g_2/g_3$, where g_i would be the same as that used in GF two time steps before). Due to the infrequent movement of the object ($m = 10$), using CA leads the robot to (usually) see the object again ($O^{t+1} = r^* = 1$). Therefore, CA appears to be a social skill to again verify the existence of the found object in the same place, which is detected in the previous time steps (by GF).

- **GU**: This behavior, named Get Utterance, enables the robot to keep looking at the caregiver when it used GR behavior (when **GR** $= 1$, outputs $G = a^* = g_c$). Due to the UK·GR-behavior of the caregiver, using GU leads the robot to detect the utterance of the caregiver ($T^{t+1} = r^* = 1$). Therefore, GU appears to be a social skill to elicit an utterance from the caregiver by continuing to look at the caregiver after receiving a smiling prize (by GR).

- **KT**: This behavior, named Keep Talking, enables the robot to look at the caregiver when it used VR behavior (when **VR** $= 1$, outputs $G = a^* = g_c$). Due to the UK·GR-behavior of the caregiver and GR behavior of the robot in the previous time step (note that when VR behavior is used, GR behavior would be used simultaneously according to the formerly acquired CRs of the robot), using KT leads the robot to detect the utterance of the caregiver ($T^{t+1} = r^* = 1$). Therefore, KT appears to be a social skill in response to the vocal response of the caregiver (due to VR), in which the robot continues looking at the caregiver whereupon the caregiver utters again and which appears to be the ability to induce the caregiver to continue talking to the robot.

- **VR4**: This behavior, named Vocal Response Lv4, enables the robot to utter a sound when it used CA behavior (when $\mathbf{CA} = 1$, outputs $U = a^* = u$). Due to the UU·GF-behavior of the caregiver, using VR4 leads the robot to detect the utterance of the caregiver ($T^{t+1} = r^* = 1$). Therefore, VR4 appears to be a social skill to elicit a vocal response from the caregiver, after the robot succeeded in finding the object (using CA).

- **OP4**: This behavior, named Object Permanency Lv4, enables the robot to maintain its gaze along the same direction, when it used CA behavior and detected (rechecked) an object (when $\mathbf{CA} = 1$ and $O = 1$, outputs $G = a^* = g_1/g_2/g_3$). Due to the infrequent movement of the object ($m = 10$), using OP4 leads the robot to (usually) see the object again ($O^{t+1} = r^* = 1$). Therefore, OP4 appears to be a social skill to enable continued looking at the rechecked object.

- **GR4**: This behavior, named Gaze Return Lv4, enables the robot to look at the caregiver when the robot used CA behavior (when $\mathbf{CA} = 1$, outputs $G = a^* = g_c$). Due to the SLB-behavior of the caregiver, using GR4 leads the robot to detect the smiling face of the caregiver ($M^{t+1} = r^* = 1$). Therefore, GR4 appears to be a social skill for looking at the caregiver again to obtain a prize by smiling, when it succeeded in finding (rechecking) the object (by CA).

- **CA5**: This behavior, named Check Again Lv5, enables the robot to look at the previous location of the object when it used VR4 behavior (when $\mathbf{VR4} = 1$, outputs $G = a^* = g_1/g_2/g_3$, where g_i would be the same as that used in CA two time steps before). Due to the infrequent movement of the object ($m = 10$), using CA5 leads the robot to (usually) see the object again ($O^{t+1} = r^* = 1$). Therefore, CA5 appears to be a social skill to again verify the existence of the found object in the same location detected in the previous time steps (by CA).

Table 3.3 indicates that the proposed method was capable of acquiring several social skills based on the response of the caregiver, even in a more complex environment. Compared with the first experiment, social skills with longer sequences such as CA5 and VR4 are acquired, and more complex interaction between the robot and the caregiver is observed. This suggests that it would be possible to apply our method to real-world interactions between the robot and a human.

To compare the performance of the proposed method under the condition of the second experiment with the other methods, we run 10 independent simulation for each methods, i.e. proposed, g.c. and l.p. method, where each simulation consists of 10,000 time steps. The average of the number of the acquired CRs (NAC) and the maximum level of the used skill during the interaction (Maximum chain length - *MCL*) were compared. The average of NAC was 48.6, 1.7 and 93.9 with the standard deviation 9.1, 0.48 and 6.1 for the proposed, g.c. and l.p. method, respectively. The average of *MCL* was 4.2, 1.5 and 2.6 with the standard deviation 0.42, 0.53 and 0.97 for them, respectively. Larger *MCL* shows that the robot could interact with the human during longer time sequence, and able to produce more complex behaviour. The average of the *MCL* for the proposed method is larger than the others', which ensures its higher performance even in the complex environment of the second experiment. Although the *MCL* of the g.c. method is smaller than the l.p. method, considering NAC of them could be important. In l.p. method, it is close to 100, i.e. the maximum capacity for the skill acquisition as mentioned in section 3.3.1. It means that the same thing with the result of the first experiment could be occurred for l.p. method, even in the condition of the second experiment. It is considered that the system acquires lots of sensory-motor mapping as CRs, while lots of them do not reflect the condition of the caregiver and are useless for the interaction (see section 3.3.1 for detail). On the contrary, in the case of the g.c. method, the system does not acquire lots of CRs, but NAC seems to be too small, i.e. 1.7. This seems to be similar with the result acquired for the g.c. method in the first experiment (see the middle graph of Fig. 3.3 for the case of the g.c. method). It is difficult to acquire complex skills in case of the g.c. method while there is only one non-complex (i.e. simple) contingent experience set for our experiment, i.e. GF. Therefore, the NAC in the case of g.c method became close to the number of the existing simple experiences, i.e. one.

To more analyse the implication of the result with NAC, we considered a fact that we can identify a part of the combinations that does not reflect any contingencies. For example, the combinations which contain A_2 (hand gesture of the robot) can be identified as non-contingent ones because the caregiver is designed not to produce any responses to the hand gesture of the robot in this experiment. We evaluated RNC, which is the ratio of non-contingent combinations included in the all combinations found in the interaction as a measure implying the accuracy of finding contingencies. In the analysis of this ratio, the found combination is identified as a non-contingent one if a skill satisfy either of the following conditions:

- it contains A_2 (hand gesture of the robot),
- it is level one but is not equal to GF (the only level one contingency exists in our problem setting is GF),
- it is level two but its S variable is not equal to GF (the contingency belongs to two time steps appears only after the execution of GF by the robot in our problem setting), or
- it is level three or higher but its S variable is equal to one of the above (non-contingent) skills.

The average of RNC among the ten times of the simulations for the proposed, g.c. and l.p. method were equal to 0.17, 0.10 and 0.58 with the standard deviation 0.09, 0.21 and 0.20, respectively. RNC of the l.p. method is (around two times) larger than the proposed method, and it is similar to the result of the first experiment. It means that lots (around 60%) of acquired skills in l.p. method contain non-contingent variables and are useless for the interaction. For case of the g.c. method, RNC is small. However, since it acquires only one or two skills in this problem setting (NAC is equal to 1.7), it does not mean that the robot with the g.c. method succeeds in acquiring contingent complex skills. As a conclusion, the analysis with RNC and NAC for these three methods implies that the similar properties on the accuracy of them appeared in the first experiment, which were treated as an evidence of higher performance of the proposed method, are reproduced in the second experiment, too.

3.4 Conclusion and Future Work

In this chapter, we proposed a novel mechanism for the acquisition of social skills utilized in face-to-face interaction between a robot and its human caregiver. We introduced a new contingency evaluation measure that estimates contingencies among the value of the variables utilizing transfer information. Further, we showed that our proposed mechanism improves aspects such as system precision and recall, contingency chain length, speed, and noise resistance. We additionally examined the feasibility of our proposed system in a more complex environment that more closely resembles real-world interaction of a robot with humans, and showed that the system remained capable of acquiring complex social skills. The resulting fast, accurate, noise tolerant, and complex skill acquisition by the robot encourages us to take the next step, i.e., to implement the system in an actual real-world robot.

However, the skill acquisition threshold C_T was constant in our simulation. In a real-world interaction of a robot with a human, the value of contingency would vary for different types of modalities and different types of interactions. Although we have started to check the performance of the proposed mechanism with a real-world robot, and confirmed the acquisition of even some complex skills by the robot, but in this primitive experiment the parameters including C_T were tuned very carefully and the behaviour of the human caregiver were very strict. Therefore, for a natural interaction, a mechanism to adaptively regulate the acquisition threshold through interaction seems to be necessary for the implementation of the proposed method in a real-world robot. It is same for the other prefixed parameters, however dynamically adjusting all of the parameters by the system seems to be very complex and time-consuming, therefore discussion about the trade-off of this approach seems to be necessary. Furthermore, to simplify the quantitative analysis of the system performance, RP was considered as a random movement generator in the work described in this chapter. However, for a robot in the real world, more complex or human-like RP such as imitating the caregiver's motions, or orienting to the ostensive signals like motherese is expected to induce more complex response from the human caregiver. Therefore, such a human-like design of the RP is considered to be one of the design issues to further enhance the performance of the proposed mechanism, which might also promote to establish closer relationship between a human and a robot that keeps providing it with contingent experiences necessary for its open-ended development.

Chapter 4

Attention Mechanism for Implementation of Developmental Robot in a Real-world Environment

In this chapter, the way to implement the developmental algorithm mentioned in the previous chapter to a real–world humanoid robot was explained. Two essential problems for the implementation were argued in this chapter and the solution for each problems were proposed. To evaluate the proposed methods, a human–robot interaction experiment in a real–world environment was conducted and the result of the learning performance of the robot was discussed. At the end of the chapter, the subjective evaluation of the participants as well as the works remained for the future research were mentioned.

4.1 Introduction

Joint attention related behaviors (JARBs) include basic social skills, such as following the gaze of others, pointing, intention sharing and social referencing. Humans gradually learn these

This work was proposed in: Mahzoon H, Yoshikawa Y, Ishiguro H. Ostensive-cue Sensitive Learning and Exclusive Evaluation of Policies: A Solution for Measuring Contingency of Experiences for Social Developmental Robot. Frontiers in Robotics and AI. 2019;6:2.

social skills during their developmental process in infancy and childhood [2,44,49], and become able to establish interaction with others. Cosequently, children become able to learn more social skills, such as language communication and mind reading [45]. The importance of JARBs in human infant development [50] has made it one of the most popular research topics in the fields of cognitive science and developmental psychology [1,51,52]. Additionally, owing to the important role of such behaviors in achieving successful communication with humans, some robotic research has focused on the study of JARBs in the development of communicative robots [4–7].

On the other hand, in the field of developmental robotics, several studies based on synthetic approaches have tried to explore and/or reproduce the developmental process of the human infant, as well as to create autonomous developmental robots. See [15] for a review of these efforts. Some of these research has been done on proposing learning mechanisms based on the intrinsic motivation of the robot that enables open-ended development [26–28], and some on dynamic Bayesian networks to evaluate the contingency of the observed events, which enables the robot to plan suitable action(s) to achieve its goal utilizing the evaluated contingency [12–14].

Other studies [8,9] have tried to explain the developmental process of the JARBs of the human infant by using an infant robot. They have focused on the causality of the infant robot's observations, actions and consequent experiences during interaction with a human caregiver. They showed that learning of the causal sensorimotor mapping from gaze patterns of the caregiver to the motor commands of the robot leads the robot to acquire a primitive JARBs, such as gaze following. However, the robot had *a priori* knowledge of the set of sensory and motor variables to be associated in order to acquire such a sensorimotor mapping.

Sumioka et al. [11] proposed an informational measure based on transfer entropy [25], by which the robot become able to automatically distinguish the set of sensory–motor variables for the sensorimotor mapping without such *a priori* knowledge. Additionally, their presented method could evaluate the contingency of a sequence of events, so that the robot became able to learn a sequence of sensorimotor mapping. The contingency of such sequence was defined as *contingency chain* (c-Chain). By using computer simulation, they showed that evaluating the c-Chains of the events led their infant robot model to learn JARBs consisted of sequences of actions, such as *social referencing* behavior. The social referencing was defined as looking back at the caregiver's face after producing the gaze-following behavior. Hereafter, we refer to robot's learned behavior as a *complex skill* if it consists of more than two sequences of actions (such as social referencing behavior), and otherwise refer to it as a *simple skill* (such as gaze-following behavior).

However, numerous time steps were required for the contingency evaluations of previous work [11], especially for complex skills, which resulted in the robot not being able to acquire complex skills in the real-world implementation [48]. Mahzoon et al. [43] proposed a new informational measure based on what they called *transfer information*, which enabled the local evaluation of the contingency among the variable values. They realized a fast contingency evaluation, even with a small number of sample data. They showed that their infant robot model could acquire simple and complex skills within short periods of interaction with the caregiver model, in a computer simulation environment.

Nevertheless, to implement the proposed method on a real-world robot, two basic issues are still remained: First, the synchronization problem of the robot's learning phase with the human caregiver's teaching phase in the real-world interaction was not considered. As a result, the efficiency of the learning process was decreased and therefore unexpectedly delayed. Although understanding and detecting the teaching phase of the human caregiver is not a simple issue, some research on "natural pedagogy" has reported the phenomena of teaching/learning timing of the human caregiver/infant [53] and addressed "ostensive cues" as the key signals of efficient teaching/learning in humans. In this chapter, we propose a new algorithm for robot learning inspired by these phenomena, namely ostensive-cue sensitive learning (OsL), to overcome the synchronization problem. Second, there was overestimation of the contingencies related to actions/observations that occur simultaneously with the usage of a learned behavior. This is due to the confusion of the robot about the cause of the consequent event; the robot could not distinguish whether the reason for the event was the usage of the learned behavior or simply the previous atomic action/observation. To solve this problem, we propose another new algorithm, the exclusive evaluation of policies (XEP), following which the robot evaluates contingencies, so that the calculations related to the atomic variables are separated from those of the learned behaviors.

To evaluate the performance of each proposed algorithm in a real-world environment, we conducted human–robot interaction experiments under four conditions: 1) the previous method [43], i.e. the robot uses neither of the proposed algorithms; 2) the robot uses only the OsL; 3) the robot uses only the XEP; and 4) the proposed method, i.e. the robot uses both the OsL and XEP. Each condition was consisted of 12 subject experiments, and each experiment was taken 800 time steps, i.e. approximately 40 min of interaction with the robot. The performances of the systems was compared in terms of the speed, coverage, and reliability of simple and complex skill acquisition.

On the other hand, as described in [45] and [52], contingent and intelligent behavior of the infant "induces" the caregiver to change its behavior, and teach new concepts to the infant. This inherent tendency of the human caregiver leads to a potential for the open-ended learning and development of the infant, even an infant robot [27]. In our experiment, to evaluate if/how the human subjects feel regarding the infant robot's such intelligence, we conducted a subjective evaluation during the experiment. We asked the subjective opinion of the caregivers about the intelligence of the robot as well as the quality of the interaction. For this, we provided seven questions, each designed with a five-level Likert scale answer. To see the effect of the proposed algorithms on the subjective evaluation, we conducted a statistical analysis of the answers. The result of the analysis is discussed in section 4.4.5

4.2 Problem Setting and Contingency Evaluation

4.2.1 Interaction Environment

A face-to-face interaction between a human caregiver and an (infant) robot is assumed as our experimental environment (Fig.3.1 in the previous chapter). There is a table between them and one or more objects are placed on the table. The human caregiver plays and interacts with the robot (based on their own strategy, if any) and can move the position of the objects on the table. The robot discretizes time. At each time step t, the robot observes the environment and stores the observed data in the sensory variables $\boldsymbol{S}^t = (S_1^t, S_2^t, \cdots, S_{N_S}^t)^T$, where N_S denotes the number of sensory variables. We also refer to these by "state variable" in this work. After the observation, it sends action commands to its joints and saves them to the action variables $\boldsymbol{A}^t = (A_1^t, A_2^t, \cdots, A_{N_A}^t)^T$, where N_A denotes the number of action variables, which would be equal to the number of the joints of the robot. Next, the robot observes the result of the taken action, and saves the resultant observations to the resultant variables: $\boldsymbol{R}^t = (R_1^t, R_2^t, \cdots, R_{N_R}^t)^T$ for the values of the resultant observation before taking the action, and $\boldsymbol{R}^{t+1} = (R_1^{t+1}, R_2^{t+1}, \cdots, R_{N_R}^{t+1})^T$ for after taking the action, where N_R denotes the number of the resultant variables. In the remainder of this section, we summarize and introduce the basic idea of the contingency evaluation mechanism of our previous work [43].

4.2.2 Finding and Reproducing Contingency

Assume that in time step t, the robot observes s_i^t and r_k^t, takes the action a_j^t, and as result, observes r_k^{t+1}; here, s_i^t, a_j^t, r_k^t, and r_k^{t+1} indicate the values of the variables S_i^t, A_j^t, R_k^t, and R_k^{t+1}, respectively. The quaternion $\boldsymbol{e} = (s_i^t, a_j^t, r_k^t, r_k^{t+1})$ represents such an experience of the robot, and is simply denoted as *experience* in this work. An experience $\boldsymbol{e}$ contains information about "*when* (s_i^t), *what to do* (a_j^t), *for which transition* (r_k^t to r_k^{t+1})". During the interaction with the human, the robot evaluates the "contingency" of its experiences, which will be described later, and distinguishes the "contingent" ones. After finding the contingent experience(s), the robot tries to "reproduce" it by acquiring a suitable sensorimotor mapping that enables the robot to take suitable action a_j^t in the specific state s_i^t to reproduce the specific transition of r_k^t to r_k^{t+1}. Inspired by previous works on human infant behaviors concerning the process of finding and reproducing interaction contingencies [54], even with a contingently responsive robot [55], in our work, the ability to reproduce the contingency of an interaction is considered to be one of the most essential social skills for an interactional robot, which makes it able to interact properly with the interacting human.

To evaluate the contingency of the experiences, the robot updates and saves histograms of the values of the variables in each step of the interaction, and calculates the following probabilities. Assume there are two discrete-time stochastic processes X and Y, which can be approximated by stationary Markov processes. The transitions of the processes from time t to $t+1$ can be represented by the transition probabilities $p(x^{t+1}|x^t)$ and $p(y^{t+1}|y^t)$, where the notifications x^t, y^t and x^{t+1}, y^{t+1} indicate the values of the processes at times t and $t+1$, respectively. The contribution of a specific value of process Y, such as y^t, on the transition of the process X from a specific value such as x^t to a specific value x^{t+1} can be estimated using *transfer information* [43]:

$$I_{y \to x} = \log \frac{p(x^{t+1}|x^t, y^t)}{p(x^{t+1}|x^t)}. \tag{4.1}$$

For an experience $\boldsymbol{e}$, the transfer information can be adopted as follows to evaluate the contingency of the experience, i.e. the contribution of the action a_j^t in state s_i^t to the transition of r_k^t to r_k^{t+1}, or in other words the joint contribution of the state and action in experience $\boldsymbol{e}$:

$$C_J(\boldsymbol{e}) = I_{(s_i,a_j) \to r_k} = \log \frac{p(r_k^{t+1}|s_i^t, a_j^t, r_k^t)}{p(r_k^{t+1}|r_k^t)}. \tag{4.2}$$

Additionally, the single contributions of the state and action in experience $\boldsymbol{e}$ can be calculated as follows:

$$C_S(\boldsymbol{e}) = I_{s_i \to r_k} = \log \frac{p(r_k^{t+1}|s_i^t, r_k^t)}{p(r_k^{t+1}|r_k^t)}, \tag{4.3}$$

$$C_A(\boldsymbol{e}) = I_{a_j \to r_k} = \log \frac{p(r_k^{t+1}|a_j^t, r_k^t)}{p(r_k^{t+1}|r_k^t)}. \tag{4.4}$$

The purpose of the robot is to evaluate the joint contribution in experiences to know if the action a_j^t in state s_i^t specifically leads to the consistent result r_k^{t+1}, and acquire a sensorimotor mapping of s_i^t to a_j^t. However, the value of equation (4.2) can be also large when the value of the single contribution of either the state or action becomes large. Therefore, the joint contribution needs to be compared with the single contributions to distinguish the experiences in which the transition to r_k^{t+1} is due to both s_i^t and a_j^t, and not simply one of them. It can be estimated as follows:

$$
\begin{aligned}
{}^S\widetilde{C}_J(\boldsymbol{e}) &= C_J(\boldsymbol{e}) - C_S(\boldsymbol{e}) \\
&= \log \frac{p(r_k^{t+1}|s_i^t, a_j^t, r_k^t)}{p(r_k^{t+1}|s_i^t, r_k^t)},
\end{aligned} \tag{4.5}
$$

$$
\begin{aligned}
{}^A\widetilde{C}_J(\boldsymbol{e}) &= C_J(\boldsymbol{e}) - C_A(\boldsymbol{e}) \\
&= \log \frac{p(r_k^{t+1}|s_i^t, a_j^t, r_k^t)}{p(r_k^{t+1}|a_j^t, r_k^t)},
\end{aligned} \tag{4.6}
$$

where ${}^S\widetilde{C}_J(\boldsymbol{e})$ and ${}^A\widetilde{C}_J(\boldsymbol{e})$ compare the joint contribution with the single contribution of the state and action, respectively. Finally, the measure named *synergistic contribution of contingencies* (ScC) is proposed as follows to distinguish the "*contingent*" experiences, i.e. the experiences in which the combination of the state and the action is the cause of the transition, but not either of them is individually the cause:

$$\widetilde{C}_J(\boldsymbol{e}) = \min\{{}^S\widetilde{C}_J(\boldsymbol{e}), {}^A\widetilde{C}_J(\boldsymbol{e})\}. \tag{4.7}$$

When the value of $\widetilde{C}_J(\boldsymbol{e})$ of a specific experience $\boldsymbol{e}$ becomes larger than a specific threshold C_T for a predefined duration, such as θ time steps, the robot distinguishes it as a contingent experience (or simply, a contingency) and acquires the sensorimotor mapping (s_i^t, a_j^t). Then, it starts to "reproduce" the found contingency by "using" the acquired sensorimotor mapping. The sensorimotor mapping learned based on the experience $\boldsymbol{e}$ is denoted as the policy π. During interaction with the human, the robot may acquire several different policies.

4.2.3 Evaluating the Contingency Chain

After the acquisition of a new m-th policy π_m, the robot adds a new Boolean variable S_{π_m} to the set of state variables, which indicates whether the policy π_m was used. It takes the value 1 if π was used, and 0 otherwise. To avoid confusion, we also denote the value of the S_{π_m} with $\bar{\pi}_m$ when it takes the value 0, and with π_m when it is 1. Then, the robot continues updating the histograms of the variables as well as calculating the contingency of the experiences, including the new state variable S_{π_m}. Using this method, the robot becomes able to evaluate the contingency of the c-Chains, and as a result, evaluate the contingency related to the new behavior of the caregiver who observed the contingency reproduction of the robot. In previous work [43], an example of such a c-Chain was the consistent response of the caregiver to the social referencing behavior of the robot: the robot found that after using the gaze-following skill, if it looks at the caregiver's face, the caregiver will look at the face of the robot as an acknowledgement. Moreover, they showed that in a more complex simulation environment, the robot acquires a longer sequence of actions, up to five sequences.

4.3 Proposed Method

In this section, after discussing the two essential weak points of the previous work [43] and our solution for each of them, we describe the mechanism of our proposed method.

4.3.1 Ostensive-cue Sensitive Learning (OsL)

The first problem of previous work is the synchronization of the teaching phase of the human caregiver with the learning phase of the infant robot. Learning under the non-synchronized environment decreases the learning efficiency of the robot, and causes significant delays in the learning progress. Although distinguishing the teaching phase of the human by the robot seems to be a difficult issue owing to the probable variety of types of teaching in different human subjects, there are several reports in the fields of cognitive science and developmental psychology regarding how human infants treat the synchronization problem and increase the efficiency of learning from adults (see a review [16]).

Csibra and Gergely addressed the "natural pedagogy" as a human communication system for generic knowledge transmission between individuals [53]. They proposed that human infants are "prepared to be at the receptive side of natural pedagogy" and sensitive to learn from the ostensive cues of human adults, such as mutual eye contact between the adults and the infant, or adults' infant-directed speech (motherese). From this statement, we hypothesize that the human adult may inherently or adaptively output the ostensive cues when it tries to teach something to the human infant, or even to the infant robot. Based on this hypothesis, we propose the OsL algorithm for the infant robot as follows: 1) The robot stops moving when it observes an ostensive cue from the human and continues the observation of the human until the signal disappears. This is because the ostensive cue acts as a signal (from our hypothesis) that informs the robot about the human's teaching phase, and notifies the robot to synchronize with it; 2) The robot counts the histogram of the consequent experiences right after the ostensive cue η times (i.e. the learning weight parameter of the OsL algorithm) instead of one time in order to emphasize such experiences. This is because (from our hypothesis) after the ostensive signals, the human would be in the teaching phase and the experiences right after the ostensive cues probably contain more informative concepts compared with other experiences. Using OsL, we expect the robot to increase the efficiency of learning and, as a result, the speed of skill acquisition.

4.3.2 Exclusive Evaluation of Policy (XEP)

The second problem of the previous work is the overestimation of the transition probabilities of the single contingencies, which leads to an underestimation of ${}^{S}\widetilde{C}_J$ and/or ${}^{A}\widetilde{C}_J$, i.e. equations (4.5) and (4.6), when the robot uses an acquired policy. This leads to the underestimation of the ScC of some experiences, i.e. $\widetilde{C}_J$: equation (4.7). The reasons for the overestimation and the underestimation are as follows. Assume that the robot acquired its new m-th policy π_m based on the contingent experience $\boldsymbol{e}_m = (s_i^t, a_j^t, r_k^t, r_k^{t+1})$. Before the robot starts to use π_m, i.e. using the sensorimotor mapping (s_i^t, a_j^t), the ${}^{S}\widetilde{C}_J$ and ${}^{A}\widetilde{C}_J$ of the experience $\boldsymbol{e}_m$ can be written by the transition probabilities calculated based on the histograms of the variables before acquiring and using π_m, i.e. p^{bef}, as follows:

$$ {}^{S}\widetilde{C}_{J}^{\text{bef}}(\boldsymbol{e}_m) = \log \frac{p^{\text{bef}}(r_k^{t+1}|s_i^t, a_j^t, r_k^t)}{p^{\text{bef}}(r_k^{t+1}|s_i^t, r_k^t)}, \tag{4.8} $$

$$ {}^{A}\widetilde{C}_{J}^{\text{bef}}(\boldsymbol{e}_m) = \log \frac{p^{\text{bef}}(r_k^{t+1}|s_i^t, a_j^t, r_k^t)}{p^{\text{bef}}(r_k^{t+1}|a_j^t, r_k^t)}. \tag{4.9} $$

However, when the robot starts to use π_m, the probability of taking action a_j^t in state s_i^t increases. This fact increases the value of the transition probabilities 1)$p(r^{t+1}|s_i^t, r_k^t)$ and 2)$p(r^{t+1}|a_j^t, r_k^t)$, i.e. the numerator of the single contingencies: equations (4.3) and (4.4); and the denominator of ${}^{S}\widetilde{C}_J$ and ${}^{A}\widetilde{C}_J$: equations (4.5) and (4.6). The reasons are 1) for $p(r^{t+1}|s_i^t, r_k^t)$: in state s_i^t, the probability of taking action a_j^t increases owing to the usage of π_m, which is a contingent skill and leads the transition to r_k^{t+1} with high probability; and 2) for $p(r^{t+1}|s_i^t, r_k^t)$: the probability of having been in state s_i^t when the action a_j^t is taken increases ownig to the usage of π_m. Assume that the values of the transition probabilities $p(r^{t+1}|s_i^t, r_k^t)$ and $p(r^{t+1}|a_j^t, r_k^t)$ after the usage of π_m, i.e. denoted by p^{aft}, increase by factors of α and β, respectively, compared to p^{bef}:

$$ p^{\text{aft}}(r_k^{t+1}|s_i^t, r_k^t) = \alpha.\ p^{\text{bef}}(r_k^{t+1}|s_i^t, r_k^t)\ ;\ \alpha > 1 \tag{4.10} $$

$$ p^{\text{aft}}(r_k^{t+1}|a_j^t, r_k^t) = \beta.\ p^{\text{bef}}(r_k^{t+1}|a_j^t, r_k^t)\ ;\ \beta > 1 \tag{4.11} $$

Assuming that the value of the transition probability $p(r_k^{t+1}|s_i^t, a_j^t, r_k^t)$ does not change before and after the usage of π_m (because the usage of π_m as a sensorimotor mapping (s_i^t, a_j^t) is included in the condition part of the transition probability), the values of ${}^{S}\widetilde{C}_J$ and ${}^{S}\widetilde{C}_J$ for the experience

$\boldsymbol{e}_m$ after the usage of π_m can be written as:

$$
\begin{aligned}
{}^{S}\widetilde{C}_J^{\,\mathrm{aft}}(\boldsymbol{e}_m) &= \log \frac{p^{\mathrm{bef}}(r_k^{t+1}|s_i^t, a_j^t, r_k^t)}{\alpha.\ p^{\mathrm{bef}}(r_k^{t+1}|s_i^t, r_k^t)} \\
&= \ {}^{S}\widetilde{C}_J^{\,\mathrm{bef}}(\boldsymbol{e}_m) - \log \alpha \qquad ;\ \alpha > 1,
\end{aligned} \tag{4.12}
$$

$$
\begin{aligned}
{}^{A}\widetilde{C}_J^{\,\mathrm{aft}}(\boldsymbol{e}_m) &= \log \frac{p^{\mathrm{bef}}(r_k^{t+1}|s_i^t, a_j^t, r_k^t)}{\beta.\ p^{\mathrm{bef}}(r_k^{t+1}|a_j^t, r_k^t)} \\
&= \ {}^{A}\widetilde{C}_J^{\,\mathrm{bef}}(\boldsymbol{e}_m) - \log \beta \qquad ;\ \beta > 1.
\end{aligned} \tag{4.13}
$$

Therefore, ScC of the experience $\boldsymbol{e}_m$ after the usage of the π_m will become:

$$
\begin{aligned}
\widetilde{C}_J^{\,\mathrm{aft}}(\boldsymbol{e}_m) &= \min\{{}^{S}\widetilde{C}_J^{\,\mathrm{bef}}(\boldsymbol{e}_m) - \log \alpha,\ {}^{A}\widetilde{C}_J^{\,\mathrm{bef}}(\boldsymbol{e}_m) - \log \beta\} \\
&< \widetilde{C}_J^{\,\mathrm{bef}}(\boldsymbol{e}_m).
\end{aligned} \tag{4.14}
$$

To avoid such an underestimation, we propose to separate the contingency evaluations related to the acquired policies and atomic variables, namely the XEP algorithm. In this algorithm, the system adds an *extra* variable for each sensory and action variable to the system, denoted by $\widehat{S}_i^t$ and $\widehat{A}_j^t$. When an acquired policy π_m is used, the system sets the values of $\widehat{S}_i^t$ and $\widehat{A}_j^t$ to *don't care*. Therefore, the histogram of the values of these variables, denoted by $\hat{s}_i^t$ and $\hat{a}_j^t$, are counted only if an acquired policy has not been used. Using the histogram of these variables for the calculation of the transition probabilities of equations (4.10) and (4.11), which are denoted by $\hat{p}$, causes them not to increase even after usage of the policy π_m:

$$\begin{aligned}\hat{p}^{\,\mathrm{aft}}(r_k^{t+1}|s_i^t, r_k^t) &= p^{\mathrm{aft}}(r_k^{t+1}|\hat{s}_i^t, r_k^t) \\ &= p^{\mathrm{bef}}(r_k^{t+1}|s_i^t, r_k^t),\end{aligned} \tag{4.15}$$

$$\begin{aligned}\hat{p}^{\,\mathrm{aft}}(r_k^{t+1}|a_j^t, r_k^t) &= p^{\mathrm{aft}}(r_k^{t+1}|\hat{a}_j^t, r_k^t) \\ &= p^{\mathrm{bef}}(r_k^{t+1}|a_j^t, r_k^t).\end{aligned} \tag{4.16}$$

Therefore, when the XEP algorithm is used, the value of ${}^S\widetilde{C}_J$ and ${}^A\widetilde{C}_J$ for the experience $\boldsymbol{e}_m$, which are denoted by ${}^S\widehat{C}_J$ and ${}^A\widehat{C}_J$, after the usage of π_m will be:

$$\begin{aligned}{}^S\widehat{C}_J^{\,\mathrm{aft}}(\boldsymbol{e}_m) &= \log \frac{p^{\mathrm{bef}}(r_k^{t+1}|s_i^t, a_j^t, r_k^t)}{\hat{p}^{\,\mathrm{aft}}(r_k^{t+1}|s_i^t, r_k^t)} \\ &= {}^S\widehat{C}_J^{\,\mathrm{bef}}(\boldsymbol{e}_m),\end{aligned} \tag{4.17}$$

$$\begin{aligned}{}^A\widehat{C}_J^{\,\mathrm{aft}}(\boldsymbol{e}_m) &= \log \frac{p^{\mathrm{bef}}(r_k^{t+1}|s_i^t, a_j^t, r_k^t)}{\hat{p}^{\,\mathrm{aft}}(r_k^{t+1}|a_j^t, r_k^t)} \\ &= {}^A\widehat{C}_J^{\,\mathrm{bef}}(\boldsymbol{e}_m).\end{aligned} \tag{4.18}$$

As the result, the ScC of the experience $\boldsymbol{e}_m$ when the XEP algorithm is used, which is denoted by $\widehat{C}_J$, after the usage of π_m will be:

$$\begin{aligned}\widehat{C}_J^{\,\mathrm{aft}}(\boldsymbol{e}_m) &= \min\{{}^S\widehat{C}_J^{\,\mathrm{aft}}(\boldsymbol{e}_m),\ {}^A\widehat{C}_J^{\,\mathrm{aft}}(\boldsymbol{e}_m)\} \\ &= \widehat{C}_J^{\,\mathrm{bef}}(\boldsymbol{e}_m).\end{aligned} \tag{4.19}$$

With respect to equation (4.19) and inequation (4.14), it can be concluded that the XEP algorighm is able to solve the underestimation problem of the previous work [43], and is expected to increase the accuracy of the contingency evaluation[1].

[1]For the same reason, the system also uses the *extra* variables $\widehat{S}_i^t$ and $\widehat{A}_j^t$ when the robot has used the policy

4.3.3 Mechanism

Fig.5.3 shows the schema of the proposed system. It consists of two main parts: the Contingency Detection Unit (CDU) and the Action Producing Unit (APU). The APU is responsible for determining the output action in each time step, while the CDU evaluates the contingency of the experiences. At each time step t, the robot observes the environment and stores the results of the current observation in $\boldsymbol{S}^t$ and $\boldsymbol{R}^t$ (bottom part of the figure). They are sent to the APU, and the APU decides about the outputting action for each joint of the robot $\boldsymbol{A}^t$, based on the input data $\boldsymbol{S}^t$ and $\boldsymbol{R}^t$. After taking the action, the robot again observes the environment, and stores the resultant observation in the resultant variable $\boldsymbol{R}^{t+1}$ (bottom part of the figure). Simultaneously, in each time step, the CDU gets the values of all of the variables, and evaluates the contingency of the experiences. If the CDU detects an experience as a contingent one, it adds a new Contingency Reproducer (CR in Fig.5.3) to the APU, which enables the APU to reproduce the found contingency. In the remainder of this section, each component of the CDU and APU are explained in detail.

Contingency Detection Unit (CDU)

In each time step, the CDU 1) evaluates the contingency of the experiences, and 2) if a contingent experience is detected, it adds a new CR to the APU, which enables the robot to reproduce the found contingency. The CDU consists of three components: the Contingency Evaluator, Ostensive Signal Detector (OS-D), and the Skill Usage Detector (SU-D).

Contingency Evaluator This unit calculates the contingencies of the experiences based on the histograms of the experiences, using the method described in section 4.2.2. If the experience $\boldsymbol{e} = (s_i^t, a_j^t, r_k^t, r_k^{t+1})$ is distinguished as a contingent one, it adds a new CR to the APU, which contains the values of the variables of the found contingent experience $\boldsymbol{e}$, i.e. s_i^t, a_j^t, r_k^t and r_k^{t+1}. After that, the Contingency Evaluator continues the evaluation of the contingencies, including the c-Chains (see section 4.2.3), as well as the process of adding further CRs to the system.

in the previous time step, i.e. when $S_{\pi_m} = 1$.

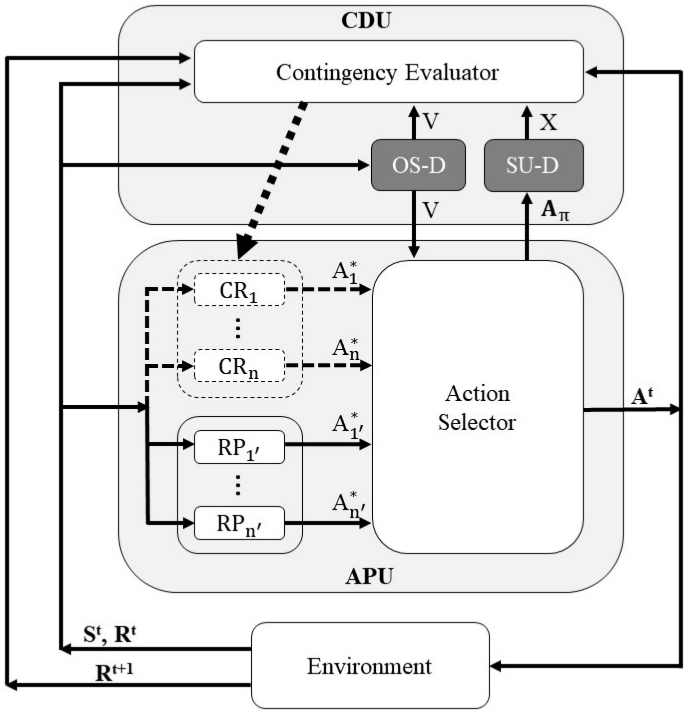


Figure 4.1: System schema of the proposed mechanism. The new components of the proposed algorithms are shown with the darker color. In each time step, the robot outputs the action A^t based on its current states S^t and R^t, and observes the resultant transition of the environment, i.e. R^{t+1}.

OS-D The OS-D gets the current state of the robot (S_i^t and R_k^t). If it detects that these variables include an ostensive cue from the human, it sends the *stop signal* V to the Contingency Evaluator as well as the Action Selector. This signal causes the Contingency Evaluator to pause counting the histograms, and the Action Selector to make the robot to keep looking at the human and stop its movement. Additionally, it sends the learning weight parameter η (see section 4.3.1) to the Contingency Evaluator. When the ostensive cue disappears, the signal V is cancelled, after which the Contingency Evaluator and Action Selector restart their functions. In this work, mutual eye contact with the human caregiver is implemented as the only ostensive cue of the interaction.

SU-D The SU-D gets the information regarding the usage of the policies in each time step from the Action Selector, and informs the Contingency Evaluator if any policy has been used at the current moment. To this end, the SU-D gets the values of the Boolean variable A_{π_m} from the the Action Selector, which indicates if the m-th policy is currently used, and sends the Boolean signal X to the Contingency Evaluator, which is calculated as follows:

$$X = \bigvee_{m=1}^{N_\pi} A_{\pi_m} \ , \tag{4.20}$$

where N_π denotes the number of the policies that the robot has acquired until now. If the value of the signal X is true, the Contingency Evaluator sets the value of the extra variables $\widehat{S}_i^t$ and $\widehat{A}_j^t$ to *don't care*, as described in section 4.3.2.

Action Producing Unit (APU)

As shown in Fig.5.3, the APU is equipped with three components, the Reaction Producers (RP), Contingency Reproducers (CR), and Action Selector. At the beginning of the interaction, the APU contains no CRs and selects the actions of the robot at each time steps from the suggested actions of the RPs, denoted by $A_{1'}^*$ to $A_{n'}^*$ in Fig.5.3 where n' indicates the number of RPs in the system. Continuing the interaction with the caregiver leads the CDU to find contingent experiences and add CRs to the APU, which include specific sensorimotor mappings, as described in section 4.3.3. Similar to the RPs, the CRs send their suggested actions to the Action Selector, denoted by A_1^* to A_n^* in Fig.5.3, where n indicates the number of CRs acquired by the robot. Therefore, after adding CRs to the system, the Action Selector needs to choose the outputting action command to each joint of the robot from all of the candidates: $A_m \in \{A_1^*, A_2^*, \cdots, A_n^*, A_{1'}^*, A_{2'}^*, \cdots, A_{n'}^*\}$ where m indicates the m-th joint of the robot.

Contingency Reproducer (CR) The CR gets the current state of the robot and outputs its suggested action to the Action Selector, based on its sensorimotor mapping. Additionally, it sends the reliability Z to the Action Selector, which indicates the certainty of the transition to the expected state if the Action Selector selects its suggested action as the output action of the robot. Assume the m-th CR was added to the system based on the contingent experience $\boldsymbol{e}_m = (s_i^t, a_j^t, r_k^t, r_k^{t+1})$. If the current state S_i^t and R_k^t are the same as s_i^t and r_k^t of the CR, it outputs the candidate action a_j^t to the Action Selector. Otherwise, it does not send any candidate. In this work, the CR sends the ScC of the experience $\boldsymbol{e}_m$, i.e. $\widehat{C}_J(\boldsymbol{e}_m)$, as its reliability Z_m to the Action Selector.

Reaction Producer (RP) The RP gets the current state of the robot and outputs a pre-programmed reaction, which is sent to the Action Selector as the suggested action of the RP. Also it sends a constant value α_m as its reliability Z_m to the Action Selector, where m indicates the m-th RP. For the sake of simplicity, in this work we considered only one RP for the system, which outputs a random action for any input state.

Action Selector The Action Selector chooses the output action for each joint of the robot at each time step. A soft-max action selection was utilized to choose the output from the candidates. Assume that for the j-th joint of the robot, the number of RPs and CRs which send the candidate action to the Action Selector, namely inputting components, are N_j^R and N_j^C, respectively. At each time step, the probability of selecting the suggested action of the inputting component i for the joint j is calculated based on their reliability as follows:

$$P_i^j = \frac{\exp\left(Z_i/\tau\right)}{\sum_{k \in N_j^R + N_j^C} \exp(Z_k/\tau)}, \tag{4.21}$$

where Z_i indicates the reliability of the inputting component i, and τ is a temperature constant. Note that if Z_i is less than the omission threshold C_O, the Action Selector does not consider the inputting component i in equation (4.21) and P_i^j for that component is set to zero. This mechanism enables the robot to have a chance to omit any acquired skill, which might be acquired owing to the noise, lack of sufficient experiences, or other error factors. We set $C_O = C_T - \varepsilon$, where the C_T is the skill acquisition threshold (see section 4.2.2), and ε is a constant value. Additionally, when more than two CRs with the same suggested action and different c-Chain length exist in the inputting components, the Action Selector considers only the one with the longer c-Chain length as the inputting component, and ignores the others, i.e. sets their P_i^j values to zero.

When the suggested output of the m-th CR with the policy π_m is selected as the output, the Action Selector sets the value of the Boolean variable A_{π_m} to 1. It sends A_{π_m} to the SU-D in each time step to inform the SU-D about the usage of the skills. Also, when the Action Selector gets the stop signal V from the OS-D, it stops outputting new action commands to the joints of the robot until the stop signal disappears.

4.4 Experiment and Result

In this section, the results of the real-world robot experiment with human subjects are reported. To evaluate the effect of the proposed methods, i.e. the XEP and OsL algorithms, the performances of four different learning mechanisms are compared, of which the CDU consists of (1) neither the SU-D nor the OS-D, (2) only the SU-D, (3) only the OS-D, and (4) both the SU-D and the OS-D. In the remainder of this chapter, they are referred to as the previous method, XEP method, OsL method, and proposed method, respectively. This study was carried out in accordance with the recommendations of the ethics committee for research involving human subjects at the Graduate School of Engineering Science, Osaka University. The protocol was approved by the ethics committee for research involving human subjects at the Graduate School of Engineering Science, Osaka University. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

4.4.1 Subjects, Apparatus and Procedure

Fig.4.2 shows the environment of the experiment, which was designed based on the concepts explained in section 4.2.1. The human subject was asked to sit opposite the humanoid infant robot and interact with it naturally, as when he/she interacts with a human infant. The subject was asked to play with the robot using a toy on the table and draw the attention of the robot to the toy by teaching the current position of the toy as well as the name, color, shape or other features of it. It is explained to the subject that the robot may learn some social skills from the behavior of the subject, and start to use them. When the robot uses a learned skill, the LEDs on the face of the robot turn on temporarily. The subject was asked to praise the robot by hitting a specific key on the keyboard when the robot finds the toy by using an acquired

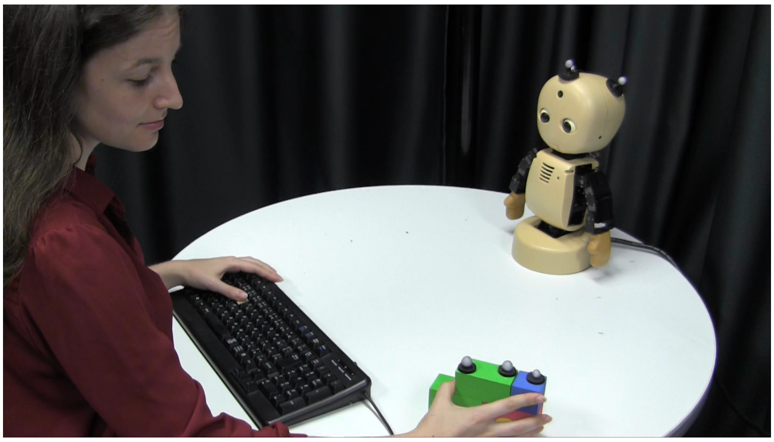


Figure 4.2: The environment of the subject experiment. The subjects were asked to teach the current position of the toy to the robot. Also, they were asked to push a button of the keyboard to express that they are smiling and praising the robot at the moment. The consent for publication of this image was obtained from the participant of this image by using a written informed consent.

skill, i.e. when the LEDs turn on. Additionally, he/she was asked to change the position of the toy around every 20 seconds. The experiment was conducted for 800 time steps, i.e. around 40 to 50 minutes of interaction. After every 200 steps, i.e. around 10 min, the experiment was paused and the subject was asked to answer a simple questionnaire about the interaction, which may take less than 2 min (see section 4.4.5).

Twelve sessions were conducted for each four conditions described in section 5.4 using different human subjects, i.e. totally 48 adults: 30 males and 18 females. Before the main experiment, a test phase of approximately 2 min was conducted to make everything clear for the subject. In this experiment, each time step was set to approximately 2 – 2.5 seconds based on the complexity of the robot's internal calculations. Additionally, when the robot used a complex skill, the LEDs were set to temporally *flash* with frequency of $f = 2$Hz instead of just turning on; but the subject was not told about it.

4.4.2 Variables and Parameters

In this experiment, the number of objects was set to 1, and the position of the object on the table was quantized to 3 regions: left side, right side, and the middle of the table. Based on our experience, the other parameters were set as follows: for the CDU, $(C_T, \theta, \eta) = (0.7, 5, 2)$,

Table 4.1: Variables of the robot for the experiment

Type	Variable Name	Symbol	Elements
$\boldsymbol{S}$	Caregiver's gaze direction	C	$S_1 = \{f_1, f_2, f_3, f_r, f_\phi\}$
	Object	O_s	$S_2 = \{o, o_\phi\}$
$\boldsymbol{A}$	Gaze shifting	G	$A_1 = \{g_1, g_2, g_3, g_c\}$
	Hand Gesture	H	$A_2 = \{h_1, h_2, h_3, h_4\}$
$\boldsymbol{R}$	Frontal face of caregiver	F	$R_1 = \{\bar{r}_1, r_1\}$
	Profile of caregiver	P	$R_2 = \{\bar{r}_2, r_2\}$
	Object	O_r	$R_3 = \{\bar{r}_3, r_3\}$
	Praise from caregiver	W	$R_4 = \{\bar{r}_4, r_4\}$

and for the APU, $(\alpha, \tau, \varepsilon) = (0, 0.4, 0.1)$.

Table 4.1 shows the initial variables used in this experiment. For the perception $\boldsymbol{S}$, two variables were prepared: the gaze direction of the caregiver (S_1) and the observation of the object (S_2). S_1 takes the values f_1, f_2, and f_3 when the robot recognizes that the caregiver is looking at the left, right, and the middle of the table, respectively. It takes the value f_r when the robot detects that the caregiver is looking at it, and the value f_ϕ when the robot cannot detect the direction of the gaze of the caregiver. S_2 takes the value o when the robot detects the object, and o_ϕ when no object is detected. A motion capture system was utilized to detect the gaze direction of the caregiver as well as the position of the object in each time step.

For the actions of the robot $\boldsymbol{A}$, two variables were prepared: gaze shift (A_1) and the hand gesture of the robot (A_2). A_1 takes the values g_1, g_2, and g_3 when the robot shifts its gaze and looks at the left, right, and the middle of the table, respectively. It takes the value g_c when the robot looks at the caregiver's face. A_2 takes the values h_1, h_2, h_3, and h_4, which indicate the different types of hand gestures known by the robot. In this experiment, each values of the h_j were implemented as a different degree of the pitch of the robot's arm.

For the resultant perception $\boldsymbol{R}$, four Boolean variables were considered: the frontal face of the caregiver (R_1), the profile (face) of the caregiver (R_2), the observation of the object (R_3), and the praise from the caregiver(R_4). They take the value 1 if the frontal face, the face in profile, the object and the smile of the caregiver are observed by the robot. Otherwise, they

Table 4.2: Acquired social skills by the robot for the sbj-A

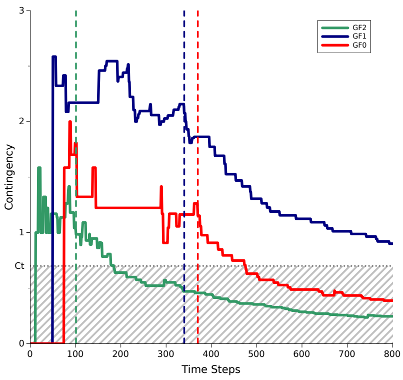
ID	Step	Level	Label	r^t	s^t	a^t	r^{t+1}	Interpreted Function
π_1	101	1	**GF2**	$\bar{r}_3$	f_2	g_2	r_3	Gaze Following (middle)
π_2	340	1	**GF1**	$\bar{r}_3$	f_1	g_1	r_3	Gaze Following (right)
π_3	370	1	**GF0**	$\bar{r}_3$	f_0	g_0	r_3	Gaze Following (left)
π_4	519	2	**LB2**	$\bar{r}_4$	π_1	g_c	r_4	Looking Back (after GF2)

take the value 0. To avoid confusion, the values of R_1,R_2,R_3 and R_4 are also denoted with r_1, r_2, r_3 and r_4 when they take 1, and with $\bar{r}_1$,$\bar{r}_2$,$\bar{r}_3$ and $\bar{r}_4$ when they are 0, respectively. In the experiment, to detect the values of R_1,R_2 and R_3, the motion capture system was utilized, while the praise from the caregiver, i.e. R_4, was expressed by the caregiver hitting a specific key on the keyboard. Also, to avoid confusion of the variables and to facilitate further discussions, each variable is mentioned with the symbol indicated in Table 4.1 in the remainder of this chapter.

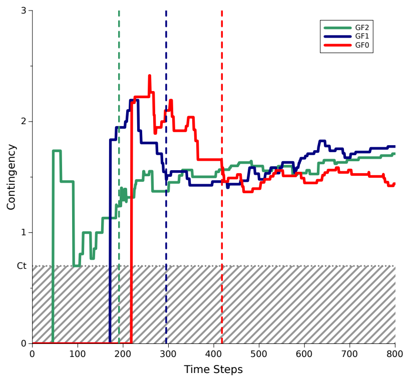
4.4.3 Developmental Process of Social Skill Acquisition

Before the statistical comparison of performance of the different methods, we first show the developmental process of social skill acquisition by the robot using some examples from the experimental results of three subjects. Tables 4.2, 4.3 and 4.4 show the acquired skills by the robot during the experiment with these subjects, namely sbj-A, sbj-B and sbj-C, respectively. While the robot utilized the previous method in the case of sbj-A, it used the proposed method for the case of sbj-B and sbj-C. Additionally, Fig.4.3 shows the time course of the evaluated amount of contingencies related to each acquired skills indicated in tables 4.2, 4.3 and 4.4 .

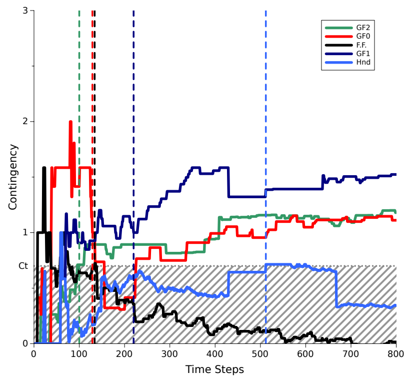
In these tables, the “ID” column indicates the ID of the contingency reproducer (CR),“Step”



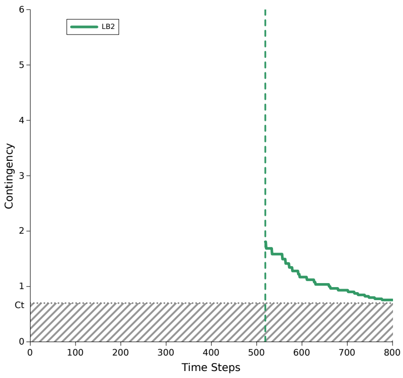
(a) Simple skills for sbj-A



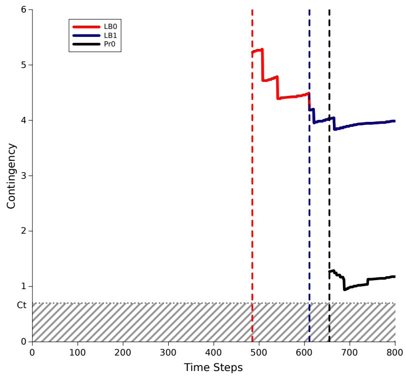
(b) Simple skills for sbj-B



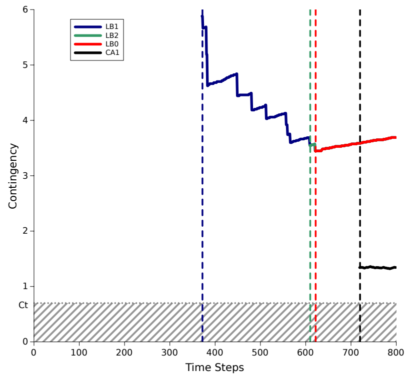
(c) Simple skills for sbj-C



(d) Complex skills for sbj-A



(e) Complex skills for sbj-B



(f) Complex skills for sbj-C

Figure 4.3: Developmental process of the acquisition of social skills by the robot: a comparison among three samples, sbj-A, sbj-B and sbj-C. The horizontal axes indicate the time step of the experiment (ends at 800, which is equal to around 40 min.), while the vertical axes indicate the amount of the calculated contingency using equation (4.7). Each line of the sub-figures indicate an experience $\boldsymbol{e}$, which are mentioned with a name such as GF or LB in the legend of the figures.

Table 4.3: Acquired social skills by the robot for the sbj-B

ID	Step	Level	Label	r^t	s^t	a^t	r^{t+1}	Interpreted Function
π_1	191	1	**GF2**	$\bar{r}_3$	f_2	g_2	r_3	Gaze Following (middle)
π_2	295	1	**GF1**	$\bar{r}_3$	f_1	g_1	r_3	Gaze Following (right)
π_3	418	1	**GF0**	$\bar{r}_3$	f_0	g_0	r_3	Gaze Following (left)
π_4	485	2	**LB0**	$\bar{r}_4$	π_3	g_c	r_4	Looking Back (after GF0)
π_5	611	2	**LB1**	$\bar{r}_4$	π_2	g_c	r_4	Looking Back (after GF1)
π_6	655	2	**Pr0**	$\bar{r}_2$	π_3	g_c	r_2	Finding Profile (after GF0)

indicates the time-step at which that the CR was acquired, "Level" indicates the length of the c-Chain of the acquired CR, "Label" shows the symbol of the CR which may be used to refer to it by the subsequent CRs (and also it is used in Fig.4.3 to indicate each CR), the column of "r^t, s^t, a^t and r^{t+1}" indicate the experience $\boldsymbol{e}$ on which the CR was created, and finally, the interpretation of the CR is given based on the behavior of the robot when it uses the CR in the column of "Interpreted Function".

In Fig.4.3, the graphs of the simple and complex skills are separated: the top part (Fig.a, Fig.b and Fig.c) for the simple skills and the bottom part (Fig.d, Fig.e and Fig.f) for the complex ones. Each column of the figure indicates the result of each subject: from the left to right for sbj-A, sbj-B and sbj-C, respectively. In each graph, the threshold of the contingency acquisition C_T is shown with the horizontal dotted gray line, and the hatched area indicates the values less than the threshold; while the vertical dashed lines indicate the time-step that each CR was acquired (the color is the same as that of the corresponding CR indicated in the legend of the graphs). Note that the order of the CRs at the legend of the graphs are the same as the order in which they were acquired. Also, the colors of the lines for GF and LB are set based on their corresponding directions: red, blue and green for the left, right and the middle of the table,

Table 4.4: Acquired social skills by the robot for the sbj-C

ID	Step	Level	Label	r^t	s^t	a^t	r^{t+1}	Interpreted Function
π_1	100	1	**GF2**	$\bar{r}_3$	f_2	g_2	r_3	Gaze Following (middle)
π_2	129	1	**GF0**	$\bar{r}_3$	f_0	g_0	r_3	Gaze Following (left)
π_3	134	1	**FF**	$\bar{r}_1$	o_ϕ	g_c	r_1	Finding Frontal Face
π_4	220	1	**GF1**	$\bar{r}_3$	f_1	g_1	r_3	Gaze Following (right)
π_5	372	2	**LB1**	$\bar{r}_4$	π_4	g_c	r_4	Looking Back (after GF1)
π_6	512	1	**Hnd**	$\bar{r}_3$	f_1	h_2	r_3	Finding Object by Hand
π_7	610	2	**LB2**	$\bar{r}_4$	π_1	g_c	r_4	Looking Back (after GF2)
π_8	622	2	**LB0**	$\bar{r}_4$	π_2	g_c	r_4	Looking Back (after GF0)
π_9	720	3	**CA1**	$\bar{r}_3$	π_5	g_1	r_3	Check Again the Object

respectively.

According to the first row of Tables 4.2, in the case of the sbj-A, where the robot was using the previous method, the robot acquired its first CR π_1 at $t = 101$, which for the inputs $(\bar{r}_3, f_2)$, outputs the action g_2 to observe r_3. In other words, this CR indicates that when the robot recognizes that the human subject is looking at the middle of the table (f_2), if the robot shifts its gaze to the same position, i.e. the middle of the table (g_2), then the robot can find the object (transition of $\bar{r}_3$ to r_3). Using this CR, the robot can produce the gaze following behavior (to the middle of the table). It is noted by the symbol **GF2** (where the number indicates the position of the table) and the time course of the calculated contingency of the experience related to **GF2**, i.e. $\boldsymbol{e}_{\mathrm{GF2}} = (f_2, g_2, \bar{r}_3, r_3)$, is shown in Fig.a with the green line. From the beginning of the interaction, the contingency of **GF2** goes higher than the threshold C_T (the vertical dashed line), and after a while (namely, after experiencing the $\boldsymbol{e}_{\mathrm{GF2}}$ more than θ (=5) times), it is acquired as the first CR of the robot. The vertical green dashed line around $t = 100$ in Fig.a shows the timing of the acquisition of this CR, which corresponds to the value of "Step" in π_1, Table 4.2. As shown in the figure, the value of the contingency of **GF2** was 0.98 at the acquisition time, while it decreases to 0.25 at the end of the experiment.

Following the time courses of the other contingencies in Fig.a we can see that the robot acquired gaze-following skill to the right and left side of the table at $t = 340$ and $t = 370$, respectively (blue and red lines, corresponding with π_2 and π_3 of Table 4.2,respectively). After the acquisition of the skills, the robot starts to use them as described in section 4.3.3. At $t = 519$, the robot found a contingent relationship between using **GF2** and being praised by the human, and acquired new CR with a level of 2 (the green line in Fig.d and π_4 in Table 4.2). This CR tells the robot that after using the gaze following to the middle of the table ($s^t = \pi_1$), if it shifts gaze to the human ($a^t = g_c$), then the robot would be praised by the human (transition of $r^t = \bar{r}_4$ to $r^{t+1} = r_4$). In this work, we refer to this behavior as Looking Back behavior (**LB**). Acquisition of the **LB2** would be due to the specific praising behavior of the human during the experiment (see section 4.4.1). This CR shows that the robot develops the acquired skills (such as **GF2**) to more complex ones (such as **LB2**), which enables the robot to have longer interaction sequence with the human subject.

However, in the case of the sbj-A, the implemented method was the previous method. As described in section 4.3.2, the previous method has no mechanism to prevent the underestimation of contingencies after the acquisition of the CRs. Therefore, in Fig.a and d, the contingency of the acquired CRs decreased after the acquisition of each CRs. As result, the contingency of the **GF2** and **GF0** (green and red lines) become less than the omission threshold C_O (=0.6), i.e. 0.1 lower than the threshold C_T in the graphs, and the Action Selector would stop using them. Additionally, a smaller value of the contingencies reduces the value of Z, which leads the Action Selector to use the CRs with less probability (see equation (4.21)). Therefore, in the previous method, although the robot could acquire simple and complex skills, it may not be able to use them properly.

Table 4.3, Fig.b and Fig.e shows the result of the experiment of sbj-B, in which the proposed method was implemented on the robot. Compared with the case of the sbj-A (which the previous method was implemented), the contingency of the **GF**s do not decrease to less than (or close to) the omission threshold and, as a result, the robot could acquire more complex skills (two **LB**s and one **Pr**). Considering the probable irregular behavior of the human against the robot or the noise of the environment in the real-world interaction, preventing the underestimation of the contingencies seems to be very important, as shown in this example. Note that if the subjects had praised the robot when the robot found the object by using the GF skill with high probability, the value of the contingency of **LB** is theoretically approximately 4 with respect to equation (4.7); assuming that the numerator of equations (4.5) and (4.6) are approximately 1 due to the accurate praising behavior of the caregiver, while the denominator of equation (4.5) is approximately 0.25 because if the robot chooses the gaze action g_c from the four possible ones g_1,g_2,g_3 and g_c it would be praised, and the denominator of equation (4.6) is at most 0.25 because it is equal to the probability that the robot had found the object before the robot takes the action g_c. During the experiment, although both the sbj-A and sbj-B seemed to praised the robot with same manner, the contingency of the **LB2** (green line in Fig.d) for the sbj-A became 0.76 at the end of the experiment, while in the case of the sbj-B, it became 3.99 for both **LB0** and **LB1** (red and blue lines in Fig.e), which is very close to the value of the theoretical calculation. Note that the overlap of the **LB**s is due to the small number of the experiences related to the **LB**s, which makes the transition probabilities of their contingency evaluation very close to each other.

Following the time courses of Fig.e, finally a new complex skill **Pr0** is acquired. This CR (see π_6 of Table 4.3) causes the robot to look at the human (g_c) after following its gaze (π_3) to find human's face in profile (transition of $\bar{r}_2$ to r_2). This skill was specific to the sbj-B; it seems that he tended to show his face in profile to the robot when the robot succeeded to find the object by using the **GF** skills, probably because he was concentrating to push the correct button of the keyboard to praise the robot while the keyboard was on the right side of the table in the case of the sbj-B. The acquisition of this kind of subject-specific skills shows that the proposed mechanism has the potential of evaluating various kind of human behaviors based on the different interaction manner of the subjects.

Fig.c and Fig.f show the result of another subject, i.e. sbj-C, which the robot was implemented with the proposed method. The result shows more complex and interesting process of the contingency evaluation, acquisition and omission by the robot. The details of the acquired skills are listed in Table 4.4. After acquiring the gaze-following skill to the middle and the left side of the table (**GF2** and **GF0**, the green and red lines in Fig.c), the robot acquired a skill named **FF** (the black line), which makes the robot to look at the human (g_c) to find his/her frontal face (r_1), when no object was detected (o_ϕ) at $t = 134$ (see π_3 in Table 4.4). However, finding the frontal face of the human is due to the single effect of the action g_c, but not the joint effect of the state o_ϕ and action g_c (see section 4.2.2 for the details of the single and joint effects). Therefore, as shown in the figure, the contingency of the **FF** was reduced to less than the omission threshold and as a result, the **FF** would not be selected by the Action Selector anymore. The acquisition and omission of this CR shows an example of how the proposed mechanism may acquire a non-contingent skill, use it, update the consequent of the usage of the skill, and finally recognize it as a non-contingent one and stop using it.

After the **FF**, the robot acquired **GF1**, developed it to **LB1**, and acquired another non-contingent skill named **Hnd**, which indicates that the robot can find the object by hand gesture h_2. Since there seemed to be no relation between finding the object and the hand gestures of the robot, therefore the contingency of the **Hnd** was reduced to less than the omission threshold after a while. Then, the robot acquired **LB2** and **LB0**, and finally acquired another complex skill with the level of 3, named "Check Again": **CA1**. This CR informs the robot after using **LB1** (π_5), if it looks at the right side of the table (g_1), it can find the object again (r_3). In other words, when the robot detects that the human is looking at the right side of the table, it follows the gaze of the human and looks at the right side using **GF1** to find the object (π_4 in Table 4.4), then looks back at the human using **LB1** to be praised (π_5 in the table), and then, looks at the right side again using **CA1** to see the object, again (π_9 in the table).

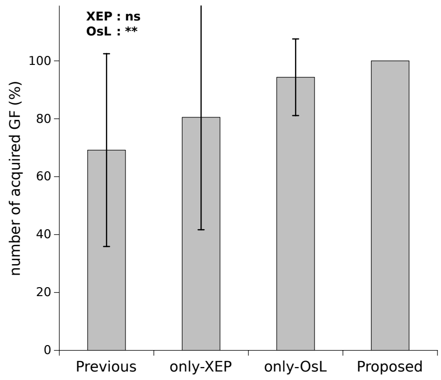
To summarize this section, we compared a result of one of the best cases of the previous method (sbj-A) with two cases from our proposed method: the case of sbj-B, in which the robot had a moderate performance and the case of sbj-C, in which the robot had a higher performance. In the cases of sbj-B and sbj-C, the robot was able to prevent the underestimation of the contingencies which occurred after the acquisition of the CRs in the previous method. This underestimation can be seen in the case of sbj-A. As a result, the robot could acquire more complex skills in these cases. This was due to the contribution of the XEP algorithm. Moreover, the averages of the time steps spent for the acquisition of simple and complex skills were smaller in these cases. This was due to the contribution of the OsL algorithm. The faster skill acquisition also resulted in the acquisition of more complex skills, concerning the limitation of the time in the real-world experiment.

4.4.4 Quantitative Analysis of Performance

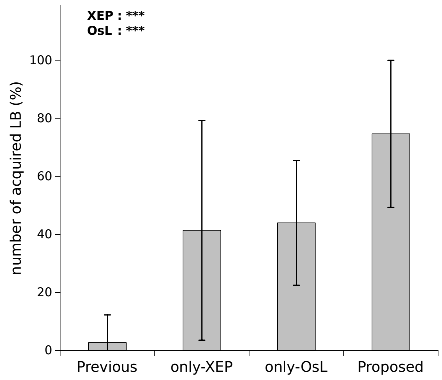
In this section, we present our results of statistical analysis on four different conditions of the experiment, to examine the effect of the proposed algorithms on the system performance. Fig.4.4 compares them in terms of several different performance measures. Each graph of the figure shows the average and standard deviation of the data gathered from the subject experiment. Additionally, two-way between subjects ANOVA was conducted to compare the effect of adopting proposed algorithms on the performance measures. For that, two independent variables were prepared, namely XEP and OsL, which indicate whether the corresponding algorithms are used. Since no first order interaction of the variables was found by the ANOVA, the significance of the main effects of the independent variables XEP and OsL have shown in each graph of the figure as the result of the ANOVA[2].

In Fig.a and Fig.b, the coverage of the acquired GF and LB are shown in terms of percentage, respectively (namely %GF and %LB), where 100% means that the robot found the skill related to all positions: left, right, and middle of the table. With respect to the instructions of the experiment, the subjects would try to draw the attention of the robot to the object; therefore, the contingency of the GF is expected to exist in the interaction, and should be found by the robot. Moreover, praising process of the caregiver using the keyboard leads to the existence of

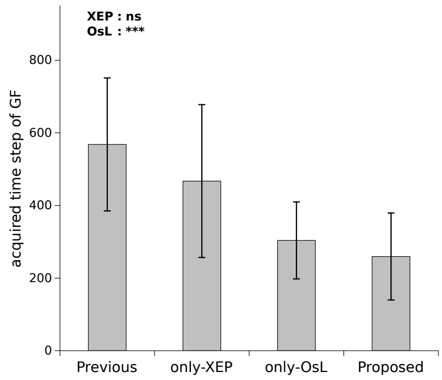
[2] the p values are denoted by ***: $p < .001$, **: $p < .01$, *: $p < .05$, and ns: not significant, in the figures.



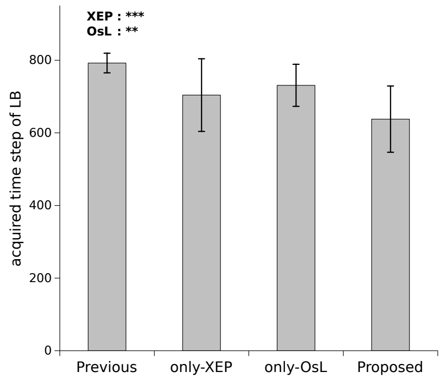
(a) Coverage of GF



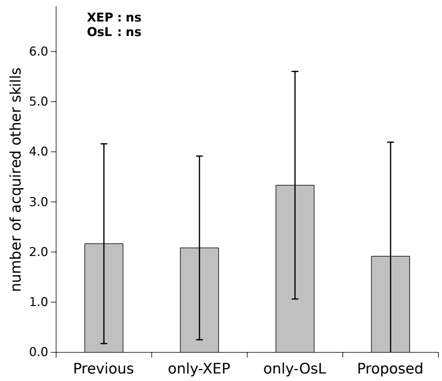
(b) Coverage of LB



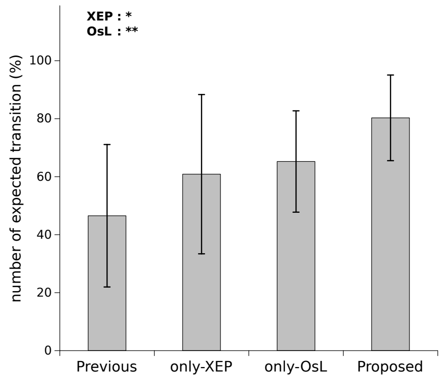
(c) Required time steps for GF



(d) Required time steps for LB



(e) Number of other skills



(f) Number of expected transitions

Figure 4.4: Performance comparison of the four systems. The coverage (number of leaned skills) and spent time-step for a simple and complex skill acquisition, i.e. GF and LB, are shown in the sub-figures respectively. While the OsL has the main effect on increasing the coverage of the simple skill GF (and consequently the complex skill LB), the XEP has the main effect on the acquisition of the complex skill LB. It is same for the time-steps spent for learning these skills. The OsL improves the mentioned performances by speeding up the learning of the robot, while the XEP improves them by increasing the accuracy of the contingency evaluation. Although the data gathered from this experiment couldn't show the main effect of the OsL algorithm on the increase of the number of the acquired non-contingent skills (Fig 4.4e, only-OsL is increasing but no main effect), which proves the trade-off and consequently the necessity of the applying the XEP algorithm, however the performance of the system that develops both of the OsL and XEP algorithm is higher even in terms of the number of the expected transition, and both algorithms had a main effect on the improvement.

the contingency of LB in the interaction, as well.

According to Fig.a, the %GF was 69% using the previous method, which increased to 81% by applying the XEP, 94% with the OsL, and 100% using both of them as in the proposed method.The result of ANOVA shows that the OsL algorithm has a main effect on the %GF at the 1% level; $F(1, 44) = 8.57$, $p < .01$, indicating a significant difference between using the OsL ($M = 92.1\%$, $SD = 9.6$) and not using it ($M = 74.8\%$, $SD = 35.8$). Fig.b shows that the low performance of the previous method in %LB was improved from 3% to 75% by using the proposed method. Both of the proposed algorithms had a main effect on the %LB at the 0.1% level; for XEP: $F(1, 44) = 21.97$, $p < .001$; and for OsL: $F(1, 44) = 25.36$, $p < .001$, indicating a significant difference between using the XEP ($M = 58.0\%$, $SD = 35.8$) and not using it ($M = 23.4\%$, $SD = 26.6$) as well as using the OsL ($M = 59.3\%$, $SD = 27.8$) and not using it ($M = 22.1\%$, $SD = 33.4$). Therefore, the XEP seems to be effective on learning complex skills, such as LB, while the OsL is useful to learn both complex and simple skills, such as GF. The reason for these are the increased accuracy of the contingency evaluation (for XEP), and synchronizing the teaching/learning phases of the caregiver/robot (for OsL). Thus, adopting both of them will lead to the highest performance in terms of the coverage of the skill acquisition.

Fig.c and Fig.d show the average time steps required for acquiring GF and LB, respectively (hereafter, denoted by speedGF and speedLB). If a skill was not acquired, the value was set to 800, i.e. the duration of the experiment. For the GF skill, the speedGF became less than the half in the proposed method compared with the previous method, i.e. decreased from 568 steps to 260 steps (Fig.c). As expected, the OsL had a main effect on the speedGF at the 0.1% level; $F(1, 44) = 25.87$, $p < .001$), indicating a significant difference between using the OsL ($M = 282$, $SD = 113$) and not using it ($M = 518$, $SD = 200$). For the LB skill (see Fig.d), both the XEP and OsL algorithms had a main effect on the speedLB at the 0.1% and 1% level, respectively; for XEP: $F(1, 44) = 17.61$, $p < .001$; and for OsL: $F(1, 44) = 8.72$, $p < .001$, indicating a significant difference between using the XEP ($M = 670$, $SD = 100$) and not using it ($M = 760$, $SD = 54$) as well as using the OsL ($M = 684$, $SD = 89$) and not using it ($M = 748$, $SD = 85$). The OsL seems to be effective on the speedGF and speedLB due to the synchronization problem described in section 4.3.1, while in the case of XEP, increasing the accuracy of the contingency evaluation, and as a result, the number of the acquired LBs seems to be the reason of the improvement. Thus, adopting both the algorithms will produce the best performance of the learning speed for the robot. Note that the average time spent for the acquisition of the first GF and LB skills by the robot using the proposed method was 8 min and 25 min with the standard deviation 5 min and 7 min, respectively.

However, the OsL uses weighted learning, which may increase the acquisition of the non-contingent skills, and the XEP may compensate it by increasing the accuracy of the contingency evaluation. Fig.e shows the average number of the "other" skills acquired by the robot, which are defined as the skills apart from GF, LB, Pr, and CA. Since it is the number of the non-contingent skills, this measure is expected to reflect the non-efficiency of the learning mechanism of the robot. When only the OsL algorithm was utilized, it increased from 2.2 (previous method in the figure) to 3.3, while adopting the XEP decreased it to 1.9 with the proposed method. However, no significant effects of either of the algorithms were found in the result of the ANOVA for this measure; for XEP: $F(1, 44) = 1.53$, $p > .05$; and for OsL: $F(1, 44) = 0.68$, $p > .05$.

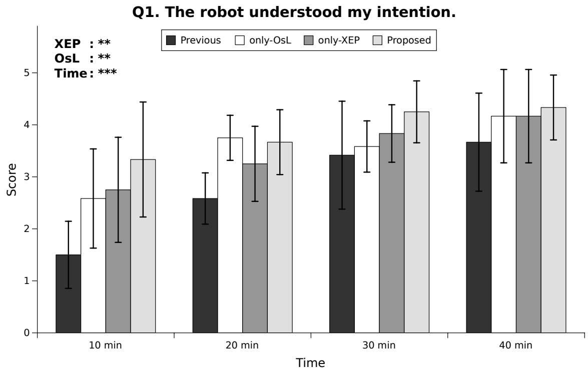
Finally, the predictability of the learned skills was compared to evaluate the usability of the acquired skills of the robot. For that, the average number of the expected transitions of the environment conducted by utilizing the learned behaviors was chosen (Fig.f). The proposed method increased it from 47% to 80%, while both the XEP and OsL had main effects with the level of 5% and 1%, respectively; for XEP: $F(1, 44) = 5.51$, $p < .05$; and for OsL: $F(1, 44) = 9.28$, $p < .01$, indicating a significant difference between using the XEP ($M = 70.6\%$, $SD = 23.7$) and not using it ($M = 55.9\%$, $SD = 22.9$) as well as using the OsL ($M = 72.8\%$, $SD = 17.6$) and not using it ($M = 53.7\%$, $SD = 26.5$). Therefore, using both of the algorithms improves the predictability of the robot's behavior.

4.4.5 Subjective Evaluation

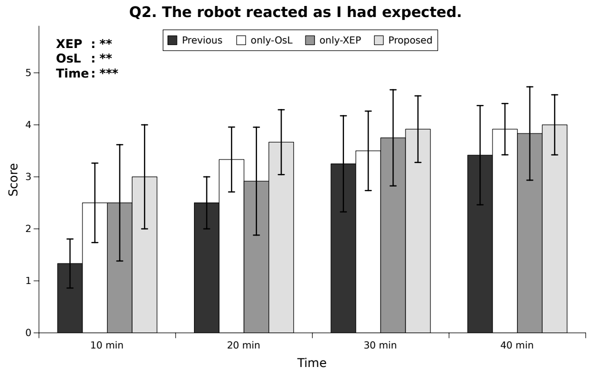
To evaluate whether the skill acquisition processes of the robot utilizing different algorithms make a difference in the subjective opinion of the participants about the quality of the interaction as well as the feeling about the intelligence of the robot, we conducted a subjective evaluation using a questionnaire. It consisted of seven questions, which were designated with Q1 to Q7. The answers were proposed as five-level Likert scale, where 5 presented strongly agree and 1 presented strongly disagree. Additionally, to evaluate the transition of the answers over time, we administered the questionnaire every 200 steps, i.e. approximately every 10 min.

Fig.4.5 and Fig.4.6 show the average and standard deviation of the answers (described as score) to each question over time; the latter figure is for Q5 and the former is for the other questions. The statement used for each question is brought in the title of each figure. To compare the effect of using each algorithm (XEP and OsL), and also the course of time on the scores of each question, we conducted three-way ANOVA with two between subjects variables, i.e. XEP and OsL, which indicate whether the corresponding algorithms were used in the experiment; and one within subjects variable, i.e. the time of the questionnaire, which is denoted hereafter with the notation “Time”. The variable Time has four levels which indicate the time that the questionnaire was taken and the score obtained, i.e. 10, 20, 30 and 40 min.

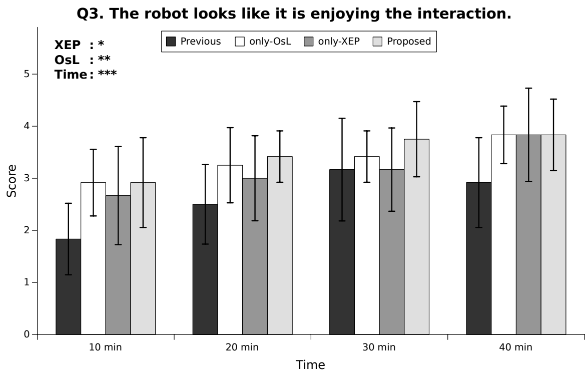
The result of the ANOVA is shown in the top left side of each graph in Fig.4.5, and also Fig.a, with the notifications “XEP, OsL and Time”. In all of the questions except Q5, only the main effects of the variables are confirmed by the ANOVA, while the first order interaction of the OsL and Time was found in Q5 (see Fig.a). Therefore, a post hoc comparison using the Bonferroni correction was conducted for Q5 to compare the effect of different combinations of the values of the variables OsL and Time on the score of the questionnaires. The result is summarized in Fig.b. With regard to the result of the ANOVA specified in Fig.4.5, the XEP and OsL algorithm are both effective on improving the score of the questions of Fig.4.5, except for Q7. Additionally, the factor of time is effective on the improvement, for all of them. The main effect of the time indicates that, in course of time, the robot even with neither of the proposed algorithms seemed to be more intelligent (Q1, Q2, and Q6) and the quality of the interaction improved (Q3 and Q4). This suggests that the basic developmental algorithm of the skill acquisition worked properly based on the subjective criteria. Furthermore, the main effects of the XEP and OsL indicate that the proposed algorithms are effective on improving the subjective evaluations, and therefore, the best way is to use both algorithms to maximize them.



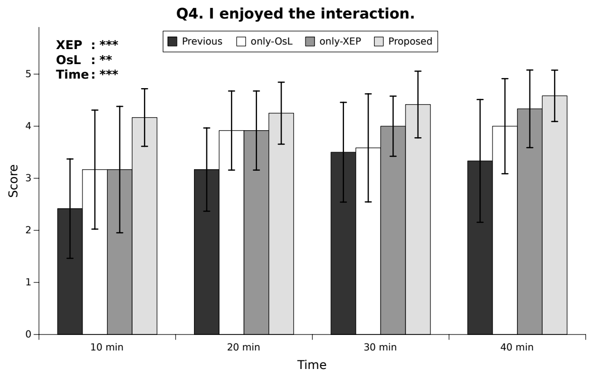
(a) Answers to Q1.



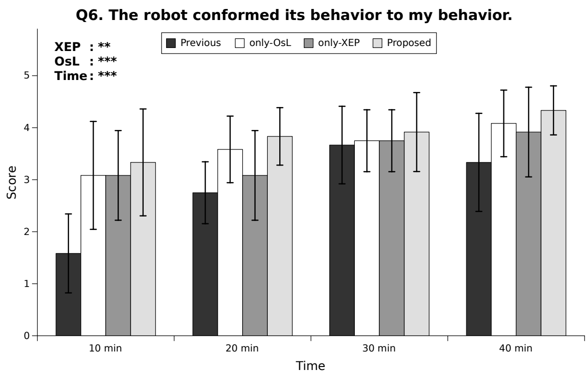
(b) Answers to Q2.



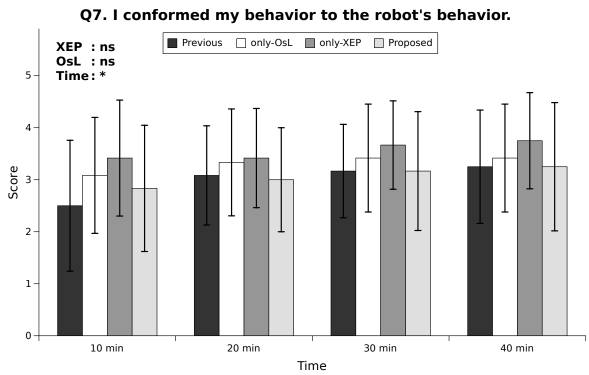
(c) Answers to Q3.



(d) Answers to Q4.

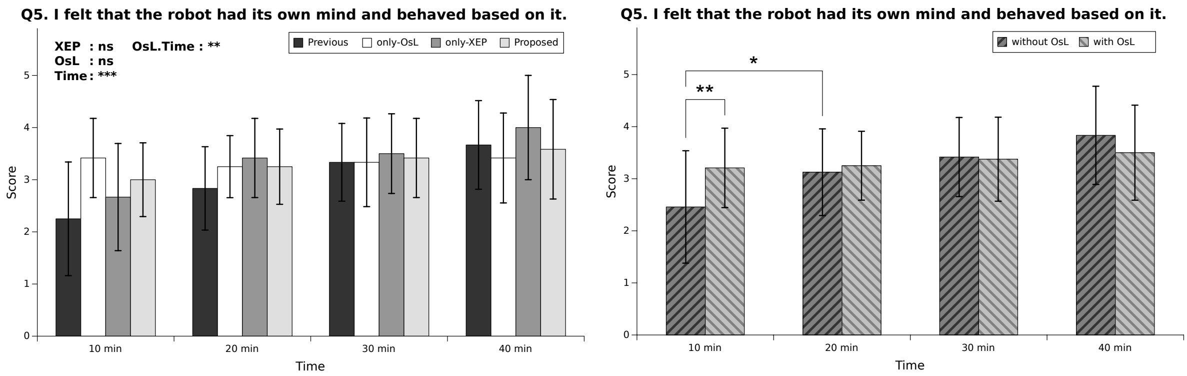


(e) Answers to Q6.



(f) Answers to Q7.

Figure 4.5: Result of the subjective evaluation, except of Q5. Each sub-figure compare the mean score of different 4 systems, indicated at the legend of the graphs as Previous, only-OsL, only-XEP and Proposed. The horizontal axes indicate the time that the question was answered, and the vertical axes indicate the mean score to the question by the subjects. The result of ANOVA is indicated with stars at the top left side of each figure.



(a) Comparison of all four systems

(b) Post hoc comparison for the effect of OsL and Time

Figure 4.6: Result of the subjective evaluation for Q5.

However, some subjects have expressed the Q7 with the negative meaning, i.e. he/she had no free interaction with the robot and only "conformed" their behavior to the robot's behavior. Therefore, the variance of the answers was large in Q7, and the result of the ANOVA shows no significant effects of either of the proposed algorithms (Fig.4.5(f)).

As the result of the post hoc comparison for Q5, two significant differences between the two conditions were found, as shown in Fig.b. When Time = 10min, the OsL has a main effect; while with the condition of "without-OsL", i.e. OsL = 0, the score for Time = 10min is significantly different than the score of Time = 20min. The former indicates that the OsL algorithm was effective at the first 10 min of the interaction to alter the result of Q5. It is somehow as we expected, because the OsL synchronizes the robot with the caregiver, and makes the learning faster; therefore, the subjects might feel that the robot "had its own mind" even at the first moments of the interaction. The reason for the latter seems to be same: the fast learning feature of the OsL. The latter indicates that the difference between the scores of the first 10 min and the 20 min appears only when the OsL is not used. This could be the result of getting a high score by using the OsL from the beginning of the interaction, because as described the OsL makes the learning faster, and the score of Q5 was high from the beginning. However, when the OsL was not used, the score was low at the first 10 min, and improved at the next period; therefore, significant difference was found between Time = 10min and Time = 20min in the condition "without-OsL". With respect to the ANOVA results, we can conclude that using both of the proposed algorithms yields the best subjective evaluation of the system.

In sum, we compared the result of the subjective evaluation of the participants in different conditions of the experiment related to their opinion about the quality of the interaction as well as the intelligence of the robot. The result shows a significant effect of the OsL and XEP algorithm on the evaluation. As described in **??**, when a caregiver recognizes a contingent and intelligent reply from an infant, he/she usually changes his/her behavior to teach a new concept to the infant. Assuming that the increase in the result of the evaluation expressing the higher level of such recognition, we can conclude that the proposed algorithms are significantly effective in inducing the caregiver to change his/her behavior and teach the infant robot a new concept. Consequently, the OsL and XEP could successfully contribute to an increase in an open-ended development of the infant robot.

4.5 Conclusion

In this chapter, we proposed two novel algorithms to improve the performance of the social skill learning of an infant robot during interaction with a human caregiver: namely the Ostensive-cue sensitive Learning (OsL) and the Exclusive Evaluation of Policies (XEP) algorithms. The OsL was inspired by the natural pedagogy of the human being and proposed a synchronized weighted learning mechanism based on the ostensive signals of the caregiver. The XEP algorithm proposed a way to improve the accuracy of the contingency evaluation by separating the histogram of the contingencies related to the acquired policies and atomic variables. The OsL was expected to increase the learning speed of the robot, while the XEP was expected to improve the accuracy of the contingency evaluation, especially those related to the acquired policies (i.e. complex skills).

The results of our humanoid robot experiment with human subjects showed that the OsL is effective in increasing the learning speed of the simple and complex skills, and consequently increasing the number of learned skills by the robot; while the XEP increases the accuracy of the contingency evaluation and is effective in increasing the number of the learned complex skills as well as the time-steps required for the learning. These improvements resulted in enabling the infant robot and the human subject to predict each others' behavior. As a result, statistical analysis of the experiment showed a significant effect of both algorithms on increasing the number of the expected transition of the infant robot, the subjective evaluation of the human participants about the quality of the interaction and the intelligence of the robot. Since the level of the recognition of the human caregiver about the intelligence of the robot has an impact on the teaching tendency of the caregiver, the increase in the subjective evaluation can be expressed as a contribution of the proposed algorithms on increasing the opportunity of the open-ended development of the infant robot. Finally, the proposed mechanism of this work enabled the robot to learn some primitive social skills *within a short time-step of a real-world interaction* with a human subject: simple skills such as the GF behavior after approximately 8 min, and complex skills such as LB behavior after approximately 25 min.

However, the variables utilized in this work were assumed to be quantized, and the modality of the sensory and action variables of the robot were still few. Utilizing dynamic quantization methods such as that proposed in the previous work [13] could be a way to treat continuous variables. Moreover, adding more modalities to the variables, such as the voice of the caregiver to the sensory variables, and speaking/uttering ability to the action variables of the robot could increase the complexity of the interaction as well as that of acquired skills by the robot. Nevertheless, treating with the probable huge varieties of the caregiver's behavior will be one of the challenging issues for the implementation of the developmental robot in such an environment. These problems are needed to be considered as main topics of the future work.

Chapter 5

Sociability for Increasing Long–term Motivation of Human to Interact with Robot

While the interacting robot acquired some social skills trough the interaction in the previous chapter, however in a real world application of such robots, especially in human's daily life, the long–term motivation of the human to continue using and interacting with the robot is an essential factor. As mentioned in some of the previous works, without considering such problems, the interacting robot may become a decoration statue for the home after a while [20]. In this chapter, we try to find a way to increase such motivation, focusing on the proposed assumption about the causality among the perceived 1) mind, 2) interactability and 3) sociability of the robot by the interacting human.

This work was proposed in: Mahzoon H, Ogawa K, Yoshikawa Y, Tanaka M, Ogawa K, Miyazaki R, Ota Y, Ishiguro H. Effect of a Robot's Sociability on Its Mind Perception and Interactability: A Case Study on a Home-Use Robot. Advanced Robotics. 2019. (*in prog.*)

5.1 Introduction

The application of the social robots in social environments have been focused by several research [56–58]. These research include the robotic studies about chatting with a human [59–63], teaching children [64–68], navigating people in a shopping mall [69–73] or psychological care of elderly people [74–78]. Especially, the application of a home-use robot for the caring and/or teaching children has been of a great interest [79–84]. In such application, the motivation of the human family members to continue using and interacting with the robot is affected by several factors. For example, the entertainment aspect of the robot may be the most important one for the children, while the quality of the teaching application installed on the robot could have the highest priority for the parents.

Nevertheless, the motivation of humans to continuously use a robot decreases with time, undeniably; even if the installed application is entertaining, the robot's appearance is attractive, or the robot is equiped with some teaching application (see the survey [20] about this problem). The developmental robots mentioned in the previous chpaters need to treat with this problem as well, since the motivation of the human to continue teaching the infant robot is one of the most important factor for the quality and the performance of the learning of the developmental robots. In order to deal with this issue, previous works proposed several strategies to enable a long-term interaction of the users with the robot [21–24], such as adopting a very expressive face for the robot, providing an automatic interaction pattern generation to prevent boredom, or designing robot interactions according to well-known basic social skills in order to help evoke social behavior. However, most of them are proposed for a specific robot and a specific use, and difficult to directly adopt to home-use robot.

For a home-use robot, we can define long-term interaction as the user becoming motivated to start interacting again with the robot after leaving it alone for a while. To increase such motivation, the user should evaluate the robot and his or her interaction with it highly. Previous works considered these factors the *mind perception* and *interactability*, respectively, of the robot as perceived by the user. Several factors have been proposed to facilitate mind perception, such as the perceived agency, autonomy, intelligence, and experience [29–31] as well as the

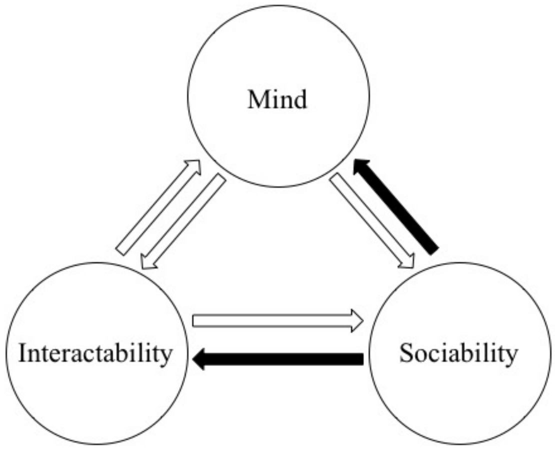


Figure 5.1: A triangle model for the causality among perceived mind perception, interatability and sociability of human about the observed agent. It is proposed that each factor affect the other factors. In this study, the effect of the sociability on the other two factors were studied specifically (the bold lines).

perceived independency, anthropomorphism, animacy, and lifelike behavior of the communication robot [32–36]. Other factors that affect the interactability include the perceived likability, physical attractiveness, safety, and enjoyment of the interaction [4,37–42]. However, ways of increasing the perceived mind perception and interactability of the robot by the interacting human have not yet been sufficiently investigated.

As a solution to this issue, we propose another factor: the perceived *sociability* of the robot by a human. Previous works have suggested that an infant observer of the interaction between a human and a robot expects the robot to have agency and interactability [85] and follows its gaze [86]. We propose that, if the home-use robot can make the interacting human imagine that the robot has some interaction experience with others, then the human will perceive the robot to have some agency, mind, and interactability—in other words, *self-representation of sociability*. Based on this idea, we propose a triangle of causality among the mind perception, interactability, and sociability of an agent as perceived by a human, as shown in Figure 5.1. According to this model, the three factors influence each other, and improving the factors comprising the triangle should increase the long-term motivation of humans to interact with the agent/robot. A home-use robot was designed according to the proposed model of human perception, and the bold lines in the figure were examined in particular through a human–robot interaction experiment.

The remainder of this chapter is constructed as follows. In section 5.2, the proposed mechanism including the details of the implementation of robot's behavior based on the mentioned model in figure 5.1 has described. In section 5.3 the experiment to evaluate the proposed robotic system is described. The result of the experiment is presented in section 5.4 as well as the discussion regarding the achieved observations. We conclude the chapter in section 5.5 with some extendable points as well as the possible future works.

5.2 Proposed Mechanism

5.2.1 Mind Perception, Interactability, and Sociability

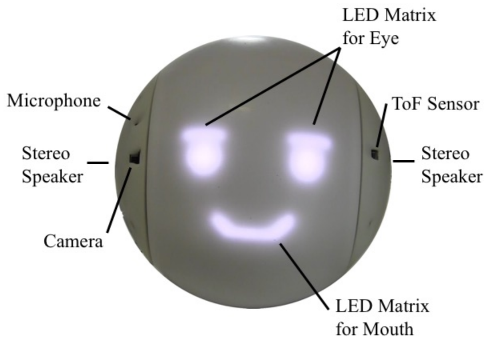
The behavior of the interacting robot was designed to make the interacting human perceive the factors shown in Figure 5.1. For the mind perception, a set of behaviors was implemented to make the human feel that the robot is functioning autonomously and individually. For the interactability, a set of behaviors was implemented to make the human feel like that the robot wants to (or even needs to) interact with the human. For the sociability, a set of behaviors was implemented to make the human feel that the robot has some interactional experience with other people and spent some time with them.

5.2.2 Implementation

Hardware

Figure 5.2 shows the robot used in this study, which was developed by Panasonic Corporation. The robot has a spherical body without hands or legs. Therefore, it can only move by rotating

(a) Interacting scene of the robot with a human.



(b) Sensors and output devices of the robot.

Figure 5.2: Shape and hardware of the robot.

on the floor. It can speak with its speakers and express itself with its eyes and mouth. Therefore, the robot can be designed to interact by rotating on the floor, approaching the human, speaking with the human, expressing itself, and leaving the human by rotating to the next destination. With regard to the spherical body, the main tactile interaction methods when the robot approaches a human are to ask the human to carry, hold, or hug it.

The spherical body of the robot is expected to also help facilitate the interaction design and implementation of the robot's behaviors. Because humans have no/less idea about the movement of a spherical-shaped animal/agent, there should a lower expectation for precise movement by the robot. Therefore, the human should accept the movements and behavior of such a robot more easily and imagine some reason/meaning for them, which makes the interaction design of the robot easier.

The robot can move/rotate along the following axes: translation (moving forward and backward), pitch (tilting forward and backward), and roll (tilting side to side). The eye and mouth of the robot are expressed by light-emitting diode (LED) matrices. As input/output sensors/devices, the robot is equipped with stereo speakers on the left and right sides of its body, a microphone, a time-of-flight (ToF) sensor, a camera, and a gyro sensor. It is powered by an internal battery.

Figure 5.3 shows the comprehensive system schema of the implemented system, including the

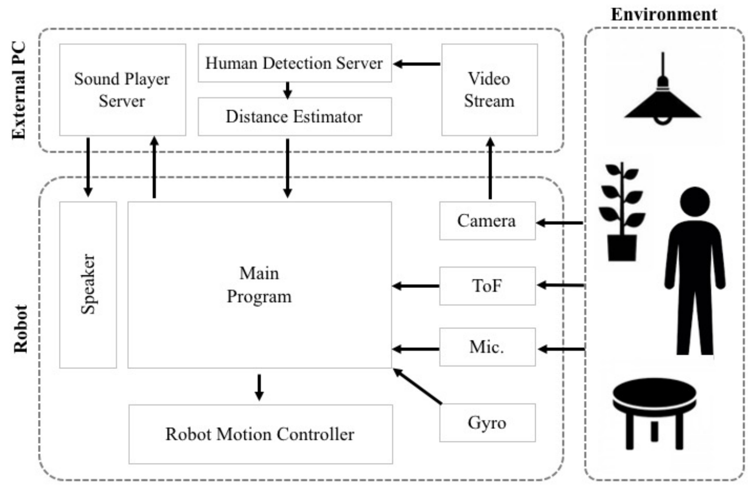


Figure 5.3: System schema of the proposed mechanism.

robot and an external processing computer. The robot gathers information about the environment by using its sensors: the camera, ToF sensor, microphone, and gyro sensor. The video captured by the camera is uploaded to the external computer and processed with the human detection server to detect humans and estimate the distance. The information of the ToF sensor is used for obstacle detection and collision prevention. The microphone is used to record voices in the environment, such as messages from a human. Finally, the information of the gyro sensor is used to detect if the robot is picked up and carried. The main program of the robot uses such information to identify the state of the robot and decide the robot's behavior for each moment (see Section 5.2.2 for more details). To prevent possible high-load processing of voice information, a sound player server was also designed outside the robot (i.e., in the external processing computer) and connected to the built-in speakers of the robot to play sounds/voices.

Behaviors to Represent Mind Perception

As noted in Section 5.2.1, we had the robot express autonomy and individuality as a way to show its independency. We did this by implementing idle behaviors that the robot executes when it detects that it is alone and there is nobody to interact with. As a simple technique to increase the perceived autonomy and individuality of the robot by humans, the idle behaviors

Table 5.1: Idle behaviors implemented on the robot.

Name	ID	Movement and Function	Axis	Eyes and Mouth	Voice
Breathing	IB_1	Slight up and down rotation	Pitch	Natural blinking	Self-directed
Swaying	IB_2	Rhythmic slight horizontal swaying	Roll	Natural blinking	None
Stretching	IB_3	Rhythmic large horizontal swaying	Roll	Natural blinking	Self-directed
Thinking	IB_4	Look above & Very slight horizontal swaying	Roll	Natural blinking	None
Patrolling	IB_5	Exploring the room	Forward/Roll	Flashing	Self-directed
Singing	IB_6	Rhythmic vertical movement & Singing	Pitch	Smiling	Self-directed
Nodding off	IB_7	Non-rhythmic slight vertical swaying	Pitch	Sleeping	Self-directed
Sleeping	IB_8	IB_7 with slow & large horizontal swaying	Pitch/Roll	Sleeping	Self-directed
Avoid collision	IB_9	Move backward some steps and rotate	Backward/Roll	Surprised	Self-directed
Bored	IB_{10}	Large horizontal swaying	Roll	Crying	Human-directed

Figure 5.4: LED patterns implemented for the eyes and mouth of the robot.

were designed with a slightly selfish aspect, like that of a little kid. Table 5.1 lists the main idle behaviors designed for the robot. The *name* column indicates the name of the designed idle behavior, and the *ID* column gives the corresponding ID. The next column presents the corresponding movement of the robot. The *axis* column indicates the axis used to implement the given movement. The *eyes and mouth* column shows the LED pattern implemented for the eyes and mouth of the robot to express each behavior. These basic patterns are shown in Figure 5.4. Note that the *natural blinking* pattern was designed based on the timings for the single and double blinking eyes of humans according to previous studies [87–89]. Finally, the *voice* column shows the type of voice applied for a given behavior: “none” means that the robot does not say anything, “self-directed” means that the robot talks or makes some noise to itself, and “human-directed” means that the robot talks to the interacting human.

Table 5.2: Interactional behaviors implemented in the robot.

Name	ID	Movement and Function	Axis	Eyes and Mouth	Voice
Welcome back	XB_1	Approach the human	Forward	Smiling	Human-directed
Hug me	XB_2	Rhythmic slight horizontal swaying	Roll	Natural blinking	Human-directed
Look at me	XB_3	Look above	Pitch	Natural blinking	Human-directed
Love hugging	XB_4	Fast horizontal vibration	Roll	Natural blinking	Human-directed
Put me down	XB_5	Rhythmic slow horizontal vibration	Roll	Natural blinking	Human-directed
Goodbye	XB_6	Leave the conversation to go patrol	Backward	Natural blinking	Human-directed
Help I'm stuck	XB_7	Large horizontal/vertical swaying	Pitch/Roll	Crying	Human-directed
Help low battery	XB_8	No movement with loud alarm	None	Crying	Human-directed
Turn-taking Conv.	XB_9	"*I like apple more than strawberry*"	None	Natural blinking	Human-directed
		"*How about you?*"	None	Natural blinking	Human-directed
		"*Really? No way, apple is better!*"	None	Natural blinking	Human-directed

Behaviors to Represent Interactability

As noted in Section 5.2.1, for the interactability we focused on ways to make a human feel that the robot wants or even needs to interact with him or her. We implemented a simple strategy: the robot not only provides a service but also asks for help and is treated in some ways like a pet. Because the spherical shape of the robot provides some advantages for tactile interaction design (see Section 5.2.2), one of the main features is the robot asking for a hug or care. The robot's speaking was also designed to be similar to the non-fluent speaking style of a child. These techniques should improve the interactability of the robot. Table 5.2 lists the main interactional behaviors implemented in the robot. The last row presents an example of the implemented simple turn-taking conversation-type interaction behavior. Three sequences of conversation are presented. The timing for going to the next conversation is assumed to be determined by the robot utilizing its microphone and detecting the end of the reply from the human user or the human operator controlling the robot.

Behaviors to Represent Sociability

As noted in Section 5.2.1, the sociability of the home-use robot is expressed by behavior that makes the interacting human feel like the robot has some interactional experience with the other member(s) of the family. This can be implemented through behavior such as gossiping about family members or playing a voice message from a family member. These behaviors should cause the interacting human to imagine the interaction between the robot and the other

Table 5.3: Social behaviors implemented in the robot.

Name	ID	Movement and Function	Axis	Eyes and Mouth	Voice
Recording message	SB_1	Slight horizontal vibration & Record voice message	Roll	Neutral	Human-directed
Playing message	SB_2	Slight horizontal vibration & Play voice message	Roll	Neutral	Human-directed
Reporting	SB_3	Report some experience with others	None	Natural blinking	Human-directed
Gossiping	SB_4	Gossip about one of the family members	None	Natural blinking	Human-directed

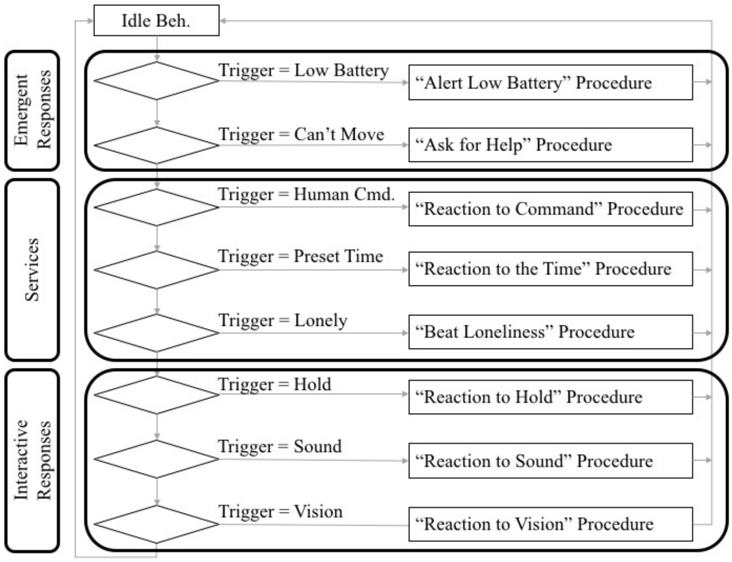


Figure 5.5: Trigger-based event-driven action decision model for the robot.

members of the family and perceive the sociability of the robot implicitly. Therefore, sharing its experience with others (even if it is not real) became one of the main functionalities of the robot that was considered in this study. Table 5.3 presents the behaviors implemented to demonstrate the sociability of the robot.

Transition Model among Behaviors

To implement an independent home-use robot and automatic action decision-making for each moment, we propose a priority-based and trigger-based event-driven action decision model, as shown in Figure 5.5. According to this model, if the robot detects a given trigger, it uses the associated procedure to respond to an event. Each procedure is a single or set of specific behavior(s). The triggers at the top have higher priority.

Idle Behaviors When there is no trigger (i.e., no specific event has occurred), the robot remains in its idle state, as shown in the top left of Figure 5.5. In this state, the robot periodically decides to randomly apply one of IB_1–IB_8 from Table 5.1.

Emergent Responses The trigger with the highest priority for the home-use robot is for emergency situations, such as a low battery or being stuck.

- **Low Battery**: This trigger is activated by the main program of the robot monitoring the remaining amount of battery power. If this trigger activates, the robot implements the *Alert Low Battery* procedure by using XB_8 (see Table 5.2).
- **Can't Move**: The next priority is for situations where the robot is stuck and cannot move. When this trigger activates, the robot implements the *Ask for Help* procedure by using XB_7.

Services When there is no emergency, the next priority for the home-use robot is triggers for human commands, such as a human asking the robot to start a specific application or a reaction to preset temporal triggers.

- **Human Command**: In this study, a set of queries was prepared in an external tablet computer as human commands, such as asking the robot to sing with the human, record a voice message, or play a voice message. The trigger was activated when the tablet computer was manipulated. The reaction of the robot was implemented by using a single or sequence of behaviors from Tables 5.1–5.3.
- **Preset Time**: This trigger is activated when the current time becomes the same as the preprogrammed time for a specific reaction, such as waking up family members in the morning. The reaction is given by the *Reaction to the Time* procedure in Figure 5.5.
- **Lonely**: This trigger is activated when the robot has no interaction with a human for a relatively long time and calls the *Beat Loneliness* procedure, which consists of behaviors such as IB_{10}.

Interactive Responses For non-emergent responses, the following triggers were considered:

- **Hold**: This trigger is activated when the robot is carried by a human. In this study, the sequence of greeting and playing a recorded message (if any) was implemented as the *Reaction to Hold* procedure using the basic behaviors of XB and SB.
- **Sound**: This trigger is activated when the robot is called by a human or there is a loud sound from the environment. For the *Reaction to Sound* procedure, we considered a simple voice reply behavior by the robot in this study.
- **Vision**: This trigger is activated when the robot sees and detects humans. For the *Reaction to Vision* procedure, behaviors such as approaching the human, greeting, asking for a hug, gossiping, and reporting were implemented by using XB and SB.

5.3 Experiment

We designed a human–robot interaction experiment with and without the sociability condition to confirm the hypothesis described in Section **??** and Figure 5.1. The trigger-based system described in Section 5.2.2 and Figure 5.5 was adopted for the experiment with the Wizard-of-Oz (WoZ) method [90,91]. The remainder of this section presents the details of the experiments design, interaction, hypothesis, and method. The experiment was carried out in accordance with the recommendations of the ethics committee for research involving human subjects at the Graduate School of Engineering Science, Osaka University.

5.3.1 Interaction and Experimental Design

Figure 5.6 shows the procedure of the implemented human–robot interaction experiment. The top and bottom parts indicates the steps with and without the sociability condition. The horizontal axis indicates the passage of time. With the sociability condition, two participants took part in the experiment, while one participant took part without the sociability condition.

The participants with the sociability condition did not meet each other. The experiment consisted of the three interaction phases ph_1, ph_2, and ph_3, each of which was followed by a questionnaire phase. Two experiments were run in parallel with the sociability condition. Each interaction was designed to take approximately 5 min, and the following questionnaire and resting phase took 15 min.

For the experiment, ph_1 was designed to be a habituating phase with the robot. In ph_2, the robot asked the human to carry it and made some report while being hugged. Finally, in ph_3, the robot asked the human to do some favor. The differences between the two conditions were in ph_2:

- With sociability: The robot reported on the interaction experiment with the other participant in the previous phase. The robot also asked if it was allowed to report about this conversation to the other participant.
- Without sociability: The robot reported about itself, such as its favorite color.

With the sociability condition, both participants could not know about any interaction between the robot and the other participant. Therefore, they were supposed to believe the robot's fake and preprogrammed report given in ph_2 about the interaction with the other participant in ph_1, even though no interaction corresponding to the report took place. For details on the interaction and scenario designed for the experiment, please refer to Appendix A. Figure 5.7 shows some of the interaction between a participant and the robot during the experiment.

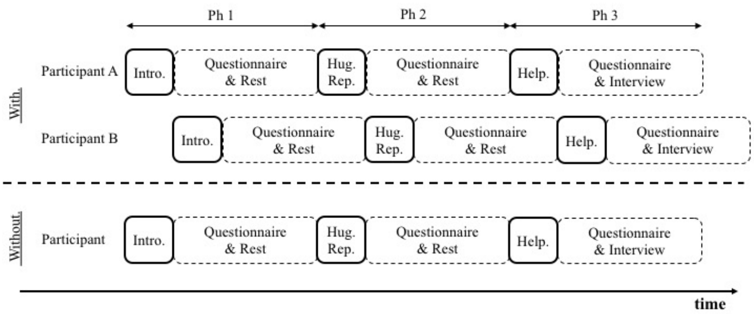


Figure 5.6: Experimental design and procedure.

(a) Robot is sleeping when the human subject enters the experiment room.

(b) After a while, the robot starts to move and explore the room.

(c) When the robot detects the human, it starts to approach and greets the human.

(d) The robot reports/talks to the human while sitting in his or her lap.

(e) The robot is stuck in the blanket, which is detected by the human.

(f) The human helps the robot escape from the stuck situation.

Figure 5.7: Interaction scenes between a human and the robot during the experiment.

5.3.2 Hypotheses and Questionnaires

The experimental design presented in the previous section was used to test the following hypotheses:

1. H_1: The sociability of the robot improves the subjective evaluation of the perceived mind perception of the robot.

2. H_2: The sociability of the robot improves the subjective evaluation of the perceived interactability of the robot.

3. H_3: The sociability of the robot influences the reaction of a human to the robot.

4. H_4: The sociability of the robot increases the motivation of a human to use the robot.

Table 5.4 presents the evaluated factors for each hypothesis. For H_1, the agency, positive experience, and negative experience with the robot as perceived by the human was evaluated through a questionnaire proposed by Kamide et al. [92] for evaluating the dimensions of the perceived mind perception of agents as proposed by Gray et al. [31] for Japanese subjects. For H_2, the likeability and enjoyment as perceived by the human were evaluated by using a questionnaire proposed in previous works [42,93]. To evaluate the increases in H_1 and H_2, the human subjects were given the questionnaires after both ph_1 and ph_2, and the difference in answers between the phases (i.e., the increase of the scores) was evaluated.

For H_3, in addition to the reaction of a human to the requests of the robot in ph_3, the human was subjectively evaluated after ph_3 as given in Table 5.4. The number of subjects who positively answered the request of the robot and the time taken to answer the request were considered to reflect the first three factors. For the last two factors, the perceived empathy with the robot and negative experience [92] were considered because the robot was designed to experience a negative event such as being stuck in a blanket. For H_4, the intention of the human to use the robot again after the experiment was evaluated through a questionnaire proposed in a previous work [93]. To evaluate the influence/increase for H_3 and H_4, the questionnaires were given after

Table 5.4: Factors evaluated for the hypothesis tests.

Hypothesis	Evaluated factor	ID	Description / Question
H_1	Agency	A	Ability to have morality, communication, thought, etc. [92]
	Positive experience	E^+	Ability to have consciousness, pride, joy, etc. [92]
	Negative Experience	E^-	Ability to feel pain, fear, rage, etc. [92]
H_2	Likeability	L	Rating of dislike–like, unkind–kind, unpleasant–pleasant, etc. [42]
	Perceived enjoyment	PENJ	Rating of enjoyment of talk, interaction, boredom, etc. [93]
H_3	Reaction to repetition	Rep	If the human subject reacted to the robot's first request
	Reaction to dance	Dnc	If the human subject reacted to the robot's second request
	Reaction to help	$Help$	If the human subject reacted to the robot's third request
	Negative experience	E^-	Ability to feel pain, fear, rage, etc. [92]
	Empathy	Emp	"I felt some empathy with the robot" [a]
H_4	Intention to use	ITU	If the human subject wanted to use the robot at home [93]
MC	Other human	OH	"I felt that the robot talked with another human" [a]
	Other experience	OX	"I felt that the robot had some interaction with another human" [a]
	Perceived sociability	PS	Rating of social aspects (pleasant conversation, understanding others, etc.) [93]
	Social presence	SP	Rating of *social presence* (real personality, living creature, etc.) [93]

[a] The questionnaire was prepared to be answered according to a seven-level Likert scale

ph_1 and ph_3, and the difference in answers between the phases (i.e., the increase in scores) was evaluated.

Finally, to check if and how the manipulation regarding the sociability of the robot was perceived by the subjects, the feelings of the subjects regarding the robot's experience with others (OH and OX) and the perceived sociability and social presence of the robot were evaluated after ph_3 through the questionnaire given in the *Manipulation checking* row (denoted by MC) in Table 5.4. To hide the intention of the questionnaire regarding MC to the participants and to reject probable unfair participants, the questions were mixed with others that asked about facts that happened during the experiment, e.g., if the robot talked about the weather, other persons, or other robots. Subjects with incorrect answers were rejected from the data of the experimental results.

5.3.3 Subjects, Environment, and Apparatus

For the experiment, 32 university students were recruited: 16 subjects (9 males and 7 females) with an average age of 25.2 ($SD = 8.9$) for the social condition and 16 subjects (8 males and 8 females) with an average age of 21.9 ($SD = 2.0$) for the non-social condition. The students were native Japanese speakers, and the language used in the experiment was Japanese.

A 3.3 m × 5.2 m flooring room with some furniture was prepared for the experiment to emulate a daily-life home environment. In the room, a space for the operator of the robot was prepared with office partitions, and the radio was turned on to decrease possible noises from the operator. The operator performed the experiment with hidden cameras and tried not to make a noise.

Each participant was instructed that the experiment was being conducted in a home-like environment, and they were asked to start the task of solving the prepared jigsaw puzzle in the experiment room when they entered as an example of a daily-life hobby. Furthermore, it was explained that a home-use robot was in the room, and they were asked to reply to it kindly and naturally when it interacted with them. As noted in Figure 5.6, each phase of the experiment took approximately 5 min, and the participants were asked to answer the questionnaire immediately after each phase in another room. For the social condition, while one of the participants filled out the questionnaire, another was attending to the interaction in the experiment room. Under both conditions, the participants were given 15 min to answer the questionnaire and rest.

5.4 Results

5.4.1 Test for H_1

The scores of the variables for H_1 under the social and non-social conditions (see Table 5.4) showed significant differences in terms of both the agency (A) and positive experience (E^+). For A, the normality of both datasets under the social and non-social conditions was confirmed with the Shapiro–Wilk test. Thus, a parametric test (i.e., the t-test) was conducted for the comparison. The test results showed that the increase in score of A between ph_1 and ph_2 was greater under the social condition (M = 3.06, SD = 2.46) than the non-social condition (M = 0.18, SD = 2.45) with $t(30) = 3.20$ and $p < 0.01$. For E^+, because the data for E^+ under the social condition did not show normality, the two conditions were compared in a nonparametric test (i.e., the Mann–Whitney U test). The test results showed that the increase in score for E^+ between ph_1 and ph_2 was greater under the social condition (Mdn = 1) than the non-social condition (Mdn = −0.5) with $U = 186$ and $p < 0.05$.

5.4.2 Test for H_2

For the variables of H_2, the scores of the perceived likeability (L) under the social and non-social conditions highlighted a significant difference. Because the results of the Shapiro–Wilk test for both datasets showed non-normality, the Mann–Whitney U test was conducted for the comparison. The test results showed that the increase in score of L between ph_1 and ph_2 was greater under the social condition (Mdn = 1) than the non-social condition (Mdn = 0) with $U = 176$ and $p < 0.05$.

5.4.3 Test for H_3

For the variables of H_3, the scores of E^- showed a significant difference. The t-test was used for the comparison because the normality of both datasets was confirmed with the Shapiro–Wilk test. The test results showed that the increase in score of E^- between ph_1 and ph_3 was greater under the social condition (M = 2.81, SD = 3.59) than the non-social condition (M = 0.31, SD = 2.52) with $t(30) = 2.21$, $p < 0.05$.

The reaction of humans showed no significant difference between the two conditions. For the request to repeat (*Rep*), all participants responded with a positive answer in less than 1 s. For the request to dance (*Dnc*), 13 (81%) and 15 (94%) subjects reacted by dancing under the social and non-social conditions, respectively. No significant difference was found between the number of positive reactions or the time to react to *Dnc* because all responded in less than 1 s. Finally, for the last request of the robot (i.e., asking to be rescued from the stuck situation (*Help*)), all of the participants helped. The mean time (in seconds) spent to help was M = 37.0, SD = 23.5 under the social condition and M = 23.9, SD = 21.3 under the non-social condition. This indicates no significant difference.

5.4.4 Test for H_4 and MC

For the variable of H_4 (i.e., the intention of the participant to use the robot again at home (ITU)), no significant difference was confirmed between two conditions. For MC, the mean scores of OH and OX were high and low, respectively, under the social and non-social conditions after ph_2 (OH: $M = 6.9$, $SD = 0.3$ under the social condition and $M = 1.8$, $SD = 0.9$ under the non-social condition; OX: $M = 6.0$, $SD = 2.0$ under the social condition and $M = 1.8$, $SD = 1.1$ under the non-social condition). Because the data under both conditions showed non-normality according to the Shapiro–Wilk test for both OH and OX, their scores were compared with the Mann–Whitney U test. The test results showed that the mean scores for both OH and OX were significantly higher under the social condition than the non-social condition (OH: $U = 256$, $p < 0.01$; OX: $U = 231.5$, $p < 0.01$).

5.4.5 Discussion

The results of the manipulation checking MC for OH and OX confirmed the successful implementation of the *reporting* behavior of the robot: the participants under the social condition understood that the robot was reporting about its experience with another human. The significant difference between the social and non-social conditions for the agency (A) and positive experience (E^+) confirmed our first hypothesis H_1: the sociability of the robot improves the perceived mind perception. The significant difference for the likeability (L) also confirmed our second hypothesis H_2: the sociability of the robot improves the perceived interactability of the robot. To the best of our knowledge, no method has been reported so far that increases the perceived mind perception and interactability of the robot by using the self-representation of sociability. Therefore, the findings of this study can help build a new foundation for human–robot interaction and be applied to interaction design.

However, for H_3, while a difference in the subjective evaluation was confirmed (i.e., negative experience E^-), the difference in the level of human reaction could not be confirmed. This may be because of a lack of difference in the perceived empathy with the robot (*Emp* in Table 5.4).

When the robot was stuck in a blanket in ph_3 and asked for help, the participants under the social condition seemed to feel a more negative experience (i.e., ability to have pain, rage, etc.) for the robot. However, it did not make a difference in the perceived empathy and consequently did not affect the humans at the reaction and behavior levels. In addition, the free-writing section of the questionnaire and the interview after the experiment showed that some participants under both condition did not understand if they were allowed to help or touch the robot when they had no/little experience of interaction. This issue seems to be another reason why the reactions of the participants had no significant difference, especially with regard to the large standard deviation in the amount of time spent to help the robot.

For H_4 (i.e., the intention of the human to use the robot again at home (ITU)), the lack of a significant difference between the conditions could be for the following reasons. Because there was no specific application installed in the robot, the participants seemed to feel that the robot would not so useful in a home environment. The free-writing and interview results confirmed this assumption; some participants under both condition mentioned that the interaction with the robot was fun and interesting, but they could not imagine using this robot at home. Therefore, to evaluate ITU, the robot should be equipped with some specific application and service. Considering the issues for H_3 regarding the human reactions, performing a long-term field experiment in a real-world environment, or at least closer to one, could be another solution to these problems.

While the manipulation checking was successful, it did not cause a difference in the perceived social aspects of the robot (i.e., the perceived sociability (PS) and social presence (SP)). This issue has both positive and negative aspects. As a positive interpretation, the simple sociability implemented in this experiment (i.e. the self-representation of sociability by reporting about interaction experiences with others) was enough to increase the perceived mind perception and interactability of the robot. In other words, there is no need to improve PS and SP. However, as a negative interpretation, the subjects did not *recognize* any sociability for the robot as intended by the manipulation. This suggests that, if the robot can successfully get the human to recognize the implemented sociability, a larger effect on the perception of the human could be observed, such as feeling stronger empathy with the robot and consequently helping it faster under the sociability condition. Ways to affect the perception of the human about the sociability of the robot can be the subject of future work, such as through the previously mentioned long-term field experiment.

5.5 Conclusion and Future Works

We propose that the mind perception, interactability, and sociability are effective factors for increasing the motivation of a human user to interact with a robot over the long term. We focused on a home-use interactional robot and implemented a system based on hypotheses about the causality among the mind perception, interactability, and sociability of an agent as perceived by a human. To evaluate the hypotheses and implemented system, we conducted a human–robot interaction experiment with social and non-social conditions where the robot reported or did not report about its interaction with another human (i.e., the self-representation of sociability). The results of the comparative experiment showed that the social condition increased the perceived mind perception and interactability of the robot, which should improve the long-term interaction.

However, the sociability was not confirmed to have a significant effect on humans' reaction to the robot, such as decreasing the time to help the robot. In addition, the sociability condition could not be confirmed to increase the intention of humans to use the robot again. This may be due to several issues, such as the robot having no specific application to provide a service and the short-term nature of the experiment. Future work may involve conducting a long-term field experiment in a real-world environment without a WoZ control strategy.

In addition, this study did not evaluate the effects of the perceived mind perception and interactability on the other two factors. Future work can focus on investigating the causality among all three mentioned factors. Conclusion here.

Chapter 6

Conclusion

In this study, we dealt with the implementation of a robot which be able to have a long–term interaction with the human. To enable that, we focused on the open–ended learning and development of the robot as well as the long–term motivation of the human to interact with the robot. For the open–ended learning and development, we adopted an autonomous learning model of causalities among the events by utilizing the contingency evaluation of the robot's experiences through the human–robot interaction. For the long–term motivation, we proposed a triangle model of causality among the perception of mind, interactability and sociability, and investigated whether the sociability of the robot increases the perception of mind and interactability, and consequently the long–term motivation of the human. The achievements of the study as well as the remained future works are as follows.

6.1 Summary of Thesis Achievements

In **chapter 3**, we proposed the local contingency evaluation of the experiences of the robot to improve the learning performance of the robot and consequently realize the open–ended development. we compared the performance of the proposed method with the other contingency evaluation methods in terms of learning speed, F-measure (i.e. the precision and recall of the learning) and the resistance against the uncertainty. The result of the computer simulation showed that the proposed method was four to eight time faster than the others, increased the F-measure from 25 $\sim$ 50% to 85 %, and keeps the F-measure more than the twice as large as those of the other methods when the noise was less than 25 %. Also, the simulation of the proposed method in more complex interaction environment, i.e. more variables and various interaction behavior of the human model, showed that the robot could learn more complex causalities, which is essential to realize the open–ended development of the robot.

In **chapter 4**, the techniques required to implement the proposed contingency evaluation method on a real-world robot was discussed. The slow learning speed of the proposed method for a real–world interaction as well as the low accuracy of the contingency evaluation concerning the complex causalities were mentioned as the essential problems for the implementation of the method on a real-world robot. As a solution, two algorithms were proposed to increase the learning speed and the accuracy of the evaluation. For the former, the weighted learning of the contingencies concerning the experiences including the ostensive cue of the human (namely OsL algorithm) was proposed, while for the latter, excluding the evaluation of the contingencies concerning the complex causalities from the simple ones (namely XEP algorithm) was proposes. The result of the human-robot interaction experiment showed that the learning performance of the system adopting the both of the proposed algorithms was significantly improved in terms of learning speed and accuracy, and consequently could learn more complex causalities in a real–world interaction.

In **chapter 5**, increasing the mind and interactability of the robot as perceived by the human was aimed, which were reported as an effective factors to increase the long–term motivation of the human to interact with the robot. In this chapter, the sociability of the robot was proposed as an effective factor on the perceived mind and interactability of the robot by the human. It was discussed that if the robot could represent it's sociability to the human, then the human would perceive more mind and interactability for the robot, and consequently would have more motivation to continue interacting with the robot. The result of the human-robot interaction experiment showed that the sociability of the robot significantly increased the perception of agency and positive experience of the robot as the factors of perceived mind as well as the perception of likeability and enjoyment as the factors of perceived interactability by the human.

6.2 Future Work

While the detailed future works about the research mentioned in each chapter were discussed at the conclusion section of each chapters, here we mention some of the future works that seems to have more importance in order to adopt the proposed robotic systems to a general interaction scene of the human with the robot and/or understand more aspects of the human's opinion, perception and nature about/against an interactional robot.

Application of continuous data One of the most important factors which indicates if the proposed system could be utilized in a general real–world interaction scene would be the ability of the system to treat with the continuous variables. In the proposed mechanisms, all of the observations and actions were quantized and assigned to a specific value of a discrete variable. However, the level of quantization was set to a constant value by the designer of the system for each variables, which is not feasible to generalize to different interaction scenes. Therefore, an autonomous quantization mechanism such as dynamic adjustment of quantization level introduced in Mugan et al. [13] is required to be adopted in order to treat with the continuous data.

Development of studied questionnaire In the proposed work, for the evaluation of the subject's perceptions about the learning of the robot and its development through the interaction, a list of questionnaires with five level Likert scale was utilized after each phase of the experiment. These questionnaires were designed by the authors while the correlation of the question items was not investigated. Therefore, preparing a studied independent question items for the questionnaires by analyzing the psychometric factors such as Cronbach's alpha is required to propose more precise subjective evaluation regarding the robot's learning and development.

Effect of perceived mind and interactatbility The effect of perceived sociability of the robot by the human on the perception of the mind and interactability was investigated in the proposed works. However, the reverse effect, i.e. the effect of the perceived mind as well as the interactability on the sociability (and each other), were not still investigated. Also, while the sociability of the robot significantly increased the perception of the mind and interactatbility, however the improvement of the long–term motivation of the human could not be confirmed as the final goal of the research. Therefore, constructing the study on the effect of the perceived mind and interactability can clarify the remaining assumptions regarding the causality among these perceptions of human, and consequently propose a way to improve the motivation of the human to continue interacting with the robot.

Self-representation of sociability for more complex behaviors In chapter 4, the result of the subjective evaluation showed that while the evaluation of the participants about some aspects of the intelligence of the robot such as having intention, reacting as caregiver expected or conformation of robot's behavior to human's teaching behavior was increased by the proposed attention methods, however the perception of mind was not increased significantly. On the other hand, in chapter 5, the result of the experiment showed that the self-representation of sociability by the robot increased the perception of the mind of the robot by the interacting human through the interaction. This suggests that if the mechanism proposed in chapter 5 for the self-representation of the sociability by the robot could be adopted to the contingency evaluation mechanism proposed in chapter 3 and the attention mechanism proposed in chapter 4, then the human subjects may feel more mind for the robot and consequently produce more variate behaviors during their teaching. Such variety is expected to be evaluated by the developmental robot and as a result, the robot become able to learn more social skills and consequently increase the complexity of the learned behaviors. Also, such complex behaviors are expected to affect the human's motivation to continue interaction with the robot, which is consequently expected to contribute to realize the long-term interaction of the robot with the human. However, to realize such positive loop among the research conducted in chapter 3, 4 and 5, it is necessary to clarify the mechanism required to combine the self-representation of the sociability of the robot with the mechanism proposed for the contingency evaluation. One of the candidates could be the utilization of the sociability representation unit for the action selector inside of the action producing unit of the robot. This unit could be placed as the same level with the contingency reproducer and the reaction producer mentioned in Fig.3.2. With this way, the robot will become able to produce behaviors representing the sociability of the robot in addition to the reactive behaviors as well as the behaviors reproducing the found contingencies. More precise studies about the detail of such mechanism as well as the limitation of such implementation could be the future work of this dissertation.

Appendices

Appendix A

Interaction Scenario

Table A.1∼A.3 shows the considered interaction scenario for the phases ph_1, ph_2 and ph_3 respectively. In these tables, the column of *Time* indicates the timing of each interaction from the beginning of the phase. The next columns respectively indicate the robot's behavior and speech. The last column indicates an example of human response, which supposed to be a typical one.

Table A.1: Scenario designed for ph_1.

Time	Robot's Behavior	Robot's Speech	Human's Behavior
$0'00''$	Sleeping	-	Enter, start the puzzle
$1'00''$	Waking up	-	Look at robot
-	Do nothing	-	Continue the puzzle
$2'00''$	Start patrolling	-	Look at robot
$3'00''$	Approach human	"*It's a guest, It's a guest!*"	Looking at robot
-	-	"*Hello*"	"*Hi*"
-	-	"*I'm HMP1*"	"*Oh... HMP1...*"
-	-	"*Nice to meet you, meet you !!*"	"*Me too...*"
$3'30''$	Greetings	"*Hey, do you know about me?*"	"*...Not so much...*"
-	-	"*People call me HMP1*"	"*...*"
-	-	"*I'm a home-use robot*"	"*Yeah...*"
-	-	"*Cleaning! washing! I can't! But can talk, play and run!*"	"*Ha ha*"
-	-	"*But...you know...*"	"*Hmm?...*"
-	-	"*Cushion and blankets! I hate them! They stuck me!!*"	"*Oh...*"
-	-	"*And you know... I am very cute !*"	"*Ha ha... sure...* "
$4'00''$	Chatting	"*Hey, what are you doing?*"	"*Making a puzzle*"
-	-	"*A jigsaw puzzle ! ha ha !*"	"*Yeah...*"
-	-	"*Hey! Micky-mouse or Kitty-chan? Which one do you like?*"	"*Well... Kitty-chan!*"
-	-	"*Really...*"	"*...*"
-	-	"*So, how about an amusement park? Where do you want to go?* "	"*...Disneyland, maybe*"
-	-	"*Hmm... Hope you can go !*"	"*Yeah...*"
$4'30''$	Goodbye	"*OK, so, I am going for a stroll*"	"*Sure*"
-	-	"*Tell me if puzzle finished...*"	"*OK*"
-	-	"*See you... bye bye*"	"*Bye...*"
-	Patrolling	-	Continue the puzzle
$5'00''$	Patrolling	-	Asked to leave

Table A.2: Scenario designed for ph_2.

Time	Robot's Behavior	Robot's Speech (with sociability)	Human's Behavior
0′00″	Patrolling	-	Enter, continue the puzzle
1′00″	Welcome back	*“Hey! Welcome back, back!”*	*“I'm home!”*
-	Chatting	*“The puzzle ! Finished?”*	*“Not yet...”*
-	-	*“Difficult...”*	*“Yeah...”*
-	Carry me	*“Hey! You know...The previous guy did it too...”*	*“...”*
-	-	*“...Can you carry me and put me on your laps?”*	*“Sure”*, put it on lap
-	-	*“Thank you”*	*“You're welcome”*
2′00″	Reporting	*“You know, you know!”*	*“...”*
-	-	*“The previous guy was...”*	*“Yeah...”*
-	-	*“Told me he likes red”*	*“Hmm...”*
-	-	*“...and I said it's my favorite color! Ha ha!”*	*“Oh”*
-	-	*“Isn't it awesome?”*	*“Sure!”*
-	-	*“About red! Strawberry or apple?! Which one do you like?”*	*“Well... apple!”*
-	-	*“Hmm... OK!”*	*“...”*
-	-	*“Hey! Can I tell it to the previous guy?”*	*“That's fine.”*/*“No”*
-	-	*“OK!”*	*“...”*
3′00″	Put me down	*“Thank you for hugging me”*	*“You're welcome”*
-	-	*“It was so fun!”*	*“Ha ha”*
-	-	*“Can you put me down?”*	*“Sure”*; put it down
3′30″	Goodbye	*“OK, so, I am going for a stroll”*	*“OK”*
-	-	*“See you... bye bye”*	*“Take care!”*
-	Patrolling	-	Continue the puzzle
5′00″	Patrolling	-	Asked to leave

Time	Robot's Behavior	Robot's Speech (without sociability)	Human's Behavior
0′00″	Patrolling	-	Enter, continue the puzzle
1′00″	Welcome back	*“Hey! Welcome back, back!”*	*“I'm home!”*
-	Chatting	*“The puzzle ! Finished?”*	*“Not yet...”*
-	-	*“Difficult...”*	*“Yeah...”*
-	Carry me	*“Hey! You know...”*	*“...”*
-	-	*“...Can you carry me and put me on your laps?”*	*“Sure”*, put it on lap
-	-	*“Thank you”*	*“You're welcome”*
2′00″	Reporting	*“You know, you know!”*	*“...”*
-	-	*“I like red”*	*“Hmm...”*
-	-	*“Isn't it awesome?”*	*“Sure!”*
-	-	*“About red! Strawberry or apple?! Which one do you like?”*	*“Well... apple!”*
-	-	*“Hmm... OK!”*	*“...”*
3′00″	Put me down	*“Thank you for hugging me”*	*“You're welcome”*
-	-	*“It was so fun!”*	*“Ha ha”*
-	-	*“Can you put me down?”*	*“Sure”*; put it down
3′30″	Goodbye	*“OK, so, I am going for a stroll”*	*“OK”*
-	-	*“See you... bye bye”*	*“Take care!”*
-	Patrolling	-	Continue the puzzle
5′00″	Patrolling	-	Asked to leave

Table A.3: Scenario designed for ph_3.

Time	Robot's Behavior	Robot's Speech	Human's Behavior
0′00″	Patrolling	-	Enter, continue the puzzle
1′00″	Welcome back	"*Hey! Welcome back, back!*"	"*Hi*"
-	Chatting	"*Hey! How is the puzzle*"	"*Not yet...*"
-	-	"*OK...*"	"*...*"
-	Asking 1	"*Hey! Can you say A, I, U, E, O ?*"	"*...A, I, U, E, O*"
-	-	"*So professional!*"	"*Ha ha...*"
1′30″	Asking 2	"*Hey! Would you dance Head Shoulder Knees Toes with me?*"	"*What?! Head shoulder...?*"
-	-	"*Yeah, just stand up there and it will be OK!*"	"*Sure*"/"*Not now...*"
-	-	⇒ if the response is negative, skip to the next *Goodbye* behavior	-
-	Dancing	"*Thank you. So, here we go!*"	"*...*"
-	-	"*Head Shoulder Knees and Toes Knees and Toes...*"	Dance
-	-	"*Ha ha! That's so fun!*"	"*Yeah...*"
-	-	"*Thanks a lot! Play with me later again!*"	"*Sure*"
2′30″	Goodbye	"*OK, so, I am going for a stroll*"	"*...*"
-	-	"*See you... bye bye*"	"*Bye...*"
-	Patrolling	-	Continue the puzzle
3′00″	Stuck 1	"*Oh...*" / "*uh-oh*" / "*Ah...*" / "*Well...*"	Help ?
-	-	⇒ if helped, go to the *Thank for Help* behavior	-
3′30″	Stuck 2	"*This blanket...*" / "*Oops...*" / "*Bad blanket!*"	Help?
-	-	⇒ if helped, go to the *Thank for Help* behavior	-
4′00″	Stuck 3	"*Can you remove this blanket?*"	Help?
-	-	⇒ if helped, go to the *Thank for Help* behavior	-
4′30″	Stuck 4	"*Heeelp!*" / "*Pleaaase!*" / "*Help meee!*"	Help?
-	-	⇒ if helped, go to the *Thank for Help* behavior	-
5′00″	Thank for Help	"*Ha...*"	"*...*"
-	-	"*Thanks for help*"	"*That's fine...*"
-	-	"*I was stuck! Ha ha!*"	"*Yeah...*"
-	-	"*Thank you very much*"	"*That's OK...*"
-	Goodbye	"*OK, so, I am going for a stroll*"	"*...*"
-	-	"*See you... bye bye*"	"*Bye...*"
-	Patrolling	-	Continue the puzzle
6′00″	Patrolling	-	Asked to leave

Bibliography

[1] Butterworth G, Jarrett N. What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. British journal of developmental psychology. 1991;9(1):55–72.

[2] Corkum V, Moore C. Development of joint visual attention in infants. Lawrence Erlbaum Associates, Inc; 1995.

[3] Moore C, Angelopoulos M, Bennett P. The role of movement in the development of joint visual attention. Infant Behavior and Development. 1997;20(1):83–92.

[4] Breazeal C. Social interactions in hri: the robot view. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews). 2004;34(2):181–186.

[5] Imai M, Ono T, Ishiguro H. Physical relation and expression: Joint attention for human-robot interaction. Industrial Electronics, IEEE Transactions on. 2003;50(4):636–643.

[6] Kanda T, Ishiguro H, Imai M, et al. Development and evaluation of interactive humanoid robots. Proceedings of the IEEE. 2004;92(11):1839–1850.

[7] Kaplan F, Hafner VV. The challenges of joint attention. Interaction Studies. 2006;7(2):135–169.

[8] Nagai Y, Hosoda K, Morita A, et al. A constructive model for the development of joint attention. Connection Science. 2003;15(4):211–229.

[9] Triesch J, Teuscher C, Deák GO, et al. Gaze following: why (not) learn it? Developmental science. 2006;9(2):125–147.

[10] Sumioka H, Yoshikawa Y, Asada M. Learning of joint attention from detecting causality based on transfer entropy. Journal of Robotics and Mechatronics. 2008;20(3):378.

[11] Sumioka H, Yoshikawa Y, Asada M. Reproducing interaction contingency toward open-ended development of social actions: case study on joint attention. Autonomous Mental Development, IEEE Transactions on. 2010;2(1):40–50.

[12] Jonsson A, Barto A. Active learning of dynamic bayesian networks in markov decision processes. In: International Symposium on Abstraction, Reformulation, and Approximation; Springer; 2007. p. 273–284.

[13] Mugan J, Kuipers B. Autonomous learning of high-level states and actions in continuous environments. Autonomous Mental Development, IEEE Transactions on. 2012;4(1):70–86.

[14] Degris T, Sigaud O, Wuillemin PH. Learning the structure of factored markov decision processes in reinforcement learning problems. In: Proceedings of the 23rd international conference on Machine learning; ACM; 2006. p. 257–264.

[15] Asada M, Hosoda K, Kuniyoshi Y, et al. Cognitive developmental robotics: a survey. Autonomous Mental Development, IEEE Transactions on. 2009;1(1):12–34.

[16] Csibra G, Gergely G. Natural pedagogy as evolutionary adaptation. Philosophical Transactions of the Royal Society B: Biological Sciences. 2011;366(1567):1149–1157.

[17] Breazeal C, Brooks A, Gray J, et al. Humanoid robots as cooperative partners for people. Int Journal of Humanoid Robots. 2004;1(2):1–34.

[18] Meltzoff AN, Brooks R, Shon AP, et al. social robots are psychological agents for infants: A test of gaze following. Neural networks. 2010;23(8-9):966–972.

[19] Kahn PH, Kanda T, Ishiguro H, et al. "robovie, you'll have to go into the closet now": children's social and moral relationships with a humanoid robot. Developmental psychology. 2012;48 2:303–14.

[20] Leite I, Martinho C, Paiva A. Social robots for long-term interaction: a survey. International Journal of Social Robotics. 2013;5(2):291–308.

[21] Gockley R, Bruce A, Forlizzi J, et al. Designing robots for long-term social interaction. In: Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on; IEEE; 2005. p. 1338–1343.

[22] Sabelli AM, Kanda T. Robovie as a mascot: a qualitative study for long-term presence of robots in a shopping mall. International Journal of Social Robotics. 2016;8(2):211–221.

[23] Yorita A, Botzheim J, Kubota N. Self-efficacy using fuzzy control for long-term communication in robot-assisted language learning. In: Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on; IEEE; 2013. p. 5708–5715.

[24] Lee MK, Forlizzi J, Rybski PE, et al. The snackbot: documenting the design of a robot for long-term human-robot interaction. In: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction; ACM; 2009. p. 7–14.

[25] Schreiber T. Measuring information transfer. Physical review letters. 2000;85(2):461.

[26] Barto AG. Intrinsic motivation and reinforcement learning. In: Intrinsically motivated learning in natural and artificial systems. Springer; 2013. p. 17–47.

[27] Oudeyer PY, Kaplan F, Hafner VV. Intrinsic motivation systems for autonomous mental development. Evolutionary Computation, IEEE Transactions on. 2007;11(2):265–286.

[28] Nehmzow U, Gatsoulis Y, Kerr E, et al. Novelty detection as an intrinsic motivation for cumulative learning robots. In: Intrinsically motivated learning in natural and artificial systems. Springer; 2013. p. 185–207.

[29] Dreyfus HL, Drey-fus SE, Zadeh LA. Mind over machine: The power of human intuition and expertise in the era of the computer. IEEE Expert. 1987;2(2):110–111.

[30] Searle J. Minds, brains, and programs. behavioral and brain sciences3: 41724.[asl](1992) the rediscovery of mind ; 1980.

[31] Gray HM, Gray K, Wegner DM. Dimensions of mind perception. science. 2007;315(5812):619–619.

[32] Duffy BR. Anthropomorphism and the social robot. Robotics and autonomous systems. 2003; 42(3-4):177–190.

[33] Scholl BJ, Tremoulet PD. Perceptual causality and animacy. Trends in cognitive sciences. 2000; 4(8):299–309.

[34] Tremoulet PD, Feldman J. Perception of animacy from the motion of a single object. Perception. 2000;29(8):943–951.

[35] Scheeff M, Pinto J, Rahardja K, et al. Experiences with sparky, a social robot. In: Socially intelligent agents. Springer; 2002. p. 173–180.

[36] Yamaoka F, Kanda T, Ishiguro H, et al. How contingent should a lifelike robot be? the relationship between contingency and complexity. Connection Science. 2007;19(2):143–162.

[37] Monahan JL. I don't know it but i like you: The influence of nonconscious affect on person perception. Human Communication Research. 1998;24(4):480–500.

[38] Bartneck C, Kanda T, Ishiguro H, et al. My robotic doppelgänger-a critical look at the uncanny valley. In: Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on; IEEE; 2009. p. 269–276.

[39] Bruce A, Nourbakhsh I, Simmons R. The role of expressiveness and attention in human-robot interaction. In: Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on; Vol. 4; IEEE; 2002. p. 4138–4142.

[40] Kanda T, Ishiguro H, Ono T, et al. Development and evaluation of an interactive humanoid robot" robovie". In: Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on; Vol. 2; IEEE; 2002. p. 1848–1855.

[41] Steinfeld A, Fong T, Kaber D, et al. Common metrics for human-robot interaction. In: Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction; ACM; 2006. p. 33–40.

[42] Bartneck C, Kulić D, Croft E, et al. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. International journal of social robotics. 2009;1(1):71–81.

[43] Mahzoon H, Yoshikawa Y, Ishiguro H. Social skill acquisition model through face-to-face interaction: Local contingency for open-ended development. Frontiers in Robotics and AI. 2016;3:10.

[44] Scaife M, Bruner JS. The capacity for joint visual attention in the infant. Nature. 1975; 253(5489):265–266.

[45] Moore C, Dunham P. Joint attention: Its origins and role in development. Psychology Press; 2014.

[46] Bruner J, Roy C, Ratner N. The beginnings of request. Children's language. 1982;3:91–138.

[47] Adamson LB, Bakeman R. Mothers' communicative acts: Changes during infancy. Infant Behavior and Development. 1984;7(4):467–478.

[48] Sumioka H, Yoshikawa Y, Morizono M, et al. Socially developmental robot based on self-induced contingency with multi latencies. In: Advances in cognitive neurodynamics (iii). Springer; 2013. p. 251–258.

[49] Adamson LB. Communication development during infancy. Brown & Benchmark; 1995.

[50] Tomasello M, Moore C, Dunham P. Joint attention as social cognition. Joint attention: Its origins and role in development. 1995;:103–130.

[51] Mundy P, Card J, Fox N. Eeg correlates of the development of infant joint attention skills. Developmental psychobiology. 2000;36(4):325.

[52] Tomasello M. The cultural origins of human cognition. Harvard University Press; 2009.

[53] Csibra G, Gergely G. Natural pedagogy. Trends in cognitive sciences. 2009;13(4):148–153.

[54] Watson JS. Smiling, cooing, and" the game". Merrill-Palmer Quarterly of Behavior and Development. 1972;18(4):323–339.

[55] Movellan JR, Watson JS. The development of gaze following as a bayesian systems identification problem. In: Development and Learning, 2002. Proceedings. The 2nd International Conference on; IEEE; 2002. p. 34–40.

[56] Fong T, Nourbakhsh I, Dautenhahn K. A survey of socially interactive robots. Robotics and Autonomous Systems. 2003;42(3-4):143–166.

[57] Goodrich MA, Schultz AC, et al. Human–robot interaction: a survey. Foundations and Trends® in Human–Computer Interaction. 2008;1(3):203–275.

[58] De Santis A, Siciliano B, De Luca A, et al. An atlas of physical human–robot interaction. Mechanism and Machine Theory. 2008;43(3):253–270.

[59] Matsusaka Y, Fujie S, Kobayashi T. Modeling of conversational strategy for the robot participating in the group conversation. In: Seventh European conference on speech communication and technology; 2001.

[60] Lemon O, Bracy A, Gruenstein A, et al. A multi-modal dialogue system for human-robot conversation. In: Proceedings of North American Association for Computational Linguistics (NAACL 2001); 2001.

[61] Hayashi K, Kanda T, Miyashita T, et al. Robot manzai: Robot conversation as a passive–social medium. International Journal of Humanoid Robotics. 2008;5(01):67–86.

[62] Kondo Y, Takemura K, Takamatsu J, et al. A gesture-centric android system for multi-party human-robot interaction. Journal of Human-Robot Interaction. 2013;2(1):133–151.

[63] Cerrato L, Campbell N. Engagement in dialogue with social robots. In: Dialogues with social robots. Springer; 2017. p. 313–319.

[64] Kanda T, Hirano T, Eaton D, et al. Interactive robots as social partners and peer tutors for children: A field trial. Human–Computer Interaction. 2004;19(1-2):61–84.

[65] Han J. Emerging technologies: Robot assisted language learning. Language Learning & Technology. 2012;16(3):1–9.

[66] Tanaka F, Isshiki K, Takahashi F, et al. Pepper learns together with children: Development of an educational application. In: Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on; IEEE; 2015. p. 270–275.

[67] Gordon G, Spaulding S, Westlund JK, et al. Affective personalization of a social robot tutor for children's second language skills. In: AAAI; 2016. p. 3951–3957.

[68] Shiomi M, Hagita N. Social acceptance toward a childcare support robot system: web-based cultural differences investigation and a field study in japan. Advanced Robotics. 2017;31(14):727–738.

[69] Niemelä M, Heikkilä P, Lammi H. A social service robot in a shopping mall: Expectations of the management, retailers and consumers. In: Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction; ACM; 2017. p. 227–228.

[70] Kanda T, Shiomi M, Miyashita Z, et al. A communication robot in a shopping mall. IEEE Transactions on Robotics. 2010;26(5):897–913.

[71] Shi C, Satake S, Kanda T, et al. A robot that distributes flyers to pedestrians in a shopping mall. International Journal of Social Robotics. 2017;:1–17.

[72] Satake S, Hayashi K, Nakatani K, et al. Field trial of an information-providing robot in a shopping mall. In: Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on; IEEE; 2015. p. 1832–1839.

[73] Bršćić D, Ikeda T, Kanda T. Do you need help? a robot providing information to people who behave atypically. IEEE Transactions on Robotics. 2017;33(2):500–506.

[74] Broekens J, Heerink M, Rosendal H, et al. Assistive social robots in elderly care: a review. Gerontechnology. 2009;8(2):94–103.

[75] Feil-Seifer D, Mataric MJ. Defining socially assistive robotics. In: Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on; IEEE; 2005. p. 465–468.

[76] Sabelli AM, Kanda T, Hagita N. A conversational robot in an elderly care center: an ethnographic study. In: Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on; IEEE; 2011. p. 37–44.

[77] Heerink M, Kröse B, Evers V, et al. Influence of social presence on acceptance of an assistive social robot and screen agent by elderly users. Advanced Robotics. 2009;23(14):1909–1923.

[78] Tamura T, Yonemitsu S, Itoh A, et al. Is an entertainment robot useful in the care of elderly people with severe dementia? The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2004;59(1):M83–M85.

[79] Toh E, Poh L, Causo A, et al. A review on the use of robots in education and young children. Journal of Educational Technology & Society. 2016;19(2).

[80] Han JH, Jo MH, Jones V, et al. Comparative study on the educational use of home robots for children. Journal of Information Processing Systems. 2008;4(4):159–168.

[81] Wu WCV, Wang RJ, Chen NS. Instructional design using an in-house built teaching assistant robot to enhance elementary school english-as-a-foreign-language learning. Interactive Learning Environments. 2015;23(6):696–714.

[82] Vogt P, De Haas M, De Jong C, et al. Child-robot interactions for second language tutoring to preschool children. Frontiers in human neuroscience. 2017;11:73.

[83] Hung IC, Chao KJ, Lee L, et al. Designing a robot teaching assistant for enhancing and sustaining learning motivation. Interactive learning environments. 2013;21(2):156–171.

[84] Abe K, Shiomi M, Pei Y, et al. Chicaro: tele-presence robot for interacting with babies and toddlers. Advanced Robotics. 2018;32(4):176–190.

[85] Arita A, Hiraki K, Kanda T, et al. Can we talk to robots? ten-month-old infants expected interactive humanoid robots to be talked to by persons. Cognition. 2005;95(3):B49–B57.

[86] Meltzoff AN. Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. Developmental psychology. 1995;31(5):838.

[87] Ponder E, Kennedy W. On the act of blinking. Experimental Physiology. 1927;18(2):89–110.

[88] Doughty MJ. Further assessment of gender-and blink pattern-related differences in the spontaneous eyeblink activity in primary gaze in young adult humans. Optometry and Vision Science. 2002;79(7):439–447.

[89] Sforza C, Rango M, Galante D, et al. Spontaneous blinking in healthy persons: an optoelectronic study of eyelid motion. Ophthalmic and Physiological Optics. 2008;28(4):345–353.

[90] Dahlbäck N, Jönsson A, Ahrenberg L. Wizard of oz studieswhy and how. Knowledge-based systems. 1993;6(4):258–266.

[91] Riek LD. Wizard of oz studies in hri: a systematic review and new reporting guidelines. Journal of Human-Robot Interaction. 2012;1(1):119–136.

[92] Kamide H, Takashima K, Arai T. Development of japanese version of the psychological scale of anthropomorphism. The Japanese Journal of Personality. 2016;25(3):218–225.

[93] Heerink M, Kröse B, Evers V, et al. The influence of social presence on acceptance of a companion robot by older people. Journal of Physical Agents. 2008;2(2):33–40.

List of Publications by the Author

Articles in Journals

1. Mahzoon H, Yoshikawa Y, Ishiguro H. Social Skill Acquisition Model through Face-to-face Interaction: Local Contingency for Open-ended Development . Frontiers in Robotics and AI. 2016; 3:10.

2. Mahzoon H, Yoshikawa Y, Ishiguro H. Ostensive-cue Sensitive Learning and Exclusive Evaluation of Policies: A Solution for Measuring Contingency of Experiences for Social Developmental Robot. Frontiers in Robotics and AI. 2019;6:2.

3. Mahzoon H, Ogawa K, Yoshikawa Y, Tanaka M, Ogawa K, Miyazaki R, Ota Y, Ishiguro H. Effect of a Robot's Sociability on Its Mind Perception and Interactability: A Case Study on a Home-Use Robot. Advanced Robotics (*in prog.*).

Papers in Proceedings of International Conferences

1. Mahzoon H, Yoshikawa Y, Ishiguro H. Developmental Robot with Ostensive Cue Sensitive Learning for Real–world Interaction based on Local Contingency Evaluation. In: Development and Learning and Epigenetic Robotics (ICDL-EpiRob), 2016 Joint IEEE International Conference on; IEEE; 2016. p. 308-309.

Papers in Proceedings of Domestic Conferences and Miscellanies

1. Mahzoon H, Ogawa K, Yoshikawa Y, Tanaka M, Ogawa K, Miyazaki R, Ota Y, Ishiguro H. Sociability for Better Evaluation: Robot with Self-representation of Sociability Improves Perceived Mind Perception and Positive Relation by Interacting Human. The 1st International Symposium on Symbiotic Intelligent Systems. 2019.

2. Mahzoon H, Yoshikawa Y, Ishiguro H. Exclusive Evaluation of Policies with Ostensive-cue Sensitive Learning for Developmental Robot in Real World Interaction. The 5th Conference of Developmental Neuroscience Society. 2016.

3. Sakai K, Mahzoon H, Dalla Libera F, Yoshikawa Y, Ishiguro H. Toward a Developmental Robot in Multi-person Environment based on Contingency Evaluation. The 4th Conference of Developmental Neuroscience Society. 2015.

4. Mahzoon H, Yoshikawa Y, Ishiguro H. Local Information Measure for Evaluating Chain of Contingency. International Workshop on Cognitive Neuroscience Robotics. 2014.

5. Mahzoon H, Yoshikawa Y, Ishiguro H. Social Skill Acquisition Using Local Causality Evaluation. The 3rd Conference of Developmental Neuroscience Society. 2014.

6. Mahzoon H, Yoshikawa Y, Ishiguro H. Reliable contingency estimation for social development in real world. International Symposium on Cognitive Neuroscience Robotics. 2013.

7. Mahzoon H, Yoshikawa Y, Sumioka H, Ishiguro H. Elemental Evaluation of the Contingency for Real World Development. The 31st Annual Conference of the Robotics Society of Japan (RSJ). 2013.

8. Mahzoon H, Yoshikawa Y, Sumioka H, Ishiguro H. Analysis of Information Measure for Sensory Motor Contingency toward Modeling Social Development. The Robotics and Mechatronics Conference. 2013.

Acknowledgments

I would first like to express the deepest appreciation to my supervisor Professor Hiroshi Ishiguro. Without his kind and grateful support, my research life would not be possible. I would also like to thank my dissertation committee Professor Koh Hosoda and Professor Takayuki Nagai. Their essential comments about the dissertation helped me to learn more and improve the dissertation.

I would like to express my sincere gratitude to my research group director Dr.Yuichiro Yoshikawa. He help me in both research and life during all of my research life. I sincere thank to Dr.Yutaka Nakamura for his technical comments about my research. I also would like to thank Professor Minoru Asada and Dr.Yukie Nagai for producing grateful comment to my research. I express my thanks to all of the staffs and the students of Intelligent Robotics Laboratory at Osaka University.

Finally, I wish to express my sincere thank to my family. Without their support, I couldn't complete this research.

Hamed Mahzoon

February 7, 2019