An Approach for a Social Robot to Understand Human Relationships: Friendship Estimation through Interaction with Robots

Takayuki Kanda[1] and Hiroshi Ishiguro[1,2]

[1] ATR, Intelligent Robotics and Communication Laboratories
2-2-2 Hikaridai Sorakugun Seikacho, 619-0288 Kyoto, Japan
`kanda@atr.jp`
[2] Osaka University, Graduate School of Engineering, Dept. Adaptive Machine Systems
Osaka, Japan
`ishiguro@ams.eng.osaka-u.ac.jp`

Abstract. This paper reports our research efforts on social robots that recognize interpersonal relationships. These investigations are carried out by observing group behaviors while the robot interacts with people. Our humanoid robot interacts with children by speaking and making various gestures. It identifies individual children by using a wireless tag system, which helps to promote interaction such as the robot calling a child by name. Accordingly, the robot is capable of interacting with many children, which would cause spontaneous group behavior of the children around it. Here, group behavior is associated with social relationships among the children themselves. For example, a child may be accompanied by his or her friends and then play together with them. We propose the hypothesis that our interactive robot prompts a child's friends to accompany him or her; thus, we can estimate their friendship by simply observing their accompanying behaviors.

We conducted a field experiment for two weeks in a Japanese elementary school to verify this hypothesis. In the experiment, two "Robovie" robots were placed where children could freely interact with them during recesses. As a result, we found that they mostly prompted friend-accompanying behavior. Moreover, we could estimate some of their friendly relationships, in particular among children who often appeared around the robot. For example, we could estimate 5% of all friendships with 80% accuracy, and 15% of them with nearly 50% accuracy. Thus, this result basically supports our hypothesis on friendship estimation from an interactive humanoid robot. We believe that this ability to estimate human relationships is essential for robots to behave socially.

Keywords: human-robot interaction, social robot, observation of interaction, human social relationships

1. Introduction

1.1 The Interaction-Oriented robots

Recent progress in robotics has brought with it a new research field known as “interaction-oriented robots.” These robots are different from traditional task-oriented robots, such as industrial robots that perform certain tasks in limited applications. Interaction-oriented robots are designed to communicate with humans and to participate in human society. We are trying to develop such an interaction-oriented robot that will function as a partner in people’s daily lives. We believe these robots will not only be used for entertainment but also provide communication support such as route-guidance and mental support tasks.

Several researchers are endeavoring to realize interaction-oriented robots, which they often refer to as interactive robots or social robots. **Table 1** describes some related research activities on developing interaction-oriented robots. Aibo was the first interactive robot to prove successful in the commercial market (Fujita, 2001) and it behaves as if it were a real animal pet. Breazeal et al. developed the face robot Kismet, and they are further exploring sociable aspects of robots (Breazeal & Scassellati, 1999). Okuno et al. developed a humanoid head robot that tracks a speaking person by using visual and auditory data. In addition, they controlled the personality of the robot by changing the tracking parameter (Okuno et al., 2002). This research work mainly focused on the fundamental interactive ability of the robots.

Human-like bodies have been used to enhance robots’ communication abilities. Humans utilize eye-gaze and pointing gestures in communication even when they are infants, which is a phenomenon widely known in developmental psychology as joint attention (Moore & Dunham, 1995). Many researchers in robotics believe that this is an essential function for both humans and robots to be social. Scassellati developed a robot with a joint-attention mechanism that follows others' gazes in order to share attention (Scassellati, 2000). Kozima et al. also developed a robot with a joint-attention mechanism (Kozima & Vatikiotis-Bateson, 2001). They have implemented their robots so that they recognize humans' eye-gaze and pointing in order to imitate joint attention behaviors, which will probably lead to further abilities such as recognition of intentions.

Moreover, several research works are exploring the possible applications of interactive robots. Burgard et al. developed a museum tour-guide robot (Burgard et al., 1998) that was equipped with robust navigational skills and behaved as a museum orientation tool. For home environment, NEC Corporation has developed a prototype of a personal robot that recognizes individuals' faces, entertains family members with its limited speech ability, and performs as an interface for television and e-mail (NEC Co., 2002). Shibata et al. successfully applied a seal-like pet robot, Paro, to the robot-assisted tasks of easing the mental stress of elderly people (Shibata & Tanie, 2001). Such a wide range of developments for interactive robots have been summarized and introduced in a recent book (Severin, 2003). Dautenhahn et al. established the AURORA project, where robots are applied to autism therapy. They have applied a simple interactive robot for autism therapy (Dautenhahn & Werry, 2002). In their later work, they utilized a small humanoid robot, Robota, for this purpose (Robins et al., 2004), which demonstrated that a robot could work as a mediator for inter-human communication. These research efforts are all devoted to social robots that are embedded in human society.

Table 1: Research activities toward social robots and interactive robots.

	Communication ability		Human-like Appearance		Task			Other functions, etc.
	Inter-active (2only)	Social (>2+)	Head	Body	Com-muni-cation	Mental sup-port	Details	
Aibo (Fujita 2001)	✓						Entertainment	Successful in markets as a pet robot
Kismet (Breazeal et al., 1999)	✓		✓		(✓*)		Infant-like interaction (Research tool)	Facial expressions
SIG (Okuno et al.,2002)	✓		✓				(Research tool)	Human localization by vision and audition
Cog (Scassellati, 2000)	✓		✓	✓			(Research tool)	Joint-attention
Infanoid (Kozima et al., 2001)	✓		✓	✓			(Research tool)	Joint-attention
Museum robot (Burgard et al., 1998)	✓				✓		Museum orientation	Navigation
PaPeRo (NEC Co., 2002)	✓		✓		✓		Entertainment, interface for TV & e-mail	Person identification
Paro (Shibata et al., 2001)	✓					✓	Care for elderly person	Seal-like pet robot
Dautenhahn's robot (Dautenhahn et al., 2002)	✓					✓	Care for autistic children	Simple enough not to confuse children
Robota for Aurora project (Robins et al., 2004)	✓		✓	✓		✓	Care for autistic children	Imitation of human arm movements
Asimo (Sakagami et al., 2002)	✓		✓	✓	✓		Route guide for visitors	Navigation, person identification
Robovie (Kanda et al., 2004a) (early version)	✓		✓	✓			Entertainment	
Robovie (Kanda et al., 2004b)	✓	✓	✓	✓	✓		To motivate foreign language learning	Person identification and adaptation

Furthermore, humanoid robots have been used to support human needs in everyday life, since their human-like bodies enable natural communication with humans as humans do. Honda's humanoid robot Asimo, which inspired much other robotics research on humanoid robots, has recently been used as a route guide for visitors in their laboratory (Sakagami et al., 2002). Robovie has been developed for interacting with children (Ishiguro et al., 2003, Kanda et al., 2004a), and has recently been applied as a peer tutor of English as a foreign language with its wireless-tag-based social communication ability (Kanda et al., 2004b), which was the first use of an interactive humanoid robot in everyday life. As in these examples, we believe that the tasks of interaction-oriented robots will be increased and enhanced along with growth in their communication abilities.

1.2 Our Research Approach toward an Interaction-Oriented Robot that Participates in Human Society

The research question we aim to address is "How can an interaction-oriented robot participate with people in their daily life, establish social relationships, and contribute to society?" In other words, our purpose is to realize a peer or partner that communicates socially with humans to support their daily lives. Although the appearance of interaction-oriented robots is becoming more human-like, their social abilities are still far from those of human beings. Most of the current interactive robots are still interactive toys that have physical bodies. Most previous social robots simply interact with one person. Although some of them establish a relationship with a certain person through personalization or adaptation mechanism, they cannot establish social relationships with more than one person. We believe that to interact with more than one person at a time is one of the essential social abilities, denoted as *social* in Table 1.

We believe that the social ability of robots will be greatly improved by putting these robots into human society. The initial tasks of the robots will be limited and perhaps not so important, since the interaction abilities of the current robots are not as high as human infants' and their social skills are very low. However, the unsolved problems and inadequate abilities of the robot, such as certain aspects of social skills, will become clearer as they work in human society, which will allow us to improve the robot's abilities. This process is most likely similar to overcoming problems human children experience in their development. An infant improves his or her interactive ability through interaction with mothers at home; children improve their social skills in human society by actions such as playing with neighbors or other students in elementary school, whereas adults work while socially communicating with each other. **Figure 1** describes our research approach toward the interaction-oriented robot as a metaphor of such human development.

Currently, robots are applied to work in our daily lives as interactive robots, such as the robots in several previous works (Fujita 2001, Burgard et al., 1998, NEC Co., 2002, Shibata & Tanie, 2001, Severin, 2003, Sakagami et al., 2002), and they are gradually growing in their interactive abilities; however, they have not yet reached the stage of social activities that require social communication with more than one person. This developmental approach for robots (Cowley & MacDorman, 1995), illustrated by the solid line in Figure 1, is already considered as a metaphor for human infants, known as cognitive developmental robotics (Asada et al., 2001). Several researchers have suggested and practiced the approach by such means as implementing triadic relationships among human-object-human through joint-attention (Moore & Dunham, 1995). While these researchers are developing robots' interactive abilities for only one person in front of the robots, we believe it is also indispensable to improve robots' social ability to make robots work in our daily lives, which is the approach indicated by the thick broken line in Figure 1. We believe that robots' tasks will emerge according to the improvement in their abilities, even though current robots are equipped with little skill for accomplishing useful tasks in human society, as indicated by the thin broken line in Figure 1.

[FIGURE 1 ABOUT HERE]

We are pursuing this approach of gradually making robots more capable of working in our daily lives to improve their social abilities as well as to explore other possible tasks. The first step in this approach was a field trial in an elementary school in which interactive robots behave as peer tutors of a foreign language, English, as previously reported (Kanda et al., 2004b). The robot, Robovie, was equipped with a person-identification function to distinguish children for such tasks as calling the names of children, and simultaneously interacted with more than one child. As a result, it was demonstrated that interactive robots have the potential to motivate children to learn a foreign language. Meanwhile, we observed group behavior among friends. For instance, a boy and his friend counted how many times the robot called their respective names, and the boy whose name was called more often proudly told his friend that the robot preferred him. If the robot could identify friendships among children, it could increase its interaction with the boys as well as promote interaction between the boys. That is, the capacity to estimate friendship can help robots mediate human interaction. Moreover, friendship is tightly connected to social relationships (described in the next section in detail). Thus, this friendship estimation is essential for accomplishing more general social relationship estimation, which might provide a future social robot that could help to solve bullying problems or the problem of rejected children. In this paper, we report our approach to estimating human friendship by using an interactive robot. We believe this ability is essential for interactive robots to be social.

1.3 Existing Research related to Friendship and Social Network Estimation

Friendship is defined as a close, mutual, and voluntary dyadic bilateral relationship, which is often formed based on similarity among individuals, common interests, and common activities (Rubin et al., 1999). It is a well-grounded finding from psychological research that children at a very young age engage in dyadic relationships, for example, in the form of pretend play. Then, with age, they increase the size, from dyadic to more than triadic, and complexity of relationships. As they grow up, they form many different peer relationships in the form of social networks. As children gradually establish social networks, each child attains a different social status (Gottman & Parkhurst, 1980, Ladd et al., 1990).

A sociometric test has been used to investigate peer relationships and social networks, and this lets a human directly answer with the names of others whom he or she likes or dislikes. This method has been validated as a reliable assessment of human-peer relationships. It categorizes each child's social status into one of the following groups: popular, average, neglected, and rejected (McConnell & Odom, 1986, Asher & Hymel, 1981). This method has been widely used to determine the relationships in a classroom or a company.

As an alternative, observation-based methods have been developed for identifying peer relations and social status. A child's behaviors toward friends are different from those toward non-friends. This difference has been investigated through more than 80 research works (Newcomb & Bagwel, 1995). For example, positive engagement (talk, smiling, laughter) was observed more often among friends rather than among non-friends. How children form groups and behave within these groups is associated with their friendly relationships. Children usually play with their peers, although there seems to be some gender difference in the size of play groups (Waldrop & Halverson, 1975). Ladd et al. investigated the associations between observed group behavior and relationships among group members. They coded videotape of children's play with four behavioral measures: cooperative play, rough play, unoccupied, and teacher-orientation. Their observations revealed that a child's cooperative play is associated with positive nominations, while his or her rough play is related to negative nominations. In addition, they revealed that past behavior successfully predicted the current peer status; for example, time spent in cooperative play was a significant predictor of positive nomination (Ladd et al., 1990). Coie et al. investigated the difference between popular and rejected children in terms of their behavior and revealed the relationship between rejected children and their aversive behaviors (Coie & Kupersmidt, 1983). We believe these findings support the notion that social robots can recognize humans' peer relationships and social status by observing their group behavior.

1.4 Existing Research related to Interpersonal Behavior Analysis from Sensor-based Observation

Our research approach is to recognize human social relationships by observing group behavior in the presence of an interactive robot, which is closely related to research works that attempt to analyze human interaction from sensors embedded in environments or attached to humans. Recently, several research works have attempted to automatically observe and analyze large-scale human behaviors by using virtual reality, wearable computing, and ubiquitous sensing technologies.

Velde et al. proposed a mixed-reality approach for supporting human activity at conferences (Velde, 1997), which is known as the COMRIS project. Here, a parrot-like physical agent perched on human shoulders supplies information connected to a backbone information infrastructure. At the same time, this system measures interactions in large groups occurring spontaneously. With recent ubiquitous sensing technology, Sumi et al. developed a ubiquitous sensor environment to capture and analyze human physical group interaction by using infrared sensors that are good at sensing movements such as eye gaze (Sumi et al., 2003). By using this sensing system, Bono et al. analyzed human social behaviors during a poster session in an exhibition, which revealed the role of non-verbal cues for conversation between an exhibitor and visitors as well as the possibility of detecting their preferences for the exhibited posters (Bono et al., 2003). Similarly, Choudhury et al. have utilized auditory-based wearable sensors for analyzing human communication networks (Choudhury & Pentland, 2003). These sensor-based approaches for analyzing human social behaviors assume that an interactive robot can feasibly use this equipment to analyze human interaction in order to recognize human social relationships. In the future, sensors embedded in environments will collaboratively capture human social behaviors with interactive robots (Hagita et al., 2003).

On the other hand, several research works have attempted to analyze how people behave toward interactive robots. For example, Dautenhahn et al. analyzed children's eye gaze and contact time with an interactive robot in the AURORA project (Dautenhahn & Werry, 2002), which aims to apply interactive robots to help autistic children to acquire social skills. Moreover, a motion capturing system enables us to automatically capture human embodied behavior during human-robot interaction. Kanda et al. measured the spatial and temporal synchrony between humans and an interactive robot to predict humans' subjective impressions of the robot (Kanda et al., 2004a).

2 Robovie: An Interactive Humanoid Robot

2.1 Hardware Configuration

2.1.1 "Robovie"

Figure 2 shows the humanoid robot "Robovie" (Ishiguro et al., 2003). The robot is capable of human-like expression and recognizes individuals by using various actuators and sensors. Its body possesses highly articulated arms, eyes, and a head, which were designed to produce sufficient gestures for communicating with humans, such as greeting, hand-shaking, and pointing and gazing in a direction for joint-attention. The sensory equipment includes auditory, tactile, ultrasonic, and vision sensors, which allow the robot to behave autonomously and to interact with humans. All processing and control systems, such as the computer and motor control hardware, are located inside the robot's body.

[FIGURE 2 ABOUT HERE]

2.1.2 Wireless Person Identification System

To identify individuals, we used a wireless tag system capable of simultaneous multi-person identification. Recent RFID (radio frequency identification) technologies have enabled us to use contact-less identification cards in practical situations. In this study, children were given easy-to-wear nameplates (5 cm in diameter) in which a wireless tag was embedded. A tag (Fig. 2, lower-right) periodically transmitted its ID to the reader installed on the robot. In turn, the reader relayed the received IDs to the robot's software system. It was possible to adjust the reception range of the receiver's tag in real-time by software. The wireless tag system provided the robots with a robust means of identifying many children simultaneously. Consequently, the robots could show some human-like adaptation by recalling the interaction history of a given person (history of executed situated modules and their results for the person) and applying it to choose the appropriate situated module for the person as well as personalizing its interaction for the person such as calling his or her name (Kanda et al., 2003).

A simple test of the wireless system shows us that it is capable of detecting a person's distance with some tolerance. The reader has eight steps of attenuation (a mechanism for increasing electrical resistance to weaken received radio signal) that reduce the maximum gain of the receiver by 12.5% with each step. We measured how stably the system could detect the tag with each attenuation parameter. The results are shown in **Figure 3**. As the attenuation parameter setting is increased, the readable area decreases. This is represented by R in the graph, where the gain is R/8 of the maximum. In addition, since the readable area with the attenuation R=8 is smaller than R=5,6,7, we did not use the lowest attenuation level (R=8); at that level it seemed that the tag system became oversensitive to the noise radiated by the robot itself.

[FIGURE 3 ABOUT HERE]

2.2 Software Architecture

"Robovie" features a software mechanism for performing consistent interactive behaviors (Kanda et al., 2002). **Figure 4** outlines the software systems that enable the robot to simultaneously identify multiple persons and autonomously interact with them based on its memory for each person. The basic components of the system are "situated modules" and "episode rules." The robot system sequentially executes situated modules according to execution orders defined by the episode rules.

[FIGURE 4 ABOUT HERE]

With respect to person identification, the architecture utilizes four kinds of databases (DB): Person ID DB to remember internal IDs for each person, episode rules to control the execution orders of the situated modules, "public" and "private" episodes to maintain communication with each person, and "long-term individual memory" to memorize information about individuals. The "module control" is in charge of the total execution of the situated modules by referring to the episode rules and episodes (history of communication).

Each situated module consists of "communicative units." The communicative units are principal elements of interactive behaviors, such as eye contact and arm movements synchronized with an utterance. By combining communicative units, the developer can easily and quickly implement new situated modules. "Reactive modules" handle emergencies in both movement and communication. For example, the robot stops when it collides with a wall and then returns to the original episode. In the situated and reactive modules, inputs from sensors are pre-processed by sensor modules such as speech recognition. Actuator modules perform low-level control of actuators.

2.2.1 Communicative Units

Humans use eye contact and arm gestures for smooth interaction. The communicative unit is an elemental unit for body movement in human-robot communication. Each communicative unit is a sensor-action unit. Specifically, we have implemented "eye contact," "nod," "positional relationship," "joint attention (gaze and point object)," and so forth. Situated modules are implemented by connecting the communicative units with other sensor-action units needed for the behavior, such as a particular utterance, and positional movements.

2.2.2 Situated Modules

In linguistics, an "adjacency pair" (Levinson, 1983, p.303) is a well-known term for a unit of conversation where the first expression of the pair requires the second expression to be of a certain type. For example, "greeting and response" and "question and answer" are considered pairs. Similarly, human-robot interaction can be divided into action-reaction pairs. That is, when a human takes an action toward a robot, the robot reacts to the human's action; and when the robot takes an action toward the human, the human reacts to the robot's action. In other words, the continuation of the actions and reactions forms the interaction.

Although the number of actions and reactions between humans and robots should be equal, at present the recognition ability of the robot is not as powerful as the humans'. Therefore, the robot actively takes actions rather than making reactions in order to sustain communication with the human. In other words, in our approach, robots are proactive in initiating interactions and let humans adaptively respond to their actions. The robots' initiation of actions poses constraints on the recognition of human reaction. Each situated module is designed to realize a certain action-reaction pair in a particular situation, where the robot mainly takes an action and recognizes the humans' reaction. The robot responds to deviation from the expected action-reaction pair in a given situation through a reactive transition to other situated modules or an activation of reactive modules.

Precondition, Indication, and Recognition Parts

Each situated module consists of precondition, indication, and recognition parts, as shown in **Figure 5**. By checking its precondition, the robot knows whether the situated module is executable. For example, the situated module that talks about the weather by retrieving weather information from the Internet is not executable (precondition is not satisfied) when the robot cannot access the Internet. The situated module that asks to shake hands is executable when a human (a moving object located near the robot) is in front of the robot.

[FIGURE 5 & 6 ABOUT HERE]

By executing the indication part, the robot takes an action to interact with humans. For example, the robot says, “Let’s shake hands,” and offers its hand in the hand-shake module. This behavior is achieved by combining communicative units for eye contact and for maintaining positional relationships (moving its body toward the human), while speaking the sentence “Let’s shake hands” and making a particular body movement to offer its hand. The recognition part is designed to recognize humans’ reactions affected by the indication part. The situated module creates the particular situation between the robot and the human; therefore, the recognition part can predict certain human responses that are highly probable for the situation. By expecting a specific set of responses, the necessary sensory processing can be tuned to the situation. Thus, the robot can recognize complex human behaviors with simple sensory data processes. When the robot performs situated recognition by sight, we call this “situated vision.”

Sequential and Reactive Transition of Situated Modules, and Reactive Modules

After the robot executes the indication part of the current situated module, it recognizes the human's reaction by using the recognition part. It then records a result value corresponding to the recognition result and moves to the next executable situated module (**Figure 6** (a)). The next module is selected by using result values and the execution history of the situated modules (episode). This sequential transition is defined by the episode rules.

Episode rules allow for consistent transitions between the situated modules. Sequential transition according to episode rules does not represent all transition patterns needed for human-robot communication. In fact, there are other types of transitions for interruption. Let us consider the following situation. When two persons are talking, a telephone suddenly rings. They stop talking and respond to the telephone call. For the robot, interruption is dealt with as a reactive transition, which causes a deviation. Reactive transitions are also defined by some episode rules (Figure 6 (b)). If a reactive transition is assigned for the current situation and the precondition of the assigned succeeding situated module is satisfied, the robot stops executing the current situated module and immediately moves to the next situated module.

The reactive modules are also prepared for an interruption, but in this case, the robot does not quit the execution of the current situated module (Figure 6 (c)). Instead of causing deviation, the robot executes the reactive module in parallel with the current situated module. For example, we implemented a reactive module to gaze at body parts of the robot when they are touched. When a human touches the arm of the robot while it is speaking, the robot gazes at the arm while continuing to speak. This control is similar to subsumption architecture (Brooks, 1986). That is, upper-hierarchy modules (situated modules) suppress lower ones (reactive modules) in our architecture, which is similar to "*subsume*" (more-meta-level module takes over control from a lower-level one, if available) in subsumption architecture.

2.2.3 Distinction of Participant and Observers

In linguistics, Clark classified talking people into two categories: participants and listeners (Clark 1996). Participants are mainly speakers and hearers, and listeners just listen to the conversation and take a passive role in it. Similarly, we classify humans located around the robot into two categories: participants and observers. Since we are concerned only with humans within the robot's scope of awareness, the categories are similar to Clark’s definitions, but our observer category does not include eavesdroppers (persons listening in without the speaker’s awareness).

[FIGURE 7 ABOUT HERE]

The person identification software simultaneously identifies persons and separates them into participant and observer categories (**Figure 7**). The distance between the robot and the humans also enables the robot to categorize them. As Hall discussed, there are several distances between talking humans (Hall, 1966). According to his theory, a distance of less than 1.2 m is “conversational,” and a distance from 1.2 m to 3.5 m is “social.” Persons who meet each other for the first time often talk from the social distance. Our robot recognizes the nearest person as a participant and others stably located within a readable distance of the wireless identification system (within a distance of about 1.2 m) as observers.

2.2.4 Episode Rules: Activation of Situated Modules

Transitions among situated modules are ruled by episode rules, each of which can be considered an activation of situated modules, but the robot executes just one situated module at a time. Conflicts among the activations are mediated by referring to the precondition of situated modules and the priority of episode rules. Episode rules also realize personalized interactive behaviors of the robot by referring to private episodes, which are the individuals’ episodes separately remembered. The following describes the detailed mechanism for the activation and coordination of situated modules realized by episode rules.

Public and Private Episodes

We define an episode as a sequence of interactive behaviors produced by the robot and humans. Internally, it is represented as a sequence of situated modules. There are two types of episodes as shown in **Figure 8**: “public” and “private.” The public episode is the sequence of all executed situated modules. That is, the robot exhibits those behaviors to the public. By contrast, the private episode is a private history of each person. By memorizing each person’s history, the robot adaptively behaves toward the person who is participating in or observing the communication.

[FIGURE 8 ABOUT HERE]

Table 2: Grammar of episode rules

1. <ModuleID=result_value>…<…>NextModule
2. (<ModuleID1=result_value1>|<ModuleID2=result_value2>)...
3. (…){n,m}…
4. !<…>NextModule
5. ^<ModuleID=^result_value>NextModule

(1:basic structure of describing executed sequence, 2: “OR”, 3: repetitions, 4: negation of *episode rule*, 5: negation of Module ID and result value)

Episode Rules for Public Episodes

The episode rules direct the robot into a new episode of interaction with humans by controlling transitions among situated modules. They also give consistency to the episode. When the robot switches the situated modules, all episode rules are compared with the current situated module and the episode (execution history of situated modules) to determine which situated module to execute next. The system performs the comparison in the background of the current situated module’s execution and prepares the next executable module list. After the current module’s execution, the robot checks the preconditions of each situated module in the list. If the precondition is satisfied, it transits to the next situated module. Each episode rule has a priority. If some episode rules conflict, the episode rule with the higher priority is used.

Table 2 indicates the basic grammar of the episode rule. Each situated module has a unique identifier called a Module ID. "<Module ID=result_value>" is the rule that refers to the execution history and the result value of the situated modules, then "<ModuleID1=result_value 1> <ModuleID2=result_value 2>…" means the referring rule of the previously executed sequence of situated modules (Table 2-1). “<…>|<…>” means a selective-group (OR) of the executed situated modules, and “(…)” means the block that consists of a situated module, a sequence of situated modules, or a selective-group of situated modules (Table 2-2). In a manner similar to that of regular expressions, we can describe the repetition of the block as "(…){n,m}", where n gives the minimum number of times matching the block and m gives the maximum (Table 2-3). We can specify the negation of the whole episode rule with an exclamation mark "!". For example, “!<…>…<…>NextModuleID” (Table 2-4) means the module of NextModuleID will not be executed when the episode rule matches the current situation specified by “<…>…<…>.” The negation of a ModuleID or a result value is written with a caret character "^" (Table 2-5).

Episode Rules for Private Episodes

We introduce two characters “P” and “O” to specify participation and observation, respectively. If there is a “P” or “O” character at the beginning of the episode rule, the episode rule refers to the private episodes of the current participant or observers. Otherwise, the episode rules refer to public episodes. If the first character in the angle bracket is “P” or “O,” it indicates that the person experienced the module as a participant or an observer, respectively. Thus, “<P ModuleID=result_value>” is a rule to represent “if the person participated in the execution of ModuleID and it resulted in the result value.” Omission of the first character means, “The person participated or observed it.”

Examples

Figure 8 is an example of public and private episodes, episode rules and their relationships. The robot memorizes the public episode and the private episodes that correspond to each person. Episode rules 1 and 2 refer to the public episode, which realizes self-consistent behaviors of the robot. More concretely, episode rule 1 realizes a sequential transition that “the robot will execute the situated module SING next, if it is executing GREET and it results in *greeted*.” Similarly, episode rule 2 realizes the reactive transition “if persons touch the shoulder, the precondition of TURN is satisfied and then the robot stops execution of SING to start TURN.”

There are also episode rules that refer to private episodes. Episode rule 3 means that if all modules in the participant’s individual episode are different from GREET, it will execute GREET next. Episode rule 4 represents “once the person hears the robot’s song, the robot will not sing the same song for a while.” As in these examples, the episode rule lets the robot adaptively behave toward individuals by referring to the private episodes.

2.3 Implementations of Interactive Behaviors

The objective behind the design of the interactive behaviors of Robovie is that it behaves as a young child to interact with children. One hundred interactive behaviors have been developed. Seventy of them are interactive behaviors such as shaking hands, hugging, playing paper-scissors-rock, exercising, greeting, kissing, singing, briefly conversing, and pointing to an object in the surroundings (shown in **Figure 9**). Twenty are idling behaviors such as scratching the head or folding the arms, and the remaining ten are moving-around behaviors. In total, the robot can utter more than 300 sentences and recognize about 50 words.

[FIGURE 9 ABOUT HERE]

Several interactive behaviors depend on the person identification function based on the RFID tag system. For example, in one interactive behavior the robot calls a child's name if that child is at a conversation distance (about 1.2 m, where he or she is recognized as a participant or observer). This behavior is useful for encouraging the child to come and interact with the robot. Another interactive behavior is a body-part game, where the robot asks a child to touch a body part by saying the part's name.

[FIGURE 10 ABOUT HERE]

The interactive behaviors appear in the following manner based on simple rules (**Figure 10**). The robot sometimes triggers the interaction with a child by saying "Let's play, touch me," and it exhibits idling or moving-around behaviors until the child responded; once the child reacted, it continued performing friendly behaviors for as long as the child responds. When the child stops reacting, the robot stops the friendly behaviors, saying "good bye," and re-starts its idling or moving-around behaviors.

3 Reading Friendly Relationships among Individuals

3.1 Friendship estimation model

Human behavior is largely based on social relationships, which can be in the form of dyadic relationships, known as friendship, or larger groups known as social networks, where there are complex peer relationships among different individuals. Since the previous research has proven the correlations between children's group behavior and the valence of their relationships (Ladd et al., 1990, Waldrop et al., 1975, Coie et al., 1983), we believe we can estimate their peer relationships and social networks by observing their group behavior. Subjective reports and the results from observational methods are, however, not always consistent even in psychological trials (Hartup, 1996). This implies that even if robots could perfectly recognize children's behavior, the estimation would not perfectly match the subjective reports. Rather, we focused on the *estimation* of peer relationships, which are the fundamental parts of a social network, as an early attempt at *recognizing* peer relationships and social networks. Yet it is not through *recognition* (finding all correct information accurately) but through *estimation* (partially finding correct information with moderate accuracy) that robots can use the obtained information to further promote human-robot interaction.

Our basic hypothesis is that when an autonomous robot interacts with several children simultaneously it stimulates their spontaneous group behavior, which then enables the robot to recognize their relationships. Our friendship estimation model is based on the association of social group behavior and social relations, which is inspired by previous psychological research such as the works mentioned above. In general, social relationships among people affect group behavior, such as accompanying, distance among members, facial expressions during conversation, and so forth. For instance, a human is accompanied by another friendly human, but does not willingly approach (accompanying and close distance) a disliked human. Sometimes, such dislike relations cause a quarrel or fight (distance is short, but facial expression will be far from friendly). Meanwhile, forced relationships sometimes trigger non-spontaneous group behavior. For instance, a teacher may organize co-working activity such as "children collaborate to carry a heavy box." The left figure in **Figure 11** illustrates these examples of associations between group behaviors and peer relations in general situations.

[FIGURE 11 ABOUT HERE]

On the other hand, according to our hypothesis, an interactive robot mostly causes spontaneous friendly behaviors. In fact, we observed such a situation where a child is accompanied by his/her friend to interact with the robot as shown in **Figure 12**. (We verify this hypothesis later in this paper.) Thus, we believe we can estimate such friendly relationships by simply observing group behavior. This idea is described in **Figure 13**-right. As the first step is the estimation, we only utilize "accompanying" behavior that can be recognized by using a wireless tag system.

[FIGURE 12 & 13 ABOUT HERE]

3.2 Algorithm

Figure 13-right indicates the mechanism of friendship estimation. From a sensor (in this case, wireless ID tags and receiver), the robot constantly obtains the IDs (identifiers) of individuals who are around it. The robot continuously accumulates its interacting time with person A (T_A) and the time that person A and B simultaneously interact with it (T_{AB}, which is equivalent to T_{BA}). We define the estimated friendship from person A to B ($Friend(A \rightarrow B)$) as

$$Friend(A \rightarrow B) = if\,(T_{AB} / T_A > T_{TH}), \tag{1}$$

$$T_A = \Sigma\ if\,(observe(A) \text{ and } (S_t \leq S_{TH})\,) \cdot \Delta t, \tag{2}$$

$$T_{AB} = \Sigma\ if\,(observe(A) \text{ and } observe(B) \text{ and } (S_t \leq S_{TH})\,) \cdot \Delta t\ , \tag{3}$$

where $observe(A)$ becomes true only when the robot observes the ID of person A, if () becomes 1 when the logical equation inside the parentheses is true (otherwise 0), and S_t is the number of persons simultaneously interacting at time t. We prepared two thresholds T_{TH} and S_{TH}; threshold T_{TH} is for the ratio of simultaneous interaction time, and the robot only accumulates T_A and T_{AB} when S_t (the number of persons simultaneously interacting at time t) is less than threshold S_{TH} (Eqs. 2 and 3). In our trial, we set Δt to one second.

4 Experiment

We conducted a field experiment in an elementary school for two weeks with the developed interactive humanoid robot, which was originally designed to promote the children’s learning of English. As we previously reported (Kanda et al., 2004b), the robots had a positive effect on the children by improving their listening score of English. In this paper, we use the interaction data during that trial as a test-set of our approach to estimating friendship from the children’s interaction.

4.1 Method

We performed the field experiment at an elementary school affiliated with Wakayama University. Two identical humanoid robots were put in the open corridor near the sixth-grade classrooms, and during the recess (approximately one hour per day) for two weeks the two robots interacted with sixth-grade students. The following subsections describe the method of the trial in more detail.

4.1.1 Setting and Participants

In general, there are six grades in a Japanese elementary school. This particular elementary school has three classes for each grade and about 40 students in each class. There were 109 sixth-grade students (11-12 years old; 53 males, and 56 females). **Figure 14**-upper shows the three classrooms. There are no walls between the classrooms and corridor, so that the corridor (which is called a workspace) is open to every sixth grader.

[FIGURE 14 ABOUT HERE]

4.1.2 Procedure

The experiment was conducted for two weeks, which was equivalent to nine school days due to a national holiday. We gave the children safety instructions before the trial. Pictures of the robot were accompanied by messages in Japanese such as “do not treat the robots roughly,” and “do not touch the joints because it is not safe.” We did not give the children any further instructions.

The two robots were put in the corridor as shown in Figure 14 (indicated as Robovie 1 and 2). The children were allowed to interact freely with both robots during recesses, which were in total about one hour per day. Every child had a nameplate with an embedded wireless ID tag (Figure 2, right-bottom) so that the robots could identify the child during interactions.

The teachers were not involved in the field trial. Two experimenters (university students) looked after the two robots. They did not help the children interact with the robots but simply ensured the safety of the children and robots. For example, when the children crowded closely around the robot, the experimenters would tell them to maintain a safe distance.

4.1.3 Data Collection

Time Spent Interacting with the Robot

Each robot was equipped with a wireless ID tag reader that detected and identified ID tags embedded in the nameplates given to the children (described in Section 2.1.2). After identifying the children's IDs, the robot made a detection log of IDs for later analysis about friendship estimation and other statistics, in addition to using it during interaction with the children.

Friendships among children

We administered a questionnaire before the experiment that asked the children to write down the names of their friends. This obtained friendship information was collected for comparison with the friendship relationships estimated by our proposed method.

Other observations of interactions

We also recorded scenes from the field trials with four cameras and two microphones. Figure 14 (upper and bottom-right) shows the arrangement of the cameras and microphones and the obtained scenes of the trial. The video was used only to verify the consistency of the wireless ID tag system and it was not analyzed further.

4.2 Results

4.2.1 Results from interaction with the robot during the two weeks

This subsection summarizes the interaction between robots and children during the two weeks. **Table 3** and **Figure 15** indicate the transition of the interaction. It seemed that plenty of excitement occurred on the first day: up to 17 children simultaneously crowded around the robot, as shown in **Figure 16**. Then, the excitement of the first day soon faded, with the average number of simultaneously interacting children gradually decreasing. In the first week, someone was always interacting with the robots, so the rate of vacant time was still quite low. The interaction between the children and the robots became more like inter-human conversation. Several children got in front of the robot, touched it, and watched its response.

Table 3: Results for the change in children's behaviors at an elementary school

	1st week					2nd week			
	1	2	3	4	5	6	7	8	9
Interacted children	75	67	27	34	47	43	30	21	43
Avg. (max) simul. interacted	7.8 (17)	3.7 (16)	2.3 (10)	4.5 (15)	5.1 (18)	2.7 (15)	2.4 (7)	3.8 (8)	4.5 (24)
Experiment time (min)	62	60	18	18	18	61	49	16	37
Rate of vacant time	0.02	0.05	0.06	0.02	0.01	0.38	0.45	0.28	0.68

[FIGURE 15, 16, 17, & 18 ABOUT HERE]

On the second week, it seemed that saturation had occurred. **Figure 17** shows a scene at the beginning of the second week. At the beginning, the vacancy time around the robots suddenly increased, and the number of children who played with the robots decreased. Near the end, there were no children around the robot during half of the daily experiment time. In the second week, the average number of simultaneously interacting children was 4.4. The way they played with the robots seemed similar to the style of play in the first week. Thus, only the frequency of children playing with the robot decreased. The decreasing interaction levels in repeated interactions has also recently been confirmed by Salter et al (2004).

4.2.2 Results for friend-accompanying behavior

As we compared answers on the questionnaire on friendships and the interaction time with the robot, we found that 72% of children's interaction time with the robot was in the company of one or more friends (see **Figures 18**). That is, this result supports our hypothesis that *our interactive robot mostly causes friend-accompanying behavior of children around it rather than the behaviors associated with other relationships, such as dislike or co-working.* This implies that we can estimate their friendship by simply observing the children's accompanying behavior.

4.2.3 Results for friendship estimation

Since the number of friendships among children was fairly small, we focused on the appropriateness (coverage and reliability) of the estimated relationships instead of the rate of correct classification. This is similar to the evaluation of an information retrieval technique such as a Web search. Questionnaire responses indicated 1,092 friendships among a total of 11,772 relationships; thus, if we suppose a classifier

Table 4: Estimation results with various parameters

		T_{TH} *(ratio of simultaneously interacting time)*					
	coverage *reliability*	0.3	0.2	0.1	0.05	0.01	0.001
S_{TH} *(num. of simultaneously Interacting children)*	2	0.01	0.02	0.03	0.04	0.04	0.04
		1.00	0.93	0.79	0.59	0.54	0.54
	5	0.00	0.02	0.06	0.11	0.18	0.18
		1.00	1.00	0.74	0.47	0.29	0.28
	10	0.00	0.00	0.04	0.13	0.29	0.31
		-	1.00	0.74	0.46	0.23	0.20

('-' indicates that no relationships were estimated, so reliability was not calculated)

that always classifies a relationship as a non-friendship, it would obtain 90.7% rate of correct classification, which seems to be good performance; however, such an evaluation is completely useless. Thus, we evaluate our estimation of friendship based on coverage and reliability, which are defined as follows.

Coverage = number of correct friendships in estimated friendship / number of all correct friendships

Reliability = number of correct friendships in estimated friendships / number of estimated friendships

Here, *correct friendships* means the friendships the children identified in the questionnaire and the *estimated friendships* are given by our estimation method. **Table 4** and **Figure 19** indicate the results of estimation with various parameters (S_{TH} and T_{TH}). In Figure 19, *random* represents the reliability of random estimation where we assume that all relationships are friendships (since there are 1,092 correct friendships among 11,772 relationships, the estimation obtains 9.3% reliability at any coverage). In other words, *random* indicates the lower boundary (or, chance rate) of estimation. Each of the other lines in the figure represents the estimation result with different S_{TH}, which has several points corresponding to different T_{TH}. There is obviously a tradeoff between reliability and coverage, which is controlled by T_{TH}. Relatively S_{TH} has a small effect on the tradeoff: although it is difficult to simply conclude which one performs the best estimation, we suppose that the upper-right line in the figure 19 as the better one, S_{TH}=5 mostly performs the best estimation of the friendship, and S_{TH}=10 performs better estimation than S_{TH} =5 when coverage is more than approximately 0.10. As a result of reading the figure 19, for example, our method estimated about 5% coverage of the friendship relationships with greater than 80% reliability (at "S_{TH}=5") and 15% coverage of them with nearly 50% reliability (at "S_{TH}=10").

[FIGURE 19 ABOUT HERE]

Table 5: Estimation results with various parameters

		T_{TH} (ratio of simultaneously interacting time)					
	coverage *reliability*	0.3	0.2	0.1	0.05	0.01	0.001
Condition (num. correct rel.)	H-H	0.02	0.09	0.22	0.35	0.65	0.65
	(250)	1.00	1.00	0.59	0.29	0.19	0.18
	H-L	0.00	0.14	0.33	0.70	0.16	0.16
	(241)	-	1.00	0.78	0.39	0.26	0.26
	L-L	0.00	0.01	0.06	0.11	0.11	0.11
	(360)	-	1.00	0.85	0.83	0.76	0.76

('-' indicates that no relationships were estimated, so reliability was not calculated)

Furthermore, we analyzed the results in detail to determine how we could improve the estimation. **Figure 19** and **Table 5** show a comparison between frequency of interaction with the robot and estimation results with various T_{TH} at S_{TH} =5. We classified the children into two groups as the higher-frequency half of 109 children (54 children; denoted as H) and the lower-frequency half (55 children; denoted as L) according to their interaction time with the robot. Figure 20 indicates the performance of estimation among H-H (friendships among children within the H group), H-L (friendship between a child in the H group and a child in the L group), and L-L (friendships among children within the L group), where "all" represents the estimation result for all relationships, which is the same as S_{TH}=5 in Figure 18. Results of the comparison revealed that our method better estimated the relationships in the H-H group. In other words, the estimation is more accurate for the relationships between children where each of them often appeared around the robot. In contrast, the upper boundary of coverage is more limited for the relationships between children who did not often appear around it. In addition, the higher reliability in the H-H group was not caused by the difference in the children's accompanying behavior. As Figure 17 shows, there is no correlation between frequency of interaction with the robot and the rate of friend-accompanying behavior, which is also verified by the fact that the Peason correlation value is -0.154, which is not significant. We believe these findings also support the effectiveness of our estimation model, since it seems that the estimation will become more accurate with an increase in the amount of observed data on inter-human relationships around the robot.

[FIGURE 20 ABOUT HERE]

5 Discussion

5.1 Contributions to Research for Interaction-Oriented Robot

The primary contribution of this paper is that it presented a novel approach to acquisition of social skills by interaction-oriented robots. Experimental results revealed that our interactive humanoid robot prompted friend-accompanying behavior that enabled the robot to estimate the friendship among children in an elementary school. These early findings support our general friendship estimation model. We believe this fundamental approach will lead to further research works for richer social skills of interaction-oriented robots, as well as research on using social skills to promote interaction with humans.

5.2 Contributions to Interaction Analysis with Ubiquitous Sensors

Recent progress in sensing technology has enabled us to automatically record large-scale spontaneous group behavior. In our study, we automatically gathered data on the interaction among two robots and 109 children, although we only measured the interaction time. Since this automatic measurement enables the robot to estimate the children's group behavior related to friendship estimation, we believe that this will provide new research tools for the analysis of human interactions in the not-too-distant future.

Robovie's observation accuracy is still far from that of humans. However, it is highly scaleable in terms of spatial size, temporal size, size of human group, and so forth. This versatility is similar to that of a computer. Although they are not as intelligent as humans, computers have extended our society by doing tasks that humans cannot do, such as quick calculations, management of enormous amounts of data, and very fast transfers of data. From this analogy, we believe that the scalability of automatic observation will provide us with information that we could not otherwise get from our own analysis. Our findings in this study, as well as those of similar recent studies such as the works of Velde et al. (1997), Sumi et al. (2003), Bono et al. (2003), and Choudhury & Pentland (2003), reveal this high potential of the new interdisciplinary approach of observing humans' interaction automatically.

5.3 Limitations

Regarding our hypothesis that “the presence of the interactive robots induces children’s friendly group behavior,” we found a higher frequency of accompanying a friend, but we could not determine whether this rate is relatively higher than the other situations. That is, we have not yet compared the frequency of accompanying behavior of a friend in different situations such as in a classroom without the robot, a corridor, somewhere with the presence of other robots, and so forth. The current findings are sufficient for the friendship estimation from the interactive robots; however, such a comparison is necessary for considering a more general estimation model.

Regarding the estimation model for friendship, we only used the accompanying behavior, not other possible elements such as facial expression and distance, whose effectiveness has not yet been verified. The psychological literature has reported that proximity is only a good predictor of those whom children say are their friends at an early age (under 6), and that other things become more predictive later (like common interests) (Rubin et al., 1999). These findings imply that use of other elements will be essential to improve estimation performance. We only estimated “friendship” (actually, the names of friends each child mentioned) as dyadic relationships; however, we have not yet attempted to analyze social networks. Meanwhile, the group behavior of humans is related to not only peer relationships but also to social networks. For example, a child might be accompanied by his/her friend as well as another child who is a friend of his/her friend, but he/she does not think of this third child as a friend.

The generality of the findings is still limited. This study focused on sixth graders in a Japanese elementary school, in the presence of our developed interactive humanoid. There remain many variables that might affect these phenomena, such as social environment, types of robots, age of participants, and the cultural differences among countries. Solving these generality problems with more rigorous experiments will be included in our future work.

5.4 Conclusions

We proposed a friend estimation model for a social robot that interacts with humans and verified the fundamental component of the model through a field experiment. In the field experiment, two identical interactive humanoid robots were placed in an elementary school for two weeks, where children freely interacted with the robots during recess. These interactive humanoid robots identified individual children by using a wireless tag system, which is used to record the time of individual and friend-related interaction as well as to promote interaction by such actions as calling a child's name. The experimental results suggest that most children were accompanied by one or more friends (72% of the total interacting time) and that the robot estimated a portion of their friendly relationships, in particular among children who often appeared around the robot, whereas it showed moderate performance for the others (for example, 5% of all relationships with an 80% accuracy). We believe that these early findings will encourage further research into the social skills of interaction-oriented robots as well as into sensing technology for autonomous observation of inter-personal and human-robot interaction.

Acknowledgments

We wish to thank the teachers and students at the elementary school for their agreeable participation and helpful suggestions. We would also like to thank four anonymous reviewers for their great efforts in patiently and constructively advising us to improve the paper. We also thank Takayuki Hirano and Daniel Eaton, who helped with this field trial in the elementary school. This research was supported by the National Institute of Information and Communications of Japan.

References

Asada, M., MacDorman, K. F., Ishiguro, H., and Kuniyoshi Y. (2001). Cognitive Developmental Robotics as a new Paradigm for the Design of Humanoid Robots. *Robotics and Autonomous Systems,* Vol. 37, 185-193.

Asher, S. R. and Hymel, S. (1981). Children's social competence in peer relations: Sociometric and behavioral assessment. In J. D. Wine & M. D. Smye (Eds.), *Social Competence* (pp. 125-157), New York: Guilford.

Bono, M., Suzuki, N., and Katagiri, Y. (2003). An analysis of non-verbal cues for turn-taking through observation of speaker behaviors, *Proceedings of the Joint International Conference on Cognitive Science (ICCS/ASCS-2003),* July 13-17, The University of New South Wales, Australia, (CD-ROM).

Breazeal, C. and Scassellati, B. (1999). A context-dependent attention system for a social robot, *Proceedings of the International Joint Conference on Artificial Intelligence,* July 31 - August 6, Stockholm, Sweden, Morgan Kaufmann Publishers, pp. 1146-1151.

Breazeal, C. (2003). Towards sociable robots, *Robotics and Autonomous Systems*, **42**(3-4), pp. 167-175.

Brooks, R. A. (1986). A Robust Layered Control System for a Mobile Robot, *IEEE Journal of Robotics and Automation,* 2(1): 14-23.

Burgard, W., Cremers, A. B., Fox, D., Hahnel, D., Lakemeyer, G., Schulz, D., Steiner, W., and Thrun, S. (1998). The interactive museum tour-guide robot, *Proceedings of the National Conference on Artificial Intelligence,* 26–30 July, Madison, Wisconsin, AAAI, pp. 11-18.

Choudhury, T. and Pentland, A (2003). Sensing and Modeling Human Networks using the Sociometer, *Proceedings of the 7th International Symposium on Wearable Computers (ISWC 2003),* 21-23 October, White Plains, NY, IEEE Computer Society, pp. 216-222.

Clark, H. H. (1996). *Using Language*, Cambridge University Press.

Coie, J. D. and Kupersmidt, J. B. (1983). A behavioral analysis of emerging social status in boys' groups. *Child Development*, Vol. 54, 1400-1416.

Cowley, S. J. & MacDorman, K. F. (1995). Simulating conversations: The communion game. AI & Society, 9(3), 116-137.

Dautenhahn, K. and Werry, I. (2002). A quantitative technique for analyzing robot-human interactions, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* September 30 - October 4, Lausanne, Switzerland, IEEE, pp 1132-1138.

Fujita, M. (2001). AIBO; towards the era of digital creatures, *International Journal of Robotics Research, 20,* 781-794.

Gottman, J. M. and Parkhurst, J. (1980). A developmental theory of friendship and acquaintanceship processes. In W. A. Collins (Ed.), *Minnesota symposia on child psychology,* (Vol. 13, pp. 197-253). Hillsdale, NJ: Lawrence Erlbaum.

Hall, E. (1966). *The Hidden Dimension*, Anchor Books/Doubleday.

Hartup, W.W. (1996). The company they keep: Friendships and their developmental significance. *Child Development,* 67, 1-13.

Hagita, N., Kogure, K., Mase, K., and Sumi, Y. (2003). Collaborative capturing of experiences with ubiquitous sensors and communication robots, *Proceedings of the International Conference on Robotics and Automation (ICRA '03),* 14-19 September, Taipei, Taiwan, IEEE, pp. 4166-4171.

Ishiguro, H., Ono, T., Imai, M., and Kanda, T. (2003). Development of an interactive humanoid robot "Robovie" -An interdisciplinary approach, R. A. Jarvis and A. Zelinsky (eds.), *Robotics Research, Springer,* 179-191.

Kanda, T., Ishiguro, H., Imai, M., Ono T., and Mase, K. (2002). A constructive approach for developing interactive humanoid robots, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* September 30 - October 4, Lausanne, Switzerland, IEEE, pp. 1265-1270.

Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H. (2003). Person Identification and Interaction of Social Robots by Using Wireless Tags, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* October 27-31, Las Vegas, Nevada, IEEE, pp. 1657-1664.

Kanda, T., Ishiguro, H., Imai, M., Ono, T., (2004a). Development and Evaluation of Interactive Humanoid Robots, *Proceedings of the IEEE*, Vol.92, No.11, pp. 1839-1850.

Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H. (2004b). Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial, *Human Computer Interaction*, Vol. 19, No. 1-2, 61-84.

Kozima, H. and Vatikiotis-Bateson, E. (2001). Communicative criteria for processing time/space-varying information, *Proc. 10th IEEE International Workshop on Robot and Human Communication,* 18 – 21 September, Bordeaux and Paris, France, IEEE, pp. 377-382.

Ladd, G. W., Price, J. M., and Hart, C. H. (1990), Preschooler's behavioral orientations and patterns of peer contact: predictive of peer status? In Asher S. R. & J. D. Coie (Eds). *Peer rejection in childhood,* 90-115. Cambridge University Press.

Levinson, S. C. (1983) *Pragmatics,* Cambridge University Press.

McConnell, S. R. and Odom, S. L. (1986), Sociometrics: Peer-referenced measures and the assessment of social competence. In P. S. Strain, M. J. Guralnick, & H. M. Walker (Eds), *Children's social behavior: Development, assessment, and modification,* 215-284, Orlando: Academic Press

Moore, C. and Dunham, P. J. eds. (1995). *Joint Attention: Its Origins and Role in Development,* Lawrence Erlbaum Associates.

NEC Corporation. (2002). Personal robot PaPeRo. http://www.incx.nec.co.jp/robot/robotcenter_e.html

Newcomb, A. F. & Bagwel, C. (1995). Children's friendship relations: A meta-analytic review, *Psychological Bulletin,* 117, 306-347

Okuno, H. G., Nakadai, K., and Kitano, H. (2002). Realizing audio-visually triggered ELIZA-like non-verbal behaviors, *PRICAI2002, LNAI 2417, Lecture Notes in Artificial Intelligence,* Springer-Verlag, 552-562.

Robins, B., Dickerson, P., Stribling P., and Dautenhahn K. (2004). Robot-mediated joint attention in children with autism: A case study in robot-human interaction, *Interaction Studies, 5:2,* pp. 161-198

Rubin, K. H., Coplan, R. J., Nelson, L. J., Cheah, C. S. L., Legace-Seguin, D. G. (1999). Peer Relationships in Childhood, *Development psychology: An advanced textbook,* pp. 451-502.

Sakagami, Y., Watanabe, R., Aoyama, C., Matsunaga, S., Higaki, N., and Fujimura, K. (2002). The intelligent ASIMO: System overview and integration, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* September 30 - October 4, Lausanne, Switzerland, IEEE, pp 2478-2483.

Salter, T., Dautenhahn, K., and Boekhorst, R. (2004). Robots moving out of the laboratory – Detecting Interaction Levels and Human Contact in Noisy School Environments, *Proceedings of the IEEE International Workshop on Robot and Human Communication*, 20-22 September, Kurashiki, Japan, IEEE, pp. 563-568,

Scassellati, B. (2000). Investigating Models of Social Development Using a Humanoid Robot, *Biorobotics,* MIT Press.

Severin, E. O. (2003). *Robot Companions: Mentorbots and Beyond,* McGraw-Hill.

Shibata, T. and Tanie, K. (2001). Physical and affective interaction between human and mental commit robot, *Proeedings of the IEEE International Conference on Robotics and Automation,* May 21-26, Seoul, Korea, IEEE, pp. 2572-2577.

Sumi, Y., Matsuguchi, T., Ito, S., Fels, S., and Mase, K. (2003). Collaborative capturing of interactions by multiple sensors, *Proceedings of the International Conference on Ubiquitous Computing (Ubicomp 2003)*, October 12-15, Seattle, Washington, Springer, pp. 193-194.

Van de Velde, W. (1997). Co-Habited Mixed Realities. *Proceedings of the IJCAI'97 workshop on Social Interaction and Communityware*.

Waldrop, M. F. and Halverson, C. F. (1975). Intensive and extensive peer behavior: Longitudinal and cross-sectional analyses. *Child Development,* Vol. 46, 19-26.

Authors address

Takayuki Kanda
Intelligent robotics and communication labs.
Advanced telecommunication research (ATR)
2-2-2 Hikaridai, Seikacho, Sorakugun, Kyoto
Japan

Hiroshi Ishiguro
Dept. Adaptive Machine Systems
Osaka University
2-1 Yamadagaoka, Suita, Osaka
Japan

About the authors

Takayuki Kanda is a senior researcher in the Intelligent Robotics and Communication Laboratories at ATR (Advanced Telecommunications Research Institute), Kyoto, Japan. His current research interests include intelligent robotics, human-robot interaction, and sociality of interactive robots.

Hiroshi Ishiguro is a Professor of Adaptive Machine Systems in the School of Engineering at Osaka University, Osaka, Japan, and a visiting group leader in the Intelligent Robotics and Communication Laboratories at ATR, Kyoto, Japan. His current research interests include computer vision and intelligent robots.

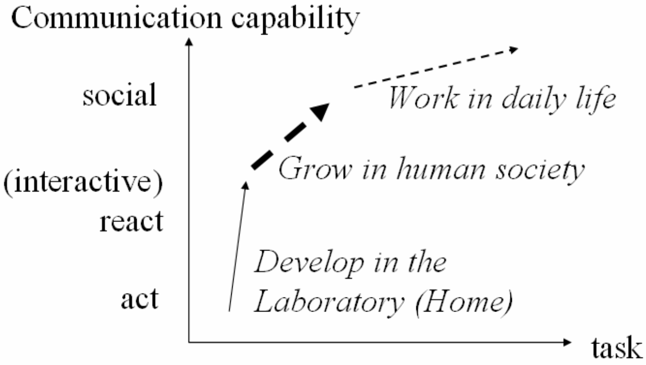


Figure 1: Developmental approach for social ability of interaction-oriented robot

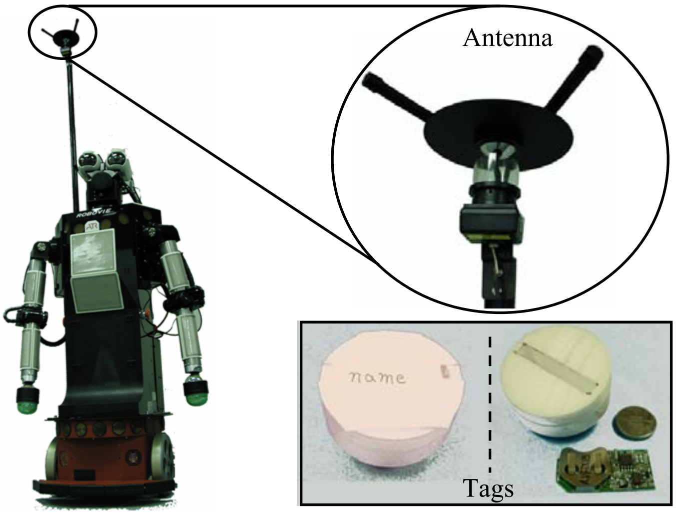


Figure 2: Robovie (left) and Wireless tag

Robovie is an interactive humanoid robot that autonomously speaks, makes gestures, and moves around. With its antenna and tags, it is able to identify individuals.

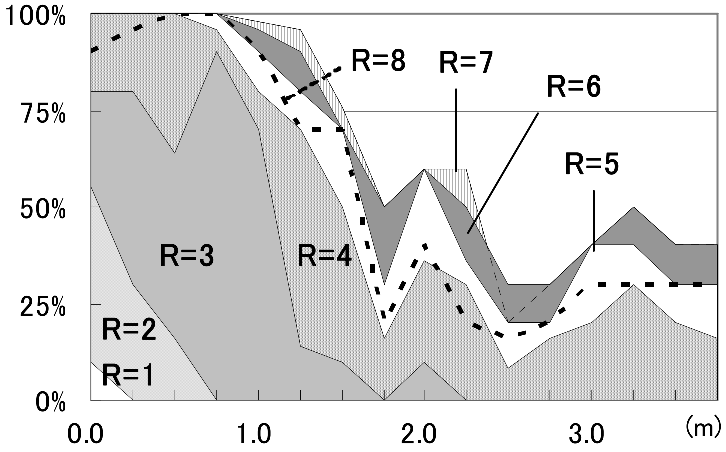


Figure 3: Read distance with different attenuation

In the graph, R indicates the attenuation parameter, the vertical axis corresponds to the rate that the robot found tags, and the horizontal is the distance from the robot.

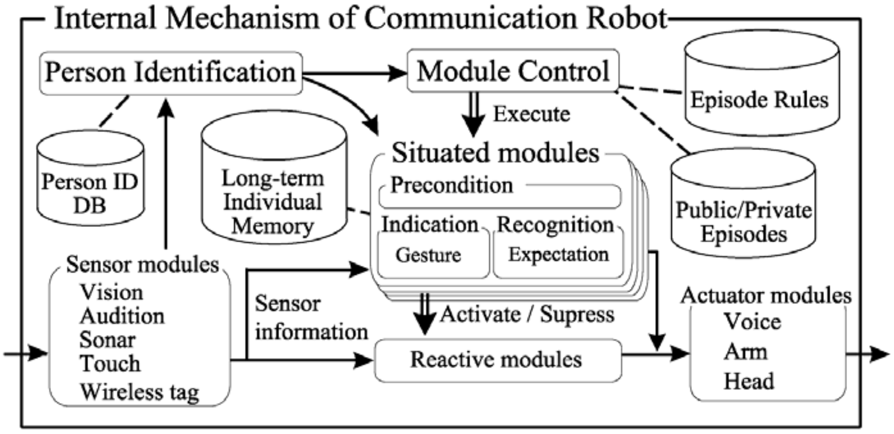


Figure 4: Software architecture of interactive humanoid robot *Robovie*

Situated modules are the essential components to perform interactive behaviors by using sensor information and actuators. The robot selects a suitable situated module for the current interaction situation with person identification, episode, and episode rules.

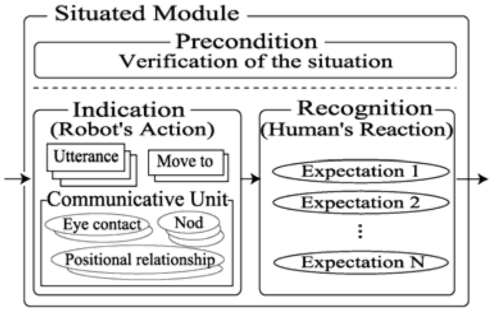
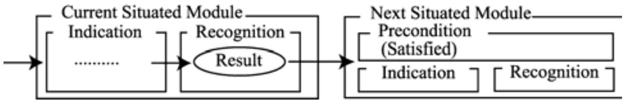
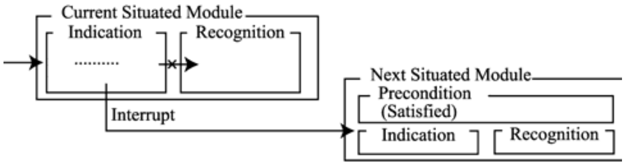


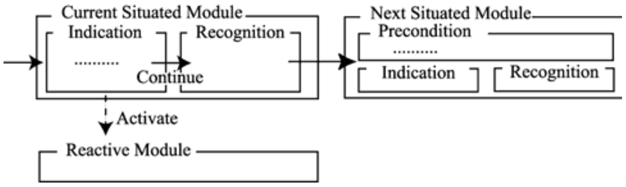
Figure 5: Situated module



(a) Sequential transition (human reacts to the robot)



(b) Reactive transition (the robot reacts to human's interruption)



(c) Activation of Reactive Modules
(robot reacts, but no reactive transition occurs for the reaction)

Figure 6: Transition of situated modules

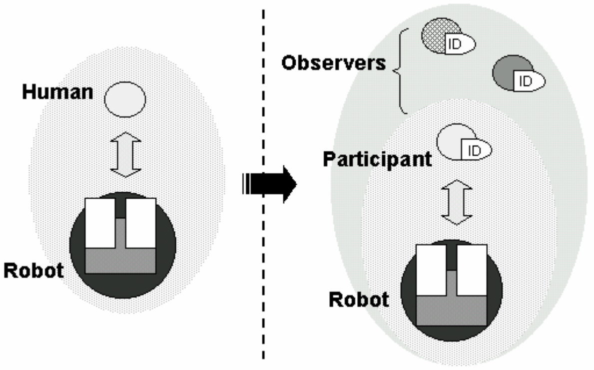


Figure 7: Multi-person communication model

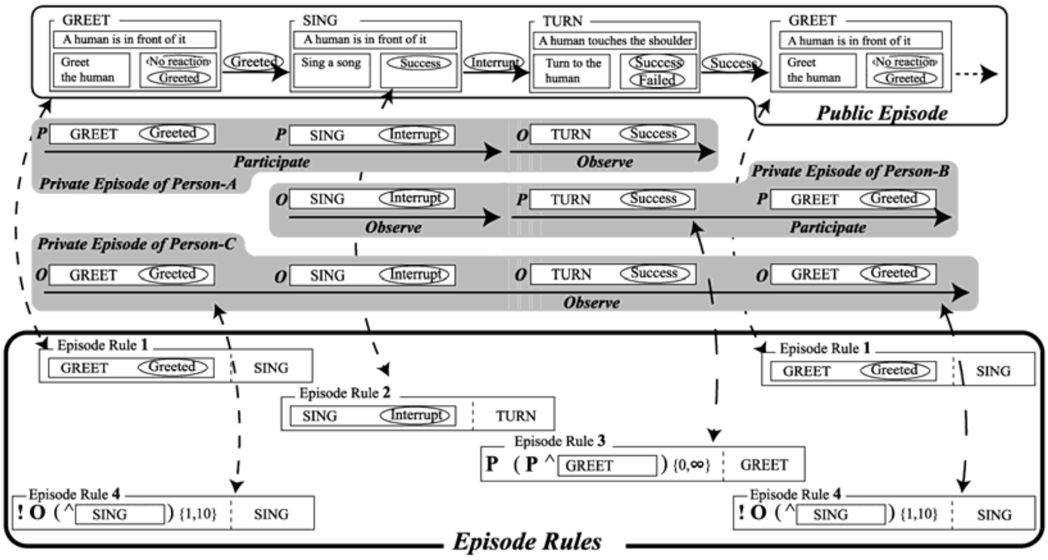


Figure 8: Illustrated example of episodes and episode rules

Episode rules refer to *private episodes* of *participants* and *observers* to adaptively interact with them as well as *public episodes* to realize consistent behavior.

(a) shake hands (b) hug (c) paper-scissors-rock (d) exercise

Figure 9: Interactive behaviors

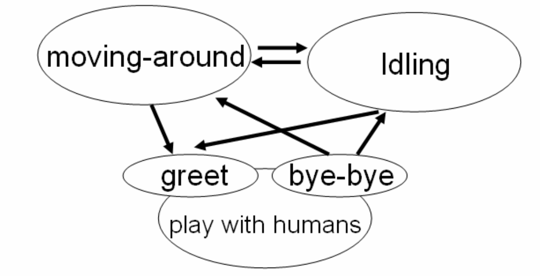


Figure 10: Meta-level transition among interactive behaviors

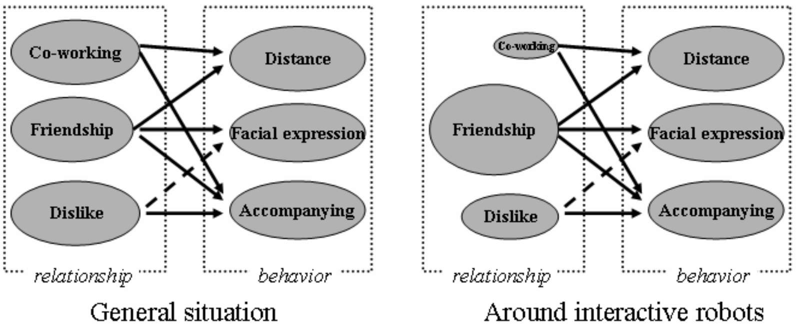


Figure 11: Relations between social relationships and group behavior

Figure 12: Scenes of friends' accompanying behavior in front of an interactive humanoid robot

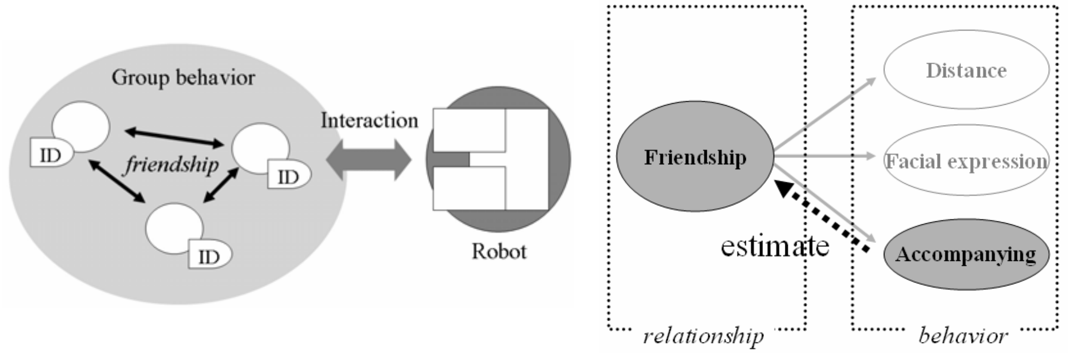


Figure 13: Current estimation model for friendship

Our hypothesis: Robot identifies several people in front of it simultaneously; as a result, it is able to estimate friendship among them, because the robot's interactive behaviors cause mostly friend-accompanying behavior.

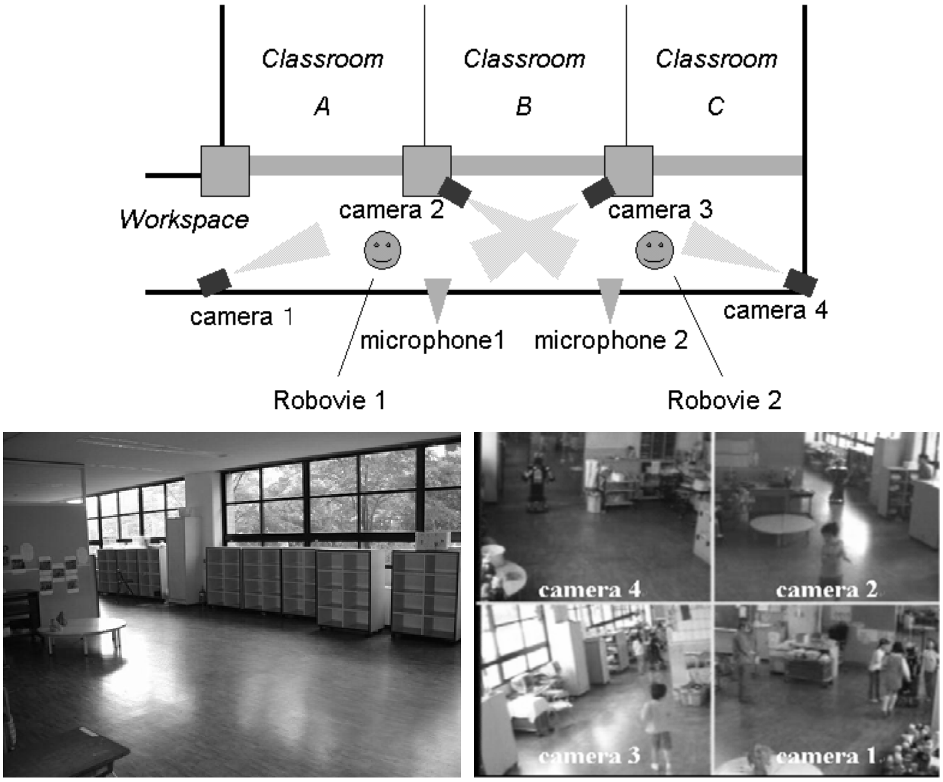


Figure 14: Environment of elementary school

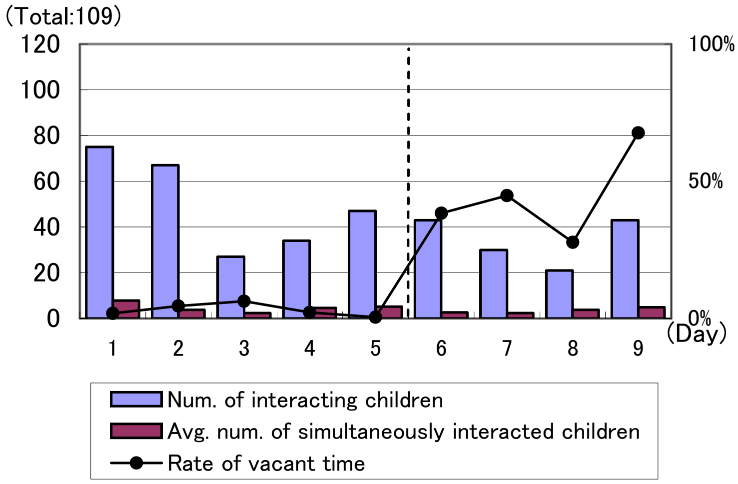


Figure 15: Illustration of Results for the change in children's behaviors at an elementary school

Figure 16: Scene of the experiment at first day

Figure 17: Scene of the experiment after first week

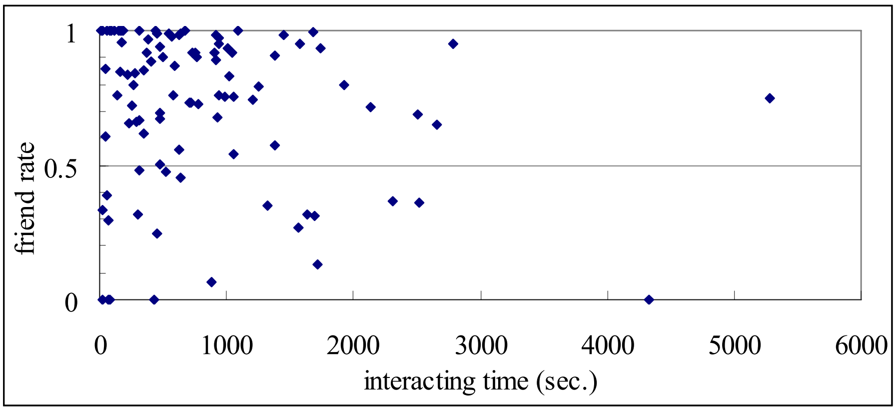


Figure 18: Rate of friend accompanying behavior

Each dot represents a child with a certain interacting time and friend rate

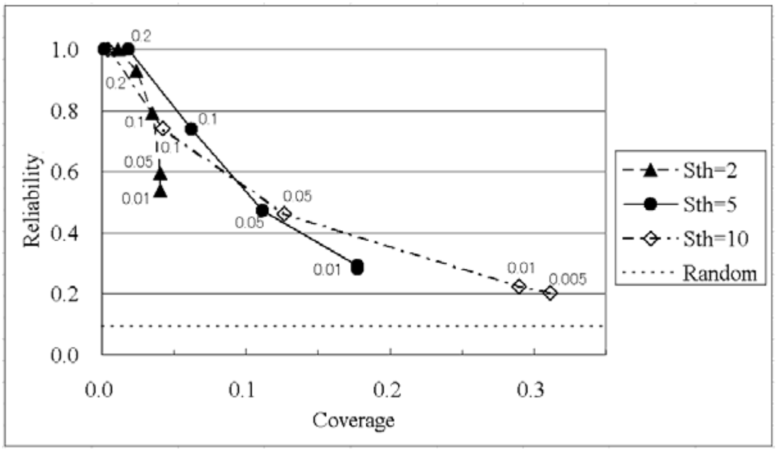


Figure 19: Illustrated estimation results with various parameters

(Each line corresponds with the S_{TH} (2, 5, and 10). Each point of these lines corresponds to each reliability and coverage derived from a certain T_{TH} denoted in Table 4. The numbers in the figure also represent the value of T_{TH}.)

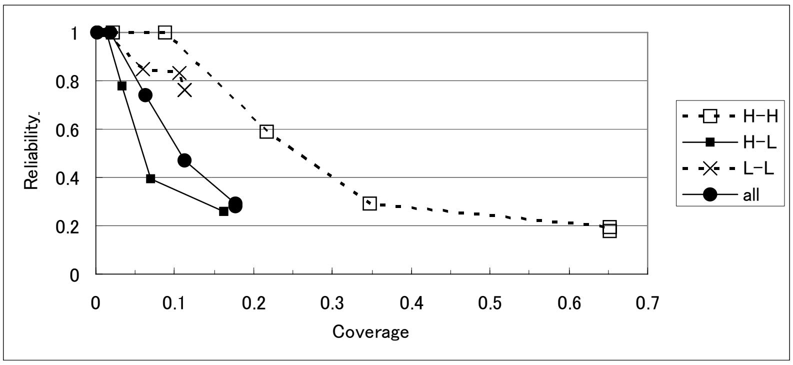


Figure 20: Relationships between friendship estimation and frequency of interaction with robot

(All line corresponds with the S_{TH}=5. Each point of these lines corresponds to a certain T_{TH} in Table 5.)