Designing Social Greetings and Proxemics in Human Robot Interaction

Brandon Heenan, Saul Greenberg, Setareh Aghelmanesh
and Ehud Sharlin
Department of Computer Science, University of Calgary
Calgary, Alberta, Canada T2N 1N4
Email: bmheenan.ux@outlook.com *or* saul.greenberg@ucalgary.ca

We operationalize on a robot a subset of social behaviors as described by Hall's proxemics theory and Kendon's observations of greetings. Our hypothesis is that basing robot behaviors on the social science of such human behaviors will make the robot appear to convey social intelligence. Specifically, we track the location and orientation of a Nao humanoid robot relative to a person, and programmed the robot to engage in a distance salutation, approach, close salutation and transition as described by theory. Overall, our design appears effective in simulating social intelligence, especially with respect to eye contact. However, mechanical limits affects the robot's ability to express necessary social nuances, including seemingly fine distinctions such as the robot's slow speed in moving into position, or its inability to direct gaze independent of head position. Our findings suggest that HRI design must consider detailed nuances of how particular expressions of social theory are realized as robotic behaviors.

Keywords: Human robot interaction; proxemics; greetings; social robots; social science

Introduction

As computers' capabilities continue to increase, the field of human robot interaction (HRI) provides the promise of integrating robots into the everyday human environment. In a number of fields, including healthcare, construction, manufacturing, education and public services [Mumm and Mutlu, 2011; Goodrich and Schultz, 2007], the ability of robots to socially integrate into those environments will be key to their acceptance. To approach this social integration, researchers in HRI (as well as popular literature and movies) have generally suggested that the design of robot behavior should be modeled after human behavior. The idea is that, if done well, humans can use their own natural social skills and expectations to recognize robotic behaviors, and ultimately to interact with the robot.

Designing robots based on human behaviors is far from easy. On the input side, a robots' sensing channels are quite different from human senses, where robots cannot yet socially sense the world as richly or as quickly as humans can. It is no easy matter to program robots to read, interpret and respond to the nuances of what they sense. On the output side, robots typically have many mechanical constraints, especially when compared to the musculature control that people

Cite as:
Heenan, B., Greenberg, S., Aghelmanesh, S. and Sharlin, E. Designing Social Greetings and Proxemics in Human Robot Interaction. Technical Report 2013-1044-11, Department of Computer Science, University of Calgary, Calgary, AB, Canada T2N 1N4.

have. This means that robots may not be able to carry out the equivalent human actions. These challenges of sensing social settings correctly and of being able to act within these social situations leads to the primary focus of this paper: how do we articulate the exact social behaviors the robot should be exhibiting, and how can we operationalize the subtleties of those behaviors?

Our overall design approach follows a four-fold method. First, we turn to social science and particular social theories or observations made of human behavior as our intellectual foundation driving the design of robotic behaviors. Second, we transform and articulate those theories or observations as a set of behaviors that can be realized on a robot. Third, we operationalize those behaviors on a particular robot with particular capabilities. Fourth, we observe and reflect on both the successes and failures of our approach.

In this paper, we are particularly interested in what we consider the very first steps of human robot interaction: how humans and robots greet one another. The problem is that robots do not currently convey behaviors that allow them to seamlessly initiate interactions with humans. This is due in part to the complex and subconscious rules that humans believe must be followed in order for this initial interaction with a social player to be natural and appropriate. We believe that such greetings are fundamental to the acceptance of robots as social players among people.

As we will see, human greetings involve nuances in proxemics [Hall, 1966], and particular body language depending on where one is in the greeting process [Kendon, 1990]. Properly enacted greetings with respect to nuances in proxemics and body language will help robots and people engage in interaction. Conversely, a robot acting inappropriately during the greeting process may cause it to be misinterpreted or ignored, which could jeopardize the interaction.

Our paper is structured along the four-fold method described above. First, we review two social science constructs relevant to greetings: proxemics theory [Hall, 1966] and body language within the greeting process [Kendon, 1990]. Because we are not the first to consider proxemics and greetings in HRI, we summarize related work in HRI. Second, we describe how we transform and articulate those social science constructs as an abstract model, which we realize as a state machine. Third, we show how we operationalized and implemented these behaviors on the Nao robot. Finally, we reflect on our research, where we consider what worked and what did not, where we pay special attention to the nuances important in designing socially accptable human-robot interfaces

The Social Science of Greetings

Our work is based on three main areas of previous research. The first area is proxemics, the study of spacing and distancing in humans, as pioneered most notably by Hall [1966]. Proxemics has been influenced by earlier work (e.g., [Summer, 1959]), and continued by many others since then (e.g., [Summer, 1969; Altman, 1975]). The second area is human greetings, specifically the work done by Kendon [1990] in describing the subconscious behaviors observed in typical greetings. The third is human-robot interaction, where some work has already been done in describing distancing between humans and robots, though very little has been applied to the design of greeting interactions.

Proxemics

The work by Hall [1966]is generally considered to be the seminal account, where he described the basic theory of proxemics. Hall, who was a cultural anthropologist, studied similarities and differences of interpersonal distancing in various cultures. At its simplest, Hall's thesis is that people equate social distance with physical distance. The caveat is that the way people do this can vary somewhat between cultures, and involve many nuances.

According to Hall, humans tend to exhibit different behaviors towards each other in accordance with four levels of "closeness." Hall labeled these four levels as intimate, personal, social and public space. The metrics provided below are typical of western cultures [Hall 1966].

- *Intimate Space* exists when people are 0 to 0.45m apart. As the name suggests, this zone tends to be reserved for people with an intimate relationship, e.g., very close friends, lovers, and so on. People within this space can sense the warmth of skin and the smell of the other individual. Interactions in this space tend to be physical, and vocalizations are minimal.
- *Personal space* exists from 0.45m to 1.2m. This zone tends to be used by people in conversation who know each other well and/or who are comfortable with each other. In the personal space, humans use a normal voice level, are able to clearly see another person's face in great detail, and tend to keep the view of another person's hands in their peripheral vision.
- *Social space* exists from 1.2m to 3.6m. This space is used for impersonal business. The closer end is used between people who know each other (e.g., acquaintances) and the far end is for more formal situations. In this space, many people shift their gaze back and forth from eye to eye when interacting. In the far boundary of the social space, people can comfortably work independently without the social obligation to interact.
- *Public space* exists beyond 3.6m. Sustained interaction is mainly in the context of presentations and public figures, i.e., the way a presenter spaces him or herself away from the audience. Humans observably change their speech patterns at this level. This distance is also used by people to space themselves away from others when they do not wish to interact with those who are nearby.

Of course, there are many other factors in Hall's theory, and others have since contributed to the understanding of proxemics. For example, for each level above, Hall distinguishes a near zone and a far zone. Furthermore, Hall described how fixed features (e.g., boundaries such as doorways) and semi-fixed features (e.g., the positioning of furniture) can affect how people perceive social distance. People's body orientation matters: facing towards one another, kitty-corner, side by side, or away from each other effects perceived social distance (e.g., [Sommer, 1959; Kendon 1990]). Personal space serves as a protective function [Aiello, 1987] somewhat akin to territories [Sommer, 1959]. Use of this space is dictated by social rules and norms [Aiello, 1987], where people take umbrage if those rules are broken [Altman 1975]. Marquardt [2013, Chapter 3] provides a good summary of other factors influencing proxemics.

For brevity we will not describe proxemics theory further. As we will see, the above description suffices to influence and begin our work in creating social models of basic robot behaviors.

Human Greetings

Adam Kendon, in his 1990 book *Conducting Interaction*, defined the term "greetings" as:

> *"that unit of social interaction often observed when people come into one another's presence, which includes a distinctive exchange of gestures or utterances in which each person appears to signal to the other, directly and explicitly, that he has been seen." [Kendon, 1990, page 153]*

Kendon stated that greetings and the way people signal one another are vital both to manage the relations between people (e.g., confirmation of friendship, degree of familiarity, belonging, social status), and to serve as a precursor leading to interaction.

Kendon detailed observations of a number of greeting behaviors of humans in a social context, which became his foundation for a model of social greetings [Kendon 1990]. He observed and videoed people as they greeted each other, where he analyzed the videos to identify people's non-verbal behaviors. What follows is a brief description of his team's observations, again oriented towards a western culture. This description is far from exhaustive, but (as with proxemics) suffices to influence our work on basic robot greeting behaviors as described in later sections. In essence, Kendon found that a typical exchange between two individuals who wish to greet each other follows a structure comprising phases; sighting, distance salutation, approach, close salutation, and finally a transition into the interaction. As we will see, there is an implicit element of proxemics in Kendon's findings: specific behaviors are observed at inexact, but predictable distances [Kendon 1990].

Precursor: sighting and decision to greet. Before a greeting can begin, at least one person must sight the other (e.g., a passing glance, by overhearing a voice). That person (or both) must perceive the other person as someone he or she wishes to greet. In addition, that person would evaluate how available the other is to receive a greeting e.g., if the other person looks busy (such as being engaged in a conversation). The decision to greet is also influenced by one's own willingness to interrupt, the importance of the expected interaction, and so on. Based on these and other factors, the person may then decide to initiate the greeting, or wait, or move on.

Distance Salutation. The greeting starts with a distance salutation, after one or both participants sight one another and at least one of them identifies a wish to engage in a greeting. If one participant is not aware of the other's presence, the latter will call attention to himself through vocalizing a name, or a subtle action, like the clearing of one's throat. If this step is not necessary, there is still an observable but tacit action taken by both participants: they orient their bodies towards each other and exchange glances in a subtle acknowledgement that the greeting is desired by both. These greetings typically only take place if the initiating party has a special obligation or right to greet the other. A distance salutation may not necessarily continue to the next greeting phase. For example, two people may quickly acknowledge each other in passing, but not engage in further interaction.

Kendon described several other physical behaviors that people tend to do to signal their distance salutation [Kendon 1990].

- *The wave* is highly varied but common in distance salutations. In all cases, the hand is raised and the palm open and oriented towards the person being greeted. How much the hand is raised, and whether it is "wagged" varies according to the distance between the parties. It may also be used to communicate excitement.
- *The head toss* occurs when the head is tilted back rapidly, and then brought forward again. It is usually accompanied by a vocalization, such as "hi."
- *The head lower* is also common, whereby the head is tilted downward, held briefly in that position, and then raised again. This is typically combined with a lesser version of a wave, in which the arm is raised slightly but not shaken side to side.
- *The nod* is similar to the head lower, but the head is immediately raised again after being lowered. It is usually observed in greetings in passing, and is not followed by an approach or close greeting.
- *The head dip* by one of the participants often follows one of the above behaviors, where a person lowers their head (i.e., looks downwards). Kendon hypothesized that the head dip marked a shift of attention, i.e., when the person was moving into the next phase of the greeting.
- *Smiling,* which may or may not continue to the next phase.

Approach. Assuming the greeting is not simply a terminal distance salutation, the two parties close the distance between themselves. In proxemics terms, they are moving towards one of the zones most appropriate to their interaction (e.g., public, personal or intimate). Kendon notes that, in his terms, "how far one goes out of one's way" as they move towards one another (perhaps unequally) has communicative significance depending on matters such as environmental factors, status, and context. During the approach, a number of subtle, but important behaviors are observed.

- *Changes in facial orientation.* While people tend to look towards one another during the distance salutation to signal that a greeting is desired, they tend to look away during the actual approach. They may also look sharply away just prior to the next phase. Kendon hypothesizes that looking away is done to increase one's behavioral distance from the other person.

- *Body cross.* People may draw one or both arms across the body.
- *Grooming.* People may adjust hair, clothing or their accessories in an act of self-grooming.

Final approach. People exhibit another set of behaviors as they move increasingly near one another (~3 meters or less). While people during the approach normally look away from one another, they will look towards each other again during the final approach, especially as they transition to the close salutation. Other behaviors may include:

- *Palm presentation.* People commonly orient their palms towards those they are greeting in an "open hand" gesture. This appears not to be formalized or intentional, but is a none-the-less observable behavior.
- *Smiling.* If a person is not yet smiling during the approach, he or she will typically smile during the final approach.
- *Head set.* People alter the way they hold their heads, although the head posture ranges considerably. Examples include the erect head, head tilted forward or back, and head cocked to the side.

Close Salutation. This greeting phase is the most formalized, generally occurring once the approach reaches 1.6 meters or less. At this point, a broad number of salutations may occur.

- *The non-contact close salutation* is one example. In this case, participants halt facing one another and exchange verbal greetings, but no distinctive non-verbal cues are observed. However, this phase is still distinct and observable; people look away sharply during the last part of the approach and move to a conversational stance after a non-contact close salutation.
- *Handshakes* vary in length and intensity, influenced significantly by the sex of the people and formality of the occasion. They are very common in male-male greetings but uncommon in female-female greetings in social contexts.
- *Embraces* are also observed in human greetings, although this depends on the relationship between the two parties.
- Other close salutations exist. Many are quite culturally dependent, for example, bowing, cheek to cheek kisses, etc.

While the above salutations show variety, commonalities exist between all of them. First, while people face one another directly during the final approach, they usually do not maintain this orientation once the close salutation is complete. Second, people fine-tune their relative body positions, albeit in a variety of ways. For example, people frequently move a step back, standing at right angles to each other once they engage in conversation. They then proceed to one or more actual salutations.

Transition. As discussed in proxemics, once the close greeting has been performed, participants tend to increase or decrease the space between them in a way that matches the nature of their interaction, where they may step back from the intimate zone to the personal or public zone. People then typically adopt a stance not directly facing each other: Kendon pursues patterns of these stances in later chapters of his book dealing with 'f-formations' [Kendon 1990]. People may even move to another location by mutual agreement. It is at this point that Kendon determined the greeting portion of the interaction was complete [Kendon 1990].

Human Robot Interaction

Various researchers in HRI have experimented with applying aspects of social theory to human robot interaction. While some have looked at proxemics, no one (to our knowledge) has considered Kendon's work.

Mumm and Mutlu [2011] analyzed the interpersonal distance humans naturally kept from robots, where they manipulated variables including the robot's gaze behavior (mutual vs. averted gaze) and their robot's likeability (i.e., where the robot's initial greeting message was polite *vs.* rude). In their experiment, they pitted four proposed models of interpersonal distance against each other and analyzed which model best predicted human's natural behavior towards such robots. They concluded that people who disliked the experimental robot naturally increased their distance to it when it was looking directly at them, but participants who liked the robot did not modify their interpersonal distance in response to the robot's gaze. They effectively showed that, like Hall's proxemics regarding humans, there is a correlation between social distance and physical distance when robots are involved. That is, a human reacts to a robot by adjusting one's inter-personal distance from it, thereby treating the robot as a social being.

Satake et al [2009] designed an interaction for mobile robots to approach humans. Their first approach of simply taking the shortest distance to the nearest person and attempting a verbal greeting did not have a high success rate. They showed that approaching the human from the front and selecting targets carefully significantly increased the chances of successfully starting a conversation with a human. One of the common failure classes they observed was a person showing interest in the robot and "testing" it for a reaction but not getting the response they expected. Usually, this would cause them to cease the interaction. In fact, in their experiment, 49% of people in a public setting were either unaware of the robot, unsure about whether they could interact with it, or intentionally rejected it. Their work showed that humans responded differently to different robot behaviors during a greeting.

Takayama and Pantofaru [2009] empirically established the interpersonal distances subjects were comfortable with when approaching and when being approached by robots. They found that for a robot approximately 1.35 m high, the average interpersonal distance they observed ranged from 0.4m to 0.6m, although participants were still comfortable with minimum distances ranging from 0.2m to 0.35m. Consistent with the researchers' hypothesis, whether the robot held a mutual gaze with the participant had an effect on distancing, but unexpectedly, women maintained a larger distance from the robot when it was looking towards their face, while men decreased their distance in the same situation. Walters et al [2009] performed a similar experiment, with comparable findings, concluding that participants maintained a mean distance of 0.57m from the robot, but that distance went as high as 0.71m if there was uncertainty or perceived uncertainty. Both these works showed that people were comfortable with robots in their personal space and even within what Hall [1966] described as intimate space, although that happened with robots that were much smaller than the participants.

Mead et. al. (2011) reviews a variety of metrics (such as human poses) commonly used in social science to analyze proxemics behaviors. Their goal is to try to automate tracking of these methods of people and robot moving in a space. To test their algorithms, they track people using a tracking system, where they were able to detect particular metrics that would signify proxemics behaviors. To our knowledge, they have not yet applied these to an autonomous robot although they have suggested various approaches as part of their future work.

Saulnier, Sharlin and Greenberg [2011] were interested in how people perceived a robot's attempt to attract their attention and interrupt a conversation in progress. Using a Wizard of Oz methodology, they crafted robotic interruption behaviors – from benign to aggressive – by manipulating how the robot exhibited various physical nonverbal cues to initiate robot-human interruption. These included: (a) speed of motion, (b) gaze, (c) head movement, (d) rotation and

(e) proximity to the person (including crossing a doorway boundary into the participant's room). Their results not only showed that people were able to interpret robots as social beings during their interruption attempts, but that they also interpreted which of these basic physical behaviors conveyed the most information regarding its sense of interruption urgency.

These are not the only works in Human-Robot Proxemics (e.g., see references in Walters et al, [2009]). What all have in common is confirming that humans are able to interact with robots that exhibit various social cues, including proxemics and select greeting cues. Our work takes this to the next stage, where we abstract particular greeting behaviors as a model, which is in turn operationalized on an autonomous robot that can sense human distance and orientation towards itself.

From Social Science to an Abstract Design Model: Social Constructs as a State Machine

While it is one thing to describe behaviors as observed by social scientists and as characterized by social science theories, it is quite another to translate those behaviors into a model usable by technologists in the design of human robot interaction.

Our approach to doing this was based on the following. First, we considered that it was not currently infeasible – nor desirable – to create a high fidelity literal translation of all the proxemics and greeting behaviors described in the literature (which includes more than what has been summarized in the review section). We knew that such behaviors – even if they could be translated – could not be applied to the design of a robot. For example, robots do not yet have the ability to sense, track and correctly read the nuances of the other person's behavior (e.g., their facial expression and subtle body language). In addition, robots do not yet have the ability to apply such knowledge in a manner that attends to the context and history of the social interaction we wish to model; this remains a difficult problem in AI. Furthermore, if we tried to be literal, we suspect that we would hit the uncanny valley problem [Mori, 1970], where even small deviations from a near-perfect caricature would exhibit a sense of strangeness.

Consequently, we decided instead to consider the robot as a caricature, where – like a cartoon character – it would exhibit only rudimentary behaviors in a simple manner. In this approach, we would stress mostly a few primary behaviors that seemed socially essential. These behaviors would serve as a first order approximation for the design of a robot capable of autonomous human greetings.

For simplicity's sake, we decided to model the flow of these behaviors as an abstract state machine. We did this because both proxemics and greetings appear to follow a progression through a series of states. For example, proxemic distances decrease from far to near during the greeting process, and each phase has its own particular behaviors associated with it. While we recognize that a state machine has limitations, it serves as a reasonable starting point for our own design explorations.

Figure 1 illustrates our translation of these behaviors as a state diagram. Robotic behavioral states are collected on the right side in the large light grey box. Behavioral activities of a robot moving between proxemic zones are shown as rose-colored boxes. All other behavioral changes in a robot's body language behaviors (excepting moving from one location to another) are the light blue boxes. Robotic sensing of a human's orientation, distance or gaze are shown in green diamonds, where they are also collected on the left side in the large light yellow area. As seen in Figure 1, the robot makes yes/no decisions based on what it senses in these states.

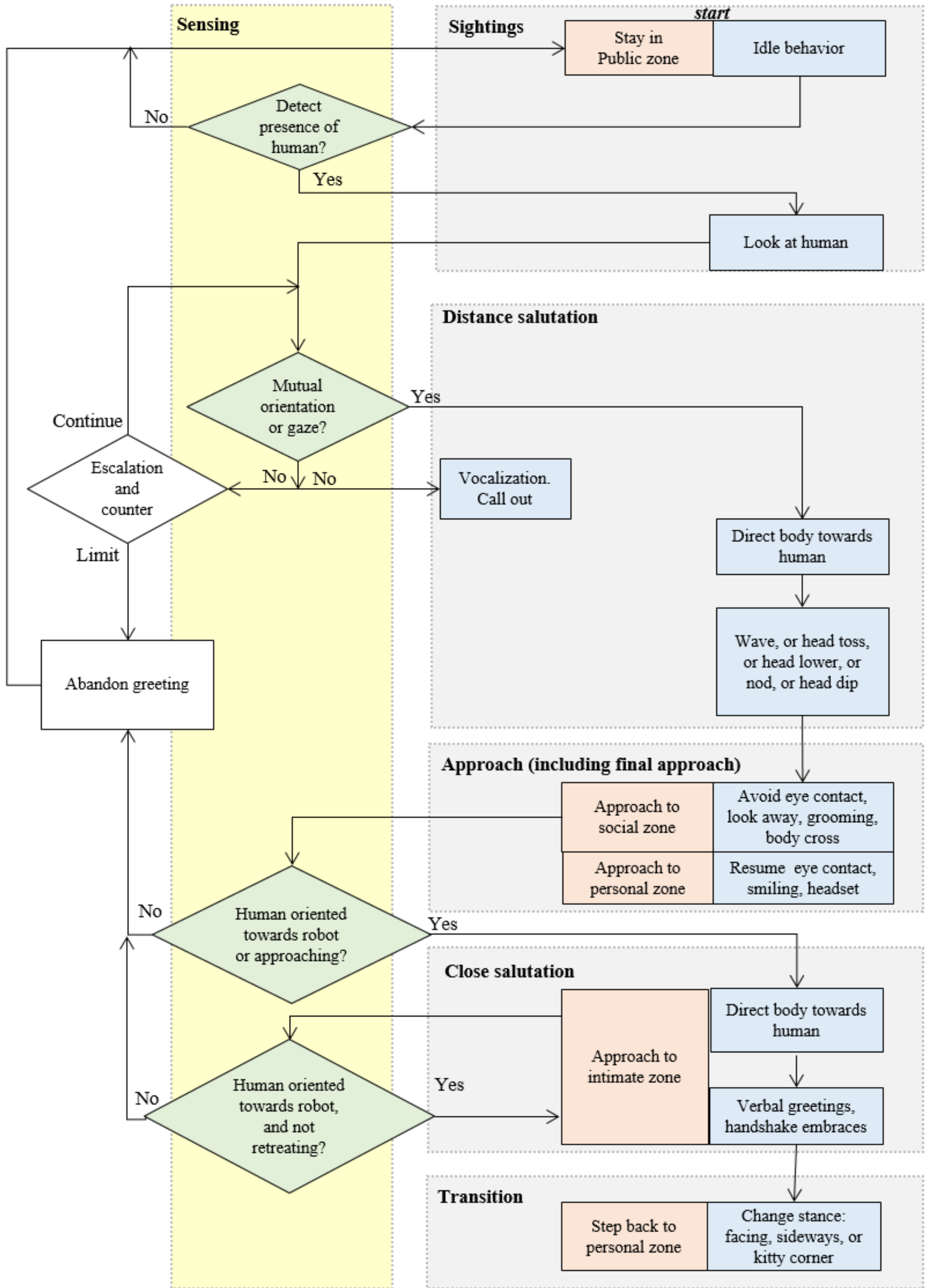


Figure 1. The Greeting Model as a State Diagram

To illustrate, consider what happens during a successful greeting, starting at the top right of Figure 1. During the Sighting phase, the robot is located in a public zone, manifesting idle behavior (i.e., some sort of observable motion activity indicative of its aliveness). When the robot detects the presence of a human, it turns to look at that person. It then transitions to the Distance Salutation phase if it sense that that person has responded by orienting his body towards the robot and/or by returning the robot's gaze. The robot then responds by directing its body towards the human, and performing another distance salutation, such as a wave, a head toss, a head lower, or a head dip. The robot then enters the Approach phase, where it moves to that person's social zone. The robot adapts to a person's approach, where it mediates its distance from that person by sensing that person's location. During this movement, the robot avoids eye contact by looking away, and performs other greeting behaviors appropriate for this stage such as grooming or body cross.

If all goes well, the robot continues into the Final Approach, entering that person's personal zone. The robot looks towards the person in an attempt to re-establish eye contact, and performs other behaviors such as smiling, and headset actions. The robot then enters the Close Salutation phase, where it reorients its body to face the human and – if the person is also oriented towards the robot and not retreating – attempts salutations such as verbal greetings, handshakes, and embraces. Finally, the robot transitions into interaction, where it steps back into the personal zone and changes its stance to one appropriate for what is to happen next.

Part of our abstract greeting model includes decisions on how to manage special cases, such as when the robot should try to attract a person's attention and/or when to abandon the greeting. According to our model, the robot does this largely by sensing the presence, distance, and orientation (possibly including gaze) of the person. For example, during the distance salutation, if the robot detects that the person has not yet return its gaze, the robot will attempt a vocal salutation to attract that person's attention by calling out to that person a certain number of times. If the gaze is still not returned, the robot abandons that particular greeting state and returns back to its idle behavior. In most other phases, the robot continues to sense if the person is approaching and / or maintaining his or her orientation towards it, where it interprets this as a cue to continue the greeting process. However, the robot abandons the greeting if the person turns or moves away from it. If this happens after the robot has already started moving towards the person, the robot will reposition itself away from the person to return to the public zone.

Operationalizing the Greeting Model on the Nao Robot

The greetings model as described in the prior section is an abstract model. We now show how this model was operationalized and implemented on the Nao robot, a small humanoid robot commercialized by Aldebaran Robotics[1].

The Nao Robot

The Nao robot is illustrated in Figure 2**.** It is significantly smaller than an adult human, only 25cm tall. Its key features include a body with 25 degrees of freedom operated by programmable electronic motors and actuators. Its sensors include 2 cameras, 4 microphones, pressure sensors, a sonar range finder, 9 tactile sensors, and others. It also includes a voice synthesizer, various lights, and speakers. As seen by the robot's joints in Figure 2, it can independently move a large number of joints, including its feet,

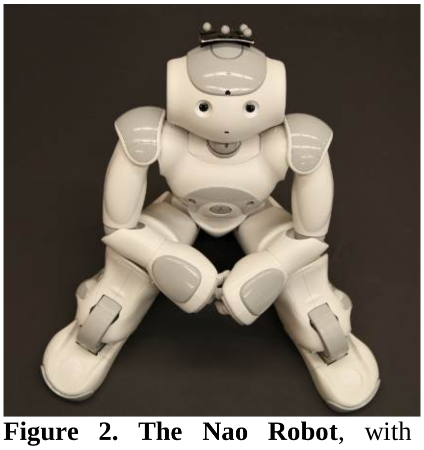


Figure 2. The Nao Robot, with reflective markers on its head

[1] www.aldebaran-robotics.com/

hands, fingers, elbows, shoulder, forearms, head, and neck rotation.

We control the robot through a custom .Net application using the NaoQi API. We preferentially use non-blocking calls in its API, as any behavior must be cancellable at any point if the robot is to react in sync to sensed human actions. To alleviate any delays due to blocking calls, we issue certain commands (such as the command to move in a certain direction at a certain speed) twice per second, rather than telling the robot to move to a specific location at any point.

The Sensing Environment

The greeting model relies on a robot being able to detect the presence, orientation and location (including distance) of a person relative to the robot in real time. We do our raw sensing using a motion tracking system, located in a room equipped with Vicon motion tracking cameras and associated hardware[2]. The Vicon motion tracking system tracks the 3D x, y and z location of specialized reflective passive markers arranged in patterns, as well as yaw, pitch and roll. By having both the robot and the person wear the markers on strategic locations, the Vicon system can locate not only the robot or person, but particular body parts. The robot wears its marker atop its head (as seen in Figure 2), which allows us to track both the location of the robot and its orientation and, because the robot is controlled in code, we know its body orientation relative to its head). The person wears a hat with the markers on it. Because the hat is worn face-forward on the head, it allows us to track the orientation of that person's head.

Using the raw output of the Vicon system involves quite complex coding. Instead, we use the Proximity Toolkit [Marquardt et. al., 2011], which is software specially designed to track various proxemics relationships between entities in an environment. The Proximity Toolkit runs on a layer in between the Vicon system and our application. Using its easy to program API, we track presence, distance and orientation relationships between the robot and the person, where we use a person's head orientation as an estimate of gaze (these metrics are a simple subset of those suggested by Mead et. al. 2011). The metrics provide all the information necessary to implement the sensing requirements of Figure 1.

Using the Proximity toolkit and the NaoQi API, our software tracks the location and orientation of both the human and the robot in a shared three dimensional space – a room – and has the robot respond to its dynamically changing situation. Our software operates a version of the state machine illustrated in Figure 1, checking for sensed conditions to move the program (and thus robot) into a different state depending on the physical relationship between the robot and human. The software operates the robot's head position and therefore its gaze, where it checks the state as well as the currently sensed conditions to determine if the gaze should be directed at the human, slightly away, or at random. Similarly, the software controls the robot's movement to particular locations, where it uses the current state to determine if, when and where the robot should move relative to a person, and what its body orientation should be relative to that person. Finally, the software instructs the robot to enact particular body gestures, such as waves, head nods, and others as indicated in the state diagram.

Design of the Nao's Greeting Interaction

We now describe the details of how the Nao operationalizes the greetings model.[3]

As a pre-cursor, we were quite aware that our Nao robot is small. Because of the difference in height, we adjusted the proxemic distances downwards to conform with the findings of Takayama and Pantofaru [2009]. The adjusted distances were:

- Public space: 1.3 m to maximum in Vicon area
- Social space: 0.75 – 1.3 m
- Personal space: 0.2 – 0.75 m

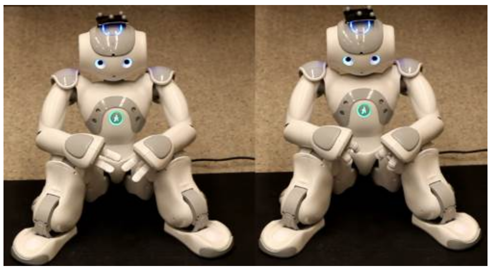
[2] http://www.vicon.com/

[3] We actually implemented several designs that vary aspects of the greetings model described here. For brevity and literary convenience, we synthesize these here as a single design.

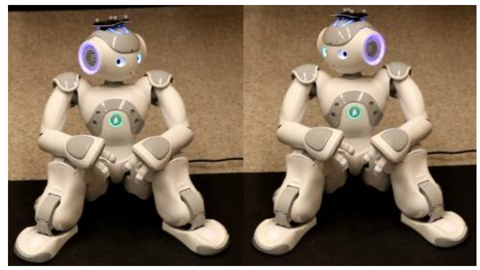
- Intimate space: Not implemented

Sighting. Our first state indicates an idle behavior. In our original design, we implemented this idle state by having the robot do nothing. However, we quickly realized that this was inadequate: people did not initially realize the robot could be active (e.g., if they sighted the robot before the robot sighted them), and were somewhat startled when the robot started moving (see also [Satake et al 2009]).

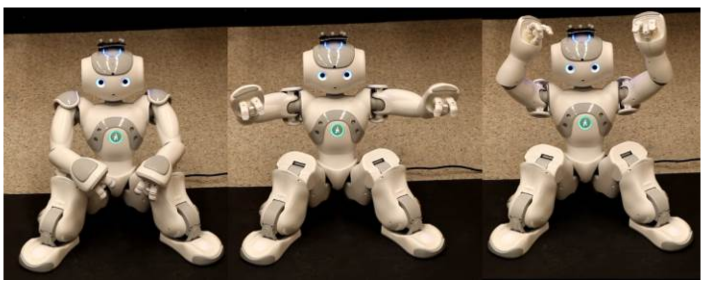
Consequently, we implemented several idle behaviors designed to show the aliveness of the robot, as illustrated in Figure 3.



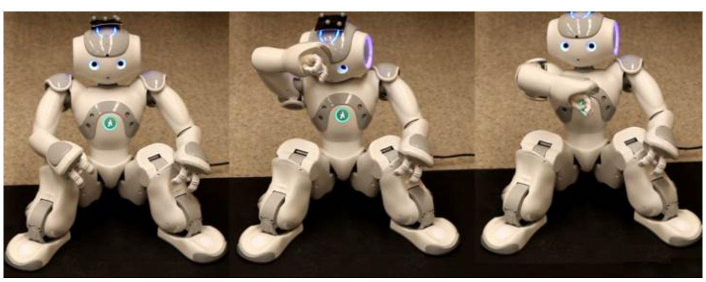
a. Fidgeting: small hand / body movements



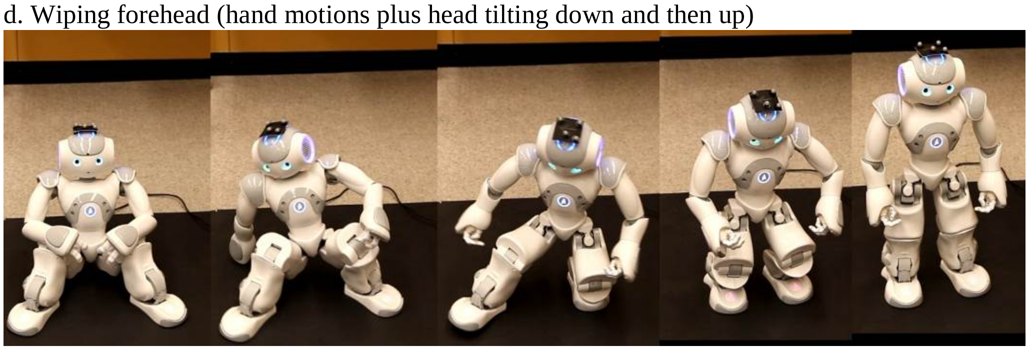
b. Looking around: side-side head movements



c. stretching (hand and upper body moving in a stretching motion)



d. Wiping forehead (hand motions plus head tilting down and then up)



e. Standing

Figure 3. Idle behaviors of the Nao Robot during the Sighting Phase.

We designed these idle behaviors based on the human behaviors when one is left alone. One of the most common human idle actions is fidgeting, where people play with their fingers when they have nothing else to do. In our implementation of fidgeting, illustrated in Figure 3a, the robot looks at its hands and opens and closes them a few times to simulate the fidgeting action. Our next idle behavior is looking around, where people who are left alone may look for someone or something to interact with. In our implementation (Figure 3b), the robot turns its head to left and right to simulate it looking around for something to do. A third action is stretching: most people stretch somewhat when they have not moved for a long time. As seen in Figure 3c, the robot simulates this by raising both its hands and bringing them back down slowly. Our next idle behavior is wiping the forehead, usually associated with the person being tired or bored. These involved continuous behaviors similar to human idle behaviors. Forehead wiping is enacted by the robot in Figure 3d. Other idle behaviors have the robot occasionally looking at random directions, and/or standing up and moving towards a random location (Figure 3e shows the robot as it is standing up). In practice, these behaviors are interleaved together over time, where the robot moves smoothly from one action to the next in an apparently natural manner. This behavior is accomplished using two repeating timers and a random number generator.

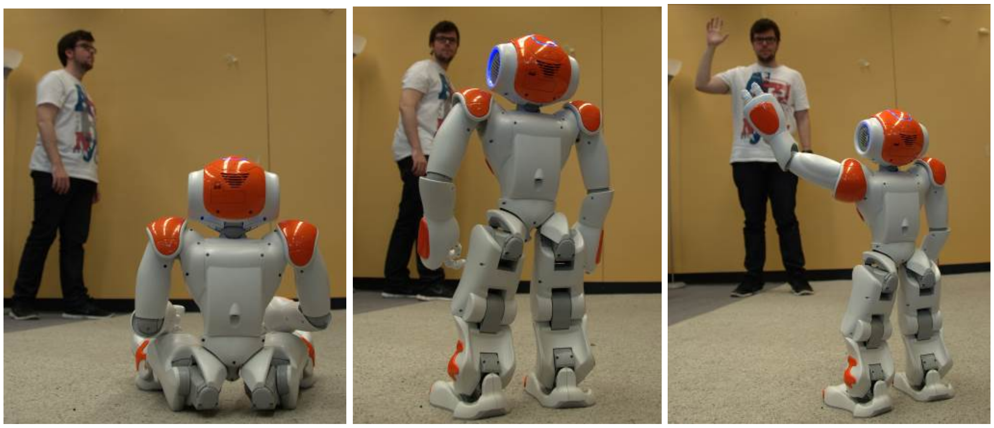
As the robot detects the presence of a person entering the room (shown in Figure 4a), it simulates attempts to make eye contact with him by looking directly at him. To do this, the robot rotates and tilts its head upward so its head is oriented towards the sensed direction and angle of that person's head. As the person moves around the robot's public space, the robot follows him with its gaze.

Distance salutation. Once the user makes eye contact with the robot, it engages in a distance salutation. It does this by sensing if the user's head orientation is within 15 degrees of the robot, which is a reasonable heuristic for assuming mutual gaze. If the robot is not standing (e.g., as in Figure 3a-c), the robot will stand up (using the same operation as in Figure 3d). The robot will then orient its body torso and face (including maintaining its head tilt angle) directly towards the person and the person's head (Figure 4b). If the person remains looking at the robot, the robot will then perform one of the distance salutations. As we will discuss later, the wave (Figure 4c) worked best for this particular robot. The robot performed this by raising its arm, with an open hand directed towards the person, as seen in Figure 4c.

Approach. After the distance salutation, the robot will then move towards the person, passing through their social zone and into their personal zone. As dictated by our model in Figure 1, the robot avoids eye contact during the initial approach. This is done through two calculations. First, the robot angles its head 20 degrees down from the person's facial orientation, as sensed by the Proximity Toolkit. Secondly, it looks 30 degrees off center from the direction of its body, in the direction away from the person. Because the robot's gaze appears to follow directly straight out from the front of its head, this gives the illusion of avoiding eye contact. The intent of this behavior is to stop the robot from seeming like it is "staring" at the person during the approach.

The robot continues to move towards the person. It adjusts its movement based on how it senses where the person is (for example, if the person is moving towards it and at an angle). As the distance between the two reaches the personal space threshold (as calculated on-the-fly by the sensing system tracking both robot and person), the final approach begins. The robot re-orients its head to simulate resuming eye contact (Figure 4d). Our robot cannot smile, so we do not implement that behavior. This head movement into the facing posture also simulates a headset behavior.

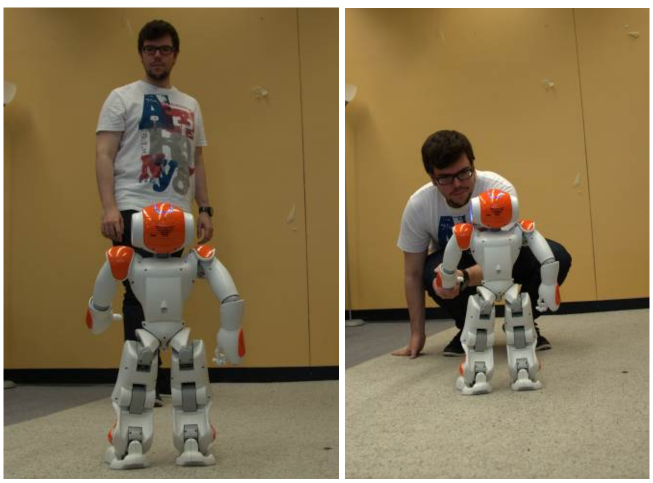
Close Salutation. The robot continues to close the distance, where it will temporarily move into the person's intimate space. The robot then engages in a close salutation, which in this case is the handshake as shown in Figure 4e. The robot does this by moving its hand forward with an open palm held sideways, as an invitation for the other person to shake its hand. As it does this, it performs a vocal close salutation as well, where it uses its voice synthesizer to say "how are you?"



a. sighting: looks at person

b. distance salutation: stands up, faces person

c. distance salutation: waves



d. Approach: moves towards person

e. Close salutation: handshake

Figure 4. The Nao Robot: selected behaviors from sighting to close salutation phase.

Other behaviors associated with a close salutation are maintained; the robot keeps a straight-on body posture, directly facing the human, and maintaining its eye contact via its head orientation and tilt.

Transition. After the close salutation, the robot steps back to the personal space zone relative to the person (not shown), where it still maintains eye contact. At this point, the conversation or purpose of the interaction would take place, but this is beyond the scope of our current research.

Failure cases. If the user does not appear interested in greeting the robot, does not appear to notice the robot, or actively avoids it, the robot's behavior as described in the previous section is modified as dictated by our state diagram in Figure 1. These situations are as follows.

The first failure case occurs when the person initially ignores the robot (Figure 1, second diamond from the top). If the person does not return the robot's eye contact, the robot will vocalize in an attempt to get their attention, where our robot says variations of the word 'hi' (to avoid repetition). If there is no response by the person even after several vocalizations (as detected by the person not looking at the robot), the robot will abandon its greeting attempt. Our robot then expresses 'sadness' through its body language, where it tilts its head and gazes downward briefly before returning to its idle behavior as dictated by the idle state. If at some point the person does re-orient itself to face the robot, the robot will leave its idle state to look at the person, which reactivates the state diagram.

The second failure case occurs when the person moves away (and loses eye contact) during the robot's approach in the remaining phases (Figure 1, third and fourth diamond from the top). As a reminder, Kendon observed that both participants in a greeting tended to move towards each other during the approach. Thus if the robot senses that the person is moving away from the robot when it tries to move through their social space towards them, the robot will take that as a sign that the human does not want to interact with it, and will abandon its greeting as described in our first failure case.

Reflection and Discussion

We very informally evaluated our robot's behaviors, where we considered how people reacted to our demonstrations of the robot as described earlier, as well as our own experiences. This is a reasonable approach for early work, as we were primarily interested in 'big effects' that were immediately obvious.

Overall, the use of proxemics and Kendon's greeting observations proved very effective in simulating a sense of social intelligence in the robot. The Nao, as operationalized from our greeting model, appeared engaged and interested by the way it acted, using eye contact, body language and distancing to effectively communicate the social aspects of human behavior during greetings. Particular robotic behaviors, such as the wave and its vocalizations, were easy for people to interpret and were well received as part of the greeting. We had little doubt that our work – as simplified as it is relative to real human greetings – is a positive starting point for the design of human robot greeting interactions based on the observations of human behavior. However, there were several nuances that emerged in our design, as described below.

Eye contact. The use of eye contact by having the robot's head and tilt angle face the person proved highly effective in simulating social behavior. Even in early versions of the implementation, where the robot simply made eye contact with any human wearing the motion trackers, a perception of social behavior was apparent.

However, this must be tempered somewhat, where the robot would appear to "stare" at the person (recall also that the Nao cannot blink). To alleviate this, we can simply have the robot turn its head slightly away from the person for a short amount of time, and then back again.

Eye contact also proved problematic in situations where the robot was required to move its head for other reasons. For example, we initially programmed the robot to perform several different explicit distance salutations; the head toss, the head nod and the wave. Yet the head toss and head nod were not perceived by people as intentional social actions. The reason is that the robot loses eye contact during these motions (remember, the robot simulates eye contact by its head angle and tilt angle). The Nao robot does not possess the capability to move its eyes and head in independent directions; instead, its eyes appear to look forward, wherever the head is facing. Thus when it performs the head nod and head tilt, it appears as if the robot is looking elsewhere. In contrast, humans direct their eyes independent of the motion of their heads. For example, when a person performs a head toss, he cocks his head upwards while still maintaining eye contact throughout the entire interaction. The robot is not capable of doing this; when its head angles upward, its gaze appears to angle upwards as well, breaking mutual eye contact. This hardware limitation may make this type of social action impossible for the Nao robot. As mentioned in the previous section, this is why we used the wave, instead of the head toss and head tilt, as our explicit distance salutation.

We strongly believe that proper use of eye contact in future human-robot interactions will be fundamental to good design. Given the above, having social robots that can independently control their eye positions would be valuable.

Intentional Gaze. During both its idle behavior and its initial approach, the robot looks in random directions to simulate distractions and curiosity. In actual human behavior, gaze is not directed randomly, but intentionally at things of interest. Having the program evaluate which parts of the environment would realistically catch the attention of the robot. and directing its gaze there during idle behavior and approach may improve the perception of the robot as a social being.

Lack of facial expression. Our robot had very limited ability to adjust any of its facial expressions. Thus behaviors such as 'the smile' could not be implemented. As a consequence, we relied on other equivalent behaviors appropriate for that particular phase in the greeting model.

Inability to implement particular physical greeting behaviors. Several greeting behaviors in our model (and of course the more complex ones described by Kendon) involve physical contact (i.e. handshaking and hugging). Our Nao robot is small and fragile, where such physical contact involves risks to it (e.g., falling over). As well, our robot could not respond appropriately to the subtleties of human physical contact (e.g., detecting and returning a hug), so such actions could not be implemented. Thus some close salutations included in our model were excluded from the implementation.

Palm Presentation and Grooming. In some of our design versions, we did not include the palm presentation and self-grooming. The absence of these actions did not prove particularly noticeable, which suggests that the greeting model has considerable tolerance in what particular gestures can be included and/or varied. Even so, inclusion of these small actions likely add depth to the behavior of the robot.

Pacing. The speed that the Nao robot is able to move, and especially walk, affected the pacing of the interaction considerably. The Nao robot is very slow at standing up, and very slow at walking (it moves in quite small steps). Thus people were required to wait and/or slow down their actions in order to stay 'in sync' with the robot. This made some points of the interaction consciously noticeable and disruptive.

For example, Kendon [1990] observed that humans orient their bodies directly towards each other during the distance salutation. The robot also exhibits this behavior, but because it can take much longer for the robot to do this, it can be more disruptive to the flow of the interaction then when the action is done between humans. What should be barely noticed becomes something the human must wait for before the interaction can continue normally. We strongly believe that social robots needs to have the capability to perform their actions at a socially appropriate speed.

Noise. Because the robot's joints are operated by motors, it produces sound whenever it moves. This is both a benefit and a problem. It is a benefit during the sighting phase, as this introduced sound tends to *attract* the person's attention. It is an annoyance any other time, where it is perceived as noise that does little to contribute to the interaction.

The robot as caricature. Our greeting model, and the way it is implemented on the robot, is just a caricature of human behavior. As mentioned, we believe we have created a 'cartoon' caricature. People were accepting of this, and indeed we saw it as a way for people to be quite tolerant of both the simplicity of the greeting model and the way the robot would exhibit its behaviors.

As research in HRI progresses, we could expect that robotic greeting behaviors could improve to the point that they closely mimic human behaviors. Even so, we have to be wary. For example, the much-debated uncanny valley [Mori, 1970] suggests a phenomenon where an anthropomorphic being (in this case, a robot), appears to have a sense of strangeness to an observer when it comes close to being humanlike, being come across as "eerie" [Mori, 1970]. We expect this phenomenon will also exist during robotic greetings, making hyper-realistic greetings a design challenge.

Implementation aspects. In order to ease implementation, our model is based on a state machine. This approach led to predictable, evaluable behavior in controlled environments. However, a more robust underlying architecture would be an improvement in order to realistically handle failure cases and unexpected situations. The design implemented in this work has very little memory of the interaction. The weaknesses of this approach were most apparent when the greeting is repeated many times; there is not enough variety or reaction to context in the robot behavior to appear realistic.

Another implementation issue concerns the sensing environment. Because our environment relies on expensive fixed cameras and hardware in the space, it is not realistically deployable. A better solution, of course, is to have the robot do all the sensing autonomously, perhaps via its on-board camera. However, this relies on several factors. First, current robots normally have the camera associated with the robot's eyes. Yet our model requires several stages where the robot needs to look away from the person, which implies losing sensing information. Thus a robot's camera should be situated in a way that its view is independent of the direction of its head (e.g., by where it is located, by using wide angle lenses, by having multiple cameras). Second, the on-board vision system must be capable of discerning and tracking not only humans, but the orientation of their face or gaze. This can be a challenge, especially given the relatively low resolution cameras and limited processing capabilities of most robots.

Conclusions

> Someday in the future we may find robots assisting humans in situations well beyond their current use. They may become assistants, workmates, entertainers, even socializers. But to achieve this will require careful consideration of the role and effect of robots in our lives. As robot technology becomes more commonplace, the design aspects will become increasingly important.
>
> *–Journal of Human-Robot Interaction, Special Issue on Design in HRI: Call for Papers*

The HRI design community is currently establishing many methods for creating engaging robots. Like others, we believe that robotic design based on social theories and observations of human behaviors will provide a fruitful and valuable way to shape their overall appeal and usefulness that goes well beyond their technical function and capabilities. We showed one application of this methodology based around Hall's theory of proxemics and Kendon's observations of human greetings. We reviewed the theory, translated it into an abstract greeting model that could be used by technologists, and implemented it on a particular robot. Our preliminary evaluation reveals that it improved our robot's autonomous "social skills" during a greeting exchange in a controlled setting. At the same time, it revealed several design and implementation nuances and challenges that either limit what the robot could do or that disrupt the illusion of sociality. Despite the challenges, these problems appear to be solvable, and future enhancements to software and hardware will improve the social behavior of robots. This will become more important as robots become common in society.

Acknowledgements

This work was funded in part by the NSERC/AITF/SMART Industrial Chair in Interactive Technologies, and by the SURFNET NSERC Network Grant.

References

Aiello, J.R., 1987. Human Spatial Behaviour. In D. Stokols & I. Altman, eds. Handbook of environmental psychology. New York: John Wiley & Sons, pp. 359–504.

Altman, I., 1975. The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding, Brooks/Cole Publishing Company, Monterey, California.

Goodrich, M.A. and Schultz, A.C. (2007) Human-Robot Interaction: A Survey. *Foundations and Trends in Human-Computer Interaction*, 1(3), 203-275. DOI= 10.1561/1100000005.

Hall, E.T. (1966) *The Hidden Dimension*, Garden City: Doubleday.

Kendon, A. (1990) *Conducting Interactions: Patterns of Behavior in Focused Encounters*. New York: Cambridge University Press.

Marquardt, Nicolai (2013) Proxemic Interactions in Ubiquitous Computing Ecologies. PhD thesis, Department of Computer Science, University of Calgary, Calgary, Alberta, Canada, July. Available at: http://grouplab.cpsc.ucalgary.ca/papers.

Marquardt, N., Diaz-Marino, R., Boring, S. and Greenberg, S. (2011) The Proximity Toolkit: Prototyping Proxemic Interactions in Ubiquitous Computing Ecologies, Proceedings of the 24th annual ACM Symposium on User Interface Software and Technology-UIST'11, pp. 315 - 326. DOI= 10.1145/2047196.2047238

Mead, R., Atrash, A. and Mataric, M. (2011) Proxemic Feature Recognition for Interactive Robots: Automating Metrics from the Social Sciences. In Social Robotics: Proceedings of the Third International Conference, ICSR, Amsterdam, The Netherlands, November 24-25, 2011., LNCS Volume 7072, pp. 52-61. DOI=10.1007/978-3-642-25504-5_6

Mori, M. (1970) The Uncanny Valley. *Energy*, vol. 7, no. 4, pp. 33-35, 1970. See English translation hosted by IEEE Spectrum: http://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley.

Mumm, J., & Mutlu, B. (2011) Human-Robot Proxemics: Physical and Psychological Distancing in Human-Robot Interaction. Proceedings of the 6th International Conference on Human-robot Interaction (HRI '11). ACM, New York, NY, USA, 331-338. DOI=10.1145/1957656.1957786.

Satake, S., Kanda, T., Glas, D. F., Imai, M., Ishiguro, H. & Hagita, N. (2009) How to Approach Humans? Strategies for Social Robots to Initiate Interaction. Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction - HRI '09, pp. 109-116. DOI=10.1145/1514095.1514117

Saulnier, P., Sharlin, E. and Greenberg, S. (2011) Exploring Minimal Nonverbal Interruption in HRI. 20th IEEE International Symposium on Robot and Human Interactive Communication-RO-MAN'11. pp. 79-86. DOI: 10.1109/ROMAN.2011.6005257

Summer, R., 1959. Studies in Personal Space. Sociometry, 22(3), pp.247–260.

Sommer, R., 1969. Personal space: the behavioral basis of design., Englewood Cliffs, N.J: Prentice-Hall.

Takayama, L. & Pantofaru, C. (2009) Influences on Proxemic Behaviors in Human Robot Interaction. IEEE / RSK International Conference on Intelligent Robots and Systems, pp. 5495-5502. DOI=10.1109/IROS.2009.5354145

Walters, M., Dautenhahn, K., te Boekhorst, R., Koay, K. L., Syrdal D. S. & Nehaniv, C. L. (2009) An Empirical Framework for Human-Robot Proxemics. Proceedings of New Frontiers in Human-Robot Interaction: Symposium at the Artificial Intelligence and Simulation of Behaviour Convention. pp. 144-149. URI: http://hdl.handle.net/2299/9670

Authors' names and contact information: Brandon Heenan (bmheenan.ux@outlook.com), 888 Western Ave #711 Seattle, WA 98104; Saul Greenberg (saul.greenberg@ucalgary.ca), Setareh Aghelmanesh (setareh.16@gmail.com) and Ehud Sharlin (ehud.sharlin@gmail.com), Department of Computer Science, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4