

Designing Social Greetings in Human Robot Interaction

Brandon Heenan, Saul Greenberg, Setareh Aghel Manesh, and Ehud Sharlin

Department of Computer Science, University of Calgary
Calgary, Alberta, Canada T2N 1N4

bmheenan@outlook.com, [saul.greenberg, saghelma, ehud.sharlin]@ucalgary.ca

ABSTRACT

We designed and operationalized a greetings model for human robot interaction as a state machine, derived from a subset of social behaviors as detailed by Kendon's observations of greetings and augmented by Hall's proxemics theory. Our premise is that designing robot greetings on the social science of human greetings will make the robot's greeting actions socially understandable. Specifically, we track the location and orientation of a Nao humanoid robot relative to a person, and programmed the robot via state transitions to engage in a distance salutation, approach, close salutation and transition as described by theory. Overall, our design appears effective in simulating social intelligence during greetings.

Author Keywords

Human robot interaction; proxemics; greetings; social robots; social science.

ACM Classification Keywords

H.5.m. Human robot interaction.

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.

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

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

DIS '14, June 21 - 25 2014, Vancouver, BC, Canada

Copyright 2014 ACM 978-1-4503-2902-6/14/06...\$15.00.

<http://dx.doi.org/10.1145/2598510.2598513>.

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 non-verbal communications, vocalizations, and inter-personal distancing depending on where one is in the greeting process [Kendon, 1990][Hall, 1966]. Properly enacted greetings with respect to these nuances will help robots and people initiate interaction. Conversely, a robot acting inappropriately during the greeting process may cause it to be misinterpreted, which could jeopardize the interaction.

Yet designing robotic greetings based on human behaviors is far from easy. On the input side, a robots' sensing channels are quite different from human senses: robots cannot yet socially sense, let alone interpret and respond to the nuances, of the world as richly or as quickly as humans can. On the output side, robots typically have many mechanical constraints, especially when compared to the musculature control that people have. This means that robots may not be able to mimic equivalent human actions. These challenges of correctly sensing and reacting to social situations lead to the primary focus of this paper: how do we articulate a reasonable subset of the exact social greeting behaviors the robot should be exhibiting, and how can we operationalize the subtleties of those behaviors?

Our overall contribution is to provide a first approach to formalize and implement greeting salutations for an autonomous robot. Our design approach and specific contributions are structured below as a four-fold method.

1. We review two social science constructs relevant to greetings as our intellectual foundation driving the design of robot-human greeting behavior. These are: body language within the greeting process [Kendon, 1990] and proxemics theory [Hall, 1966]
2. We transform and articulate those theories as behaviors that can be realized on a robot. In particular, we construct an abstract state machine model to capture the essence of the greeting process as described by theory.
3. We operationalize those behaviors on the particular capabilities of the Nao robot.
4. We observe and reflect on both the successes and failures of our approach, where we pay special attention

to the nuances important in designing socially acceptable human-robot greeting interfaces.

THE SOCIAL SCIENCE OF GREETINGS

Our work is based on three main research areas. First is the social science of human greetings, specifically the work of Kendon [1990] who describing the subconscious behaviors observed in typical greetings. Because greeting behavior is influenced by proxemics (the study of interpersonal distancing in humans), we briefly review that research as well. The third research area is human-robot interaction, where we address current work related to robotic greetings.

Human Greetings

Adam Kendon 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, p153]

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, oriented towards a western culture. In essence, Kendon found that a typical exchange between two individuals who wish to greet each other follows a structure of mostly non-verbal communications comprising phases; sighting, distance salutation, approach, close salutation, and finally a transition into the interaction.

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

- *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 continue to the next phase.

Approach. Assuming the greeting is not simply a terminal distance salutation, the two parties close the distance between themselves. 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 their 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.* A person will typically smile if not yet smiling.
- *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:* participants halt facing one another and exchange verbal greetings without additional non-verbal signals. This phase is still distinct, as 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 formality of the occasion. They are very common in male-male greetings but uncommon in female-female greetings.
- *Embraces* are also observed in human greetings, although this depends on the relationship between the two parties.
- *Other close salutations* exist. Many are culturally dependent, e.g., bowing, cheek to cheek kisses, etc.

While these 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. Once the close greeting has been performed, participants tend to increase or decrease the space between them in a way that matches the proxemic relation of their interaction. People then typically adopt a stance not directly facing each other: Kendon [1990] pursues patterns of these stances in later chapters of his book dealing with ‘f-formations’. 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’s description of human greetings is far from exhaustive, but suffices to influence our work on basic robot greeting behaviors as described in later sections.

Proxemics

Within Kendon’s greeting process phases, people adjust their interpersonal distance, i.e., where specific greeting behaviors are observed at inexact, but predictable distances [Kendon 1990]. This builds upon Hall’s [1966] seminal account of interpersonal distancing through his theory of *proxemics*. At its simplest, Hall’s thesis is that people equate social distance with physical distance. According to Hall [1966], humans tend to exhibit different behaviors towards each other in accordance with four levels of “closeness” as listed below. The actual distances provided below are typical of western cultures.

- *Intimate Space* exists when people are less than 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.
- *Personal space* (0.45–1.2m) tends to be used by people in conversation who are comfortable with one other. Humans use normal voice levels, and are able to clearly see another person’s face in great detail.
- *Social space* (1.2–3.6m) 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.
- *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 oneself away from the audience. 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.

There are many other nuances to Hall’s theory (see [Marquardt 2013, Chapter 3], [Walters et. al. 2009]). For example, fixed features (e.g., boundaries such as doorways) and semi-fixed features (e.g., the positioning of furniture) affect how people perceive social distance [Hall, 1966]. Body orientation affects perceived social distance: facing towards one another, kitty-corner, side by side, or away [Sommer, 1969; Kendon 1990]. Personal space serves as a protective function [Aiello, 1987] somewhat akin to territories [Sommer, 1969]. Use of this space is dictated by social rules and norms [Aiello, 1987], where people take umbrage if those rules are broken [Altman 1975].

Within greetings, proxemics are relevant in all phases. Sighting typically happens in the public zone. During the approach, people move progressively through zones and signal accordingly. During a final approach, they can momentarily step into each other’s intimate zone. During the transition, they likely step back to the personal or public zone (depending upon their social level of engagement).

HUMAN ROBOT INTERACTION

HRI researchers often motivate robot design around social behaviors. While a few researchers have incorporated greeting behaviors within human robot interaction, their designs appear to be driven by intuition rather than by

following specific formal social science greeting theories. Still, several have considered social science aspects to design complementary robotic behaviors, such as approach, interruption and proxemics. To our knowledge, no one has yet considered how we could apply Kendon's detailed greetings description to HRI design.

Several projects explored how people react to a robot that greeted them. Bainbridge et. al. [2008] had an experimenter introduce a subject to a non-mobile physical robot, whereupon the robot would wave at them. About half the subjects produced a greeting response back, such as a return wave or a verbal response. Trovato and Zecca (2013) observed how people from two different cultures (Japanese and Egyptian) reacted to greetings made by what appeared to be Japanese and Egyptian non-mobile robots over a simulated teleconference call. The Japanese robot would bow and vocalize a Japanese greeting, while the Egyptian robot raised its hand and vocalized in Egyptian. Participants from the Egyptian culture reacted more favorably to the Egyptian robot, where reasons included understandability, sense of familiarity, and non-verbal communications. Japanese participants were somewhat neutral to either. Both projects were Wizard of Oz studies, where an experimenter actuated robot actions. Gockley et. al. (2005) created a non-mechanical virtual receptionist - a graphical face on a screen - that could (amongst other things) greet people. Our work differs from these projects, as it embodies a more complex greeting model performed by an autonomous and mobile robot.

Saulnier et al. [2011] investigated how people perceived a mobile robot's attempt to attract their attention and interrupt a conversation in progress. Using Wizard of Oz, 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 physical behaviors conveyed the most information regarding its sense of interruption urgency. Our work differs as it considers conventional greetings rather than interruptions.

Satake et al [2009] was interested in how people perceived approaching mobile robots. Their first approach of simply taking the shortest distance to the nearest person (including approaching from behind) and attempting a verbal greeting did not have a high success rate. They showed that approaching the human from the front significantly increased the chances of a person successfully starting a conversation. Dautenhan et. al. (2006) added that people preferred robots whose frontal approach was from the right side rather than directly frontwards. Satake et. al. [2009]

also observed that people usually ceased interacting with the robot, when they "tested" it for a reaction but then did not get the expected response. Our work differs, as our robot both approaches and responds to the person in a manner that conforms to expected greeting protocols.

Others considered how proxemic factors within HRI. Takayama and Pantofaru [2009] empirically established that people were comfortable with an interpersonal distance ranging from 0.4–0.6m when approaching or being approached by a 1.35 m tall robot. Distancing was affected by other factors, such as mutual gaze and prior experiences with robots. Walters et al [2009] did a similar experiment that produced comparable findings, where they identified other factors affecting distancing (e.g., robot height, voice style, gender, appearance, etc). Mumm and Mutlu [2011] also 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). Mead et. al. (2011, 2013) disclosed that other factors known to affect human-human proxemics also apply to HRI. All these works generally confirm that proxemic theory is applicable to HRI. Humans react to a robot as a social being by adjusting their inter-personal distance from it. Humans are generally comfortable with robots in their personal space, although various factors affected the precise distancing. Our work builds upon this: proxemics and distancing metrics are a subset of our broader greeting 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. We considered that it was not currently feasible – nor desirable – to create a high fidelity literal translation of all the proxemics and greeting behaviors described in the literature. We knew that such behaviors – even if they could be translated – could not be applied to current robots. 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.

Consequently, we decided instead to consider the robot as a caricature. It would exhibit a primary behaviors that Kendon's descriptions suggest are socially essential, but would leave out some of the finer-grained nuances. These behaviors would serve as a first order approximation for the design of a robot capable of autonomous human greetings.

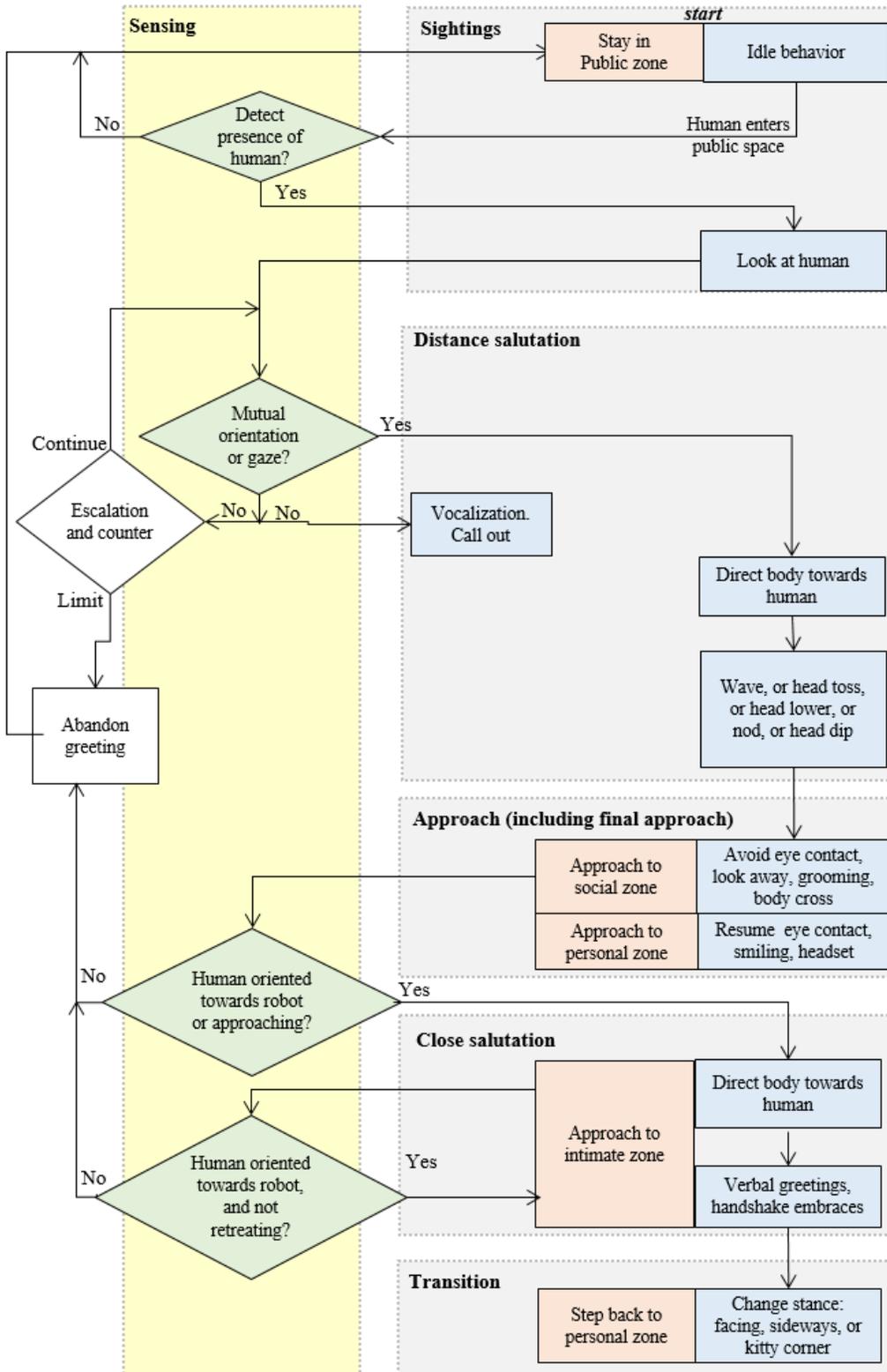


Figure 1. The Greeting Model as a State Diagram

For ease of implementation, we model the flow of these behaviors as an abstract state machine. We did this because Kendon's greeting process appears 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.

Figure 1 illustrates our translation of greeting behaviors as a state diagram. Robotic behavioral states are collected on the right side in the large light grey areas. 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.

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 and to make the interaction process more natural, as advocated by Bainbridge et al., 2008). 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 the 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

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 (assuming the robot is physically capable of these actions). The robot then enters the Close Salutation phase. 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: it steps back into the personal zone and changes its stance to be 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

The greetings model as described in the prior section is an abstract model. We now show how we operationalized this model on a robot: the Nao by Aldebaran Robotics (www.aldebaran-robotics.com).

The Nao Robot

The Nao robot is illustrated in Figures 2 and 3. It is significantly smaller than an adult human, only 58cm 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 includes a voice synthesizer, various lights, and speakers. It

can move a large number of joints including feet, hands, fingers, elbows, shoulder, forearms, head, and neck.

We control the robot through a custom .Net application using the NaoQi API. We preferentially use non-blocking calls, as any behavior must be cancellable at any point if the robot is to react in sync to sensed human actions.

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 (<http://www.vicon.com>). The system tracks the 3D x, y and z location of specialized reflective passive markers arranged in patterns, as well as yaw, pitch and roll. The robot and person wear markers on strategic body parts to allow tracking. The robot wears its marker atop its head or shoulders, which allows us to track both the location of the robot and its orientation. Because the robot is controlled in code, we know its relative head and body orientations. 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.

We use the Proximity Toolkit [Marquardt et. al., 2011] to calculate various proxemics relationships, derived from the raw data produced by the motion tracker. The Proximity Toolkit produces 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 autonomously 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 the 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 some random spot. Similarly, the software controls the robot's movement to particular locations, where it uses the current state to determine, when and where the robot should move relative to the person, and what its body orientation should be relative to the 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 Robot's Greeting Interaction

We implemented several designs that vary aspects of the greetings model described here. For brevity, we describe these as a single integrated design, where we detail how the Nao operationalizes the greetings model.

As a pre-cursor, we know that the people's proxemic zones would be likely different from Hall's estimates, as they would be interacting with a robot [Takayama and Pantofaru 2009], that was also quite small. Thus we adjusted our distances downwards as an estimate of where reasonable zones could actually lie: from the intimate (<2m), to personal (0.2 – 0.75 m), to social = (0.75 – 1.3 m), to public zone (>1.3m).

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 [Bainbridge et. al., 2008; [Satake et al 2009]).

Consequently, we implemented several idle behaviors designed to show the aliveness of the robot, with several examples illustrated in Figure 2. We designed these idle behaviors based on known 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 (not shown), 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 [Bainbridge et. al., 2008], where people who are left alone may look for someone or something to interact with. In our implementation (Figure 2a), the robot turns its head to left and right to simulate it looking around for something to do. A third action is wiping the forehead (Figure 2b), usually associated with the person being tired or bored. Our next idle behavior is stretching: most people stretch somewhat

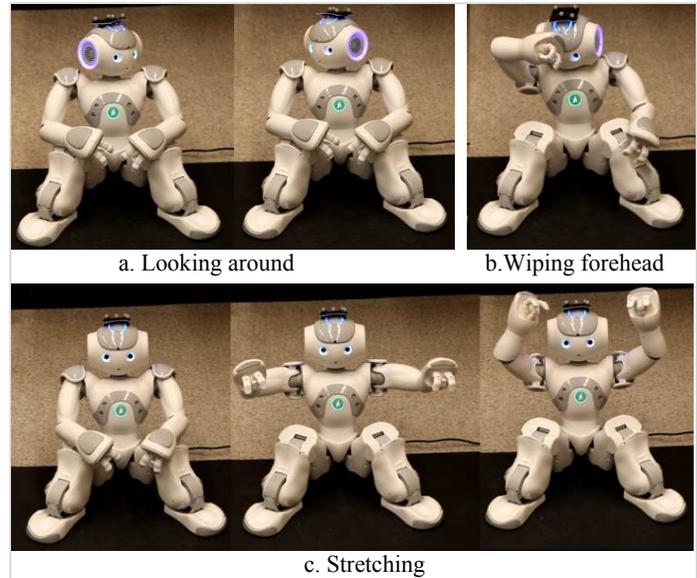


Figure 2. Idle behaviors of the Nao Robot during Sighting Phase.

when they have not moved for a long time (somewhat similar to Bainbridge et. al. [2009]'s arm-swinging). As seen in Figure 2c, the robot simulates this by raising both its hands and bringing them back down slowly. Other idle behaviors have the robot shift its weight, occasionally looking around the room, and/or standing up and moving towards a random location. 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 similar to human idle behaviors. 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 3a), 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.

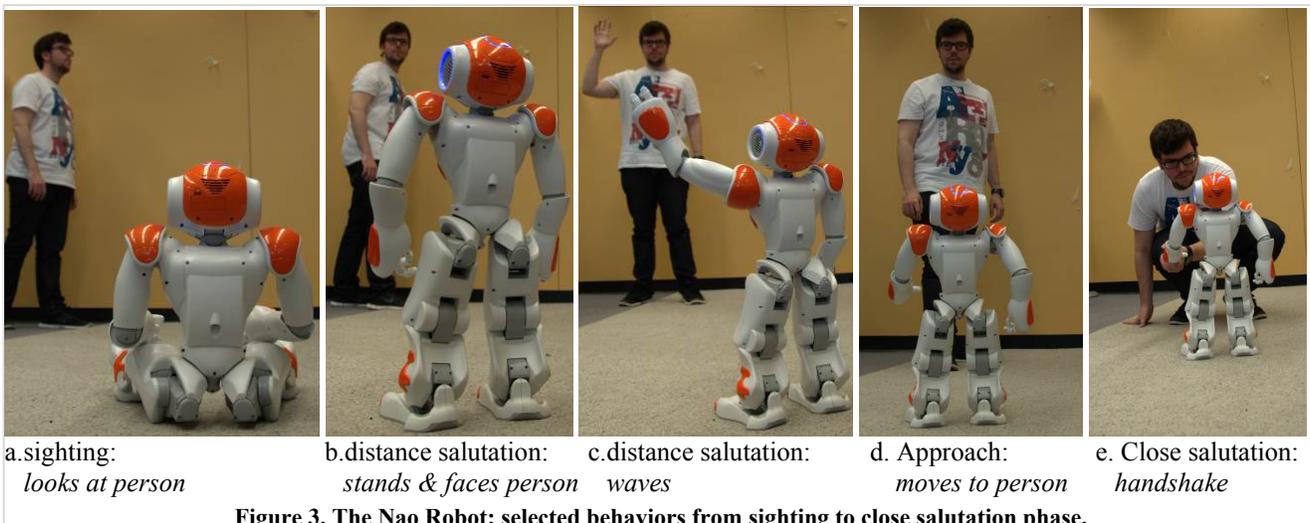


Figure 3. The Nao Robot: selected behaviors from sighting to close salutation phase.

Distance salutation. Once the user makes eye contact with the robot, it engages in a distance salutation. Our system senses 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 seated (e.g., as in Figure 2a-c), the robot will stand up. 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 3b). 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 worked best for this particular robot. As seen in Figure 3c, the robot waves by raising its arm with an open hand directed towards the person.

Approach. After the distance salutation, the robot will then move towards the person, passing from the social to the 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-orient its head to simulate resuming eye contact (Figure 3d). 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 3e. 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?" 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 top). Kendon had 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 informally observed how people reacted to our demonstrations of the robot. Combined with our own experiences, we reflect on both the successes and failures of our greeting model and implementation. This is a reasonable approach for early work, as we were primarily interested in 'big effects' that were immediately obvious. Future work should, of course, redesign the robot behaviour based on these observations, and conduct a formal evaluation to consider people's fine-grained reactions.

Overall, 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. People reacted favorably. 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 greeting model – 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, several nuances emerged in our design, described below. While some arise from particular constraints of the

Nao robot, they are relatively generalizable to other robots that share similar or analogous constraints.

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, where people naturally felt that the robot was responding to their presence. However, persistent eye contact can become uncomfortable, where the robot would appear to "stare" at the person (recall also that the Nao cannot blink). To temper this, we can simply have the robot avert its gaze away from the person for a short amount of time, and then back again.

Eye contact also proved technically 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 human-robot interactions is fundamental. Given the above, having social robots that can independently control their eye positions would be crucial.

Intentional Gaze. During its idle behavior and its initial approach, the robot looks in random directions to simulate distractions and curiosity. In actual human behavior, gaze is directed intentionally at things of interest. Having the program evaluate which parts of the environment should 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. As with most off-the-shelf robots, the Nao 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). Larger robots would have the opposite effect of safety effects, where people could feel (or be) threatened by physical contact. As with most robots, 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 such actions will add depth to how robots are perceived.

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 took the Nao much longer to do this, it somewhat disrupted the continuity of the interaction flow. What should be barely noticed becomes something the human must now wait for before the interaction can continue. We strongly believe that social robots need to have the capability to perform their actions at socially appropriate speeds. This includes large robots – if they approach too quickly, they could appear menacing.

Noise. Because robot 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. 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.

Implementation aspects. Our state machine model led to predictable, evaluable behavior in controlled environments. However, a more robust underlying modeling approach

(perhaps a Markov model) is needed to realistically handle failure cases and unexpected situations. Our design 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 loses 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

The HRI design community is currently establishing many methods for creating engaging robots. Like others, we believe that robotic design based on formal 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 Kendon's observations of human greetings and supported by Hall's theory of proxemics. We reviewed the theory, translated it into an abstract greeting model that could be used by technologists, and implemented it on a particular autonomous robot. Our reflection of the robot-in-action is that the greeting model improved the 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 somewhat disrupt the illusion of sociality. Despite the challenges, these problems are solvable, where future software and hardware enhancements will improve the robot's ability to enact social behavior. This will be increasingly important as robots become commonplace.

Acknowledgements. Funding by NSERC/AITF/SMART Industrial Chair in Interactive Technologies, and NSERC's Discovery Grant and Surfnet Network.

REFERENCES

1. Aiello, J. (1987) Human Spatial Behaviour. In D. Stokols & I. Altman, eds. *Handbook of environmental psychology*. John Wiley & Sons, 359–504.
2. Altman, I. (1975) *The Environment and Social Behavior*. Brooks/Cole Publishing Company.
3. Bainbridge, W., Hart, J. Kim, E. & Scassellati, B. (2008) The Effect of Presence on Human-Robot Interaction. *Proc. IEEE Ro-Man*, 701-706.
4. Dautenhahn, K., Walters, M., Woods, S., Koay, L., Nehaniv, C., Sisbot, A., Alami, R. & Sim, T. (2006). How May I Serve You? A Robot Companion Approaching a Seated Person in a Helping Context. *Proc ACM Human-Robot interaction*, 172-179.
5. Goodrich, M. & Schultz, A. (2007) Human-Robot Interaction: A Survey. *Foundations and Trends in Human-Computer Interaction*, 1(3), 203-275.
6. Hall, E. (1966) *The Hidden Dimension*, Doubleday.
7. Kendon, A. (1990) *Conducting Interactions: Patterns of Behavior in Focused Encounters*. Cambridge Univ Press
8. Marquardt, N. (2013) *Proxemic Interactions in Ubiquitous Computing Ecologies*. PhD Thesis, Dept Computer Science, Univ. of Calgary, Calgary, Canada, July. <http://grouplab.cpsc.ucalgary.ca/papers>.
9. Marquardt, N., Diaz-Marino, R., Boring, S. & Greenberg, S. (2011) The Proximity Toolkit: Prototyping Proxemic Interactions in Ubiquitous Computing Ecologies, *Proc. ACM UIST*, 315 - 326.
10. Mead, R., Atrash, A. and Mataric, M. (2011) Proxemic Feature Recognition for Interactive Robots. *Proc Social Robotics*, LNCS Volume 7072, pp. 52-61, Elsevier.
11. Mead, R., Atrash, A. and Mataric, M. (2013) Automated Proxemic Feature Extraction and Behavior Recognition: Applications in Human-Robot Interaction. *Int J Soc Robot*, 5(3)May, Springer
12. Mumm, J., & Mutlu, B. (2011) Human-Robot Proxemics: Physical and Psychological Distancing in Human-Robot Interaction. *Proc ACM HRI*, 331-338.
13. Satake, S., Kanda, T., Glas, D., Imai, M., Ishiguro, H. & Hagita, N. (2009) How to Approach Humans? Strategies for Social Robots to Initiate Interaction. *Proc ACM HRI*.
14. Saulnier, P., Sharlin, E. and Greenberg, S. (2011) Exploring Minimal Nonverbal Interruption in HRI. *Proc RO-MAN*, 79-86.
15. Sommer, R. 1969. *Personal Space: The Behavioral Basis of Design*, Prentice-Hall.
16. Takayama, L. & Pantofaru, C. (2009) Influences on Proxemic Behaviors in Human Robot Interaction. *IEEE Int'l Conf Intelligent Robots and Systems*, 5495-5502.
17. Trovato, G. & Zecca, M. Towards Culture-Specific Robot Customization: A Study on Greeting Interaction with Egyptians. *Proc. IEEE Ro-Man*, 447-452.
18. Walters, M., Dautenhahn, K., te Boekhorst, R., Koay, K. L., Syrdal D. S. & Nehaniv, C. L. (2009) An Empirical Framework for Human-Robot Proxemics. *Proc New Frontiers in Human-Robot Interaction: Symp. AI and Simulation of Behaviour*. pp. 144-149.
19. Gockley, R., Bruce, A., Forlizzi, J., et. al. (2005) Designing Robots for Long-Term Social Interaction. *IEEE/RSJ Conf Intelligent Robots and Systems (IROS)*.