Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®

ELSEVIER

Robotics and Autonomous Systems 42 (2003) 167–175

Robotics and Autonomous Systems

www.elsevier.com/locate/robot

Toward sociable robots

Cynthia Breazeal*

MIT Media Lab, 77 Massachusetts Ave, NE 18-5FL, Cambridge, MA 02139, USA

Abstract

This paper explores the topic of social robots—the class of robots that people anthropomorphize in order to interact with them. From the diverse and growing number of applications for such robots, a few distinct modes of interaction are beginning to emerge. We distinguish four such classes: socially evocative, social interface, socially receptive, and sociable. For the remainder of the paper, we explore a few key features of sociable robots that distinguish them from the others. We use the vocal turn-taking behavior of our robot, Kismet, as a case study to highlight these points.

Keywords: Sociable robots; Human–robot interaction; Humanoid robots; Facial expression; Social interaction

1. Introduction

Recent commercial applications are emerging where the ability to interact with people in an entertaining, engaging, or anthropomorphic manner is an important part of the robot's functionality. A new generation of robotic toys have emerged (such as Tiger Electronic's hamsters-like Furby or Sony's robotic dog, Aibo) whose behavior changes the more children play with it. Video games, such as Creatures, allow the participant to "genetically" design graphical critters and then interact with them. Lego Mindstorms takes a more engineering approach, providing people with a robot toolkit.

Although the ability of these products to interact with people (and people's ability to interact with them) is limited, they are motivating the development of increasingly life-like and socially sophisticated robots. Mediated communication through robotic avatars (http://www.tele-actor.com/original.htm) would allow one to have a physical and social presence to others despite being geographically distant. Location based entertainment applications such as museum tour guide robots [13] offer not only entertainment value but also provide visitors with information of interest. Health-related applications are being explored, such as robot nursemaids that help the elderly (http://www-2.cs.cmu.edu/~nursebot), or robotic pets (such as Omron's NeCoRo) that are intended to provide some of the health-related benefits of pet ownership. NASAs humanoid robot, Robonaut, developed at the Johnson Space Center is envisioned to be an astronaut's assistant. The success of these robots hinges not only on their utility but also on their ability to be responsive to and interact with people in a natural and intuitive manner.

2. Paradigms of social robots

It is important to recognize that humans are a profoundly social species. Our social-emotional intelligence is a useful and powerful means for understanding the behavior of, and for interacting with, some of the most complex entities in our

* Tel.: +1-617-452-5601; fax: +1-617-225-2009.
E-mail address: cynthiab@media.mit.edu (C. Breazeal).

0921-8890/03/$ – see front matter © 2003 Published by Elsevier Science B.V.
doi:10.1016/S0921-8890(02)00373-1

world—people and other living creatures [12]. Faced with non-living things of sufficient complexity (i.e., when the observable behavior is not easily understood in terms of its underlying mechanisms), we often apply a social model to explain, understand, and predict their behavior as well [15]. For instance, we are all familiar that people anthropomorphize all sorts of technologies (e.g., cars, computers, etc.). The studies of Premack and Premack [14] show that people attribute mental states (i.e., intents, beliefs, feelings, desires, etc.) to describe the behavior of interacting shapes on a screen. Even Braitenberg [4] often uses such mentalistic terms to describe the behavior of his ingenious vehicles. Right or wrong, people rely on social models (or fluidly switch between using a social model with other mental models) to make complex behavior more familiar and understandable and more intuitive with which to interact. We do this because it is enjoyable for us, and it is often surprisingly quite useful.

We argue that people will generally apply a social model when observing and interacting with autonomous robots. Autonomous robots perceive their world, make decisions on their own, and perform coordinated actions to carry out their tasks. As with living things, their behavior is a product of its internal state as well as physical laws. Augmenting such self-directed, creature-like behavior with the ability to communicate with, cooperate with, and learn from people makes it almost impossible for one to not anthropomorphize them (i.e., attribute human or animal-like qualities). We refer to this class of autonomous robots as social robots,[1] i.e., those that people apply a social model to in order to interact with and to understand.

This definition is based on the human observer's perspective. One can further distinguish several subclasses of social robots by viewing matters from the robot's perspective. Namely, given that people will apply a social model in order to understand and interact with such robots, to what extent does the robot design support and validate this model? To what extent do these anthropomorphic attributes map to computational entities within the robot? To what extent does the robot apply a social model to understand people? Basically, to what extent is the robot a full-fledged social participant?

This is fundamentally an "appearance versus reality" question—does the robot only appear to be socially intelligent or is it genuinely so? (We argue that the robot's design does not have to be indistinguishable from human analogs to serve the same function or to exhibit the same competence, i.e., to be genuine.) One could imagine various subclasses of social robots as lying along an appearance verses reality continuum.

Another way to look at this distinction is in terms of the brittleness of the design—namely, when does the robot's behavior no longer adhere to the person's social model for it? Once this occurs, the usefulness of the person's social model for the robot has been marginalized—it breaks down. A robot that only appears to be socially intelligent is "believable" [2]. This may be completely acceptable for a sufficiently structured scenario such as theme park entertainment where the audience's interaction with the robot is highly constrained.

As the complexity of the environment and the scenario increases, however, the social sophistication of the robot will clearly have to scale accordingly. If the robot's observable behavior adheres to a person's social model for it during unconstrained interactions in the full complexity of the human environment, then we argue that the robot is socially intelligent in a genuine sense. Basically, the person can engage the robot as one would another socially responsive creature, and the robot does the same. At the pinnacle of performance, this would rival human–human interaction. However, we take the position that dogs, for instance, are another socially responsive species—socially intelligent in a genuine sense, although their social sophistication is less than that of a human. Hence, this criterion for success should not be confused with trying to build a robot that is indistinguishable from a human inside and out—the appearance of the robot and its internal design details can be quite different from

[1] Traditionally, the term "social robots" was applied to multi-robot systems where the dominant inspiration came from the collective behavior of insects, birds, fish such as flocking, foraging etc. For this reason, the author coined the term "sociable" to distinguish an anthropomorphic style of human–robot interaction from this earlier insect-inspired work. The author has learned (after recent discussions with Terry Fong) that the term "social" has apparently changed over the years to become more strongly associated with anthropomorphic social behavior. Hence, we shall adopt this more modern use of the term "social" in this paper, but still distinguish "sociable" as a distinct subclass of social robots (to remain consistent with the original intention behind coining this term).

the human counterpart, what matters is how it interacts face-to-face with people, and how people interact with it in the human environment. One could imagine devising a sort of test for sociability, where the behavior of both the human and the robot are evaluated in order to determine success.

Although the field of social robots and their design is still fairly new, we can begin to distinguish several subclasses of social robots from existing applications and examples. In all cases, people are inclined to anthropomorphize these robots in order to interact with them. As one moves successively through the list, the design of the robot increases in its ability to support the social model in more complex environments (i.e. approaching the human social environment) and for more complex scenarios (i.e., approaching human style face-to-face interaction).

Socially evocative. As the term implies, this subclass is designed to encourage people to anthropomorphize the technology in order to interact with it, but goes no further. This is quite common in toys, where a nurture model is leveraged to yield an entertaining interaction (e.g., Tomogotchis and a wide variety of robotic "pets"). Also popular, typically in video games such as Creatures is to use a creator/inventor model where participants "breed" or engineer animated creatures to interact with them. The act of "creating" these simple creatures encourages the participant to feel more invested in their creation's "lifespan". In short, the human attributes social responsiveness to the robot, but the robot's behavior does not actually reciprocate.

Social interface. This subclass of robots uses human-like social cues and communication modalities in order to facilitate interactions with people (i.e., to make the interactions more natural and familiar). For instance, one might adopt a performance model to communicate with others from far away using a robot avatar (giving the distal person both a physical presence and social presence to others). In this case, such a robot would need sufficient social intelligence to appropriately convey (or perform) a person's message to others, complemented with gaze, gestures, facial expression, etc. More commonly, an interface model is used, such as robot museum tour guides, where information is communicated to people using speech and sometimes with reflexive facial expressions. Because this class of robot tends to value social behavior only at the interface, the social model that the robot has for the person tends to be shallow (if any) and the social behavior is often pre-canned or reflexive.

Socially receptive. Whereas the benefits of socially communicative applications (e.g., those described in the previous paragraph) are dominantly measured from the human's perspective, socially receptive robots also benefit from interactions with people. Such examples often involve robots that learn from interacting with people through human demonstration (following a training model), such as acquiring motor skills [1], or a proto-language [3]. Cognitive modeling of a particular competence is more popular in this subclass. Interactions with people affect the robot's internal structure at deeper levels, such as organizing the motor system to perform new gestures, or associating symbolic labels to incoming perceptions. People can shape the robot's behavior through other social cues, such as using gaze direction or head pose to direct the robot's attention to a shared reference [16]. This class of robots tends to be more perceptive of human social cues, allowing people to shape the robot's behavior in richer ways. They are socially passive, however, responding to people's efforts at interacting with them but not pro-actively engaging people to satisfy internal social aims.

Sociable. Sociable robots are socially participative "creatures" with their own internal goals and motivations. They pro-actively engage people in a social manner not only to benefit the person (e.g., to help perform a task, to facilitate interaction with the robot, etc.), but also to benefit itself (e.g., to promote its survival, to improve its own performance, to learn from the human, etc.) [6]. Hence, social interactions with people are valued not just at the interface, but at a pragmatic and functional level as well. Such robots not only perceive human social cues, but at a deep level also model people in social and cognitive terms in order to interact with them. The design of the robot maps the human's social model for it to underlying computational entities. Hence the robot's social behavior is a product of its computational social "psychology".

3. Our sociable robot, Kismet

This remainder of this paper focuses on the last paradigm, robot as sociable creature. We highlight a few core attributes of sociable robots by means of a

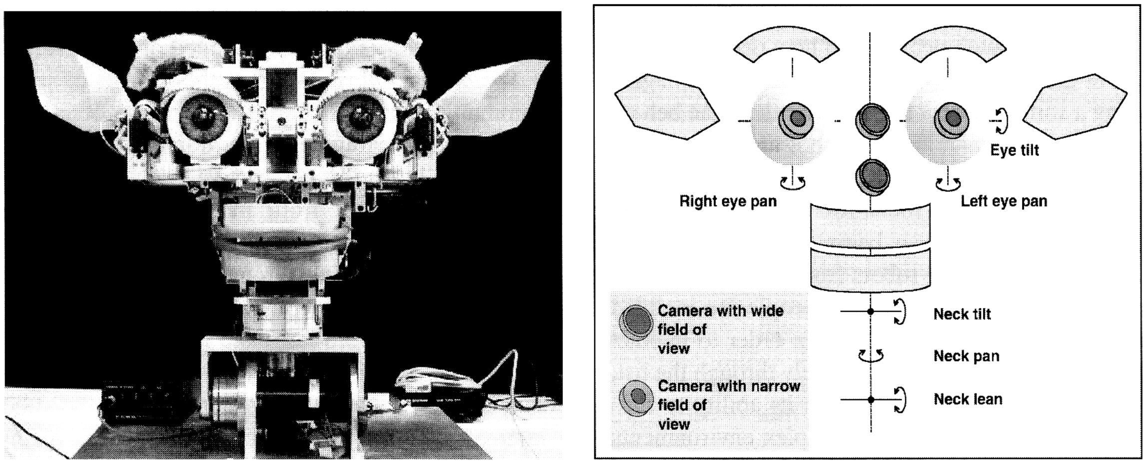

Fig. 1. Kismet is an expressive anthropomorphic robot head. Three degrees of freedom direct the robot's gaze, another three control the orientation of its head, and the remaining 15 move its facial features (e.g., eyelids, eyebrows, lips, and ears). To visually perceive the person who interacts with it, Kismet is equipped with a total of four color CCD cameras. A lavalier microphone worn by the person is used to process their vocalizations.

small case study—the design and evaluation of the vocal turn-taking behavior of our robot, Kismet (see Fig. 1). One does not use Kismet to perform a task. Instead, Kismet is designed to be a robotic creature that can interact physically, affectively, and socially with humans in order to ultimately learn from them. Accordingly, our robot is designed to elicit interactions with the human that afford rich learning potential. We have endowed Kismet with a substantial amount of infrastructure that we believe will enable the robot to leverage from playful, infant-like interactions to foster its social development.

These skills and mechanisms help it to cope with a complex social environment, to tune its responses to the human, and to give the human social cues so that he/she is better able to tune him/herself to Kismet. This allows the robot to be situated in the world of humans without being overwhelmed or under-stimulated [5]. Currently, these skills include the ability to direct the robot's attention to establish shared reference [10,16], the ability to give readable, expressive feedback to the human [7], the ability for the robot to recognize expressive feedback such as praise and prohibition [9], the ability to take turns to structure the learning episodes [8], and the ability to regulate interaction to establish a suitable learning environment [7]. Due to space limitations, we refer the interested reader to [6] for a detailed presentation of Kismet's overall design.

4. Regulating the exchange of speaking turns

The ability to exchange turns during face-to-face interactions is the cornerstone of human style communication and instruction. The tempo and rhythm of conversational turn-taking is flexible, robust to perturbation, and mutually regulated by the interlocutors. In a teaching scenario, the ability to exchange turns allows the instructor to structure the interaction, such as providing variations upon a theme. Allowing the human to respond with immediate contingency to the robot (and vice versa) gives the robot finer control over the human's behavior, allowing it to discover which actions on its part gives rise to a particular response from the human.

Well studied by discourse theorists, humans employ a variety of para-linguistic social cues, called envelope displays, to regulate the exchange of speaking turns [11]. They involve facial displays, gestures, shifts in gaze, and eye blinks—such as raising one's brows and establishing eye contact to relinquish one's speaking turn, or looking aside to hold one's speaking turn even when speech is paused.

Given that a robotic implementation is limited by perceptual, motor, and computational resources, we have found that these social cues are useful in regulating the turn-taking of humans and robots. This has proven particularly important for Kismet because its speech processing limitations force the robot to exchange speaking turns at a slower rate than is typical for human adults (for humans, this takes place after a 0.25 second pause once speech has ended. However, Kismet does so after a minimum of a 0.5 second pause). However, humans seem to intuitively read Kismet's cues and use them to regulate the rate of exchange at a pace where both partners perform well.

To avoid a "canned" performance, Kismet does not exhibit displays according to a rigid schedule, but rather applies them more as a rule-of-thumb. Note that we refer to this vocal exchange as a proto-dialog because although the human's utterances are spoken in natural language (i.e., in English), the robot uses a Kismet-esque babble for its speaking turn. Hence, the envelope displays are used to regulate the dynamics of interaction during the exchange of speaking turns, rather than focus on the content of what is said. The same holds true for proto-dialogs between human caregivers and their pre-linguistic infants.

Kismet's envelope displays are as follows:

- To acquire the floor: break eye contact and/or lean back a bit.
- To start one's speaking turn: vocalize a Kismet-esque babble.
- To stop one's speaking turn: stop vocalizing and re-establish eye contact.
- To hold the floor: look to the side.
- To relinquish the floor: raise brows and/or lean forward a bit.

Blinking tends to occur at the end of a vocalization.

4.1. Vocal turn-taking experiments

To investigate Kismet's turn-taking performance during proto-dialogs, we invited four naive subjects to interact with Kismet. Subjects ranged in age from 12 to 28 years old. Two male and two female subjects participated. In each case, the subject was simply asked to carry on a "play" conversation with the robot. The exchanges were video recorded for later analysis and annotated according to Table 1. The subjects were told that the robot neither speaks nor understands English, but babbles in a characteristic manner. Due to space limitations here, only a small portion of the proto-dialog carried out by Subject C (a female subject) is presented in Appendix A. The time codes are those that appear on the videotape used to record the sessions. A turn is defined with respect to the speaker who holds the floor and consists of four phases: acquire the floor (Aq), start the utterance (St), stop the utterance (Sp), and relinquish the floor (Rq). The speaker may also hold the floor (Hd), i.e., maintain their speaking role during a silent pause.

Table 1
Annotations for proto-dialog experiments (see Appendix A for sample of annotated video)

Type	Option	Annotations
Listener, speaker	Human	H
	Robot	R
Turn phase	Acquire floor	Aq
	Start speech	St
	Stop speech	Sp
	Hold floor	Hd
	Relinquish floor	Rq
Cue	Avert gaze	
	Eye contact	
	Elevate brows	
	Lean forward	
	Lean back	
	Blink	
	"Utterance"	
Turns	Clean turn	#
	Interrupt	I
	Missed	M
	Pause	P

4.2. Evaluation

We evaluate these human–robot interactions with respect to three criteria. These are inherently subjective, yet quantifiable, measures that evaluate the quality and ease of interaction between human and robot. They address the behavior of both partners, human and robot. The evaluation criteria are as follows:

- Do people intuitively read and naturally respond to Kismet's social cues?
- Can Kismet perceive and appropriately respond to these naturally offered cues?
- Does the human adapt to the robot, and the robot adapt to the human, in a way that benefits the interaction? Specifically, is the resulting interaction natural, intuitive, and enjoyable for the human; can

Kismet perform well despite its perceptual, mechanical, behavioral, and computational limitations?

Often the subjects begin the session by speaking longer phrases and only using the robot's vocal behavior to gauge their speaking turn. They also expect the robot to respond immediately after they finish talking. Before the subjects adapt their behavior to the robot's capabilities, the robot is more likely to interrupt them. For instance, it is often the case that the robot interrupts them within the first couple of exchanges. In general, there tends to be more frequent delays in the flow of "conversation" where the human prompts the robot again for a response. Often these "hiccups" in the flow appear in short clusters of mutual interruptions and pauses (often over 2–4 utterances of the speaker) before the turn phases become coordinated and the flow of the exchange of speaking turns smoothes out. We call these clusters significant flow disturbances.

Video analysis of these human–robot "conversations", provide evidence that people entrain to the robot (see Table 2). They often start to use shorter phrases, wait longer for the robot to respond, and more carefully watch the robot's turn-taking cues. For instance, the robot prompts the person to take their speaking turn by either craning its neck forward, raising its brows, or establishing eye contact when it is ready for them to speak. It will hold this posture for a few seconds until the person responds. Often, within a second of this display, the subject does so. When the subject stops speaking, Kismet tends to lean back to a neutral posture, assume a neutral expression, and perhaps shift its gaze away from the person. This cue indicates that the robot is about to speak. The robot typically issues one utterance, but it may issue several. Nonetheless, as the exchange proceeds, the subjects are more likely to wait until prompted by the relinquish turn display.

As the subjects seem to adjust their behavior according to Kismet's envelope displays these "hiccups" within speaking turns become less frequent. As can be seen in Table 2, for each subject there are progressively

Table 2
Data illustrating evidence for entrainment of human to robot (as time progresses there are increasing number of clean turns before a "hiccup" in the flow occurs)

	Time stamp (min:s)	Clean turns (s)
Subject 1		
Start 15:20	15:20–15:33	13
	15:37–15:54	21
	15:56–16:15	19
	16:20–17:25	70
End 18:07	17:30–18:07	37+
Subject 2		
Start 6:43	6:43–6:50	7
	6:54–7:15	21
	7:18–8:02	44
End 8:43	8:06–8:43	37+
Subject 3		
Start 6:47	6:47–6:54	3
	6:55–7:21	7
	7:22–7:57	11
End 8:44	8:03–8:44	16
Subject 4		
Start 4:52	4:52–4:58	10
	5:08–5:23	15
	5:30–5:54	24
	6:00–6:53	53
	6:58–7:16	18
	7:18–8:16	58
	8:25–9:10	45
End 10:40	9:20–10:40	80+

Table 3
Kismet's turn-taking performance during proto-dialogue with four naive subjects (significant disturbances are small clusters of pauses and interruptions between Kismet and the subject until turn-taking becomes coordinated again)

	Subject 1		Subject 2		Subject 3		Subject 4		Average (%)
	Data	%	Data	%	Data	%	Data	%	
Clean turns	35	83	45	85	38	84	83	78	82.5
Interrupts	4	10	4	7.5	5	11	16	15	10.9
Pauses	3	7	4	7.5	2	4	7	7	6.3
Significant flow distribution	3	7	3	5.7	2	4	7	7	6
Total speaking turns	42		53		45		106		

longer runs of cleanly exchanged turns as time progresses. This suggests that the flow of communication becomes smoother (e.g., fewer interruptions, pauses, and significant flow disturbances) as people read and entrain to Kismet's envelope displays. At this point the rate of vocal exchange is well matched to the robot's perceptual limitations. Table 3 shows that the robot is engaged in a smooth proto-dialog with the human partner the majority of the time (about 82.5%).

5. Discussion

This case study highlights the importance of mutually regulated exchanges, expressive feedback, and readable social cues in the design of sociable robots.

Pro-actively regulate interaction. Kismet takes a pro-active role in regulating its exchanges with people so that it is neither overwhelmed nor under-stimulated—a scenario suitable for learning. It has several different mechanisms to accomplish this each intuitively tunes the person's behavior so that it is appropriate for the robot. Two of these regulatory systems help Kismet to maintain itself in a state of "well-being", namely its drives and emotive responses [7]. In this paper, we presented Kismet's para-linguistic envelope displays—another case in point that helps the robot to modulate the exchange of speaking turns. Clearly to be effective for each case, the robot's expression and behavior must be easily interpreted by the human and well matched to his/her expectations as specified by the mental model the person has for the robot.

Feedback and readable behavior are critical. The importance of feedback and the readability of expression in this process cannot be underestimated. As the human applies the social model to understand Kismet, they are constantly observing the robot's behavior and manner of expression to infer its internal states. This allows the person to predict and understand the robot's behavior only if the robot's expression is readable (the intended signal is appropriately interpreted by the human). The robot's expression reliably maps to the internal state being expressed, and this internal state adheres to the human's mental model for the robot.

This holds for both emotive expression as well as communicative expression. In the case of emotive expression, it is important that the robot's internal model emotions are well matched to their evolutionary counterpart. If so, then a human observer has an intuitive understanding of what makes the robot "angry", "sad", "happy", etc. and can better predict its behavior [7]. Another interesting case is the robot's visual behavior and how it relates to the state of its attention system. A person can infer quite a lot about the robot's internal state by interpreting its gaze and the manner in which it moves its eyes—i.e., what Kismet is interested in or what it is reacting toward [10].

As discussed in this paper, the direction of gaze serves a number of other social functions, such as helping to regulate the exchange of speaking turns. For instance, when the robot broke eye contact during a speaking turn, the person intuitively understood that Kismet was still holding the floor even when not talking. However, if Kismet looked into their eyes during that pause, he/she intuitively understood that the robot was relinquishing its turn to him/her.

Humans are uncannily good at perceiving the direction of another's gaze with surprising accuracy. The ability to establish and maintain eye contact with another is used as a signal to show that you are "in communication" with that person. With Kismet, there was a tremendous difference in a person's sense of social connectedness with the robot when it gazed into his/her eyes rather that just looking at his/her face. It signaled the robot's engagement towards that person, and in turn, it seemed to make the person far more engaged in the robot.

Social interaction is a dance. It is also important to emphasize that it is not just a matter of displaying a readable expression to the person that supports their mental model for the robot. The timing of when and how this is done in relation to the human's behavior is just as important. Social interaction is not just a scheduled exchange of content, it is a fluid dance between the participants. Our interaction studies showcase one of the significant contributions of Kismet's design; namely, the ability to engage people in face-to-face, dynamic, mutually regulated, and closely coupled affective interactions. The resulting interactions (as demonstrated in numerous experiments regarding the communication of affective intent [7,9] and the vocal turn-taking experiments [8]) are quite engaging because the robot's expressive behavior is timely and appropriately synchronized with the human's behavior at fine-grained time scales (i.e., less than a second). Note that in our turn-taking experiments, the human entrains to the robot as well.

This attention to temporal detail and its synchrony with real-time human behavior is critical in order to establish a natural flow and rhythm to the human–robot interaction that is characteristic of human–human interaction. As a result, the interaction is not only stimulating for the robot, but it is also compelling for the person who interacts with it because the robot is "in tune" with them. In short, to offer a high quality (i.e., compelling and engaging) interaction with humans, it is important that the robot not only do the right thing, but also at the right time and in the right manner.

6. Conclusion

Taking this body of work as a whole, we argue that endowing a robot with social skills and capabilities has benefits far beyond the interface value for the person who interacts with it. The ability for robots to interact with people and to leverage from these interactions to perform tasks better, to promote their self-maintenance, and to learn in an environment as complex as that of humans is of tremendous pragmatic and functional importance for the robot. The performance and the benefits that sociable robots bring to us will still need to be evaluated, of course, but from the human's perspective and that of the robot. It is our hope that our experience in this case study (among those described in our other works) will be helpful to others in future efforts to design and evaluate sociable robots.

Table A.1
A short sample of envelope displays during a proto-dialog exchange between Kismet and a human subject (Subject C)

Time code	Speaker			Listner		Turns
	S	Ph	Cue	L	Cue	
07:13:05	H	Aq	Eye contact	R	Eye contact	11
		St	"Did you ask me how I am? I'm fine. How are you?"			
07:14:25		Sp:Rq				
07:17:09	R	Aq	Avert gaze	H		12
07:17:10		St	Babble			
07:18:03		Sp	Eye contact			
07:20:05		Hd	Avert gaze			
07:21:24		Rq	Eye contact			
			Raise brows			
07:22:23	H	Aq		R	Eye contact	13
		St	"Are you speaking another language, Kismet?"	Babble		I
07:24:23		Sp:Rq				14
07:24:06	R	Aq:St	Babble	H		15
07:25:04		Sp	Blink			
		Rq	Elev brows			
07:25:14	H	Aq:St	"Sounds like you're speaking Chinese"	R	Eye contact	16
07:27:10		St:Rq				
07:27:20	R	Aq	Lean forward	H		17
07:27:45		St	Babble			
07:28:03		Sp	Eye contact			
07:28:25		Rq	Elev brows			
07:30:08	H	Aq:St	"Hey!"	R	Avert gaze	18
07:30:15		Sp:Rq	Lean forward		Eye contact	
07:31:08	R	Aq:St	Babble	H	Eye contact	19
07:33:01		Sp	Blink			

Acknowledgements

The development of Kismet was funded by NTT and DARPA contract DABT 63-99-1-0012. The author would like to give special thanks to Sherry Turkle of MIT's Science, Technology, and Society Department for our many useful and enjoyable conversations regarding human relationships with interactive technologies. Medial Lab consortium funds (from the Things that Think and Digital Life consortia) support the author's current research interests.

Appendix A

Table A.1 shows an annotation of a brief portion of a video segment during a proto-dialog between Kismet and a human subject. It uses the annotations as summarized in this table.

References

[1] C. Atkeson, S. Schaal, Robot learning from demonstration, in: Proceedings of the International Conference on Machine Learning (ICML97), Morgan Kaufman, San Francisco, CA, 1997, pp. 12–20.

[2] J. Bates, The role of emotion in believable characters, Communications of the ACM 37 (7) (1994) 122–125.

[3] A. Billard, K. Dautenhahn, Grounding communication in autonomous robots: an experimental study, Robotics and Autonomous Systems 24 (1998) 71–81.

[4] V. Braitenberg, Vehicles: Experiments in Synthetic Psychology, MIT Press, Cambridge, MA, 1984.

[5] C. Breazeal, A motivational system for regulating human–robot interaction, in: Proceedings of the 15th National Conference on Artificial Intelligence (AAAI98), Madison, WI, 1998, pp. 54–61.

[6] C. Breazeal, Designing Sociable Robots, MIT Press, Cambridge, MA, 2002.

[7] C. Breazeal, Emotion and sociable humanoid robots, International Journal of Human Computer Interaction, in press.

[8] C. Breazeal, Regulation and entrainment in human–robot interaction, International Journal of Robotics Research, in press.

[9] C. Breazeal, L. Aryananda, Recognition of affective communicative intent in robot-directed speech, Autonomous Robots 12 (1) (2002) 83–104.

[10] C. Breazeal, P. Fitzpatrick, B. Scassellati, Active vision systems for sociable robots, in: K. Dautenhahn (Ed.), IEEE Transactions on Systems, Man, and Cybernetics: Special Issue Part A, Systems and Humans, 2001.

[11] J. Cassell, Nudge nudge wink wink: elements of face-to-face conversation for embodied conversational agents, in: J. Cassell, J. Sullivan, S. Prevost, E. Churchill (Eds.), Embodied Conversational Agents, MIT Press, Cambridge, MA, 1999, pp. 1–27.

[12] D. Dennett, The Intentional Stance, MIT Press, Cambridge, MA, 1987.

[13] I. Nourbakhsh, J. Bobenage, S. Grange, R. Lutz, R. Meyer, A. Soto, An affective mobile educator with a full-time job, Artificial Intelligence 114 (1–2) (1999) 95–124.

[14] D. Premack, A. Premack, Origins of human social competence, in: M. Gazzaniga (Ed.), The Cognitive Neurosciences, Bradford, New York, 1995, pp. 205–218.

[15] B. Reeves, C. Nass, The Media Equation, CSLI Publications, Stanford, CA, 1996.

[16] B. Scassellati, A theory of mind for a humanoid robot, in: Proceedings of the First IEEE-RAS International Conference on Humanoid Robots, Cambridge, MA, 2000.

Cynthia Breazeal directs the Robotic Presence Group at the MIT Media Lab. She has developed numerous autonomous robots, from planetary micro-rovers, to upper-torso humanoid robots, to highly expressive robotic faces. Always inspired by the behavior of living systems, scientific models and theories as well as artistic insights factor heavily into the hardware and software design of her robotic creations. Her current interests focus on social interaction and socially situated learning between people and life-like robots. Her graduate and postdoctoral research was carried out at the MIT Artificial Intelligence Lab. She received her Sc.D. and S.M. degrees from MIT in the Department of Electrical Engineering and Computer Science with specialization in robotics and artificial intelligence, and a BS in electrical and computer engineering from UC Santa Barbara.