WORKSHOP REPORT

Toward social mechanisms of android science

A CogSci 2005 Workshop
25 and 26 July 2005, Stresa, Italy

Karl F. MacDorman and Hiroshi Ishiguro
School of Informatics, Indiana University, USA / Department of Adaptive Machine Systems, Osaka University, Japan

Toward Social Mechanisms of Android Science was the first workshop at which researchers from various disciplines joined to discuss the application of very humanlike robots to the study of interaction and cognition and the social impact of this technology. It was the only workshop of the 27th Annual Conference of the Cognitive Science Society. The main conference was held at the Convention Centre, in Stresa, Italy on the 21st through to the 23rd of July. On the 25th and 26th, the Android Science workshop took place next door, at the Regina Palace Hotel on picturesque Lake Maggiore.

For the purposes of android science, an android is an artificial system designed with the ultimate goal of being indistinguishable from humans in its external appearance and behavior. The original premise of the workshop is that androids that look and act like people can elicit from human subjects a range of social responses that only other people had been previously able to elicit (Ishiguro, 2005; MacDorman et al., 2005). This ability to elicit social responses enables androids to provide not just a well-controlled experimental apparatus for studying human interaction but a testbed for developing theories about how neural or cognitive processes influence interaction.

Android development extends beyond the scope of engineering because, to make androids humanlike, it is necessary to investigate human interaction, and to evaluate theories of human interaction accurately, the theories need to be implemented in androids (Minato et al., 2004). Since mechanisms supporting human interaction come into play when people relate to machines, it is essential to examine their human likeness.

Interaction Studies 7:2 (2006), **289–296**.
ISSN 1572–0373 / E-ISSN 1572–0381

The workshop laid a foundation for research in android science, a new interdisciplinary framework that integrates the incremental development of robots with the empirical methodologies of the social sciences (Ishiguro, 2005). Given the workshop participants' range of interests, it was thought android science should be construed more broadly to include all the effects of engineered human likeness. Thus, android science studies the significance of human likeness in human-machine relationships by exploiting the tools of science and engineering. This in turn leads to a better understanding of what it means to be human.

The workshop organizers, Stephen Cowley, Kerstin Dautenhahn, Hiroshi Ishiguro, and Karl MacDorman, received 25 submissions, from which 15 papers were accepted. With three invited talks, this made for a total of 18 talks. The disciplines of computer science, engineering, psychology, neuroscience, philosophy, and ethnography were represented among the presenting authors and participants, although most engaged in research that was clearly multidisciplinary. Thus, the workshop attracted the broad range of expertise called for by android science.

Hiroshi Ishiguro kicked off the workshop with a plenary talk proposing android science as a new cross-disciplinary framework in which roboticists develop humanlike robots based on insights gleaned from cognitive science, while cognitive scientists use these robots in experiments to better understand human beings (Ishiguro, 2005). The talk charted his laboratory's development of three androids in collaboration with Kokoro Co., Ltd.: Repliee R1, Q1, and Q1Expo. Repliee Q1Expo is driven by 42 air actuators, and can make eye movements, facial expressions, gestures, and head, neck, and torso movements. Because the air actuators are compliant and nearly silent, Repliee's movements can seem quite natural. Ishiguro discussed a number of methods being used in his laboratory to implement humanlike motions, including one in which markers on a performer's body surfaces are mapped to the android; the android learns a feedforward controller using a neural network, and human and android motion similarity are evaluated by a motion capture system (Matsui et al., 2005).

Ishiguro demonstrated the importance of nonconscious and autonomic movement in an experiment in which a participant is seated three meters from the android and a blind is opened for two seconds. In the control group, the android is still, but in the experimental group it is moving. Seventy percent of participants in the study believed the moving android to be human, but only thirty percent believed the still android to be human. More surprisingly, some people who knew about the android from television or the Internet were still fooled by it. Ishiguro also presented two experiments that showed how eye contact and gaze are influenced by a person's cognitive states. The first experiment compares gaze behavior under cognitive load in human-human and human-android interactions (MacDorman et al., 2005). The second experiment compares gaze behavior while telling lies (Minato et al., 2005).

The interrelatedness of people's thought processes and how they orchestrate their bodies was also elucidated in a talk by Morana Alač. Ishiguro pointed out that android science needs an adequate definition and understanding of human communication. To tackle this Alač analyzed video-taped records of interactions among scientists, Repliee Q1Expo, and other technologies involved in the process of designing the android's movements. Observations of minute aspects of practice reveal that, to solve the task they are facing, scientists engage their bodies in the process of design. They define the problem and search for solutions by moving, observing, and touching their own bodies. Importantly, to be meaningful, these movements are always social in character: they reveal processes of identification and human-machine symbiosis at the level of multimodal interaction. Alač's analysis suggests that thinking/communicating is an embodied, intersubjective, and dynamical process in which meaning is distributed, shared, and negotiated through coordination among social actors and their material world of practice. Her approach, based on long-term observation and multimodal microanalysis of interaction (Alač, 2005), combines insights drawn from distributed cognition (Hutchins, 1995; Cowley & Spurrett, 2003), conceptual integration theory, and a practice-based theory of knowledge and action (Goodwin, 1994).

In a session on behavioral studies of interaction, Michael Walters presented experimental work comparing individuals' comfortable approach distances with a mechanical-looking robot to previous experiments with people only (Walters et al., 2005). He and his colleagues found that the approach distances of sixty percent of participants are compatible with interpersonal norms; however, forty percent took up significantly closer approach distances, suggesting that they were not treating the robot as a social entity. Upon assessment of the participants' personalities, those who scored higher on a *proactiveness* factor approached less closely to the robot. In the same session, Noriko Suzuki presented results showing how Japanese subjects align their speech to computer-generated speech along such prosodic dimensions as volume and switching pause duration (Suzuki & Katagiri, 2005). Michael Brady (2005) demonstrated an analog vocal tract for a humanoid robot head that he developed. Participants in an experiment rated characters' facial expressions along various affective dimensions and their closeness of fit to vocalizations of the robot.

Perceived arousal was correlated with glottal open quotient and valence with fundamental frequency.

In a session on brain processes during interaction, Boris Velichkovsky presented functional magnetic resonance imaging (fMRI) and electromyogram (EMG) results showing similar brain activity regardless of whether a virtual character smiled at the subject or at an out-of-view "other" despite significantly reduced eye contact in the latter case (Schilbach et al., 2005). Therefore, an analysis of eye movements was conducted revealing differentiating effects on a finer timescale. The study points to an apparent discrepancy between visual attention, which depended on the observer's level of involvement, and its neurophysiological correlates.

In the session on implemented social mechanisms, Brian Duffy gave a presentation on the social robot architecture, which encompasses a design methodology that combines hardware abstraction, a synthesis of reactive and deliberative control, and explicit social interaction (Duffy et al., 2005). The social robot architecture supports soccer play between Nomad Scout II robots and children, a waltz between a physical and a virtual robot using the Virtual Robotic Workbench, and human-robot interaction with the humanoid "Joe." In the same session, Andrea Thomaz presented a computational model of social referencing (Thomaz et al., 2005). The model has been embedded in the robot Leonardo at the MIT Media Lab's Robotic Life Group, directed by Cynthia Breazeal. The model is inspired by how human infants take into account the emotional reactions of others in learning to appraise situations, and consists of systems that support empathic imitation of facial expressions, shared attention, and affective memory. This research is meant to lead to a deeper understanding of infant social development, but it is also an important step toward the development of robots that can commune with people, share in their culture, and develop relationships.

A major theme concerned how to develop androids that could at least act as if they were moral and empathetic agents. Wallach and Allen pointed out that the presence of androids in society will depend in part on their public acceptance: people will need to believe that androids will not harm them but respect their norms and values. So the question arises how to engineer ethics into android design, either by conformance to ethical theories or a more task-specific adherence to goals and standards (Wallach & Allen, 2005). But Calverley (2005) expressed concern about the morality of constructing such a moral being. Christopher Ramey's talk made an important point about the uniqueness of androids as a technology in their close resemblance to human beings. To treat a humanlike creation as dispensable has the potential to dehumanize our relations with each other, since social cognition and behavior are rooted in the fact that we exist as beings for the sake of others (Ramey, 2005). However, Billy Lee pointed out that an android must also be a being for its own sake, one that exists for its own reasons and purposes. Otherwise, its emotional responses, however humanlike, would seem inauthentic.

Billy Lee gave a talk that reported a lack of a significant relationship between perceptual performance at detecting lies and empathy, intimacy, and felicity (Lee, 2005). The results, nevertheless, showed that it is easier for women to give and receive comfort, suggesting that endowing robots with a feminine appearance can be useful for entraining people into closer relationships with them. Nevertheless, in her studies in two nursing homes, Sherry Turkle found that even such nonhuman robots as Paro, which is modeled on a baby seal, can elicit relational, nurturing responses (Taggert et al., 2005; Turkle, 2005). Such devices may call for a new psychology involving the relations among people and their animate artifacts.

Turkle's paper presented the case of an elderly woman who felt Paro had understood her feeling of having been abandoned by her son. The workshop participants expressed a diversity of opinions on this incident, ranging from praise for the clearly therapeutic role of robot in comfort giving to concern that the woman's dignity had been compromised by a situation that allowed her to project empathy on to a device that understood nothing of its role in this encounter. Had the woman shared this moment of intimacy with a friend or counselor, there would have been no ethical issue as the outcome was clearly positive. But should the woman's feelings be the only consideration, when her experience itself lacked authenticity? The absence of consensus among workshop participants revealed the breadth of a debate that is only likely to grow as robot technology advances and becomes more pervasive.

Susan Schneider argued that relevance determination, also known as Fodor's version of the frame problem (Fodor, 2000), does not impact the development of androids by precluding a computational model of human reasoning (Schneider, 2003). However, in a session on embodiment and social learning, Jessica Lindblom argued that intentional agency should emerge from embodied cognitive development (Lindblom & Ziemke, 2005). To make an android mind humanlike, it is not enough to construct a humanlike robot body and endow it with cognitive capacities — or even to let it develop those capacities on its own — because embodied cognition is shaped by the experience of developing within the physical and social realm through the body. She discussed how a child's development of joint attention and the self may spring from the interplay between mimesis and self-locomotion.

For androids to be integrated into social roles, it is important for them to look human, just as it is important for people to look human. Since partial facial paralysis owing to such diseases as Parkinson's can cause even the most gregarious of people to be shunned as sullen introverts (Cole, 2001), androids need sufficient facial expressiveness to avoid the same fate. But expressiveness beyond human norms becomes increasingly disturbing with the realism of human simulation (Vinayagamoorthy et al., 2005). At this point, the android may be running against what Masahiro Mori identified as the *uncanny valley*, the apparent tendency of machines to seem more familiar as they appear more humanlike until a point is reached at which subtle defects appear unnerving. Karl MacDorman explored in his talk the possibility that androids, when presented in states of imperfect or incomplete assembly, seem eerie because they are reminders of our own mortality. Thus, one would feel like a defunct android, coming apart at the seams (MacDorman, 2005). Hiroshi Ishiguro closed the workshop by summarizing the conclusions from our discussions throughout the event.

The workshop stimulated dialogue among researchers from the relevant disciplines on how the experimental application of androids can deepen our understanding of human beings, their thoughts and interactions, and their underlying mechanisms. It also initiated collaborative projects among some of the authors, organizers, and program committee members. The growth of android science depends on greater access to this platform in the social and neurosciences, and so in future workshops we will invite the participation of more people who engineer androids.

Online versions of the workshop contributions are available at http://www.androidscience.com/.

Mori's uncanny valley is scheduled for more detailed discussion at the Humanoids 2005 workshop "Views on the Uncanny Valley" on December 5, 2005, in Tsukuba, Japan. In addition, a second Android Science workshop is planned for the 27th Annual Conference of the Cognitive Science Society on July 26, 2006 in Vancouver, Canada.

References

Alač, M. (2005). From trash to treasure: Learning about the brain through multimodality, *Semiotica, 156*(1/4), 177–202.

Brady, M.C. (2005). Synthesizing Affect with an Analog Vocal Tract: Glottal Source. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 45–49.

Calverley, D.J. (2005). Android Science and the Animal Right's Movement: Are There Analogies. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 127–136.

Cole, J. (2001). Empathy needs a face. *Journal of Consciousness Studies, 8*(5–7), pp. 51–68.

Cowley, S. & Spurrett, D. (2003). Putting apes (body and language) together again. *Language Sciences, 25*(3), 289–318.

Duffy, B. R., Dragone, M., & O'Hare, M.P. (2005). The Social Robot Architecture: A Framework for Explicit Social Interaction. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 18–28.

Fodor, J. (2000). *The mind doesn't work that way: The scope and limits of computational psychology.* London: MIT Press.

Goodwin (1994). Professional vision. *American Anthropologist, 96*(3), 606–633.

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, Mass.: MIT Press.

Ishiguro, H. (2005). Android Science: Toward a New Cross-Disciplinary Framework. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 1–6.

Lee, B. (2005). Interpersonal Perception and Android Design. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 50–55.

Lindblom, J. & Ziemke, T., "Body-in-Motion" and Social Scaffolding: Implications for Human and Android Cognitive Development. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 87–95.

MacDorman, K. (2005). Androids as an Experimental Apparatus: Why is There an Uncanny Valley and Can We Exploit it? *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 106–118.

MacDorman, K.F., Minato, T., Shimada, M., Itakura, S., Cowley, S., and Ishiguro, H. (2005). Assessing human likeness by eye contact in an android testbed. *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*. Stresa, 21–23 July.

Matsui, D., Minato, T., MacDorman, K. & Ishiguro, H. (2005). Generating natural motion in an android by mapping human motion. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. August 2–6, 2005. Edmonton, Canada.

Minato, T., Shimada, M., Ishiguro, H. and Itakura, S. (2004). Development of an Android Robot for Studying Human-Robot Interaction, Innovations in Applied Artificial Intelligence; *Proc. of the Seventeenth International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE),* Ottawa, Canada, May 17–20, 2004, pp. 424–434.

Ramey, C. H. (2005). "For the Sake of Others": The "Personal" Ethics of Human-Android Interaction. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 137–148.

Schneider, S. (2003). Fodor's version of AI's frame problem (the globality problem): A solution. *Proceedings of the American Philosophical Association, Pacific Division* (Vol. 76, Issue 3). San Francisco, March 26–30, 2003.

Schilbach, L., Vogeley, K., Helmert, J. R., Pannasch, S., Velichkovsky, B. M. & Mojzisch, A. (2005). Visual Attention, Facial Expression, and Brain Processing while Meeting a Virtual Other. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 74–86.

Suzuki N. & Katagiri, K. (2005). Prosodic Adaptation in Human-Computer Interaction. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 38–44.

Taggart, W., Turkle, S., & Kidd, C. D. (2005). An Interactive Robot in a Nursing Home: Preliminary Remarks. *Toward Social Mechanisms of Android Science: A CogSci 2005 Worksho*p. July 25–26, Stresa, Italy, pp. 56–61.

Thomaz, A. L., Berlin, M. & Breazeal, C. (2005). Robot Science Meets Social Science: Embodied Computational Model of Social Referencing. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 7–17.

Turkle, S., (2005). Relational Artifacts, Children, and Elders: The Complexities of Cyber-Companions. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 62–73.

Vinayagamoorthy, V. Steed, A. & Slater, M. (2005). Building Characters: Lessons Drawn from Virtual Environments. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop.* July 25–26, Stresa, Italy, pp. 119–126.

Wallach, W. & Allen, C. (2005). Android Ethics: Bottom-up and Top-down Approaches for Modeling Human Moral Faculties. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 149–159.

Walters, M.L., Dautenhahn, K., Koay, K.L., Kaouri, C., Woods, S., Nehaniv, C., te Boekhorst, R., Lee, D., Werry, I. (2005). The Influence of Subjects' Personality Traits on Predicting Comfortable Human-Robot Approach Distances. *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. July 25–26, Stresa, Italy, pp. 29–37.