Social Robots as Persuasive Agents

Evgenios Vlachos and Henrik Schärfe

Aalborg University, Department of Communication and Psychology,
Nyhavnsgade 14, 9000, Denmark
{evlachos,scharfe}@hum.aau.dk

Abstract. The topic of human robot interaction (HRI) is an important part of human computer interaction (HCI). Robots are more and more used in a social context, and in this paper we try to formulate a research agenda concerning ethical issues around social HRI in order to be prepared for future scenarios where robots may be a naturally integrated part of human society. We outline different paradigms to describe the role of social robots in communication processes with humans, and connect HRI with the topic of persuasive technology in health care, to critically reflect the potential benefits of using social robots as persuasive agents. The ability of a robotic system to conform to the demands (behaviors, understanding, roles, and tasks) that arise from the place the robot is designed to perform, affect the user and his/er sense of place attachment. Places are constantly changing, and so do interactions, thus robotic systems should continually adjust to change by modifying their behavior accordingly.

Keywords: human-robot interaction, persuasive agent, social robots, ethics, place attachment.

1   Introduction

Until very recently, robots were limited to industrial environments, and research facilities. Only lately did they migrate to our daily life, and became more social, user friendly, communicative and interactive. In most of the cases of social HRI the user is not able to distinguish clearly the entities that are embodied within the robot when interacting with it. Is the robot completely autonomous and self-oriented, or is it semi-autonomous? Is it controlled by a human, by a team of humans, or by other robots? What kind of information is the robot storing? Why and for what purpose? Does it share information with third parties, and who are they? A precondition needed for establishing a trusting relationship between the user, and the robotic system prior to interaction, is to have an answer for each of the risen questions above, meaning that the robotic system should be completely overt. For that reason, HRI could borrow the Code of Ethics from Information and Communication Technology (ICT) called PAPA, acronym of privacy, accuracy, intellectual property, and access [1]. Indeed, HRI and ICT share a plethora of common ethical issues; however, a robot's physical representation is a decisive factor in the argument in favor of determining new ethics for robots, the robo-ethics.

The majority of the published research findings in HRI deal with the target group of children and elderly people. On the one hand, there is a growing body of research presenting fruitful interactions between children and robots in the home, and in the classroom, specifically when the subject matter is related to science and engineering [2, 3]. Robots have also been shown to have a positive outcome in therapeutic applications for children [4]. On the other hand, according to the Population Division of the United Nations the population ageing is unprecedented, pervasive, enduring, and has profound implications for many facets of human life [5]. Therefore, ageing population is expected to need physical and cognitive assistance. Moreover, space and staff shortages at health care facilities are already an issue. Assistive robotic systems and companion robots for the elderly could be a solution provided that technologies are capable of being commanded through natural communication (e.g., facial expressions, speech, non-verbal communication, gestures,), of grasping and lifting items, and of assisting with daily chores and tasks (e.g., moving, feeding) if they are to improve the physiological and psychological health of the ageing population.

In the following sections we will discuss the notion of social robotics, present the key principles of persuasive technology, explain how the displayed behavior of a robot can affect the requirements for place attachment, and finally investigate the relationships between a user, a robot, and the robot's operator in a HRI scenario taking place in a health facility.

2   Social Robotics

Robots that are able to interact and communicate with humans in a human-like manner, but also with other robots, as well as with their environment, respecting the existing social and cultural norms are called social robots [6]. When interacting with such robots we apply social rules, and act on inherited behavioral guidelines, expecting that the robots will have the ability to understand, and follow them. Notable instances of social robots around us are the Geminoid-DK android when it took up the role of a university lecturer, or the role of a business man offering deals in an office [7], the Geminoid-F android when it performed in theatrical plays around the world [8], the Telenoid teleoperated humanoid when used for facilitating communication with elderly people suffering from dementia [9], the Kaspar humanoid robot when fostering cooperative dyadic play among children with autism [10], the Rofina teleoperated robot when it helped children with special needs to understand play behaviors [11], and the Zeno child-like humanoid when it assisted physical therapists to treat sensor-motor impairments [12]. Several researchers have also explored interactions with zoomorphic robots like the robot dog AIBO that uses body language and simple musical melodies to communicate with people [13], the robotic creature Kismet that engages physically, affectively, and socially with humans so as to learn from them [14], the seal robot Paro when used to improve the lives of elderly dementia patients [15], the Nabatzag rabbit that augments audio messages with display of non-verbal expressions [16], or the robotic cat NeCoRo whose behavior depends on the history of its interactions and can recognize its name [17].

Robots were not initially created to deceive, but to be trusted. After all, the term robot comes from the Czech word "*robota*" which means forced labor, or servitude and firstly appeared in the play R.U.R. (Rossum's Universal Robots) by the author Karel Čapek in 1921 [18]. Their purpose was, and still is, to serve the human either by handling situations, data, or by dealing with various tasks.

3   Persuasive Technology

If we assume that social robots bare strong similarities to traditional media, they should be defined as a medium that connects users to a source of a message. Under the view of computer-as-medium paradigm [19], the robot is the mediator of communication between the users, and the robot programmer or/and the robot operator. Therefore, HRI can be considered as human-human interaction. On the contrary, paraphrasing the computer-as-source unmediated perspective users respond to social robots as a source of information by following unintentionally the Media Equation formula (convenience to perceive robots as humans) [20]. There is strong evidence that in HCI users communicate directly with the computer, and not with a vague persona of a programmer behind it. If we apply these results to the area of HRI, then users should relate directly to the social robot, and not to the person behind it, either this person holds the position of a programmer, a designer, an operator, or embodies a whole organization, or a brand. But, is that the case? Ambient intelligence, multi-Intelligence (draw on multiple sources of intelligence, including big data, cloud and crowd resources), and networked robotics (share sensory input, solutions and problems across many locations and application areas), are three popular research topics among roboticists that enable robots to be more than mediating artifacts. Hence, robots include in their definition and the other two aspects of Fogg's Functional Triad [21] namely the notion of social actor, and the notion of tool. Robots, and especially social robots, encompass much more qualities than a computer does, and since they are a relatively new field of technology, people have not yet conceptualized their full range of abilities. In order to minimize the amount of false information such robots transmit, either intentionally for the greater good of the mixed initiative team comprised of the user, the robot and its operator, or unintentionally due to the effects of the Media Equation formula, the user should in advance be informed of what the robotic system is capable of doing, and equally important of not doing.

According to the Greek philosopher Aristotle persuasion was the art of convincing people to accept something, or do something they would normally not otherwise. The three modes of persuasion introduced by Aristotle are Ethos (ethical character of the source of information), Pathos (emotional state of the receiver), and Logos (argument) [22]. For a persuasive message a blend of all three is needed. The definition of persuasive technology (PT) includes robotic systems that are *"designed to change people's attitudes or behaviors or both without using coercion or deception"* [21]. Gass et al. [23] proposed that "*persuasion involves one or more persons who are engaged in the activity of creating, reinforcing, modifying, or extinguishing beliefs,*

attitudes, intentions, motivations, and/or behaviors within the constraints of a given communication context". The ultimate goal of PT is to promote wellbeing, health, quality of life, and a more sustainable lifestyle, but requires awareness of the user that it is an intentional act, and at all times he/she has the choice to decline. Trust in robotic systems directly influences both the interactions, and the overall acceptance of robotic agents. A user's trust, or distrust, towards a robot is expanded towards the entities that are embodied in the robot, which might be the programmer, the operator, the organization, or company the robot is located in, and even the brand of the robot (e.g., Honda Motor Co. in the case of Asimo, Kokoro Co. ltd in the case of the Geminoids, and Hanson Robotics in the case of Zeno).

What nearly all of the social robots have in common is the –most of the times-false message they transmit concerning two features that make their character being perceived more believable, and encourage interaction; the freedom of their actions, and their degree of autonomy. In [24], due to false attribution of robot capabilities, the children were expecting the robot to play along with them, while researchers were expecting the children to play along with the robot. Hiding or showing false information can be regarded as manipulation of the truth. This kind of deceit takes unintentional advantage of the effects of the Media Equation [20] and tricks the human mind by letting it treat machines in the same way as towards other people.

4   Place Attachment

"*The structure of the space around us moulds and guides our actions and interactions*", S. Harrison and P. Dourish [25]. The place a user is located frames his/er behavior, and automatically creates a mental icon concerning both the properties of the robotic system, and the type of HRI that would take place in case a robot was present. Therefore, even before real time HRI occurs, the user might have already categorized the expected-to-be-there robot by following unintentionally a robotic version of the HCI Paradigms; robot categorized as a tool (extending the abilities, strength, competence, intelligence of the human), robot categorized as a medium -or avatar [6]- (being a mediator of interpersonal communication and intentionality), robot categorized as a partner (embodying anthropomorphic features, humanlike properties, behavioral characteristics, and emotional/mental states) [26-28]. It seems that the HCI paradigms are identical to the Fogg's Functional Triad that was discussed earlier.

Places are constantly changing as they are continuously enacted by people [29]. Being part of the material topography of a place, the robotic system should be readjusted according to these changes either by being reprogrammed manually, or by being able to detect them through its sensory input and modify its behavior. The character of HRI can only be understood, and thereafter evaluated, if linked to a place. Thus, the evaluation criteria for a competent robotic system are limited to only one; how well it conforms to the user's perspective, meaning how well it fits the place. This criterion can be subdivided into smaller segments that represent each task, and

each task can also be decomposed to smaller components in relation to the objectives of the task.

While place plays an essential role in human life, it is equally important in robotics, and often takes precedence over all other aspects of HRI. Place attachment is defined as the bond a person develops for a place that "*evolves from specifiable conditions of place and characteristics of people*" [30]. Extending that definition towards the field of HRI we suggest that robotic behavior should be added to the key factors that affect place attachment. Social robots are structurally coupled with their operational environment, and are connected to it with channels of mutual perturbation [31].

Hence, negative attitudes toward a robot, might lead to negative attitudes toward the people, the organization, the brand, or the company the robot embodies. We hope our statement to stimulate further research in order to avoid being confronted with the phenomenon of place aversion (including brand, company, and organization aversion) due to prejudice against interacting with robots in the near future.

5   Ethical Concerns in a HRI Scenario

Let us consider the scenario where a hospital makes use of a toy-robot companion for hospitalized children to play with. The robot is monitoring the child, observing every move, collecting personal information and, sending them to a system supervised by an operator. Figure 1 depicts the interactions between the user, the robot, and the robot's operator. The default situation would be the autonomous circle, where the
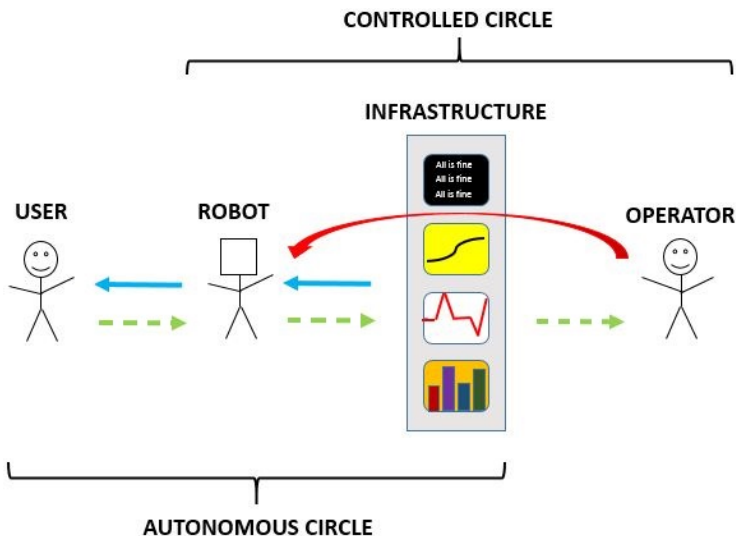
Fig. 1. Interactions between the user, the robot, and the operator featuring the Autonomous Circle (operator does not intervene in the HRI) and the Controlled Circle (operator intervenes in the HRI)

robotic system supervises and communicates with the user without the intervention of an operator. The operator would only override the autonomous circle in case of emergency, and take control. The toy-robot holds three roles; (i) is the mediator of communication between the operator (the doctors, and the hospital;) and the hospitalized child, (ii) is a tool measuring temperature, pulse rate, blood pressure, and whatever else is needed according to the situation, and (iii) serves as a companion partner to the child, where the child is speaking to, sharing personal stories, and maybe information that even his/her parents are not aware of. The toy-robot exhibits all of the persuasive characteristics in Fogg's Functial Triad, but the cognitive and/or mental state of the child might not be sufficient to understand the roles of the involved stakeholders, including that of the robot. Is a consensus between the parents and the hospital about using a toy-robot enough? The toy-robot should neither be used as a justification for leaving the child on its own for longer time since it could lead to malformed development and emotional problems [2], nor as an excuse to migrate the responsibility from humans to robots [32]. One can imagine the same scenario with people suffering from dementia.

Another ethical concern is the liability of the robotic system. In case of a malfunction we have to be consistent with where to place responsibility. It could be the ethics of the operator, the ethics of the designer, the embedded ethical system of the robot, or the ethics of the user. According to [33], the wisest decision is to either avoid blaming anyone, or blame everyone. To prevent such a dramatic turn of events from happening, the stakeholders (user, operator, organization) should form a mixed initiative team having one common goal aligned and oriented towards one direction; to **protect the user, and secure his/er interests**.

6   Conclusion

Throughout this paper we focused on ethical concerns raised when humans communicate with robots. We posed questions regarding the privacy, accuracy, intellectual property, access, and liability of a robotic system aiming to formulate an ethical research agenda for issues related to human robot social interaction. In a brief overview we have linked the notion of social robotics to that of persuasive agents, and proposed that social robots could be more than just mediating artifacts. Social robots could also act as tools, and social actors, and thus have all the characteristics in Fogg's Functional Triad. The ability of a robotic system to adjust to the behaviors, understanding, roles, and tasks that arise from the place the robot is designed to perform, affects the user and his/er degree of place attachment, the bond a person develops for a place.

An aligned perspective among the stakeholders of HRIs through the formation of a mixed initiative team could be the first step towards ensuring that robots do actually benefit, and protect the users, and are not just designed to alleviate guilt from parental personalities, or reduce operational costs of an organization. In spite of the fact that robotic systems are designed by humans, and have more human values inherited than expected, the engagement of a human operator is a reassuring act indicating that robots are not here to substitute us, but to assist us.

References

1. Veruggio, G., Operto, F.: Roboethics: a bottom-up interdisciplinary discourse in the field of applied ethics in robotics. International Review of Information Ethics 6, 2–8 (2006)
2. Sharkey, N., Sharkey, A.: The crying shame of robot nannies: an ethical appraisal. Interaction Studies 11(2), 161–190 (2010)
3. Kanda, T., Nishio, S., Ishiguro, H., Hagita, N.: Interactive Humanoid Robots and Androids in Children's Lives. Children, Youth and Environments 19(1), 12–33 (2009)
4. Liu, C., Conn, K., Sarkar, N., Stone, W.: Online affect detection and robot behaviour adaptation for intervention of children with autism. IEEE Transactions on Robotics 24(4), 883–896 (2008)
5. World Population Ageing: 1950-2050, `http://www.un.org/esa/population/publications/worldageing19502050`
6. Breazeal, C.: Social interactions in HRI: the robot view. IEEE Transactions on Systems, Man, and Cybernetics 34(2), 181–186 (2004)
7. Vlachos, E., Schärfe, H.: The Geminoid Reality. In: Stephanidis, C. (ed.) HCII 2013, Part II. CCIS, vol. 374, pp. 621–625. Springer, Heidelberg (2013)
8. Paré, Z.: Robot Drama Research: From Identification to Synchronization. In: Ge, S.S., Khatib, O., Cabibihan, J.-J., Simmons, R., Williams, M.-A. (eds.) ICSR 2012. LNCS (LNAI), vol. 7621, pp. 308–316. Springer, Heidelberg (2012)
9. Yamazaki, R., Nishio, S., Ogawa, K., Ishiguro, H.: Teleoperated android as an embodied communication medium: A case study with demented elderlies in a care facility. In: RO-MAN 2012, pp. 1066–1071 (2012)
10. Wainer, J., Dautenhahn, K., Robins, B., Amirabdollahian, F.: Collaborating with Kaspar: Using an autonomous humanoid robot to foster cooperative dyadic play among children with autism. In: IEEE-RAS 2010 Inter. Con. on Humanoid Robots, pp. 631–638. IEEE Press (2010)
11. Wong, A., Tan, Y.K., Tay, A., Wong, A., Limbu, D.K., Dung, T.A., Chua, Y., Yow, A.P.: A User Trial Study to Understand Play Behaviors of Autistic Children Using a Social Robot. In: Ge, S.S., Khatib, O., Cabibihan, J.-J., Simmons, R., Williams, M.-A. (eds.) ICSR 2012. LNCS (LNAI), vol. 7621, pp. 76–85. Springer, Heidelberg (2012)
12. Ranatunga, I., Rajruangrabin, J., Popa, D.O., Makedon, F.: Enhanced therapeutic interactivity using social robot Zeno. In: Proc. of the 4th International Conference on Pervasive Technologies Related to Assistive Environments, pp. 57–62. ACM (2011)
13. Kaplan, F.: Talking AIBO: First experimentation of verbal interactions with an autonomous four-legged robot. In: Learning to behave: interacting agents CELE-TWENTE Workshop on Language Technology, pp. 57–63 (2000)
14. Breazeal, C.: Toward sociable robots. Robotics and Autonomous Systems 42(3), 167–175 (2003)
15. Calo, C.J., Hunt-Bull, N., Lewis, L., Metzler, T.: Ethical Implications of Using the Paro Robot. In: 2011 AAAI Workshop (WS-2011-2012), pp. 20–24 (2011)
16. Rist, T., Wendzel, S., Masoodian, M., André, E.: Creating awareness for efficient energy use in smart homes. In: Intelligent Wohnen. Zusammenfassung der Beiträge zum Usability Day IX, pp. 162–168 (2011)

17. Libin, A., Cohen-Mansfield, J.: Therapeutic robocat for nursing home residents with dementia: preliminary inquiry. American Journal of Alzheimer's Disease and Other Dementias 19(2), 111–116 (2004)
18. Gupta, A.K., Arora, S.K.: Industrial automation and robotics. Laxmi Publications (2007)
19. Sundar, S.S., Nass, C.: Source Orientation in Human-Computer Interaction Programmer, Networker, or Independent Social Actor. Communication Research 27(6), 683–703 (2000)
20. Reeves, B., Nass, C.: The Media Equations: How People Treat Computers, Television, and New Media Like Real People and Places. CLSI Publications (2002)
21. Fogg, B.J.: Persuasive Technology: Using Computers to Change what We Think and Do. Morgan Kaufmann (2003)
22. Aristotle's Ethics, `http://plato.stanford.edu/entries/aristotle-rhetoric`
23. Gass, R.H., Seiter, J.S.: Persuasion, social influence, and compliance gaining, 2nd edn., p. 34. Allyn & Bacon/Longman, Boston (2003)
24. Strommen, E.: When the interface is a talking dinosaur: learning across media with ActiMates Barney. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 288–295. ACM Press/Addison-Wesley Publishing Co. (1998)
25. Harrison, S., Dourish, P.: Re-place-ing space: the roles of place and space in collaborative systems. In: Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work, pp. 67–76. ACM, New York (1996)
26. Beaudouin-Lafon, M.: Designing Interaction, not Interfaces. In: Proceedings of the Working Conference on Advanced Visual Interfaces, pp. 15–22. ACM (2004)
27. Kuzuoka, H., Yamazaki, K., Yamazaki, A., Kosaka, J., Suga, Y., Heath, C.: Dual Ecologies of Robot as Communication Media: Thoughts on Coordinating Orientations and Projectability. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2004), vol. 6(1), pp. 183–190 (2004)
28. Epley, N., Waytz, A., Cacioppo, J.T.: On Seeing Human: A Three-Factor Theory of Anthropomorphism. Psychological Review 114(4), 864–886 (2007)
29. Cresswell, T.: Place. Blackwell Pub. (2004)
30. Shumaker, S.A., Taylor, R.B.: Toward a clarification of people-place relationships: A model of attachment to place. In: Feimer, N.R., Geller, E.S. (eds.) Environmental Psychology: Directions and Perspectives, pp. 219–251. Praeger, New York (1983)
31. Dautenhahn, K., Ogden, B., Quick, T.: From embodied to socially embedded agents–implications for interaction-aware robots. Cognitive Systems Research 3(3), 397–428 (2002)
32. Sharkey, A., Sharkey, N.: Granny and the robots: ethical issues in robot care for the elderly. Ethics and Information Technology 14(1), 27–40 (2012)
33. Reason, J.T.: Human error. Cambridge University Press (1990)