ON ELECTRIC OWLS
Implicit life-stories of robots and their impact on human empathy

PALASH NANDY

B.S. in Computer Science, Stanford University, 2003

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology.

JUNE 2016

Author:
Palash Nandy
Program in Media Arts and Sciences
May 6, 2016

Certified by:
Cynthia Breazeal
Associate Professor of Media Arts and Sciences
Thesis Supervisor

Accepted by:
Pattie Maes
Academic Head
Program in Media Arts and Sciences

ON ELECTRIC OWLS
Implicit life-stories of robots and their impact on human empathy

PALASH NANDY

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, on May 6, 2016 in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology.

ABSTRACT

Robots are moving from factories to people's homes taking on the roles of artificial pets, tutors and companions. If we are to have emotionally engaging robots, we must understand how we can design robots that people feel empathy towards. In this work, I explore one design criteria for such robots: implicit life stories or the ability for a robot to experience the world we share, be transformed by that experience and communicate that experience to us. Through the construction of novel robots that can have implicit life-stories, and through human subject studies I show that such robots can evoke empathy. I also show that empathy for robots can impact empathy for other human beings.

Thesis Supervisor:
Professor Cynthia Breazeal
Associate Professor of Media Arts and Sciences
Program in Media Arts and Sciences

The following people served as readers for this thesis:

Thesis Reader:
Edith Ackermann
Professor of Developmental Psychology
University of Aix-Marseille
Research Affiliate
MIT School of Architecture

Thesis Reader:
Tod Machover
Muriel R. Cooper Professor of Music and Media
MIT Media Lab

ACKNOWLEDGMENTS

I am grateful to Cynthia for giving me the opportunity to pursue this work, and for her insights and feedback on the research. I would also like to thank my dear readers, Tod and Edith for the wonderful discussions that shaped my thinking.

I have been fortunate to be part of the Personal Robots Group with its all star crew. I will miss Polly's love, hugs and punches. I want to thank my fellow investigator, Jin Joo, for welcoming me and showing me the ropes, and of course my research partner, Kate, for being a bad influence and introducing me to violence against robots. I am grateful to the amazing post-docs, Brad, Hae Won, and especially Goren for all the crazy ideas and feedback.

My family has always encouraged me to pursue my curiosity. I am thankful for my parents who taught me to think, and my sister who always has my back. I want to thank Gabi. Without her love and support for my dreams, I wouldn't have been here. A huge thanks to Audra for keeping me sane. Last but not the least, Sylvia for all the love and the inspiration.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Figure 1: Death of a replicant in the movie *Blade Runner*

What makes us human? This question is explored in the science fiction movie *Blade Runner*. Set in a future-noir world against a backdrop of neon lit skyscrapers, we take on the persepctive of Deckard who is tasked with hunting down and killing replicants, an organic artificial humanoid. The replicants are assumed to be not human, a premise supported by their lack of moral rights. The movie suggests that what makes us human lies in the difference between the two.

However, from this beginning, the movie systematically eliminates possible candidate differences: physical form, intelligence, motivation to survive, emotional experiences, and so on. The last ontological distinction is lost at the death of the lead replicant. In one of cinema's most memorable scenes, just before his death, the replicant says:

> "I've seen things you people wouldn't believe. Attack ships on fire off the shoulder of Orion. I watched C-beams glitter in the dark near the Tannhäuser Gate. All those moments will be lost in time, like tears...in...rain. Time to die."

As he dies, in an allegorical reference to the ascension of a soul, a dove leaves his hands and flies up into the light. What was lost at death was, in his own words, his unique experiences that we could only glimpse but not fully know. At that moment, we cannot help but feel moved by the death of the replicant [49]. The movie merges the ontological categories by having Deckard, who we had identified with, realizing after this scene that he might be a replicant himself [41].

Relating to and being emotionally engaged by life-stories of machines is a recurring theme in many science fiction works: *2001: A*

Space Odyssey, *Ex Machina* and *Otherspace*, to name a few. This thesis is an exploration of this theme.

1.1 RESEARCH QUESTION

The main research question of this thesis is whether the implicit life-stories of a robot can invoke empathy for it. By implicit life-stories I mean the ability of the robot to experience the world we live in, to be transformed through that experience, and to communicate the experience to us. By empathy for a robot, I mean understanding and experiencing the robot's perceived emotions as if they are our own.

1.2 WHY EMPATHY FOR ROBOTS?

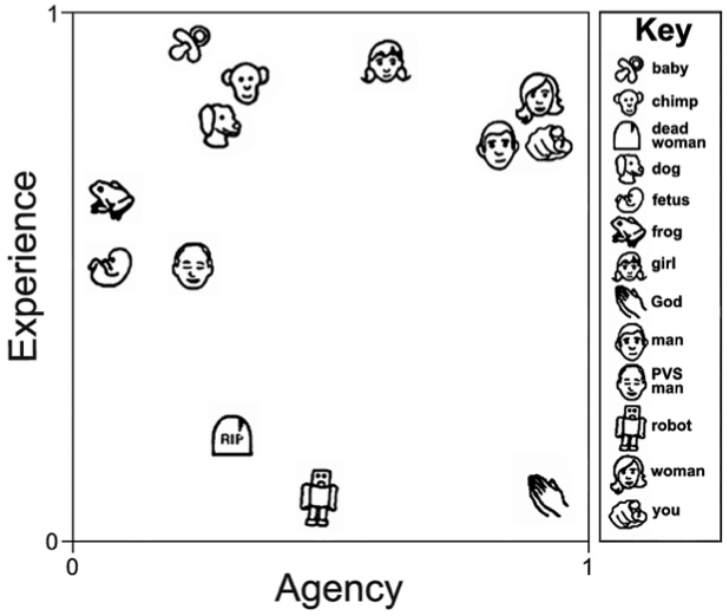

Figure 2: Gray et al.'s analysis of people's perception of minds of various character showed that robots are considered to be lacking in experience compared to other characters [27]

Empathy for artificial agents has been proposed as a test of believability; that is to say, the greater the ability an agent has to invoke empathy, the more it is perceived to be life-like [44]. A common conception of robots is that they do not have the capacity for emotions. Gray et al. examined people's perception of minds by asking participants to rank characters on various mental capacities (for example, out of a robot and a baby, which has greater capacity for pride, for telling right from wrong, etc.) and personal judgements (for example, who should be more deserving of blame if they were to cause harm to others) [27]. Analysis of the responses showed that while robots were perceived to have agency that is comparable to other living en-

tities, they were most lacking in affective experiences (See Figure **??**). If we have empathy for robots, that is to say, we feel that we experience their emotions as if our own, then I argue that we implicitly believe that such robots have emotional experiences. Such capability could elevate the perception of robots to be closer to that of other living creatures. The process of bringing robots closer to us in our perception is, by construction, a pursuit of trying to understand who we are.

Figure 3: Residents at a retirement home find comfort in Paro, a therapeutic robotic seal[52]

A second case for empathy towards robots is to create robots that can have roles similar to companion animals. There is wide evidence that companion animals have a beneficial effect on our physiological and psychological health. Interaction with companion animals has been shown to reduce loneliness, blood pressure, cholestorol, depression, anxiety, and stress [55]. An important aspect of our relationship with animals is that we tend to feel empathy toward them. [45]. In addition, empathy toward companion animals can lead to greater empathy toward humans, which is a highly prosocial trait [53]. It is worth exploring how to design robots to elicit empathy which could bring some of the benefits of companion animals to those who are unable to have them.

1.3 IMPLICIT LIFE STORIES

Living entities grow and change from experience over the course of their lives. Dautenhahn and Nehaniv described life as a complex story embedded in biology [13]. Changes in morphology or behavior tell the story of the object. For example, consider encountering an one-

eyed orange cat sunning itself on a wall. From that brief encounter, one extrapolates a checkered past of scuffles in alleyways and summers enjoying warmth. While I leave more to the reader's imagination, the point here is that through morphology and behavior, living creatures implicitly tell the stories of their lives.

However, these stories are told imperfectly. They are not simple reproductions but rather underspecified ambiguous descriptions. Underspecificity, Gaver argued is a virtue, as it engages us as particpants [23]. We become particpants in constructing the story. Ackermann wrote that viewers engage with objects by "reconstructing them through the lens of their own interests and experiences." [1] We color in the details of the stories using our own life-stories. We wince at the pain of the imagined fight that cost our orange cat his eye from past experience of our own pain. Empathy is generally defined as the ability to understand and feel what another feels [17]. If, through our own experience, we understand and are engaged by the experiences of another, then it stands to reason we would feel empathy for them. I posit that if we can build robots with implicit life-stories, we will feel empathy towards them.

Based on the argument I outlined, I consider the following properties necessary for a robot to have an implicit life-story:

- Autonomous agent: Robot can mean different things. For the purpose of implicit life-stories, all that is required is the robot is perceived to be capable of sensing and acting on the world driven by an internal process. This is necessary to be able to attribute experience to it.

- Perceptible change: The robot must change from experience. Without change, experience wouldn't leave an imprint on the robot. The change needs to be human perceptible so we would know of its existence and of the experience that caused it.

- Relatable experience: The change to the robot must come from an experience that we could imagine ourselves having. A machine learning algorithm, learning to classify a dataset, is changing from experience but an experience we cannot share. Such an agent wouldn't qualify as having an implicit life-story by my definition. Embodiment and/or having human-like perception would make it easier for an agent to have a shareable experience.

- Ambiguous Experience: As discussed earlier, the change to the robot must ambiguously reflect, or evoke, the experience it had. A video camera recording a movie is changing from a perceptual input that we could have. However, the perfect reproduction doesn't leave room for us to imagine a social model much less being able to put ourselves in its place. On the other hand

given an input, the resulting change cannot be incomprehensible as that would preclude us from reconstructing the experience.

Over the next few chapters, I share what I learned from building robots according to these ideas and testing my thesis with human subject studies.

1.4 THESIS OVERVIEW

Chapter **??**: I provide background for social robotics and empathy and describe the relationship of this work to these fields. I also note related work done on stories for artificial agents.

Chapter **??**: I describe a minimal robot that I construct as a preliminary exploration of implicit life-stories. Consider this as a thought experiment.

Chapter **??**: I test my thesis with a human robot interaction study where participants interact with a toy robot that has been given a fictional implicit life-story. After the interaction, the participants are asked to strike the robot. From measuring their hesitation to strike and through psychometric tests, I show that life-stories invoke empathy.

Chapter **??**: Based on the results from the initial human subject study, I build a sound robot that can have a life-story. In this chapter, I describe design considerations, the process of building the mechanical systems, and the architecture of the software necessary to animate the robot and to learn sounds.

Chapter **??**: I then conduct pilot studies with human subjects to understand perception of the robot and to refine validation study design. I describe my experience of iterating on the design.

Chapter **??**: I conduct controlled human subject study with both online and in person participants. I show that implicity life-stories create greater empathy for the new robot. I also find that empathy for robot can impact subsequent empathy for humans.

Chapter **??**: I summarize my findings from my studies and my contributions to the field. I discuss ideas for extending this work and possible applications for such robot.

1.5 CONTRIBUTION

As a preview of the final chapter, I submit these as my contributions:

- Proposed a new design element, implicit life-stories, for engendering empathy for social robots
- Demonstrated empathy for robots with implicit life-stories through human robot interaction experiments across two different platforms

- Designed and built a novel sound interaction social robot to embody and test these ideas
- Demonstrated that the implicit life-stories can improve perception of social robot on animacy, anthropomorphism, likeability and intelligence measures
- Showed that empathy for robot has an impact on subsequent empathy for humans. This is the first work to examine the connection between the two.
- Open-sourced design and code for the new robot

BIBLIOGRAPHY

[1] Edith Ackermann. "Experiences of Artifacts: People's Appropriation/Object's 'Affordances'." In: *Keyworks in radical constructivism* (2007), pp. 249–259.

[2] Joscha Bach. "The micropsi agent architecture." In: *Proceedings of ICCM-5, international conference on cognitive modeling, Bamberg, Germany*. Citeseer. 2003, pp. 15–20.

[3] Wilma A Bainbridge, Justin Hart, Elizabeth S Kim, and Brian Scassellati. "The effect of presence on human-robot interaction." In: *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on*. IEEE. 2008, pp. 701–706.

[4] Mark A Barnett, Jeffrey A Howard, Laura M King, and Geri A Dino. "Helping behavior and the transfer of empathy." In: *The Journal of social psychology* 115.1 (1981), pp. 125–132.

[5] Christoph Bartneck, Michel Van Der Hoek, Omar Mubin, and Abdullah Al Mahmud. "Daisy, Daisy, give me your answer do!: switching off a robot." In: *Proceedings of the ACM/IEEE international conference on Human-robot interaction*. ACM. 2007, pp. 217–222.

[6] Christoph Bartneck, Marcel Verbunt, Omar Mubin, and Abdullah Al Mahmud. "To kill a mockingbird robot." In: *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on*. IEEE. 2007, pp. 81–87.

[7] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots." In: *International journal of social robotics* 1.1 (2009), pp. 71–81.

[8] Joseph Bates et al. "The role of emotion in believable agents." In: *Communications of the ACM* 37.7 (1994), pp. 122–125.

[9] Timothy Bickmore, Daniel Schulman, and Langxuan Yin. "Engagement vs. deceit: Virtual humans with human autobiographies." In: *Intelligent Virtual Agents*. Springer. 2009, pp. 6–19.

[10] Valentino Braitenberg. *Vehicles: Experiments in synthetic psychology*. MIT press, 1986.

[11] Cynthia Breazeal. "Designing Socially Intelligent Robots." In: *Frontiers of Engineering: Reports on Leading-Edge Engineering from the 2004 NAE Symposium on Frontiers of Engineering*. The National Academies Press, 2005, pp. 123–130.

[12] Kate Darling, Palash Nandy, and Cynthia Breazeal. "Empathic concern and the effect of stories in human-robot interaction." In: *Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on*. IEEE. 2015, pp. 770–775.

[13] Kerstin Dautenhahn and Chrystopher Nehaniv. "Artificial life and natural stories." In: *International Symposium on Artificial Life and Robotics*. Citeseer. 1998, pp. 435–439.

[14] Mark H Davis. "Measuring individual differences in empathy: evidence for a multidimensional approach." In: *Journal of personality and social psychology* 44.1 (1983), p. 113.

[15] Richard Dawkins. *The selfish gene*. 199. Oxford university press, 2006.

[16] Frans BM De Waal. "Putting the altruism back into altruism: the evolution of empathy." In: *Annu. Rev. Psychol.* 59 (2008), pp. 279–300.

[17] Jean Decety and Claus Lamm. "Human empathy through the lens of social neuroscience." In: *The Scientific World Journal* 6 (2006), pp. 1146–1163.

[18] Daniel Clement Dennett. *The intentional stance*. MIT press, 1989.

[19] Kenji Doya and Terrence J Sejnowski. "A computational model of birdsong learning by auditory experience and auditory feedback." In: *Central auditory processing and neural modeling*. Springer, 1998, pp. 77–88.

[20] Charles R Figley. "Compassion fatigue: Psychotherapists' chronic lack of self care." In: *Journal of clinical psychology* 58.11 (2002), pp. 1433–1441.

[21] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. "A survey of socially interactive robots." In: *Robotics and autonomous systems* 42.3 (2003), pp. 143–166.

[22] Chris D Frith and Uta Frith. "Mechanisms of social cognition." In: *Annual review of psychology* 63 (2012), pp. 287–313.

[23] William W Gaver, Jacob Beaver, and Steve Benford. "Ambiguity as a resource for design." In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2003, pp. 233–240.

[24] E von Glasersfeld and Edith K Ackermann. "Reflections on the Concept of Experience and the Role of Consciousness. Unfinished Fragments." In: *Constructivist Foundations* 6.2 (2011), pp. 193–203.

[25] Rachel Gockley, Allison Bruce, Jodi Forlizzi, Marek Michalowski, Anne Mundell, Stephanie Rosenthal, Brennan Sellner, Reid Simmons, Kevin Snipes, Alan C Schultz, et al. "Designing robots for long-term social interaction." In: *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*. IEEE. 2005, pp. 1338–1343.

[26] Alvin I Goldman. "Empathy, mind, and morals." In: *Proceedings and addresses of the American philosophical association*. Vol. 66. 3. JSTOR. 1992, pp. 17–41.

[27] Heather M Gray, Kurt Gray, and Daniel M Wegner. "Dimensions of mind perception." In: *Science* 315.5812 (2007), pp. 619–619.

[28] Tobias Greitemeyer and Silvia Osswald. "Effects of prosocial video games on prosocial behavior." In: *Journal of personality and social psychology* 98.2 (2010), p. 211.

[29] Louise C Hawkley and John T Cacioppo. "Loneliness matters: A theoretical and empirical review of consequences and mechanisms." In: *Annals of Behavioral Medicine* 40.2 (2010), pp. 218–227.

[30] Guy Hoffman and Wendy Ju. "Designing robots with movement in mind." In: *Journal of Human-Robot Interaction* 3.1 (2014), pp. 89–122.

[31] Takayuki Kanda, Rumi Sato, Naoki Saiwaki, and Hiroshi Ishiguro. "A two-month field trial in an elementary school for long-term human–robot interaction." In: *Robotics, IEEE Transactions on* 23.5 (2007), pp. 962–971.

[32] Frédéric Kaplan. "Artificial attachment: Will a robot ever pass ainsworth's strange situation test." In: *Proceedings of humanoids*. 2001, pp. 125–132.

[33] Dennis Krebs. "Empathy and altruism." In: *Journal of personality and social psychology* 32.6 (1975), p. 1134.

[34] John Lasseter. "Principles of traditional animation applied to 3D computer animation." In: *ACM Siggraph Computer Graphics*. Vol. 21. 4. ACM. 1987, pp. 35–44.

[35] Kwan Min Lee, Namkee Park, and Hayeon Song. "Can a robot be perceived as a developing creature?" In: *Human Communication Research* 31.4 (2005), pp. 538–563.

[36] Iolanda Leite, Samuel Mascarenhas, André Pereira, Carlos Martinho, Rui Prada, and Ana Paiva. "" Why Can't We Be Friends?" An Empathic Game Companion for Long-Term Interaction." In: *Intelligent Virtual Agents*. Springer. 2010, pp. 315–321.

[37] Matthew D Lieberman. *Social: Why our brains are wired to connect*. OUP Oxford, 2013.

[38] C Neil Macrae and Lucy Johnston. "Help, I need somebody: Automatic action and inaction." In: *Social Cognition* 16.4 (1998), pp. 400–417.

[39] Stacy Marsella, Jonathan Gratch, and Paolo Petta. "Computational models of emotion." In: *A Blueprint for Affective Computing-A sourcebook and manual* (2010), pp. 21–46.

[40] Masahiro Mori, Karl F MacDorman, and Norri Kageki. "The uncanny valley [from the field]." In: *Robotics & Automation Magazine, IEEE* 19.2 (2012), pp. 98–100.

[41] Stephen Mulhall. "Picturing the human (body and soul): A reading of Blade Runner." In: *Film and philosophy* 1.1994 (1994), pp. 87–104.

[42] "Online Etymology Dictionary." In: (2016). URL: http://www.dictionary.com/browse/empathy.

[43] Steffi Paepcke and Leila Takayama. "Judging a bot by its cover: an experiment on expectation setting for personal robots." In: *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*. IEEE. 2010, pp. 45–52.

[44] Ana Paiva, Joao Dias, Daniel Sobral, Ruth Aylett, Polly Sobreperez, Sarah Woods, Carsten Zoll, and Lynne Hall. "Caring for agents and agents that care: Building empathic relations with synthetic agents." In: *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1*. IEEE Computer Society. 2004, pp. 194–201.

[45] Clive Phillips. "Empathy Towards Animals." English. In: *The Welfare of Animals*. Vol. 8. Animal Welfare. Springer Netherlands, 2009, pp. 47–54. ISBN: 978-1-4020-9218-3. DOI: 10.1007/978-1-4020-9219-0_3. URL: http://dx.doi.org/10.1007/978-1-4020-9219-0_3.

[46] Astrid M Rosenthal-von der Pütten, Nicole C Krämer, Laura Hoffmann, Sabrina Sobieraj, and Sabrina C Eimler. "An experimental study on emotional reactions towards a robot." In: *International Journal of Social Robotics* 5.1 (2013), pp. 17–34.

[47] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y Ng. "ROS: an open-source Robot Operating System." In: *ICRA workshop on open source software*. Vol. 3. 3.2. 2009, p. 5.

[48] Byron Reeves and Clifford Nass. *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press, 1996.

[49] Mark Rowlands. *The Philosopher at the End of the Universe*. Random House, 2005.

[50] Jürgen Schmidhuber. "Art & science as by-products of the search for novel patterns, or data compressible in unknown yet learnable ways." In: *Multiple ways to design research. Research cases that reshape the design discipline, Swiss Design Network-Et al. Edizioni* (2009), pp. 98–112.

[51] Stela H Seo, Denise Geiskkovitch, Masayuki Nakane, Corey King, and James E Young. "Poor Thing! Would You Feel Sorry for a Simulated Robot?: A comparison of empathy toward a physical and a simulated robot." In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2015, pp. 125–132.

[52] Nidhi Subbaraman. *My robot friend: People find real comfort in artificial companionship*. [Online; accessed 03-April-2016]. 2013. URL: http://www.nbcnews.com/technology/my-robot-friend-people-find-real-comfort-artificial-companionship-6C10146787.

[53] Kelly L Thompson and Eleonora Gullone. "Promotion of empathy and prosocial behaviour in children through humane education." In: *Australian Psychologist* 38.3 (2003), pp. 175–182.

[54] R Walker and J Glenn. *About the Significant Objects Project*. [Online; accessed 10-June-2015]. 2009. URL: http://significantobjects.com/about/.

[55] Froma Walsh. "Human-animal bonds I: The relational significance of companion animals." In: *Family process* 48.4 (2009), pp. 462–480.

[56] Ryan Mark Wistort. "TofuDraw: choreographing robot behavior through Digital Painting." PhD thesis. Massachusetts Institute of Technology, 2010.