Leveraging Robot Embodiment to Facilitate Trust and Smoothness

Samantha Reig  
Human-Computer Interaction Institute  
Carnegie Mellon University  
Pittsburgh, USA  
sreig@cs.cmu.edu

Jodi Forlizzi  
Human-Computer Interaction Institute  
Carnegie Mellon University  
Pittsburgh, USA  
forlizzi@cs.cmu.edu

Aaron Steinfeld  
Robotics Institute  
Carnegie Mellon University  
Pittsburgh, USA  
steinfeld@cmu.edu

Abstract—Interactions with social robots in public and private spaces are becoming more and more common and varied. As this trend continues, it is important to understand how a robot’s embodiment influences its ability to calibrate trust and comfort with its users and behave in accordance with social norms. This is especially true when one social intelligence embodies multiple physical robots (re-embodiment). We have conducted two studies—one quantitative and one qualitative—which shed light on the way robots should be embodied and re-embodied by intelligences during different types of social interactions. This paper outlines our previous work on elucidating the role of embodiment in social interactions and experimenting with re-embodiment as a design paradigm, and it describes the directions in which we plan to take this research in the near future.

Index Terms—Embodiment, Empirical Study, Human-Computer Interaction, HCI, Trust

I. INTRODUCTION

Social robots are on the rise, but have not quite come into their own as true social partners. Still, recent advances like Anki’s Vector robot [1], Google Assistant’s new ability to respond to requests in multiple languages [2], and the increasing presence of university dormitory robots that use Amazon Alexa [3] demonstrate that robots and IoT devices that socialize and serve are continuing to find their places in our day-to-day routines and increasingly becoming part of our professional and personal lives.

While principles and precedents exist for creating robots and intelligent agents that meet users’ expectations for interaction, the role of embodiment in designing for different types of interactions is not well understood. Guidelines for design and behavior may also break down when we consider intelligences that move from one physical housing to another, maintaining their “essence” or “substance” while modifying their physical presentations to adapt to new contexts with new demands. This paradigm, which we call re-embodiment, raises numerous questions: Should your Google assistant be trusted to drive your autonomous car? If you use Siri at someone else’s house through their home robot, is it your Siri? Our team is employing surveys, lab and field studies, and technical development to identify the new social norms, probe the possibilities, and describe potential futures of a world wherein intelligent agents migrate across embodiments.

The impact of a robot’s embodiment in social interactions where no specific physical characteristics are obviously required (e.g., being strong enough to move large objects but lightweight enough to be portable) is fairly well-explored in the HRI and HCI literature, yet still confusing. Some studies suggest that the physical presence of a robot increases its social presence [4], [5]—which is key to positive and productive interactions [6]–[8]—and makes interactions more engaging for humans [4], [9]. Others suggest that virtual agents may be preferred over physical agents in situations that do not involve physical touch [9] or in interactions that are particularly task-oriented [11]. Still others find that additional factors, such as social behavior and participants’ pre-existing attitudes toward robots, play a larger role than embodiment type in determining social attractiveness [12] and engagement [13]. A few researchers have attempted to characterize embodiment to give the HRI community a common language and set of guidelines to study and design embodied agents [14], [15].

In addition to understanding embodiment, our work explores the concept of software intelligence that migrates across robot bodies (re-embodiment). This concept made one of its first appearances in the literature in the form of Duffy’s Agent Chameleon Project [16], which was an early attempt to propose an architecture that could support migration across multiple physical systems. Researchers have extended this by conducting experiments that test the psychological boundaries of robots [17], running discussion groups to assess children’s ability to conceptualize migrating intelligence [18], and prototyping agent migration in long-term human-robot interaction by installing migrating intelligence in robots that inhabit mock smart homes [19]. This work has raised questions about the “ideal” configuration of intelligence and embodiment [18].

II. RESEARCH APPROACH

Our efforts so far have focused on distilling the meaning of being embodied. Our goal has been to test the way various forms of embodiment influence aspects of the complex interactions that people are likely to encounter with service robots and social robots in the near future.

In our empirical study, an intelligent agent gave verbal “guidance” to a human whose task was to solve a puzzle in a limited amount of time. Here, we were interested in people’s interpretations of an agent’s failure to provide useful

This work was funded by the National Science Foundation (SES-1734456).
whether or not behavior [21], [22] even when it is designed to seem dishonest. This result was interesting to us because robot deception is often perceived as a malfunction [20] or otherwise acceptable. We predicted that the quality of information and the agent’s embodiment would affect participants’ interactions with the agent, assumptions about the agent, task performance, and thoughts and feelings about both the agent and the task. Participants were asked to complete a search task that required looking around a room. During the search, an agent periodically gave suggestions for finding clues. The task was designed to mimic real-life situations in which a user moves in physical space while interacting with an intelligent agent that may or may not be embodied; for example, collaboratively manufacturing a product, cooking a meal, or navigating to a location. Our study found key differences in perceptions of the agent stemming from both the agent’s embodiment and the quality of information it provided. For example, a robot and a voice were perceived as easier to interact with than a tablet. A robot also had less influence over the participants’ behavior than a tablet or a voice. Whether or not the agent provided Good or Bad information played a measurable role in participants’ perceptions that the agent had betrayed them. The study used the method of user enactments [23] to expose participants to futuristic situations involving various forms of agent-body presentation. Our study design encouraged participants to openly evaluate and reflect on their experiences with agents that used different embodiments in different contexts. In this experiment, we made minimal use of immobile screens; we instead designed low-fidelity prototypes of robots and smart objects that used voices and movements to communicate with users. We uncovered key insights about the way people interpret one agent embodying different physical robots, including when re-embodiment is promising and when it may be alarming, how people assume it works, and what boundaries it violates. People were generally receptive to re-embodiment and reported that it made interactions with multiple devices feel more efficient and seamless. Re-embodiment was less well-received in domains requiring specific expertise, intense concentration, or the exchange of private information.

III. Future Work

Our first study analyzed humans’ impressions of agent and robot behaviors when it is not clear whether the motivations for those behaviors are in the interest of the user, and our second study probed the broad and largely unexplored design space of re-embodiment. Our future work will extend our investigation of ambiguous robot behaviors and re-embodiment.

First, we will explore how people may re-calibrate their trust of a robot following an experience indicative of robot failure. “Failure” behavior in this context might be interpreted as an intentional violation (as in the first study described), a one-time malfunction, or a sign of general incompetence. These interpretations could be impacted by the robot’s embodiment, appearance, and communicative capabilities, and mediated by contextual factors such as the task at hand, the user’s history with the robot, and the presence of other people and/or robots. In extreme cases, a failure may come in the form of a physical (or physical-seeming) problem that is unrecoverable using traditional recovery methods like acknowledging the difficulty of the task [5], [24] or social practices like acknowledging the error, explaining why the error occurred, and apologizing for it [5], [25]. Examples of this sort of irreparable failure include a battery or motor dying in the middle of an interaction, a piece of the robot breaking in a visible way, or an excess of demands from multiple users causing the robot to fail to manage all of the requests at once. Such occurrences may detract significantly from any trust that has been cultivated and whatever positive social attributions the human has made to the robot. This is especially true if the failure prevents the system from continuing to interact in multiple modalities (e.g., a dead battery prevents verbal and nonverbal communication; an arm moving around in an erratic ways gives visual, auditory, and potentially tactile feedback that is confusing and not useful; ignoring requests results in no user-directed feedback at all). We propose that in these situations, continuing the interaction with a second robot that picks up the interaction where the first robot left off—maintaining the same “personality” and memory of the first part of the interaction—may be an effective way to mitigate the negative effects of failure on enduring trust and social perceptions. We plan to test this idea using studies that expose participants to re-embodiment as a strategy for recovery after failure.

Our second effort is focused on understanding the assumptions humans make about robots’ abilities, limitations, status, and capacity. As the number and complexity of interactions between humans and social agents increases, people may perceive agents and robots as having cognitive resources that can be exhausted, similarly to how humans managing multiple tasks can experience cognitive overload [26]. It is important for robots to be able to indicate to human users that they are nearing their cognitive limits, that they are being asked to do something outside their scope, or that they are close to experiencing a failure. This part of the project will use empirical studies to explore how robot design and task influence human assumptions about the robot’s cognitive load. Inspired by prior work exploring the idea of re-embodifying agents communicating and autonomously responding to near-failures and near-fatal states [16], we will also utilize participant workshops and user enactments to generate and test strategies for agents to communicate the impact of task demands.
REFERENCES


