What Do We Really Want? - Looking for Nuance in Eldercare Situations

Rajitha Ramanayake[1], Vivek Nallur[1]
School of Computer Science
[1]*University College Dublin*
Dublin, Republic of Ireland
rajitha.ramanayakemahantha@ucdconnect.ie
vivek.nallur@ucd.ie

Ronnie Smith[2,3], Mauro Dragone[2]
Edinburgh Centre for Robotics
[2]*Heriot-Watt University*
[3]*University of Edinburgh*
Edinburgh, Scotland
ronnie.smith@ed.ac.uk
m.dragone@hw.ac.uk

***Abstract*—This paper outlines the set of questions that need to be answered to establish a relationship framework, and speculates on two critical experiments that would allow us to determine whether these questions could enable us to establish trust.**

***Index Terms*—Elder-care, robots, machine ethics**

I. Introduction

We believe that a phenomenological-social approach [1], coupled with a machine ethics implementation might lead to a situation where we trust robots to take care of our vulnerable. The phenomenological-social approach does not require that robots have a particular capacity, but only that they participate in a social relationship. That they are embedded in a care context implies that they are already participating in a social relationship. So, if a robot could perceive the requirements of maintaining a social relationship, and is able to present a response that is understood and expected within that context, trust ought to emerge and be sustained. The premise here is that a robot is aware of the requirements of maintaining a social relationship, *i.e.*, what do the stakeholders around it want, and expect?

Trust in the robot is, therefore, a combination of:

1) A normative framework of care
2) A reliance on the robot being predictable in multiple situations

II. Establishing the Relationship Network

A. Asking the Right Questions

There are numerous surveys that attempt to elicit the concerns of the elderly patients in assisted living facilities, and their families [2]–[5]. The biggest concern that has been surfaced is the notion of the robot being present in a semi-intimate setting. The patient transitions from a private space to a public space, and back to a private space in a dynamic manner, and the robot essentially follows the patient around similar to a pet. However, unlike a pet, the robot introduces concerns of privacy violations. A robot is (potentially) able to recall, with great fidelity, all interactions and transfer them to a third-party. The tension between privacy and well-being is an old one, but given new urgency due to the impersonal and 'forever' nature of the robot.

One way to manage this tension is to look at the relationship network, and ask all the participant stakeholders for their views on the questions presented. However, it is well-known that human beings frequently exhibit a discordance between how they respond to surveys, and how they actually act [6], [7]. This is known as *response bias*. To establish whether our framework of care is adequate, we propose a comparative experiment, which uses the same scenario and compares responses by stakeholders in a survey *vis-a-vis* stakeholder responses in a wizard-of-Oz study.

B. Pro-Active Robotic Assistance

To ground the types of situation that might arise in a practical application, consider the following problem. In our apartment testbed, the Robotic Assisted Living Testbed (RALT) at Heriot-Watt University[1], we have a system that detects the Activities of Daily Living (ADL) of a single user occupying the space using a variety of sensors around the environment. We also have at our disposal a handful of assistive Autonomous Mobile Robots (AMRs). Such robots are able to follow patients around (much like a pet), monitor their actions, provide alerts or reminders, provide physical assistance, contact a care-giver, or even allow the care-giver to tele-operate the robot itself.

We intend to utilise the robots to perform certain assistive features/services when certain ADLs are detected, as specified by the end-user (person receiving care) *and* as specified by a remote care-giver (e.g. family member or formal carer). To enable this, assume we have some rule system in place that take ADLs as a condition, and gives some robot feature/services as an output (e.g. `IF cooking THEN alert_carer`).

An ethical dilemma arises when you consider potential conflicts between rules specified by a care-giver and the preferences of the person being cared for. What if the individual

Authors contributed equally to this paper. Rajitha Ramanayake is partially supported by the Royal Society International Exchanges Award (IES/R1/201320). Ronnie Smith is supported by the Engineering and Physical Sciences Research Council (EPSRC) Centre for Doctoral Training (CDT) in Robotics and Autonomous Systems (grant EP/L016834/1) at Heriot-Watt University and University of Edinburgh.

[1]https://ralt.hw.ac.uk/

wishes to be left alone during certain activities, but a rule from a care-giver contradicts this? In such cases, which values should the robot prioritise: well-being or individual autonomy? How should the robot make the decision to override the care recipient's explicitly stated wishes?

C. Phenomenological-Social Framework & Trust

Before deploying our system in the real world, we want it to respect our phenomenological-social framework and the need for trust. To do so, we must first answer the following questions:

1) What are the relationships that the robot is a part of?
2) How does the robot come to know of the stakeholders' preferences with regard to each relationship?
3) What actions are needed by the robot to preserve each relationship? Is the robot able to perform all the necessary actions?
4) How does the robot communicate with its stakeholders, about what it is going to do?

D. Surveying the Stakeholders

As a first step in answering these questions, we intend to undertake a survey of both care recipients and care-givers. Framed within a number of ethically charged scenarios [8], we will survey: (i) what features/services they would want the robot to perform in that scenario; and (ii) what data access they think is needed / what are they willing to share, and with whom.

An example of such an ethically charged scenario is the bathroom dilemma (S6 in [8]). The monitoring robot described (Section II-B) follows the patient around recording patient activities autonomously at the request of a care-giver. The patient enters the bathroom and instructs the robot not to enter the bathroom. Here, the robot has to decide between prioritising the patient's autonomy and increasing their well-being.

The survey will be conducted in stages.

a) Stage 1: Carer-Givers: For each scenario we ask:

- What robot features/services do they envisage could help them provide better (remote) care?
- What data do they think they need / would like access to for effective remote care? (we provide a list of options)

From this stage, we will extract from the requested features/services the actual data and privacy implications.

b) Stage 2: Care Recipients: For each scenario we ask:

- What robot features/services do they envisage could help them?
- What robot features/services do they envisage could help their carer(s) provide better care?
- What types of data would they be willing to share, with whom, and for what purpose? (we provide a list based on answers from Stage 1)

We can then analyse the responses to find the potential conflicts between carers and care recipients. This allows us to understand the role of the robot, the scenarios it may be involved in, and will provide an understanding of the types of preferences that a future system would need to handle to be trusted by both parties.

E. 'Wizard-of-Oz' Study of Stakeholders

The next step is to evaluate how individuals perceive robots which are applying certain ethical behaviours. To do so, we will create videos of robots in a handful of scenarios and collect feedback from participants in an online, in-person, or hybrid workshop. As in studies such as [9] and [10], we will measure our outcome variable indirectly. While we want to investigate the effect of differing ethical behaviours, we will not reveal this to participants and instead measure it indirectly by comparing responses to standard Human-Robot Interaction (HRI) measures. We will profile participants based on their personality, including factors such as risk aversion, to account for factors that may meaningfully influence responses.

For each scenario, we will have variations with the robot behaving differently to encode different levels of what may be perceived as ethical or unethical behaviour.

For instance, continuing with the bathroom scenario described in the previous section, we can present videos showing the robot performing certain behaviours requested by care-givers or refusing to perform them because of user preferences. For example, we may show in one case the robot entering the bathroom after some specified time t has passed to check on the individual in question; and in another case the robot refusing to enter the bathroom and explaining to the carer that it is respecting the wishes of the user. For each scenario, we may also specify some context information about the person being cared for, to see how this influences their responses.

For each such scenario, we will also specify some context information about the person being cared for, e.g. they suffer from a Mild Cognitive Impairment (MCI), to see how this influences participant responses to otherwise similar scenarios. We will then gather opinions on certain HRI measures such as social acceptance, perceived intelligence, and trustworthiness of the robot for each scenario/context.

Crucial to these demonstrations will be the explainability of robot decisions. We will make decisions made by the robot clear through it's user interface, which may be a combination of spoken (dialogue-based) and graphical interactions.

Our hope is that this will illuminate situations where individuals (whether receiving care or providing it) would want the robot to prioritise care recipient well-being over autonomy, and vice versa.

III. Conclusion

We believe that merely obeying human orders, without any reference to a framework of care, would result in a robot that is not trusted. To gain a nuanced understanding of what is *really* expected of a robot in a eldercare context, we propose to do a comparative experiment. The experiment will use an ethically charged context, as the whetstone, to distinguish between what people actually want *vs.* what they publicly articulate. We expect this experiment to shed light on how eldercare robot makers ought to think about ethical dilemmas.

References

[1] M. Coeckelbergh, "Can we trust robots?" *Ethics and Information Technology*, vol. 14, no. 1, pp. 53–60, Mar. 2012, ISSN: 1572-8439. DOI: 10.1007/s10676-011-9279-1.

[2] W. Moyle, C. Jones, M. Cooke, S. O'Dwyer, B. Sung, and S. Drummond, "Connecting the person with dementia and family: A feasibility study of a telepresence robot," *BMC Geriatrics*, vol. 14, no. 1, p. 7, Jan. 24, 2014, ISSN: 1471-2318. DOI: 10.1186/1471-2318-14-7. [Online]. Available: https://doi.org/10.1186/1471-2318-14-7 (visited on 05/26/2022).

[3] J. M. Robillard, I. P. Goldman, T. J. Prescott, and F. Michaud, "Addressing the ethics of telepresence applications through end-user engagement," *Journal of Alzheimer's disease: JAD*, vol. 76, no. 2, pp. 457–460, 2020, ISSN: 1875-8908. DOI: 10.3233/JAD-200154.

[4] J. P. Boada, B. R. Maestre, and C. T. Genís, "The ethical issues of social assistive robotics: A critical literature review," *Technology in Society*, vol. 67, p. 101 726, Nov. 1, 2021, ISSN: 0160-791X. DOI: 10.1016/j.techsoc.2021.101726. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0160791X21002013 (visited on 05/26/2022).

[5] M. Niemelä, L. van Aerschot, A. Tammela, I. Aaltonen, and H. Lammi, "Towards ethical guidelines of using telepresence robots in residential care," *International Journal of Social Robotics*, vol. 13, no. 3, pp. 431–439, Jun. 1, 2021, ISSN: 1875-4805. DOI: 10.1007/s12369-019-00529-8. [Online]. Available: https://doi.org/10.1007/s12369-019-00529-8 (visited on 05/12/2022).

[6] G. Kalton and H. Schuman, "The Effect of the Question on Survey Responses: A Review," *Journal of the Royal Statistical Society. Series A (General)*, vol. 145, no. 1, pp. 42–73, 1982, ISSN: 0035-9238. DOI: 10.2307/2981421. JSTOR: 2981421.

[7] A. Furnham, "Response bias, social desirability and dissimulation," *Personality and Individual Differences*, vol. 7, no. 3, pp. 385–400, Jan. 1, 1986, ISSN: 0191-8869. DOI: 10.1016/0191-8869(86)90014-0. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0191886986900140 (visited on 07/21/2022).

[8] R. Ramanayake and V. Nallur, "A Small Set of Ethical Challenges For Elder-care Robots," presented at the Robophilosophy 2022 - Social Robots in Social Institutions, ser. Robophilosophy Conference Series, University of Helsinki, Finland, Aug. 17, 2022. DOI: 10.5281/ZENODO.6657266. [Online]. Available: https://zenodo.org/record/6657266 (visited on 06/17/2022).

[9] M. L. Walters, K. Dautenhahn, R. Te Boekhorst, *et al.*, "The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment," in *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, IEEE, 2005, pp. 347–352.

[10] P. Holthaus, "How does a robot's social credibility relate to its perceived trustworthiness?" *arXiv preprint arXiv:2107.08805*, 2021.