Design of a Robotic Therapy Companion from a Value-Sensitive and Technical Perspective

Vivienne Jia Zhong, Phillip Gachnang, Theresa Schmiedel

School of Business, FHNW University of Applied Sciences and Arts Northwestern Switzerland
viviennejia.zhong@fhnw.ch,
phillip.gachnang@fhnw.c
theresa.schmiedel@fhnw.ch

Abstract. The nursing sector is facing a profound shortage of skilled workers. Social robots have the potential to relieve the nursing staff. However, given that social interaction is indispensable in the nursing sector, it is essential to equip social robots with comprehensive artificial social interaction capabilities to support the staff effectively. Human values are an essential building block of social interactions and drive individual's behaviors. Therefore, human values need to be reflected in human-robot interaction. In this conceptual paper, we present a value-sensitive approach to the use case of a humanoid social robot in elderly care and conceptually outline how we intend to build to a cognitive system using control theory to enable the humanoid social robot exhibiting value-sensitive actions.

Keywords: Humanoid Robots; Social Robots; Human-Robot-Interaction; Elderly Care; Value Sensitive Design; Use Case; Control Theory; Cognitive System.

1 Relevance of Value-Sensitive Robot Design in the Care Sector

According to the OECD, long-term care is facing a profound shortage of skilled workers worldwide [1], which might lead to unsatisfying quality for care receivers. Social robots have the potential to address this issue by autonomously taking repetitive, non-nursing tasks off the hands of the often-overworked nurses.

However, for a successful deployment in elderly care, it is crucial to design social robots as a participatory and socially interactive entity in everyday life [2]. Such design requires the necessary technical functions to account for the social interaction capabilities that enable the robot to relieve overworked professionals in care institutions effectively.

Human values are an essential building block of social interactions and should, thus, be considered in robot design. However, despite their relevance, values have only received little attention in the robotics industry [3]. Value sensitive design (VSD) is an approach that considers human values in the design of technology and has increasingly raised robot researcher's interest in recent years [4].

In our research, we design a humanoid social robot to be able to function as a therapy companion. The remainder of this conceptual paper is structured as follows: Section 2 provides an overview of robots in the care context. Section 3 introduces the VSD approach, which guides our research. Section 4 outlines our research approach, and in Section 5, we envision a cognitive system with values in mind as a potential technical solution. Finally, we discuss our research in Section 6 before concluding our work in Section 7.

2 Social Robots in the Care Sector

Robots represent systems that have the "ability to perform intended tasks based on current state and sensing, without human intervention" [5]. They are increasingly used for applications in the healthcare sector. As a specific type of robots, social robots are able to communicate socially and interact with humans and other robots. These robots come in various appearances, with the ability to recognize, react, and activate certain emotions of human users based on artificial intelligence (AI) [6].

In light of the ongoing global shortage of skilled care workers, researchers have started investigating the potential of social robots in the nursing sector. The application contexts of social robots are versatile. Research has examined the use of social robots to improve elderly people's psychological well-being [7 and cognitive functions [8]. Further, social robots can provide companionship for lonely and socially isolated elderly people [9 and offer valuable services such as reminding of medication to support older adults' independent living [1 While a large body of research focuses on zoomorphic robots with animal-like shapes, such as the robotic baby harp seal Paro, which appears to have a positive impact on the quality of life of the elderly [11], the use of humanoid robots has also shown benefits for older adults and caregivers in elderly care [1

However, it is important to note that the breadth of social robotics research underscores that the mere presence of functionalities is not sufficient for users to adopt social robots in the long run, and therefore, other social factors need to be considered in the design of social robots [13]. In particular, researchers have increasingly recognized the mutual shaping relationship between technology and society [14]. While scholars have begun explicitly incorporating values into social robot design [15], more research that demonstrates a concrete approach to the design and development of value-sensitive social robots is still needed.

3 Value Sensitive Design

To account for the mutual shaping relationship between technology and society, value sensitive design adopts an interactional stance on technology and values [4]. Values are subconscious needs, that members of a group share (e.g., autonomy, security, or collaboration) [16], and which influences their behavior (e.g., intending to use certain technologies or not) [17]. Indeed, research has indicated that a misalignment between social robots and values might present a barrier to their adoption [18–20].

When designing values in technology, a core question is whose values should be considered. Compared to other human-centered methods, an advantage of VSD is that it extends the focus from the values of mere end-users to those of stakeholders that are affected by the use of a technology [21]. Additionally, as values are not isolated but rather interrelated, VSD recognizes the emergence of value tensions that can occur on various levels: within an individual, within a stakeholder group, among different stakeholder groups. VSD stresses that even though values might be in opposition, at the same time this tension can be solved by balancing values in relation to each other through identifying design requirements that help solve the tension [4].

To ease researchers to investigate and incorporate stakeholder values in the technology, VSD proposes a tripartite methodology that compasses three different types of investigation: Conceptual, empirical, and technical investigations [4]. While conceptual investigation explores the research object from various standpoints (e.g., analytic and theoretical) — particularly identifying stakeholders and their values, the context in which the technology is deployed is examined in empirical investigation. This investigation typically addresses how stakeholders experience values in a specific context by applying methods from social science research. Finally, in technical investigation, researchers examine the alignment between the values and an existing technology design. Moreover, technical investigation also concerns the development of technology according to the identified values [4].

4 Approach to Value-Sensitive Social Robot Design

4.1 Use Case of a Therapy Companion

Motivated by the vision of freeing up nurses for their core nursing tasks, we study the use case of a therapy companion, which serves the purpose of reminding residents of their therapy appointments and eventually accompanying them. This use case encompasses several interaction scenarios, such as searching residents within the nursing home, engaging in conversations with residents to remind them of their appointments and to motivate them for their therapy if necessary, and accompanying residents to the therapy room, etc. Realizing this use case implies two main challenges, namely understanding which values the social robot should act upon and how to realize this technically.

The first challenge concerns designing values for the social robot. So far, existing works have mainly focused on the evaluation of social robots according to a set of predefined values. For instance, there are models assessing the value of trust on user acceptance of social robots [22]. Furthermore, frameworks are proposed to examine how predetermined moral elements are manifested in robot-supported care practices [23]. Although these post-evaluations advance the understanding of the role of value in user adoption of social robots, the design of social robots that incorporates stakeholder values, particularly the operationalization of values into the various interaction features of the social robot, is under-researched.

The second challenge concerns the technology. Thanks to the extensive research on individual AI systems such as navigation, orientation, recognition, and conversation exist, all these AI systems have achieved substantial improvements within the last decade and have been successfully applied individually. However, the technical challenge of integrating all the individual AI technologies into a complex cognitive system to achieve a common goal with values in mind remains unsolved.

4.2 Research Approach

In our research, we employ VSD to design values in the social companion robot. VSD encourages the application of the three investigation types (conceptual, empirical and technical) in an iterative and integrative way. To facilitate this, we adopt the wave approach proposed in [21], which provides specific recommendations for iterating these three investigation types in an integrative manner throughout our design process.

We start with a context analysis to investigate the current practice of therapy companion, which enables us to identify several stakeholders for this use case. Among them, we classify residents, care givers and therapists as direct stakeholders and resident's relative and the management of the nursing home as indirect stakeholders. Following this, we conceptually examine stakeholder-relevant values through existing research (e.g., [24]). The results of the conceptual analysis are then leveraged in our empirical investigation through various qualitative methods. For instance, in value-oriented interviews [4], we ask stakeholders to explicitly assess values identified from the conceptual investigation with regard to the use case, while also implicitly eliciting values from their conception of the robot-supported therapy companion.

For the operationalization of values, we utilize the value hierarchy [25] to translate identified values into design requirements for the use case. Additionally, we collect design requirements from stakeholders through empirical methods (e.g., workshops). When new values emerge from the empirically collected design requirements, we revisit the value hierarchy [25 to derive design requirements for these newly identified values. To guide the technical development, we establish an understanding of the various interaction scenarios between the robot and residents. We suggest using design patterns [26] to create these interaction scenarios and specify human-robot interaction (HRI) based on the derived design requirements. Design patterns have been used in previous works to enable effective and pleasant HRI [27, 28].

Implementing a technical artifact based on the derived design requirements and HRI specifications means creating a cognitive architecture that orchestrates all the individual AI systems necessary for this use case. This alone is a complex endeavor. Enabling the architecture for value sensitivity further raises the bar of complexity. To tackle this challenge, we envision a cognitive interaction system utilizing control theory.

5 Control Theory as Concept of a Cognitive Interaction System

Values are mainly visible through the actions a robot takes while interacting with a human. Therefore, we focus on HRI and see everything the robot does as an action, such as speaking kindly, moving side by side with people, showing empathic facial expressions, etc. and the humans' responses as reaction.

In order to translate value-sensitive behavior to a technical level, we apply control theory, because a system that reacts on actions can be modeled in control theory to control the system's state. Control theory has the objective to develop a model governing a system input to drive the system to a desired state. Such a system can be described in the following Eq. 1:

$$A(y) = f(v) \tag{1}$$

Whereby y represents the unknowable state of the system we are willing to control. On the other hand, $v \in \mathrm{U}_{ad}$ is the control, the variable we can choose to act and change the state. If $A: D(A) \subset Y \longmapsto Y$ and $f: \mathrm{U}_{ad} \longmapsto Y$ are two linear or nonlinear functions where the operator A determines the equation that must be satisfied by the state y, D the set of decision, Y the solution space and all possible states, and U_{ad} the set of all controls. Therefore, the function f indicates how the control v effects the state of the system [29].

The main goal of control theory is to model the system to be controlled by finding the controls v, that lead to an associated state $y(v)$. Such a model allows to build a control system managing, manipulating, or directing the state of the system to be controlled into a desired state. There are two fundamental types of control systems: the open-loop control system and the closed-loop control system. We focus on the closed-loop control because it includes the feedback concept that is essential in HRI to reflect which values to emphasize in the robot's behavior based on the humans' reaction to the robot. Therefore, the state of a system is fed back to the control system, which continuously determines the control v based on the feedback of the system to affect the state y towards a desired state.

Such a control system could also be applied to HRI, if we see the *interaction* as the system to be controlled, the *set of values to be respected* as the state of the system and the *robot's actions* as control variables. Important is that human (re)action also influences the state of the interaction in regard to the respected values and therefore, creates a new state within the interaction system.

The observation of the human's (re)action using sensors builds the feedback as depicted in Fig. 1. In such an interaction system, the robot's behavior or actions are decided by the control system based on the current state, represented by the respected values by the human's reaction, with the goal to reach a new desired state of respected values. Modern control systems are carried out in the state space and can deal with multiple-input and multiple-output (MIMO) systems, which is important in the application of control theory in HRI and control multiple values by multiple actions. Therefore, actions may affect several respected values. With state spaces, a robot is able to respect multiple values with tensions in the action to take and read multiple values out of the humans reaction to find a new desired state of the interaction to choose the next action.

Imagine an interaction between a robot and a human, where the robot is introduced to the human. The robot first approaches the unknown human with low emphasis on the value of security but high emphasis on the value of openness. However, the camera notices that the human backs away and the control system adjusts the desired values of security and openness for this interaction, interpreting the backing away as an indication for unsecure feelings of the human and, thus, as a need to increase attention to the value of security. Instead of forcing a certain distance to the human, the robot actions are adjusted to the human's reaction and the robot stops to keep a secure distance and emphasizes the value of security more than before while keeping the value of openness as high as before in the actions.

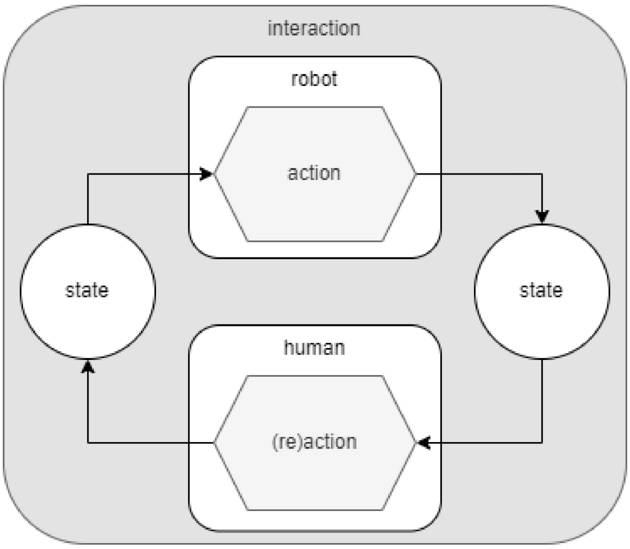

Fig. 1. Interaction feedback loop

In the context of HRI, it is important to have room for fluctuations of the values within the interaction system to appropriately represent a desired value set in the states. Forcing the state of the interaction into certain set of respected values can lead to negative or opposite effects. A robot cannot just act happy to force happiness into an interaction with an upset human. The robot should rather try to express respect for values in small steps to incrementally lead the interaction into a desired set of respected values and avoid any enforcement of values. Fernandez-Cara and Zuazua [29] stating in their summary to control theory the following:

> *"To control a system arising in Nature or Technology, we do not have necessarily to stress the system and drive it to the desired state immediately and directly. Very often, it is much more efficient to control the system letting it fluctuate, trying to find a harmonic dynamics that will drive the system to the desired state without forcing it too much."*

We propose to design a cognitive system with values in mind as a decision-making system (DMS) based on control theory and control systems. Perception can be defined as the human fundamental form of cognitive contact to the world, while it refers in robotics to the ability to perceive the environment [30]. Therefore, the robot can estimate on the sentiment of (re)action of the human how the robot's action affected the system. Cognition is defined as the ability of humans to know, learn, and understand things. Hence, designing cognitive systems implies making them capable of reasoning about their actions [30]. Therefore, the robot must choose the next action accordingly to get closer to the desired state, the prescribed set of values to respect.

6 Discussion

In the present work, we highlighted the need for considering human values in social robot design and outlined our approach to incorporating values in a social robot application in a nursing home. Finally, we present a novel cognitive interaction system based on VSD and control theory.

In particular, we employ VSD to account for human values. A major differentiation of our research from other VSD robotics work (e.g., [15]) is that we strongly strive for an integrative approach as recommended by VSD. We deliberately iterate the value analysis through conceptual and empirical investigations because it is particularly important to understand how the conceptually identified values are manifested in the usage scenario and allows stakeholders to express context-specific values. The same applies to the derivation of design requirements. Through conceptually deriving requirements by using the value hierarchy [25] and empirically with stakeholders, we can triangulate these derived design requirements, thereby creating a design that better reflects the stakeholder values. Furthermore, while previous works only derive design requirements but do not specify how to design HRI based on these requirements [15, 23], our research proposes the use of design patterns, which facilitate the specification of HRI across various robotics platforms and ease the technical development.

Control theory has primarily found applications in physical human robot interaction [30–33]. In contrast, our research explores the possibilities to apply this theory in social interactions by defining the interaction as the system to be controlled and the respected values within the interaction as the states. Notably, the application of control theory and the implementation of a control system in social HRI comes with its own challenges. The value preferences underlying the state of an HRI, must be prescribed and measurable. Sentiment analysis of the human's behavior (speech, mimics, gestures) might allow for a measurement, but concluding from human behavior to a desired set of values including their appropriate emphasis is a difficult endeavor. While subtle emotional human twists may probably be impossible to detect by the robot in the first place, we intend that the robot is able to distinguish states that indicate, for example, emergency situations, boredom, fear, and satisfaction to appropriately react to these. We assume we can address this challenge with the application of fuzzy logic, but further research must be carried out to validate this assumption.

7 Conclusion

In light of the pressing issues presented in long-term care, social robots have the potential to support the care givers in their daily tasks and enhance the quality of life for care receivers. Given the fundamental role of values in user acceptance of robots, we urge all researchers and practitioners to incorporate values in robot design as a whole. The presented conceptual paper gives an insight into our research approach to value-sensitive social robots and outlines how to technically tackle the value integration based on control theory. As part of our future work, we aim to validate our approach to operationalizing values in HRI and refine the proposed technical concept.

References

1. OECD: Who Cares? Attracting and Retaining Care Workers for the Elderly, https://www.oecd-ilibrary.org/content/publication/92c0ef68-en, (2020).
2. Breazeal, C.: Socially intelligent robots. Interactions. 12, 19–22 (https://doi.org/10.1145/1052438.105245
3. Cheon, E., Su, N.M.: Integrating Roboticist Values into a Value Sensitive Design Framework for Humanoid Robots. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 375–382 (2016). https://doi.org/10.1109/HRI.2016.7451775
4. Friedman, B., Hendry, D.G.: Value Sensitive Design: Shaping Technology with Moral Imagination, (2019).
5. International Organization for Standardization: ISO 8373:2021, Robotics – Vocabulary, https://www.iso.org/obp/ui/en/#iso:std:iso:8373:ed-3 v1:en.
6. Sarrica, M., Brondi, S., Fortunati, L.: How many facets does a “social robot” have? A review of scientific and popular definitions online. Inf. Technol. People. 33, 1–21 (2019). https://doi.org/10.1108/ITP-04-2018-0203
7. Moyle, W., Jones, C.J., Murfield, J.E., Thalib, L., Beattie, E.R., Shum, D.K., O’Dwyer, S.T., Mervin, M.C., Draper, B.M.: Use of a robotic seal as a therapeutic tool to improve dementia symptoms: a cluster-randomized controlled trial. J. Am. Med. Dir. Assoc. 18, 766–773 (2017
8. Kim, G.H., Jeon, S., Im, K., Kwon, H., Lee, B.H., Kim, G.Y., Jeong, H., Han, N.E., Seo, S.W., Cho, H.: Structural brain changes after traditional and robot-assisted multi-domain cognitive training in community-dwelling healthy elderly. PloS One. 10, e0123251 (2015).
9. Banks, M.R., Willoughby, L.M., Banks, W.A.: Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. J. Am. Med. Dir. Assoc. 9, 173–177 (2008).
10. Jayawardena, C., Kuo, I.-H., Broadbent, E., MacDonald, B.A.: Socially Assistive Robot HealthBot: Design, Implementation, and Field Trials. IEEE Syst. J. 10, 1056–1067 (2016). https://doi.org/10.1109/JSYST.2014.2337882
11. Wang, X., Shen, J., Chen, Q.: How PARO can help older people in elderly care facilities: A systematic review of RCT. Int. J. Nurs. Knowl. 33, 29–39 (2022). https://doi.org/10.1111/2047-3095.1232
12. Persson, M., Redmalm, D., Iversen, C.: Caregivers’ use of robots and their effect on work environment - a scoping review. J. Technol. Hum. Serv. 40, 251–277 (2022). https://doi.org/10.1080/15228835.2021.20
13. de Graaf, M.M.A., Ben Allouch, S., van Dijk, J.: Why Do They Refuse to Use My Robot? Reasons for Non-Use Derived from a Long-Term Home Study. In: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. pp. 224–233. Association for Computing Machinery, Vienna, Austria (2017). https://doi.org/10.1145/2909824.3020236.
14. Šabanović, S.: Robots in Society, Society in Robots. Int. J. Soc. Robot. 2, 439–450 (2010). https://doi.org/10.1007/s12369-010-0
15. Umbrello, S., Capasso, M., Balistreri, M., Pirni, A., Merenda, F.: Value Sensitive Design to Achieve the UN SDGs with AI: A Case of Elderly Care Robots. Minds Mach. 31, 395–41 (2021). https://doi.org/10.1007/s11023-021-09561-y.

16. Ros, M., Schwartz, S.H., Shoshana, S.: Basic Individual Values, Work Values, and the Meaning of Work. Appl. Psychol. 48, 49–71 (1999).
17. Maio, G.R.: The Psychology of Human Values. Routledge, London (2016).
18. Persson, M., Redmalm, D., Iversen, C.: Caregivers' use of robots and their effect on work environment - a scoping review. J. Technol. Hum. Serv. 40, 251–277 (2022). https://doi.org/10.1080/15228835.2021.2000554.
19. Turja, T., Aaltonen, I., Taipale, S., Oksanen, A.: Robot acceptance model for care (RAM-care): A principled approach to the intention to use care robots. Inf. Manage. 57, 103220 (2020). https://doi.org/10.1016/j.im.2019.103220.
20. Carros, F., Meurer, J., Löffler, D., Unbehaun, D., Matthies, S., Koch, I., Wieching, R., Randall, D., Hassenzahl, M., Wulf, V.: Exploring Human-Robot Interaction with the Elderly: Results from a Ten-Week Case Study in a Care Home. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–12. ACM, Honolulu HI USA (2020). https://doi.org/10.1145/3313831.3376402.
21. Schmiedel, T., Zhong, V.J., Eyssel, F.: Towards a Wave Approach for Value Sensitive Design in Social Robotics. In: Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction. pp. 592–596. Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3568294.3580154.
22. eerink, M., Kröse, B., Evers, V., Wielinga, B.: Assessing Acceptance of Assistive Social Agent Technology by Older Adults: the Almere Model. Int. J. Soc. Robot. 2, 361–375 (2010). https://doi.org/10.1007/s12369-010-0068-5
23. van Wynsberghe, A.: Designing Robots for Care: Care Centered Value-Sensitive Design. Sci. Eng. Ethics. 19, 407–433 (2013). https://doi.org/10.1007/s11948-011-9343-6.
24. Rider, E.A., Kurtz, S., Slade, D., Longmaid, H.E., Ho, M.-J., Pun, J.K., Eggins, S., Branch, W.T.: The International Charter for Human Values in Healthcare: An interprofessional global collaboration to enhance values and communication in healthcare. Patient Educ. Couns. 96, 273–280 (2014). https://doi.org/10.1016/j.pec.2014.06.017.
25. van de Poel, I.: Translating Values into Design Requirements. In: Michelfelder, D.P., McCarthy, N., and Goldberg, D.E. (eds.) Philosophy and Engineering: Reflections on Practice, Principles and Process. pp. 253–266. Springer Netherlands, Dordrecht (2013). https://doi.org/10.1007/978-94-007-7762-0_20.
26. Kahn Jr., P.H., Freier, N.G., Kanda, T., Ishiguro, H., Ruckert, J.H., Severson, R.L., Kane, S.K.: Design Patterns for Sociality in Human-Robot Interaction. In: HRI 2008 - Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction: Living with Robots. pp. 97–104 (2008). https://doi.org/10.1145/1349822.1349836.
27. Sauppé, A., Mutlu, B.: Design patterns for exploring and prototyping human-robot interactions. In: Conference on Human Factors in Computing Systems - Proceedings. pp. 1439–1448 (2014). https://doi.org/10.1145/2556288.2557057.
28. Pollmann, K., Ziegler, D.: A Pattern Approach to Comprehensible and Pleasant Human–Robot Interaction. Multimodal Technol. Interact. 5, 49 (2021). https://doi.org/10.3390/mti5090049.
29. Fernández Cara, E., Zuazua Iriondo, E.: Control theory: history, mathematical achievements and perspectives. (2003).
30. Maroto-Gómez, M., Alonso-Martín, F., Malfaz, M., Castro-González, Á., Castillo, J.C., Salichs, M.Á.: A Systematic Literature Review of Decision-Making and Control Systems for Autonomous and Social Robots. Int. J. Soc. Robot. 15, 745–789 (2023). https://doi.org/10.1007/s12369-023-00977-3.
31. Yu, H., Huang, S., Chen, G., Pan, Y., Guo, Z.: Human–Robot Interaction Control of Rehabilitation Robots With Series Elastic Actuators. IEEE Trans. Robot. 31, 1089–1100 (2015). https://doi.org/10.1109/TRO.2015.2457314.

32. Li, Y., Sena, A., Wang, Z., Xing, X., Babič, J., Asseldonk, E. van, Burdet, E.: A review on interaction control for contact robots through intent detection. Prog. Biomed. Eng. 4, 032004 (2022). https://doi.org/10.1088/2516-1091/ac8193.
33. Mukherjee, D., Gupta, K., Chang, L.H., Najjaran, H.: A Survey of Robot Learning Strategies for Human-Robot Collaboration in Industrial Settings. Robot. Comput.-Integr. Manuf. 73, 102231 (2022). https://doi.org/10.1016/j.rcim.2021.102231.