Master projects in social robotics (and more)

Paul Vogt

p.a.vogt@rug.nl

1 Introduction

Master projects for AI and CCS can be done in the field of social robotics or human-robot interaction (HRI). The projects I propose mostly concern research (and development) about the way social robots can communicate effectively and meaningfully with humans in an autonomous manner for specific groups of people in several application domains. Think, for example, of robots interacting with children in educational settings (e.g., second language learning, cf. http://l2tor.eu), or with elderly people in healthcare settings.

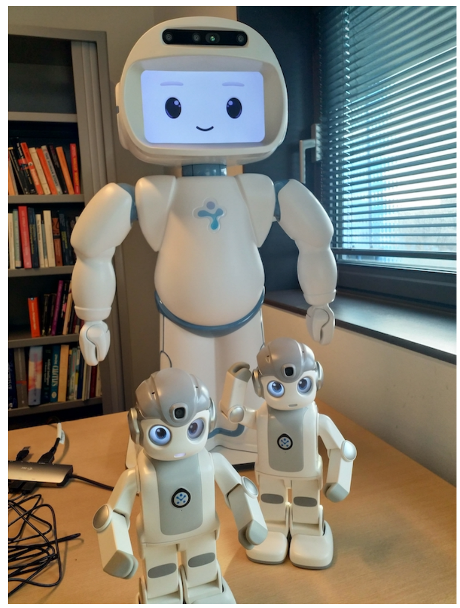

Figure 1: QT and Alpha Mini robots.

The development of social robots is advancing fast, but their ability to interact with humans –especially with a-typical populations– is still limited in scope, form and effectiveness. For example, automatic speech recognition (ASR) is often flawed, except for most adult speech when processed through the cloud, which slows down interactions and thus effectiveness and acceptability of the communication. Also, non-verbal interaction (e.g., through hand gestures), considered essential for effective communication with humans, is hard to achieve in a meaningful manner. Another challenge is to design the HRI such that the robot can understand the human behaviour and can understand whether the human understands the robot, and that the robot responds in an appropriate manner. Furthermore, the robot's environment is changing continuously either due to its own behaviour, that of the human or through some other external source. The robot should be able to adapt to this dynamic environment. Finally, especially to facilitate long-term interaction, the robot should learn to interact with humans in a personalised manner for which it should be able to develop some form of common ground.

Below, these issues are clarified in more detailed themes. The work will –in most cases– be done with either the Alpha Mini robot or the QT robot (see figure). Where the Alpha Mini robot is programmed in either Note that the Alpha Mini robot can be programmed in Python, Javascript or Android, and the QT robot in ROS. Strong programming skills in ROS are required for working with the QT robot.

2 Themes

2.1 Grounding meaningful and effective communication

When communicating with humans meaningfully and effectively, social robots face a major challenge, namely *grounding* whether the interaction was successful. As in human-human interaction, the robot must understand the human's verbal and non-verbal signals and should make sure that the human knows that it understood these signals. Vice versa also holds: the robot should make sure that it knows whether or not the human understood the communication acts produced by the robot. This is what Herb Clark calls grounding in communication [3, 4].

The main research question for this theme is: How can we design algorithms for robots that allow the robot to ground their communication with humans (in the sense of Clark)? This may be related to a research project on social robots for elderly care [1], in child-robot interaction [12, 14], but also more fundamentally with lab experiments involving other people. An interesting approach that could serve as a starting point is provided by Kopp and Krämer (2021) [8] who suggest that simulating *common ground* is a key to achieve this. Issues that may be addressed include modelling:

- Theory of Mind / common ground.
- repair strategies in communication
- non-verbal signals (e.g., backchannels, beat gestures)
- turn-taking

2.2 Understanding human behaviour - a data-driven approach

Somewhat related is the challenge of understanding of human behaviour from perception, as this is to a large extent required to ground communication. While humans behave in certain standard ways, variation is always there. Humans behave in numerous ways using many different body parts. Variation occurs on an individual level, even when they repeat a certain behaviour there are nearly always some detectable differences. Variation is enhanced when considering different individuals. Nevertheless, humans can easily recognise other humans' behaviour.

For effective human-robot interaction, the robot should be able to recognise those behaviours essential to respond appropriately. How can a robot recognise that the user is engaged with it? How can the robot recognise the user's gestures? How can the robot recognise that the user responds appropriately or inappropriately to its instructions? To investigate such questions, it important to investigate how humans behave when interacting with robots and use this information to apply machine learning techniques to allow the robot to recognise their behaviour.

To achieve this, annotated human-robot interaction data sets would be very useful. Previously, we have developed a game of charades with a NAO robot to collect data of non-verbal human behaviour, and use machine learning to recognise this [5]. While interesting, these data contains some very specific behaviours that do not occur frequently in normal interactions. One way to go is to develop another interactive, and playful strategy to elicit machine-trainable data of human-robot interactions.

2.3 Language grounding in social robots

The Large Language Models (LLMs) such as ChatGPT, offer new opportunities to improve human-robot conversation [7, 6]. However, these offer also many new challenges, since the LLMs are typically not trained to have face-to-face interactions as expected for social robots. Conversations between humans and robots based on LLMs are therefore often considered superficial and boring. Moreover, the conversations are not grounded in the physical/social context in which the conversations take place. How can we improve the conversations using LLMs? How can we connect LLMs to the robot's physical world (e.g., reading human's state or by perceiving some task environment)?

2.4 Co-speech gesturing for social robots

Humans use gestures while they are speaking to support their own speech processing or to support the addressee understanding their own intentions. Typically, such co-speech gestures are well synchronised with speech. Now that Large Language Models provide opportunities to model human-robot interactions to a higher level, question is how to generate gestures to accompany the speech. What gestures to accompany which words? How to turn them into motor programmes? How to synchronise them with speech?

2.5 Designing social robots for the elderly

The world's population is growing older, which poses all kinds of challenges to facilitate the well-being of the elderly people. This group of people often feel isolated or lonely, and/or have physical, cognitive and emotional issues, for which they need support (e.g., [2]). However, (in)formal care is hard to organise due to a lack of time or lower numbers of professional carers. Social robots may help to alleviate some of the issues this populations face.

In this theme, research is carried out how to design social robots, such that they can support the well-being of the elderly people. This may be done through a user-centered design where one studies how robots should function to support elderly with certain issues, and to iteratively develop a prototype and evaluate this with the projected users, which –in addition to elderly people– may include (in)formal carers. Another approach could be to improve technology to facilitate the quality of interactions between a robot and the elderly.

2.6 Common ground

When robots engage with humans for prolonged periods of time, the robot and human create a history with each other. Moreover, even before they start interacting with each other, both robot and human may have shared knowledge about the world and –possibly– each other. Roughly this shared knowledge is called *common ground* [3, 15] if both are aware that they share this knowledge. This knowledge can be very rudimentary, e.g., the knowledge that robots and humans can talk, but also quite complex, e.g., have knowledge how the other (robot or human) typically behaves in certain situations. The premise that both robot and human should be aware that the other shares the knowledge and that they are aware that the other knows that the other is aware of this sharing, requires them to have some *Theory of Mind*.

In this theme, the question we try to answer is how to model this in a robot? How can a robot build up a form of common ground? For this, research is required into what kind of knowledge do people have about robots, and –more challenging– into methods how a robot can learn to memorise interactions with humans.

2.7 Other

Other themes that relates to my research interest concern the following:

- **Language development in robots** How can we design a robot that learns language in a similar way that human children do? See, e.g., [9]
- **Agent-based modelling of language acquisition** How can simulated child-agents learn language from language-competent agents? The idea here is to design an ABM based on an annotated corpus of multimodal interactions surrounding infants, cf. [13].
- **Modelling language emergence** For ideas, see the following papers [10, 11].

References

[1] R. Balalic, S.and Sanderman, P. Vogt, N. Degens, and M. Hagedoorn. Deployment of a social robot in the care for older adults. In *The First Workshop on Care Robots for Older Adults*, 2022.

[2] Cynthia L Breazeal, Anastasia K Ostrowski, Nikhita Singh, and Hae Won Park. Designing social robots for older adults. *Natl. Acad. Eng. Bridge*, 49:22–31, 2019.

[3] H. H. Clark. *Using Language*. Cambridge University Press, 1996.

[4] Herbert H Clark and Susan E Brennan. Grounding in communication. In *Perspectives on socially shared cognition.*, pages 127–149. American Psychological Association, 1991.

[5] Jan de Wit, Emiel Krahmer, and Paul Vogt. Introducing the NEMO-Lowlands iconic gesture dataset, collected through a gameful human–robot interaction. *Behavior Research Methods*, 53(3):1353–1370, 2021.

[6] Bahar Irfan, Sanna-Mari Kuoppamäki, and Gabriel Skantze. Between reality and delusion: Challenges of applying large language models to companion robots for open-domain dialogues with older adults. 2023.

[7] Weslie Khoo, Long-Jing Hsu, Kyrie Jig Amon, Pranav Vijay Chakilam, Wei-Chu Chen, Zachary Kaufman, Agness Lungu, Hiroki Sato, Erin Seliger, Manasi Swaminathan, et al. Spill the tea: When robot conversation agents support well-being for older adults. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, pages 178–182, 2023.

[8] Stefan Kopp and Nicole Krämer. Revisiting human-agent communication: The importance of joint co-construction and understanding mental states. *Frontiers in Psychology*, 12:580955, 2021.

[9] Akira Taniguchi, Tadahiro Taniguchi, and Angelo Cangelosi. Cross-situational learning with bayesian generative models for multimodal category and word learning in robots. *Frontiers in neurorobotics*, 11:66, 2017.

[10] P. Vogt. The emergence of compositional structures in perceptually grounded language games. *Artificial Intelligence*, 167(1–2):206–242, 2005.

[11] P. Vogt. Modelling interactions between language evolution and demography. *Human Biology*, 81:237–258, 2009.

[12] Paul Vogt, Mirjam De Haas, Chiara De Jong, Peta Baxter, and Emiel Krahmer. Child-robot interactions for second language tutoring to preschool children. *Frontiers in Human Neuroscience*, 11:73, 2017.

[13] Paul Vogt and J. Douglas Mastin. Anchoring social symbol grounding in children's interactions. *Künstliche Intelligenz*, 27:145–151, 2013.

[14] Paul Vogt, Rianne van den Berghe, Mirjam De Haas, Laura Hoffman, Junko Kanero, Ezgi Mamus, Jean-Marc Montanier, Cansu Oranç, Ora Oudgenoeg-Paz, Daniel Hernández García, et al. Second language tutoring using social robots: a large-scale study. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 497–505. IEEE, 2019.

[15] Deanna Wilkes-Gibbs and Herbert H Clark. Coordinating beliefs in conversation. *Journal of Memory and Language*, 31(2):183–194, 1992.