

1.2.1 – Considering human/machine interactions

Practical guidance – social care

Author: Dr Catherine Menon, University of Hertfordshire

Assistive robots support independent living by aiding humans to conduct basic activities. As part of this they must perform functions important to safety, such as alerting a user if an appliance has been left on, alerting a user if medication has not been taken, or encouraging a user to perform necessary rehabilitative or medical actions. To do this effectively, assistive robots must also demonstrate behaviour which is sufficiently socially acceptable for the end-user to fully engage with them. This link between social credibility of the robot, the user's engagement with the robot, and effective performance of safety-related behaviours has been demonstrated in an experimental study supported by the AAIP through a demonstrator project.

Example Assistive Robot

The Care-O-Bot [1] is an up-to-date example of a mobile assistant robot with the capacity for social interaction. The Care-O-Bot is typically expected to perform a range of functions including:

- Alerting a user if an electrical appliance is malfunctioning
- Reminding a user to take their medication
- Reminding a user if an appliance has been left on

In addition to these care-giving behaviours, the Care-O-Bot is expected to encourage the user to engage and interact by offering entertainment and companionship.

The Care-O-Bot, like all assistive robots, is designed with reablement as a priority [2]. Reablement is defined as the drive to "Support people to do rather than doing to / for people" and is an important characteristic for service and assistive robots. Designing with reablement in mind means that the assistive robot is not intended to carry out the tasks itself (e.g., administering medicine to a user), but is instead intended to encourage the user to complete the task themselves. A direct result of this design principle is that the assistive robot will alert the human user to a potential hazard (e.g. an appliance that has been left on) but the user must take action themselves to complete the mitigation of any safety risk. That is, the human user must switch off the oven, repair the appliance, take the medication etc. as necessary, rather than relying on the robot to do this. In this way, the mitigation of safety risks is split between the assistive robot and the human user, and hence the effectiveness of safety performance is directly related to the extent which a user is willing to engage with the robot.

Effects of Social Credibility

Social credibility is a measure of how well an assistive robot obeys the social norms relevant to its environment. These social norms will be specific to the environment, and for a domestic assistive robot may include:

- Frequency and urgency of any interruptions
- Nature and intensity of interaction, engagement and interruptions
- Responsiveness of the robot to verbal and non-verbal feedback
- Appropriate physical movement and distance maintained from end-user

Social credibility is an evolving measure that is dependent on the actions of the robot. Social credibility may be temporarily lost by an inappropriate action, and gained back by subsequent actions. A loss of social credibility (from any cause) can lead to an end-user disengaging with the robot, choosing either to ignore its prompts or to switch it off. Studies have shown that users are more willing to switch off robots if they consider them to be solely robotic devices, instead of intelligent, social beings [3]. This is exacerbated when the mode of engagement with this robot becomes arduous. In [4] drivers concluded that they would prefer to be able to turn off a speed warning system that was judged irritating, even where they agreed that use of the technology would be helpful.

Because assistive robots, by design, rely on the end user to complete mitigation of an identified safety risk, user disengagement compromises the ability of these robots to perform their safety-critical functions. For example, a robot reminding the user that the oven has been left on has no effect unless the user engages with the robot, and moves to switch the oven off. This is particularly true where the robot alert contradicts some existing mental model that the user has of the environment (“I didn’t leave the oven on”). In the aviation domain – where autonomous cockpit systems are not considered to be social entities – pilots have been observed to attempt to debug the automation when its actions deviate from those they expected [5]. This is also a risk for assistive robots when the end user considers them to be monitoring devices only, instead of social, intelligent beings.

An experiment has been conducted to identify and characterise the link between social credibility and safety, and a summary of this is provided in the Appendix.

Means to Address Social Credibility

There are a number of potential methods to address loss of safety-critical functionality resulting from lowered social credibility. Each of these methods trades a slight decrease in the robot's overall capability in return for maintaining an adequate level of social credibility. Since social credibility is a requirement for effective safety critical performance, this corresponds to decreasing the robot's capabilities in order to gain confidence that safety-critical engagements will be performed effectively when needed. When the social credibility drops below a threshold value (termed the disengagement threshold), the robot should alter the nature of its alerts and reminders to stop social credibility loss. For example, the robot may identify those alerts which are not safety-critical and choose to:

- Avoid performing the alert entirely
- Delay the alert or perform it less frequently
- Slow its physical movements when coming to interrupt a user
- Decrease the volume of any audible alerts

This allows a robot to omit routine behaviours (such as interrupting the user with the offer of food or drink) in order to retain sufficient social credibility to ensure that any safety-critical behaviour (such as notification the oven is on) will be engaged with by the user.

References

- [1] R. Kittmann, T. Fröhlich, J. Schäfer, et al., “Let me introduce myself: I am care-robot 4, a gentleman robot,” in *Mensch und Computer 2015 – Proceedings*, S. Diefenbach, N. Henze, and M. Pielot, Eds., Berlin: De Gruyter Oldenbourg, 2015, pp. 223–232.
- [2] F. Amirabdollahian, R. op den Akke, S. Bedaf, et al., “Assistive technology design and development for acceptable robotics companions for ageing years.,” *Paladyn: Journal of Behavioral Robotics*, vol. 4, no. 2, pp. 94–112, 2013.
- [3] C. Bartneck, T. Kanda, O. Mubin, et al., “Does the design of a robot influence its animacy and perceived intelligence?” *International Journal of Social Robotics*, vol. 1, no. 2, pp. 195–204, 2009.
- [4] J. Wall, V. Cuenca, L. Creef, et al. “Attitudes and opinions towards intelligent speed adaptation,” in *Intelligent Vehicles Symposium Workshops (IV Workshops)*, 2013 IEEE, IEEE, 2013, pp. 37–42.
- [5] N. B. Sarter and D. D. Woods, “Team play with a powerful and independent agent: Operational experiences and automation surprises on the airbus a-20,” *Human Factors*, vol. 39, no. 4, pp. 553–569, 1997.
- [6] C. Carpinella, A. Wyman, M. Perez, S. Stroessner. “The robotic social attributes scale (ROSAS): Development and validation”, in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 254-262, 2017.
- [7] C. Menon, P. Holthaus. “Does a loss of social credibility impact robot safety?”, in *Proceedings of the 9th International Conference on Performance, Safety and Robustness in Complex Systems*, pp. 18 – 25, 2019.