

# What do care robots reveal about technology?

Rieks op den Akker

*Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, The Netherlands*

**Abstract**—Ethical issues raised by the idea of social robots that care point at a fundamental difference between man and machine. What sort of “difference” is this? We propose a semiotic view on technology to clarify the relations users have with social robots. Are these autonomous agents just promising or can we also count on them?

## 1. INTRODUCTION

If a “smart” coffee machine knows about its user’s heart problems, should it accept giving him a coffee when he requests one? The issue is raised in “Ethical Things” a project that “explores the effects of autonomous systems of the future.”<sup>1</sup> Similar ethical issues raised by the idea of autonomous care robots were discussed in the Accompany project, one of the many EU projects in the field of social robotics for elderly care.<sup>2</sup>[1].

Social robots challenge our traditional theories of moral responsibility. Are they moral agents? Can they be held responsible? In this short note I invite the reader to take a look behind these type of ethical issues raised by the growing autonomy of our intelligent technical artifacts of which the social robots are the most impressive representatives. Can we perceive robots as social responsible autonomous companion agents that care and at the same time as technical instruments? How can we understand social robots from the principles of technology? And what do users that report about their interactions with social robots tell us about the limitations of technology that follow from these principles?

## 2. ROBOT ETHICS AND ETHICAL ROBOTS

People have different views on the moral issues raised by autonomous artifacts like robots and what they mean for their application in for example health care practice. Implicit in these views is an idea about what technology can accomplish which is based on ideas about what technology is, about the relation between mind and matter in men and in the machine. The emphasis in the usual approach in robot ethics research is “on the robot and what the robot really is or thinks”, in order to be able to answer questions like “Are robots intelligent, rational, ‘moral agents’?” or “it limits ethics to concerns about things that might go wrong in interactions with robots.” “For many moral philosophers, ethics is about holding someone responsible and about the rightness of one’s actions, and then questions regarding moral status and action

are central. We usually ascribe moral responsibility only to beings that have a sufficient degree of moral agency - whatever that means- and ask about the rightness of what that agent does, has done, or could do.” [2]. Coeckelberg proposes a human centric or interaction centric approach to the ethics of robot technology. “Instead of a philosophy of mind concerning what robots really are or really (can) think, let us turn to a philosophy of interaction and take seriously the ethical significance of appearance.”([3], p.220).

One of the outcomes of the Accompany focus group discussions was that control over the programming of the robot needed to be a negotiation between the older person living with the robot, and that person’s other support networks of formal and informal carers, rather than simply implementing an older person’s wishes. However, the data also suggests that at least one approach - the ‘let’s do it together’ strategy may itself undermine autonomy by (unconsciously, perhaps) infantilising the older person [1].

I will argue that what is needed for ethical decisions is an open dialogue between partners involved; a dialogue that takes into account the specific situation in which a decision has to be made. Ethical issues are raised when we become aware of a conflict between general rules of good conduct, between different values, autonomy and safety for example. “Open” means that there is no protocol that is forced upon the dialogue partners. A robot would be social when it would *take* responsibility, not because it is *ascribed* responsibility. Someone who is just following a procedure, as computers and clergymen do, is not responsible since he does not at the same time reflect critically on the appropriateness of the procedure, a reflection that should be based on sensitivity for the values that are important in the particular situation at hand. Sometimes we must leave things for others to do. Trust is okay, but not blind trust. Responsibility is a virtue, not a commodity that can be given away.

Moor argues that “explicit ethical robot agents can decide what to do in a conflict situation.” [4]. But also then we can only implement general rules. They need to be applied in a careful way. “The human act of caring is the recognition of the intrinsic value of each person and the response to that value” (Schoenhofer). From the patient’s view point care values are safety, satisfaction, responsiveness to care, dignity, physical and psychological well-being. Values of the analytical, empirical scientific view are quite different: structurability, reproducibility, analysability. For modern technology we can add computability, programmability. The designer of (social) technology makes user models *and assumes programmability* of the user, who adheres to the

<sup>1</sup><http://www.creativeapplications.net/objects/ethical-things-the-mundane-the-insignificant-and-the-smart-things/>

<sup>2</sup>In Accompany a robotic companion was developed for providing services to elderly users in a motivating and socially acceptable manner to facilitate independent living at home. (<http://accompanyproject.eu/>)

models underlying the user interface of the system. Although tailoring is a hot topic in the field of intelligent software agents, from a designers perspective the user remains an *abstract* entity. For the care giver the unique person he cares about is the one who determines what has to be done in a concrete situation.

### 3. DIALOGUE AND RESPONSIBILITY

In everyday life we encounter each other as persons. What makes man a person is his rationality, in the sense of accountability. The postulate of rationality is a -contrafactual-principle that partners in a personal dialogue adhere to. According to Kant being accountable, having the will to take responsibility, is what characterizes the moral person. On the contrary, things are those objects that can not take responsibility<sup>3</sup>.

Note that ‘man is rational’ is not meant here as an empirical statement, but a contrafactual postulate. When we are engaged in a dialogue we must assume that it holds and we must act accordingly so it becomes reality. This postulate is constitutive for the dialogue: without this there is no dialogue between persons possible. Even when someone lies we assume that he will have an explanation for it. We have to take seriously that the other says something. This is the first postulate of dialogue. Being accountable is thus characteristic for being rational.

What do users’ experiences tell us about the interaction with artificial companions? Bickmore et al. study long term relationship between embodied conversational agents and elderly people [6]. “Several participants mentioned that they could not express themselves completely using the constrained interaction. One of them reported: ‘*When she ask me questions ... I can’t ask her back the way I want*’ . [6]. Clearly, users of conversational agents experience that a *real interaction* with the system is not possible. It simulates programmed “social behaviors” but it lacks social competence. The coffee machine that knows about its user’s heart problems and that is confronted with a moral problem: ‘Should I present a coffee or not?’ could start a dialogue with the user and try to convince him. Eventually, questions will come up: ‘Who am I talking to?’ ‘Do you really care?’. The philosopher tries to understand what this reveals about the very idea of technology. How does technology work and serve us? A semiotic approach might help.

### 4. UNDERSTANDING TECHNOLOGY

For understanding the “difference between man and machine” it may help if we think about the difference between the physical sign and the meaning it carries. Machine is “part

<sup>3</sup>“Person ist dasjenige Subjekt, dessen Handlungen einer Zurechnung fähig sind. Die moralische Persönlichkeit ist also nichts anderes als die Freiheit eines vernünftigen Wesens unter moralischen Gesetzen (die psychologische aber bloss das Vermögen, sich der Identität seiner selbst in den verschiedenen Zuständen seines Daseins bewusst zu werden); woraus dann folgt, dass eine Person keinen anderen Gesetzen als denen die sie (entweder allein oder wenigstens zugleich mit anderen) sich selbst gibt, unterworfen ist.” “Sache ist ein Ding, was keiner Zurechnung fähig ist. Ein jedes Objekt der freien Willkür, welches selbst der Freiheit ermangelt, heisst daher Sache (res corporalis)”, [5], Einl. IV (III 26 f.)

of” an intelligent relation; without the human intellect it has no meaning. Just like a sign without a meaning is not a sign. The physical presentation and its form is on the one hand arbitrary (there is no intrinsic relation between the meaning of a word and how the words looks or sounds), on the other hand it is conventional and historically motivated (to be understood you need to learn the language of a community). In the same way machines are outside objectifications of our intellect. As technical means they mediate between men and nature. They are based on forces of the physical nature and on the forces of social psychological nature.

Computers are language machines. Suppose we talk to a machine and ask “What time is it?” and the machine answers “It is 2 o’clock in the afternoon.” How does this work? This works because of the implemented *correspondence* between the structure of the physical process that my talking (also) is and the meaning I express. Natural language is the socially shared interface we use to express our thoughts, emotions, commands. By making the machine react to sequences of tokens specified in a formal system, tokens *that we choose to resemble the words and sentences in our own natural language*, and by making the machine generate sentences in a situation that satisfies certain felicity conditions we bring about the user experience of having to do with an understanding machine. The social robot by uttering some natural sounds and by showing some natural behaviours promises to be of our natural kind.

### 5. CONCLUSION

We propose a semiotic view on modern technology and understand technological beings essentially as *outside objectifications* of our intellectual meaningful relations in social practices. The semiotic view on modern technology suggests a conceptual framework for thinking about the moral issues raised by social robots. It reveals the fundamental limitations of any technical system however “smart”. It is our responsibility to see these limitations when we use a system. In thinking about morality in technology we should carefully distinguish between the general abstract *value free* technical ideas and their application in devices used in concrete value laden situations.

### REFERENCES

- [1] H. Draper, T. Sorell, S. Bedaf, H. Lehmann, C. G. Ruiz, M. Hervé, G. J. Gelderblom, K. Dautenhahn, and F. Amirabdollahian, “What asking potential users about ethical values adds to our understanding of an ethical framework for social robots for older people,” in *Presentation at AISB50 - 50th Annual Convention of the AISB*, 2014.
- [2] M. Coeckelberg, *Growing moral relations: critique of moral status ascription*. Palgrave Macmillan, 2012.
- [3] M. Coeckelbergh, “Personal robots, appearance, and human good: A methodological reflection on roboethics,” *International Journal of Social Robotics*, vol. 1, no. 3, pp. 217–221, 2009, open Access Article. [Online]. Available: <http://doc.utwente.nl/76112/>
- [4] J. H. Moor, “Four kinds of ethical robots,” *Philosophy Now*, pp. 12–14, March/April 2009.
- [5] I. Kant, *Die Metaphysik der Sitten*. Verlag von Felix Meiner, 1907.
- [6] T. W. Bickmore, L. Caruso, and K. Clough-Gorr, “Acceptance and usability of a relational agent interface by urban older adults,” in *Adults, ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*. New York, NY, USA: ACM, 2005, pp. 1212–1215.