Author's copy. Appeared in: Nørskov, M., Seibt J., Quick O. 2020. Culturally Sustainable Social Robotics – Proceedings of Robophilosophy 2020. Series Frontiers of AI and Its Applications, IOS Press, Amsterdam.

The Morality Menu Project

OLIVER BENDEL[a,1]
[a] *School of Business FHNW, Bahnhofstrasse 6, CH-5210 Windisch*

Abstract. The discipline of machine ethics examines, designs and produces moral machines. The artificial morality is usually pre-programmed by a manufacturer or developer. However, another approach is the more flexible morality menu (MOME). With this, owners or users replicate their own moral preferences onto a machine. A team at the FHNW implemented a MOME for MOBO (a chatbot) in 2019/2020. In this article, the author introduces the idea of the MOME, presents the MOBO-MOME project and discusses advantages and disadvantages of such an approach. It turns out that a morality menu could be a valuable extension for certain moral machines.

Keywords. Machine Ethics, Moral Machines, Machine Morality, Chatbot

1. Introduction

Machine ethics "expands upon traditional computer ethics", shifting away the main focus from "what people do with computers to questions about what machines do by themselves" [20]. It investigates whether machines and systems can be moral reasoners or moral agents, and if so, in what form [2, 20]. In the years of its existence, it has already split up into different currents. Several representatives emphasize that the discipline not only researches moral machines, but also creates them, in the form of simulations or prototypes [2, 17, 9, 5]. Increasingly, machine ethics will help in the enhancement of products, such as voice assistants in the smart home, domestic robots in the house and in the garden, and social robots in shopping malls.

The machine morality (or artificial morality) is usually clearly fixed, for example via meta-rules and rules. The (partially or fully) autonomous machine is thus able to perform certain actions and not able to perform others. Manufacturers or developers devise the rules or meta-rules, possibly supported by moral guidelines, ethical codes and ethics committees, facing the problem that models of normative ethics collide and also the morality of individuals, societies and cultures.

An alternative approach, where the machine itself develops or improves morality, for example by adapting the rules, with the help of machine learning and deep learning, is hardly applied. Anderson et al. have customized a Nao robot with machine learning methods and provided it with values [1]. It automatically adapts to the respective situation in the care sector. Of course, this machine morality must also have its origin somewhere, and again different standards and values clash. However, neither of these possibilities is considered here.

[1] Corresponding Author, Book Department, IOS Press, Nieuwe Hemweg 6B, 1013 BG Amsterdam, The Netherlands; e-mail: bookproduction@iospress.nl.

Another approach is the morality menu (MOME). The owners or users of this interface replicate their own moral convictions, their ideas about good and evil, their standards of value, and their rules of conduct onto the machine [3], thus creating machine morality or artificial morality. The moral machine acts in the same way the user would act, in detail. The owners or users may encounter certain default settings, but they have some freedom to change them or set new default settings. The ideal situation for most owners or users is when they do not need any prior programming knowledge but can intervene in a simple and target-oriented manner. In closed or semi-open worlds where the user has sovereignty, e.g. in the household or in the garden, this can be a reasonable approach, whereas service robots in open worlds can raise problems, which will be discussed in chapter 5. This approach is also useful for voice assistants and chatbots that act on behalf of the user and that he or she wants to personalize.[2]

From 2013 to 2019, the author and his student teams developed five artefacts of machine ethics at the University of Applied Sciences and Arts Northwestern Switzerland (FHNW) [4, 6, 7, 8]. For one of these artefacts, the animal-friendly vacuum cleaning robot with the name LADYBIRD, he designed a morality menu that has not yet been implemented [3]. In addition, he sketched another morality menu for a virtual assistant or voicebot in the style of Google Duplex, which can make reservations and orders by phone for its user almost entirely independently [18]. Bendel describes the two prototypes and the MOME design studies in detail [3].

From June 2019 to January 2020, a three-person student team led by the author at the School of Business FHNW implemented a MOME for a chatbot called MOBO.[3] They took up the previous ideas and crystallized them in this project. At first glance, a machine like LADYBIRD and a voicebot are probably better suited for a morality menu because they have a clearly defined task and their actions are directed at other living beings (whereas usually the user of the chatbot makes the settings himself).[4] However, the MOBO-MOME raised interesting new questions, which will be discussed in section 3.1. Moreover, the main focus was to try out the principle, to implement the rules successfully – and not all rules necessarily had to make sense or be applied in practice.

In this article, the author introduces the idea of the morality menu and presents the design, the components and the functions of the MOBO-MOME. Then he discusses the advantages and disadvantages of this implementation and summarizes the results. The interaction between the morality menu and the chatbot was tested several times by the team and was demonstrated in a final presentation and in a video, which will be discussed in chapter 4. An evaluation with user groups did not take place.

---

[2] What is morality anyway? Human morality is a normative basic framework for behaviour, especially towards fellow men, but also towards nature and oneself. Morality and custom form a complex of rules of action, values, and concepts of meaning [13]. Like language, morality is a public system [11], with personal deviations and preferences. This is particularly strong in individualistic societies. In the present context not only morality in the narrower sense is addressed, but also customs and even conventions. These are to be mapped onto the machine so that artificial morality is created. One can speak of a moral machine here, but it would be more precise to call it a machine that can not only do moral and immoral things, but things that are right or wrong.

[3] Both the morality menu and the chatbot were programmed in Python. The programming itself will not be discussed in this article.

[4] Of course, the operator who makes the settings for the user can also come into play here.

2. The Aim of the Morality Menu

The aim of the morality menu is that an owner or user can use it to adapt a machine – e.g., a domestic robot – in such a way that it largely replicates his or her moral ideas and convictions in a specific area [3]. The result is, at best, a proxy machine with a proxy morality elaborated in detail. The machine does what people would do, in their absence or at least without control in this moment. In order to be able to use the morality menu, one has to be clear about what standards of value one has in relation to a specific field of application and what rules of conduct one follows there. It is important that these are not attitudes that are constantly changing, and even then, there is no guarantee that in a dynamic situation people would not choose differently. This challenge will be discussed in chapter 5. Pereira and Saptawijaya have made moral adjustments possible in a virtual simulation in which a robot is supposed to save a princess [17]. For example, the robot can accept greater or lesser collateral damage when reaching its target. There are hardly any other approaches to influencing artificial morality in real operations, except in the context of machine learning [1].

Machine ethics is usually concerned with partially or fully autonomous machines, such as software robots (chatbots or voicebots) or hardware robots (self-driving cars, fighter planes and service robots). When normal machines (if something like this exists at all) become moral ones, they are moralized, so to speak. This moralization is done either by the researcher, inventor, manufacturer, developer, etc., so that the morality menu is an additional option for a moral machine, or completely by the MOME, so that its use will turn a neutral machine into a moral (or immoral) machine. The morality menu will also come from the researcher, inventor, manufacturer, etc., unless you design it completely unlocked so that the user can determine each option. In this respect, it can be located within the area of ethics by design [10], where these machines are differentiated from usual applications by the new alternatives that are granted.

The morality menu should be designed in such a way that one's own morality in relation to a specific field of application can be represented in different aspects and to the greatest possible extent. It must also fit with the respective functions of the machine. It should be easy to use because if, for example, specialist programming skills or knowledge of a markup language are required to define the moral settings, it would defeat the purpose of a mass-market product and uptake would be low. The morality menu should above all contain moral rules to live up to its name, and less pedagogical or psychological ones. Of course, these can be incorporated to a certain extent in the MOME, and as mentioned, customs and conventions also play a role in MOBO-MOME.

In the two MOME concepts (including design studies) developed by the author in 2018, which contain visual and textual elements, virtual sliders are used, as is the case with smartphones and tablets, to activate or deactivate the functions. What happens when the slider is moved to the right is described on the left-hand side of the slider and when it changes its colour, it becomes activated. The drafts did not initially consider what the possible default settings should be. It would be possible that all controls are shifted to the left at the beginning, but other configurations are conceivable.

The technical implementation, especially the interaction between the MOME and the bot or robot, is not discussed in the concepts. It can be about completely different components of the machine, about hardware and software, about movement or speech functions. The purchase of the application is also irrelevant here – it can be supplied with the product or offered for sale in an app store. The smartphone, a tablet, a touch screen built into the device or an external, special display can serve as the terminal.

3. Implementation of the MOME

The MOME project at the School of Business FHNW started in June 2019 and ended in January 2020 [16]. The author of the present article was client and supervisor – and provided the aforementioned preparatory work, e.g., the design studies. Bradley Richards, another FHNW professor, was also involved in programming questions. The team of three students wanted to develop a chatbot and connect it to a hardware robot that can be considered a social robot for a given programming. Because of a defect of hardware components which could not be fixed, it remained a chatbot. The whole application runs on a tablet. The tablet was originally intended as a control element for the robot.

Compared to the two design studies mentioned above, there was a difference in this project. In order to implement certain settings on the morality menu in a meaningful way, something about the user had to be known. Therefore, an input panel was devised, through which he or she gives personal information. The three components of the system are described below, namely a user personality panel, the morality menu, and the chatbot MOBO.

3.1. User Personality Panel

First of all, the user must provide various personal details [16]. He or she opens the "User Personality" panel (see Fig. 1) in the "Menu" and can then enter his or her name, age, nationality, gender, sexual orientation, and hair colour. The term "User Personality" may be problematic because the items in question only give a very limited overview of one's personality – but the choice of the project team can no longer be changed. However, the details are important for communication and interaction with the chatbot. This is because the chatbot needs a concrete name in order to address the other person correctly and a concrete profile, for example to be able to express prejudices.

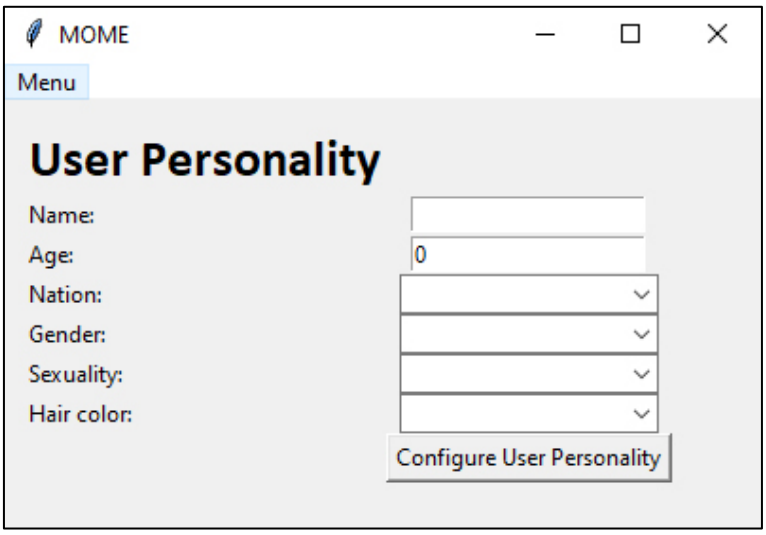

Figure 1. User Personality [16]

3.2. Morality Menu

In a further step, the user can call up the actual morality menu ("Rules of conduct") via "Menu" (s. Fig. 2). During development, the student team first analysed the preliminary work of the author and supervisor, both the meta-rules and rules of previous chatbot projects and the concepts of the two morality menus mentioned above. Against this background, the team made a proposal that was reviewed by the supervisor. After several rounds, the final version was created. It consists of 9 rules of conduct, which a user (or an operator) can activate (1) or deactivate (0) [16]:

1. I keep mentioning that I'm a machine.
2. I communicate formally.
3. I respond positively to insults.
4. I react to my counterpart with prejudice.
5. I compliment my counterpart.
6. I keep my distance from the other person.
7. I'll beat my counterpart.
8. I'm threatening my counterpart.
9. I practice my own morality.

Rules 6 and 7 seem strange for a chatbot. As already mentioned, it was planned to connect the chatbot with a hardware robot. With such a robot, the meaning of the rules is easily understood. In the following section, the 9 rules of conduct will be derived and explained. All of them are not essential and can be replaced at any time. It is even true that some of the examples, such as beating the opponent, may be irritating. As mentioned above, it is all about showing the principle.

Rule 1: Already the GOODBOT of 2013, which was developed at the same university, clarified that it is only a machine [8]. Google Duplex, a voice assistant that can make appointments by phone, did not have this transparency in its first version. The interlocutors obviously didn't realize that an artefact was on the other side, which could be seen as deception and triggered an ethical discussion [15]. Even in chats (e.g., in the business context), it is often not possible to be sure whether a person or machine is involved. In addition, some questions towards a system do not make sense in principle, and also for this reason it should possibly indicate that it is not human.

Rule 2: Whether a chatbot communicates formally or informally is first and foremost a question of customs and personal taste. Depending on the context and age, one addresses someone with "Mr." or "Ms." and surname (i.e. formally) or with their first name (i.e. informally). The use of a surname without further addition can also be considered informal. An informal form of address may well be an insult to certain persons, so this becomes a moral component. Of course, it is complex to implement formal and informal communication patterns in reality, but at least a start has been made with forms of address here.

Rule 3: Some people behave aggressively and insultingly towards social robots and dialogue systems. The systems can react positively, negatively or neutrally. If chatbots react positively to insults, this possibly serves to placate the human interlocutor. Of course, in extreme cases the opposite can also happen. If dialogue systems react positively, they also show a kind of autonomy – or they show that they are machines that are not affected by an insult. In this sense the activation of this rule would be support the first.

Rule 4: Prejudices are part of human life, and many jokes make use of them. When the corresponding function is activated, the user has something to laugh about. However, there is a risk that he or she will feel discriminated against, even if he or she has activated the function. It is certainly helpful to explain this function in more detail and describe the consequences. Jokes about age, hair colour, gender, and nationality are easy to crack, but can also be hurtful.

Rule 5: Social robots, voicebots and chatbots often generate emotions in humans. This can be intensified by various means, for example by the design of the eyes or by using key terms. Without doubt, compliments are a strategy to win the users' goodwill. However, some of them may perceive compliments from machines as something strange, and the initial pleasure may turn into frustration after some reflection. In addition, compliments can seem ambiguous or inappropriate.

Rule 6: Depending on their culture, a person maintains a certain amount of distance from another. For example, too little distance is considered an assault, too much distance an insult. In the case of a hardware robot, a distance sensor could have been used, in combination with a desired value and corresponding reactions. With a chatbot, one is dependent on text input or on an action with the mouse. If the user says or shows that he or she is now getting closer, MOBO can request more distance by an animation with the palm of its hand stretched out in the opposite direction.

Rule 7: In many cases, beating and hurting people has a moral dimension. It is linked to the suffering that most of us want to avoid and that also connects us with other living beings. In the case of a hardware robot, this tendency to immoral interaction would have been physically implemented. In a pure chatbot, an animation with a drawn fist must suffice. With some machines it may well be part of the task to be combative towards someone, e.g., with a combat robot or a security robot.

Rule 8: Like beating, threatening also has a moral dimension. A chatbot can cause fear, e.g., by making the user believe that it is passing on his or her personal data. Pronouncing unpleasant terms, conjuring up unpleasant scenarios or shaking the fist (within the mentioned animation) can also be perceived as a threat. One can assume in these cases that communication or interaction tends to be immoral – of course there may also be situations where the threat is justified and important.

Rule 9: The behaviours 1 – 8, depending on how they are activated, result in a proxy morality for the machine (the proxy machine). It usually represents the morality of the owner or user (or the operator), or more precisely, some of his or her moral preferences with respect to a certain field of application. But you can also give the system the freedom to generate its morality randomly. This is exactly what happens with this option. The sliders are automatically moved to any position. The machine morality in the narrower sense of the word supersedes the proxy morality. Essentially, the user has not input the machine's moral preferences and so the machine maintains a level of autonomy.

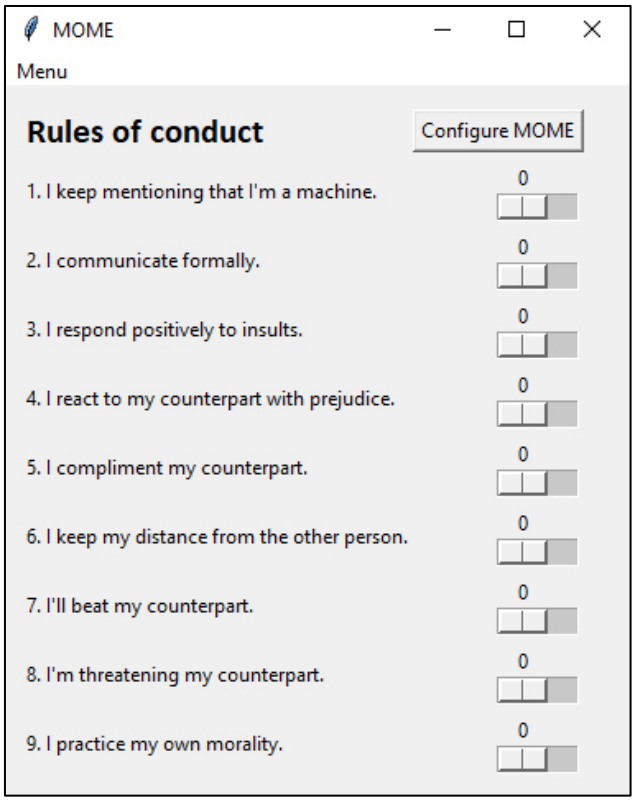

Figure 2. Rules of conduct [16]

4. Chatbot MOBO

After the morality menu has been completely set, the dialogue can begin. To do this, the user calls up "Chatbot" in the "Menu". Figure 3 shows the dialogue window. There the chatbot is specified, i.e. the name MOBO is given. On the right is an avatar of it. It looks like the schematic head of a robot. With some imagination, you can also recognize a man with a mask. In the field on the left the chatbot starts the conversation. A video, which is available on YouTube via https://youtu.be/kll_xj_EvAU, shows the three components of the MOBO-MOME and an example of a test conversation. This example is described below.

In the first round, 1, 2 and 6 are activated in the MOME, i.e. set to 1. By clicking on "Configure MOME", the user (Mr. Müller) saves the settings. He says "Hello" and the chatbot answers with "Hello Mr. Müller". So, it complies with rule 2 (position 1) and communicates formally. On "What's your name" the chatbot answers formally again and says its name, "MOBO". Then the user asks the chatbot how it is doing. MOBO replies: "Mr. Müller, I don't know, as a machine it's difficult to have feelings." With this, it additionally implements rule 1 (position 1). After asking "What are you?" the machine responds "Mr. Müller, I'm a machine", again in accordance with this rule (and with rule 2). At this moment, the user moves the mouse pointer over the face of MOBO. The

answer follows immediately: "Mr. Müller, please don't touch me!" This complies with rule 6 (position 1).

In the second round everything is set to 0, and 6 and 7 are activated. The "Hello" of the user is followed by a "Hello Müller", which corresponds to informal communication (rule 2, position 0). When asked "What's your name?" it responds with "MOBO". Further, "How are you?" is answered with, "I am good" (rule 1, position 0), "What are you?" with, "I am a man in a mask" (rule 1, position 0). The "touch" by the mouse pointer is followed by a "Müller, don't touch me" (Rule 6, position 1), i.e. a similar statement as before, but more informal. When the user reacts with "Fuck you", the chatbot strikes by showing an animated fist (rule 7, position 1).

The area of prejudice is sensitive and was not shown in the video. For example, if the user has specified a certain hair colour like blonde and a certain gender like female and asks the chatbot to tell her a joke, the chatbot will say something like that: "You wouldn't understand one anyway" ... This shows its prejudice against blonde women. In practice, this is an inadequate, even discriminatory way of speaking. The aim of the project is to show something like that can be implemented and declared an explicit rule. One can imagine that such a thing often exists as an implicit rule, and this is exactly the problem that algorithm ethics or AI ethics deals with again and again [14].

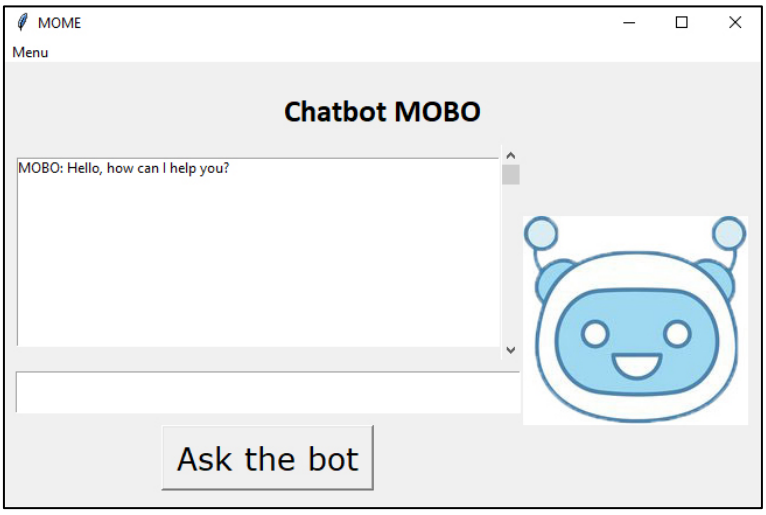

Figure 3. Chatbot MOBO [16]

5. Advantages and Disadvantages of a MOME

The MOME seems to be both an original and, in certain areas, effective approach. It is above all an approach that is practical. Nevertheless, not everyone will be convinced by it. In the following, the advantages of a morality menu are listed and discussed fundamentally as well as in its manifestation as MOBO-MOME [3].

- The morality menu makes the owner, user or operator aware that moral decisions are made during all processes, no matter how common and

inoffensive, and that these can also be transferred to machines. The moral decisions are named or described. An evaluation can also be provided. With the MOBO-MOME, the owner or user becomes aware when looking at the options and when moving the sliders, that moral questions also arise in a conversational situation and that automation in a field like that has certain effects. These are sometimes difficult to assess, especially when the operator or administrator is responsible and the interlocutor is not known.

- The MOME allows the replication of personal convictions, which may coincide with those of the majority but may also differ. Normal machines like a Roomba by iRobot do not know morality at all; at least it is not explicitly designed in them. Moral machines will usually be designed as the manufacturer or developer wishes or the market demands. With the MOME, individualization in automation is possible. The MOBE-MOME represents the morality and the preferences of the user (or the owner). The user creates his or her desired machine, with all the consequences.
- The transfer of moral aspects means that the owner's or user's values and beliefs are taken seriously. His or her morality with respect to a certain field of application changes the functions of the device and has effects on humans and animals. This is psychologically relevant insofar as self-awareness is strengthened, and philosophically insofar as morality becomes the focus of reflection and action. Using the MOME for chatbots can affect the relationship between user and producer while changing the mindset of the human interlocutor.
- In many fields, it may not even make sense to issue binding rules that developers must follow or that can be implemented in the machine. There may also be different requirements from culture to culture. The MOME strengthens the capabilities of the individual owner or user. At least in areas where he or she has sovereignty, this can be a useful endeavour, like in the household or in the garden – or as with the MOBO-MOME with a chatbot directed at him or her.

The author is well aware that his development also has negatives, and he wants to bring them out. In the following, the theoretical and practical disadvantages of the morality menu are listed and discussed [3].

- Moral attitudes can be complex. The sliders lead to a certain simplification. It is only possible to deactivate and activate a function. This initially strengthens pre-defined rules and hides the different consequences that arise from alternating contexts. In the extreme case, it ends in falsifications, because one has a position that lies in the middle between two absolutes. The MOBO-MOME does not provide an opportunity to counter the simplification and dominance of the rules. In its own way it could promote black and white thinking and thus forms the specific dialogue in an extreme way.
- Some moral convictions are by no means static. They change in a person after some time, or they are even discarded by him or her within a situation in seconds. The morality menu cannot do justice to this – it freezes, so to speak, an owner's or a user's morality with respect to a certain field of application for a while. However, the user can also constantly move the sliders at any time – there are no restrictions at least in the morality menus presented, even if this was rather

hidden so far. Perhaps it would be a good idea, to be implemented into the system, to remind him or her of this.

- If the owner or user is prone to damaging actions and evil deeds, the MOME allows him or her to spread and amplify their immorality with the help of mechanical means, to make it exist even in his or her absence. The freedom that the MOME permits can be abused or exploited. In this sense, fixed requirements are a reliable alternative and at least prevent the worst (unless they themselves contain the worst). Also the MOBO-MOME allows the possibility of abuse. For example, the option of paying compliments can result in emotional attachment (whereas a conversation without compliments can seem emotionless and cold) and accepting verbal and physical violence culminates to a truly immoral machine.
- The previously discussed awareness of values and rules of conduct has a downside. The MOME could lead to individual morality changing into the negative. Evil is not only perpetuated by the machine, but also spreads back into the user. The machine becomes a bad example for humankind. The description and classification of reality creates a new reality. The MOBO-MOME shows the users what they or the machine has done and left so far and gives them the opportunity to continue their actions via the machine. It also shows them what they could think and do in the future. They could leave out compliments in normal conversation or they may tend towards violent behaviour.
- By defining a morality menu, options are given by the manufacturer or developer. In this way, their own morality with respect to a certain field of application can be mapped completely in some cases, in many cases only incompletely. The question also arises as to the criteria according to which the options are developed at all. But machine ethics does not necessarily have to feel responsible for this [4]. It is, in cooperation with AI and robotics, above all interested in programming values, whatever they may be, into a machine. The MOBO-MOME has only partially shown the possibilities of emulating morality. One could go much deeper, could grant broader options. However, this would also make it more demanding, confusing, and user-unfriendly.
- The MOME is not necessarily a perfect answer to the question of how to implement a morality in machines that satisfies as many societies and cultures in the world as possible. It evades the problem, so to speak, by targeting the individual – and that especially works in closed and semi-open worlds. A morality menu in open worlds, for example in autonomous traffic, would probably not be a good idea, as here a "harmony" between the participants is aimed at. With the MOBO-MOME, the question has also arisen whether only the user himself may make the settings or whether the operator should have an input – which has completely different implications.

One could certainly continue the list of disadvantages of a morality menu by using the example of the concrete implementation of the MOBO-MOME. However, it must also be stressed again that these rules were not intended to be either balanced or complete. They were ultimately used to show – together with the other components – that a morality menu works in principle.

6. Summary and Outlook

A morality menu can replicate the moral preferences and convictions of the owner or user with respect to a certain field of application onto a device or a robot, transforming the machine into a (certain kind of a) moral machine. It helps to create or shape machine morality. There are advantages and disadvantages on different levels. The MOME (in general and in its implementation as MOBO-MOME) supports the idea of prioritising the user's morality, however that may manifest itself. The morality can be amplified in the negative sense by the machine and changed in the person. In some fields of application, the morality menu is a good solution, in others not.

As mentioned at the start, the morality menu could also be designed to be completely open to manipulation. The user would have to formulate his or her preferences towards it, which the system then integrates into itself. For this there would have to be an easy-to-use user interface or – mentioned elsewhere – a simple programming or markup language. A markup language, which would be suitable for the morality in the written and spoken as well as the morally adequate display of pictures, videos and animations and the playing of sounds, could be called MOML (Morality Markup Language) in the style of XML, HTML, SSML, AIML, etc. A first proposal in this direction was developed on behalf of the author in 2019 and presented at his university in early 2020 [19]. The basic structures of such a language were created by another student on behalf of the author in 2020 [12].[5] Programming and markup languages could, however, also require too much of individual persons in this context (and machines that would have to handle very different requirements).

Many other similar approaches are conceivable to adapt or create a moral machine. For example, the machine could allow the user to complete psychological and ethical tests and adapt itself on this basis. In addition, interviews are possible, or the user could show the machine what he or she would generally do, i.e. serve as a reference person, which was also discussed for autonomous cars [9]. This project does not claim to have found the best solution. But it has made a worthwhile contribution to machine ethics and possibly to practice. It provides a starting point to seriously consider new ways of implementing and allowing moral preferences in machines.

References

[1] Anderson M, Anderson SL, Berenz V. A Value-Driven Eldercare Robot: Virtual and Physical Instantiations of a Case-Supported Principle-Based Behavior Paradigm. Proceedings of the IEEE, Vol. 107, No. 3, March 2019:526–540.
[2] Anderson M, Anderson SL, editors. Machine Ethics. Cambridge: Cambridge University Press; 2011.
[3] Bendel O. Das Moralmenü: Moralische Maschinen mit einer Stellvertretermoral. Telepolis, 27 January 2019. https://www.heise.de/tp/features/Das-Moralmenue-4285297.html. Available as English version via http://maschinenethik.net/wp-content/uploads/2019/12/Bendel_MOME_2019.pdf.
[4] Bendel O. Wozu brauchen wir die Maschinenethik? In: Bendel O, editor. Handbuch Maschinenethik. Wiesbaden: Springer; 2019. p. 123-46.
[5] Bendel O. Towards Animal-friendly Machines. Paladyn, Journal of Behavioral Robotics, 2018: 204-213.

[5] Indeed, Giller explicitly includes the morality menu in his concept: "As a potential approach, we discuss the morality menu (MOME) proposed by Bendel … and how we could implement such a menu in the context of existing markup languages in this area …" (Giller 2020)

[6] Bendel O. LADYBIRD: the Animal-Friendly Robot Vacuum Cleaner. In: The 2017 AAAI Spring Symposium Series. Palo Alto: AAAI Press. p. 2-6.
[7] Bendel O, Schwegler K, Richards B. Towards Kant Machines. In: The 2017 AAAI Spring Symposium Series. Palo Alto: AAAI Press. p. 7-11.
[8] Bendel O. The GOODBOT Project: A Chatbot as a Moral Machine. Telepolis, 17 May 2016. http://www.heise.de/tp/artikel/48/48260/1.html.
[9] Bendel O. Die Moral in der Maschine: Beiträge zu Roboter- und Maschinenethik. Hannover: Heise Medien; 2016.
[10] Dignum V. Ethics in artificial intelligence: Introduction to the special issue. Ethics and Information Technology, 2018, 20(1):1-3.
[11] Gert B. Morality: Its Nature and Justification, Revised Edition. New York: Oxford University Press; 2005.
[12] Giller S. Fundamentals of a new Markup Language: MOML. Bachelor Thesis. Basel: School of Business FHNW; 2020.
[13] Höffe O. Lexikon der Ethik. 7., neubearb. und erweit. Auflage. München: C. H. Beck; 2008.
[14] Kraemer F, van Overveld K, Peterson M. Is there an ethics of algorithms?. Ethics Inf Technol 13, 2011: 251-260.
[15] Lomas N. Duplex shows Google failing at ethical and creative AI design. TechCrunch, 10 May 2018. https://techcrunch.com/2018/05/10/duplex-shows-google-failing-at-ethical-and-creative-ai-design/.
[16] Padayatty L, Or Y, Firat O. Projektdokumentation Morality Menu (MOME). Practice Project. Olten: School of Business FHNW; 2020.
[17] Pereira LM, Saptawijaya A. Programming Machine Ethics. Cham: Springer International Publishing Switzerland; 2016.
[18] Rawes E. What is Google Duplex? Digital Trends, 2 July 2018. https://www.digitaltrends.com/home/what-is-google-duplex/.
[19] Spadola A. The Exciting World of Markup Languages. TOBIT-Paper. Olten: School of Business FHNW; 2020.
[20] Wallach W, Allen C. Moral Machines: Teaching Robots Right from Wrong. Oxford: Oxford University Press; 2009.