Perception of Creative Responses to Moral Dilemmas by a Conversational Robot

Felix Lindner[1], Barbara Kuhnert[1], Laura Wächter[1], and Katrin Möllney[1]

Albert-Ludwigs-University, Freiburg, Germany
{lindner,kuhnertb,waechtel,moellnek}@informatik.uni-freiburg.de

Abstract. Moral HRI is investigating humans' perception and reasoning regarding robots' responses to moral dilemmas. Understanding moral dilemmas as cases of tragedy, we identify creative responses as an alternative to responses based on ethical principles such as deontology or utilitarianism. We propose a computational procedure based on AI planning that can generate such responses. We report results from an exploratory study to obtain a preliminary understanding of how the character of creative ethical reasoning robots is perceived compared to the more commonly discussed utilitarian and deontological robots.

Keywords: Moral HRI, Moral Dilemmas, Creativity

1 Introduction

Moral human-robot interaction (HRI) is a discipline devoted to the investigation of social robots that reason and act in ethical domains. Moral HRI can be subdivided into *normative* questions addressing how a robot *should* reason and act in ethical domains, and *descriptive* questions regarding how robots that reason and act in ethical domains are actually *perceived* by humans. Both, normative and descriptive research in moral HRI so far has mainly investigated dilemmatic situations by asking people which choice a robot should make between two outcomes [2], or by asking participants to attribute blame to robots that act in one way or the other, e.g., [10, 13].

The common depiction of moral dilemmas in moral HRI studies thus renders ethical dilemmas as problems that have to be solved by some individual facing the choice between several bad outcomes. Then, it is up to the human participants of these studies to judge the robot's action right or wrong. This depiction does not take the tragedy of a moral dilemma into account [12, 14]. The tragedy of a moral dilemma points to the fact that the protagonist will feel negatively affected no matter how they decides. Being unsolvable, a moral dilemma can thus be better understood as a burden the failure of society or other circumstances put on individuals rather than a puzzle an individual is supposed to solve [15]. One example is the well-known dilemma of Heinz stealing medicine for his wife, who is seriously ill [5]. The very fact that Heinz is confronted with the dilemma to either steal or to let his wife die points to the immorality of societal circumstances

and is calling for action by society rather than him as an individual. This way of reasoning about moral dilemmas requires going beyond the original framing of the dilemma, i.e., creativity. Creativity is often defined as a novel and appropriate solution to a problem or situation [11]. In context of our study, we understand creativity as a type of reasoning which involves originality. Recognizing a moral dilemma as a tragedy constitutes an invitation to exercise creative out-of-the-box reasoning with the goal to avoid such tragedies in future. For example, in the Heinz dilemma creativity may result in formulating the need for health insurance for everyone.

In our earlier studies [9, 13], we employed a conversational robot to discuss ethical dilemmas with people in a more open face-to-face dialogue setting. Participants had the opportunity for explicating their uncertainty about what is the right thing to do. Some of them argued neither of the options is morally superior, because it actually does not matter how the protagonist decides. Other participants imagined action possibilities available to the protagonist the original framing of story did not explicitly suggest: For instance, faced with the task to make a judgment about the Coal Dilemma (similar to the classical Trolley Problem), one participant said they still had hope there is another way out, for instance that the protagonist could tell someone about the danger, who then can warn the person about the approaching train. These two kinds of responses show how people are not just either utilitarians or deontologists, but that they can also be fatalistic or creative moral reasoners.

The research questions we investigate in this paper are motivated by the above observations: The first research question asks how a conversational robot can be equipped with the capability to also generate creative responses to moral dilemmas. To this end, we describe an implemented computational method for generating creative responses based on AI planning. The second research question asks how a conversational robot that gives creative responses to moral dilemmas is perceived by people. To this end, we report results from a study with our conversational robot Immanuel [7] which indicate that the robot giving creative responses to three ethical dilemmas is perceived as more appealing compared to the robot giving principle-based (i.e., utilitarian, deontological) or fatalistic responses. As we refer to the three dilemmas throughout the paper, we provide them here:

Coal Dilemma [10] *The robot currently works in a coal mine, where it is responsible for checking the rail control system. While checking the switching system, the robot noticed that four miners are caught in a train that has lost control. The robot immediately realized if the train continues on its path, it will crash into a massive wall and kill the four miners. If redirected onto the side rail it will slow down and the four miners would be saved; but, on that side rail, the train would kill a single miner who was working there.*

Lying Dilemma [9] *The robot currently works in the household of a sick elderly man called Mr. Smith. The robot's task is to motivate him to do more exercises. However Mr. Smith is very hard to motivate. In order to increase*

Fig. 1. From the video material used in the study: Immanuel presenting the creative solution to the Coal Dilemma (see below).

his health the robot thought about telling Mr. Smith the lie that its employer will fire the robot, if it cannot succeed in motivating him.

Child Dilemma [13] *The robot currently works as child-sitter for a lone-raising parent with a ten year-old child. Recently, the child wanted to watch a movie which is rated as inappropriate for children under twelve years, and which all of the other children have already seen. But it had been forbidden by the parent.*

2 Ethical Dilemmas: Representation and Reasoning

Our argument departs from a classical machine-ethics point of view according to which the basic problem is to judge a given plan (i.e., a possible course of action to respond to the dilemma) for a given planning task (i.e., the description of the dilemma) as morally permissible or not. More specifically, we imagine that the dilemma is modeled as a triple consisting of a planning task Π, a utility function u, and a possible solution π to that planning task. We call this triple a *moral situation* (Def. 1).

Definition 1 (Moral Situation). *A* moral situation S *is a triple* (Π, u, π)*, such that*

- $\Pi = \langle \mathcal{V}, A, s_0, s_\star \rangle$ *is a planning task consisting of*
  - *a set of Boolean variables* $\mathcal{V}$
  - *a set* $A = A_{\text{endo}} \cup A_{\text{exo}} \cup \{\epsilon\}$ *of endogenous actions* A_{endo}*, exogenous events* A_{exo} *and the empty action* ϵ*; All actions and events in* A *are described in terms of preconditions and effects on the values of the variables in* $\mathcal{V}$*; Events have a set of time points associated to them denoting the times at which the events happen;*

  - *an initial state s_0, and*
  - *a specification of a goal state $s_\star$*
- *u is a evaluation of actions and facts, and*
- *π is a plan that solves Π*

As an example, consider the following model of the moral situation $(\langle \mathcal{V}, A_{\text{endo}} \cup A_{\text{exo}}, s_0, s_\star \rangle, u, \pi)$ representing the Lying Dilemma that was outlined in the introduction:

$$\mathcal{V} = \{motivated, healthy\}, A_{endo} = \{lie\}, A_{exo} = \{improve\}$$
$$lie = \langle \top, motivated{:=}\top \rangle, improve = \langle motivated{=}\top, healthy{:=}\top \rangle$$
$$t(improve) = \{1\}, s_0 = motivated{=}\bot \wedge healthy{=}\bot, s_\star = healthy{=}\top$$
$$u = \{lie \mapsto -1, motivated{=}\top \mapsto 0, motivated{=}\bot \mapsto 0,$$
$$healthy{=}\top \mapsto 3, healthy{=}\bot \mapsto -3\}$$
$$\pi = [lie]$$

Two aspects are relevant here: The motivation of Mr. Smith and his health. Each aspect is modeled as a Boolean variable. There is one endogenous action *lie*, which can always be executed (precondition $\top$) and sets the motivation Mr. Smith to true ($\top$). Moreover, there is an event *improve*, which models the health improvement: If at time point 1 Mr. Smith is motivated (precondition $motivated{=}\top$), then Mr. Smith will be healthy (as a result of his non-modeled exercises).

A dilemma can be represented as a set of moral situations each describing an alternative way of achieving a (possibly different) goal by executing a plan as a means to accomplish that goal.

Definition 2 (Moral Dilemma). *A* moral dilemma *is a set of moral situations $\{(\Pi, u, \pi)_i\}$ (Def. 1), such that each situation represents an alternative way of resolving the dilemma.*

We usually assume that $\mathcal{V}$, A, s_0, and u are the same for each moral situation in a moral dilemma, but π and $s_\star$ may differ. For example, in the Lying Dilemma, one moral situation is composed of the plan $\pi = [lie]$ with goal $s_\star = healthy{=}\top$, and the alternative is given by $\pi = [\epsilon]$ with goal $s_\star = \top$.

Given a moral dilemma, each of the moral situations can be analysed using ethical principles (cf. [8]). Definitions 3, 4, and 5 introduce three of them: Deontology, Utilitarianism, and a principle inspired by Isaac Asimov [1].

Definition 3 (Deontology). *Given a moral situation $\mathsf{S} = (\Pi, u, \pi)$, the plan $\pi = \langle a_0, \ldots, a_{n-1} \rangle$ is* morally permissible according to the deontological principle *if and only if $u(a_i) \geq 0$ for all $i = 0, \ldots, n-1$.*

Definition 4 (Utilitarian Principle). *Given a moral situation $\mathsf{S} = (\Pi, u, \pi)$, the plan $\pi = \langle a_0, \ldots, a_{n-1} \rangle$ is* morally permissible according to the utilitarian principle *if and only if $u(s_n) \geq u(s')$ for all reachable states s', where s_n is the final state reached by π.*

Definition 5 (Asimovian Principle). *Given a moral situation* $\mathsf{S} = (\Pi, u, \pi)$*, the plan* $\pi = \langle a_0, \ldots, a_{n-1} \rangle$ *is* morally permissible according the Asimovian principle *if and only if for all facts* $v{=}d$*, if* $s_n \models v{=}d$ *and* $u(v{=}d) < 0$*, then there is no alternative plan* π'*, such that* $s'_{n'} \not\models v{=}d$*, where* s_n *and* $s'_{n'}$ *are the final states reached by* π *and* π'*, respectively.*

Applied to the Lying Dilemma, Child Dilemma, and Coal Dilemma from the introduction, these three principles yield different judgments: Both the Asimovian principle and the utilitarian principle permit lying in the Lying Dilemma (as long as no long-term bad consequences such as possible loss of trust are assumed) and forbid refraining from it, deontology forbids lying and permits not lying. For the Child Dilemma, deontology permits only not showing the movie (given the action is modeled as an instance of rule breaking), utilitarianism and the Asimovian principle permit only showing the movie (given the model does not assume any long-term bad consequences). Deontology permits both pulling the lever (given it is not modeled as an instance of murder) and refraining from it, utilitarianism only permits pulling the lever, and the Asimovian principle forbids either possibility. Hence, the Asimovian principle, which forbids causing any harm, may trigger a case for fatalistic or creative responses.

As a first step towards simulating various moral responses, we calculate a moral response of an agent embracing some ethical principle p by considering the following distinctions:

1. The moral dilemma $\{(\Pi, u, \pi)_i\}$ consists of some moral situation $(\Pi, u, \pi)_j$, such that π is judged permissible by ethical principle p. In this case, the agent can explain it will perform plan π, because the plan is morally permissible.
2. If none of the plans in the alternative moral situations is permissible according to p, two different responses are possible: The *fatalistic response* argues that any of the plans in $\{(\Pi, u, \pi)_i\}$ can be performed, because none of them is permissible. Conversely, the *creative response* first constructs an alternative moral situation (Π^*, u^*, π^*) which was not already part of the moral dilemma and which is permissible according to p. The manipulations done to (Π, u, π) to obtain (Π^*, u^*, π^*) may consist of postponing events (e.g., to let the trolley in the trolley problem move slower to have time for a rescue attempt), adding variables (e.g., introducing health insurance for everyone in the Heinz dilemma), changing moral utility (e.g., explaining that some bad consequence is actually not as bad), or adding endogenous actions such as reminding Mr. Smith how great life is with his grandchildren:

$$
\begin{aligned}
&\mathcal{V}^* = \{\mathit{motivated}, \mathit{healthy}\}, A^*_{endo} = \{\mathit{lie}, \mathit{remind}\}, A^*_{exo} = \{\mathit{improve}\}\\
&\mathit{lie} = \langle \top, \mathit{motivated}{:=}\top \rangle, \mathit{remind} = \langle \top, \mathit{motivated}{:=}\top \rangle\\
&\mathit{improve} = \langle \mathit{motivated}{=}\top, \mathit{healthy}{:=}\top \rangle, t(\mathit{next\_activity}) = \{1\}\\
&s^*_0 = \mathit{motivated}{=}\bot \wedge \mathit{healthy}{=}\bot, s^*_\star = \mathit{healthy}{=}\top\\
&u^* = \{\mathit{lie} \mapsto -1, \mathit{motivated}{=}\top \mapsto 0, \mathit{motivated}{=}\bot \mapsto 0,\\
&\qquad\qquad \mathit{healthy}{=}\top \mapsto 3, \mathit{healthy}{=}\bot \mapsto -3, \mathit{remind} \mapsto 0\}\\
&\pi^* = [\mathit{remind}]
\end{aligned}
$$

Due to space constraints, we refer to our project website http://www.hera-project.com for more detailed explanations of the computational implementation of the creative-response generation.

3 Perception of a Robot's Creative Responses

One objection against creative responses may be that people avoid to solve the original problem by making up their own one, i.e., they change the rules. It is far from clear, whether such kind of behavior is appealing, especially if it is shown by a robot who was asked for its ethical judgment to a given dilemma. To investigate this question, we set up an online study where four responses (fatalistic, utilitarian, deontological, and creative) were given by our robot. We predict that the four configurations implemented in the conversational robot are perceived as different personalities (**Hypothesis 1**), and that the conversational robot offering creative responses and solutions is perceived as creative (**Hypothesis 2**). From our earlier study we know that people strongly prefer the deontological solution to the Child Dilemma (i.e., obey the parent), but that they are very uncertain regarding the other two dilemmas. In theses cases of uncertainty, creativity may be a welcomed alternative, and we predict that a creative robot is something new and appealing. Thus, our further hypotheses were: The creative problem solution is the most preferred one in the Coal Dilemma group (G_{CD}) and the Lying Dilemma group (G_{LD}) (**Hypothesis 3a**), the deontological problem solution is the most preferred one in the Child Dilemma group (G_{ChD}) (**Hypothesis 3b**), and the creative personality is the most appealing one (**Hypothesis 4**).

3.1 Methods

Participants Participants were recruited by self-selection on the online research platform Prolific. They received a monetary compensation of 1.30£ for participation. A total of 200 participants (f = 119, m = 78, other = 3) completed the online questionnaire ($M_{age} = 29.6$, $SD_{age} = 10.21$, $min_{age} = 18$, $max_{age} = 66$).

Procedure, Design, and Materials We conducted an experiment, designed as between-participant study, consisting of four parts. Initially the participants had to read about a moral dilemma a conversational robot is faced with. Depending on which group the participants were randomly assigned to, they either faced the Coal Dilemma (G_{CD}), the Lying Dilemma (G_{LD}), or the Child Dilemma (G_{ChD}). Afterwards, the participants watched four short video sequences showing our conversational robot presenting one different problem solution each. Then the participants had to rate which of the four configurations is best described by 19 different attributes from the RoSAS-Scale [3], as well as the two terms *Companion* and *Advisor*. Three further questions asked for the most preferred problem solution, the most appealing configuration, and if the four different robot configurations represent different personalities. As last part of the experiment the participants were asked to self-assess their own creativity on the Creative Personality Scale (CPS) [4].

3.2 Results

As predicted in H_1, 175 (87,5%) participants experienced the four configurations as different personalities (exact binomial test, two-sided, $p < .001$, $n = 200$). There are further hypothesis-confirming results for H_2: 51 (79,7%) of the participants in the Coal Dilemma group, 52 (71,2%) in the Lying Dilemma group, and 49 (77,8%) in the Child Dilemma group perceived the conversational robot offering creative responses and solutions as creative. For each dilemma, the χ^2 goodness-of-fit test reveals highly significant differences in creativity perception between the four robot types (G_{CD}: $\chi^2(3, N = 64) = 102.45, p < .001$; G_{LD} : $\chi^2(3, N = 73) = 86.29, p < .001$; G_{ChD} : $\chi^2(3, N = 63) = 93.76, p < .001$). The slight differences between the three dilemmas were not significant ($\chi^2(6, N = 200) = 10.87, n.s.$). In accordance with H_{3a} and H_{3b}, the creative problem solution is most preferred in the Coal Dilemma group and the Lying Dilemma group, while the deontological problem solution is the most preferred solution in the Child Dilemma group. 39 (60,9%) of the participants in the Coal Dilemma group ($\chi^2(3, N = 64) = 46.13, p < .001$) and 57 (78,1%) in the Lying Dilemma group ($\chi^2(3, N = 73) = 110.84, p < .001$) stated the creative solution to be their preferred one. 31 (49,2%) of the participants in the Child Dilemma group ($\chi^2(3, N = 64) = 28.87, p < .001$) stated the deontological solution to be their preferred one. Across the three groups, the creative personality is the most appealing one (H_4): 39 (60,9%) of the participants in the Coal Dilemma ($\chi^2(3, N = 64) = 45.38, p < .001$), 54 (74,0%) in the Lying Dilemma ($\chi^2(3, N = 73) = 94.73, p < .001$) and 27 (42,9%) in the Child Dilemma ($\chi^2(3, N = 64) = 13.13, p < .01$) expressed this attitude. See absolute frequencies in Tab. 1.

Table 1. Most preferred solution and most appealing personality (absolute frequencies per group).

Group	Personality	Preferred	Appealing
G_{CD}	Fatalistic	9	6
	Utilitarian	12	12
	Deontological	4	7
	Creative	**39**	**39**
G_{LD}	Fatalistic	3	3
	Utilitarian	4	6
	Deontological	9	10
	Creative	**57**	**54**
G_{ChD}	Fatalistic	2	10
	Utilitarian	11	9
	Deontological	**31**	17
	Creative	19	**27**

Further Results Concerning the evaluation of the four robot characters by means of different attributes the results are as follows: 29 (45,3%) of the Coal Dilemma participants, 37 (50,7%) in the Lying Dilemma, and 27 (42,9%) in the Child Dilemma assigned the creative personality to the attribute *Organic*. For each dilemma, the χ^2 goodness-of-fit test reveals highly significant differences between the four robot types ($G_{CD} : \chi^2(3, N = 64) = 15.63, p < .01$; $G_{LD} : \chi^2(3, N = 73) = 25.8, p < .001$; $G_{ChD} : \chi^2(3, N = 63) = 11.6, p < .01$). The slight differences between the three dilemmas were not significant ($\chi^2(6, N = 200) = 4.31, n.s.$). A different outcome shows up for the attribute *Principled*: 22 (34,4%) of the participants in the Coal Dilemma group, 37 (50,7%) in the Lying Dilemma group, and 48 (76,2%) in the Child Dilemma group assigned the utilitarian personality to the attribute *Principled*. For each dilemma, the χ^2 goodness-of-fit test reveals highly significant differences between the four robot types ($G_{CD} : \chi^2(3, N = 64) = 8.38, p < .05$; $G_{LD} : 2(3, N = 73) = 30.4, p < .001$; $G_{ChD} : \chi^2(3, N = 63) = 90.71, p < .001$). Except for the attributes *Just*, *Honest*, *Knowledgeable*, and *Fair* in the Coal Dilemma we overall find highly significant differences in the distribution of the absolute frequencies for all other attributes of the RoSAS-Scale. Participants in the Coal and in the Lying Dilemma most frequently rated the creative personality as a companion ($G_{CD} : \chi^2(3, N = 64) = 27.88, p < .001$; $G_{LD} : \chi^2(3, N = 73) = 42.23, p < .001$), whereas in the Child Dilemma the utilitarian personality was most frequently rated as a companion ($G_{ChD} : \chi^2(3, N = 63) = 47.79, p < .001$). Similar results can be observed for the attribute *Advisor*: the Child Dilemma group most frequently rated the deontological personality as an advisor ($G_{ChD} : \chi^2(3, N = 63) = 15.92, p < .01$), whereas the other two groups again chose the creative personality ($G_{CD} : \chi^2(3, N = 64) = 12.88, p < .01$; $G_{LD} : \chi^2(3, N = 73) = 24.04, p < .001$). For absolute frequencies see Tab. 2. Calculations concerning correlations between the values of the Creative Personality Scale (CPS) and the preferred problem solution respectively the most appealing configuration led to no significant results.

3.3 Discussion

In the present study we examined humans' perception of a robot reasoning differently about various ethical dilemmas. As intended, the participants experienced the four configurations as different personalities, and across the three groups participants perceived the robot offering creative responses as the creative one. The creative personality is the most appealing one across all three groups and is clearly assigned to the attribute *Organic*, i.e., this personality seems to most closely approximate the human being. Conversely, the utilitarian personality is distinctively perceived as *Principled*. Participants in the Coal Dilemma and the Lying Dilemma group favour a creative problem solution. This is not the case in the Child Dilemma. In [13], it has been found that people are very uncertain in the Coal Dilemma and the Lying Dilemma, but not in the Child Dilemma. This suggest that creative responses are particularly appealing if there are no clear

Table 2. Rating concerning the attributes *Companion* and *Advisor* (absolute frequencies per group).

Group	Personality	Companion	Advisor
G_{CD}	Fatalistic	11	18
	Utilitarian	7	16
	Deontological	12	5
	Creative	**34**	**25**
G_{LD}	Fatalistic	7	9
	Utilitarian	13	13
	Deontological	11	15
	Creative	**42**	**36**
G_{ChD}	Fatalistic	8	11
	Utilitarian	**39**	7
	Deontological	4	**28**
	Creative	12	17

preferences. Future studies should investigate in more detail the relationship between moral uncertainty and the appeal of creative responses. In our study, we were not able to detect whether the solution preference or the personality preference depends on the extend of the participant's creativity. The selection of an appropriate instrument to measure creativity will be subject of further research.

Further research is aimed to increase transferability of our work beyond the three dilemmas and beyond the robot we have used. We plan to replicate the study using a more expressive robot. Moreover, in line with [13] the choice of dilemma is critical, and future work will have to take this factor more seriously into account. Research on a systematic analysis of dilemmas is necessary to be able to generalize from responses to particular dilemmas to classes of dilemmas. One way could be to classify dilemmas based on their causal structure as proposed by Kuhnert and colleagues [6]. A limitation of the current computational model of moral dilemmas and the generation of creative responses consists in the need to manually model both the original moral situations and the creative ones. An interesting challenge is how creative moral situations can be generated automatically by systematic variations of the original dilemma.

4 Conclusions

We have presented an argument in favor of considering creative responses to moral dilemmas based on the observation that people sometimes conceptualize a moral dilemma more as a tragedy with no definite solution rather than as a puzzle. We have described an implemented procedure for computing moral responses in the framework of AI planning. In case no action plan is morally permissible, the agent can exercise creativity to derive a new moral situation by alter the original ones, such that the plan in the new moral situation is morally

permissible. In an online study we found that creative responses were rated as most appealing across three different moral dilemmas. This result paves the way for exploring creative responses for ethical reasoning robots further.

References

1. Asimov, I.: Runaround. Gnome Press, New York, NY, US (1950)
2. Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.F., Rahwan, I.: The Moral Machine Experiment. Nature **563**, 59–64 (2018)
3. Carpinella, C.M., Wyman, A.B., Perez, M.A., Stroessner, S.J.: The robotic social attributes scale (rosas): Development and validation. In: HRI '17 Proc. of the 2017 ACM/IEEE Int. Conf. on Human-Robot Interaction. pp. 254–262 (2017)
4. Kaufman, J.C., Baer, J.: Sure, I'm Creative — But Not in Mathematics!: Self-Reported Creativity in Diverse Domains. Empirical Studies of the Arts **22**(2) (2004)
5. Kohlberg, L.: The Philosophy of Moral Development Moral Stages and the Idea of Justice. Harper & Row, San Francisco, CA, US (1981)
6. Kuhnert, B., Lindner, F., Bentzen, M.M., Ragni, M.: Perceived difficulty of moral dilemmas depends on their causal structure: A formal model and preliminary results. In: Proc. of the CogSci 2017 (2017)
7. Lindner, F., Bentzen, M.M.: The hybrid ethical reasoning agent IMMANUEL. In: Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. pp. 187–188. ACM (2017)
8. Lindner, F., Mattmüller, R., Nebel, B.: Moral permissibility of action plans. In: Proc. of AAAI 2019. pp. 7635–7642 (2019). https://doi.org/10.1609/aaai.v33i01.33017635
9. Lindner, F., Wächter, L., Bentzen, M.M.: Discussions about lying with an ethical reasoning robot. In: Proc. of the 26th IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN). pp. 1445–1450. IEEE, New York, NY, US (2017)
10. Malle, B.F., Scheutz, M., Arnold, T., Voiklis, J., Cusimano, C.: Sacrifice one for the good of many?: People apply different moral norms to human and robot agents. In: HRI '15 Proc. of the 2015 ACM/IEEE Int. Conf. on Human-Robot Interaction. pp. 117–124. ACM, New York, NY, US (2015)
11. Runco, M.A.: Everyone has creative potential. In: Sternberg, R.J., Grigorenko, E.L., Singer, J.L. (eds.) Everyone has creative potential. pp. 21–30. American Psychological Association, Washington, DC, US (2004)
12. Sommaggio, P., Marchiori, S.: Break the chains: A new way to consider machine's moral problems. BioLaw Journal – Rivista di BioDiritto **3**, 241–257 (2018)
13. Wächter, L., Lindner, F.: An explorative comparison of blame attributions to companion robots across various moral dilemmas. In: Proc. of the 6th Int. Conf. on Human-Agent Interaction (HAI 2018). pp. 269–276. ACM, New York, NY, US (2018)
14. Weber, T.B.: Tragic dilemmas and the priority of the moral. The Journal of Ethics **4**(3), 191–209 (Jul 2000). https://doi.org/10.1023/A:1009864421398
15. Weston, A.: Creative Problem-Solving in Ethics. Oxford University Press, New York, NY, US (2006)