

Robot rights? Towards a social-relational justification of moral consideration

Mark Coeckelbergh

Published online: 27 June 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Should we grant rights to artificially intelligent robots? Most current and near-future robots do not meet the hard criteria set by deontological and utilitarian theory. Virtue ethics can avoid this problem with its indirect approach. However, both direct and indirect arguments for moral consideration rest on ontological features of entities, an approach which incurs several problems. In response to these difficulties, this paper taps into a different conceptual resource in order to be able to grant some degree of moral consideration to some intelligent social robots: it sketches a novel argument for moral consideration based on social relations. It is shown that to further develop this argument we need to revise our existing ontological and social-political frameworks. It is suggested that we need a social ecology, which may be developed by engaging with Western ecology and Eastern worldviews. Although this relational turn raises many difficult issues and requires more work, this paper provides a rough outline of an alternative approach to moral consideration that can assist us in shaping our relations to intelligent robots and, by extension, to all artificial and biological entities that appear to us as more than instruments for our human purposes.

Keywords Moral consideration · Artificially intelligent robots · Ethics of robotics · Rights · Relations · Social ecology

Introduction

Current robots are not only used in industrial contexts; they find applications in many new domains, including the personal and social sphere. We now have entertainment robots, toy robots, nursing robots, sex robots, and educational robots, amongst others, and the prevalence of such robots is expected to increase strongly over the next few decades (Veruggio 2006). It is likely that in the near future many of these robots will be artificially intelligent and have the ability to interact with humans in a human-like way. For instance, it is projected that we will live with ‘social robots’ (Breazeal 2003) or ‘artificial companions’ that may act as artificial pets, provide information services, take care of home security, assist health care, or perform household tasks (Floridi 2008; Dautenhahn et al. 2005). Moreover, several artificially intelligent robots are used and being developed for military use, many of which have humanoid characteristics (see for instance DARPA humanoid robot projects in the US).

These developments raise many ethical issues. Usually these concern the implications of intelligent robots for humans (Sharkey 2008). Robots are seen as (potential) moral agents, which may harm humans and therefore need a ‘morality’. It is argued that robots should follow laws that prevent them from harming humans, as the famous ‘Laws of Robotics’ testify (Asimov 1942), or that we should build ‘moral machines’ that do not only follow ‘top-down’ laws given to them by humans but that also have the capacity to develop into moral machines ‘bottom up’, that is, by moral learning emerging from their intelligence (Wallach & Allen 2008). However, sometimes robots come into focus as ‘moral patients’. How should we treat *them*? Given that some robots will become more human-like, should (some) robots be protected from abuse by humans? For instance,

M. Coeckelbergh (✉)
Department of Philosophy, University of Twente, P.O. Box 217,
7500 AE Enschede, The Netherlands
e-mail: m.coeckelbergh@utwente.nl

Whitby has called for urgent action on the ethics of mistreatment of human-like artefacts (Whitby 2008). The ethical question I ask here is more general: should we take artificially intelligent robots into moral consideration at all?

Often this issue is framed in terms of rights. In 2000 Rodney Brooks, director of the Artificial Intelligence Lab at M.I.T., wrote a *Time* article entitled ‘Will Robots Rise Up and Demand Their Rights?’ in which he says that robots will become more human-like, which ‘will eventually lead to robots to which we will want to extend the same inalienable rights that humans enjoy’ (Brooks 2000). Recently there is a proliferation of calls for robot rights. ‘A report commissioned for the UK government, containing a paper on ‘Robo-rights: Utopian Dream or Rise of the Machines?’ (2006), suggested that robots may demand rights in the future. Peter Asaro thinks along the same lines. Identifying robots rights as one of the issues robot ethics should be concerned with, he asks ‘How should robots treat people, and how should people treat robots? Should robots have rights?’ (Asaro 2006, p. 10). He speculates: ‘At some point in the future, robots might simply demand their rights.’ (Asaro 2006, p. 12). In a recent conference paper David Levy has argued that artificially conscious robots should have rights (Levy 2009). And Steve Torrance, after noting that the *UN Universal Declaration of Human Rights* does not cater for the rights of artificial agents or humanoids (Torrance 2008, p. 496), asks if there are ‘any circumstances under which it might be morally appropriate for us to consider extending such rights specifically to humanoids that are taken to be devoid of phenomenal consciousness?’ (Torrance 2008, p. 501).

However, how adequate is this ‘rights’ approach for the issue at hand? Robots with consciousness or the ability to ‘demand their rights’ seem to belong to the realm of science-fiction or at least the far future. Does this mean that current and near-future artificially intelligent robots should be excluded from our moral world entirely? Are there perhaps other ways of granting them moral consideration? In this paper, I argue that the rights approach is not flawed per se but that we should expand the range of arguments for moral consideration of current and near-future artificially intelligent robots in at least three ways. In the course of my arguments I discuss analogies with the ‘animal rights’ discussion and emancipation claims.

First I will draw attention to the observation that moral consideration comes in degrees, that giving ‘rights’ to an entity is a particularly strong form of moral consideration, and that there are other forms of moral consideration that may be more appropriate and relevant to the kind of artificially intelligent robots we may use and live with in the near future. Then I will note that apart from rights, a concept which is directly related to deontology, there are other theoretical frameworks which we might want to use

to argue for giving robots moral consideration: utilitarianism and virtue ethics. However, I will show that all three kinds of arguments (deontological, utilitarian, and virtue) rely on ontological, non-relational features of the robot (and the human) and that this approach incurs several problems. In response to these problems, I explore alternative arguments for moral consideration. First I discuss arguments that—like virtue ethics—are based on indirect moral status (based on human personhood, human property rights, and human values). However, they turn out to be vulnerable to the previous objections and do not seem to be very ‘moral’ given that their justification is exclusively rooted in the subject of moral consideration. Then I offer a novel argument that employs a social-relational justification of moral consideration. This approach takes seriously the moral-social significance of appearance and is based on an ecological social ontology instead of a contractarian, utilitarian, or communitarian one. I suggest that in order to further develop such an ontology, we may want to take inspiration from environmental ethics and Eastern or Asian philosophy. I then discuss some implications of this approach for moral consideration of robots, animals, and humans.

Standard direct and indirect arguments for moral consideration

Direct arguments for moral consideration

Giving rights to an entity is a particularly strong form of moral consideration: it implies that the entity in question has inherent worth and that therefore the entity needs to be treated as such irrespective of all other (human or non-human) considerations. It is connected to the deontological and natural rights traditions in moral theory, which in turn can be interpreted as secular continuations of the Jewish-Christian idea that the person has an inherent worth as a creature of God.¹ Hence, rights are given a kind of quasi-sacred status: any violation of them counts as a secular version of sin. In the Kantian tradition, the spiritual nature of man is replaced² by the rational nature of persons. According to Kant, anyone who violates rights does not

¹ It is believed that humans are created in the image of God (the so-called *imago Dei* doctrine, based on *Genesis*).

² Perhaps ‘replaced’ is not exactly right: as said, there was also continuation. There is an influential (Aristotelian) current of thought in the Christian tradition that stresses the rational nature of the person. Thomas Aquinas viewed the person in terms of rational nature. He did not oppose rationality to spirituality. He followed Boethius, who in his *Liber de Persona et Duabus Naturis* defined the person as an individual substance of a rational nature. Boethius was a translator of Aristotle, who held that man was a rational animal.

take into consideration that, ‘as rational beings’, others should at the same time be treated as ends (Kant 1785, p. 92).³ In such Enlightenment arguments, direct references to religion or spirituality are absent. However, in modern deontological ethics the idea is retained that rights should therefore be respected at (literally) all costs. For example, human rights are seen as *inalienable* rights that always should be respected and that ground other goods. The first lines of the Preamble of the *Universal Declaration of Human Rights*, adopted by the UN in 1948, holds that ‘recognition of the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world’. To make a claim for robot rights in the deontological tradition, then, is to make a claim for strong moral consideration. But there are many more possible forms of moral consideration and other kinds of arguments. As Gruen remarks: ‘Being morally considerable is like showing up on a moral radar screen—how strong the signal is or where it is located on the screen are separate questions’ (Gruen 2003).

In order to explore alternative forms of moral consideration of robots and different justifications, let me make an analogy to arguments for moral consideration of animals. Regan, from a deontological position, has argued that some animals have rights since they are experiencing ‘subjects of a life’: they have wants, preferences, beliefs, feelings, memories, and expectations, and their welfare matters to them (Regan 1983). However, this is not the only kind of position in the debate. Utilitarians adopt a different approach. They reject deontological rights and if they use the language of rights they understand them as derived from utilitarian principles. For them, what counts is not the protection of inalienable rights but the promotion of happiness, pleasure or interests and the avoidance of suffering or frustration of interests. Singer has argued that any being—human or non-human—that is sentient and therefore has an interest in not suffering deserves to have that interest taken into account (Singer 1975, 1993). This implies that utilitarians can give moral consideration to humans and non-humans alike, provided that they are sentient. In practice, this means that the boundary of moral consideration is wider than in the rights approach, since there may be many animals that cannot be considered subjects-of-a-life but that are sentient.

In spite of the many differences between these positions, which I will not discuss here, there are many similarities as

well. Both arguments demand moral consistency: it is argued that the human/non-human distinction is not morally relevant when it comes to treatment of animals. The reasoning may go as follows. In human history, animals have always been enslaved and mistreated because they were (and are) considered to be non-human. However, why should we restrict rights to humans? It is speciecist to do so. ‘Speciecism’ is a term Singer ascribed to his opponents and refers to the view that a certain species, here humans, is superior to other species (Singer 1975). But why should the species barrier be a moral barrier? Singer has argued that we should take sentience (and therefore an interest in not suffering) as a criterion, not being a member of a certain species—be it the human species or another one. Similarly, Regan has questioned the human/non-human distinction as a moral distinction for ascribing rights (Regan 1983). Thus, both Singer and Regan appeal to the demand for consistency: they share the view that if we take a certain characteristic to be grounds for ascribing rights in humans, we should be consistent and do the same for non-humans that have the same characteristic in common, e.g. sentience (Singer) or being the subject of your life (Regan).

For robots, one can make a similar demand for consistency coupled with an *emancipatory* claim that can also be found in the animal rights movement broadly understood (based on deontological and utilitarian arguments): if (in the future) it turns out that robots share features with humans such as rationality or consciousness, then if we hold these features as a basis for human rights, why restrict those rights to humans? If they might one day become sentient, then why neglect their interests in avoidance of suffering? Why continue to treat artificially intelligent robots as things we can use or abuse if we have good reasons to include them in our community of moral consideration and rights? We have emancipated slaves, women, and some animals. First slaves and women were not treated as ‘men’. However, we made moral progress and now we consider them as human. In the past animals were treated as things, but we learned that many of them can feel and think. Now, or in the near future, it is time to give certain robots what is due to them on account of their intelligence, consciousness, or whatever feature we humans share with them we believe is a basis for ascribing rights or for taking their interest into account.

In the ethics of robotics literature, the emancipatory mission that motivates the call for robot rights is not always present or made explicit; however, the analogy to animal rights can always be drawn and is sometimes made explicit. Calverley, for instance, has explored the analogy between the debate on rights for androids and animal rights. Discussing the similarities between animals and androids, he has pointed to the issue of the relation between our ideas of consciousness and our concept of rights

³ Unfolding his ‘Formulate of the End in Itself’ in the *Groundwork*, Kant argued that in contrast to non-rational beings, who have only ‘relative value as means and are consequently called things’, rational beings are persons ‘because their nature already marks them out as ends in themselves—that is, as something which ought not to be used merely as a means’ (Kant 1785, pp. 90–91).

(Calverley 2006). And McNally and Inayatullah (1998) have put the discussion about robot rights within a broader historical and cultural framework. But regardless of the particular perspective of the authors, the arguments always draw on ontological features of the entity that are seen as morally significant.

This reveals a deeper argumental similarity between the rights and utilitarian position in the discussions about robots and about animals. Both arguments for consistency (and emancipation) rely on holding a particular ontological feature of the entity in question to be sufficient for granting moral consideration to that entity: rationality, consciousness, sentience, and so on.

This approach is problematic for two kinds of reasons. Apart from well-known problems of *application* (for instance, what exactly counts as respecting an entity's rights or capacity for suffering?), deontological and utilitarian accounts face problems with regard to their *justification* of moral consideration. Since their approach to moral consideration is based on ontological features of the entity, they incur at least the following problems.

High thresholds and relevance

Before outlining problems of justification, let me first mention an important problem of application. Today robots are neither conscious nor sentient. It is even questionable if any of them really are (artificially) intelligent. This renders arguments based upon such features irrelevant to the problem of how to think about giving moral consideration to currently existing intelligent robots. Similar problems with 'too high' thresholds happen in environmental ethics: a tree is neither conscious nor sentient, yet there might be other reasons why we want to grant it some moral consideration.⁴ One option is to lower the threshold, as Floridi and Sanders did when they proposed the features of interactivity, autonomy and adaptability (Floridi and Sanders 2004). However, I doubt if this does enough justice to the new *social* functions of these robots and how they appear to us (I will return to this issue below). Moreover, very young human infants are neither very interactive nor highly autonomous, yet intuitively we wish to give a high moral status to them. This limitation takes us to the next issue.

⁴ For instance, according to deep ecology, trees have inherent value as part of the ecosystem. Leopold would have considered them as parts of what he called the 'biotic community' (Leopold 1949). But one might also use an indirect argument for moral consideration: trees should be granted moral consideration to the extent that they contribute to human well-being. I will discuss indirect arguments below.

Argument from marginal cases

Arguments based on ontology are often vulnerable to what in animal ethics is known as the argument from marginal cases⁵: if particular properties are agreed upon as being sufficient for moral status and if not all humans share these properties (all the time), does that imply that these humans are not worthy of our moral concern (at the time)? Small children or aged adults do not always meet the criteria. And does our moral status disappear when we sleep? Deontological and utilitarian theories each try to deal with this problem in different ways, but it is not always easy to uphold the demand for consistency given their adherence to the moral significance of ontological properties of an entity.

Determination problems and moral epistemology

Provided that the previous difficulties can be overcome (and they partly have been overcome), then there are the following epistemological problems: it is difficult to agree which ontological property is morally relevant and to provide proof that the entity in question has that property. Consider robots again. Why should we agree with Levy, who proposes consciousness as a criterion? Why should we agree with Floridi and Sanders, who propose different criteria? Our moral intuitions differ on what criteria are the relevant ones. Moreover, even if we could agree on that, then can we provide 'hard' evidence for the presence of these criteria in a particular robot, given that the criteria are rather abstract? The debate on the moral status of animals faces a similar difficulty: the two main influential figures in that debate argue that moral status should be granted on the basis of an animal being a 'subject-of-a-life' (Regan 1983) or being sentient (Singer 1975). However, how do we know that these are the relevant features? And, faced with a particular entity, how can one provide proof of, say, 'being the subject-of-a-life'? What is the moral-epistemic basis for making such claims? (And why is it permitted to go from 'is' (ontology) to 'ought' (moral consideration) anyway? I will return to this problem below.)

Individuals and society

The rights approach focuses on the individual rights and the utilitarian approach on (the sum of) individual interests. In doing so, they make assumptions about the relation between individual entities and the wholes these entities are part of (systems, societies, communities) and these assumptions are problematic. By focussing on individual

⁵ For an overview of the discussion about the argument see for example Dombrowski 1997.

features they tend to neglect the moral relevance of relations between entities and of the wholes they are part of. For instance, by reducing ethics of robots to the debate about moral *status* of robots and humans (as based in their ontology), it leaves out of sight how (changes in) relations and social wholes at least contribute to (changes in) moral consideration. I will return to this point below.

Indirect arguments for moral consideration

One may object now that I have unnecessarily limited my discussion to deontological and utilitarian theory. Why not consider virtue ethics? At first sight, virtue ethics seems to avoid the problems mentioned above since it employs an indirect argument for moral consideration: if we (humans) wish to be virtuous persons, we should treat animals well. Abusing them is wrong not because it is a violation of rights or because on balance more suffering is created than with another act, but because we, as members of a moral community, do not exercise virtues such as compassion when abusing them. For instance, Hacker-Wright (2007) has argued that our answer to the question concerning moral consideration of animals depends on our conception of the demands of the virtue of justice.

By shifting the focus from the object of moral consideration to the subject of moral consideration, virtue ethics seems to avoid the problems mentioned above. The threshold is set by the virtues, there are no marginal cases within the (human) moral community where everyone is equal, we no longer need to know particular features of the entity and their relation to moral consideration, and surely the importance of ‘community’ is accounted for. However, this impression is false given the problem of *application* this theory faces. Not only is it unclear how we can know what the virtues are (which is a problem of justification); from the general principle, it is not clear at all (1) in relation to what entities we should exercise our (human) virtues and (2) what the application of that virtue consists in. Deontological and utilitarian approaches to moral consideration share these problems, except that they provide (ontological) criteria to respond to the first problem (which entities) whereas virtue ethics does not. Therefore, it remains unclear how it can avoid ontological justification and its problems. Moreover, it is very likely that, when pressed for it, virtue ethicists will provide that type of justification since their ethics is based on the moral status of humans and to defend *that* side of the argument they assume a similar ontological foundation of its human-centred ethics. In its neo-Aristotelian version, virtue ethics roots its ethics in a particular view of human nature: man is a rational animal and a ‘political animal’, that is, a being which only develops itself, and can only flourish, as part of a moral and social community. For example, Hursthouse’s

(1999, p. 68) answer to the problem of justification is that virtue is ‘a character trait that a human being, given her (human) nature, biological and psychological, needs (...) to flourish or live well’.⁶ MacIntyre (1999) defines human nature in terms of (social) dependence: we can only flourish through developing social relationships and communities. Such neo-Aristotelian views are far more relational than the social philosophies that are home to rights and utilitarianism, but in so far as they remain Aristotelian they by definition restrict themselves to humans. Humans, for Aristotle the only rational and political beings, are the top of the pyramid of beings. Thus, in order to provide a foundation for its claims about moral consideration, a virtue ethicist will have to provide an ontology and will provide one which seems to exclude all non-humans from moral consideration. The only way to avoid this, it seems, is to put forward ontological criteria that are ‘lower’ and allow us to include some animals, robots, and other entities. But this solution leaves the theoretical boundaries of neo-Aristotelianism and (still) incurs the problems mentioned above—problems related to the reliance on ontological features.

One may object, however, that although there are problems with virtue ethics’ indirect argument, virtue ethics could say that it has a straightforward *consequentialist* reason why this focus on the human is acceptable: if all humans were to exercise their virtue of compassion (a virtue *justified* by reference to *humans*), then as a consequence certain animals would fare better than they do now. This argument could be applied to highly intelligent robots as well. I concede this; nevertheless I wonder how broad and systematic protection of non-humans would be. For me to develop myself into a morally high-standing being that lives the flourishing life, the best life there is for beings like us, it might be enough that I develop and exercise my compassion in relation to one animal or one robot, for example my pet animal or pet robot. If everyone does that, many animals will be treated well as a consequence. However, something seems to be missing here that renders it less ‘moral’: it goes against the intuition that the motivation for and justification of moral consideration should not have its source in our own well-being or our own moral status alone (the subject or giver of moral consideration) but at least also in the well-being or status of the *object* or receiver of moral consideration—an intuition shared with deontological and utilitarian theories of moral consideration.

Other indirect arguments for moral consideration face the same limitations. For instance, one can base moral consideration of animals and robots on the property rights *of humans*, or, more broadly, on the value *we humans* give

⁶ Hursthouse then goes onto develop a Humean interpretation, which I shall not discuss here.

to things. Indeed, we already give *some* moral consideration to many objects, including robots, not because we believe that they have rights or that they are sentient but because we own them or because we value them in other ways. For instance, one may protect a tree in one's garden because one loves the view, because it contributes to the aesthetic value of the garden. Or we may protect very intelligent robots because we value them as objects to perform experiments with. The range of human values is wide and many non-humans benefit from the related practices of moral consideration. However, these arguments face similar problems as the approaches discussed above. First, in order to determine the value of a non-human it seems that we have to know its ontological status. Second, it remains doubtful how 'moral' this moral consideration is given that its justificatory source is entirely non-relational and non-social: it resides in the subject of moral consideration. Is there an alternative way of thinking about the issue of moral consideration?

An alternative, social-relational argument for moral consideration

Towards a social-relational approach: appearance and social relations

So far I discussed two sorts of arguments for moral consideration. Direct arguments focus on the moral status of the object, whereas indirect arguments centre on the moral status of the subject. Both accounts of moral status are based on the ontological features of the entity, an approach which invites epistemological scepticism and incurs other problems. The alternative approach I propose attempts to avoid the scepticism by replacing the requirement that we have certain knowledge about *real* ontological features of the entity by the requirement that we experience the features of the entity as they *appear* to us in the context of the concrete human-robot relation and the wider social structures in which that relation is embedded. This requires further explanation.

First, moral consideration is no longer seen as being 'intrinsic' to the entity: instead it is seen as something that is 'extrinsic': it is attributed to entities within social relations and within a social context.

Second, this does not imply that features of the entity are morally insignificant. We will continue to use them as criteria on which we base our moral consideration. However, in this approach they are given a different status: they are *apparent* features,⁷ features-as-experienced-by-us.

⁷ See also Coeckelbergh 2009a, b, 2010.

Third, this experience is not context-independent and not subject-independent. It is *context-dependent*: in line with feminist objections to standard moral theories,⁸ this approach acknowledges the limitations of argumentation that aims at general and abstract moral principles alone. Instead, it asks more attention for the ways in which moral consideration is granted to entities in various concrete social relations and social contexts. Moreover, the experience is *subject-dependent*: in its response to the epistemological problem the approach learns from the phenomenological tradition in philosophy, which has proposed an interesting answer to the (false) dilemma between idealism and realism. As far as I understand the basic message of Husserl, Heidegger, Sartre, and others, they hold that consciousness is always directed towards objects (for which phenomenologists use the technical term 'intentionality'⁹) and that we can only have knowledge of objects as they appear to us. Applied to moral consideration, it means that moral significance resides neither in the object nor in the subject, but in the relation between the two. Objects such as robots do not exist in the human mind alone (this would amount to idealism); however, it is also true that we can only have knowledge of the object and its features as they appear in our consciousness. There is no direct, unmediated access to the robot as an objective, observer-independent reality or 'thing-in-itself'.¹⁰

The next step I then propose is to see this subject-object relation as being shaped in social relations. What happens between subject and object takes on a form in the concrete social context, which is 'prior' to the moral argumentation offered by traditional theories. Deontological and utilitarian reasoning cannot be divorced from the social-historical context in which these theories emerged: modern society with its emphasis on the value of individuals and their preferences. This does not mean that these theories get it wrong or that we have no choice but to accept whatever theory is prevalent in a social context. Instead, recognising this link between theory and practice opens up an additional repertoire of moral vocabulary that helps us to distance ourselves from our usual ways of thinking and doing. In this case, it helps us to better understand the issue of moral consideration of robots by making comparisons with how we treat other non-human entities at different times

⁸ Consider Gilligan 1982, Noddings 1984, Ruddick 1989, and Kittay 1999. By 'standard moral theories' these writers usually mean deontology and utilitarianism.

⁹ For example, Sartre (1943, p. xxxvii) interprets intentionality in terms of consciousness of an object. In Sartre's words: 'Consciousness is consciousness of something', it must 'produce itself as a revealed-revelation of a being that is not it' (p. xxxviii).

¹⁰ Kant argued that we cannot know the thing-in-itself. The phenomenological tradition and its concept of intentionality must be understood as a response to Kant, i.e. a further development of Kant.

and different contexts. In contrast to what the term ‘moral status’ suggests, moral consideration must be seen as subject to change. There are also differences within societies and between societies and cultures. Of course we can and must critically evaluate these differences and changes. However, if we do not go beyond the three traditional theories summarized above, then we do not pay sufficient attention to differences and changes since we occupy ourselves with general criteria alone.

On a sub-societal level the approach focuses on moral considerations in human-robot relations rather than on the moral status of humans and robots alone. The implication is that both the human and the robot are not so much considered as atomistic individuals or members of a ‘species’, but as relational entities whose identity depends on their relations with other entities. The idea is that if we live with artificially intelligent robots, we do not remain the same individuals and the same humans as we were before. Thus, the relational theory of moral consideration proposed here must be connected with a relational theory of identity and, in the end, a relational (social) ontology. Let me explain this in the next sub-section.

The relation between individuals and society: social ecology beyond contractarianism, utilitarianism and communitarianism

The alternative approach proposed above may bracket individual ontology by turning to appearance, it still needs a *social* ontology given its focus on social relations and their social context. A fully developed account would require much more work; however, let me make the following suggestions about where I would like to position my approach in relation to some influential social philosophies.

In social philosophy, the rights approach is connected with a contractarian social ontology. Individuals are prior to the social, which comes only into being by agreement.¹¹ Utilitarianism has a similar instrumentalist view of society: what counts is the happiness, pleasure, preferences, and interests of the *individual*. Society must safeguard and increase the total amount of happiness, pleasure, or other goods. For utilitarians, it is the total amount of happiness that counts, not the good of society. It is, like the contractarian

society, a community of (individual) interests.¹² Even Marxism, which is usually seen as a form of collectivism, contains an individualist strand to the extent that it views society in an instrumental way. Marxists argue for collective ownership of the means of production. However, the end-goal remains formulated at an individual level: the unalienated, happy and free individual¹³ that remains at the heart of our imagination in modern times from Romanticism to contemporary consumerism. In this sense, real collectivism has never existed in the West. Even the totalitarianisms of the twentieth century (e.g. Nazism, Fascism, Stalinism, Maoism), which promoted the subordination of individuals to the state, did not embrace a radical collectivism given their leadership cults (a phenomenon that is not exclusive to totalitarianism; it can also be seen in current democratic societies). Thus, for a real contrast we must look elsewhere, beyond Western modern liberal society.

Communitarians (and virtue ethicists) have a substantially different view of the relation between individuals and society: opposing liberal individualism, they ascribe value to the community itself and see individuals as members of the community, as being shaped by that community. Authors such as MacIntyre (1984) and Taylor (1989) appear to hold a relational ontology: the point of being virtuous is the building of a moral community and there is no fundamental difference between fostering individual morality and fostering the morality of the community. In this way, both the member of the community and the community itself are not mere means to an end but ends in themselves. Thus, communitarians are neither individualist nor collectivist.¹⁴ However, communitarianism typically

¹² An alternative interpretation of utilitarianism would argue that increasing the total amount of happiness cannot be done without help from others and that the cultivation of mind and other ‘higher’ pleasures require or even *presuppose* others and society. But a ‘societal’ interpretation is not encouraged by utilitarianism’s founders. For instance, Mill insists in Chapter 2 of *Utilitarianism* that ‘it is a misapprehension of the utilitarian mode of thought, to conceive of it as implying that people should fix their minds upon so wide a generality as the world, or society at large. The great majority of good actions are intended, not for the benefit of the world, but for that of individuals, of which the good of the world is made up’ (Mill 1867, p. 27).

¹³ As Marx (1844) puts it in his *Comments on James Mill*, the goal of ‘production as human beings’ is to objectify my *individuality*, that is, to enjoy the ‘individual manifestation of my life’ and to have the pleasure of ‘knowing my personality to be objective’. And as Nordahl (1987, p. 780) argues, in Marx’s view of the communist society, the individual is not absorbed into the whole but instead human beings structure their social relations in order to ‘to maximize their enjoyment as creators’.

¹⁴ One may well question if contemporary communitarians really hold this view or rather subscribe to modern Romantic individualism with a communitarian flavour. I believe this is the case if they focus on ‘the good life’ and virtue of the individual without really ascribing to the communal relationalism described here.

¹¹ Searle’s (1995, 2006) social ontology can be interpreted as belonging to this tradition. An alternative approach would be to understand the social as prior, giving shape to individuals and individual lives—perhaps even to what we call ‘reality’ (consider phenomenological and social constructionist approaches). I will refer to social constructionism below but I will not further discuss the differences between Searle and social constructionism here.

restricts the boundaries of the moral and social community to the human world. While there have been efforts to connect ‘community’ and ‘ecology’ by authors with ecological concerns¹⁵ and while as said above virtue ethics can employ *indirect* arguments for moral consideration of non-humans, following its Aristotelian roots ‘classic’ communitarianism and virtue ethics are directed at the moral quality of humans and their *human* communities.

A similar anthropocentric limitation can be found in most social constructionist theories. They move beyond individualist and contractualist theories by understanding individuals as dependent on social interactions and, more generally, on a social order or context that transcends them (see for instance Berger and Luckmann 1966). However, the individuals and the social context in which they shape themselves remain within the boundaries of the human. (An interesting exception is Latour’s work: it attempts to cross the nature/society distinction, the human/non-human distinction, and other conceptual distinctions Latour (1993, 2004, 2005) attributes to modern thinking by introducing the notion of the collective as a hybrid assembly of humans and non-humans. However, I will not further discuss his work here.)

Other candidates for relationalism may be found in non-modern, non-Western cultures (keeping in mind that ‘pure’ forms of these cultures do not exist, if only because today most of the countries where we would locate them, for instance in Africa or Asia, the societies are soaked with Western, modern values). For instance, it appears that traditionally Chinese and Japanese cultures involve relational ontologies: humans are not ‘individuals’ but are related to other humans and other entities.¹⁶

However, perhaps one of the most relational views available to us can be found in current Western societies: ecology. I mean not only ‘deep ecology’, the normative view that challenges anthropocentric approaches, but also ecology defined as a branch of natural science, which studies relations between organisms and relations between organisms and their environment. Both kinds of ecology have usually little to say on the social. Usually they come in the form of a natural ecology which is not, by itself, a *social* ontology. However, natural ecology can be used as a model to construct a social ontology that includes some

artificial entities (which is my interest here) and perhaps a universal ontology (which seems to be Floridi’s aim when he uses ‘ecology’ in relation to the infosphere). Of course I do not have the space to do anything like that in this paper; I limit myself to making some suggestions of what an ecological social philosophy would look like. This exploration is important since within the limited space of this paper I wish to give more substance to the idea of a social-relational approach to moral consideration: if it is not built on an individual-ontological foundation but on a social-relational ontology, then what is this relational ontology?

The concept of ‘relations’ is vague and allows for much variety in the way it informs the construction of an ontology. On the one hand, the emphasis can be put on the *relata*, here the entities. For example, Platonic and Aristotelian views of the human are only relational in a weak sense given their essentialism, which amounts to the assumption that there are intrinsic, alienable features of the entity. On the other hand, the emphasis can be put on the relations. However, what does this amount to? Does it imply collectivism? Eastern worldviews, though originally very relational, have in practice fused with modern nationalism in a way that has changed their relational character: they put so much emphasis on the collectivity that its ontology looks more like a collectivist version of ontological essentialism: not the individual but the *collectivity* has intrinsic features and both the relations and their (human and non-human) *relata* tend to become less important. The nation has essential features and below are only organs, cells, and smaller units. However, does emphasis on relations imply that the *relata* are less important (morally and ontologically) or even do not exist?

Ecological thinking (deep ecology and ecology as science) does seem to put emphasis on relations. In doing so it is fundamentally different from individualist, essentialist, collectivist, and totalitarian thinking. But what is the status of the *relata*? One answer is: they do not exist. To explore this direction of thought, we must radicalize existing concepts of natural ecology. Often it still uses early-modern organicist notions: ecology (as a branch of evolutionary biology and natural science) talks about dependencies between ‘organs’ or between ‘organisms’ and their environment. And some deep ecologists see ‘the earth’ as an organism (a misleading term by which they mean all life on earth or the ecosystem). However, if everything is truly related and interdependent, why make such a strict distinction between the border of ‘organisms’ or ‘organs’? A radically relational ecology would not accept such an ontology. There are relations between *relata*, but these *relata* have no fixed ontological reality. They might appear to us as wholes; however, this might be only appearance. But that does not matter; we can work and live with appearances. A different answer is that the *relata* exist, although

¹⁵ See for example Clark 1997 for a Marxist view and Midgley 1984 for the notion of a ‘mixed community’.

¹⁶ In information ethics it has been recognized that these cultural differences are relevant for our moral thinking. Consider for instance the 2005 special issue on ‘Privacy and Data Privacy Protection in Asia’ in this journal (Ess 2005). Note also that it is not obvious that these cultures are entirely non-individualistic. One may argue, for instance, that in Japanese Shintoism ‘individual’ objects—natural and artificial—are seen as having spirits. However, this is a very different kind of ‘individualism’ than Western anthropocentric individualism.

we cannot give an essentialist definition of them but only a naturalist explanation, using evolutionary theory for example. However, this by itself does not solve the question about the ontological status of that-which-evolves.

Developing a more comprehensive social ecology will have to answer these questions. For the purpose of this paper, let me give a brief definition and description of the social ecology I have in mind, one which could serve as a basis for giving moral consideration to some robots.

A social ecology is about relations between various entities, human and non-human, which are inter-dependent and adapt to one another. These relations are morally significant and moral consideration cannot be conceived apart from these relations. Again, there is a question about the status of the relata. The social relata are individuals, groups, societies, communities, cultures, and indeed some robots provided that they participate in the social life. What is their status? We talk about them in the relata in an individualist and essentialist way since this is how they appear to us. However, the social and natural sciences show us how entities are inter-related and suggest a less individualist and non-essentialist view of the world. We may need our appearances in order to live and in order to live together. But appearances can change; we can change them as much as they change us. And in the end we do not know if these entities are real and if their boundary is fixed. This view is distinct from Buddhist ontology, which assumes that we can answer questions like these, that we can know the ultimate truth about what is and can clearly distinguish between reality and appearance.¹⁷ It is also distinct from ancient Greek essentialism and its contemporary heir: the scientific quest for ‘elementary particles’ which are supposed to be the most elementary relata.

This (social) ontology needs more work. At this point I can only say that making a relational argument cannot avoid such discussions. Arguments for moral consideration always make a jump from ‘is’ to ‘ought’, not in the sense that they mistake the normative for the descriptive, but in the sense that they rest on a view of the world, including the social world. Both relational and non-relational views of moral consideration need to make explicit their ontologies.

What follows from a social-relational ontology for moral consideration of non-human entities? Within relational views, there are no longer a priori and ontological hierarchies between entities like the ancient Aristotelian or modern neo-Aristotelian ones we are used to. This does not imply that we can no longer make moral distinctions;

rather, it is not a priori decided what (apparent) moral status we give to entities on the basis of a hierarchical and essentialist ontology. Reasoning about moral consideration of other entities, then, can only be done within a relational context, one which we experience in practice or in imagination and which is always open to change. However, what can be said in addition to this? What does the social-relational approach articulated above mean for the discussion about moral consideration of intelligent robots?

Implications for robots, animals, and humans

Let me list some of the implications:

First, in this approach to moral consideration it no longer makes sense to talk about moral consideration of *robots* in general, for example *robot* rights. Such a manner of speaking about robots is not only misleading since it puts all robots into one category but also since it suggests that moral consideration is entirely non-relational. Instead, this approach acknowledges that moral consideration is bound up with social relations between humans and robots. Therefore, it can pay attention to similarities *and* differences between treatments of robots.

Second, whether or not one day we will have conscious and sentient robots, there will be a long stage in the development of robots during which artificially intelligent robots do not meet ‘high’ or ‘hard’ criteria of standard approaches to moral consideration. This situation leaves us with two options: either we deny any moral consideration to such robots or we grant them *some* moral consideration—that is, a *different* degree of moral consideration—on a different basis. Within the standard approach, the latter option may imply defining less demanding criteria, such as those proposed by Floridi and Sanders (2004). However, as a non-relational approach it does not sufficiently take into account the new *social* functions of intelligent robots: they appear to us as social entities, not only as machines or systems with certain features. The alternative approach proposed in this paper claims to do that by proposing a relational account of moral consideration based on a social ecology that includes humans and robots. Floridi (2008) also supports an ecological approach. However, that is an *informational* ecology: it is about relations between information (and ‘inforgs’ as carriers or instantiations of information). Such an approach is similar to the ‘elementary’ and essentialist approach of modern physics and ancient Greek metaphysics: it tries to find an elementary reality ‘behind’ the appearances. My approach retains the common sense idea that the world consists of a wide variety of entities—and, of course, relations between these entities—without reducing these entities to information. The implication is that attention can be paid to social relations between humans and robots and to the social

¹⁷ Although there are differences between Buddhist teachings on reality, they all seem to assume that we can know *that* perceived reality is an illusion or that we perceive ourselves as separate from the rest of the world while *actually* we are part of it.

structures within which these relations are shaped. This supports reflection on what (degree of) moral consideration we wish to give to those robots. On this basis, we can arrive at a range of forms of moral consideration that will be less strong than robot rights but still imply some obligations towards robots in the context of particular human-robot relations. It is not possible to say what these obligations are a priori, that is, apart from and abstracted from those relations and contexts. This does not mean that one can no longer generalize and compare between different kinds of relations and contexts. Rather, it implies that it is an illusion to think that we can have a moral theory divorced from the social world in which that theory makes sense and in which it is practiced and lived.

Of course, to talk of ‘obligations’ we put the emphasis on the giver of moral consideration. This will probably remain so as long as we feel that social robots must be given some moral consideration but not the same degree as humans, who can literally *claim* their rights in social contexts. However, such intuitions can and will change when some robots will appear differently to us. If that happens and one then wishes to put the emphasis on the receiver of moral consideration, one could use the language of rights but widen the moral vocabulary. For instance, one could talk about ‘soft rights’: rights given to some robots on account of their participation in the social life. These rights would not be as ‘hard’ as for example human rights given the lower degree of moral consideration¹⁸ and given their social-relational basis. However, to use the terms ‘soft’ and ‘hard’ would suggest that the standard approach to moral status is the best one and/or provides epistemic certainty, two assumptions which I have questioned in this paper. So we might want to create an altogether different vocabulary.

Third, this approach can also be applied to animals and vice versa. We can learn from analogies. First, we might apply the ‘hard’ rights/‘soft’ rights vocabulary. We could grant ‘soft’ rights to some animals that participate in the social life. However, if we take this justification seriously we can also try to go beyond the language of rights altogether and adopt an approach that is more radically social-relational. Consider meat production in industrial societies. Instead of asking first *what kind of animal* a pig is, we must study and evaluate relations between humans and pigs within meat production systems and within industrial society and compare this with other human-animal relations such as human-pet relations. This allows us to understand and question our ideas about

moral consideration of animals. Similarly, we must study and evaluate moral consideration of social robots as different from consideration of industrial robots not on account of their intelligence or other features (*what kind of robot it is*), but on account of the human-robot and other relations within a larger social ecology. In this way we can critically reflect on our current views of moral consideration.

Fourth, this approach can also be applied to moral consideration of humans. Consider the concept of human rights. Of course there have been well-known moderately relational objections to, for instance, human rights, such as the feminist or Marxist critique.¹⁹ In practice, a focus on rights sometimes obscures structural (power) issues (related economic and gender differences) and other reasons why in spite of agreement on the moral-legal concept of human rights there is still much human suffering due to injustice. One could also make a utilitarian argument for alleviating human suffering that is not based on the idea of human rights. One might also employ the capability approach to clarify why human rights do not necessarily *empower* people to live their lives in dignity.²⁰ However, existing (Western) criticisms share a justificatory basis in ontological features of humans. Marxists, utilitarians, and neo-Aristotelians have views of ‘human nature’ that are not relational or not *radically* relational. They still assume a human essence that stands apart from the non-human world. As said before, for a more radically relational view one has to learn from natural ecology or perhaps non-Western worldviews if it were possible to purify them from modern-nationalist influences (e.g. ancient East Asian philosophies). These views resist and sometimes do not even *understand* the concept of human rights since it is so much linked up with dominant individualist and/or essentialist, non-relational Western worldviews.²¹ Thus, if we want to rethink moral consideration of humans and related moral and political concepts such as human rights, we need to engage with different kinds of social ontologies and

¹⁸ Note that my suggestion that some robots would be given a *lower* degree of moral consideration than humans would be challenged by ‘singularitarians’ like Kurzweil (2005), who argue that technology will accelerate in such a way that robots and other artificial agents will outstrip us in capacities—which, I presume, would justify a *higher* degree of moral consideration according to the standard approach to moral status.

¹⁹ Marx already criticized human rights for assuming egoistic, inward-looking individuals separated from others. Rather than universal rights, they are the rights of the *bourgeois* who want to protect their private property. Feminists add that they are the rights of *male* bourgeois.

²⁰ See for example Nussbaum (2000, 2006). The idea is that ascribing human rights or dignity to people remains abstract and perhaps even meaningless unless those rights and that dignity are interpreted as requiring that we improve what people are actually capable of doing. In Nussbaum’s sufficitarian view of justice, this means that we should not permit that capabilities of human beings fall below a minimum threshold.

²¹ Of course our societies are not internally homogenous (they are always hybrids), there are differences between these societies, and ‘the West’ does not have a clear border or definition. But one cannot deny that there are general tendencies in the way most people in Europe and the US see the world and that in other parts of the world there are other tendencies.

make explicit our position. We cannot ‘do ethics’ without occupying ourselves with these wider philosophical issues and practical problems.

Finally, note that those who are unwilling to depart from the familiar, non-relational approach to moral consideration could of course try to combine the two approaches. For instance, Warren (1997) has developed a multi-criteria view of moral status: moral status depends on relational and non-relational properties. However, apart from the fact that her account applies only to living things—a limitation that can be overcome—her account of relational criteria is not systematic. It calls for attention to social relations (which it shares with ethics of care and feminist ethics), but unlike the framework sketched here, it stops there and suggests that its turn to relations has no fundamental consequences for the non-relational account. Warren’s account does too little to avoid the impression that we can have relational and non-relational criteria side by side without further problems, that we can simply *add* relational criteria. However, the issue of moral consideration cannot be dealt with by accumulating criteria: multiple criteria as such are acceptable, but there are strong tensions between a more relational approach and a more individual-ontological approach. A combination or ‘synthesis’ seems difficult since both approaches are connected to different views about the social. Instead Warren’s view seems to suggest that we can (and should) avoid choosing one of these moral-social directions, that we can have it all. It asks us to use multiple criteria in order to ‘represent all the relevant considerations’ (Warren 1997, p. 177) but fails to pay sufficient attention to tensions between the ‘criteria’ in terms of their social ontologies. Moreover, her use of the language of ‘properties’ (of the entity), which is common to most existing approaches, tends to deny the nature and potential of the ‘paradigm shift’ offered by a relational approach. Can we speak of ‘properties’ at all once we adopt a relational approach? On the one hand, it seems to me that within a radically relational ontology, relational properties do not exist since one can no longer make sense of the idea of a ‘property’, something that belongs to an entity. One is not conscious in the way one owns a house or a car. What appears to belong to an entity is always open to change and makes only sense in relation to that entity. On the other hand, given the importance of appearance, one could still speak of properties as long as it is understood that we mean properties-as-they-appear-to-us within a social-relational, social-ecological context.

Conclusion

Whether or not it is acceptable to grant *rights* to some robots, reflection on the development of artificially intelligent robots reveals significant problems with our existing

justifications of moral consideration. This forces both defenders and opponents of robot rights to reconsider their conceptual frameworks. In this paper I have offered an alternative, social-relational approach to moral consideration, which reframes the issue of moral consideration by shifting the focus from rights and properties to relations. This approach invites us to explore radically relational, ecological ontologies. It has implications beyond theory of moral consideration and applies to artificial as well biological entities.

Let me unpack this conclusion in order to further clarify this paper’s contribution to discussions about robot rights and moral consideration. Most current and, as we can expect, near-future robots will not meet the ‘hard’ or ‘high’ criteria set by the deontological, rights approach or the utilitarian approach. However, this paper shows that there are enough other conceptual resources available to grant some kind and some degree of moral consideration to robots. My reflection on these alternative resources involved an analysis of the standard direct and indirect arguments for moral consideration. This analysis revealed those arguments as based on ontological features of the entity in question, which incurs particular (and often well-known) problems. In response to these problems I offered a first sketch of a novel argument for moral consideration based on social relations.

This approach might be used by both defenders and opponents of robot rights, provided that they justify their views by relying on social-relational arguments. However, if they really digest the relational approach, they will feel invited to reframe the issue as a question about how to shape our relations to robots instead of a question about moral status, properties, or rights. Rights appear as meta-properties: moral properties based on properties such as consciousness or sentience. By contrast, the relational approach suggests that we should not assume that there is a kind of moral backpack attached to the entity in question; instead, moral consideration is granted within a dynamic relation between humans and the entity under consideration.

Moreover, I have shown that such an approach to moral consideration does not stand on its own but implies that we should also revise our ontological and social-political frameworks. Liberal-individualist, collectivist, and communitarian conceptions of the social are challenged by the idea of a (radically) relational ontology, which draws our moral attention to relations—with other humans as well as with non-human entities. I have suggested that we may want to turn to social ecology to develop this approach and have made brief comparisons with existing Western and Eastern political theory in order to make explicit what I have in mind.

Of course this ‘relational turn’ raises many issues—more questions remain than I can answer here. In

particular, while as I have argued a simple addition of relational and non-relational criteria should be avoided, it is still not entirely clear if a turn to (social) relations demands a ‘paradigm shift’ that completely abandons existing ethical and social theories or if it requires a more moderate revision of these theories. For example, is this approach (in)compatible with a virtue ethics and communitarian framework? Do we need to accept a radically relational ontology and what does that mean exactly? What are the implications of the approach for well-known ethical and political concepts? I have briefly explored the issue of (human) rights but what are the implications for other concepts such as justice and equality²²? What exactly can we learn from natural ecology and from East Asian philosophy? Much more work has to be done to develop this argument and this incomplete account, to further compare it with other relational and with non-relational accounts (e.g. to further specify its relation to phenomenology and feminism), to spell out and elaborate its ontological and social-philosophical assumptions and implications, and to apply it to robots and other entities, including humans. However, here is a rough outline of an alternative approach to moral consideration that may assist us in shaping our relations to intelligent robots and, by extension, to all artificial and biological entities that appear to us as more than instruments for our human purposes.

Acknowledgments I thank the anonymous reviewers for their critical comments and suggestions, which helped me to further develop this paper. I am also grateful to Steve Torrance for his comments and copyediting advice in response to the penultimate version of the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- UK Government. (2006). *Robo-rights: Utopian dream or rise of the machines?* (report) Office of Science and Innovation’s Horizon Scanning Centre, Dec.
- Asaro, P. M. (2006). What should we want from a robot ethic? *International Review of Information Ethics*, 6, 9–16.
- Asimov, I. (1942). *Runaround*. (short story) *Astounding Science Fiction*, March.
- Berger, P. L., & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. Garden City, NY: Anchor Books.
- Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous Systems*, 42, 167–175.
- Brooks, R. (2000, June 19). Will robots rise up and demand their rights? *Time*.
- Calverley, D. J. (2006). Android science and animal rights, does an analogy exist? *Connection Science*, 18(4), 403–417.
- Clark, J. (1997). A social ecology. *Capitalism Nature Socialism*, 8(3), 3–33.
- Coeckelbergh, M. (2009a). Persona robots, appearance, and human good: A methodological reflection on roboethics. *International Journal of Social Robotics*, 1(3), 217–221.
- Coeckelbergh, M. (2009b). Virtual moral agency, virtual moral responsibility. *AI & Society*, 24(2), 181–189.
- Coeckelbergh, M. (2010). Moral appearances: Emotions, robots, and human morality. *Ethics and Information Technology*. doi: 10.1007/s10676-010-9221-y.
- Dautenhahn, K., et al. (2005). What is a robot companion—friend, assistant, or Butler? Intelligent robots and systems, IEEE/RSJ International Conference on Intelligent Robots and Systems.
- Dombrowski, D. A. (1997). *Babies and beasts: The argument from marginal cases*. Urbana-Champaign: University of Illinois Press.
- Ess, C. (2005). Lost in translation?: Intercultural dialogues on privacy and information ethics. *Ethics and Information Technology*, 7, 1–6.
- Floridi, L. (2008). Artificial intelligence’s new frontier: Artificial companions and the fourth revolution. *Metaphilosophy*, 39(4–5), 651–655.
- Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and Machines*, 14(3), 349–379.
- Gilligan, C. (1982). *In a different voice: Psychological theory and women’s development*. Cambridge, MA: Harvard University Press.
- Gruen, L. (2003). The moral status of animals. *Stanford encyclopaedia of philosophy*. Retrieved from <http://plato.stanford.edu/entries/moral-animal/>.
- Hacker-Wright, J. (2007). Moral status in virtue ethics. *Philosophy*, 82, 449–473.
- Hursthouse, R. (1999). Virtue ethics and human nature. *Hume Studies*, 25(1–2), 67–82.
- Kant, I. (1785/1991). *Groundwork of the metaphysics of morals*. (H. J. Paton, Trans.) London: Routledge.
- Kittay, Eva F. (1999). *Love’s labor: Essays on women, equality, and dependency*. New York: Routledge.
- Kurzweil, R. (2005). *The singularity is near*. New York: Viking.
- Latour, B. (1993). *We have never been modern* (C. Porter, Trans.). Cambridge, MA: Harvard University Press.
- Latour, B. (2004). *Politics of nature: How to bring the sciences into democracy* (C. Porter, Trans.). Cambridge, MA: Harvard University Press.
- Latour, B. (2005). *Reassembling the social: An introduction to Actor-network theory*. Oxford/New York: Oxford University Press.
- Leopold, A. (1949/1966). *A sand county almanac*. New York: Oxford University Press.
- Levy, D. (2009). The ethical treatment of artificially conscious robots. *International Journal of Social Robotics*, 1(3), 209–216.
- MacIntyre, A. (1984). *After virtue: A study in moral theory* (2nd ed.). Notre Dame, IN: University of Notre Dame Press.
- MacIntyre, A. (1999). *Dependent rational animals: Why human beings need the virtues*. Chicago: Open Court.
- Marx, K. (1844/1975). Comments on James Mill. In K. Marx & F. Engels (Eds.), *Collected works* (C. Dutt, Trans.). London: Lawrence and Wishart.
- McNally, P., & Inayatullah, S. (1998). The rights of robots: Technology, law and culture in the 21st Century. *Futures* 20:

²² For instance, if we endorse the approach presented here and thus extend the social to (other) biological and artificial entities, can we maintain and justify the principle of equal respect for humans? And does a social-relationist approach commit us to impartiality between humans and other entities that are part of the social sphere? Trying to answer these questions is not just a matter of ‘application’ but demands further development of the approach itself—in relation to existing moral, social, and political theory.

- 119–136. Available at <http://www.metafuture.org/Articles/TheRightsofRobots.htm>.
- Midgley, M. (1984). *Animals and why they matter*. Athens: University of Georgia Press.
- Mill, J. S. (1863/1867). *Utilitarianism* (3rd ed.). London: Longmans, Green, Reader, and Dyer.
- Noddings, N. (1984). *Caring: A feminine approach to ethics and moral education*. Berkeley: University of California Press.
- Nordahl, R. (1987). Marx and utopia: A critique of the “Orthodox” view. *Canadian Journal of Political Science/Revue canadienne de science politique*, 20(4), 755–783.
- Nussbaum, M. C. (2000). *Women and human development: The capabilities approach*. Cambridge: Cambridge University Press.
- Nussbaum, M. C. (2006). *Frontiers of justice: Disability, nationality, species membership*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Regan, T. (1983). *The case for animal rights*. Berkeley: The University of California Press.
- Ruddick, S. (1989). *Maternal thinking: Toward a politics of peace*. New York: Ballantine Books.
- Sartre, J. -P. (1943/1969). *Being and nothingness: An essay on phenomenological ontology*. Abingdon: Routledge.
- Searle, J. R. (1995). *The construction of social reality*. New York: Free Press.
- Searle, J. R. (2006). Social ontology: Some basic principles. *Anthropological Theory*, 6(1), 12–29.
- Sharkey, N. E. (2008). The ethical frontiers of robotics. *Science*, 322, 1800–1801.
- Singer, P. (1975). *Animal liberation*. New York: Random House.
- Singer, P. (1993). *Practical ethics* (2nd ed.). Cambridge: Cambridge University Press.
- Taylor, C. (1989). *Sources of the self: The making of the modern identity*. Cambridge, MA: Harvard University Press.
- Torrance, S. (2008). Ethics and consciousness in artificial agents. *AI & Society*, 22, 495–521.
- Veruggio, G. (2006). EURON roboethics roadmap (release 1.1). EURON Roboethics Atelier, Genua.
- Wallach, W., & Allen, C. (2008). *Moral machines: Teaching robots right from wrong*. Oxford: Oxford University Press.
- Warren, M. A. (1997). *Moral status: Obligations to persons and other living things*. Oxford: Oxford University Press.
- Whitby, B. (2008). Sometimes it’s hard to be a robot: A call for action on the ethics of abusing artificial agents. *Interacting with Computers*, 20(3), 326–333.
- United Nations. The Universal Declaration of Human Rights. Retrieved from <http://www.un.org/en/documents/udhr/>.