

Distributive Justice and Co-Operation in a World of Humans and Non-Humans: A Contractarian Argument for Drawing Non-Humans into the Sphere of Justice

Mark Coeckelbergh

Published online: 28 March 2009
© Springer Science+Business Media B.V. 2009

Abstract Various arguments have been provided for drawing non-humans such as animals and artificial agents into the sphere of moral consideration. In this paper, I argue for a shift from an ontological to a social-philosophical approach: instead of asking what an entity is, we should try to conceptually grasp the quasi-social dimension of relations between non-humans and humans. This allows me to reconsider the problem of justice, in particular distributive justice. Engaging with the work of Rawls, I show that an expansion of the contractarian framework to non-humans causes an important problem for liberalism, but can be justified by a contractarian argument. Responding to Bell's and Nussbaum's comments on Rawls, I argue that we can justify drawing non-humans into the sphere of distributive justice by relying on the notion of a co-operative scheme. I discuss what co-operation between humans and non-humans can mean and the extent to which it depends on properties. I conclude that we need to imagine principles of ecological and technological distributive justice.

Keywords Distributive justice · Contractarianism · Animals · Artificial agents · Rawls · Liberalism · Imagination

Introduction

Various arguments have been provided for drawing non-humans such as animals and artificial agents into the sphere of moral consideration: animals have been attributed rights or equal consideration of interests; plants and the ecosystem have been attributed intrinsic value; and artefacts have been attributed agency (e.g.

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

artificial agents) or ‘morality’. Corresponding justifications given for these attributions include the inherent value or capacity to suffer (animals; Regan 1983; Singer 1975), being alive or being part of a spiritual-ecological whole (plants), or a high degree of system interactivity, autonomy and adaptability (artificial agents; Floridi and Sanders 2004).

In this paper, I explore a less usual and perhaps more radical route towards drawing non-humans into the sphere of moral consideration: I tap into the vocabulary of *justice* and ask the question of distributive justice with regard to non-humans. What justification can be given for drawing non-humans into the sphere of justice, in particular for making principles of distributive justice relevant to them? Justice for biological non-humans has already been discussed under the label of ‘environmental justice’ (Dobson 2000) or ‘ecological justice’ (Baxter 2005; Bell 2003; Low and Gleeson 1998). But what about non-biological and non-ecological entities? Can we construct a conceptual framework concerning justice that applies to biological and artificial non-humans alike? Engaging with the work of Rawls and Nussbaum, I explore if and how the contractarian framework of human distributive justice can be expanded to include non-humans.

I focus on distributive justice since this is a domain of justice which allows me to engage with one of the most influential theories of justice (Rawls), but the scope of much of the argument in this paper is wider: it is also applicable to other dimensions of justice and to moral consideration in general.

My paper is structured as follows. First, I argue for a shift from an ontological to a social-philosophical approach: we should try to grasp conceptually the quasi-social dimension of relations between non-humans and humans. Discussing the work of Rawls, I show that an expansion of the contractarian framework of human distributive justice that includes non-humans causes trouble for liberalism but *can* be justified by a contractarian argument. I comment on the relation between liberalism and ecological justice, paying particular attention to Derek Bell’s interpretation of Rawls. I also respond to Nussbaum’s criticism of Rawls and argue that we can justify drawing non-humans into the sphere of distributive justice by relying on the notion of a co-operative scheme. I then further discuss the meaning of dependency and co-operation with regard to relations between humans and non-humans. Finally, I explore which principles of distributive justice we need with regard to such relations. I conclude that we need to develop our imagination in order to discover and acknowledge the co-operative links between humans and non-humans and to arrive at principles of ecological and technological distributive justice.

From Ontology Towards Social Philosophy

Tell Me What You are and I Will Tell You If You Have Rights or Value

Debates about the moral inclusion of biological and artificial non-humans are usually strictly separated. But they share a focus on ontology: they share the assumption that the crucial criterion that allows us to draw non-humans into the

moral sphere is about what these non-humans *are*, in particular which features they have and how they can be classified accordingly. For instance, animals have been attributed rights or equal consideration of interests, plants have been attributed intrinsic value, and artificial agents and artefacts have been attributed morality. As I already suggested in my introduction, in most cases these attributions have been justified by reference to features of the organisms or artefacts, such as intrinsic value (Regan 1983 on animals), the capacity to suffer (Singer 1975 on animals), being alive, being part of a spiritual-ecological whole, exhibiting a high degree of system interactivity, autonomy and adaptability (Floridi and Sanders 2004 on artificial agents).

The approach I will now propose shifts attention away from what non-humans are towards what 'we' (humans/non-humans) *do together*. I highlight their (quasi) *social* dimension: the moral relevance of their relations with humans and other non-humans. I argue that we should draw non-humans into the sphere of moral consideration not only because of what they are (features) or do as such (consequences) but because of their relations with us, in particular their (quasi-co-operation) with us. Instead of using the language of rights or value that focuses on the moral relevance of individual capacities non-humans, I wish to put the emphasis on the quasi-social dimension of non-humans, that is, their relations with other non-humans and with humans. Let me develop my argument by distinguishing it from other radical approaches to the issue.

From Ontology Towards Quasi-Social Philosophy of the World of Humans and Non-Humans

My argument turns from theoretical attention from ontology (what these entities really are; the inherent features of certain species or enteritis) towards social philosophy (how they 'live' together). But what kind of social philosophy is needed? Existing social philosophy is tailored to relations between humans. How can we incorporate non-humans, make them part of the social body, the social *corpus*? Let me distinguish between three strategies of incorporation.

A first way of understanding the relation between humans and non-humans is using a *network* metaphor. Latour's descriptive actor-network theory (ANT) could be used for this purpose. ANT is developed by Latour and Callon to map material and semiotic networks: relations between material things, people, and ideas (concepts, meanings). Agents in the network are so-called 'actants', which can be human or non-human (Latour 1993, 2005). ANT has been developed and used to study scientific and technological research, but its scope can be broadened to study various relations. ANT has the advantage of taking seriously the hybrid nature of our world, by which I mean a world that consists of humans and non-humans. And if we were to argue, with Latour, that non-humans should be politically represented on account of their being part of the network (or networks) (Latour 2004), this could be followed by the question of whether such entities would be justified to 'voice' claims of justice and if we have the duty to respond to their claims. However, such a manner of speaking about non-humans demands radical shift in thinking about the place of humans within the larger world of humans and non-humans. The notion of

‘actants’ suggests that humans and non-humans are morally equal participants in the network. I would like to search for a less symmetrical approach, one that does not suggest or presuppose moral equality.

A second approach would be to embrace a *deep ecology* (Næss 1973) which also advocates a radical paradigm change but—and this is seldom noted—requires a slightly less non-anthropocentric move from us. Organic metaphors (e.g. organic food, the ecosystem as an organism) and family relation metaphors (e.g. Mother Earth) are anthropomorphic. It is easier for us to talk about non-humans in this way, since we are used to talking about ourselves in this way. But this approach contrasts sharply with our modern way of thinking about politics and economy, and seems to advocate a world of harmony that is lost forever—and perhaps never existed at all. Furthermore, its ideal is a ‘biospheric egalitarianism’ which again puts humans and non-humans on a morally equal level. We are not just members of a ‘biotic community’ (Leopold) but *equal* members. Such an eco-centrism is difficult to accept and even harder to realise in practice, since it supposes that humans and non-humans share a lot of features and values. Finally, deep ecology seems to depend on the notion of intrinsic value, which leaves us once again with an ontological approach.

A third approach stays closer to our common understanding of the human sphere and explores widening its scope to include non-humans. It is consciously and deliberately ‘anthropocentric’—not in scope but in its use of vocabulary and concepts. However, it avoids terms such as ‘rights’ and ‘inherent worth’, since these depend on an (isolationist) ontology: entities of the kind X have these properties and have these capacities, therefore we attribute moral status A to them. Instead, it considers the relations between humans and non-humans as *social* relations. *We* (living, non-living things, and humans) do things *together*. I propose to use the term *co-operation* to refer to at least part of the relations between these natural and artificial, human and non-human entities.

Co-operation is not an all-or-nothing affair: there are gradations of co-operation. But a major advantage of using this term is that it is a social term: it is modelled on human relations and is in that sense ‘anthropocentric’ (I prefer the term ‘anthropomorphic’: the form is comparable). However, it does not strongly depend on the co-operands (all) being humans, moral agents, having certain capacities, etc. It does not depend on what the co-operands *are* but on the relations between them. But in contrast to ANT and biospheric egalitarianism (and also in contrast to traditional contract theory—see below), there is no need to assume or demand that the co-operands be moral or ontological equals.

If this brief sketch of a turn towards a ‘(quasi-)social philosophy’ for a world of humans and non-humans is plausible, then we can use this descriptive account as a basis to discuss problems of *justice* in that world. Can non-humans be part not only of the sphere of moral consideration, but also of the sphere of justice? My main argument runs as follows: To the extent that various entities in this hybrid world ‘live’ together and ‘do’ things together, that is, to the extent that they co-operate, they constitute what the contractarian tradition calls ‘a co-operative scheme’, which is in that tradition a sufficient reason for discussing justice and distributive justice. I will conclude that once we come to understand our world partly in terms of

co-operation between various entities and life forms, we no longer have a good reason to exclude non-humans from the moral sphere and our related conceptual frameworks.

Let me now substantiate my contractarian argument by engaging with the literature on distributive justice and inclusion of non-humans, in particular with the discussion about ecological justice and with the tension between Rawls(ians) and Nussbaum on inclusion of non-humans.

(Distributive) Justice and Non-Humans

To develop my main argument, I need to engage with the most influential political theory that concerns *human* distributive justice: the work of John Rawls. What are the possibilities and limits of his theory in relation to distributive justice for non-humans? What are the main points of resistance, and how can they be overcome? In this paper, I shall discuss two points of resistance. In the first section I consider an important *liberal* aspect of Rawls's theory, and discuss that aspect in the light of Bell's arguments concerning Rawls and ecological justice. In the next section I start from the *contractarian* dimension of Rawls's theory, in particular its ontological assumptions, and discuss these in the light of Nussbaum's objections to contractarianism. These discussions will provide me with a liberal and contractarian argument for drawing non-humans into the sphere of justice. I will pay particular attention to issues concerning distributive justice.

Starting from Liberalism: A Liberal Argument for Drawing Non-Humans into the Sphere of Justice

A major obstacle for those who rely on Rawls in the debate about non-humans is that his theory of justice is applicable to humans and human relations only. Part of the explanation for this lies in Rawls's *liberalism*. Let me engage with the discussion about 'ecological justice'. On the way, I benefit from and respond to Derek Bell's objections to Dobson's claim that ecologism and liberalism are incompatible (Dobson 2000, p. 165; Bell 2002, 2003).

In *A Theory of Justice* (1971) Rawls has argued that we are not required to give 'strict justice' to creatures lacking 'the capacity for a sense of justice' (Rawls 1971, p. 512). However, he also argues that this does not mean that we have no 'requirements' in regard to them:

But it does not follow that there are no requirements at all in regard to them, nor in our relations with the natural order. Certainly it is wrong to be cruel to animals and the destruction of the whole species can be a great evil. The capacity for feelings of pleasure and pain and for the forms of life of which animals are capable clearly impose duties of compassion and humanity in their case. (Rawls 1971, p. 512)

However, for Rawls such 'considered beliefs' are 'outside the scope of the theory of justice, and it does not seem possible to extend the contract doctrine so as to

include them in a natural way' (Rawls 1971, p. 512). As Derek Bell has pointed out, there is some ambiguity in Rawls's view on animals and justice, since (1) his use of the words 'strict justice' seems to leave open the possibility of justice to animals, and (2) at the end of the passage he suggests that his account of justice among persons 'cannot be too far wrong when these broader relationships are taken into consideration' (Rawls 1971, p. 512; see also Bell 2003, p. 5). But Rawls's dominant view is that the protection of animals should not be justified in terms of *justice* and should be disconnected from the theory of justice. In *Political Liberalism* (1993) he even suggests that we should protect animals and plants for 'human-instrumental reasons': for our sake, since they are useful to us.

Is this point of resistance insurmountable? I agree with Bell that Rawls's view is based on a *liberal* perspective: Rawls relies on a distinction between, on the one hand, the political, procedural theory of justice (belonging to the public sphere), and on the other hand, comprehensive views about the status of the natural world and our relation to it (belonging to the private sphere; see also Bell 2003, p. 8).

Now I see two responses to this problem. The first one is to bite the bullet: let us accept that our relations to animals—and, by extension, non-humans—could never be governed by a theory of justice and let us provide 'non-political', comprehensive good reasons why we should treat them well. This I understand to be Bell's view: liberalism is compatible with ecology, as long as we consider ecological justice claims as metaphysical and ontological claims. (He even says that we can use the language of justice as part of our comprehensive doctrine (Bell 2003, p. 10).) 'If enough people hold comprehensive doctrines that support nature preservation (generally or in particular cases), a liberal state can legitimately adopt such a policy' (Bell 2003, p. 10). Thus, according to Bell we can be liberal ecologists, although of course then we have to be 'liberals first and ecologists second' since the political conception of justice takes priority over comprehensive doctrines (Bell 2003, p. 11).

At first, this seems an acceptable solution. However, there is a serious problem with this reliance on the liberal distinction between the political and the comprehensive. Rawlsians assume that political principles of justice are not derived directly from any comprehensive doctrine. But on closer inspection, Rawls's view of the limits of the theory of justice depends on such a doctrine. His own argument for excluding animals from the sphere of justice is based on a comprehensive doctrine. I interpret this term as referring to any moral outlook that goes beyond the 'thin' core of the liberal doctrine. The very idea that morality concerns humans alone is a value-laden belief that is not itself required for making possible the social life. It is part of the dominant moral and (post)religious tradition in the West, according to which there is a sharp moral boundary between humans and non-humans. Rawls relies on an ontological argument that is steeped in that tradition: non-humans have or do not have these properties, and therefore they should be excluded from the theory of justice. Given the contradiction between the (comprehensive) views that define the political theory's scope and the ecological comprehensive view(s), Bell cannot have his compatibilist solution.

The failure of the 'liberal' solution, however, does not leave us unable to draw non-humans into the sphere of justice. A second option is to examine and re-interpret the *contractarian* roots of Rawls in order to intervene at the heart of the

theory of justice, that is, to re-interpret the political. For liberals, political liberalism ‘begins from the idea of society as a fair system of co-operation and the idea of citizens as free and equal’, and its principles apply only to ‘society’s basic structure—that is, its main political and social institutions and the way they hang together as one system of co-operation’ (Bell quoting Rawls; Bell 2003, p. 8). In the next section I outline an argument for including non-humans into the sphere of justice by focussing on the element of *co-operation* that is central to the contractarian basis of liberalism. It is not a ‘liberal’ argument if ‘liberal’ requires us to hold on to the purification strategy that divides the political and the comprehensive, but it *is* liberal in its conception of society as a (fair) system of co-operation.

Starting from Contractarianism: A Contractarian Argument for Drawing Non-Humans into the Sphere of Justice

In the previous sections I said that the Rawlsian framework of distributive justice seems to exclude non-humans from the realm of justice, since they lack the capacities and features required for discussing the principles of justice. This impression is not only due to Rawlsian liberalism, but can also be traced back to the theory’s *contractarian* roots. The parties to the social contract are rational, self-interested humans, who discuss the principles of justice on a basis of moral equality between them. Consider one of the criticisms of contractarianism Nussbaum voices in *Frontiers of Justice* (2005):

Because social contract theories start from the allegedly crucial importance of human rationality, defining both reciprocity and dignity in terms of it, they deny that we have obligations of justice to nonhuman animals We should correct such views in two ways: by recognizing the extent of intelligence in nonhuman animals, and by rejecting the idea that only those who can join in the formation of the social contract are full- fledged subjects of a theory of justice. (Nussbaum 2005, p. 93)

Is Nussbaum’s Criticism Justified?

On the one hand, we may wish to defend contractarianism and Rawlsian theory, and respond to Nussbaum that she confuses the parties who decide upon the principles of justice with the parties who benefit from these principles. Apart from the fact that there are other reasons that justify improving the well-being of *non-humans* such as animals without having to make claims of *justice*, why not say that we humans can decide to do justice to animals? What does Nussbaum mean by ‘subjects of justice’? Surely one should not confuse the political question ‘who decides’ with questions regarding a procedure that is to generate the principles of justice (such as the Rawlsian), a procedure which can in principle be followed by an individual, or perhaps even a non-human entity. The Rawlsian procedure is more a thought-experiment than a political method. In other words, we do not need to make assumptions about the rationality or intelligence of animals (or other non-humans)

in order to treat them justly: we only need rational and intelligent humans to make the judgment and decision to treat animals in that way.

On the other hand, Nussbaum's criticism—even though somewhat misdirected if my arguments above are plausible—does show a weakness of the Rawlsian model. Although there may be nothing wrong with the restrictions required of the parties in the original position procedure (which is, to use Habermas's distinction, a matter of *justification*), there is something odd about the sphere of *application*. In the Rawlsian theory of justice, justice seems only relevant to present humans living in one society. It appears to exclude matters of justice relevant to (1) the relation between present humans and future humans, (2) the relation between societies (global justice), and (3) the relation between humans and non-humans (animals, artefacts—including artificially intelligent agents and robots).

With regard to the first two relations, this impression is not entirely correct. Rawls has an account of intergenerational justice (the 'just savings' principle, see Rawls 1971, 1993, 1999, 2001) and international justice (Rawls 1999). (It is questionable, however, how consistent both accounts are with (the original position procedure of) his theory of justice. For instance, Pogge and others have argued, against Rawls, that distributive justice principle should be applied at the global level. I shall not discuss this issue here.) But concerning the third relation, there seems to be a significant lacuna. An adequate theory of justice should also have something to say about distributive justice in the context of these relations. We must at least *ask the question* concerning distributive justice in the relation between humans and non-humans.

For Rawlsians, however, it is misguided to ask such a question: (1) they assume that non-humans are not part of social co-operation and (2) argue that those who are not part of social co-operation are not addressed when it comes to distributive justice. Social co-operation is a necessary condition for the question of justice to arise and for the theory of justice to be relevant and applicable. Rawlsians make only distinctions *within* the human world. But what about non-subjects? What about artefacts? What about natural life forms? Can or should they be included in the domain of justice at all?

I see at least three possible critical responses to the Rawlsian objection and the problem concerning non-humans. A first is to argue that the exclusion is not right, as Nussbaum does, and propose an alternative by making non-humans part of the realm of justice. Nussbaum does that by focussing on capabilities rather than primary goods. But what does the concept of capability mean in relation to non-humans? It is not clear to me how we could think of animals or robots as having capabilities. A second is to agree with the Rawlsian exclusion and argue that we therefore should apply a principle of charity, not of justice, to animals, to nature, and perhaps to some kinds of robots. (This is in tune with Rawls's own solution, as I have said above.) A third and more interesting response, however, is to argue for inclusion of non-humans in the theory of justice by using a Rawlsian argument, which I shall do here.

My interpretation of Rawls resorts to the contractarian concept of a co-operative scheme. If we take seriously the fact that the human world *depends* on the non-human world, then it is not even necessary to blur the categorical line between the moral status of humans and that of non-humans to understand non-humans as an

integral and necessary part of a wider co-operative (quasi-)social scheme. Humans and non-humans are interdependent in various ways. And on closer inspection, what we call a ‘social’ scheme (our, human social scheme) is rather a social-artefactual-ecological scheme: its co-operation is not restricted to inter-human co-operation but depends on more hybrid forms of co-operation, for instance between humans and (certain) non-humans. And we can say this without assuming an ontological hybridisation between humans and non-humans: humans and non-humans need not be hybrid entities themselves in order to co-operate. Instead, the co-operative scheme is hybrid in the sense that it involves different kinds of entities; the parties to the scheme retain their ontological status. Now if this concept of a hybrid co-operative scheme makes sense, then distributive justice, usually applied to ‘social’ justice alone and thus to the ‘merely’ human sphere, should also be applied to this complex conglomerate of co-operation we sometimes call the ‘world’ and co-operative relations within that world. Then it becomes at least *thinkable* that we speak, as some do, about what we ‘owe’ to nature or to animals. These entities are not drawn into the sphere of justice because they are rational, self-interested, or because they have capabilities, but because they are part of a larger co-operative scheme that is not completely within our control but that exists and on which human society depends for its operation. And then we may be justified in using the language justice with regard to these entities, and perhaps hope for ‘a just world’ in this specific sense.

Does this move imply that we have to drop the equality assumption implied in traditional contractarianism? And if we do, what gets lost of the core contractarian foundation of the theory of *justice*? If we interpret contractarianism as requiring moral and ontological equality between the parties to a contract, then indeed this element gets lost. Of course there is a huge difference between what we (humans) are and what non-humans are. If, however, we interpret the equality requirement as meaning ‘formal equality’, then we can remain perfectly loyal to the contractarian equality requirement. With ‘formal equality’ I refer to the assumption that sustains our practices of *human* contracting: the parties are notoriously unequal—for instance a bank giving a loan to an individual—but they are still considered ‘equal parties’ to the contract (e.g. a mortgage contract). The equality is a myth, but a necessary one for the contract to exist. (Note that freedom is a ‘necessary myth’ as well in this context.) Thus, there are two options. Either we save the requirement of equality by interpreting it in this ‘formal’ way, or we drop it altogether. Justice is not equality. Either way, someone who accepts the application of a contractarian theory of justice to non-humans does not have to be committed to biospherical egalitarianism or any other substantial (as opposed to formal) egalitarianism.

Consider current human-animal relations. Humans and some kinds of animals are mutually dependent on each other for their food and living. Now there are circumstances imaginable under which problems of injustice arise: if their co-operation resulted in a wide gap between advantaged (humans) and disadvantaged (some kinds of animals), then it would make sense to discuss this problem in terms of justice or injustice. For instance, using our innate capacity for intelligence and control, we have domesticated animals and use them for our purposes (transport, food, clothing, etc.). We could argue that when animals suffer from these activities,

this is an unfair situation since we are only able to use or abuse them in this way in virtue of our specific natural capacities such as our higher intelligence. Since animals lack these capacities they cannot resist our practices. Just as the relation between intelligence and social class is an issue of justice in the social, human world, one could consider the same problem in the human/animal world.

A similar contractarian argument can be made for artificial agents and some kinds of robots. If our world changes in such a way as to include these entities in our co-operative scheme, then questions and claims with regard to justice with regard to human/robot relations can be justified, without requiring ontological equality.

Note that I do not mean to suggest that this alternative criterion—which does not rely on the features of entities as such but on their being part of co-operative scheme—makes it any easier to decide about an entity's inclusion in the sphere of justice; I only claim that it is a more *adequate* criterion given the contractarian basis of the Rawlsian theory of justice.

Surely my approach differs in at least one crucial way from contractarianism. Consider Nussbaum's description of what classical contractarians wanted to achieve in their time:

Most people believed that aristocrats and peasants were virtually different species of beings, with different proper stations in life, and different political roles. The two groups even looked different, often, on account of differential nutrition and health care. What contractarians wanted to accomplish was to get people to understand the artificiality of these differences and, once they saw that, to see that such artificial differences were not a proper foundation for political society. (Nussbaum 2006, p. 493)

I share with these contractarians a concern for building a more inclusive society (or acknowledging certain entities as being already included in society), which in this paper means including non-humans in the sphere of justice. However, whereas contractarians focused on natural, ontological similarity between the two groups in order to make their case, I point to the shared capacity for co-operation only—regardless of other capacities or features of the entities in question. Of course one needs certain capacities in order to engage in co-operation. But I do not make the 'assumption of roughly equal physical and mental powers' (Nussbaum 2006, p. 493) in order to argue for mutual advantage. Instead, I argue that the parties—human and non-human—are often *already* in a position of mutual advantage and co-operation, *in spite of* or regardless of the wide disparity in capacities ('powers'). Here we need not be concerned with the question of whether such differences are 'natural' or 'artificial'. I point to the hybrid nature of the co-operation (the social question); hybridisation at the entity level (what I called the ontological question) is not an issue for here—at least not as far as justice is concerned. In so far that they are already part of a co-operative scheme, issues of justice can arise.

An additional advantage of my non-ontological approach is that it provides those of us who think about both non-human and *human* justice with a different story of emancipation—if any. Instead of having endless discussions about what properties slaves, women, animals, children, foreigners etc. must have to be granted the right to be part of our moral and political community—which is the traditional approach

and which naturally leads to racism and anti-racism or speciesism and anti-speciesism given its focus on properties—we better consider the inclusion of ‘strangers’ to that community on the basis of the extent to which they stand in a relation of co-operation to us. Instead of requiring that we see, or want to make, the ‘newcomers’ as beings-like-us (for instance see them as, or make them ‘men’: *e-man-cipation*), we can allow for (ontological) differences while acknowledging their co-operative status and therefore our duty to do justice to them, regardless of their specific properties. Similarly, when discussing artificial agents, we need not ask questions such as ‘Are they conscious?’, ‘Do they have feelings?’ etc. in order to treat them justly. Moreover, in this framework justice is done on the basis of what already is the case: the entities are or are not already part of a co-operative scheme. If they are already part of our social world, they do not need to be ‘integrated’ in society. We only need to acknowledge their being-part-of and act justly towards them accordingly. Thus, if ‘emancipation’ is needed at all, it is an emancipation of ‘us’ rather than ‘them’, that is, an emancipation of those who do not sufficiently recognise that both ‘us’ and ‘them’ are already part of a social contract. As far as *justice* is concerned, we need not recognise their being-such-and-such; we ‘only’ need to see them as co-operands.

A further strength of contractarianism here is that it puts forward a ‘we’ that is not the ‘we’ of a particular community but of a particular co-operative scheme, thereby avoiding many difficulties related to non-liberal accounts of society. It is not communitarian: it does not aim for or presuppose a community of humans and non-humans. And it is a liberalism indeed, but one that does not depend on ontology for defining the borders of the sphere of justice (see above), but on the strongest social-philosophical roots it has: contractarianism.

Co-Operation, Properties, and Non-Humans

Let me further clarify the meaning of co-operation and its relation to properties by discussing some more possible objections to my approach.

My use of the term ‘co-operation’ is unorthodox in at least two ways. First, it does not include a notion of voluntariness. Non-humans cannot (dis)agree to be part of a co-operative scheme. Second, the contractarian tradition—especially the Kantian branch—places emphasis on having reasons for co-operation. Most non-humans do not have the capacities required for voluntary or reason-based activity. However, these qualifications offer no good reasons for dismissing the notions of co-operation and a co-operative scheme, since (1) not all humans possess these capacities but we still act as if they are part of the co-operative scheme and (2) the social contract must be interpreted as hypothetical, that is, it does not require that all humans have actually signed it, or have given their permission in any other way. It is a heuristic tool to justify and understand principles of justice. More, in reality we (humans) are not even able to withdraw from the co-operative scheme given our dependencies. Now if we allow this gap between a hypothetical notion and reality for humans, it would be inconsistent to be stricter in the case of non-humans.

Argument (1) says that it is inconsistent to include humans that lack certain capacities while excluding non-humans that lack the same capacities. This argument can be restated as the so-called *argument from marginal cases*, which is a well-known objection to Rawls. Rawls argues that rationality is an essential criterion for the contracting parties. But as Garner puts it, the consequence is that ‘we must treat marginal humans as morally inferior to normal humans, and, equally, we ought to grant an equivalent moral status to marginal humans and the many animals with levels of autonomy broadly the same as them’ (Garner 2003, p. 7). As Garner shows, Rawls fails to answer this objection satisfactory. To conclude, the marginal cases argument challenges ‘the assumption that rationality or autonomy should be the key benchmark for moral considerability’ (Garner 2003, p. 9) and urges us to include animals into the categories of those who benefit from the social contract. More generally, it shows that an approach based on properties is often not in line with our moral intuitions.

Note that these conclusions do not imply that we cannot employ Rawls’s ‘original position’ thought experiment. While it is true that Rawls excludes non-rational parties, there are good reasons for saying that we can dispense with such an assumption. Garner convincingly argues that the exclusion of animals by Rawls is an ‘indiscriminate decision that has nothing to do with his heuristic device’ (Garner 2003, p. 10). But he says this is also true for arguments that assert justice for animals (e.g. by appealing to sentience). In both cases ‘the important normative work is done before the contract is put into motion’ (Garner 2003, p. 10). This suggests that the focus on properties does not follow from the contractarian thought-experiment, but is arbitrarily attached to it by Rawls and others. Garner’s argument shows that there is no logical connection between the contract heuristic, on the one hand, and the assertions about what property is a sufficient criterion for what moral status, on the other hand. If this is true, we must look for another kind of argument, which I have tried to provide.

This point can also be made by reference to the *intuitive equality argument* as discussed by Rowlands (1997). The argument says that if a property is undeserved, then it is morally arbitrary and one is not morally entitled to whatever benefits stem from the possession of the property. If this is true, then properties such as rationality or even having a human status should also be bracketed off in the original position. Therefore, Rowlands concludes, ‘the sphere of justice should not be restricted to human beings’ (Rowlands 1997, p. 243). Now if this is correct then this applies to all properties discussed by those who want to assert or deny moral status or justice to non-humans: they should be bracketed in the original position. Thus, this argument gives another good reason to look for an alternative way of arguing for extending the sphere of moral consideration and justice to non-humans. Or it can be seen as itself constituting an alternative.

Now one may still object that while an entity may not need a will or rationality, it nevertheless needs the properties required for it to engage in co-operation. For instance, it needs to ‘have’ social skills. At first sight this seems to bring back ontological considerations. But this is only apparently so. In contrast to properties such as (self-)consciousness or free will, engagement in co-operation (and dependency) can be observed. Social relations and co-operative activities or their

consequences are visible. No proof is needed that there are mental capacities that make possible this co-operation. Moreover, because of the public character of co-operation it is likely that there is more agreement on this than on the presence of a mental property. Of course, difficulties remain. A different approach means different kinds of problems: we have to define relations of dependency and co-operation. But if no proof of mental properties is needed and agreement on relations of dependency and cooperation can be reached more easily, we can conclude that a social-philosophical approach is more 'economical' than the ontological one. To avoid confusion, I propose to change the requirement 'that the entity have a capacity for co-operation' (an ontological requirement) into 'that the entity co-operates' (something that can be observed or experienced).

However, this argument presupposes that we have a clear idea about what co-operation is. One may rightly ask for an account of co-operation. Which interactions with what kind of (non-human) entities count as co-operative? Let me distinguish between four senses of co-operation. One is a broad notion that includes all kinds of 'doing things together' understood in a rather metaphorical way. In this sense I co-operate with my computer, my desk, etc. A second is the very restrictive sense usually implied: only humans can co-operate. A related sense includes only beings with a will, rationality, etc. Here some (future) artificial agents and some higher animals could be included. The meaning of co-operation I want to construct here is less restrictive in its implications than these latter notions, but more restrictive than the metaphorical notion. However, to specify the meaning in the way the objection is put (What kind of entities would be included?) would be again to take the ontological, properties approach, which focuses on what *kind* of entities would be included. My claim starts from *relations* between entities—relations of dependency and co-operation. Thus, I have suggested we change the question. The same entity, e.g. a certain robot or a certain animal, can engage in co-operation with humans in one situation or role, and find itself outside relations of (strong) dependence and outside the co-operative scheme in other situations or roles. What matters for inclusion into the sphere of morality and justice is (the extent to which) humans and non-humans depend on one another and co-operate, that is, do things together. For instance, pets and humans live in a relation of dependency and often co-operate, whereas a wild animal of the same species (hence, having roughly the same ontological properties) would not usually qualify for engaging in such a relation, and could therefore not be drawn into the sphere of justice *on contractarian grounds*. (My argument does not exclude the possibility that certain non-humans could be drawn into the moral sphere, including that of justice, for other reasons.)

Thus I do not claim that we should include all non-humans in the moral sphere or the sphere of justice at all times and in all situations. If there is very little or no dependence and if the notion of cooperation proposed above cannot be applied at all, then the particular non-human cannot be considered as part of the sphere of justice at that time and in that situation. Moral consideration is and should be a matter of degree. And as relations and activities can change, the boundaries of the moral sphere (more generally) and the sphere of justice (in particular) are also apt to change.

Note that I use the notions of dependency and co-operation rather than benefit as criteria to determine inclusion into the sphere of justice. Dependence and

co-operation are not necessarily beneficial to all parties to the same degree. After all, in our attitude towards humans we do not require that all our dealings with them are always beneficial in order to take them into moral consideration. On the contrary, the liberal tradition has it that exactly because someone is able to harm me, there is a moral question with regard to our relation. And if, within a cooperative scheme, a human does something that harms me or does not contribute to the common aim, then we can have a discussion about the moral significance of what happened and about the consequences (e.g. some kind of punishment or reconciliation). Whatever the outcome of that discussion, we regard the human other as part of our moral and social world. Likewise, if we consider a particular non-human as part of a co-operative scheme, then we have some conceptual resources available for dealing with potential moral problems. And benefit is not necessarily the best criterion to determine the extent to which we owe something to non-humans (or humans for that matter).

A similar response can be made to the possible objection that my approach would imply that non-humans have only extrinsic value. Let me offer three replies. First, this objection assumes that co-operation is defined as an activity that is done for the sake of humans (for their benefit) only, but I have just argued that this assumption is false. If two humans co-operate, we do not infer that therefore they only have extrinsic value or that there is necessarily only benefit on one side. Second, the question of whether or not an entity has intrinsic value is independent of that entity's relations of dependency and co-operation. Third, I can make a stronger claim as well: to ask the question of value is a typical ontological, properties-based approach to the problem of how to relate to non-humans (and humans). The language of value is a language of property: an entity 'has' moral value or not. Intrinsic value is a property, a moral property in this case. The approach proposed in this paper offers an alternative moral framework.

Note also that some other moral vocabularies may be compatible with this alternative. One could think of relations of care between humans and non-humans that could develop in addition to or next to co-operation. Without have to embrace the argument that compassion or sympathy is due *rather* than 'justice', it is possible to give compassion or sympathy a role within a social-philosophical approach. Perhaps they are not only motivational elements that can possibly 'accompany' social relations with humans and non-humans, but can also be constitutive of certain kinds of relations and of their moral significance. Relations between humans and non-humans can take on a 'warmer' dimension. However, the approach proposed here conceives of social relations between humans and non-humans in terms of co-operation; that is, inspired by its contractarian roots, it views co-operation without sympathy or care as sufficient for the social life (of humans and of humans and non-humans). But nothing said here excludes the possibility that other kinds of bonds between humans and non-humans develop and flourish. A contractarian approach to the social proposed here is not against community, relations of friendship, etc. It only claims, against communitarians, that the constitution of the social does not depend on it. This makes possible the argument I propose in this paper. If the social depended on sharing values and on having warm and intimate relationships, it would not be possible to conceptually draw non-humans into the moral and political

sphere. It would be hard to make sense of the (moral dimension of) many relations of dependency and co-operation we have with them. In fact, it would be hard to make sense of that dimension with regard to many human/human relations as well. For instance, such a theory (e.g. a communitarian theory of justice) could not cope with the intuition that issues of (distributive) justice can arise at all between people that share little and do not consider themselves as part of the same community.

But is the language of rights compatible as well? Should we discuss ‘animals rights’ or ‘robot rights’? While I do not see why we must reject the language of rights altogether, in the light of my main argument I believe it is better to use the language of justice for the following reason. Rights tend to be seen as attached to a certain entity (e.g. the notion of inalienable rights), whereas the notion of (distributive) justice I proposed depends on relations (of dependency) and activities (of co-operation): it depends on what an entity does with other entities, rather than on what that entity is. Thus, within the approach proposed here the language of rights could only be used if those rights are justified on the basis of social relations, not (individual, ontological) properties. But in principle there is no objection to using other moral concepts, as long as they do not rely on the ontological approach.

Finally, let me emphasize again one of the implications of the contractarian character of my approach: it rests on a contractarian intuition about the moral significance of co-operation. Someone who rejects contractarianism and who does not share this assumption does not start from the same premises and will not accept my conclusion for contractarian reasons or might accept it for other reasons. For instance, one might advocate animal rights or robot rights based on an ontological argument concerning properties. I have only shown that we can argue in favour of drawing non-humans into the sphere of justice by taking a social-philosophical approach and by relying on an interpretation of contractarianism that manages to bypass the problems of the usual ontological approach.

Towards Ecological and Technological (Difference) Principles of Justice

If these are good arguments for including non-humans in the sphere of justice, how can this approach be applied to practical problems, and which *principle* of justice should regulate the distribution of *what to whom or what*? To end this paper, let me briefly explore these issues.

Consider the following practical question: Is eating animals for food morally wrong? This is a hard ethical question. But is it within the scope of my argument to decide this? My claim is that some non-humans in certain cases and at certain times should be included in the sphere of moral consideration and justice. My argument is about the boundaries of moral consideration and justice. And just as the arguments concerning properties do not themselves define the content of the precise duties and arrangements that follow from this definition of the sphere of morality and justice, I cannot say more within the scope of my argument. I can only say that if our relations with animals we use for food satisfy the criteria of mutual dependence and cooperation, they fall within the sphere of morality and justice. I believe this is indeed the case: we depend on them for food and they depend on us for their lives.

We co-operate towards the goal of sustaining human lives and—to some extent—their animals lives. However, my account does not itself specify what follows from that, for instance which principle of justice should apply.

Let me nevertheless offer some suggestions with regard to the question concerning principles of distributive justice.

With regard to non-humans, it is tempting to go for a sufficientarian principle, for example that (some) animals be treated in such a way that a decent level of their well-being is guaranteed, that their physical integrity is not violated, or perhaps just that they be allowed to live. Now to demand such a principle would require making measurements of the needs of other species and entities, which may be hard in some cases and which draws us towards an ontological approach and its associated problems. But more importantly, such a principle fails to do justice to the relational aspect that is so prominent in Rawls's theory of justice and that I have highlighted in contractarianism. As in 'strictly human' society (if such a thing exists), it is often, but not always, the case that the improvement of the life of the already advantaged implies a worsening of the position of the least advantaged. Consider humans that flourish by using animals for food or viruses that live in human bodies. Their advantage means a disadvantage for others: farm animals that live miserable lives or humans who become ill. A kind of difference principle—an ecological and perhaps in the future a technological difference principle—would demand that a better situation for the advantaged would not be at the expense of the disadvantaged. It could even ask to maximise the position of the disadvantaged, as Rawls's difference principle requires (Rawls 1971, p. 75–80). In so far as all parties involved here are engaged in co-operation, as in the 'strictly human' society, it makes sense to at least consider or *discuss* such a principle. Surely, given the complexity of our world the difference principle is hard to apply: how shall we determine the position of the advantaged and disadvantaged? And it may be difficult in practice to define which entities are the most advantaged and which entities are the worst off in the first place. However, such definition problems also occur in cases concerning humans only. For instance, what does relative poverty mean? Who is the worst off in our Western societies? If one takes a contractarian approach and is sympathetic to the difference principle in the human world, therefore, this objection would not count as a good argument against applying the same principle to relations between humans and non-humans.

However, even if we face these difficulties, we need not remain speechless. As in the human world, we are likely to find at least some very clear cases of 'injustice' that allow for the application of an ecological and/or technological difference principle. For instance, if we breed animals for (our, human) consumption *and* treat them very badly in the course of that process, then these cases (1) fall within the scope of problems of justice (as argued above) and (2) would warrant the application of a difference principle since increases in the advantages humans get from the co-operation (we are clearly highly dependent on them for sustaining our current consumption habits) do nothing to maximize the position of these animals, which can be considered the 'worst-off', the most disadvantaged in human/animal society. I conclude that whatever other issues of justice (and other moral problems) there may be in these cases, better treatment of such animals can be justified by reference to the difference principle as a principle of distributive justice.

Conclusion: Imagining Co-Operation in the Hybrid Society

In this paper I have suggested one way of drawing non-humans into the sphere of moral consideration. I argued that to the extent that they can be considered as part of a co-operative scheme between humans and non-humans, non-humans must be included in the sphere of justice. Although this proposal may at first appear as counter-intuitive or radical as arguments for animal rights, Latourian or deep ecology approaches, it has the advantage that it by-passes some hard ontological discussions about the nature of humans and non-humans such as animals and robots (it only requires that the entity have a capacity for co-operation with other entities), takes seriously the (quasi)social dimension of human/non-human relations, and re-interprets an already very influential theory concerning human social relations: contractarianism.

Of course hard questions remain, for instance with regard to applying this theory in practice. We may be reluctant to include some animals, robots, or (other) artificially intelligent agents in the sphere of justice. However, if we try to shift our focus from 'what these entities really are' to an investigation of the precise relations of dependence and co-operation between *particular* animals and *particular* humans (if there are any), and if we can define advantaged and disadvantaged parties in terms of quasi-social co-operation, we may arrive at a conclusion that there is indeed an issue of justice in these particular situations. Then we should engage in a Rawlsian-style thought experiment in order to derive a principle of justice with regard to that particular co-operative scheme. This requires an exercise of imagination that may seem very exotic when viewed from an ontological perspective but which is entirely appropriate if and *in so far* a co-operative scheme is in place: we imagine being in that least advantaged position within a particular co-operative scheme, which may or may not be a position occupied by a (non-)human.

I grant that making this turn this is not easy, especially when artificial non-humans are concerned; in that domain we are used to ontological thinking rather than sociological thinking. We still think too much in terms of properties rather than social relations; we generally do not perceive and acknowledge the co-operative links between humans and non-humans. (As indeed many of us still think too much in terms of properties when it comes to humans.) Thus, we need to perceive and imagine (1) our relations of dependency and co-operation with non-humans (and humans) and (2) the social position of the worst-off (human or non-human) parties to that co-operation. The latter is the kind of imagination required by the original position heuristic.

Exercising this kind of imagination is very different from trying to imagine *being* a different kind of entity. Such exercise of ontological imagination is not required by the proposed approach. And in many cases, it is impossible or at least very difficult, as Nagel has argued (Nagel 1974). Luckily for the bat, I do not need to know how it is to *be* a bat in order to morally evaluate human relation(s) with that bat in a specific context. And luckily for those humans who have been liberated from slavery, those who argued for their liberation have shifted their moral attention from properties such as skin colour to what it must be to occupy the social position

of a slave. Both less-advantaged non-humans and less-advantaged humans could benefit and have benefited from this shift in approach.

So far, we have not always been very good at developing and using such a social imagination. But in so far as we are contractarian and if the arguments advanced in this paper are convincing, we better start training it.

Acknowledgments Thanks to the anonymous reviewers for their comments; they were of great help to me in improving the quality of my arguments. I also wish to thank Julie Bytheway for her advice on matters of language.

References

- Baxter, Brian. 2005. *A theory of ecological justice*. Abingdon/New York: Routledge.
- Bell, Derek. 2002. How can political liberals be environmentalists? *Political Studies* 50: 703–724.
- Bell, Derek. 2003. Political liberalism and ecological justice. Paper presented at the ECPR General Conference, Marburg, 19 September 2003.
- Dobson, Andrew. 1998. *Justice and the environment: Conceptions of environmental sustainability and theories of distributive justice*. Oxford: Oxford University Press.
- Dobson, Andrew. 2000. *Green political thought*. London: Routledge.
- Floridi, Luciano., and J.W. Sanders. 2004. On the morality of artificial agents. *Minds and Machine* 14: 349–379.
- Garner, Robert. 2003. Animals, politics, and justice: Rawlsian liberalism and the plight of non-humans. *Environmental Politics* 12: 3–22.
- Latour, Bruno. 1993. *We have never been modern*. Cambridge, MA: Harvard University Press. Trans. C. Porter.
- Latour, Bruno. 2004. *Politics of nature: How to bring the sciences into democracy*. Cambridge, MA: Harvard University Press. Trans. C. Porter.
- Latour, Bruno. 2005. *Reassembling the social: An introduction to actor-network-theory*. Oxford: Oxford University Press.
- Low, Nicholas., and Brendan. Gleeson. 1998. *Justice, society, and nature: An exploration of political ecology*. London: Routledge.
- Næss, Arne. 1973. The shallow and the deep, long-range ecology movement. *Inquiry* 16: 95–100.
- Nagel, Thomas. 1974. What is it like to be a bat? *The Philosophical Review* 83: 435–450.
- Nussbaum, Martha.C. 2005. *Frontiers of justice: Disability, nationality, species membership*. Cambridge, MA: Harvard University Press.
- Nussbaum, Martha.C. 2006. Reply to Richardson. *Journal of Ethics* 10(4): 490–506.
- Rawls, John. 1971. *A theory of justice*. Cambridge, MA: Harvard University Press.
- Rawls, John. 1993. *Political liberalism*. New York: Columbia University Press.
- Rawls, John. 1999. *The law of peoples*. Cambridge, MA: Harvard University Press.
- Rawls, John. 2001. *Justice as fairness*. Cambridge, Mass: Harvard University Press.
- Regan, Tom. 1983. *The case for animal rights*. Berkeley, CA: University of California Press.
- Rowlands, Mark. 1997. Contractarianism and animal rights. *Journal of Applied Philosophy* 14: 235–247.
- Singer, Peter. 1975. *Animal liberation: A new ethics for our treatment of animals*. New York: Random House.