

# Liberty, Games and Contracts

Jan Narveson and the Defence of Libertarianism

*Edited by*

MALCOLM MURRAY

*University of Prince Edward Island, Canada*

© Malcolm Murray 2007

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior permission of the publisher.

Malcolm Murray has asserted his moral right under the Copyright, Designs and Patents Act, 1988, to be identified as the editor of this work.

Published by  
Ashgate Publishing Limited  
Gower House  
Croft Road  
Aldershot  
Hampshire GU11 3HR  
England

Ashgate Publishing Company  
Suite 420  
101 Cherry Street  
Burlington, VT 05401-4405  
USA

Ashgate website: <http://www.ashgate.com>

#### **British Library Cataloguing in Publication Data**

Liberty, games and contracts : Narveson and the defence of libertarianism

1. Narveson, Jan, 1936– 2. Libertarianism 3. Contractarianism (Ethics)

I. Murray, Malcolm

320.5'12

#### **Library of Congress Cataloging-in-Publication Data**

Liberty, games, and contracts : Jan Narveson and the defence of libertarianism / edited by Malcolm Murray.

p. cm.

Includes bibliographical references and index.

ISBN 0-7546-5681-0 (hardcover : alk. paper)

1. Libertarianism. 2. Social contract. 3. Contractarianism (Ethics) 4. Narveson, Jan, 1936– —Political and social views. I. Murray, R. Malcolm (Robert Malcolm), 1959– II. Narveson, Jan, 1936–

JC585.L4245 2006

320.51'2—dc22

2006011547

ISBN 978-0-7546-5681-4

Typeset by IML Typographers, Birkenhead, Merseyside and  
printed and bound in Great Britain by MPG Books Ltd, Bodmin, Cornwall.

ASHGATE

## Chapter 2

# Contracting Justice<sup>1</sup>

*John T. Sanders*

In *The Libertarian Idea*, Jan Narveson explains his interpretation of contract theory this way:

The general idea of this theory is that the principles of morality are (or should be) those principles for directing everyone's conduct which it is reasonable for everyone to accept. They are the rules that *everyone* has good reason for wanting everyone to act on, and thus to internalize in himself or herself, and thus to reinforce in the case of everyone.<sup>2</sup>

It is plain, here, that Narveson believes that social contract is to provide justification – a foundation, in fact<sup>3</sup> – for 'the principles of morality'. The burden assumed in this chapter is to examine how far Narveson has succeeded in making this foundational claim plausible.

### **Contract As the Foundation of Morality**

Narveson contends, as noted above, that the idea of social contract provides foundations for all of morality. What morality *is*, though, while understood in general terms by all, is notoriously difficult to pin down neatly. As a start, one might suggest that the realm of morality is simply the realm of what one should or shouldn't do. Then the question of 'foundations' would be about the reason or reasons underwriting shoulds and shouldn'ts.

Before ever beginning the foundational project, though, one vital reservation springs to mind: insofar as one seeks justificatory foundations for morality, the moral realm cannot helpfully be thought to include *all* shoulds and shouldn'ts. If you want to get to the post office, you should drive back south on the road to Lahaina and turn left at the traffic lights at Wahikuli Beach. If you want a really good margarita, you should use Cointreau instead of Triple-Sec. Surely there is nothing particularly moral about *these* shoulds except in the broadest possible sense: the sense in which 'morality' may be held to encompass any and all purposive activity. Immanuel Kant is famous in this philosophical domain for arguing that the telling feature of *moral* shoulds is their 'categorical' character: according to Kant, there is nothing iffy about them at all. There are some things you should just do (and some you shouldn't do) regardless of what you may want to

achieve. It is not that Kant held that this categorical character attaches *magically* to some shoulds. Rather, he argued that the properly moral claims have a quasi-logical feature that sets them apart: there is supposed to be a sense – elaborated by Kant – in which it would simply be *irrational* to violate the demands made by the properly moral shoulds.

Narveson distinguishes the moral shoulds from the others in a way that appears initially to be quite different. In the first place, Narveson wants to step away from the idea that moral reasons to act are a different *kind* of reason from all others:

When morality makes demands upon us, the question why we should go along with them, what reason there is to impose them, is *always* apt. There are no ‘bedrock’ moral reasons: that is an illusion. To say that some act is called for by morality is to say that there are good reasons for insisting that people do that kind of thing. Those reasons are not ‘moral’; they are, simply, reasons.<sup>4</sup>

Further, the relevant questions are ones of social policy. While acknowledging that morality in general may be held to embrace what he calls ‘personal’ morality,<sup>5</sup> Narveson urges that the ‘social’ sense of the term is not only the more usual sense,<sup>6</sup> but the one that gets us to the core of the matter: ‘moral requirements can, and often do, “override” individual inclinations to act otherwise.’<sup>7</sup> This, indeed, is their primary point. But this is not – as it was with Kant, because moral requirements represent a different, categorical (quasi-logical) sort of demand – to be distinguished from *hypothetical* imperatives that depend on a person’s interests and inclinations. Rather, Narveson suggests that the moral imperatives are as dependent upon people’s interests as any other. The key similarity between Narveson and Kant, at least in this first comparison, is that both point to the fact that moral requirements can and do override individual inclinations to act otherwise. The big difference between them appears to be that Kant thinks this marks them as belonging to a substantially different *kind* of requirement, while Narveson disagrees. Narveson’s diagnosis:

Kant supposes that people do not simply act in pursuit of their interests. They have, as it were, rationally split personalities. On the one hand, they are concerned to promote their interests. ... On the other, however, they are concerned to obey a transcendental law invoking some kind of universalizability requirement known as the Categorical Imperative. ... So the two aspects of reason can come into conflict, and when they do, the moral one ought to win. But as Kant depicts the situation, there are insoluble problems. For one thing, in his view ‘inclinations’ belonged in the phenomenal world and were subject to universal causation, and so the Laws of Nature decreed that people must do what their inclinations dictate – which is incompatible with their doing what morality dictates if it ever differs, as Kant thought it often (always?) would. So he is driven to a hopeless and unintelligible doctrine of contra-causal freedom.<sup>8</sup>

And we certainly wouldn’t want to be driven (or dragged) there.

At least at first glance, though, it would seem that things are not radically more coherent for Narveson. He has got people inclined, in their own interest, to do

things they are not inclined to do. More fairly, perhaps: when people act morally, according to Narveson, they do things that are really in their own interest even though they may seem at first glance not to be. While the reasons for moral behaviour are not some different ‘categorical’ kind of reason, they still override other interests and other considerations. A clue to the real difference between this overriding kind of reason and the kind that gets overridden appears to arise from the fact that the former reasons are of the ‘social’ kind rather than the ‘personal’ kind. The way to understand the importance of this difference and to see why ‘social interests’ of a certain sort might reasonably override many or most ‘personal’ interests, Narveson thinks, is to see that morality is based on a ‘kind of contract’.

It cannot be a real contract, after all. As Narveson points out, there is no reason to think that any real people have ever actually entered into an agreement of the envisioned kind. And even if there *were* an historical ‘social contract’, he clearly sees that such a thing could have no normative bearing on anyone who was not party to it.

In what is perhaps the clearest, fullest exposition of his case,<sup>9</sup> Narveson takes as his theme a matter raised by David Hume that appears to offer a serious challenge to the ‘contractarian account of morals’. As Hume pointed out in his *Inquiry Concerning the Principles of Morals*, ‘the observance of promises is itself one of the most considerable parts of justice; and we are not surely bound to keep our word because we have given our word to keep it.’<sup>10</sup> Some have taken points of this kind to undermine, rather severely, the contractarian programme. Narveson, like other contract theorists, does not think so.<sup>11</sup> But he does agree that some qualification is in order:

The contractarian holds that the fundamental principles of morals are in some way the objects of an ‘agreement’, which in turn is the outcome of what is in some sense a ‘bargain’. In trying to make clear and plausible this approach to morals, it is obviously of crucial importance to identify the sense(s) or way(s) in which these things are supposed to be so. For one thing is surely clear enough, and that is that the quotation marks are appropriate: the ‘agreements’ and ‘bargains’ of which the contractarian moral theorist speaks are not your ordinary examples of these types. The Grand Social Contract is, obviously, a model or idealisation of some kind.<sup>12</sup>

Further,

If we are to make social contract theory work, it is clearly essential to understand the ‘agreement’ in some other sense than that of a formal promise or literal contract, be it historical or hypothetical. ... [W]e would want the principle of keeping one’s contracts to be itself founded on The Social Contract.<sup>13</sup>

So the answer to Hume is: the Social Contract is not really a contract. Thus there is no circularity.

What *is* the ‘Grand Social Contract’, then? Following David Gauthier,<sup>14</sup> Narveson argues that

the 'Social Contract' is a 'contract' in the following sense: people discern that if they interact under the unrestricted motivation of their various separate interests, then they will do worse than could be done if instead they were to adopt certain restrictions or constraints on the pursuit of those interests. A main context in which such restraint is called for is that of social interaction. Here it is often the case that the unlimited pursuit of advantage by each leaves all worse off than if they had instead adopted restrictions on individual pursuit of utility. Constraining conduct promotes maximization, even though maximization is precisely what it constrains. The secret is in mutuality: each must shape her behaviour by refraining from certain erstwhile maximizing actions, *on condition that* the other does likewise. It is the necessity of this condition that makes the 'Social Contract' a *contract*. 'I shall do this, provided he does that'; 'I shall do that, provided she does this' – and that goes for both parties.<sup>15</sup>

That this requires no explicit multi-party agreement, no real negotiation and not even any communication between people, is not only clear from the explanation quoted above. It is reflected also in Narveson's portrayal of himself as Hobbesian as regards the very essence of morality. For Narveson, as for Hobbes, morality (in the more important 'non-personal' or 'social' sense, at least) does not even exist for solitary persons. It arises only in social circumstances. Morality gains relevance because of the sorry state of affairs that would prevail in its absence: no industry, no cooperative ventures, no Cubs' games – nothing that otherwise makes human life not only worth living, but possible in anything but the sparsest terms. The point is not supposed to be that there have been grim historical times out of which morality emerged for practical reasons; rather, the point is that morality serves human interests by making industry and all the rest possible. And the fact that no real explicit agreement is necessary is made yet clearer by frequent Narvesonian references to game theory. To put the matter briefly and without reference to the game theoretic details: persons in real life having to choose between making decisions consistent with moral dictates, on the one hand, and violating those dictates, on the other, have strong reason to choose morally, *if others do too*, because doing so advances their interests. Back in abbreviated game-theoretic language, this amounts to the observation that the dominant strategy in iterated prisoner's dilemma settings is tit-for-tat.<sup>16</sup> Such results, whether in game theory or moral theory, have nothing at all to do with real decisions in real situations. They derive instead from abstract consideration of relevantly constrained situations.

So there are good reasons to behave morally. The 'foundations' of morality are to be found in these reasons. These reasons make reference to the interests of persons who are called upon to behave morally, so they are not 'transcendental' in the way Narveson portrays Kant's as being. They may be understood to be modelled on 'agreement', even though no explicit agreement is contemplated, because they have a reciprocal character: I agree to behave morally if and only if I have reason to think you are disposed that way too. In game theoretic tit-for-tat, the idea is to start by cooperating in new situations, then continuing with cooperation with those who respond positively while harshly discontinuing cooperation with those who do not. No explicit verbal agreement is envisioned here either; 'agreement' is instead tacit and observable only in behaviour. So much for the

argument, then. But how does the contractarian approach stand in relation to its competitors?

### Contractarianism Versus the Alternatives

Narveson has regularly been moved, in his many writings on the subject, to defend contractarianism on the ground that none of its competitors can really hold a candle to it in doing what counts: showing how people might be motivated to constrain their behaviour along the lines prescribed by morality. But his specification of exactly what the competition in this enterprise *is* leaves his reader in some doubt about whether he is dealing with the *real* competition. Here is what he says, for example, in an essay entitled 'Contracting for Liberty':

Political and moral theories have been argued for on such bases as ideals of virtue, the will of God, self-evidence, and in one way or another, Nature. Why reject all these? Actually, as we will see, we needn't reject them all, quite: the contractarian view may be aligned with one kind of 'natural law'.<sup>17</sup>

Narveson then goes on, in this particular piece, to explain why these several candidate foundations will not work, and begins to advance positive arguments for 'the contractarian idea'. In *The Libertarian Idea*, he devotes an entire chapter to discussion of the inadequacies of 'intuitionism' in moral theory by way of preparation for his advocacy of contractarianism. But the real competition, one might think, would be standard 'foundational' approaches like utilitarianism, or Kantianism (for want of a better name). We have considered Narveson's objection to Kant's approach: the reason certain 'social' interests take precedence over many 'personal' ones is not that the former belong to a different transcendental or quasi-logical category; it is rather that they represent more important, more systematic interests. These 'social' interests involve improving the whole framework for pursuit of interest generally.

Narveson's objection to intuitionism, in the end, is fairly simple: where intuitions compete, they are unable to persuade. Since 'foundations' for morality are sought that all can agree upon, intuitionism is simply not up to the task.

As for utilitarianism, the discussion in *The Libertarian Idea* suggests that the reason for subordinating it to intuitionism as a potential foundational competitor to contractarianism is that utilitarianism is unwittingly (except in the more witting case of Henry Sidgwick, who Narveson says made the same point) dependent upon a central moral intuition.<sup>18</sup> If this were true, it would perhaps legitimate moving directly to a consideration of intuitionism in all its forms rather than worrying about particular versions like utilitarianism.

For present purposes, let us allow the arguments against Kantianism and intuitionism to pass. More telling are Narveson's arguments against utilitarianism, especially since he used to defend utilitarianism himself before he became a contractarian.<sup>19</sup>

Utilitarianism in its most general form says that right and wrong depend, in the end, on what promotes the general good. Perhaps we should understand 'the general good' in terms of 'happiness', but perhaps not. Perhaps we can actually add and subtract utils (or some other measure) of happiness, but perhaps not. Maybe we can make cardinal comparisons of different persons' degrees of happiness, but then maybe not. Perhaps we should each be making a judgment about what promotes the common good in our every decision about what to do. Or perhaps, on the other hand, we (or society, in our behalf) should (on utilitarian grounds) instead guide our lives in accordance with rules that themselves tend to promote the common good in the long run. These are matters that utilitarians argue over, and it is clear in the words Narveson occasionally still devotes to the subject that he has strong views on most of these issues.<sup>20</sup>

But how does utilitarianism, in general terms, fare against contractarianism, as the latter is cast by Narveson? Narveson's case against utilitarianism, usefully summed up in *The Libertarian Idea*, has several key features. First, he decries the conceptual confusions we are left with as we contemplate the many questions left unresolved by utilitarianism 'in general' (for example, the uncertainties reviewed a paragraph ago). He implies that these difficulties, if not insuperable in principle, are at least sufficiently puzzling as to encourage us to look elsewhere for clarity on foundational questions. Second, he argues that the utilitarian claim that the 'general good' trumps 'individual good' in cases of conflict can finally rely only on either definition or intuition. He concedes that not all utilitarians would agree with this claim of dependency, but appears to be confident that he is right and they are wrong. If he is right, then utilitarianism is impotent against those persuaded by different definitions or intuitions. Third, while acknowledging the broad attractiveness of the utilitarian principle as a foundational principle, Narveson urges that this attractiveness may not be due to morality's actually being *based* on the promotion of general well-being. Instead, while it might be true that 'justice' and 'happiness' are likely to be positively correlated in any society, this might just as well be because a significant component of happiness resides in the perception that one's society is just. That is, happiness may be (partially) based on justice, rather than the other way around. If this were true, then utilitarianism would fail in its foundational ambitions. Finally, as what is nominally a side remark (although it is placed where one might normally expect the conclusion of the argument to appear), Narveson adds that

Whenever we are in real-world situations closely resembling a 'veil of ignorance' in some respects, then a utilitarian solution is likely to be a good one. Suppose that some system must be devised for catering to the comparable interests of a large number of people, and that those people are essentially randomly situated with respect to the sources of the benefits. It will then be true that each person maximizes her expected utility by accepting a utilitarian format. ... But this works because it is a limiting case, the case where individual and general utility coincide. Many cases certainly don't seem to be like that, and when they aren't, we need a different principle.<sup>21</sup>

If this is right, then utilitarianism, while useful in rare situations in the real world, fails us in many others.

We are faced, then, with four arguments concerning the inadequacy of utilitarianism and (presumably in the background) the superiority of contractarianism: (A) Utilitarianism's conceptual problems are too difficult; (B) Utilitarianism is dependent upon definitions or intuitions about morality that may not be shared by all; (C) Utilitarianism's plausibility may be due to a dependency of happiness on justice, rather than a dependency of justice on happiness; (D) Utilitarianism's applicability may be too narrow.

It is not the objective of this chapter to defend utilitarianism or any other foundational competitor to contractarianism. Instead, the focus here is on clarifying Narveson's reasons for preferring the latter. With that in mind, it is worth remembering from the preceding section that the Social Contract alluded to in contractarianism as providing foundations for morality is not really a contract. Indeed, the morality of keeping one's contracts is supposed, by Narveson, to be non-tautologically based on the Social Contract. Hume's indirect allegation of circularity is refuted by the suggestion that the Social Contract is really no contract at all. Let us return to a consideration of what it really is, as we think about Narveson's several objections to utilitarianism.

As portrayed by Narveson, the Social Contract is an 'agreement' only insofar as it is to be identified with the mutual restraint of several interacting (or potentially interacting) persons, each of whom continues to act on behalf of their own self-interest. There is really no agreement, as noted earlier; there is instead an inclination, on the part of each of several individuals, to constrain behaviour in certain ways ('moral' ways) provided others do the same. This inclination is owed to restraint being in each person's perceived long-term best interest (as long, anyway, as the condition of *mutual* constraint continues to hold).

But wait: need these envisioned people be doing anything even so grand as perceiving that restraint is in their long-term best interest? Such mutual constraint among humans is comparable to similar restraints within animal groups. Are wolves, in cooperating and submitting in certain ways in their packs, doing so *because* it is in the long-term best interest of each? How about bees? Ants?<sup>22</sup> How about human sexual behaviour? Is *perceived* long-term best interest even relevant?

Well, it is and it isn't. It seems implausible to say that non-human animals behave the way they do because doing so is *perceived by them to be* in their long-term best interest. In the case of non-human animals, it seems highly unlikely that they do much thinking at all about long-term interest. Yet the behavioural tendencies to restrain behaviour are present quite as much among non-human animals as they are among humans. In the case of at least much of human behaviour, one could surely say the same: behaviour is morally constrained in social situations even though not much thought is directed toward long-term best interest. It is just not plausible to contend that people, any more than non-human animals, restrain their behaviour in conformity with social rules because they have thought the matter through and have concluded that such restraint is in their long-term best interest.

So if that is what we are to understand by the claim that 'people behave morally because it's in their interest', then the claim is quite dubious. But it is *not* so strange, on the other hand, to observe that it *is in fact* in people's and non-human animals' long-term interest to restrain themselves in these ways, or to suggest that the practices and behaviours (constraints and all) are what they are because they do indeed serve the long-term best interests of these several creatures, regardless of whether they know it (or perceive it) or not.<sup>23</sup>

First of all, then, the bare fact of mutual restraint among interacting beings, people included, implies not only nothing about agreements made between them, it implies nothing at all about perceptions or reasoning processes they may have experienced. This means, in turn, that mutual restraint implies nothing at all about perceptions of interests, in particular, whether individual or social, whether long-term or short.

This consideration, all by itself, may not be regarded as causing serious damage to contractarianism, though. For humans, anyway, can and do raise questions about why they should behave in this way rather than that. Contractarianism offers an answer: it is because behaving this way is in fact in your interest. That is how you explain to people why they should be moral, according to Narveson. The explanation might be a broad evolutionary one: these cooperative arrangements, where they have been adopted, have contributed to the well-being (at least considered in broad terms) of the people who have adopted them. Or the explanation can be cast in game-theoretic terms: under the general circumstances confronted by people in typical social circumstances, tit-for-tat dominates. You should therefore choose that strategy too. All this is fine. But it really comes no closer to explaining these things to those who see temporary advantage (at least) in neglecting the rules. Sometimes personal interest and social interest diverge; sometimes it pays to gamble; sometimes non-cooperation, especially when dealing with people you will never see again, pays off.

Contractarianism, it seems, risks falling prey to a particularly interesting version of the naturalistic fallacy. Contractarian narratives, like evolutionary explanations, can be extraordinarily plausible in making sense of why people do what they do: behaving in those ways makes sense because *if* we were to give thought to what is in our long-term best interest, we would so behave. It does not matter at all what thought processes or perceptions real people have as they make their way through the world, cooperating here and failing to cooperate there, behaving morally here and cheating there. What matters is what we say to those who actually ask 'Why cooperate?' Evolutionary theory, social contract theory and game theory have something to say about this. What they have to say is surely not argumentatively decisive, mind you; it is more like general advice, at least in those situations where what they have to say is more than just bare description. The whole justificatory game differs from description precisely in that justification is relevant when someone asks 'But why should I?' while description offers accounts and explanations of how things are or how they have come to be. But in that light, consider what contractarianism would allow us to do: just as is the case with other theories arising mostly from descriptive explanations of why things are the way

they are, they can be converted to justificatory application in surprising ways. Precisely as plausibly as we can use contract theory to explain to a hypothetical recalcitrant neighbour (or to ourselves) why they (we) should be moral, we could use it also to explain to the hypothetical recalcitrant ant why it should not wander off and try to fend for itself. Such a plan just wouldn't be reasonable, given typically survival requirements. Cooperation is surely the best policy. It would be much wiser for each ant to align its behaviour with that of its fellows, provided those other ants do the same. But does the fact that we can apply considerations about ant interactions in this way give us any reason at all to invoke the conceptual trappings of contract? If not, what is there in the corresponding tale about people that gives us any more reason to do so?

Perhaps the contractarian would want to point out that the person can, while the ant cannot, consider reasoning of this kind and then go on to bring behaviour into line with it. But that should not really matter. If the foundation of morality is to be found in what it would be reasonable to agree to, rather than what is *actually* agreed upon, then it makes no difference that ants and people differ in their respective capacities to follow the dictates of reason. Perhaps ants are hard-wired to do what is reasonable/moral (at least most of the time), while humans are not. But if morality really is based on what is reasonable, then ants and humans are both subject to its rules.

There should at least be something puzzling, though, about the fact that this kind of argument works for ants as much as it does for people, and this puzzlement takes us back to Narveson's arguments against utilitarianism. Let us now see, after the foregoing musings, how contractarianism stands up in the competition.

#### A. Utilitarianism's Conceptual Problems Are Too Difficult

It is no small conceptual problem that contractarianism depends on no contract; that the 'agreement' envisioned among people involves no agreement; that the perception that cooperation is in their interest need not be any part of the envisioned participants' reckoning; that cooperation may not, indeed, serve the particular perceived interests of particular persons in particular situations, and that cooperation may, instead, be quite automatic (rather than considered in the light of self-interest) in most situations where it is actually observed. The confusion here stems in large part from moving between a conceptual situation in which one is trying to understand why certain behaviour patterns emerge in populations, on the one hand, and a wildly different conceptual situation in which one is trying to explain to people why they should be moral. It would appear that contractarianism may be able to offer interesting advice, but such advice is only good for people whose perceived interests are normal or 'rational', which brings us to the next point.

#### B. Utilitarianism Is Dependent upon Definitions or Intuitions about Morality That May Not Be Shared by All

If contractarianism is to work theoretically in providing 'foundations' for morals, it

has to be talking the right language. If one assumes that practical rationality consists in choosing effective means for one's ends, and if one further assumes that the ends of a particular individual are the 'reasonable' ends that any of us 'should' (naturally?) have, one may then be able to give practical advice about how one should behave.<sup>24</sup> You are in an iterated prisoner's dilemma situation. You should adopt a policy of tit-for-tat, because that will give you a better chance of positive outcome than any other policy. Positive outcome, it is interesting to note, is in such settings presumed to be transitive across individuals. You cannot work out this game-theoretic stuff if the units of reward are not commensurable. The nice thing about game theory is that it does not matter at all what the units are units of; but whatever they are, they do have to be the same from one person to another. Thus arise game-theoretic parallels of some of the 'conceptual' problems of interpersonal comparison that supposedly condemn utilitarianism. But never mind that. Does this sort of thing – practical advice in obtaining what you want – generate morality?

I suspect that depends on your 'intuitions' about morality. One disturbing thing about Narveson's own contractarianism, for example, has to do with the treatment of animals. Narveson is committed to the idea that mistreatment of animals is wrong, all right; but it is not *directly* wrong, so to speak. According to Narveson, whatever wrongness there is in such behaviour has to do with our interactions and quasi-contractual negotiations with our fellow humans, our own several tastes and so forth.<sup>25</sup> Why? Because (simply put) morality is based on contract, and people cannot make contracts with animals. That this should be the way to arrive at moral conclusions seems wildly counter-intuitive.<sup>26</sup> I am no radical on this particular subject. Indeed, I am not at all sure what is right and wrong concerning the treatment of animals (apart from my abhorrence of plainly cruel treatment, which I appear to share not only with animal-rights' activists but with contractarians such as Narveson). But the idea that what is wrong regarding the treatment of animals should end up being fully dependent on what you and I perceive to be best for ourselves just seems mistaken. There is no problem with agreeing that actual moral and legal codes drawn up by or evolved by real communities do arise mostly through agreements made by humans (although the role of real 'agreement' in such matters needs to be scrutinized case by case). But the question whether such moral and legal codes actually capture what is right must remain open. The evidence of this is that they so often *don't* capture what they putatively aim at, and this is especially clear in cases where the content of the codes is plainly biased in favour of the interests of a minority. That majorities may in other cases be the beneficiaries of more democratically formulated codes does not alter the fact that they can be wrong too.

### C. Utilitarianism's Plausibility May Be Due to a Dependency of Happiness on Justice, Rather than a Dependency of Justice on Happiness

Contractarianism's plausibility, as noted above, may be due to a misapplication to contexts of justification of theoretical equipment that is appropriate only in contexts

of description. Further: while it may well be that it is reasonable to agree to be moral (because it is in your best long-term interest), it is no clearer in this case than in the case of utilitarianism that the dependency is as contractarianism needs it to be. For there may be (and likely is) something in behaving morally that renders it rational to adhere to, rather than the other way around. That is, it may be that far from arising from 'Social Contract', morality's 'foundations' are entirely independent of 'agreement'. Instead, 'agreement' is rational because of the morality (based on other foundations) of the behaviour.

### D. Utilitarianism's Applicability May Be Too Narrow

Narveson's remark to this effect was directed particularly at Rawls-like invocations of 'veils of ignorance' and their constraints. This seems an odd complaint for Narveson to make, given that his Social Contract involves no contract, no agreement, no consideration of particular humans in particular circumstances,<sup>27</sup> but instead relies on considerations that seem no less abstract than those invoked by Rawls. It is hard to see why anyone should prefer Narveson's framework, in principle, to Rawls's. Indeed, it seems that Narveson might just as well have preserved the veil-of-ignorance contraption while lobbying for its repositioning. We need to take into consideration, according to Narveson, some of the things that Rawls argued should be obscured. So why not just move back the veil a bit? We are still in pretty abstract territory, after all, when compared to real human beings making real decisions and making (or failing to make) real agreements and contracts.

More than that, though: as indicated above, Narvesonian contractarianism allows us to draw moral conclusions only about situations that involve, in one way or another, human interactions with other humans. The realm of morality and justice is thus narrowed to a particular realm of behaviour – the realm of human interaction – and can be extended to putatively moral obligations regarding non-human aspects of the world only indirectly, if at all, by way of reduction to human quasi-agreements with other humans. This rendering of contract theory, therefore, in providing a contractual justification for moral claims, does so by contracting the realm of justice to a base much narrower than the realm of common moral judgment that is to be explained. While a reduction of some kind or other might well be acceptable in principle, this one, as indicated above, seems unwarranted. The name 'contract theory', while intending to focus attention on quasi-agreements, may thus inadvertently highlight one of the theory's more problematic features: it contracts morality and justice to an inappropriately narrow base.

### Conclusion

Whether one accepts contractarianism of this kind or not, it must be conceded in closing that there is much that is attractive about the line of thought that supports it. Even if it is unreasonable to think of it as involving a 'contract' or 'agreement' (that

is, even scare-quoted ones), it must truly be the case that morality, in broad terms, somehow serves the interests of the beings in whose lives it plays such a large role. The particular naturalistic considerations invoked by Narveson (or invocable in principle) in defence of this version of contractarianism – evolutionary considerations suggesting that certain constraints exhibited by people (or by non-human animals) make sense in terms of survival, or game-theoretic considerations that suggest that choosers in certain decision situations would be left worse off if they acted in terms of a narrowly construed self-interest than if they moderated their behaviour socially – are suggestive. They are correctly regarded as promising areas for further inquiry into possible ‘naturalistic’ foundations for morality. But they do not justify the language of contract.<sup>28</sup> This is because they leave us only with the fairly secure conclusion that there are good reasons for the kinds of restraint we observe in real populations of interacting animals, whether human or non-human. It is of course possible to contend that ‘Social Contract’ provides the basis of non-human behaviour as much as it provides foundations for morality, but this stretches the notion of ‘contract’ much too thin. And that is what we would have to be willing to do if we were to accept Narveson’s version of contractarianism for humans.

#### Notes

- 1 I am grateful to Jesús Aguilar, Timothy Engström, Malcolm Murray, Wade Robison, Evan Selinger, David Suits, Katie Terezakis and Victoria Varga for their encouragement and for their helpful suggestions concerning more primitive drafts of this chapter.
- 2 Jan Narveson, *The Libertarian Idea* (Philadelphia, 1988), 131.
- 3 The quoted passage appears in a section in Narveson’s book entitled ‘The Idea of the “Contract” Approach to Foundations’, 131–4.
- 4 *Ibid.*, 126.
- 5 *Ibid.*, 123.
- 6 *Ibid.*, 125.
- 7 *Ibid.*, 124.
- 8 *Ibid.*, 127.
- 9 Jan Narveson, ‘The Agreement to Keep Our Agreements: Hume, Prichard, and Searle’, *Philosophical Papers*, XXIII/2 (1994).
- 10 David Hume, *Inquiry Concerning the Principles of Morals* (Indianapolis, 1957), Appendix III, 122. See also his *Treatise of Human Nature*, L. A. Selby-Bigge (ed.) (Oxford, 1978), bk III, pt II, sect. II, 484–501.
- 11 Nor does Hume, really. Like Narveson, Hume may be understood as a ‘contractarian-of-sorts’. Unlike Narveson, though, Hume would probably not have cared to insist on the ‘rationality’ of contract as a justifying factor. It seems likely that his considered position would be much more naturalistic, along the lines of his well-known analysis of ‘inductive reasoning’. We do indeed evolve moral codes as a function of their success in interpersonal activity, just as we develop expectations of the future based on our experience of the past. There is nothing particularly ‘rational’ about this, though. It just reflects the output of natural dispositions which we’ve come to have through means that have nothing to do with reason. See especially the passage about ‘Two men who pull on the oars of a boat’ in Hume, *Treatise of Human Nature*, 490. Hume’s discussion, in this passage, concerns the way in which social conventions – like ‘the rule concerning the stability of possession’ – arise in human societies. While much of this discussion can indeed be interpreted in terms of rational choice (and some of it admittedly seems hard to interpret in any other way), the emphasis on the slow, gradual process of growth of such conventions cannot be interpreted that way. See also, on this point, Brian Skyrms, ‘Game Theory, Rationality and Evolution of the Social Contract’, in Leonard D. Katz (ed.), *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives* (Bowling Green, OH, 2000), especially 272–3. Skyrms makes some especially interesting suggestions in this article concerning the occasionally antagonistic relationship between rational choice applications of game theory and evolutionary applications. While Hume’s approach is closer to modern evolutionary game theory in most respects, Narveson appears to want to use both approaches from time to time. For more on the contrast between rational choice game theory and evolutionary game theory, see Skyrms, ‘The Stag Hunt’, *Proceedings and Addresses of the American Philosophical Association*, 75:2, 2001: 31–41.
- 12 Narveson, ‘The Agreement to Keep Our Agreements’, 76–7.
- 13 *Ibid.*, 77.
- 14 David Gauthier, *Morals by Agreement* (New York/Oxford, 1986).
- 15 Narveson, ‘The Agreement to Keep Our Agreements’, 78.
- 16 Robert Axelrod, *The Evolution of Cooperation* (New York, 1984).
- 17 Jan Narveson, ‘Contracting for Liberty’, in T. R. Machan and D. B. Rasmussen (eds), *Libertarianism for the 21st Century* (Lanham, MD, 1995).
- 18 See Narveson, *Libertarian Idea*, 150–53.
- 19 See especially Jan Narveson, *Morality and Utility* (Baltimore, MD, 1967).
- 20 For a sketch of some of these views, and for the arguments of the next paragraph, see Narveson, *Libertarian Idea*, 150–53.
- 21 *Ibid.*, 153.
- 22 Ants? In ‘The Agreement to Keep Our Agreements’, Narveson describes the Social Contract as ‘an agreement in the sense of a co-ordinated set of conditional dispositions’, 84. It is important to Narveson’s overall argument that this is what the Social Contract is, and that it not depend on a prior or a more basic agreement. If so, then wherever one finds such coordinated sets of conditional dispositions, there one ought also to have a Social Contract. Ants.
- 23 It is, of course, important to note that the ‘interests’ promoted by restrained behaviour on the part of individual animals, human or not, may not always be those of the self-same individuals. It might be, for example, that what is promoted in some Social Contracts (that is, coordinated sets of conditional dispositions – see note 22, above) is the ‘interest’ of the tribe or group or species. Such interests, of course, can conflict with those of the individual, opening again the motivational question that Social Contract theory is supposed to resolve. Why should I care about the Social Contract? Perhaps a contractarian would argue that some Social Contracts are more reasonable than others to abide by, but the question then would be ‘Which ones?’ and the answer seems likely to depart from references to contracts, social or otherwise.
- 24 Peter Stemmer, while reluctantly going along with the tradition of using the expression ‘hypothetical contract’ to refer to a justificatory mechanism like the one proposed by Narveson, observes cogently that what actually does the justificatory work appears to be no more than a ‘configuration of interests’ among the members of some group, and that what makes contractarian terminology plausible is simply the fact that it is just such configurations that often, under normal circumstances, provide the framework for

- actual contracts. See Stemmer, 'Moralischer Kontraktualismus', *Zeitschrift für philosophische Forschung*, 56/1 (2002): 1–21, esp. 16–17.
- 25 Jan Narveson, 'Animal Rights Revisited', in H. Miller and W. Miller (eds), *Ethics and Animals* (Clifton, NJ, 1983). See also Jan Narveson, 'A Case Against Animal Rights', in *Advances in Animal Welfare Science* (Washington, DC, 1986); and Jan Narveson, *Moral Matters* (2nd edn) (Peterborough, Ontario, 1999), ch. 6: 'Morals and Animals', 133–42.
- 26 This idea also seems to misapply the logic of Social Contract as set out elsewhere in Narveson's work (see previous citations), which depends upon what it would be reasonable to agree upon, rather than upon what anyone actually agrees to.
- 27 Narveson writes, on occasion, as if his contractarianism is superior to other foundational approaches precisely because it does indeed emphasize particular persons rather than abstractions. He presses this point with particular emphasis in the opening paragraphs of an as yet unpublished essay with the working title 'Baseline, Baseline, Where Art Thou?' But Narveson does not actually accomplish this in any of his published work, and for good reasons. His objective, after all, is a universal moral code, so he must inevitably deal with people in terms of what they have in common, rather than in terms of their individual differences. Narveson's solution is to speak about what 'everyone' would agree to, rather than what anyone agrees to.
- 28 It is ironic that some of the reasons offered by Narveson for thinking of his proffered 'foundations' for morality as contractarian appear to be precisely the reason some thinkers offer by way of *criticism* of contract theory. Hegel, for example, argued that there are good reasons to opt into social institutions whether or not we have given our consent to them, and these reasons can justify those institutions quite independently of consent. Because consent is not necessary – or even morally relevant independently of social institutions, according to Hegel – the idea of 'contract' has no justificatory role to play. Narveson's concession that the 'Social Contract' is not a contract, but may instead be inferred or read from behavioural tendencies, together with his willingness to infer a 'Social Contract' from what it is reasonable for people to agree to whether they actually give their consent or not, seems to bring his view into what Narveson might regard as an uncomfortably close alignment with Hegel's anti-contractarian view. See Georg Wilhelm Friedrich Hegel, *Die Vernunft in der Geschichte*, J. Hoffmeister (ed.) (Hamburg, 1955), 118. See also Alan Patten, 'Social Contract Theory and the Politics of Recognition in Hegel's Political Philosophy', in Robert R. Williams (ed.), *Beyond Liberalism and Communitarianism: Studies in Hegel's Philosophy of Right* (Albany, NY, 2001), 167–84.