

A Royal Road to Consequentialism?

Martin Peterson

Division of Philosophy, Royal Institute of Technology, Stockholm

November 16, 2008

Abstract. To consequentialise a moral theory means to account for moral phenomena usually described in nonconsequentialist terms, such as rights, duties, and virtues, in a consequentialist framework. This paper seeks to show that all moral theories can be consequentialised. The paper distinguishes between different interpretations of the consequentialiser's thesis, and emphasises the need for a cardinal ranking of acts. The paper also offers a new answer as to why consequentialising moral theories is important: This yields crucial methodological insights about how to pursue ethical inquires.

1. Introduction

Euclid famously told King Ptolemy that, 'there is no royal road to geometry'.¹ But what about other theoretical constructs? Is there any royal road to any other theory worth thinking about? Some moral philosophers believe there is a royal road to consequentialism in ethics. On their view, every nonconsequentialist moral theory can be *consequentialised*. To consequentialise a moral theory means to account for moral phenomena usually described in nonconsequentialist terms, such as rights, duties, and virtues, in a consequentialist framework. In a recent article, Portmore proposes the following general recipe for consequentialising nonconsequentialist theories:

Take whatever considerations that the nonconsequentialist theory holds to be relevant to determining the deontic status of an action and insist that those considerations are relevant to determining the proper ranking of outcomes. In this way, the consequentialist can produce an ordering of outcomes that when combined with her criterion of rightness yields the same set of deontic verdicts that the nonconsequentialist theory yields. (Portmore 2007:39)

To see how consequentialising works in practice, consider the Kantian claim that lying is always wrong. If you face a choice between lying and telling the truth, the consequentialist could mimic the advice offered by Kantians by assigning a high utility to telling the truth and a low utility to lying. According to Dreier (1993:23) rival moral theories can be consequentialised in similar ways, because we can always, 'take the

features of an action that the theory considers to be relevant, and build them into the consequences.’²

If true, the claim that every moral theory can be consequentialised reveals an important asymmetry between consequentialism and other theories. Consequentialists can account for phenomena that are usually thought of in nonconsequentialist terms, such as rights, duties, and virtues, whereas the opposite is false of nonconsequentialist theories. Rights, duty or virtue-based theories cannot account for the fundamental moral importance of consequences.³ Because of this asymmetry, it seems it would be preferable to become a consequentialist – indeed, it would be virtually impossible not to be a consequentialist. Call this the asymmetry argument.

This paper defends two claims about the asymmetry argument. The first is about its truth. I argue that every moral theory can indeed be consequentialised. However, previous accounts of how to consequentialise moral theories, such as those proposed by Dreier and Portmore, have either produced ordinal rankings of acts, which I argue is unsatisfactory, or cardinal rankings based on technical assumptions that many non-consequentialists find implausible. I therefore develop two new accounts, and I argue that the new accounts avoids the flaws attributed to the previous ones.

The second claim I defend is about the relevance of the asymmetry argument. Even if every moral theory can be consequentialised, it is far from obvious that consequentialists should welcome this insight. For example, Brown (2004) argues that this threatens to make consequentialism vacuous, since no particular conclusion then follows from consequentialism about what ought to be done in any given situation. Or, differently put, what is the point of accepting a moral theory that has virtually no normative implications? I argue that the value of the consequentialiser’s thesis does not lie in its practical implications, but rather in the methodological insights it gives us about how to pursue ethical inquiries.

2. Four interpretations

How should we understand the claim that nonconsequentialist moral theories can be consequentialised? It seems that this idea can be interpreted in at least four different ways, which are obtained by combining two distinctions. The first is between weak and strong interpretations of the claim, and the second between general and restricted interpretations.

According to *strong* interpretations, every moral theory is deontically equivalent (in a sense to be explained in the next section) to some version of consequentialism. This is because all moral theories – including duty ethics, rights-based theories, and virtue ethics – can be represented in some utility function as claims about consequences. If correct, this shows that the theoretical divide between, say, duty ethics and utilitarianism is no greater than the divide between hedonistic utilitarianism and preferentialism. People advocating rival moral theories just make slightly different claims about how to evaluate consequences.

According to *weak* interpretations, there is a genuine difference between consequentialism and non-consequentialism that gives rise to an asymmetry in explanatory power. nonconsequentialist theories can be consequentialised in the sense that consequentialists can accommodate and explain intuitions about rights, duties, and virtues, but the opposite is false of rival theories. For example, rule utilitarians typically claim that intuitions about rights can be justified and explained by the fact that all things considered, the long-term consequences of respecting certain rights tend to be favourable. However, this is not to say that people really *have* rights in the absolute sense proposed by Locke and Nozick. On a weak interpretation, consequentialism merely provides an alternative justification of a number of widespread moral intuitions, but the theory itself is essentially incompatible with the nonconsequentialist theories usually taken to justify those intuitions.

Restricted interpretations hold that the possibility of consequentialising a moral theory reveals an asymmetry between consequentialism and *some* of its rivals, whereas *general* interpretations claim that this asymmetry holds between consequentialism and *all* of its rivals. Many scholars defend restricted interpretations. For example, Brandt (1984), Hare (1982), Pettit (1988), and Sen (1982, 2004) argue that consequentialist theories can recognise moral rights. For instance, Sen argues that a violation of a right is an evil that ought to be counted as *intrinsically* bad when evaluating a state of affairs. His point is not that the violation of a right makes people unhappy or frustrates someone's preference. The violation of a right is bad in itself, and this badness can be accounted for in a consequentialist framework. Sen's point is thus that the traditional utilitarian way of accounting for consequences is too narrow. Rights also matter, intrinsically, when evaluating a state of affairs. Clearly, this account of rights is meant to support to a strong, but restricted, interpretation of the consequentialiser's thesis. Sen *construes* rights in purely consequentialist terms, by making a normative claim about how to evaluate rights-violations.

Naturally, it is also possible to account for duties in consequentialist terms by using the line of thought appealed to by Sen in his justification

of rights. In this version of duty ethics, duties have intrinsic value. The duty not to lie and the duty not to commit suicide can thus be accounted for in consequentialist terms. If you tell a lie, a consequence of this act is that a certain amount of intrinsic value in the resulting state of affairs is lost. This holds true even if no one is ever made aware of the fact that you lied.⁴

Consequentialists have also attempted to account for the moral importance of virtues. Driver (1996, 2001a, 2001b) argues that, ‘On my theory the value of all these [character] traits resides in their tendency to produce good consequences’.⁵ Therefore, to adopt certain character traits will lead to good consequences and, consequently, this is a good thing to do. Since Driver construes virtues in consequentialist terms, she is in effect defending a strong interpretation. The correlation between virtues and good consequences is not supposed to be a mere empirical coincidence. The correlation is, according to Driver, an essential one. A virtue consists in bringing about good consequences. As pointed out by Bradley (2005), this is by no means an uncommon view among virtue ethicists, not even among scholars advocating ‘classic’ non-consequentialist versions of the theory. For example, Foot (1978) famously claimed that, ‘virtues are in general beneficial characteristics, and indeed ones that a human being needs to have, for his own sake and that of his fellows’.⁶ Arguably, this also rings true to the consequentialist. Thus, it is a mistake to think that consequentialism and virtue ethics are fundamentally inconsistent moral views. Perhaps the only major difference is that consequentialism focuses on the moral status of acts, but remains silent when it comes to answering the virtue ethicist’s question about what kind of person one ought to be.

Drier (1993), Brown (2004), and Portmore (2007) are mainly concerned with general and strong interpretations of the consequentialiser’s thesis. On their view, every moral theory, or at least every ‘remotely plausible’ theory, can be consequentialised, in the sense that it is deontically equivalent to some version of consequentialism. In what follows we shall focus on the strong and general interpretation, mainly because it seems to be the most challenging and interesting interpretation.

3. The strong and general interpretation

According to the strong and general interpretation every moral theory is deontically equivalent to some version of consequentialism. In order to assess this claim, we need to clarify at least three issues: What is a ‘moral theory’; what does it mean to say that two theories are

‘deontically equivalent’; and what counts as a ‘consequentialist’ moral theory?

A moral theory can be conceived of as a device that attaches deontic predicates to acts. By a deontic predicate I mean predicates such as ‘right’, ‘wrong’, ‘permissible’, ‘forbidden’, ‘obligatory’, and ‘supererogatory’. The proposed account of moral theories tallies well with the familiar utilitarian claim that it is ‘permissible’ to perform an act just in case no other act produces more wellbeing, and that no act is ever ‘supererogatory’. Moreover, the proposed account can account for the Kantian claim that lying is always ‘forbidden’, as well as for similar claims made by virtually any other moral theory.

In order to render this account more precise, let A_i be a set of alternative acts $\{a, b, c, \dots\}$, and let \mathbf{A} be a set of sets $\{A_1, A_2, A_3, \dots\}$. Intuitively, \mathbf{A} is the set of all choice situations. Moreover, let D be a set of deontic predicates, such as $\{\text{permissible, wrong, obligatory, supererogatory, } \dots\}$. Now, a moral theory can be conceived of as a function that takes an element A_i in \mathbf{A} as its argument and attaches a predicate in D to each element in A_i . If moral theories are conceived of in this way, it follows that the utility of every act can be represented numerically in an ordinal ranking. Consider, for instance, Brutus’s decision to murder Caesar, in which he faced a choice among five alternatives (see below). By hypothesis, murdering Caesar was morally forbidden, which we represent by assigning a low number, say 0, to this act. Kindly asking him to resign voluntarily would have been supererogatory, which we represent by assigning a higher number to this act, 2. Putting Caesar in prison falls somewhere between murdering him and asking him to resign; we represent this by assigning the number 1 to this act. As pointed out above, the numbers merely carry ordinal information – nothing can be concluded about how *much* better it would have been to put Caesar in prison rather than to kill him.

<i>Act</i>	<i>Deontic property</i>	<i>Ordinal ranking</i>
1. Murder Caesar	Forbidden	0
2. Force Caesar to resign without killing him	Permissible	1
3. Put Caesar in prison	Permissible	1
4. Kindly ask Caesar to resign voluntarily	Supererogatory	2
5. Do nothing	Forbidden	0

Two theories are *deontically equivalent* if and only if both theories produce the same (ordinal) ranking no matter which set of acts they are applied to. However, the mere fact that every moral theory can be associated with numbers does not prove that every theory can be

consequentialised. It remains to show that the numbers are somehow determined by the *consequences* of each act. Fortunately, this additional requirement is not as far-fetched as it might first appear to be. The consequences of an act can be conceived of as the entire state of affairs brought about by an act. Hence, if a duty ethicist promises to return a book to the library, but instead sets off for a holiday in the Caribbean, a consequence of this decision seems to be that the promise to return the book will remain unfulfilled.

This broad understanding of consequentialism is endorsed by Parfit in his widely circulated manuscript *Climbing the Mountain* (2006). He defines consequentialism as the view that, ‘the rightness of acts depends only on facts about how it would be best for things to go.’⁷ His characterisation thus supports the claim that virtually any fact can be relevant for determining the moral value of a consequence. Note that this even includes things that happened in the past, such as promises or legal contracts signed a long time ago. On Parfit’s view it would thus be incorrect to think that consequentialism is essentially future-oriented:

Consequentialists can believe that the goodness of some outcomes depends in part on facts about the past. On two such views, for example, it would be better if benefits went to people who had earlier been worse off, and it would be bad if people were punished for crimes that they did not commit. That is one reason for talking, not of the goodness of outcomes, but of how well things go, or will have gone. (Parfit 2006:233)

Parfit’s points about the relevance of past events shows just how broad the consequentialist’s ethical framework really is. However, at least two objections have been raised in the literature against the idea that every moral theory can be consequentialised.⁸ The first objection is about moral dilemmas and the second about supererogation.⁹ By definition, a moral dilemma is a situation in which every act is morally wrong, or forbidden. No matter what you do, you will do something you ought not to do. Most consequentialists find it hard to see how moral dilemmas could ever arise, because they believe every set of acts must comprise at least one act whose consequences are at least as good as those of all alternatives. However, note that this reasoning relies heavily on the assumption that the ranking of consequences is non-cyclical and complete. If the ranking is cyclical all acts would be wrong, since for every act there would exist some act that produces better consequences. Furthermore, a moral dilemma could arise if some consequences are incomparable, i.e. if it is false that one consequence is better than the other and false that they are equally good. In such a case, the ranking

would be incomplete, and it would thus be false that every set of acts comprises at least one act whose consequences are at least as good as those of all alternatives. Hence, given that an act is right just in case its consequences are at least as good as those of all alternatives, and wrong if and only if it is not right, a choice among incomparable acts constitutes a moral dilemma.

So what about supererogation? By definition, an act is supererogatory just in case it goes above and beyond what is morally required. From a consequentialist point of view, it is natural to link supererogation to satisficing consequentialism. Some consequentialists, like Slote (1985), believe that one is not always morally required to maximise. Sometimes one merely has to do whatever produces sufficiently good consequences – doing what has best consequences is supererogatory. It seems that this idea nicely captures a consequentialist notion of supererogation. Whenever an act goes beyond and above what is morally required, the consequentialist theory attaches the predicate ‘supererogatory’ to the act in question.

The preliminary conclusion is that consequentialists can account for many of the features that non-consequentialists wish to emphasise. However, it remains to show that this broad understanding of consequentialism really captures the central ideas we normally associate with consequentialism. Unlike the authors discussed above, I do not think this is case. Briefly put, the problem lies in the fact that the account given so far merely provides an ordinal ranking of acts and consequences, whereas we need a cardinal ranking.

4. The need for a cardinal ranking

Consequentialism is attractive mainly because of its conceptual clarity and precision, not because of its normative recommendations. As pointed out above, consequentialism is compatible with every possible view about what one ought to do in a particular situation. However, in order to take advantage of the conceptual advantages of consequentialism we need more than just an ordinal ranking of acts. Contrary to what is sometimes claimed, the clarity and precision of consequentialism comes alive once we obtain a cardinal ranking of acts.¹⁰

For brevity, let ‘cardinal consequentialism’ refer to consequentialist theories that acknowledge the need for cardinal measurement, and let ‘ordinal consequentialism’ refer to views that reject this requirement. By definition, cardinal consequentialists maintain that numbers assigned to acts and consequences reflect distances in utility, and sometimes even ratios.

The conceptual advantages of cardinal consequentialism fall into three categories. First, cardinal consequentialism can – unlike ordinal consequentialism – make sense of the distinction between act and rule consequentialism. Some consequentialists maintain that an act should be performed because its actual consequences are good, whereas others maintain that it should be performed because the average consequences of this type of act in the long run tend to be good. Arguably, if the thesis that every moral theory can be consequentialised is to be of philosophical interest, the consequentialist theory one ends up with must be sensitive to this distinction. A consequentialiser who cannot account for the difference between act and rule consequentialism has not succeeded to deliver a theory that deserves the label ‘consequentialism’. However, only cardinal consequentialism can account for this distinction. Rule consequentialism presupposes that one is able to calculate averages (or at least sum up the utility of different consequences into a sum total) and this requires that we measure utility on a cardinal scale.

The second conceptual virtue that speaks in favour of cardinal consequentialism is that it can be successfully applied to cases in which the morally relevant features of a decision are partially known. Suppose, for instance, that you know the probabilities of a set of consequences. Then, in order to distinguish between views holding that an act is right just in case it maximises expected utility, and views holding that one ought to maximise actual utility, we again need cardinal consequentialism. The principle of maximising expected utility can only be applied if utility is measured on a cardinal scale.

The third argument is related to sequential decision making. Sometimes it makes sense to divide a decision into a sequence of separate decisions taken at different points in time. However, in order to choose a sequence of decisions that maximises the sum total of good consequences, utility has to be measured on a cardinal scale. Otherwise one cannot tell whether, for example, two acts with pretty good consequences together produce more good consequences than a single act with really good consequences.

For the reasons outlined above, I will in the following focus on cardinal consequentialism. The thesis I wish to examine is thus: Is it true that every moral theory is deontically equivalent to some version of cardinal consequentialism? To the best of my knowledge, Oddie and Milne (1991) are the only authors who have defended this claim. They argue that every moral theory M can be *represented* by some other moral theory M^* , and that M^* is certain to be a version of cardinal consequentialism. More precisely put, they argue that no matter which moral theory one subscribes to, it is always possible to assign cardinal numbers to acts in a way that conforms to the consequentialist dictum

that an act is right if and only if its consequences are at least as good as those of all alternatives.

Let us take a closer look at Oddie and Milne's account. They start from the assumption that every moral theory is concerned with a finite number of deontic categories. The purpose of formulating a moral theory is, in their view, to sort all alternative acts into exactly one of these deontic categories.¹¹ The deontic ordering produced by the moral theory M is represented by the binary relation \geq_M . In line with what we have said above, two moral theories M and M^* are *deontically equivalent* just in case \geq_M and \geq_{M^*} order the alternative acts $\{a_1, \dots, a_n\}$ in the same way.

According to Oddie and Milne, a moral theory is a version of consequentialism just in case: (i) there is a function that assigns utility to states of affairs, (ii) the deontic features of an act are determined solely by the utility of the act, and (iii) the utility of an act is determined by a function u that takes the utilities and probabilities of the possible consequences of the act as its argument. Oddie and Milne impose three structural conditions on the function u , which entail that the utility of an act equals its objective expected utility of its consequences. Now, the main theorem proved by Oddie and Milne states that for every M , if the expected utility axioms hold for all acts, then there exists a deontically equivalent theory M^* that fulfils conditions (i) – (iii).

However, a major problem with Oddie and Milne's representation theorem is that many non-consequentialists would not accept the premises on which it is based. Oddie and Milne *stipulate* that the utility of every act can be represented by some real number.¹² From a nonconsequentialist point of view, this is an unacceptable assumption. To start with, some non-consequentialists believe that consequences might be incomparable. For instance, an agent facing a genuine moral dilemma, in which he will do something wrong whatever he does, might say that such alternatives cannot be morally compared. It is neither the case that one consequence is better than the other, nor are they equally good. Hence, since the existence of moral dilemmas is excluded by a real-valued representation of consequences, some non-consequentialists have good reasons for rejecting Oddie and Milne's approach.

An additional problem with stipulating that the utility of every act can be represented by some real number is that this presupposes the Archimedean property of the reals. (For all x, y , if $x > y$ then there exist some integer n such that $n \cdot y > x$.) This means that no act can be *infinitely* better than another. Hence, if one consequence is assigned a high number and another consequence a low number, it then follows that the first consequence is a finite number of units better than the latter. From a nonconsequentialist point of view, this is

an unacceptable assumption. For instance, defending his rights-based theory, Nozick would claim that violating a right is *infinitely* worse than not violating a right, and Kantians would make the same point about the moral difference between lying and not lying. Therefore, the distance in utility between violating a right instead of respecting it, or between lying instead of telling the truth, cannot be represented as a distance between real numbers.

The weakness of Oddie and Milne's cardinal approach can be further clarified by considering the von Neumann-Morgenstern axiomatisation of the expected utility principle.¹³ Imagine an agent who is facing a choice between a set of alternative acts $\{f, g, h, \dots\}$. Assume that you are an external observer, observing the agent's behaviour. Occasionally, there is some uncertainty with respect to which act will be performed by the agent (as viewed from the perspective of the external observer). Perhaps the external observer is sitting on an ethics committee advising others what to do. By assumption, the external observer knows all there is to know about the acts.

Naturally, defending his favourite moral theory, the external observer agrees that it is better that an act deemed to belong to the highest deontic category is chosen with a high probability, rather than with low probability. The external observer is even prepared to express preferences over probabilistic mixtures of alternatives. Hence, the observer may claim that, say, the lottery in which ' f is performed with probability p and g with probability $1 - p$ ' is better than the lottery in which ' g is performed with probability q and h with probability $1 - q$ '.¹⁴ What structural conditions, if any, should these preferences over probabilistic mixtures of acts satisfy?

Enter the vNM axioms. In order to state these purely structural conditions, let $F = \{f, g, h, \dots\}$, and let G be the set of functions $p : F \rightarrow [0, 1]$ such that $\sum_{f_i} p(f_i) = 1$. We assume that for every number $\alpha \in (0, 1)$ and every f and g in G , the composite act $\alpha \cdot f + (1 - \alpha) \cdot g$ is also in G . Hence, $G = \{f, g, h, \dots\}$ is the set of all possible (binary) probabilistic mixtures between the elements of F . Let the relation \succ on G denote a comparative moral evaluation between a pair of its elements, i.e. $f \succ g$ means that f is morally better than g . In this framework, the vNM axioms can be formulated as follows.¹⁵

- (i) The relation \succ is complete and transitive in G .
- (ii) If $f \succ g$, then $\alpha \cdot f + (1 - \alpha) \cdot h \succ \alpha \cdot g + (1 - \alpha) \cdot h$.
- (iii) If $f \succ g \succ h$, then there exist $\alpha, \beta \in (0, 1)$ such that $\alpha \cdot f + (1 - \alpha) \cdot h \succ g \succ \beta \cdot f + (1 - \beta) \cdot h$.

Condition (i) is an ordering condition, which implies that all acts have to be fully comparable from a moral point of view. This assumption is discussed in more detail below. Condition (ii) is the infamous independence condition, which has been criticised by decision theorists.¹⁶ Condition (iii) is an Archimedean condition. It claims that it will always be possible to construct a lottery between something good and bad, which is strictly better than something in between, given that the probability for the good consequence (act) is high enough.

Taken together, conditions (i) – (iii) imply that the moral value of each act f, g, h, \dots in G can be represented by a utility function that is unique up to a positive linear transformation. This means that the utility of each alternative act can be measured on an interval scale. (Note that utilities are assigned both to probabilistic acts as well as to normal, non-probabilistic acts in G .) Formally put, the representation and uniqueness theorem states that there exists a function u from G to $[0,1]$ such that:

1. If $f \succ g$, then $u(f) > u(g)$, and
2. $u(\alpha \cdot f + (1 - \alpha) \cdot g) = \alpha \cdot u(f) + (1 - \alpha) \cdot u(g)$, and
3. for every other function u' satisfying (1) and (2), there are numbers $c > 0$ and d such that $u' = c \cdot u + d$.

A proof of this result can be found in any good textbook on decision theory.¹⁷ Unfortunately, non-consequentialists have strong reasons to reject conditions (i) and (iii). The problem with Condition (i) is that it is incompatible with the existence of moral dilemmas, because if the agent faces a moral dilemma this seems to be a case in which the relation \succ is *not* a complete order. If you, for instance, face a moral dilemma in which you cannot avoid violating either your brother's rights or your sister's, it is neither better, nor worse, nor equally as good to prefer one option over the other.

Condition (iii), the Archimedean condition, is also unacceptable from a nonconsequentialist point of view. Both Kant and Nozick would agree that the probabilistic trade-off between harms inherent in the Archimedean condition is morally unacceptable according to many versions of non-consequentialism. Nozick explicitly claims that 'Imposing how slight a probability of a harm that violates someone's rights also violates his rights'.¹⁸ However, the Archimedean condition contradicts this intuition. If f and g are two fairly good acts such that f is slightly better than g , and h is some morally horrible act, such as the act of

torturing a baby for fun, then the Archimedean condition guarantees the existence of a lottery in which f is performed with a high probability and h with a very low probability, which is strictly better than the performance of g with certainty. For the non-consequentialist, this is absurd.

What should one make of these objections? Some non-consequentialists obviously believe in moral dilemmas, so if the consequentialiser cannot model this idea he seems to be doomed. The same is true of the non-Archimedean feature of many prominent nonconsequentialist theories. However, I think neither objection poses any genuine threat. In fact, not even all consequentialists can accept the vNM axioms. The consequentialist endorsing Rawls's maximin condition must, for instance, reject the Archimedean condition. Of course, it would be incorrect to conclude that such non-Archimedean theories are not versions of consequentialism. It would also be incorrect to conclude that such theories cannot be mathematically represented by utility functions. All that follows is that not every moral theory can be represented by a von Neumann-Morgenstern function. In the literature of measurement theory, several alternative measurement systems have been proposed. A non-Archimedean (lexicographic) expected value function was even briefly considered by von Neumann and Morgenstern (1947: 631), and full-blown non-Archimedean analyses have been detailed by others. For example, by dropping the Archimedean condition and replacing it with a strengthened version of the independence condition, Hausner (1954) obtained a vector representation of utility.¹⁹ In a similar vein, Carlsson (2006) has shown that the completeness assumption can be dropped, given that we are prepared to accept a vector representation.

Naturally, in order to show that every moral theory can be consequentialised, it will suffice to show that there is *some* way in which every moral theory can be represented by a utility function. It is not necessary that there is a *single* way in which all theories can be represented. In the next two sections we shall, nevertheless, consider two different ways of representing all moral theories in a unified way. A unified representation is attractive mainly because this is simpler than using different representations for different theories.

5. First Proposal: A probabilistic approach

Moral philosophers sometimes distinguish between what an agent ought to do given her present state of information, and what a fully informed and ideally rational agent ought to do. In this section we propose a way of consequentialising all possible answers to the first type of

moral question. In the next section we show how all possible answers to what fully informed and ideally rational agents ought to do can be consequentialised.

Briefly put, the main idea in the probabilistic approach is to distinguish between two notions of acts, viz. *real-world acts* and *descriptions* of real-world acts. When the moral philosopher sets up a thought experiment in which killing one in order to save four has better consequences than refraining from actively killing anyone, he is merely describing the abstract contours of some real-world act. Although statements about the moral status of descriptions of acts can perhaps be made with absolute certainty, statements about the moral status of real-world acts will always be more or less uncertain. There is a huge difference between *stipulating* that some act in a thought experiment would lead to better consequences overall than all alternatives, and saying that a particular real-world act would lead to better consequences overall. The second statement will always be riddled with some uncertainty. For example, if I pull the trigger of my gun the bullet will probably kill you, and therefore it is highly probable that the act in question is morally wrong. However, it is *possible* that the bullet would instead have accidentally killed the person standing on your right, who might one day have become a cruel dictator, in which case the act might have been morally right. In similar ways, there may always be some doubt about the actual motive triggering one's act, or some doubt about whether a promise was actually made ten minutes ago.

In what follows, the possibility of moral doubt will play a crucial role. It is helpful to make this assumption as explicit as possible: *All* claims about the moral ranking of real-world acts are susceptible to error. Moral uncertainty can be modelled by ascribing probabilities to which act has the highest deontic status. In what follows, we shall ascribe probabilities to statements such as, 'act f has the highest deontic status in the set of alternatives B according to the moral theory M '. This probabilistic notion roughly corresponds to Oddie and Milne's deontic relation \geq_M . If you accept a moral theory M according to which you have a duty never to break a promise, and you are fairly certain that you promised the librarian to return *Naming and Necessity* today, the probability might be, say, .99 that you are obliged to return the book instead of keeping it for another week. However, there is a small probability that you are not obligated to return the book, since it is not entirely certain that you actually promised the librarian to return it. These probabilities can be interpreted either subjectively or objectively. If interpreted subjectively, they tell us what the agent subscribing to M *believes* about the moral status of some act. If interpreted objectively, the most plausible interpretation is, arguably, to maintain that

the probabilities reveal information about what an ideal, fully rational agent *ought* to believe about the moral status of some act. The moral theory M attaches deontic predicates to abstract descriptions of acts, which invariably prompts the agent accepting M to ascribe deontic probabilities to the corresponding real-world acts. Thus, on this version of the consequentialiser's thesis, every version of non-consequentialism will be equivalent to some version of consequentialism on the most concrete level, i.e. when it comes to making moral evaluations of real-world acts.

This probabilistic framework is compatible with non-Archimedean moral theories. This is because one merely ascribes probabilities to qualitative comparisons between acts, and one can never be sure that one alternative is deontically better than another. The agent is thus never justified in claiming with full certainty that, say, not murdering a person is certain to be infinitely better than doing so. In the present framework, the non-Archimedean conviction that murdering people is infinitely worse than not doing so will simply be represented by a probability close to one that one ought not to murder. The probabilistic approach is moreover compatible with the existence of moral dilemmas. No matter whether the agent faces a genuine moral dilemma, or a situation that just appears to be a dilemma, it holds true for each dilemmatic alternative that there is some probability (greater than zero) that it has a higher deontic status than the other alternative according to theory M . The insight that neither option is permissible does not affect the fact that a choice has to be made. Of course, an agent facing a moral dilemma may feel horrible 'on the inside' compared to an agent facing a set of non-dilemmatic options, but when it comes to choosing what to do, there is no fundamental difference.²⁰

The deontic probability of an act (the probability that it has a higher deontic status than some alternative act relative to a moral theory M) is directly linked to its utility. Let $p(f \succ_M B)$ denote a probability of p that act f has the highest deontic status in B according to the moral theory M . Let $p(A \succ_M B)$ denote the probability of the following conditional: If A is a subset of B , then the probability is p given M that the top element in B is also an element of A .

Now, given this notation, the choice axiom proposed by Luce (1959/2005) seems to be a plausible structural constraint on deontic probabilities. The choice axiom holds that if A is a subset of B , then the probability that f would be chosen from B equals the probability that f would be chosen from A multiplied by the probability that the chosen alternative in B is also an element of A . In symbols,

CHOICE AXIOM: Let $A \subset B$. Then $p(x \succ B) = p(x \succ A) \cdot p(A \succ B)$.

To grasp what kind of assumption is at stake in this axiom, it is helpful to consider an example from the decision theoretical context in which it was originally proposed. Suppose a lady visiting a posh restaurant is about to choose a wine from a list containing two red and two white wines. The axiom states that it should not matter if the woman divides her choice into two stages, that is, first choose between red and white wine, and then between the wines in the chosen subset, or choose directly which of the four wines to order. Hence, if the woman is indifferent to the choice between red and white wine in general, as well as to the choice between the two red wines and the two white ones at hand, the probability that a particular bottle will be chosen is $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. In the present context, the probabilities, of course, refer to how certain it is that one option is morally better than another.

The choice axiom implies the existence of a cardinal utility scale.

THEOREM 5.1 (Luce 1959/2005). *If the choice axiom and the axioms of the probability calculus hold for B and all its subsets, then there exists a positive real-valued function u on B such that for every $A \subset B$, $p(f \succ_M A) = \frac{u(f)}{\sum_{g \in A} u(g)}$, and for every other function u' satisfying this condition there is a constant k such that $u = k \cdot u'$.*

The proof is straightforward and follows Luce (1959/2005:23-24).²¹ As indicated above, the probabilistic approach only applies to real-world acts, i.e. to cases in which the probability that each act ought to be chosen is strictly greater than zero. For moral theories that seek to tell us what an agent ought to do given her present state of information, this is no doubt a reasonable restriction. However, when moral theories are used for non-practical purposes, such as theoretical explanations of the nature of moral obligations, moral philosophers routinely wish to discuss cases in which it is certain that some acts is, say, wrong. The probabilistic approach cannot consequentialise such moral verdicts. It is, thus, still an open question whether *all* nonconsequentialist claims can be reformulated in consequentialist terms.

6. Second Proposal: Hyperreal numbers

In this section we shall propose a more general way of consequentialising moral theories, which seems to avoid the difficulty outlined above. The new proposal uses hyperreal numbers.²² The hyperreals is an extension of the real numbers introduced by Robinson (1966). The main advantage of the hyperreal number system is that it allows us to give a rigorous mathematical characterisation of infinite

and infinitesimal numbers, without losing any results about the reals. (The so-called transfer principle guarantees that every true first order statements about the reals can be reinterpreted as a true first order statement about the hyperreals.) Intuitively put, an infinitesimal number is a number that is smaller than all numbers in the sequence $1/2, 1/3, 1/4, 1/5, \dots$, but greater than 0, whereas an infinite number is a number that is greater than all integers in the sequence $1, 2, 3, \dots$. More precisely put, the hyperreal number system differs from the real number system in that it comprises numbers that have the following properties:

Infinitesimals: For some x , $|x| < 1/n$ for all integers n .

Infinite Numbers: For some x , $x > n$ for all integers n .

The hyperreals can be constructed in different ways. However, the most straightforward approach is to use the ultrapower construction. Consider the following infinite sequences of reals:

$$a = [1, 2, 3, 4, 5, \dots,]$$

$$b = [0, 0, 0, 4, 5, \dots,]$$

$$c = [2, 4, 6, 9, 10, \dots,]$$

$$d = [1, 1, 1, 1, 1, \dots,]$$

$$e = [2, 2, 2, 2, 2, \dots,]$$

$$f = [1/2, 1/3, 1/4, 1/5, \dots,]$$

$$g = [1/4, 1/9, 1/16, 1/25, \dots,]$$

The ultrapower construction uses a so called ultrafilter (i.e. a certain kind of nonempty sub-set) on the set of natural number for constructing and comparing hyperreals from infinite sequences of reals. Several different ultrafilters can be used, and the hyperreals are unique only up to the choice of ultrafilter. This means that each ultrafilter may produce a different number system. However, the hyperreal numbers that we are likely to use or encounter will be same no matter which ultrafilter is chosen, so for present purposes we need not worry about the mathematical details. Let us just keep in mind that no matter

which ultrafilter is chosen every infinite sequence of reals corresponds to a hyperreal number, and that two sequences that differ only in their first n members, but are equal after that, represent the same hyperreal number. Hence, $a = b$. Moreover, since the sequence of reals in c approaches infinity ‘faster’ than the numbers in a , it follows that c is a greater hyperreal than a . For similar reasons, g is a smaller infinitesimal than f . Finally, $e > d$. Hence, $c > a = b > e > d > f > g$. Finally, it is useful to point out that the hyperreals satisfy closure, commutativity, associativity, and distributivity.²³

So why are hyperreal numbers useful for consequentialising moral theories? The main point is that all deontic verdicts derived from any moral theory can be represented by hyperreal numbers, including deontic verdicts about supererogation, non-archimedean values, and moral dilemmas. In order to see how this can be done, we return to Brutus’s decision to murder Caesar, in which he faced a choice among a set of alternatives (see below). By hypothesis, murdering Caesar was morally forbidden, which we represent by assigning a low number, say d , to this act. Putting Caesar in prison would have been morally right and infinitely better, which we represent by assigning the infinite number a to this act. Furthermore, kindly asking him to resign voluntarily would have been even better and supererogatory, which we representant by assigning a higher infinite number to this act, c . Note that the numbers merely carry cardinal information – they tell us how *much* better it would have been to put Caesar in prison rather than to kill him.

- (i) $u(\text{Murder Caesear}) = d$
- (ii) $u(\text{Put Caesar in prison}) = a$
- (iii) $u(\text{Kindly ask Caesar to resign voluntarily}) = c$

It remains to account for moral dilemmas. There are two alternatives to consider. The first holds that moral dilemmas should be accounted for as cases in which all alternatives are assigned numbers below a certain threshold (say, e or lower), that is, in roughly the same way as in the ordinal approach. The problem with this suggestion, both when applied to cardinal and ordinal contexts, is that it is incompatible with the consequentialist principle saying that an act is permissible if no other act leads to better consequences. If all acts have utilities that fall below the threshold, the one(s) with the highest utility ought to be permissible. However, since the proposed account of moral dilemmas is not compatible with this extremely plausible principle, we shall not discuss it any further.

The second alternative is more sophisticated.²⁴ To start with, note that moral dilemmas pose a problem for the consequentialiser because the hyperreal numbers obey the trichotomy principle: For every pair of hyperreals, it holds that one is greater than the other, or they are equal. This means that there is no conceptual space for the claim that moral dilemmas are cases in which alternatives are incomparable, i.e. in which the trichotomy thesis fails. That said, we may nevertheless introduce incomparability by generalising the thoughts underlying the hyperreal number system. Let D be a set of distances, such that distances in utility between acts corresponds to a hyperreal or real element in D . We then stipulate that an agent facing a choice among, say, three alternatives, a_1 , a_2 , and a_3 , should compare the first alternative with the two other alternatives and conclude that a_1 is permissible if and only if the distance in utility between a_1 and all other alternatives is positive or equal to zero. On this proposal, a moral dilemma can be conceived of as a case in which the distance in utility between some alternatives is not an element in D , that is, in which it is false that the distance is greater than zero, smaller than zero, or equal to zero.

Note that we assign utility to distances (differences) between acts, rather than to the acts themselves. This tallies well with the intuitions that led von Neumann and Morgenstern to construct a similar scale for measuring distances in terms of binary preferences among lotteries.

7. Concluding discussion

The upshot of this paper is an argument to the effect is that it is always possible to measure how well things go by assigning utilities on a cardinal scale to (comparisons between) acts. However, no substantial conclusion follows from this about what ought to be done in a particular situation.

So why is this result important? If every moral theory can be mimicked by some version of consequentialism, it seems that little or nothing has been gained – this theoretical manoeuvre seems to be of little help. Brown (2004: 3) even suggests that, ‘If every nonconsequentialist theory may be ... mimicked [by some version of consequentialism], then consequentialism is vacuous’.

However, I believe the significance of the thesis that every moral theory can be consequentialised is not that it offers any action guidance. Its significance rather stems from the fact that it helps us gain a better understanding of moral theories, especially of theories that have previously been thought of as nonconsequentialist views. This point is worth pushing a bit further. Imagine a future in which moral theo-

rists engage in ranking alternative acts as described above. Naturally, utilitarians and duty ethicists will continue to disagree over the moral value of breaking a promise. However, this disagreement would be no deeper than disagreements within utilitarianism itself. Some utilitarians believe that happiness matters, whereas others focus on objective lists of intrinsically valuable items. Both versions of consequentialism can be rendered very precise by applying the conceptual machinery available within the consequentialist framework.

Joanna Burch Brown has objected that a consequentialised theory need not be any more precise than a nonconsequentialist theory.¹ Her argument starts from the observation, which I think is correct, that no theory can be consequentialised unless it already is very precise. In fact, it seems that no nonconsequentialist theory can be consequentialised unless the nonconsequentialist is able to explain in detail what the theory would prescribe in all (or nearly all) possible situations. Hence, it is unclear whether we actually gain any precision by consequentialising a nonconsequentialist theory that is so precise that it *can* be consequentialised. In response to this objection I would like to remind the reader of three conceptual tools discussed above, all of which contribute to increased precision, which are currently available to the consequentialist but will remain unavailable to traditional nonconsequentialists until their theories have been consequentialised. The three conceptual tools are: 1) the distinction between act and rule based versions moral theories, 2) the distinction between the actual consequence and expected consequences of an act, and 3) the distinction between sequential and non-sequential decision making. For example, if we consequentialise duty ethics we will be able to distinguish between versions of duty ethics according to which one ought to act such that as many duties as possible are *actually* fulfilled, and versions according to which it is the *expected* duty-fulfilment that matters. Traditional accounts of duty ethics cannot make much sense of this useful distinction. This indicates that something important is actually gained by consequentialising a nonconsequentialist theory such as duty ethics.

Therefore, instead of trying to render theories usually thought of as nonconsequentialist more precise by using the rather blunt conceptual tools available within those frameworks, the insight that every moral theory can be consequentialised suggests that more precision can be gained by expanding the consequentialist circle to encompass views that have traditionally been thought of as non-consequentialist.

An additional advantage of consequentialising ethical theories is that we will achieve a form of conceptual unity. By learning that all

¹ Personal communication, Nov. 10 2008.

moral theories can be formulated in a common framework, we reduce the number of primitive concepts needed for stating different ethical views. This is an important achievement, since it makes it easier for proponents of different ethical views to communicate in a fruitful way with each other.²

Notes

¹ According to legend, Euclid rejected King Ptolemy's request for an easier way of learning geometry by pointing out to him that, 'there is no royal road to geometry'.

² A number of scholars, including Broome (1991), Oddie and Milne (1991), Carlson (1996), and Brown (2004), have defended similar views.

³ How do we show that not every version of consequentialism can be conceived of as a version of some nonconsequentialist moral theory? I believe this is easy. An important feature of many versions of consequentialism is that *every* kind of bad consequence can be weighted against some sufficiently large good consequence. Even murdering a friend might be right, if this prevents the death of many others. But standard interpretations of major nonconsequentialist theories, such as rights, duty and virtue ethics, imply that murdering a friend is always wrong, no matter the consequences. Therefore, consequentialism can never be shown to be a version of some nonconsequentialist theory.

⁴ Hare (2000:8.4) famously argues that Kantian duty ethics and consequentialism 'does not stand at opposite poles of moral philosophy' and that 'Kant ... *could* have been a utilitarian, though he *was* not'. Hare's main point is that the notion of impartiality inherent in the categorical imperative supports utilitarianism. On Hare's reading of the categorical imperative, 'I shall be able to will only such maxims as do the best, all in all, impartiality, for all those affected by my action. *And this, again, is utilitarianism*'. That said, there is an important difference between Hare's defence of duty ethics and the defence of outlined here: Hare does not assign intrinsic value to individual duties. He merely claims that the overall structure of Kant's theory, and in particular the notion of impartiality derived from the notion of universalisability, can be accounted for in consequentialist terms. According to Hare, Kant's famous claim that we have a duty not to commit suicide is simply false, and the same goes for many similar duties. Hare argues that his version of utilitarianism can account for the theoretical spirit of Kant's theory, but not for all its practical implications. Because of these circumstances, Hare's position cannot be classified as belonging to any of the four versions of the argument outlined above.

⁵ Driver (1996:122-3).

⁶ Foot (1978:3)

⁷ Parfit (2006:252)

⁸ See Brown (2004) and Portmore (2007) for excellent overviews of the literature.

² This paper was present at the conference organised by the British Society for Ethical Theory in Edinburgh in July 2008. I wish to thank the participants and organisers for helpful comments. I would also like to thank Joanna Burch Brown, Rafaela Hillerbrand, Iwao Hirose, Karsten Klint Jensen, and Toby Ord for helpful comments.

⁹ A third objection, not to be discussed here, is that consequentialists cannot make sense of agent-relative values, which are often cherished by non-consequentialists. However, as convincingly argued by e.g. Broome (1991), Skorupski (1995), Bykvist (1996), and Portmore (2001, 2003), it seems that some versions of consequentialism can actually take agent-relative values into account.

¹⁰ Brown (2004) and Portmore (2007) argue, incorrectly I think, that consequentialisers merely need an ordinal ranking.

¹¹ In a comment on the paper by Oddie and Milne, Carlsson (1995:30) argues that this view of a morality is over-simplified, ‘since a deontic ordering does not itself induce an ordering of [the agent’s] alternatives’. I think this complaint is mistaken. Arguably, one could expand the set of deontic categories such that each element in the ordering of alternatives corresponds to exactly one deontic category. Cf. Oddie and Milne (1991:52).

¹² Oddie and Milne (1991:56).

¹³ (Oddie and Milne make no direct reference to the vNM axioms, since they propose a slight different axiomatisation, but because the vNM axioms are all necessary consequences of the expected utility principle they are nevertheless committed to these axioms.

¹⁴ In order to link these preferences to things that matter for the observer, one may imagine some ingenious story in which the observer is an unselfish agent concerned with performing morally good acts; perhaps the agent takes advice from the observer, thereby giving the observer some probabilistic influence over what act will be performed by the agent.

¹⁵ von Neumann and Morgenstern expressed their axioms in a slightly different way; see their (1947: 24-27). The formulations given here, which are more attractive from a technical point of view, can be found in e.g. Kreps (1988: 43-4) and Schmidt (1998: 4-6).

¹⁶ See e.g. Kreps (1988) for an overview.

¹⁷ See e.g. Peterson (2009).

¹⁸ Nozick (1974: 74)

¹⁹ See also Skala (1975) and Schmidt (1998)

²⁰ That said, there may however be a difference in degrees here. If you face a genuine moral dilemma you may expect all deontic probabilities to be about the same, whereas in a situation that merely appears to be a dilemma the probabilities shall be closer to zero or one.

²¹ Since it was assumed that $p(A \succ B) \neq 0$, the choice axiom implies that:

$$p(x \succ A) = \frac{p(x \succ B)}{p(A \succ B)} \quad (1)$$

In order to prove the existence part, saying that a utility function u exists, let $u(x) = k \cdot p(x \succ B)$, where $k > 0$. Then, since the elements of A are mutually exclusive, the probability axioms guarantee the truth of the following equation.

$$p(x \succ A) = \frac{k \cdot p(x \succ B)}{\sum_{y \in A} k \cdot p(y \succ B)} = \frac{u(x)}{\sum_{y \in A} u(y)} \quad (2)$$

In order to prove the uniqueness part, suppose that u' is another function defined as above. Then, for every $x \in B$, the following holds.

$$u(x) = k \cdot p(x \succ B) = \frac{k \cdot u'(x)}{\sum_{y \in B} u'(y)} \quad (3)$$

By letting $k' = k / \sum_{y \in B} u'(y)$, it immediately follows that $u(x) = k' u'(x)$. Q.E.D.

²² I wish to thank Toby Ord for drawing my attention to the hyperreals.

²³ These properties are defined in most textbooks on elementary analysis.

²⁴ This idea was suggested to be by John Cantwell.

References

- Bradley, B. 2005. 'Virtue Consequentialism', *Utilitas* 17: 282-298.
- Brandt, R. 1992. *Morality, Utilitarianism, and Rights*. Cambridge: Cambridge University Press.
- Broome, J. 1991. *Weighing Goods*. Oxford: Basil Blackwell.
- Brown, C. 2004. 'Consequentialise This', unpublished manuscript.
- Bykvist, K. 1996. 'Utilitarian deontologies? On preference utilitarianism and agent-relative value', *Theoria*, 62:124-143.
- Bykvist, K. 2007. 'Violations of normative invariance: some thoughts on shifty oughts', *Theoria*, 73:98-120.
- Carlsson, E. 1995. *Consequentialism Reconsidered*. Theory and Decision Library, Kluwer Academic Publishers.
- Carlsson, E. 2006. 'Incomparability and Measurement of Value'. In K. McDaniel et al. (ed.), *The Good, The Right, Life, and Death* Aldershot: Ashgate.
- Dreier, J. (1993). 'Structures of Normative Theories', *The Monist* 76:2240.
- Driver, J. 1996. 'Virtue and Human Nature', in R. Crisp (ed.) *How should one Live?*, Oxford.
- 2001a. *Uneasy Virtue*. New York: Cambridge University Press.
- ed. 2001b. 'Character and Consequentialism'. Special Issue of *Utilitas*, 13:2.
- Foot, P. 1967. 'Abortion and the Doctrine of Double Effect', *Oxford Review*, 5: 28-41.
- Foot, P. 1978. *Virtues and Vices and Other Essays in Moral Philosophy*. Oxford: Blackwell.
- Foot, P. 1983. 'Utilitarianism and the Virtues', *Proceedings and Addresses of the American Philosophical Association*, 57, 2: 273-83.
- Hare, R. M. 2000. 'Could Kant have been a utilitarian?', chapter 8 in his *Sorting Out Ethics* Oxford: Oxford University Press.
- Kreps, D. 1988. *Notes on the Theory of Choice*, Westview Press.
- Locke J. 1993. *Two treatises of government*. London: Dent.
- Luce, D. 1959/2005. *Individual Choice Behavior. A Theoretical Analysis*. John Wiley and Sons.
- Nozick, R. 1974. *Anarchy, State, and Utopia*. Oxford : Blackwell.
- Oddie, G. & Milne P. 1991. 'Act and Value: Expectation and the Representability of Moral Theories', *Theoria* 57: 42-76.
- Parfit, D. 2006. *Climbing the Mountain*, manuscript.
- Peterson, M. 2009. *An Introduction to Decision Theory*, Cambridge University Press.
- Pettit, P. 1997. 'The Consequentialist Perspective' in *Three Methods of Ethics*, by M. Baron, P. Pettit, and M. Slote. Oxford: Blackwell.
- Portmore, D. 2001. 'Can an Act-Consequentialist Theory be Agent-Relative?' *American Philosophical Quarterly*, 38: 363-77.

- Portmore, D. 2003. 'Position-Relative Consequentialism. Agent-Centered Options, and Supererogation', *Ethics*, 113: 303-32.
- Portmore, D. 2007. 'Consequentializing Moral Theories', *Pacific Philosophical Quarterly* 88: 39-73.
- Robinson, A. 1966. *Nonstandard Analysis*, Princeton University Press.
- Schmidt, U. 1998. 'Axiomatic Utility Theory under Risk: Non-Archimedean Representations
- Sen, A. 1982. 'Rights and Agency', *Philosophy and Public Affairs*, 11:3-39.
- Sen, A. 2004. 'Elements of a Theory of Human Rights', *Philosophy and Public Affairs*, 32:315-56.
- Skala, H. 1975. *Non-Archimedean Utility Theory*, Springer.
- Slote, M. 1985. *Common-Sense Morality and Consequentialism*, Routledge.
- von Neumann, J. & Morgenstern, O. (1947) *Theory of Games and Economic Behavior*, 2nd edition, Princeton University Press.

