From consequentialism to utilitarianism

Citation for published version (APA):

Document status and date:
Published: 01/01/2003

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 14. Sep. 2023
FROM CONSEQUENTIALISM TO UTILITARIANISM*

In this article, we show that total act utilitarianism can be derived from a set of axioms that are (or ought to be) acceptable for anyone subscribing to the basic ideals of consequentialism.

I. INTRODUCTION

Total act utilitarianism prescribes that an act is morally right just in case the total sum of utility for all individuals is at least as high as that of all alternatives. The term ‘consequentialism’ denotes a more general set of ethical theories, according to which the moral status of an act is entirely determined by its consequences. Derek Parfit’s priority view is an example of a nonutilitarian version of consequentialism, which prescribes that benefits to the worse off should count for more than benefits to the better off. Another example is the maximin principle, which entails that an act is morally right just in case the consequences for the worst off individual are as good as possible. Further examples include different versions of egalitarianism, prescribing that the relative differences in well-being among individuals are ethically important (in most cases together with the total amount of well-being).

A plausible point of departure for a defense of total act utilitarianism is the question: Why is total act utilitarianism better than other versions of consequentialism, such as the priority view, the maximin principle, and different forms of egalitarianism? This is equivalent to asking how we know that utilitarianism is correct given that some version of consequentialism is known to be correct. Needless to say, a satisfactory answer to this question does not resolve all ethical problems. A substantial gap between consequentialism on one hand, and rights, duty, and virtue ethics, on the other hand, will remain.

Unlike John C. Harsanyi’s famous utilitarian theorem and John Broome’s improved version of this theorem, the axiomatization pre-

---

*I wish to thank Sven Ove Hansson, John Cantwell, Gunnar Björnsson, Hans Mathlein, Torbjörn Tannusjo, and an anonymous referee for valuable comments on earlier drafts.


2 The most well-known application of the maximin principle to an ethical context is John Rawls’s A Theory of Justice, (Cambridge: Harvard, 1971).

sented here does not rely on Bayesian decision theory. Hence, the new axiomatic system will appeal also to people who do not find the axioms of Bayesian decision theory acceptable.

A difficulty in Harsanyi’s and Broome’s Bayesian approaches, which has been recently discussed by Wlodek Rabinowicz, is that their axioms only imply that, “The overall goodness is an additively separable function of the individual goodness values, but they do not guarantee that each individual is treated equally in this addition.” In order to get all the way to utilitarianism, Harsanyi and Broome also have to assume that we somehow can calibrate the utility functions of different individuals and thereby make sure that everyone’s benefits count equally much. This would, obviously, presuppose cardinal comparisons of interpersonal utility. The axiomatic system presented here does not solve the problem of interpersonal utility comparisons, but cardinal comparisons of interpersonal utility play a more salient role in the new axiomatization, compared to Harsanyi’s.

In section ii, a set of ethical axioms are proposed, and in section iii, two utilitarian theorems are stated. The proofs of the theorems are given at the end.

II. ETHICAL AXIOMS

Somewhat roughly put, we are going to derive utilitarianism from the following five consequentialistic intuitions:

1. (The value of) consequences form an extensive structure.
2. In every ethical problem, at least one act is morally right.
3. If an act is split into two other acts, which both have exactly the same consequences as the original act, then the two new acts are morally right just in case the original act is morally right.
4. If an act yields strictly better consequences for all individuals than another, then the latter act is not morally right.
5. If the consequences of an act are better for individual \( i \) than for \( i' \), then there is some nonzero amount by which the consequences for \( i \) can be diminished, such that there is some (possibly large) increase in the consequences for \( i' \) that compensates for this.

We shall now render these consequentialistic intuitions more precise. Let \( A = \{a, a', a'', \ldots\} \) be a nonempty and finite set of alternative acts. We conceive of acts as particular acts rather than as act-types, and naturally all alternative acts should be performable by the same agent.

---

5 It should be emphasized that acts are choices about what total amount of good and bad consequences each individual will experience from now until he or she dies. That is, by choosing an act we choose each individual’s final amount of good and bad consequences.
Furthermore, let $I = \{i, i', i'', \ldots\}$ be a nonempty and finite set of individuals. In an ethical context such as this, it is reasonable to assume that the set of individuals includes all individuals in the world—not just some specific subset, such as white South Africans. For each pair of acts and individuals in $A \times I$ there are consequences $c_{ai}, c_{ai'}, \ldots$. The set of all possible consequences individuals may face is denoted $C$. It is not concrete states of affairs that matter from an ethical point of view for consequentialists (at least not according to the versions considered here). Instead, a ‘consequence’ should rather be conceived of as the entity that the consequentialist claims to be valuable, for example, preference satisfaction or amounts of hedonistic pleasure.

Causal and other relational aspects of consequences will not be modeled. The following example illustrates this point: suppose that the consequence no enemy soldiers killed is preferred to the consequence many enemy soldiers killed. Then, if no causal or other relational aspects are modeled, it will be reasonable to also assume that the consequence no enemy soldiers killed $\&$ war is lost is preferred to the consequence many enemy soldiers killed $\&$ war is lost, since we have added the same consequence to both initial consequences. Some (nonutilitarian) consequentialists, however, would probably consider it better that many enemy soldiers are killed in case the war is lost, which contradicts the assumption introduced above. Arguably, the best way to make sense of this noncausal and nonrelational notion of consequences is to think of consequences in a more abstract way than we usually do, for example, by focusing on their value properties, as suggested above.

In any axiomatic treatment of utilitarianism, one must at some point introduce the idea of interpersonal utility comparisons. One option is simply to assume that there is a cardinal function assigning real numbers to all consequences. This would not be entirely unreasonable, since adherents of different versions of consequentialism disagree about which rule should be used for evaluating a given distribution of consequences, not about the cardinal function itself. In the present exposition, however, we shall take another approach, and instead start from the assumption that the agent is able to make ordinal interpersonal comparisons. More precisely, we assume that the agent’s preferences among the elements of $C$ form a weak ordering. This means, among other things, that a consequence $c_{ai}$ either is better for $i$ (is preferred to) than consequence $c_{ai'}$ is for $i'$, or of

---

6 In the end, this assumption also implies that cardinal comparisons of interpersonal utility are possible, as will soon become evident, so this assumption is no radical improvement over Harsanyi’s.
equal value, or worse. Axiom 1 (extensive structure) can then be applied for constructing a cardinal utility function. An ethical problem can be conceived of as a triple $\delta = \langle A, I, C_{ai} \rangle$, in which $A$ and $I$ are defined as above and $C_{ai}$ is the set of consequences corresponding to $A \times I$, that is, all consequences that are obtainable for the individuals $I$ with the acts $A$. We then define a solution to an ethical problem as follows:

**Definition 1.** Let $\delta = \langle A, I, C_{ai} \rangle$. Then, $\Phi(\delta)$ is a solution to $\delta$ if and only if

(i) $\Phi(\delta) \subseteq A$, and

(ii) every member in $\Phi(\delta)$ is morally right to perform.

The solution $\Phi(\delta)$ is, of course, not supposed to be known in advance by the agent. If anything, it is the aim of normative ethics to determine the members of $\Phi(\delta)$. The triple $\langle C, \succ, \sim \rangle$ is a comparison structure for consequences, in which $\succ$ and $\sim$ are relations on $C$ representing strict preference respectively indifference.8 (As usual, $c \equiv c' \iff c \succ c' \lor c \sim c'$.) We assume that there is a binary operation $\cdot$ on $C$. Intuitively, $c \cdot c'$ means “both consequence $c$ and $c'$ obtain.” Furthermore, we wish to model the idea of an “interval” of consequences that have a bounded (small) negative value, for example, “Anne loses ten cents,” “Mark arrives one minute earlier to work than necessary,” and so forth. Consider the following definition.

**Definition 2.** Let $D(\subseteq C)$ be a set of bounded undesirable consequences just in case there are some $c, c' \in C$ such that

(i) for every $d \in D$ it holds that $c \succ d \succ c'$, and

(ii) for every $c'' \in C$ and every $d \in D$, if $c'' \succ d$ then $c'' \succ c'' \cdot d$, and

(iii) for every $d, d' \in D$ in $D$ there is some $d''$ such that $d \succ d'' \succ d'$.

Let us now consider the following ethical axioms, which I propose hold for every ethical problem $\delta$. The consequences $c, c', c'', c'''$ are arbitrary elements in $C$, and $C_{ai}$ is the concatenation of all consequences faced by $i$ in case $a$ is performed (in the ethical problem $\langle A, I, C_{ai} \rangle$).

**Axiom 1.** (Extensive Structure) For all $c, c', c'', c''' \in C$:

(i) $\langle C, \succ, \sim \rangle$ is a weak order.

---

7 As stated below, Axiom 1 will allow $C$ to be infinitely large. At the expense of considerable technical complexity, however, Axiom 1 could be relaxed such that only a set $C$ of more realistic size has to be considered. A good starting point for such a modification is to apply the results in David H. Krantz et al., in *Foundations of Measurement, Volume 1: Additive and Polynomial Representations* (New York: Academic, 1971), chapter 3.4 and 3.5.

8 The concept of comparison structures is introduced and investigated in detail by Sven Ove Hansson in *The Structure of Values and Norms* (New York: Cambridge, 2001).
(ii) \( [c \cdot (c' \cdot c'')] \sim [(c \cdot c') \cdot c'] \)

(iii) \( e \succ c' \iff (c \cdot c') > (c' \cdot c'') \iff (c'' \succ c) \)

(iv) If \( e \succ c' \), then there is a positive integer \( n \) such that \( (nc \cdot c'') \succ (nc' \cdot c'') \), where \( nc \) is defined inductively as: \( lc = c, (n + 1)c = nc \cdot c \).

Axiom 2. (Action Guidance) \( \Phi(\delta) \neq \emptyset \)

Axiom 3. (Split of Acts) Suppose that \( \delta = \langle A, I, C_{\alpha} \rangle \) and \( \delta' = \langle A', I, C_{\alpha} \rangle \), with \( A' = (A - \{a\}) \cup \{a', a''\} \). Then, if \( c_{\alpha} \sim c_{\alpha} \sim c_{\alpha} \) for all \( i \in I \), it holds that \( \{a', a''\} \subseteq \Phi(\delta') \) if and only if \( a \in \Phi(\delta) \).

Axiom 4. (Weak Pareto) If there is an \( a \) such that for some \( a' \) it holds that \( c_{a'} \succ c_{a} \) for all \( i \), then \( a \not\in \Phi(\delta) \).

Axiom 5. (Trade-Off) There is a set of bounded undesirable consequences \( D_{\alpha} \) such that for every \( d_{\alpha} \in D_{\alpha} \) there is a (possibly much better) consequence \( c_{d_{\alpha}} \) such that for all ethical problems \( \delta' \), if \( a \in \Phi(\delta) \) and \( c_{\alpha} \succ C_{\alpha} \), it holds that \( \Phi(\delta) = \Phi(\delta') \), where \( \delta' \) is the ethical problem obtained from \( \delta \) by substituting \( C_{\alpha} \) with \( C_{\alpha} \cdot d_{\alpha} \) and \( C_{\alpha} \cdot c_{d_{\alpha}} \).

Axiom 1 enables us to assign real numbers to consequences representing utility. This is the only axiom that is not a necessary consequence of utilitarianism. In principle, it could be replaced by any set of assumptions that allow for a numerical representation of utility (that is unique at least up to a positive linear transformation).

Axiom 2 would, of course, not be accepted by people who believe in moral dilemmas. Consequentialists, however, typically show little sympathy to that idea, and therefore this is not a serious problem here.

Axiom 3 asserts that if an act is split into two other acts, which both have exactly the same consequences as the original act, then the two new acts are morally right just in case the original act is morally right. In order to illustrate what it means to split an act into two, suppose that you have decided to go for a walk in the park. This activity can be performed in two ways, namely by carrying the umbrella in either your left or in your right hand. If the consequences of these two new acts are the same as the consequences of the original act, then the moral status of all three acts should be the same. Axiom 3 follows from the basic consequentialistic idea that consequences alone determine the moral status of an act.

Axiom 4 is a version of the well-known Pareto principle, which plays a central role in Harsanyi’s as well as in Broome’s axiomatizations.\(^9\) Note, however, that Axiom 4 is weaker than the traditional Pareto principle.

\(^9\) Axiom \( C \) in Harsanyi, "Bayesian Decision Theory, Rule Utilitarianism, and Arrow’s Impossibility Theory," p. 293. Broome discusses his use of the Pareto principle on p. 92 (op. cit.).
Axiom 5 is the most controversial of the five axioms. It asserts that a withdrawal of a small amount of utility from someone who is well off can be ethically compensated by giving some amount of utility (perhaps much more) to a worse-off individual. Of course, this axiom will not be accepted by someone subscribing to, for example, ethical egoism. One of Harsanyi’s axioms, however, is also incompatible with ethical egoism, namely Axiom D, which states that one should give equal weight to everyone’s utility functions (ibid., p. 294). Furthermore, it is also worth noticing that ethical egoism cannot even be formulated in the technical framework used in this paper, since it does not allow us to tell which \( i \in I \) is identical to the agent. Already this fact is, arguably, a partial reason for regarding ethical egoism as a pathological form of consequentialism.

It is trivial to show that Axioms 2 and 3 are necessary consequences of at least four versions of consequentialism, namely utilitarianism, prioritarianism, the maximin view, and egalitarianism.\(^\text{10}\) Axiom 4 (Weak Pareto) is also a necessary consequence of these versions of consequentialism, with the exception of the most extreme version of egalitarianism; people who are concerned with nothing except value differences do not have to accept Axiom 4 (or Harsanyi’s or Broome’s stronger versions of the Pareto principle). Parfit’s levelling down objection is a well-known argument to the effect that such an extreme form of egalitarianism is unreasonable (op. cit., pp. 17–18). Axiom 5, or the weaker Axiom 5’ stated in the next section, is entailed by utilitarianism, prioritarianism, egalitarianism, and the maximin view, given that there is a set of bounded undesirable consequences \( D\) as stated in Definition 2 and given that \( \mathcal{C} \) is fine-grained (see below). I leave to the reader verification of my claims about Axioms 2–4. For Axiom 5, however, a formal observation is motivated, since this proof requires a few technical assumptions.

\textit{Observation 1.} Suppose that \( \mathcal{C} \) is fine-grained (for every \( c_{ai} \succ c_{ai} \in \mathcal{C} \) there is some \( c_{ai}' \) such that \( c_{ai} \succ c_{ai}' \succ c_{ai} \)) and also suppose that there is a set of bounded undesirable consequences \( D(\subseteq \mathcal{C}) \). Then utilitarianism, prioritarianism, and egalitarianism all entail Axiom 5 or Axiom 5’, and the maximin view entails a version of Axiom 5 according to which \( d_{ai} \) and \( c_{ai} \) have to be chosen such that the worst consequence of the act in question is not diminished.\(^\text{11}\)

\(^{10}\) For technical formulations of these ethical theories, see the proof of Observation 1.

\(^{11}\) Expressed in technical terms, \( d_{ai} \) and \( c_{ai} \) should be chosen such that \( \mathcal{C}_{ai} \succ d_{ai} \succ \mathcal{C}_{ai} \) and \( \mathcal{C}_{ai} \succ c_{ai} \succ \mathcal{C}_{ai} \).
The version of Axiom 5 entailed by the maximin view is trivially implied by the original version of Axiom 5. Furthermore, since this version is also sufficient for deriving the theorem stated in the next section—I leave it to the reader to verify this\(^{12}\)—it can be reasonably maintained that the axiomatization constructed here is also relevant for advocates of the maximin view.

Thus, given that Axiom 1 can be accepted, or some other axioms that allow us to construct a cardinal utility function, it seems that we have a compelling argument in favor of utilitarianism, namely that this theory follows (see section III and the proofs) from premises that are accepted by advocates of all major positions in the debate about consequentialism.

### III. TWO UTILITARIAN THEOREMS

 Taken together, Axioms 1–5 imply that utility numbers can be assigned to all consequences, and that the solution to an ethical problem is the set of acts having the highest sum of utility.

**Theorem 1.** Let Axioms 1–5 hold for a set \( \Delta \) of ethical problems. Then there is a real-valued function \( u \) on \( C \) such that for every \( \delta \in \Delta \):

1. \( u(c \circ c') = u(c) + u(c') \) and \( u(c) > u(c') \iff c \succ c' \)
2. If \( u' \) is another function satisfying (1), then for some real number \( r > 0 \), \( u = ru' \).
3. \( \Phi(\delta) = \{ a : \sum_{i \in I} u(C_{ai}) \geq \sum_{i \in I} u(C_{a'i}) \text{ for all } a' \in A \} \)

Parts (1) and (2) of Theorem 1 assert that the utility of consequences are measurable on a ratio scale. Part (3) is the utilitarian property, recommending us to choose alternatives that maximize utility.

The strongest axiom employed in Theorem 1 is Axiom 5, as noted in the previous section. While that axiom requires the value of a good consequence compensating for a poor consequence to be entirely determined by the poor consequence in question, the following weakening (Axiom 5\(^{\prime}\)) allows the compensation to be a function of more factors, as long as it can be described by a positive monotone function.

**Axiom 5\(^{\prime}\). (Weak Trade-Off)** There is a set of bounded undesirable consequences \( D_{ai} \) such that for every \( d_{ai} \in D_{ai} \) and every ethical problem \( \delta \), there is a (much better) consequence \( c'_{ai} \) determined by some positive monotone function \( (\alpha, d_{ai} \text{ and } c'_{ai}) \) such that, if \( a \in \Phi(\delta) \) and \( C_{ai} \succ C_{a'i} \), it holds that \( \Phi(\delta) = \Phi(\delta') \), where \( \delta' \) is the ethical problem obtained from \( \delta \) by substituting \( C_{ai} \) for \( C_{ai} \circ d_{ai} \) and \( C_{a'i} \) for \( C_{a'i} \circ c'_{ai} \).

\(^{12}\) Note that the proof given below leaves the minimal level of act \( a'' \) in \( \delta'' \) at least as high as the minimal level of act \( a' \) in \( \delta'' \).
Theorem 2, stated below, shows that also Axioms 1–4 and 5’ are sufficient for obtaining the desired support for utilitarianism.

Theorem 2. Let Axioms 1–4 and 5’ hold for a set \( \Delta \) of ethical problems. Then there is a real-valued function \( u \) on \( C \) such that for every \( \delta \in \Delta \), (1)–(3) in Theorem 1 are satisfied.

In my view, axiomatic arguments are valuable mainly because they show how certain central ideas are connected to each other. If a set of axioms is less controversial than some theorem, then a proof that the theorem follows from the axioms indicates that either one should give up at least one of the axioms, or accept the theorem. As I see it, the conjunction of Axioms 1–5, respectively Axioms 1–4 and 5’, are less controversial than the utilitarian doctrine. Therefore, it is an interesting piece of information that utilitarianism can be deduced out of those axioms. As pointed out in the introduction, the present axiomatization differs from Harsanyi’s and his successors’ axiomatizations in at least two aspects. First, Harsanyi constructed his axiomatization by adding a set of axioms to Bayesian decision theory. That is, his ethical conclusion is derived from what is fundamentally a set of rationality requirements for individual decision making under risk. Anyone who considers that theory of rationality to be unacceptable, for example, because he or she feels that the much criticized independence postulate (or “the sure-thing principle”) lacks intuitive justification, will also refute Harsanyi’s argument for utilitarianism.\(^\text{13}\)

The second difference, which is perhaps less important, has to do with interpersonal utility comparisons. Harsanyi’s Axiom \( D \) requires that agents can judge whether the utility functions for a set of \( n \) individuals are expressed in equal utility units.\(^\text{14}\) This, he admitted, presupposes an ability to make cardinal comparisons of interpersonal utility. In the present axiomatization, we instead assume that ordinal comparisons of interpersonal utility are possible, and we show that a sufficiently rich structure of ordinal comparisons allows us to assign cardinal numbers (by applying Axiom 1). This is a well-known maneuver in measurement theory, and I do not want to claim that it is a radical improvement over Harsanyi’s proposal. (For example, we still lack a satisfactory theory about what sort of evidence can legitimately be taken to indicate that one person enjoys a higher utility level than another.) A small advantage of the suggested approach, however, is that it does not, unlike Harsanyi’s Axiom \( D \), require direct numerical

\(^{13}\) In Broome’s version of Harsanyi’s axiomatization, the most controversial Bayesian axiom is the impartiality axiom.

\(^{14}\) Harsanyi, “Bayesian Decision Theory, Rule Utilitarianism, and Arrow’s Impossibility Theory,” p. 294.
comparisons; the numbers used for describing different utility levels are derived from the axiomatization itself.

Royal Institute of Technology/Sweden

PROOFS

Proof of Observation 1. Utilitarianism: According to this theory, there is a function $u$ that assigns real numbers to all elements in $C$, and for every ethical problem $\delta$ it holds that $\Phi(\delta) = \{a: u(C_{a1}) + \ldots + u(C_{an}) \geq u(C_{a'1}) + \ldots + u(C_{a'n})\}$ for all $a, a' \in A$. In order to derive Axiom 5, choose for every $d_{a1} \in D_{a1}, a \in C_{a1}$ such that $u(d_{a1}) = u(C_{a1})$. (That there is such a consequence $C_{a1}$ follows from the assumption that $C$ is fine-grained.) Then the total sum of utility of the modified acts in $\delta$ and $\delta'$ will be equal; hence, these ethical problems have the same solution.

Prioritarianism: Prioritarians distinguish between utility and moral value; the moral value $v(a)$ of an act can be described by the equation $v(a) = r(u(C_{a1}) + \ldots + u(C_{an}))$, where $r$ is a strictly increasing and concave function. In order to derive Axiom 5 choose, for every $d_{a1} \in D_{a1}, a \in C_{a1}$ such that $r(u(C_{a1} \circ d_{a1})) = r(u(C_{a1})) + r(u(C_{a1}'))$. (That there is such a consequence $C_{a1}'$ follows from the assumption that $C$ is fine-grained.) Then the total sum of moral value of the modified acts in $\delta$ and $\delta'$ will be equal; hence, these ethical problems have the same solution.

Egalitarianism: According to this theory, the moral value $v(a)$ of an act can be described by the equation $v(a) = e(u(C_{a1}), \ldots, u(C_{an}))$, where $e$ is a function that is strictly increasing with respect to some appropriate measure of equality. In order to derive Axiom 5 choose, for every $d_{a1} \in D_{a1}, a \in C_{a1}$ such that $e(\ldots, u(C_{a1}), u(C_{a1}')) = e(\ldots, u(C_{a1} \circ d_{a1}), u(C_{a1}' \circ d_{a1}), \ldots)$. (That there is such a consequence follows from the assumption that $C$ is fine-grained.) Then the total sum of moral value of the modified acts in $\delta$ and $\delta'$ will be equal; hence, these ethical problems have the same solution.

The Maximin View: This view does not presuppose any cardinal measurement of consequences. As noted in Observation 1, we consider a slightly weakened version of Axiom 5, according to which $d_{a1}$ and $c_{a1}'$ are chosen such that $C_{a1} \circ d_{a1} \geq C_{a1} \circ c_{a1}' \geq C_{a1}'$. In order to derive this version of Axiom 5 it is sufficient to note that the maximin view singles out the set of morally right acts by just taking into consideration the worst consequences of each act; hence, since the worst consequences of each act were assumed to be unaffected, and we know that there is a set of bounded undesirable consequences,
any \( c_{a'\prime} \) that satisfies the condition stated in Observation 1 can be chosen.

**Proof of Theorem 1.** Parts (1) and (2) rely entirely on Axiom 1, and are essentially an application of a standard theorem in measurement theory that has been carefully proved by others. Therefore, we will just provide a rough sketch of the main idea.\(^{15}\)

In order to construct the function \( u \) we select a consequence \( c^* \) in \( C \), which will serve as the unit of the ratio scale. For any other \( c \) in \( C \) and for any positive integer \( n \), the Archimedean axiom 1:(iv) implies that there is an integer \( m \) such that \( mc^* > nc \). Let \( m_n \) be the smallest integer for which this is true, and observe that \( m_n c^* > nc \geq (m_n - 1) c^* \). Hence, \( m_n \) copies of the unit consequence \( c^* \) are approximately equal to \( n \) copies of \( c \). By selecting a larger and larger \( n \) the approximation becomes better and better, and, given that the limit exists, we can define \( u \) as:

\[
\lim_{n \to \infty} \frac{m_n}{n}
\]

In order to complete the proof, one has to show that the limit exists and that it satisfies the properties stated in (1) and (2).

Part (3) is easier to prove if we first prove the following lemma.

**Lemma 3.** (Equal Trade-Off) There is a closed interval of positive numbers \([ a, b ]\) such that for all \( \alpha \in [ a, b ] \) and for all ethical problems \( \delta \), if \( a \in \Phi(\delta) \) and \( u(C_{\omega}) > u(C_{\omega'}) \), it holds that \( \Phi(\delta) = \Phi(\delta') \), where \( \delta' \) is the ethical problem obtained from \( \delta \) by subtracting \( \alpha \) from \( u(C_{\omega}) \) and adding \( \alpha \) to \( u(C_{\omega'}) \).

Because of parts (1) and (2) in Theorem 1 we know that utility numbers can be attached to all elements in \( C \). Thus, by applying parts (1) and (2) to Axioms 3, 4, and 5, they can be restated in numerical terms. It is fruitful to state the numerical version of Axiom 5 as a trivial lemma. In Lemma 4 below, the concept of ‘bounded undesirable consequences’ in Axiom 5 has been replaced by a closed interval of real numbers, which property (1) allows us to do.

**Lemma 4.** There is a closed interval of positive numbers \([ a, b ]\) such that for all \( \alpha \in [ a, b ] \) there is a number \( \beta \) such that for all ethical problems \( \delta \), if \( a \in \Phi(\delta) \) and \( u(C_{\omega}) > u(C_{\omega'}) \), it holds that \( \Phi(\delta) = \Phi(\delta') \), where \( \delta' \) is the ethical problem obtained from \( \delta \) by subtracting \( \alpha \) from \( u(C_{\omega}) \) and adding \( \beta \) to \( u(C_{\omega'}) \).

We are going to prove Lemma 3 by showing that \( \alpha = \beta \) whenever

---

\(^{15}\) For full details, see, for example, Krantz (*op. cit.*, pp. 77–81); see also F. Roberts, *Measurement Theory*, Volume 7 of Gian-Carlo Rota, ed., *Encyclopedia of Mathematics and Its Applications* (Boston: Addison-Wesley, 1979). For the sketch provided here, see Krantz, p. 75.
Lemma 4 is applied to an ethical problem. Assume for reductio that \( \alpha \neq \beta \) as Lemma 4 is applied to the act \( a \) and the individuals \( i \) and \( i' \). Let us first consider the case in which \( i \) and \( i' \) are the only individuals in \( \delta \), and \( a \in \Phi(\delta) \). Let \( u(C_w) = u_i \) and let \( u(C_{w'}) = u_{i'} \). \( u_i > u_{i'} \). By applying Axiom 3 twice, we find that \( \delta \) can be transformed into an ethical problem \( \delta' \) with three new (identical) acts \( a', a'', a''' \) such that \( \Phi(\delta') = \{ \Phi(\delta) - a \} \cup \{ a' \cup a'' \cup a''' \} \).

We now apply Lemma 4 to \( a'' \) and \( a''' \) in \( \delta' \) exactly \( k \) times with some \( \alpha' \) and \( \beta' \), such that \( 0 < u_i - k_{\alpha'} - [u_{i'} + k_{\beta'}] < \alpha \). (It might hold that there are no \( k, \alpha', \beta' \) that satisfy this inequality, for example, if \( \beta' \gg \alpha' \). In that case, let \( k = 1, \alpha' = \alpha, \beta' = \beta \), and ignore the next sentence.) Then we apply Lemma 4 to \( a''' \) one more time, this time subtracting \( \alpha \) respectively, and adding \( \beta \). It now holds that \( u(C_{w''}) < u(C_{w'''}) \) in \( \delta'' \). Finally, we apply Lemma 4 again, but this time we subtract \( \alpha \) from \( u(C_{w'''}) \) and add \( \beta \) to \( u(C_{w''}) \).

\[
\begin{array}{cccc}
\delta' & \delta'' \\
\hline
i & i' & i & i' \\
\hline
\ldots & \ldots & \ldots & \ldots \\
a' & u_1 & u_2 & a' & u_1 & u_2 \\
a'' & u_1 & u_2 & a'' & u_1 - k_{\alpha'} & u_2 + k_{\beta'} \\
a''' & u_1 & u_2 & a''' & u_1 - k_{\alpha'} - \alpha & u_2 + k_{\beta'} + \beta - \alpha \\
\ldots & \ldots & \ldots & \ldots \\
\end{array}
\]

We initially assumed that \( \alpha \neq \beta \). Hence, \( -\alpha + \beta \neq 0 \). Thus, by Axiom 4, \( a'' \in \Phi(\delta'') \) or \( a''' \in \Phi(\delta'') \), which contradicts the result that \( \{ a', a'', a''' \} = \Phi(\delta') = \Phi(\delta'') \). Consequently, our initial assumption must be false.

Let us now consider the case in which \( a \not\in \Phi(\delta) \). Then, by Axiom 2, there is some other act \( b \) such that \( b \in \Phi(\delta) \). Furthermore, since \( \beta \) is a function of \( \alpha \), it follows that \( \alpha \) and \( \beta \) can be applied to \( u(C_{w}) \) and \( u(C_{w'}) \) in the same way as above, yielding the analogous contradiction.

Finally, in order to handle the general case with more than two individuals, it is sufficient to note that for every individual \( j \in \{ I \setminus \{ i \cup i' \} \} \), it holds that \( u(C_w) = u(C_{w'}) = u(C_{w''}) = u(C_{w'''}) \). Hence, the maneuver from \( \delta \) to \( \delta'' \) can be carried out in a similar way as above. Q.E.D. (Lemma 3).

We are now in a position to prove part (3) of Theorem 1. Let \( \delta \) be an arbitrary element in \( \Delta \). We then apply to \( \delta \) the function \( \text{to}^\ast \) (trade-off), which we define as the following function to iterated \( n \) times for some \( n \) such that \( \text{to}^n(\delta) = \text{to}^{n+1}(\delta) \). Let \( \text{high}(\delta, a) \) be a function that returns an ordered pair \( \langle a, i \rangle \) such that \( u(C_w) \geq u(C_{w'}) \) for every \( i' \), and \( \text{low}(\delta, a) \) a function that returns an ordered pair \( \langle a, i \rangle \) such that
\( u(C_a) \geq u(C_{a'}) \) for every \( i' \). Let \( \alpha \) be a small positive nonzero number obtained from Lemma 2 such that for every \( u(C_a) \) there is an even number \( t \) such that \( \alpha \cdot t = u(C_a) \). (In case there are no \( \alpha \) and \( t \) satisfying this condition, we also have to choose some additional number \( \alpha' < \alpha \), that we apply the “last time” utility is subtracted; that is, \( \alpha' \) should be chosen such that \( \alpha \cdot t + \alpha' = u(C_a) \).

\[
A_i \defeq A \setminus \{i, i'\}
\]

\[
\delta = \begin{cases} 
\delta & \text{if } u(C_a) = u(C_{a'}) \text{ for all } a, i, i' \\
\langle A', I', u' \rangle & \text{otherwise, with:}
\end{cases}
\]

(i) \( A' = A \)
(ii) \( I' = I \)
(iii) For some \( a \) such that \( h = \text{high}(\delta, a), l = \text{low}(\delta, a) \) and \( u(h) \neq u(l) \), for all \( r, t, v, w \):

\[
u'(c_a) = u(c_a) - \alpha \text{ if and only if } \langle a, i \rangle = h
\]

\[
u'(c_{a'}) = u(c_{a'}) + \alpha \text{ if and only if } \langle a, i' \rangle = l
\]

\[
u'(c_{a''}) = u(c_{a''}) \text{ for all } \langle a', i'' \rangle \neq h, l
\]

Since rule to* is obtained from an iterated application of Lemma 2 (equal trade-off), it holds that \( \Phi(\delta) = \Phi(\text{to}^*(\delta)) \). It follows from the definition of \( \alpha \) above that to* will converge such that \( u(C_a) = u(C_{a'}) \) for all \( a, i, i' \) in the ethical problem to* \( (\delta) \).

We now apply Axiom 4 and Axiom 2 to \( \Phi(\text{to}^*(\delta)) \), which gives us a set \( A^* = \{a: \sum_{i \in A} u(C_a) \geq \sum_{i \in A} u(C_{a'}) \text{ for all } a' \in A\} \). Hence, \( \Phi(\delta) = \Phi(\text{to}^*(\delta)) = A^* = \{a: \sum_{i \in A} u(C_a) \geq \sum_{i \in A} u(C_{a'}) \text{ for all } a' \in A\} \). Q.E.D. (Theorem 1).

Proof of Theorem 2. We prove Theorem 2 by showing that Theorem 1 holds even if Axiom 5 is substituted by Axiom 5', that is, that property (3) of Theorem 1 follows from Axioms 1–4 and 5'. As before, we apply property (1) and (2) of Theorem 1 and obtain the following numerical version of Axiom 5':

Lemma 5. There is a closed interval of positive numbers \([a, b]\) such that for all \( \alpha \in [a, b] \) and for all ethical problems \( \delta \), there is a (much larger) number \( \beta \) determined by some monotone function \((\alpha \text{ and } u_a, u_{a'})\) such that if \( a \in \Phi(\delta) \) and \( u(C_a) > u(C_{a'}) \), it holds that \( \Phi(\delta) = \Phi(\delta') \), where \( \delta' \) is the ethical problem obtained from \( \delta \) by subtracting \( \alpha \) from \( u(C_a) \) and adding \( \beta \) to \( u(C_{a'}) \).

It will be sufficient to prove Lemma 3, since the proof of property (3) of Theorem 1 only relies on that lemma in conjunction with Axiom 4. Lemma 3 is proved by showing that \( \alpha = \beta \) whenever Lemma 5 is applied to an ethical problem. Assume for reductio that \( \alpha \neq \beta \) as Lemma 5 is applied to the act \( a \) and the individuals \( i \) and \( i' \). Let us first consider the case in which \( i \) and \( i' \) are the only individuals in
\( \delta \), and \( a \in \Phi(\delta) \). Let \( u(C_{a_i}) = u_1 \) and let \( u(C_{a_j}) = u_2; u_1 > u_2 \). By applying Axiom 3 twice, we find that \( \delta \) can be transformed into an ethical problem \( \delta' \) with three new (identical) acts \( a', a'', a''' \) such that

\[
\Phi(\delta') = \{ \Phi(\delta) - a \} \cup \{ a' \cup a'' \cup a''' \}.
\]

We now apply Lemma 5 to \( a'' \) and \( a''' \) in \( \delta' \) exactly \( k \) times with some \( \alpha' \) and \( \beta', \beta'', \ldots \), such that \( 0 < u_1 - k\alpha' - [u_2 + (\beta' + \beta'' + \ldots)] < \alpha \). The difference between \( u_1 \) and \( u_2 \) is now very small. (If, for example, \( \beta' \gg \alpha' \) it might hold that there are no \( k, \alpha', \beta', \beta'' \) that satisfy this inequality; in that case, let \( k = 1, \alpha' = \alpha, \beta' = \beta \), and ignore the next sentence.) Then we apply Lemma 5 to \( a'' \) one more time, this time subtracting \( \alpha \) respectively adding \( \beta \). It now holds that \( u_{a'''} < u_{a''} \) in \( \delta' \).

\[
\begin{align*}
\delta' & \quad i & i' \\
\ldots & \quad & \\
a' & u_1 & u_2 \\
a'' & u_1 - k\alpha' & u_2 + (\beta' + \beta'' + \ldots) \\
a''' & u_1 - k\alpha' - \alpha & u_2 + (\beta' + \beta'' + \ldots) + \beta \\
\ldots & \quad & \\
\delta'' & \quad i & i' \\
\ldots & \quad & \\
a' & u_1 & u_2 \\
a'' & u_1 - k\alpha' & u_2 + (\beta' + \beta'' + \ldots) \\
a''' & u_1 - k\alpha' - \alpha + \beta_0 & u_2 + (\beta' + \beta'' + \ldots) + \beta - \alpha \\
\ldots & \quad & \\
\end{align*}
\]

We now apply Lemma 5 to \( a''' \) again and subtract \( \alpha \) from \( u_{a'''} \) and add \( \beta_0 \) to \( u_{a'''} \). This yields the ethical problem \( \delta'' \). Since \( \beta \neq \alpha \) there are two cases to consider, namely, \( \alpha > \beta \) and \( \beta > \alpha \). Let us first consider the case when \( \beta > \alpha \). Then, \( \beta_0 > \beta > \alpha \), because of monotonicity. Hence, \( -\alpha + \beta_0 > -\alpha + \beta > 0 \). Thus, by Axiom 4, we have \( a'' \notin \Phi(\delta'') \) or \( a''' \notin \Phi(\delta'') \), which contradicts the result that \( [a', a'', a'''] = \Phi(\delta') = \Phi(\delta'') \). Consequently, our initial assumption must be false. The case when \( \alpha > \beta \) is handled in the analogous way.

The case in which \( a \notin \Phi(\delta) \) is handled exactly the same as in Theorem 1, as is the general case with more than two individuals. Q.E.D. (Theorem 2).