Might Theory X Be a Theory of Diminishing Marginal Value?[*]

Abstract

Act Utilitarianisms divide into Total and Average versions. Total versions seem to imply Parfit's "Repugnant Conclusion". Average versions are proposed in part to avoid the Repugnant Conclusion, but these are subject to "Mere Addition" arguments as detailed by Hudson in "The Diminishing Marginal Value of Happy People". Thus, various intermediate versions of utilitarianism, such as the one investigated by Hurka in "Value and Population Size", take on interest. But Hudson argues that such compromise theories are subject to the mere addition arguments, and are therefore no improvement over Average Utilitarianism. I disagree: some such compromise solutions escape Hudson's Mere Addition arguments.

1. Totalism

Utilitarianisms require us to maximize value. We can develop this basic idea in the following way.[1] At any time, some possible worlds are *accessible* to me. These worlds may be good or bad to various degrees. Intuitively, act utilitarianism claims that I ought to behave as I do in the best of these worlds. Equivalently:

> AU: it is right at some time t for S to see to p iff p is true in some world w such that i) w is accessible to S at t, and ii) there is no world w' which is accessible to S as of t, and better than w.

AU is incomplete, for we have said nothing about which worlds are better than others. Only when coupled with a theory of the *better than* relation for possible worlds does AU generate results. The true theory of this relation, "Theory X", has proved to be elusive.[2]

One family of such theories is familiar. Assume that for every person x existing at world w, there is a number, $U_{x,w}$, which represents her *life utility* at

---

[1]See Feldman (1986, chapter 2).

[2]Parfit introduces this expression in Parfit (1984, p. 361).

w—all of the utility ever experienced by her at w.[3] This may involve happiness, preference satisfaction, or something else; let us remain neutral. Then, the *total value* of a world w is given by summing the life utilities over $P(w)$, the set of all people who ever exist at w:

$$\mathrm{TV}(w) = \sum_{x \in P(w)} U_{x,w}$$

Totalism is the claim that possible world w_1 is better than possible world w_2 iff $\mathrm{TV}(w_1) > \mathrm{TV}(w_2)$. *Total Act Utilitarianism* (TAU) is simply the conjunction of totalism and AU.

The chief problem with this approach is that totalism implies the dreaded "Repugnant Conclusion". Consider world w_1, which contains a billion people each with a very high life utility. Let k be a very small positive number that represents a life of drudgery—a life with life utility k is worth living, but just barely. Clearly there is a world w_2 in which everyone has a life utility of k, but which, because of its enormous population, has a higher total value than w_1. Totalism implies the *Repugnant Conclusion*: w_2 is better than w_1. But, many think, it is plain that w_1 is the better world.

2. Averagism

In the face of the Repugnant Conclusion we may retreat to "averagism". The *simple average value* of a world w, $\mathrm{SAV}(w)$, is the result of dividing $\mathrm{TV}(w)$ by the total number of people that ever exist at w. *Simple averagism* is the thesis that w_1 is better than w_2 iff $\mathrm{SAV}(w_1) > \mathrm{SAV}(w_2)$. Finally, *Simple Average Act Utilitarianism* (SAAU) is the conjunction of simple averagism and AU. According to SAAU, I ought to see to it that proposition p is true iff p is true in all the worlds accessible to me which have the highest average utility per person.

A virtue of simple averagism is that it does not weight *quantity* of utility exclusively. A vice is that it does not weight quantity enough. Consider the scenario of Adam and Eve:[4] Adam and Eve are alone in the world living very happy lives, and they face a choice: should or should they not create one hundred people, each of whom will have a life utility of +99? Adam and Eve

---

[3]I assume the zero point of the life utility function to be the point at which life is neither worth living nor worth ending.

[4]This is Derek Parfit's example (1984, p. 420).

will each have a life utility of +100 regardless of their choice. SAAU has the consequence that it would be wrong for Adam and Eve to create the people, since the simple average value of the world accessible to them in which they *do* create the people is about +99, whereas the world in which they don't create them has a simple average value of +100. But surely it would not be wrong for Adam and Eve to create the people simply because they will be below average, utility-wise. The new people would lead lives well worth living. Quantity seems to have *some* value, particularly when the population is so low. SAAU ignores the value of quantity.

3. Compromise Theories

Totalism and TAU seem to run into trouble at high population levels. SAAU has problems at low population levels. This has suggested to some writers that a compromise theory might be true.[5] By a "*compromise axiological theory*" I will mean a theory about the values of possible worlds that:

    i) avoids the Repugnant Conclusion by allowing the value contributed by additional happy people at a given utility level to approach zero as the population grows, and

    ii) holds that at low population levels the values of worlds approximate their total values, and that at high population levels values approximate simple average values.
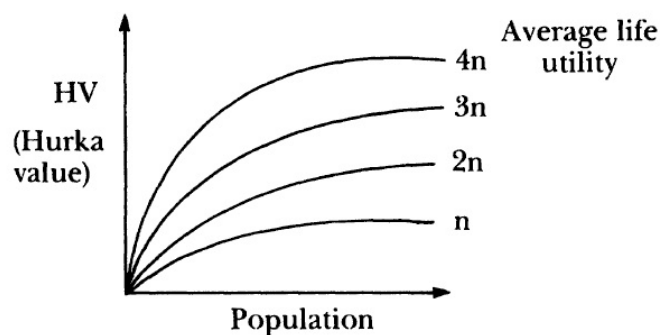
A *compromise act utilitarian theory* is the conjunction of a compromise axiological theory and AU.

    Thomas Hurka attempts to garner intuitive support for the idea that the value of members of an animal species decreases as the total number of members of that species increases.[6] He notes that we will expend a lot of energy to save the lives of the last few members of an endangered species simply because there are few of them. Perhaps this carries over to people. If there are few people, we may raise the value of the world by adding more, but if there are many, we must increase the average quality of life to bring about large changes in the value of the world. People have "diminishing marginal value".

---

[5]Parfit (1984, sections 137 and 138) suggests and rejects related views (views "(5)" and "(6)"). Hurka (1983) proposes a compromise view.

[6]Hurka (1983, pp. 496,7).

There are many compromise axiological theories, since there are many mathematical functions which generate the appropriate results. Hurka proposes one compromise axiological theory.[7] He asks us to consider a population graph representing the dependence of value on population size, for various simple average values:

As I understand it, the theory is this. For any world w, calculate the average utility of the world, and find the appropriate curve for that average on The Graph. Find the point on that curve whose x-value is the population of w. The *Hurka Value* of the world, $HV(w)$ is the y-value of that point. *Hurka's Theory* is that a world w_1 is better than a world w_2 iff $HV(w_1) > HV(w_2)$. Hurka's Act Utilitarianism (HAU) is the conjunction of Hurka's Theory and AU.

Hurka's Theory is a compromise axiological theory. It lacks some of the defects of TAU and SAAU, for it does not imply the Repugnant Conclusion and it avoids the problem of Adam and Eve. But both Hurka's Theory and simple averagism imply the following:

(MA1) sometimes the world can be made worse by simply adding a new person with a life utility that is positive, but below the simple average level.

Though we cannot tell this simply by looking at The Graph, Hurka assures us that his theory implies (MA1).[8] Presumably, this is because we move to a lower curve when we add the person, thereby outweighing the utility she contributes.

---

[7] 1983, pp. 501–503. I only consider his simpler version, "V1".
[8] 1983, p. 505.

Hurka suggests that this consequence is acceptable.[9] However, James Hudson has a convincing argument to show that this is not so.[10] We will need the concept of "mere addition". Suppose that I create exactly one person with a positive life utility. Further, suppose that neither my doing this nor the person's existence affects the utilities of others. Finally, suppose that, when I do so, my only other option for action was to do nothing. My time could not have been better spent. This act of mine is one of *mere addition*.[11]

Hudson's argument is clearest when we consider its application to SAAU. Suppose I commit an act of mere addition and create an utter hermit who interacts with no one. Let us see how the morality of this act is affected by the *ancient Egyptians*.[12] If there are very many Egyptians and they live much better lives than my hermit, then worlds accessible to me in which I don't create the hermit have higher simple average values than the actual world, where I do create the hermit. The hermit would be dragging down the average of the actual world. But if there are many ancient Egyptians with extremely *low* life utilities, then the actual world has the higher average. The hermit, being above average, would boost the average. SAAU says that in the first case my action is wrong, and in the second case it is right. So if SAAU were true, then the moral status of my action depends on the existence and welfare of the ancient Egyptians. But surely this is absurd.

SAAU falls to the case of the hermit. Hurka's Act Utilitarianism is refuted by this case as well, since it implies (MA1).[13] This is of little interest in itself, since Hurka's Act Utilitarianism is but one of many compromise theories. But Hudson makes a much bolder claim, that *every* compromise act utilitarian theory is refuted by the case of the hermit:

> Let us construct a Compromise theory which will be like [TAU] when the population is small and like [SAAU] when population is large... [The

---

[9]1983, pp. 505–507.

[10]His argument is in Hudson (1987, pp. 129–130), and is a simplified version of Parfit's "mere addition" problem (1984, section 143).

[11]It should be noted that this concept of mere addition differs slightly from Parfit's (1984, p. 420).

[12]The example is based on Hudson's example (see previous note) and one of Parfit's (1984, p. 420).

[13](MA1) states that *sometimes* mere addition can lower the value of a world. Let w be such a world. Now let w' be exactly like w, but with ancient peoples with *very* low utilities. Mere addition would *not* lower the value *of w'*. An ancient Egyptians argument may be based on this pair of worlds.

> Compromise theory] falls prey to all the objections to [SAAU]... The prospect of producing a variant theory that avoids [the Repugnant Conclusion] is indeed attractive; but not if the new theory has to bear the crushing weight of the objections to [SAAU]...[14]

The idea of constructing compromise theories is attractive, so Hudson's claim that all such theories are refuted by the hermit argument is important. However, it is demonstrably false, for there is a compromise act utilitarian theory that implies the following:

(MA2) Mere addition is always right

4. Geometrism

Consider any world w. Group the population of w into two ordered sets:

> $G : 1 \ldots i \ldots n$ (the people with positive or zero lifelong utility, *in order of descending utility*—in the case of ties, any order for those tied will suffice)
>
> $B : 1 \ldots j \ldots m$ (the people with negative lifelong utility, in order of ascending utility)
>
> let $u_i =$ the life utility of the ith person from G
>
> let $v_j =$ the life utility of the jth person from B
>
> let r be some real number greater than 1

Let the *geometric value* of w be defined:

$$\text{GV}(w) = \sum_{i=1}^{n} \frac{u_i}{r^{i-1}} + \sum_{j=1}^{m} \frac{v_j}{r^{j-1}}$$

Geometrism says that a world w_1 is better than a world w_2 iff $\text{GV}(w_1) > \text{GV}(w_2)$. (Note: different choices for r yield different *better than* relations on the possible worlds. The most plausible versions will choose r close to one.[15]) Geometrism is similar to Hurka's Theory, particularly for cases where all those existing at

---

[14] 1987, pp. 131,133.

[15] r represents how quickly the values of additional happy people approach zero. The larger r is, the more quickly these decline.

a world have the same life utility. The Graph gives a rough representation of the geometric values for these worlds. *Geometric Utilitarianism* (GU) is the conjunction of AU and *geometrism*.

First, we note that geometrism is a compromise axiological theory. Since the life utilities of happy people get divided by r^j for increasingly large j if the population is large, these quotients approach zero at high populations. And it is easy to show that geometrism does not imply the Repugnant Conclusion (see Appendix). Furthermore, at low populations, if r is chosen close to one, then people's life utilities are divided by numbers close to one; the GVs of such worlds are close to their total values. Finally, as a simple example of how GV behaves at high population levels, consider worlds in which everyone has the same life utility. For any number k, the GVs of worlds containing people whose life utilities are all k will approach $k/(1 - \frac{1}{r})$ in the high population limit.[16] Thus, for large worlds of this kind value is nearly proportional to the average utility; geometrism approximates simple averagism in the high population cases.

Thus, geometrism meets the conditions for being a compromise axiological theory, and so GU is a compromise act utilitarian theory.

Second, we note that GU implies (MA2). Consider any world w_1 containing a group of people. Suppose there is another world w_2 with the same utilities distributed over the same people, but which contains one extra person x. It may be shown that $GV(w_2) > GV(w_1)$ iff the life utility of x is positive (see Appendix). So GU implies (MA2).

5. Conclusion

Hudson's claim that all compromise theories are refuted by the hermit example is false, since GU is such a theory and it avoids this case.[17] Furthermore, GU avoids both the Repugnant Conclusion and the case of Adam and Eve.

I do not advocate GU, for it has other problems.[18] Nonetheless, the point

---

[16]See Arfken (1985, 278–279) for a treatment of the geometric series.

[17]Hudson has other versions of the Hermit argument directed against different, more complicated versions of averagism (1987, p. 130). Presumably, he thinks that compromise theories constructed in ways analogous to the different averagistic theories will fall to the corresponding Hermit arguments. But versions of geometrism could be constructed to parallel these averagistic theories, and these versions would escape the arguments.

[18]For example, it generates rather extreme results with respect to distributive justice. Furthermore, it succumbs to a variant of the Hermit argument. Suppose I create two hermits: the

still stands: Hudson has not shown that the idea of constructing compromise theories is a bad one. Perhaps Theory X is just around the bend, and is a compromise axiological theory.

Appendix

A. Geometrism avoids the Repugnant Conclusion

Consider any world with n people all with life utilities of k, where $k > 0$. The GV of such a world is the sum from 1 to n of k/r^{i-1}, where $r > 0$. Note that this series being summed is a geometric series. The *infinite* geometric series is convergent with sum $k(1 - \frac{1}{r})$, hence the sum of this finite geometric series cannot exceed this limit.[19] Hence, the value of such a world cannot be made arbitrarily large merely by adding to the population. The Repugnant Conclusion is avoided.

B. Geometrism and the addition of persons to worlds

Consider any sequence $\langle a_i \rangle$ of length n given by $a_i = u_i/r^{i-1}$, such that $r > 1, u_i \geq 0$, and $u_{i+1} \leq u_i$ for all i. Let $\langle a_i' \rangle$ be a new sequence of length $n+1$ given by: $\forall i (1 \leq i < k \rightarrow a_i' = a_i)$, $a_k' = u/r^{k-1}$, and $\forall i (k < i \leq n+1 \rightarrow a_i' = u_{i-1}/r^{i-1})$, where u is some number > 0, and k is the least number such that $u > u_k$ (let $k = n+1$ if there is no such number). It is easily verified that

(i) $\langle a_i' \rangle$ is the result of "inserting" the term u/r^{k-1} into the original sequence in a place that preserves the descending order of the numerators of the terms of the sequence, and

(ii) $\sum a_i' > \sum a_i$

Part (i) follows immediately from the definition of k, and the fact that $u_{i+1} \geq u_i$. To show (ii), note the following. The sequences are identical until k; by definition of k, $u > u_k$ so $a_k' > a_k$; from k to n, since $\langle u_i \rangle$ is a positive descending

---

welfare of the ancient Egyptians will affect the relative weightings of the life utilities of the hermits.

[19]See Arfken (1985, pp. 280,1) for details on the geometric series.

sequence $a'_i = u_{i-1}/r^{i-1} \geq u_i/r^{i-1} = a_i$; finally, the a' sequence has an extra (positive) term, a'_{n+1}.

Let us apply this result to mere addition. Let us compare a world w_1 with a world w_2 that differs utility-wise from w_1 only in that it has one extra person, with life utility u. Recall that the GV of a world is the sum of two summations, the sum representing the G people (the "G-sum") and the "B-sum". We may now show our desired result:

$$\mathbf{GV}(w_2) > \mathbf{GV}(w_1) \textbf{ iff } u > 0$$

If $u = 0$, then obviously $\mathrm{GV}(w_1) = \mathrm{GV}(w_2)$.

If $u > 0$, then let $\langle a_i \rangle$ represent the terms of the G-sum of w_1; by i) $\langle a'_i \rangle$ contains the terms of the G-sum of w_2, and by ii), the G-sum of w_2 is greater. The B-sum of w_2 is the same as that of w_1, so $\mathrm{GV}(w_2) > \mathrm{GV}(w_1)$.

If $u < 0$, then factor out -1 from each term in the B-sum of w_1. Let $\langle a_i \rangle$ represent the remaining sequence (all of the terms are positive); by (i) $\langle a'_i \rangle$ is a positive descending sequence containing the terms of $\langle a_i \rangle$ as well as $(-u/r^{k-1})$. Therefore, the terms of this sequence multiplied by -1 are the terms of the B-sum of w_2. By (ii), $\sum a'_i > \sum a_i$; therefore the B-sum of w_2 is *lower* than the B-sum of w_1. The G-sums are identical, so $\mathrm{GV}(w_2) < \mathrm{GV}(w_1)$.

References

Arfken, George (1985). *Mathematical Methods for Physicists*. Orlando, FL: Academic Press.

Feldman, Fred (1986). *Doing the Best We Can*. Dordrecht: D. Reidel.

Hudson, James (1987). "The Diminishing Marginal Value of Happy People." *Philosophical Studies* 51: 123–37.

Hurka, Thomas (1983). "Value and Population Size." *Ethics* 93: 496–507.

Parfit, Derek (1984). *Reasons and Persons*. Oxford: Clarendon.