Population Axiology

Ryan Doody

February 16, 2022

The Search for Theory X

Because of the Non-Identity Problem, Parfit concludes that our best theory of beneficence—what he calls **Theory X**—will not take a person-affecting form: it will not (solely) appeal to what is good or bad for those people whom our acts affect.
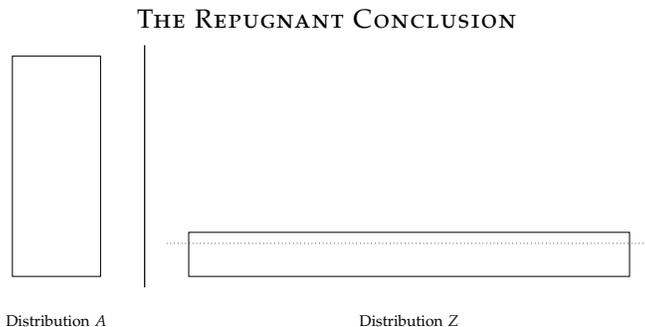
What is **Theory X** given that our actions can affect both *who* as well as *how many* people will exist?

...if we believe that causing to exist doesn't benefit.

We might also want **Theory X** to explain *The Asymmetry*:

○ We have a moral duty *not* to have the Wretched Child.

○ We lack a moral duty to have the Happy Child.

Is there a moral theory that solves the Non-Identity Problem, avoids the Repugnant Conclusion, and explains the Asymmetry?

Totalism & Averagism

Totalism: X is better than Y if, and only if, *total* well-being in X is higher than *total* well-being in Y. They're equally good if the totals are the same.

$$V_{\text{TOTAL}}(X) = \sum_i^n v_i(X)$$

The problem with Totalism is that it entails:

The Repugnant Conclusion

Distribution *A*   Distribution *Z*

The Repugnant Conclusion (RC): For any population with sufficiently many people, all with a high quality of life, there must be some much larger populations whose existence, other things equal, would be better, even though its members have lives that are barely worth living.

One way to avoid the Repugnant Conclusion is by evaluating populations in terms of their *average* level of well-being.

Averagism: X is better than Y if, and only if, *average* well-being in X is higher than *average* well-being in Y. They're equally good if the averages are the same.

$$V_{\text{AVERAGE}}(X) = \frac{\sum_i^n v_i(X)}{n}$$

But Averagism has serious problems too. It entails:

The Sadistic Conclusion (SC): It can be better to add persons with lives that aren't worth living than to add persons with positive, but below average, levels of well-being.
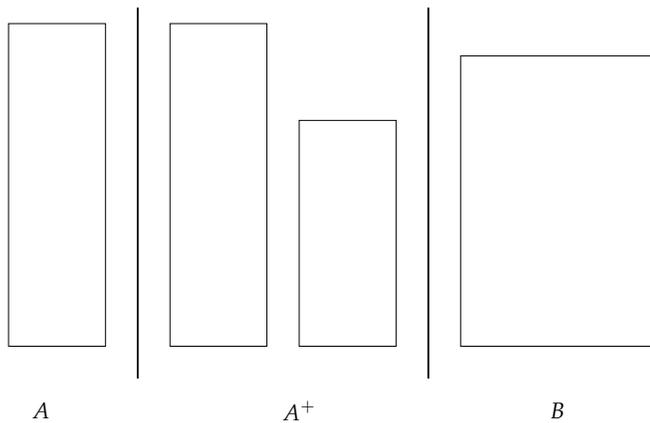
THE SADISTIC CONCLUSION
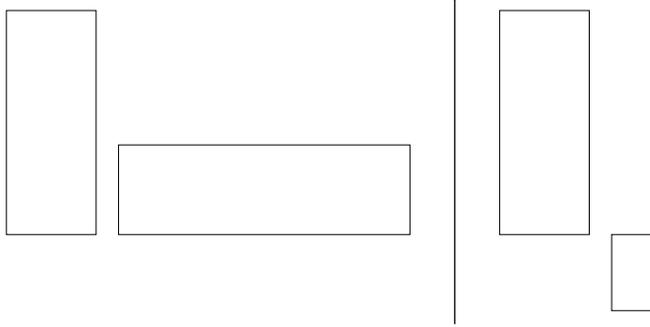
Can you think of other problems with Averagism? Totalism? Are either of these views plausible candidates for Parfit's **Theory X**?

The Mere Addition Paradox

Even if you reject Totalism, it is hard to resist the Repugnant Conclusion. It follows from a couple of plausible claims.

THE MERE ADDITION PARADOX

A	*A*⁺	*B*

According to the Mere Addition Principle, A^+ is not worse than A. According to Non-AntiEgalitarianism, B is better than A^+. So, by transitivity, B is not worse than A.

Non-Anti-Egalitarianism: If (1) X and Y contain the same people, (2) Y has higher total (and, thus, average) well-being than X, and (3) everyone in Y is equally well-off, then Y is better than X.

We can repeat the argument above, starting with population B, to show that population C is not worse than B, and then that population D is not worse that C, and then that ...

Averagism also violates:

The Mere Addition Principle. For any population X, let Y be a population just like X except that some additional people with lives worth living exist in Y who don't exist in X. Then Y is not worse than X.

If the additional lives are below the average, the resulting population will have a lower average level of well-being.

There are other possible views, too. Here are some.

○ *Variable Value Views.* Adding a person with a fixed well-being level has diminishing marginal value.

$$V_{\text{VARIABLE}}(X) = V_{\text{AVERAGE}} \cdot g(n)$$

It avoids RC, but not SC.

○ *Critical Level Views.* Adding a person increases the value of the world provided her well-being is above some "critical" threshold (α).

$$V_{\text{CRITICAL}}(X) = \sum_{i}^{n} v_i(X) - \alpha$$

It either avoids RC but not SC, or *vice versa*, depending on how high α is set.

○ *Person-Affecting Views:* e.g., Presentism (only presently existing people matter), Actualism (only actually existing people matter), Necessitarianism (only people existing in all the worlds under consideration matter).

○ *Harm-minimization Views.* The best world is the one with the lowest total comparative harm, summed over all persons who exist in that world.

... until we reach population Z: a population of very many people all living lives only barely worth living. By transitivity, A is not worse than Z: *The Repugnant Conclusion.*