Ethics & Infinity

December 26, 2017

A Puzzle for Aggregative Ethics with Infinity

Consider the following two plausible principles.

Pareto: If worlds w_x and w_y contain all the same people, and, if everyone is at least as well-off in w_x as they are in w_y, and there is someone is who strictly better-off in w_x, then w_x is better than w_y.

Anonymity: If worlds w_x and w_y contain none of the same people, and, if there's a one-to-one mapping from the former to the latter such that each person in w_x is at least as well-off as their counterpart in w_y, then w_y isn't strictly better than w_x.

If the worlds contain an infinite number of people, however, we run into a problem.

People:	...	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	...
w_A:	...	1	1	2	1	1	2	1	1	2	...
w_A^*:	...	1	2	2	1	2	2	1	2	2	...

People:	...	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9	...
w_B:	...	1	2	1	2	1	2	1	2	1	...

By **Pareto**, w_A^* is better than w_A. By **Anonymity**, w_B is as good as w_A^* and is as good as w_A. This violates Transitivity.

The Extensionist Program (Kagan & Vallentyne)

$w1$:	...	1	1	1	1	1	1	1	1	1	...
$w2$:	...	2	2	2	2	2	2	2	2	2	...

Example 1

Basic Idea: If worlds w_x and w_y have all the same locations, and, if relative to any finite set of locations, w_x is better than w_y, then w_x is better than w_y.

The **Basic Idea** is very weak. In particular, the principle is silent about the overall goodness of the worlds in the following example.

$w2$:	...	2	2	2	2	2	2	2	2	2	...
$w3$:	...	1	3	1	1	1	1	1	1	1	...

Example 2

Compare this principle to another: *The Pareto Principle.*

It says, roughly, that if everyone is at least as well-off in w_x as in w_y, and someone is strictly better-off in w_1 than in w_2, then w_x is better than w_y.

In Example 1, w_2 *Pareto-dominates* w_1 (assuming that the "locations" are people). In fact, the former *strongly* Pareto-dominates the latter. What's the relationship, in general, between **Basic Idea** and the Pareto Principle?

But, intuitively, $w2$ is better than $w3$. (Notice: for any finite set of locations containing at least three locations, $w2$ will be better than $w3$.)

SBI1: If worlds w_x and w_y have all the same locations, and, for any finite set of locations, there is a finite expansion such that for all further expansions, w_x is better than w_y, then w_x is better than w_y.

One world can be better than another *even if it is worse at a finite number of locations* so long as it sufficiently better at the other locations to compensate.

SBI1 (Strengthened Basic Idea 1) is still fairly weak. Consider the following example, where the "locations" are times, and your choosing between living two different immortal lives.

Times:	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	$\dots$
$w4$:	5	1	5	1	5	1	5	1	5	$\dots$
$w5$:	3	2	3	2	3	2	3	2	3	$\dots$

EXAMPLE 3

Intuitively, $w4$ is a better immortal life than $w5$. But there are finite expansions such that for all further expansions, $w4$ is better; and there are finite expansions such that for all further expansions, $w5$ is better. So, **SBI1** is silent in this case.

Why think that $w4$ is better than $w5$? Any time after the second day (t_2), you'll have enjoyed more well-being overall in $w4$ than $w5$. But what if the lives are immortal in both directions?

If the lives are immortal in both directions — everlasting in both the past and the future — there is no time at which the immortal has enjoyed a great amount of overall well-being at that point. At any point, the immortal has enjoyed an infinite amount of well-being in both worlds.

Times:	$\dots$	t_{-2}	t_{-1}	t_0	t_1	t_2	t_3	t_4	t_5	t_6	$\dots$
$w4^*$:	$\dots$	5	1	5	1	5	1	5	1	5	$\dots$
$w5^*$:	$\dots$	3	2	3	2	3	2	3	2	3	$\dots$

EXAMPLE 3.1

SBI2: If worlds w_x and w_y have all the same locations, and, for any bounded region of expansions, there is a bounded regional expansion such that for all further bounded regional expansions, w_x is better than w_y, then w_x is better than w_y.

A *bounded region* is a set of locations that are all inside a boundary. In order for this notion to be well-defined, the locations must have some natural ordering. Plausibly, times and spatial locations have such an ordering. People, on the other hand, do not.

Problems for the Program

1. **Only Partially Orders Worlds.** There are very many cases in which Kagan & Vallentyne's proposal remain silent. For example, what about cases in which the "locations" are not the same?

2. **Requires a Natural Ordering of Locations.** Some "locations" of value (e.g. people) have no natural ordering. Even among those sets of locations that do have a natural ordering (e.g. time and space), why think that it's ethically relevant?

Does **SBI2** conflict with *temporal neutrality*, for example?

3. **No Cardinal Rankings.** This approach (at best) tells us whether one world is better (or worse, or as good as) another. We cannot say *how much* better (or worse). We cannot, then, evaluate actions in terms of their *expected value*.