

Too Much of a Good Thing: Decision-Making in Cases with Infinitely Many Utility Contributions

Christopher J. G. Meacham

Forthcoming in Synthese

Abstract

Theories that use expected utility maximization to evaluate acts have difficulty handling cases with infinitely many utility contributions. In this paper I present and motivate a way of modifying such theories to deal with these cases, employing what I call “direct difference taking”. This proposal has a number of desirable features: it’s natural and well-motivated, it satisfies natural dominance intuitions, and it yields plausible prescriptions in a wide range of cases. I then compare my account to the most plausible alternative, a proposal offered by Arntzenius (2014). I argue that while Arntzenius’s proposal has many attractive features, it runs into a number of problems which direct difference taking avoids.

1 Introduction

Consider theories of permissible action that tell you to maximize expected utility, where the utility of an outcome is taken to be the sum of its good-making features. Such theories include versions of decision theory, subjective utilitarianism, and the like.

There are four sources of infinity problems for such theories.¹

1. One source of problems comes from cases with infinitely many acts. For example, consider a case in which you can choose any natural number, and receive that much utility. Since there is no largest natural number, there is no act that maximizes expected utility. Thus there is no permissible act.
2. A second source of problems comes from cases with infinitely many states; i.e., cases in which particular acts can have infinitely many possible outcomes. For example,

¹These issues will also arise for theories that take uncertainty into account and take the sum of the good-making features of outcomes to be *one* of the factors that bears on whether an act is permissible (e.g., theories which appeals to both utility and rights to determine what acts are permissible). That said, to simplify the discussion, I’ll focus on theories which appeal only to utility in this paper.

consider a case where you can take a bet on how many times in a row a fair coin will land tails, and where the payoffs are tailored so as to make the expected utility contributions of these outcomes equal to the members of the alternating harmonic series: $1 - 1/2 + 1/3 - 1/4 + \dots$ ² Since the alternating harmonic series only conditionally converges, this game has no well-defined expected value. Thus expected utility maximization yields no verdict regarding the permissibility of this bet.

Let a *utility contribution* be something which contributes to the utility of an outcome; e.g., the satisfaction of an agent's desire (given decision theory), or a moment of happiness (given hedonic utilitarianism).

3. A third source of problems comes from cases with outcomes containing individual utility contributions of infinite value. For example, assume some form of utilitarianism is true, and consider a case where only three people exist: you, Sad Susie, and Happy Hannah. You're choosing whether to give some chocolate to Sad Susie. If you do, Sad Susie's utility will be 2 instead of 1. Either way, your utility will be 10, and Happy Hannah's utility will be ∞ . Since the expected utility seems infinite no matter what you do, utilitarianism appears to entail that both options are permissible. This seems like the wrong verdict.
4. A fourth source of problems comes from cases with outcomes containing infinitely many utility contributions. For example, assume some form of utilitarianism is true, and consider a case where there are countably infinitely many people with utility 1. You're choosing whether to give one of them chocolate; if you do, they will have a utility of 2 instead of 1. Since the expected utility seems infinite either way, utilitarianism appears to entail that both options are permissible. Again, this seems like the wrong verdict.

This fourth source of problems is the focus of this paper. It's one of the most serious of these four problems, for there are people who are actually in cases of this kind. To see this, note that many contemporary physical theories assign a non-zero probability, per unit of time, to a sentient individual forming in any reasonably sized region of space. Although the probability of such an event is vanishingly small on the time scales we're used to, given enough time, the probability of such an event happening goes to 1. Indeed, given enough time, the probability of such an event happening any number of times goes to 1. Note also that many contemporary cosmological theories suggest that the universe has infinite temporal extension. Together, these theories entail that there's a probability of 1 that there will be infinitely many individuals, and thus infinitely many utility contributions. So agents who have a non-zero credence in these theories will face decision problems containing outcomes with infinitely many utility contributions.³

²This is the Pasadena Game from Nover & Hájek (2004).

³See Bostrom (2011) and Arntzenius (2014) for versions of this point. (Arntzenius (2014) notes that we can partially mitigate these worries by noting that most agents like us will tend to only believe their acts will bear on the utility of subjects in a finite region. But, of course, this won't help in cases like the chocolate case described above, which only requires the agent to believe they're changing the utility of a single subject.) For a more detailed discussion of how these cosmological theories yield these results, see Carroll (forthcoming) and the references therein.

There are also other ways to motivate the importance of this kind of infinity problem. For example, if one

How should we deal with such cases? One thought is to try to assign values to outcomes using some non-standard account of value.⁴ But as others have noted, this approach is unsatisfactory for a number of reasons.⁵ A different approach is to forgo trying to assign a single value to outcomes, and to instead compare the patterns of utility contributions at different outcomes. Both the economics and philosophical literatures have discussed using proposals along these lines to rank outcomes.⁶

Unfortunately, we can't just appeal to these outcome-ranking proposals. For as Bostrom (2011) and Arntzenius (2014) note, proposals for how to rank outcomes aren't enough to determine how one should act. For example, suppose we rank outcomes such that $o_1 > o_2 > o_3$, and we have a choice between (i) bringing about o_2 with certainty, or (ii) a 50/50 chance of bringing about o_1/o_3 . What should we do? As a little thought shows, the ranking of these outcomes doesn't suffice to answer this question. For if o_1 is vastly better than o_2 , and o_3 is only slightly worse than o_2 , then we should take the gamble. Whereas if o_1 is only slightly better than o_2 , and o_3 is vastly worse than o_2 , then we should not.

While these outcome-ranking proposals won't give us what we want, one could try using a similar approach – comparing patterns of utility contributions – to rank acts instead of outcomes. Arntzenius (2014) presents the first detailed proposal along these lines. In many ways Arntzenius's proposal is an attractive one, yielding plausible and discriminating verdicts in a number of scenarios. Nonetheless, I'll argue that it also faces a number

follows preference utilitarians and takes utility contributions to consist of satisfied/unsatisfied desires of a subject at a time, then such problems might arise immediately (e.g., if we individuate desires in a fine-grained way, subjects like us might end having infinitely many desires, and thus yield infinitely many utility contributions; since time is dense, a temporally extended subject might yield infinitely many utility contributions; and so on). Likewise, some economists have suggested that we should model questions regarding our obligations to future generations as infinite decision problems, even if this is unlikely to actually be the case (for examples of this kind of approach, see Chichilnisky (1996) and Chichilnisky (1997)).

⁴For discussions of extensions of standard decision theory along these lines, see for example Hájek (2003a), Bostrom (2011), and Chen & Rubio (forthcoming). Bartha (2007) develops a related but distinct approach, that shares many of the drawbacks of these approaches (e.g., it doesn't tell us how to figure out what one should do (or what preferences one should have) in cases with outcomes with infinitely many utility contributions; see Colyvan & Hájek (2016) for discussion of a related point). Note that most of these authors weren't trying to respond to the fourth kind of infinity problem described above, so it would be unreasonable to fault them for failing to do so.

⁵Here are two of the reasons this approach is unsatisfactory. First, we can need to choose whether to assign coarse-grained or fine-grained values to these outcomes, and both options are problematic. If we assign coarse-grained values to such outcomes (e.g., " ∞ " and " $-\infty$ "), then these assignments yield implausible verdicts (e.g., that we have no reason to increase infinitely many people's utility from 1 to 2, since this won't change the utility of the outcome). Whereas if we assign fine-grained values to outcomes (e.g., hyperreal numbers), then it's not clear it's possible to assign these values to outcomes in a way that avoids arbitrary prescriptions. (See Arntzenius (2014) for a discussion of these difficulties for accounts employing hyperreal values.)

Second, the strategy of expanding the range of values assigned to outcomes won't help in many cases. For in many cases the problem isn't that we lack the values we want to assign to these outcomes, it's that it's unclear what value *should* be assigned. (For example, see the Egalitarian Sweetmeats case discussed in section 3.3.)

⁶The problem is first raised by Ramsey (1928). Serious discussion in the economics literature begins with Koopmans (1960) and vonWeizsacker (1965); for a up-to-date survey of this literature, see Lauwers (2014). Serious discussion in the philosophical literature begins with Segerberg (1976); later discussions include Nelson (1991), Vallentyne (1993), Garcia & Nelson (1994), Liedekerke (1995), Cain (1995), Vallentyne & Kagan (1997), Hamkins & Montero (2000), Fishkind et al. (2002), and Lauwers & Vallentyne (2004).

of serious problems.

In this paper I defend a different proposal – Direct Difference Taking – for how to make decisions in cases with outcomes with infinitely many utility contributions. Like Arntzenius’s proposal, it follows the outcome-ranking literature in adopting a comparative approach to render verdicts. But Direct Difference Taking avoids the problems facing Arntzenius’s view.

The rest of this paper will proceed as follows. In section 2, I’ll introduce some background and lay out some assumptions. In section 3, I’ll motivate my proposal by considering a series of cases involving infinitely many utility contributions. Each case poses a different challenge for expected utility based views, and in each case I’ll endorse a way of revising decision theory to avoid these problems. Together, these revisions yield the proposal I’ll advocate, Direct Difference Taking. In section 4, I’ll present a concise formulation of Direct Difference Taking. In section 5, I’ll present Arntzenius’s competing proposal. In section 6, I’ll present objections to Direct Difference Taking that Arntzenius’s proposal avoids. In section 7, I’ll discuss objections to Arntzenius’s proposal that Direct Difference Taking avoids. In section 8, I briefly summarize my results.

2 Background Assumptions

For the purposes of this discussion, I’ll take a *decision problem* to be an ordered 5-tuple (A, O, cr, C, u) where:

- A is a set of available acts,
- O is the set of outcomes those acts might bring about,
- $cr : \mathcal{P} \rightarrow [0, 1]$ is a probabilistic function from the minimal Boolean algebra \mathcal{P} that contains the elements of A and O as members, to a number in the real interval $[0, 1]$ representing the subject’s degree of confidence in that proposition,
- C is the set of sets C_o for all outcomes $o \in O$, where C_o is the set containing all of the utility contributions at outcome o ,
- $u : O \times C_o \rightarrow \mathbb{V}$ is a function that takes an outcome and a utility contribution at that outcome, and yields the value of that utility contribution.⁷

This characterization of decision problems is relatively standard except for the characterization of the utility function and the inclusion of C . Utility functions are usually taken to just be functions from outcomes to values that correspond to the utility of that outcome. But in the cases we’re concerned with – e.g., cases where we want to distinguish between outcomes in which infinitely many people each have a utility of 1 instead of 2 – it’s hard to know how to assign utilities to outcomes in a satisfying way. So I’ll employ a more fine-grained utility function that assigns values to individual utility contributions at outcomes, instead of to outcomes as a whole.

I’ll take the task here to be that of figuring out what acts are permissible in decision problems like this.

⁷I use the generic name “ \mathbb{V} ” here in order to keep this characterization of a decision problem neutral with respect to what kinds of values a utility contribution can have.

In what follows, I'll make the following further assumptions. First, I'll assume that all the cases we're concerned with have outcomes with either finitely many or countably infinitely many utility contributions. That is, I'll assume that:

Assumption 1. There are no outcomes with uncountably many utility contributions.

This ensures that we can aggregate utility contributions using summation.

Next, in order to put aside the other three kinds of infinity problems discussed in section 1, I'll make the following assumptions:

Assumption 2. There are finitely many acts.

Assumption 3. There are finitely many states.⁸

Assumption 4. Individual utility contributions take real-numbered (and thus finite) values.

These assumptions allow us to focus our attention on the fourth kind of infinity problem: cases with infinitely many utility contributions.

In order to remain neutral between evidential and causal decision theory, it will be convenient to introduce the term " $cr(o : a)$ ", where (i) if we're evidential decision theorists, this stands for the credence of o conditional on a , and (ii) if we're causal decision theorists, this stands for the credence of o imaged on a .⁹ Using this terminology, we can provide a theory-neutral characterization of expected utility. Given a decision problem (A, O, cr, C, u) , let the *expected utility* of an act $a \in A$, $EU(a)$, be:

$$EU(a) = \sum_{(i|o_i \in O)} cr(o_i : a) \cdot \sum_{(j|c_j \in C_{o_i})} u(o_i, c_j).$$

This is the standard characterization of expected utility, except that it replaces the utility assigned to an outcome with a sum over the individual utility contributions at an outcome.

Finally, I'll assume that in finite cases we want our account to yield the same prescriptions as expected utility maximization. That is:

Assumption 5. In finite cases (cases in which there are only finitely many utility contributions at each outcome), an act is permissible *iff* it maximizes expected utility.

This assumption does two things. First, it focuses our attention on theories that (in finite cases) take uncertainty into account in a risk-neutral way, by assessing acts via expected utilities (in contrast to risk-averse or risk-seeking theories).¹⁰ Second, combined with the

⁸Given this way of characterizing decision problems given above, one can think of states as functions from acts to outcomes, and the assumption that there are finitely many states as the assumption that there are only finitely many distinct functions of this kind.

⁹The probability of o conditional on a is usually defined as $\frac{cr(o \wedge a)}{cr(a)}$ (though see Hájek (2003b) for a case for taking conditional probabilities to be primitive). Letting w^a be the world most similar to w where a is true, the probability of o imaged on a is usually defined as $\sum_w cr(w) \cdot \frac{cr(o \wedge w^a)}{cr(w^a)}$. (This assumes that there is a unique world that is most similar to w ; for formulations of imaging that don't assume this, see Joyce (1999), chapter 6.)

¹⁰For a discussion of these alternatives, see Buchak (2013).

definition of expected utility given above, it entails that we're working with what Lauwers & Vallentyne (2004) call "finitely additive" theories of value – theories which, for outcomes with only finitely many utility contributions, assign utilities to those outcomes by simply summing over the utility contributions.¹¹ The definition of expected utility entails that we sum over utility contributions to determine the effective utility of an outcome when making expected utility calculations, and assumption 5 requires our prescriptions to go via these expected utilities in finite cases.

3 Motivating Direct Difference Taking

In this section I'll go through a series of cases with outcomes with infinitely many utility contributions, each of which raises a different problem. After presenting each case, I'll then present what I take to be the most plausible way of revising decision theory in order to handle the problem.

Going through these cases serves two purposes. First, my positive proposal, Direct Difference Taking, is simply the combination of these revisions. So by using this sequence of cases to motivate these revisions, I thereby motivate Direct Difference Taking. Second, it can be hard to make out how Direct Difference Taking works, and what kinds of verdicts it yields, from the concise formulation of Direct Difference Taking provided in section 4. By going through these cases and building up Direct Difference Taking step-by-step, I hope to provide a more lucid view of how Direct Difference Taking functions, and the choice points that lead to it having the consequences it does.

In order to make it easier to provide concrete examples, I'll make some assumptions about what utility contributions are in this section. For the purposes of this section, I'll assume that utility contributions are values representing the overall happiness of an individual's life.

3.1 Finite Improvements

Stealing Candy. There are countably infinitely many babies, each of whom has some candy. You have two options. The first (Do Nothing) is to leave the babies as they are, in which case each will have a utility of 2. The second (Steal Candy) is to steal candy from the first baby, decreasing their utility from 2 to 1.

	Baby 1	Baby 2	...	Sum
Do Nothing	2	2	...	undefined
Steal Candy	1	2	...	undefined

The standard way of handling these cases is to find out what the sum of the individual utility contributions at an outcome converges to, take that to be the utility of the outcome, and then use those values to figure out which acts maximize expected utility. But in

¹¹Note that this includes theories which take there to be holistic utility contributions, such as various forms of egalitarianism. As long as such theories assign utilities to outcomes in finite cases by summing over these features (e.g., by adding up the utility stemming from individual happiness and the utility stemming from the value of the outcome's (in)egalitarian features), they'll fall within the scope of our discussion.

this case, the sums of utility contributions don't converge to any real number, and so the utility of these outcomes is undefined. So the standard approach falls silent. But the correct prescription in this case seems clear: dominance considerations indicate that Steal Candy is impermissible.

3.1.1 Suggested Change: Difference-Taking

One natural way to deal with this problem is to evaluate a pair of acts A and B by taking the differences between their utility contributions and summing over these differences.¹² We can then conclude that A is impermissible if the difference between A and B is negative; likewise, we can conclude that B is impermissible if this difference is positive. More generally, in cases in which there are more than two acts, we can assess each pair of acts in this manner to determine which acts are impermissible, and then let any acts not deemed impermissible be permissible.

In Stealing Candy, the sum of the differences between utility contributions at the Do Nothing and Steal Candy outcomes will be 1:

	Baby 1	Baby 2	...	Sum
Do Nothing	2	2	...	undefined
Steal Candy	1	2	...	undefined
Difference	1	0	...	1

Thus we get the desired verdict that Steal Candy is impermissible.¹³

3.2 Infinite Improvements

In the case above, the sum over differences converges to a finite value. Now let's consider cases where the sum diverges by growing infinitely high or low.

Stealing All The Candy. There are countably infinitely many babies, each of whom has some candy. You have two options. The first (Do Nothing) is to leave the babies as they are, in which case each will have a utility of 2. The second (Steal All Candy) is to steal candy from all of them, decreasing the utility of each baby from 2 to 1.

¹²Appealing to something like difference-taking is common in the literature regarding how to compare outcomes (instead of acts), though it's usually implemented after some kind of preliminary aggregation of subsets of utility contributions; for example, see vonWeizsacker (1965), Vallentyne (1993), Vallentyne & Kagan (1997), and Lauwers & Vallentyne (2004). Arntzenius (2014) likewise appeals to something like difference-taking when comparing acts, applied after preliminary aggregations of utility contributions (see section 5). Interestingly, some have also suggested employing difference-taking to address some of the problems that arise in infinite state cases. Colyvan (2008) has suggested that we apply difference-taking to states (instead of utility contributions) in order to derive the desired verdicts in cases where one act dominates another, and yet neither act has a well-defined finite expected value, and both Lauwers & Vallentyne (forthcoming) and Meacham (forthcoming) have defended extensions of this approach.

¹³I'm assuming here that there is a natural way of pairing utility contributions in different outcomes that satisfies certain desirable properties: symmetry, transitivity, one-to-one, etc. We'll look at how to proceed if we reject this assumption in section 3.5.

	Baby 1	Baby 2	...	Sum
Do Nothing	2	2	...	undefined
Steal All Candy	1	1	...	undefined
Difference	1	1	...	undefined

In this case, mere difference-taking won't help. This is because the infinite sum over the differences between Do Nothing and Steal All Candy won't converge to any real number. Thus, given the standard way of evaluating such sums, the sum of these differences is undefined. But the correct prescription in this case seems clear: dominance considerations indicate that Steal All Candy is impermissible.

3.2.1 Suggested Change: Infinite Convergence

One natural way to deal with this problem is to extend the values we allow infinite sums to take.

Here's one way to do this. Let the *extended real numbers* $\overline{\mathbb{R}}$ be the union of the real numbers \mathbb{R} and two additional elements, ∞ and $-\infty$, where the ordering and arithmetic operations over the reals are extended over these new elements in more or less the way one would expect.¹⁴ We can extend the notion of convergence to extended real numbers as follows:

Convergence to $\overline{\mathbb{R}}$: Say an infinite series S converges to $r \in \overline{\mathbb{R}}$ iff

1. $r \in \mathbb{R}$, and for every $\epsilon \in \mathbb{R}$, there exists a $n \in \mathbb{N}$ such that every partial sum of n or more terms is within $\pm\epsilon$ of r ,
2. $r = \infty$, and for any $\alpha \in \mathbb{R}$, there exists a $n \in \mathbb{N}$ such that every partial sum of n or more terms is larger than α ,
3. $r = -\infty$, and for any $\alpha \in \mathbb{R}$, there exists a $n \in \mathbb{N}$ such that every partial sum of n or more terms is smaller than α .

Then instead of taking the values of infinite sums to be the real numbers they converge to, we can take the values of infinite sums to be the extended real numbers they converge to.¹⁵

In Stealing All The Candy, the sum of the differences between Do Nothing and Steal All Candy will now be ∞ :

	Baby 1	Baby 2	...	Sum
Do Nothing	2	2	...	∞
Steal All Candy	1	1	...	∞
Difference	1	1	...	∞

¹⁴More precisely, the ordering and arithmetic operations are extended over these new elements as follows: *Order:* $\forall r \in \mathbb{R}, -\infty < r < \infty$. *Addition and Subtraction:* 1. $\forall r \in \mathbb{R}, r \pm \infty = \pm\infty$. 2. $\infty + \infty = \infty$. 3. $-\infty + -\infty = -\infty$. 4. $\infty + -\infty = \text{undefined}$. *Multiplication:* 1. $\forall r \in \mathbb{R}^+, r \cdot \pm\infty = \pm\infty$. 2. $\forall r \in \mathbb{R}^-, r \cdot \pm\infty = \mp\infty$. 3. $\pm\infty \cdot \pm\infty = \infty$. 4. $\pm\infty \cdot \mp\infty = -\infty$. 5. $0 \cdot \pm\infty = 0$. *Division:* 1. $\forall r \in \mathbb{R}, \frac{r}{\pm\infty} = 0$. 2. $\forall r \in \mathbb{R}^+, \frac{\pm\infty}{r} = \pm\infty$. 3. $\forall r \in \mathbb{R}^-, \frac{\pm\infty}{r} = \mp\infty$. 4. $\frac{\pm\infty}{\pm\infty} = \frac{\pm\infty}{\mp\infty} = \text{undefined}$. 5. $\frac{\pm\infty}{0} = \text{undefined}$.

¹⁵Arntzenius (2014) likewise assumes an extended notion of convergence to make sense of infinite-valued sums (see Arntzenius (2014), p54), though he does not commit himself to all of the details given here. In a similar vein, much of the literature on how to compare outcomes (instead of acts) assumes some way of making sense of infinite-valued sums (e.g., see Lauwers & Vallentyne (2004), section 5).

Since ∞ is positive, we get the desired verdict that Steal All Candy is impermissible.

3.3 Order-Dependence

In the above cases, the order in which we sum over the differences doesn't matter. Now let's consider cases where the order does matter.

Egalitarian Sweetmeats. There are countably infinitely many babies, each named after one of the natural numbers, who have varying quantities of sweetmeats. You have two options. The first (Do Nothing) is to leave the babies as they are, in which case the utilities of these babies will correspond to the alternating harmonic series: $1, -1/2, 1/3, \dots$, and so on. The second (Egalitarianism) is to ensure every baby has the same amount of sweetmeats, resulting in each having a utility of 0.

	Baby 1	Baby 2	Baby 3	...	Baby n	...	Sum
Do Nothing	1	$-\frac{1}{2}$	$\frac{1}{3}$...	$\frac{(-1)^{n+1}}{n}$...	$\ln 2?$
Egalitarianism	0	0	0	...	0	...	0
Difference	1	$-\frac{1}{2}$	$\frac{1}{3}$...	$\frac{(-1)^{n+1}}{n}$...	$\ln 2?$

Given the proposal sketched in sections 3.1 and 3.2, our evaluation of this case will depend on the way in which we order these terms. If the sum of differences in utility contributions is performed in the order indicated in the diagram above – i.e., in order of the natural numbers that name the babies – then the sum will converge to $\ln 2$, or ≈ 0.69 . This is positive, so we'll get the verdict that Egalitarianism is impermissible. By contrast, if we order these terms so that we sum over four even numbered babies for every odd numbered one – i.e., sum over the differences between babies 1, 2, 4, 6, 8, 3, 10, 12, 14, 16, 5, ..., then the sum will converge to ≈ -0.11 . This value is negative, so we'll get the verdict that Do Nothing is impermissible.

So our verdict in this case will depend on how we order these terms. This seems problematic, because there doesn't seem to be a morally or prudentially privileged way to order these babies. There may well be features of the case that we could use to provide an ordering – the order of the numbers we've used to name them, the order in which they were born, and so on – but none of these facts seem morally or prudentially relevant.

3.3.1 Suggested Change: Permutation Invariance

I hold that there is no morally or prudentially privileged way of ordering utility contributions. As a result, to obtain the verdict that an act A is impermissible, we should require the sum over differences in utility contributions between A and another act to be negative regardless of the way in which we order these terms. That is, we should require these sums to be permutation invariant.¹⁶

¹⁶It's worth highlighting that this permutation invariance requirement is distinct from the (much stronger) Isomorphism requirement that if there's a value-preserving bijection between the utility contributions in a pair of outcomes, then those outcomes should be interchangeable. I discuss Isomorphism in more detail in appendix

In Egalitarian Sweetmeats, the valence of the sum of differences is not permutation invariant. So our comparison of the two acts doesn't render either impermissible. Since these are the only two options, it follows that Do Nothing and Egalitarianism are both permissible. This seems like a reasonable verdict.¹⁷

3.4 Unpaired Utility Contributions

In the cases above, all of the utility contributions in each outcome are paired with utility contributions in every other outcome. Now let's consider cases where there are unpaired utility contributions.

Siblings. You have been given the task of deciding what kind of sibling to create for a lonely child. You have two options. The first (Happy Sister) is to create a happy sister for the child, in which case both siblings will have a utility of 2. The second (Unhappy Brother) is to create an unhappy brother for the child, in which case the child will have a utility of 2 and the brother a utility of -2.

A, and consider a case in which Isomorphism and permutation invariance come apart (a case where DDT violates Isomorphism even though it satisfies permutation invariance).

That said, it's worth noting that even the relatively weak permutation invariance requirement I defend in this section is controversial. As we'll see in section 6, Arntzenius's (2014) proposal violates this requirement. And in the literature on how to compare outcomes (instead of acts), many of the economists and philosophers who have written on this topic reject the analog of permutation invariance. For example, vonWeizsacker (1965) and Vallentyne (1993) effectively suggest that we should order terms temporally when making our assessments, contra permutation invariance. In a similar vein, Vallentyne & Kagan (1997) and Lauwers & Vallentyne (2004) suggest that if there's a "natural ordering", then we should use that to order terms when making our assessments, even if this violates permutation invariance. That said, there have also been several defenders of something like permutation invariance, such as Garcia & Nelson (1994), Liedekerke (1995), Hamkins & Montero (2000) and Fishkind et al. (2002), though these authors generally argue for much stronger principles than the one advocated here (principles like the Isomorphism requirement described above).

¹⁷Another natural thought regarding how to address this problem is to try to take one of the order-independent ways of evaluating infinite sums suggested in the infinite states literature (the second source of infinity problems mentioned in section 1), and adapt it to apply to cases like Egalitarian Sweetmeats. For example, one might try to adapt something like Gwiazda's (2014) approach, by insisting on ordering utility contributions by the magnitude of their utilities. Or one might try to adapt something like the "principal value" approach described by Easwaran (2014) to handle these cases, by considering truncated sums which omit utility contributions of magnitude greater than n , and then seeing what values these truncated sums take in the limit as n goes to infinity. Unfortunately, none of these adapted proposals help with the Egalitarian Sweetmeats case. The adaptation of Gwiazda's proposal won't help because there is no smallest magnitude utility contribution to start with. And the adaptation of Easwaran's proposal won't help because the values of these truncated sums will themselves be order-dependent. (With Easwaran's original proposal and infinite states cases, the truncated sums we consider are over expected utility contributions of the relevant states, and these truncated sums will always yield order-independent results because the sum of the magnitudes of these terms will always be finite. With this adaptation of Easwaran's proposal to infinitely many contribution cases, the truncated sums we consider are over the relevant utility contributions, and the sum of the magnitudes of *these* terms can be infinite and order-dependent (as in the Egalitarian Sweetmeats case).)

	Child	Sister	Brother	Sum
Happy Sister	2	2	—	4
Unhappy Brother	2	—	-2	0
Difference	0	—	—	0

In this case, the sum of the differences in utility contributions between the two outcomes will only yield a single term: 0. Thus we'll get the verdict that both options are permissible. But this seems like the wrong verdict: Unhappy Brother should be impermissible.

3.4.1 Suggested Change: Zero Point Pairing

I propose that we employ a utility scale with a privileged zero point, corresponding to the point at which whether or not we include that contribution makes no difference to our verdict. Doing so provides us with a natural way to deal with this problem: we can treat utility contributions that aren't paired with anything as if they were paired with a contribution with 0 utility. This will allow us to take all of the utility contributions at an outcome into account, even if they aren't paired with anything.¹⁸

In Siblings, the sum of differences between Happy Sister and Unhappy Brother will now be 4:

	Child	Sister	Brother	Sum
Happy Sister	2	2	'0'	4
Unhappy Brother	2	'0'	-2	0
Difference	0	2	2	4

Since 4 is positive, we get the verdict that Unhappy Brother is impermissible, as desired.

¹⁸In the context of ranking outcomes (instead of acts), Vallentyne & Kagan (1997) briefly mention the possibility of adding "empty virtual locations" in order to compare outcomes in cases where one outcome has a proper subset of the utility contributions of another outcome (see Vallentyne & Kagan (1997), footnote 18). This is similar to zero point pairing, though it won't apply to cases like Siblings, in which neither outcome has utility contributions that are a proper subset of the other's. While discussing an alternative to his proposal (the Weak Person Condition discussed in appendix B), Arntzenius (2014) also briefly mentions the possibility of using something like zero point pairing with respect to temporally extended individuals.

It's worth noting that zero point pairing is controversial. For example, in the literature on ranking outcomes, Lauwers & Vallentyne (2004) join a number of economists in suggesting that our ranking shouldn't depend on where we put the zero point of the scale we use to measure utility. Thus, for example, if we measure utility using a scale that assigns every contribution a value that's 10 units higher, our ranking of outcomes shouldn't change. The analog of zero point pairing in this context (i.e., zero point pairing applied to outcomes instead of acts) will violate this requirement. For example, if we consider an outcome with one person with 3 utility, and another outcome where both that person and another person have 1 utility, then zero point pairing will suggest that the first outcome is better than the second. But if we consider an outcome with one person with 13 utility, and another outcome where both that person and another person have 11 utility, then zero point pairing will suggest that the second outcome is better than the first.

3.5 Pairing-Dependence

So far, I've been assuming that our method of pairing utility contributions in different outcomes satisfies certain desirable properties: symmetry, transitivity, and being one-to-one. But there are cases where the standard ways of pairing utility contributions in different outcomes won't satisfy these properties. And these cases raise difficulties for the difference-taking procedure described above. For example:

Fusion. You have been given the task of deciding whether to perform fusion on a pair of children. You have two options. The first (Do Nothing) is to leave the two children as they are, in which case the first will have a utility of 1, and the other a utility of 3. The second (Fusion) is to fuse the two children into one child who will have a utility of 2.

	Child 1	Child 2	Sum
Do Nothing	1	3	4
Fusion	2		2
Difference	?		?

In this case, one natural stance is to pair Children 1 and 2 in Do Nothing with their fusion in Fusion. However, this pairing isn't one-to-one. As a result, it's unclear how to assess the sum of the differences in utility contributions between Do Nothing and Fusion. Should one take the sum of differences to consist of two terms, each comparing one of the children in Do Nothing to the fused child in Fusion? Or take the sum of differences to consist of one term, using the average of the utilities of Children 1 and 2 to make the comparison? Or do something else entirely? Similar worries arise for cases involving fission instead fusion.

There are two standard ways of pairing individuals in different outcomes. One is to appeal to a primitive *transworld identity* relation. The other is to appeal to similarity-based and context-dependent *counterpart* relations.¹⁹ These one-to-one violation problems arise regardless of whether one uses transworld identity facts or counterpart theory to pair subjects in different outcomes. But further potential problems for difference-taking arise if we employ counterpart theory.

One worry is that counterpart relations are generally taken to be context-dependent. Indeed, it's generally assumed that there are infinitely many different counterpart relations that one could employ, depending on the conversational context. This is problematic if we want an account in which what's permissible doesn't vary from context to context. But there's a natural response to this worry available: we can follow Arntzenius (2014) and take the counterpart relation we'd use in ordinary contexts to be the one we should use when evaluating which acts are permissible.

A second worry is that counterpart relations, unlike identity relations, needn't be symmetric or transitive. And this can make trouble for the kind of difference-taking described above. For example, consider the following case:

¹⁹For discussions of these approaches, see Divers (2002), Melia (2003), and the references therein.

Rich and Poor. There are countably infinitely many babies, each named after one of the natural numbers, and each of which has some candy. The odd numbered babies are rich, the even numbered babies are poor. You have two options. The first (Do Nothing) is to leave the babies as they are, in which case each odd-numbered baby will have a utility of 2, and every even-numbered baby will have a utility of -1. The second (Steal All Candy) is to steal candy from all of them, resulting in odd-numbered babies with a utility of 1 and even-numbered babies with a utility of -2.

Suppose our counterpart relation pairs subjects with the same name in the two outcomes:

	Baby 1	Baby 2	Baby 3	Baby 4	...	Sum
Do Nothing	2 ↑ ↓	-1 ↑ ↓	2 ↑ ↓	-1 ↑ ↓	...	undefined
Steal All Candy	1 ↑ ↓	-2 ↑ ↓	1 ↑ ↓	-2 ↑ ↓	...	undefined
Difference	1	1	1	1	...	∞

Then the sum of differences in utility contributions will be ∞ , and we'll get the verdict that Steal All Candy is impermissible.

Suppose instead our counterpart relation pairs subjects in a cross-hatched fashion, as follows:

	Baby 1	Baby 2	Baby 3	Baby 4	...	Sum
Do Nothing	2 ↙ ↘	-1 ↙ ↘	2 ↙ ↘	-1 ↙ ↘	...	undefined
Steal All Candy	1 ↙ ↘	-2 ↙ ↘	1 ↙ ↘	-2 ↙ ↘	...	undefined
Difference	4	-2	4	-2	...	undefined

Then the sum of differences in utility contributions will be undefined, and neither act will be judged impermissible. Thus we'll get the verdict that both options are permissible.

Finally, suppose that our counterpart relation is not symmetric – that babies in the Do Nothing outcome have as counterparts the babies in the Steal All Candy outcome with the same name, while the babies in the Steal All Candy outcome have as counterparts the cross-hatched babies in the Do Nothing outcome:

	Baby 1	Baby 2	Baby 3	Baby 4	...	Sum
Do Nothing	2 ↙ ↘	-1 ↙ ↘	2 ↙ ↘	-1 ↙ ↘	...	undefined
Steal All Candy	1 ↙ ↘	-2 ↙ ↘	1 ↙ ↘	-2 ↙ ↘	...	undefined
Difference (DN-SAC)	1	1	1	1	...	∞
Difference (SAC-DN)	2	-4	2	-4	...	undefined

With this non-symmetric counterpart relation we'll get inconsistent verdicts. If we start with the Do Nothing outcome, and take the difference with respect to the Steal All

Candy outcome, we'll get a sum that converges to ∞ , and the verdict that Steal All Candy is impermissible. If we start with the Steal All Candy outcome, and take the difference with respect to the Do Nothing outcome, we'll get a sum that's undefined. And since these are the only two acts, the procedure described in section 3.1.1 will entail that Steal All Candy is permissible.

So non-symmetric counterpart relations can make trouble for the difference-taking account. In a similar fashion, counterpart relations that aren't transitive can make trouble for the difference-taking account.²⁰

3.5.1 Suggested Change: Acceptable Pairings

One natural way to deal with this problem is to impose certain restrictions on the ways in which we can pair individuals in different outcomes in this context. In particular, while we'll want our way of pairing contributions to closely resemble the transworld identity facts (if we're transworld identity theorists) or the ordinary counterpart relations Arntzenius (2014) appeals to (if we're counterpart theorists), we'll also want our way of pairing contributions to satisfy certain constraints that make them well-behaved, like symmetry, transitivity, and being one-to-one.

Acceptable Pairings: Let an *acceptable pairing*, relative to a decision problem, be a way of pairing utility contributions in different outcomes such that:

1. With respect to the restricted domain consisting only of the utility contributions in the outcomes of this decision problem,²¹ this method of pairing contributions is symmetric, transitive and one-to-one.²²

²⁰Arntzenius (2014) assumes some form of counterpart theory when formulating his proposal, but does not address the issues that arise from the ways in which counterpart relations can diverge from identity relations, or issues that arise from cases of fission and fusion (I discuss this worry for Arntzenius's proposal in section 7.1). In the context of ranking outcomes (instead of acts), Vallentyne & Kagan (1997) also discuss ways of accommodating counterpart theory. But like Arntzenius, they do not address the issues that arise from the ways in which counterpart relations diverge from identity relations (they assume the counterpart relation is one-to-one), or issues that arise from cases of fission or fusion.

²¹This highly restricted domain minimizes the amount by which an acceptable pairing will deviate from the pairings suggested by the transworld identity facts or the ordinary counterpart relation. First, this requirement only holds with respect to whatever one takes to be providing utility contributions (e.g., mental states at a time). One may continue to pair everything else in the standard way, even though such pairings may fail to be symmetric, transitive, or one-to-one. Second, this requirement only holds with respect to the utility contributions that appear in the outcomes of the relevant decision problem. So most utility contributions will still be paired in the standard way, even though, again, such pairings may fail to be symmetric, transitive, or one-to-one.

²²We can spell out these conditions more precisely as follows. (I'll focus here on how to spell out these conditions for counterpart theorists; it should be clear how to spell out these conditions for transworld identity theorists.) Let O be the outcomes of this decision problem, and let D be the restricted domain of utility contributions of outcomes in O . Following Lewis (1968), let $I(x, y)$ be the relation that holds *iff* x is in possible world y , and let $C(x, y)$ be the relation that holds *iff* x is a counterpart of y . For all $o, o^* \in O$, and all $w, x, y, z \in D$, we require the counterpart relation to be such that:

- (a) $C(w, x) \rightarrow C(x, w)$ (symmetry)
- (b) $C(w, x) \wedge C(x, y) \rightarrow C(w, y)$ (transitivity)
- (c) $I(w, o) \wedge I(y, o^*) \wedge o \neq o^* \wedge C(w, y) \rightarrow \neg \exists x \neq w (I(x, o) \wedge C(x, y)) \wedge \neg \exists z \neq y (I(z, o^*) \wedge C(w, z))$ (one-to-

2. Of the ways of pairing utility contributions that satisfy 1, this method of pairing best fits the transworld identity facts (if we're transworld identity theorists) or best fits the ordinary counterpart relation (if we're counterpart theorists).

I propose that we require our difference-taking procedure to pair contributions using acceptable pairings.

(What about cases in which there are ties with respect to 2 – cases where there are multiple pairings that satisfy 1 and which fit the transworld identity facts or ordinary counterpart relation facts equally well? Then we require all of the acceptable pairings to agree in order to derive an impermissible verdict. That is, we conclude that *A* is impermissible only if the sum of the difference between *A* and *B*'s utility contributions, according to every acceptable pairing, is negative.)

Return to the Fusion case. In this case, the one-to-one clause requires us to pair the child in the Fusion outcome with only one of the children in the Do Nothing outcome. Since there don't seem to be any reasons to favor one pairing over the other, both pairing Child 1 in Do Nothing with the child in Fusion, and pairing Child 2 in Do Nothing with the child in Fusion, are acceptable pairings. If we adopt the first acceptable pairing, the sum of differences between Do Nothing and Fusion will be:

	Child 1	Child 2	Sum
Do Nothing	1 ↓	3 ↓	4
Fusion	2	'0'	2
Difference	-1	3	2

If we adopt the second acceptable pairing, the sum of differences between Do Nothing and Fusion will be:

	Child 1	Child 2	Sum
Do Nothing	1 ↑	3 ↑	4
Fusion	'0'	2	2
Difference	1	1	2

Either way, the difference is positive, so Fusion is impermissible.

Now let's turn to the Rich and Poor case. These kinds of symmetry violations won't arise for transworld identity theorists (since transworld identity is symmetric), so we can focus on counterpart theorists. In the Rich and Poor case, the description of the case naturally suggests pairing babies with the same name; this would seem to be the ordinary counterpart relation. Since this ordinary counterpart relation is symmetric when restricted to the contributions in the outcomes of this decision problem, and best fits the

one)

ordinary counterpart relations (since it is the ordinary counterpart relation), it will be the (unique) acceptable pairing. As we've seen, this way of pairing subjects yields the verdict that Steal All Candy is impermissible. Thus we get the desired result that Steal All Candy is impermissible.

3.6 Uncertainty

In the cases we've looked at so far, there's no uncertainty about what outcome an act will bring about. Now let's consider cases in which agents are unsure about what outcome their options will bring about.

Candy Gamble. There are countably infinitely many babies, each of which has some candy. You have two options. The first option (Do Nothing) is to leave them as they are, resulting in an outcome in which each has 2 utility. The second option (Gamble) is to gamble with all of their candy; if you do so, you have a credence of 1/3 that you'll win and get more candy for each baby (in which case each baby will have a utility of 3), and a credence of 2/3 that you'll lose, and each baby will be left with no candy (in which case each baby will have a utility of 1).

	Credence	Baby 1	Baby 2	...	Sum
Do Nothing	1	2	2	...	∞
Gamble	1/3 (win)	3	3	...	∞
	2/3 (lose)	1	1	...	∞

In this case, where the Gamble act might yield either of two different outcomes, it's not clear how to apply the difference-taking procedure described above.

3.6.1 Suggested Change: Expected Outcomes

A natural way to deal with this problem is to apply the difference-taking procedure to *expected* utility contributions instead of utility contributions. More precisely, let an *expected utility contribution* (given a certain act) be the weighted average of the values of that utility contribution at the outcomes the act might bring about (where the weights are determined by the agent's credence that the act will bring that outcome about). Then we can assess pairs of acts by summing over the differences between their expected utility contributions.²³

In Candy Gamble, the sum of the differences between expected utility contributions will be ∞ :

	Credence	Baby 1	Baby 2	...	Sum
Do Nothing	1	2	2	...	∞
Gamble	1/3 (win)	3	3	...	∞
	2/3 (lose)	1	1	...	∞
Do Nothing (Expected)	—	2	2	...	∞
Gamble (Expected)	—	5/3	5/3	...	∞
Difference	—	1/3	1/3	...	∞

²³This approach to handling uncertainty follows Arntzenius (2014) (see section 5).

Since ∞ is positive, we get the desired verdict that Gamble is impermissible.

4 Direct Difference Taking

I'll now present my positive proposal for how to handle cases with outcomes with infinitely many utility contributions, Direct Difference Taking. Direct Difference Taking is just the amalgamation of the changes described in section 3. With the definitions of *convergence to extended reals* (section 3.2.1) and *acceptable pairings* (section 3.5.1) in hand, we can present it in a more concise fashion as follows.

First, we introduce the notion of an acceptable numbering:

Acceptable Numbering: Let an *acceptable numbering* N , relative to a decision problem (A, O, cr, C, u) , be a function $N : O \times \mathbb{N} \rightarrow C_o$ that assigns a natural number x to every utility contribution in every outcome of the decision problem such that, for some acceptable pairing:

1. Contributions that are paired with one another get assigned the same number.
2. Contributions which aren't paired with one another are assigned different numbers.

An acceptable numbering is a way of assigning labels (natural numbers) to utility contributions such that utility contributions in different outcomes get assigned the same number *iff* they're paired with each other.

Second, we introduce the function $u_N : O \times \mathbb{N} \rightarrow \mathbb{R}$, where

$$\begin{aligned} u_N(o, n) &= u(o, N(o, n)) \text{ (if a contribution at } o \text{ is assigned } n \in \mathbb{N} \text{ by } N) \\ &= 0 \quad \text{(otherwise)} \end{aligned}$$

So u_N is a function (relativized to an acceptable numbering N) that takes an outcome and a number, and yields the utility of the contribution assigned that number at that outcome by N (or yields 0, if nothing is assigned that number at that outcome by N).

We can then characterize Direct Difference Taking as follows:

Direct Difference Taking (DDT): For a given decision problem (A, O, cr, C, u) , an act $a \in A$ is impermissible *iff* there exists some act $b \in A$ such that, for every acceptable numbering N , the sum:

$$\sum_{n \in \mathbb{N}} \left(\sum_{(i|o_i \in O)} cr(o_i : a) \cdot u_N(o_i, n) - \sum_{(j|o_j \in O)} cr(o_j : b) \cdot u_N(o_j, n) \right)$$

converges to a negative extended real number.

4.1 Utility Contributions

We've now seen the form of DDT. But in order to get concrete verdicts, we also need an account of what utility contributions are. Interesting questions arise here, regarding both what kind of thing contributes utility – e.g., happiness, satisfied desires, art, health,

knowledge, biodiversity, etc. – and what the instantiators of utility contributions are – e.g., temporal parts of sentient subjects, temporally extended sentient subjects, objects, entire worlds, etc. And different answers to these questions will yield different verdicts. This is obvious in the case of what kind of thing contributes utility – taking happiness to be the only thing that contributes utility will lead to different prescriptions than taking artwork to be the only thing that contributes utility. But it’s also true with respect to what the instantiators of utility contributions are.

Consider again the Egalitarian Sweetmeats case from section 3.3. Our initial assessment of this case assumed that the utility contributions depended on the lifetime happiness of a subject, and it resulted in the verdict that both Do Nothing and Egalitarianism were permissible options. But suppose instead that we took the utility contributions to be determined by the lifetime happiness of a pair of subjects, in such a way as to pair the first and second babies, the third and fourth babies, and so on. And suppose that the utility contributions of a pair of subjects is equal to sum of the contributions that would be assigned to each child individually on an account which assigns contributions to individual subjects. Then DDT will assess this case as follows:

	Baby 1+2	Baby 3+4	Baby 5+6	...	Sum
Do Nothing	$1 - \frac{1}{2} = \frac{1}{2}$	$\frac{1}{3} - \frac{1}{4} = \frac{1}{12}$	$\frac{1}{5} - \frac{1}{6} = \frac{1}{30}$...	$\ln 2$
Egalitarianism	0	0	0	...	0
Difference	$\frac{1}{2}$	$\frac{1}{12}$	$\frac{1}{30}$...	$\ln 2$

And in this case, the sum of these terms *is* permutation invariant, so we’ll get the verdict that Do Nothing is obligatory instead of merely permissible.

So what should we take the utility contributions to be? This is an important question, but it’s not a question for DDT to answer. DDT is a proposal for how to take theories like standard decision theory or hedonic utilitarianism and extend them so that they can yield plausible verdicts in cases with infinitely many utility contributions. And what the utility contributions are in such cases is determined by the theory DDT is trying to extend. E.g., if we’re extending hedonic utilitarianism of the kind defended by Feldman (2004), the utility contributions will be something like moments of pleasure. If we’re extending a version of decision theory paired with a Fletcher (2013)-style theory of prudential good, then utility contributions will be something like the lifetime agglomeration of the agent’s achievement, friendship, happiness, pleasure, self-respect, and virtue. And so on.

5 Regions and Persons: Arntzenius’s Proposal

Arntzenius (2014) presents a different proposal for how to handle cases with infinitely many utility contributions.²⁴ Like Direct Difference Taking, Arntzenius’s proposal is based on a way of comparing pairs of acts in decision problems.

Here’s the rough idea behind Arntzenius’s proposal for how to compare acts. Pick some finite spatiotemporal region, and, taking only utility contributions in that region

²⁴As Arntzenius (2014) notes, one can think of his proposal as a modified version of Vallentyne & Kagan’s (1997) proposal, extended to apply to acts instead of outcomes.

into consideration, calculate the difference in expected utilities between the two acts. Then uniformly expand this region, and, taking only utility contributions in this larger region into consideration, calculate the difference in expected utilities between the two acts. Do this repeatedly for larger and larger regions. This yields a sequence of expected utility differences between the two acts. If at some point in this sequence one of the acts pulls into the lead and stays there, then it's the better act.

Now, let's spell out Arntzenius's proposal more carefully. Arntzenius's proposal for comparing pairs of acts can be thought of as a two step procedure.

The first step is to determine the utility associated with every finite spatiotemporal region of every potential outcome of a decision problem.²⁵ In the examples Arntzenius describes, this is determined by summing over all of the utility contributions in that region.

(One question that arises here is how to determine what the utility contributions "in" a spatiotemporal region are. For example, suppose we're preference utilitarians, who take positive/negative utility contributions to be satisfied/unsatisfied desires at a time. Suppose a subject in a small spatiotemporal region *S* – say, a 10 meter cube that persists for 10 seconds – has the desire at that moment to be remembered 100 years from now. And suppose that this desire is in fact satisfied – 100 years from now, there will be someone who remembers them. Should we take this positive utility contribution to be "in" *S*, and thus to count toward the utility associated with *S*? On the one hand, the subject having the desire is in this region, which suggests that it should contribute to *S*'s utility. On the other hand, whether the subject's desire is satisfied isn't determined by how things are in *S*.

Although Arntzenius doesn't address this issue, I take the latter approach – on which the utility contribution stemming from the subject's satisfied desire is not "in" *S* because *S* alone doesn't determine whether the desire is satisfied – to be preferable. This is because by spatiotemporally "spreading out" utility contributions more, it allows Arntzenius's proposal to be fruitfully applied to a broader range of cases.²⁶ So in what follows I'll say that a utility contribution is *in* a region *iff* the proposition describing the intrinsic features of that region entails that this utility contribution obtains. And for conciseness, I'll use

²⁵The restriction to finite regions is important. If we could determine the utility associated with infinite spatiotemporal regions – i.e., the entire outcome – then there'd be no need for the machinery Arntzenius brings in for the second stage.

²⁶E.g., suppose there are infinitely many planets, each labeled with a natural number. And suppose there is a narcissistic agent who desires of each planet that there be a statue of themselves on it (where each unsatisfied desire contributes 0 utility, and each satisfied desire contributes 1 utility). Finally, suppose this agent has two options: (a) making all planets have a statue of themselves on it, or (b) making only the even-numbered planets have a statue of themselves on it. Obviously, the agent should choose option (a) given their desires. And if we take each of the agent's desires to only contribute to regions that entail that the utility contribution obtains – e.g., a region containing both the agent and the planet in question – then every finite region will only contain finitely many utility contributions, and Arntzenius's proposal will yield the desired result that (a) is obligatory. But if we instead take the agent's infinitely many desires to all contribute to the region containing the agent, then both acts will lead the region containing the agent to have infinite utility, and Arntzenius's proposal will be unable to distinguish between the two acts. (Put another way, by taking these utility contributions to all lie in the region containing only the agent, we increase the prevalence of problematic cases in which there are finite regions with infinitely many utility contributions (see section 7.6).)

“the region entails x ” as shorthand for “the proposition describing the intrinsic features of the region entails x ”.)

Let R be the set of all of the finite spatiotemporal regions of outcomes in O . In what follows it will be convenient to encode the results of this first step with a function $u_{ST} : R \rightarrow \mathbb{R}$, where $u_{ST}(r) = x$ when region r is associated with utility x .

The second step of Arntzenius’s proposal for comparing pairs of acts uses these utility assignments and a way of generating paired expanding sequences of regions in different outcomes to rank these acts. Following Arntzenius, let’s assume that counterpart theory is the right way to pair regions in different outcomes. On Arntzenius’s proposal, any acceptable way of generating an expanding sequence of regions in the possible outcomes of a case can be characterized as follows:

Acceptable Expanding Sequence: Let an *acceptable expanding sequence* s , given a set of outcomes O , be a function from outcomes to sequences of spatiotemporal regions at those outcomes that satisfies the following conditions:

1. Let m_1, \dots, m_n be the spatiotemporal metrics of the spacetime of these outcomes.²⁷ The initial regions in each sequence are all the same shape. And there is a fixed distance d_i for each metric m_i such that, for every sequence, each successive region in the sequence contains all and only the points that are within distances d_i (for all i) of some part of the previous region in the sequence.^{28,29} (E.g., in the case of Newtonian spacetime, there are two metrics, a spatial distance metric m_s and a temporal distance metric m_t . Thus the initial regions are required to have the same spatial and temporal extensions at each outcome. And there is some fixed spatial distance d_s and temporal distance d_t such that each successive region in the sequence contains all and only points within d_s and d_t of some part of the previous region in the sequence.)
2. The sequences assigned to each outcome are such that the i th member of each sequence is an ordinary counterpart of the i th member of every other sequence.

Thus, roughly, an acceptable expanding sequence is a sequence of uniformly expanding spatiotemporal regions at each possible outcome such that the members of these sequences all line up with each other (in the sense of being the same size and being ordinary counterparts of each other).

Let $s_i(o)$ be the i th element that the acceptable expanding sequence s assigns to o . We can then characterize Arntzenius’s proposal as follows:

Arntzenius’s Proposal: Given a decision problem (A, O, cr, C, u) , let the *expected utility* of act $a \in A$ relative to an acceptable expanding sequence s and a stage in that sequence

²⁷For the purposes of this paper, I’ll use the term “metric” loosely to refer to any spatiotemporal distance function, be it a (proper) metric function, a pseudometric function, or some other nearby relative.

²⁸That is, there should be fixed distances d_i such that each successive region in a sequence contains all and only the points x such that for some point y in the previous region, $m_i(x, y) \leq d_i$ for all i .

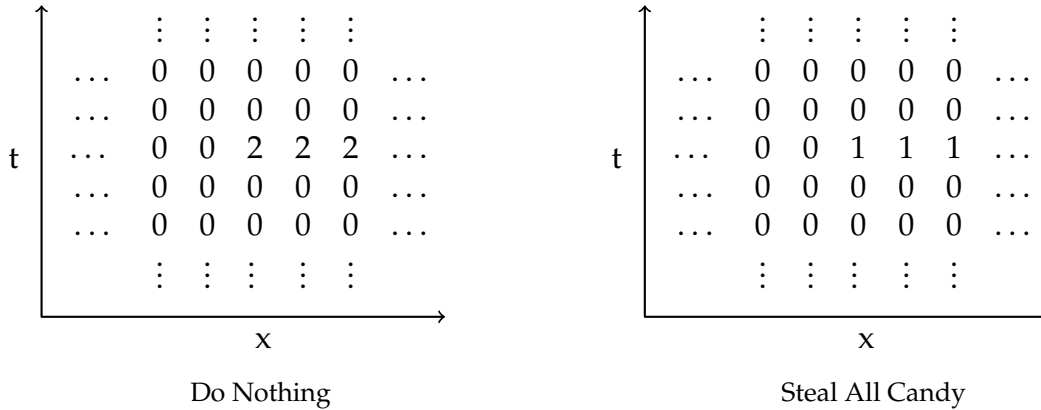
²⁹This description is a bit of a simplification; it describes Arntzenius’s suggestion for how to pick out regions in cases involving classical spacetimes, like Newtonian spacetime. Arntzenius suggests a different way of picking out regions in Lorentzian spacetimes (because the region containing all of the points within (say) 1 unit of spacetime interval of some point will be infinite). But these complications are orthogonal to the issues we’re concerned with, so I gloss over these complications in the text.

$i, EU_{s,i}(a)$, be:

$$EU_{s,i}(a) = \sum_{(j|o_j \in O)} cr(o_j : a) \cdot u_{ST}(s_i(o_j)). \quad (1)$$

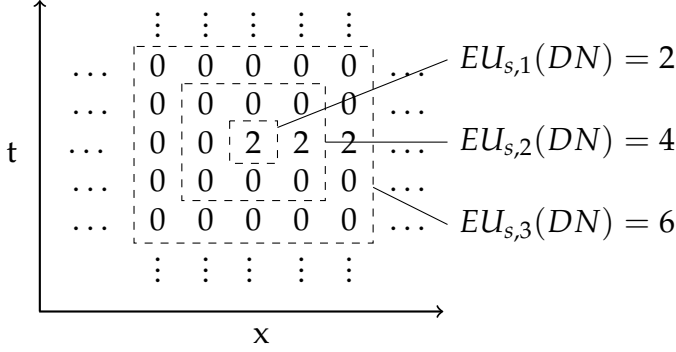
Given a decision problem, an act $a \in A$ is better than an act $b \in A$ ($a > b$) iff, for any acceptable sequence s , there is some point i in the sequence such that, for some real number x , and all $j > i$, $EU_{s,j}(a) > EU_{s,j}(b) + x$.³⁰

To get a better feel for this proposal, let's apply it to the Stealing All The Candy case discussed in section 3.2. In this case, recall, there are infinitely many babies with candy who will have a utility of 2 if you do nothing, and a utility of 1 if you steal all of their candy. Arntzenius's proposal requires details regarding the locations of the babies, so let's add to the case that all of these babies exist at the same time, and are spatially spread out uniformly in an infinite line starting at your current location. Thus we can present the two possible outcomes as follows (with the first baby located at the center of the graph):

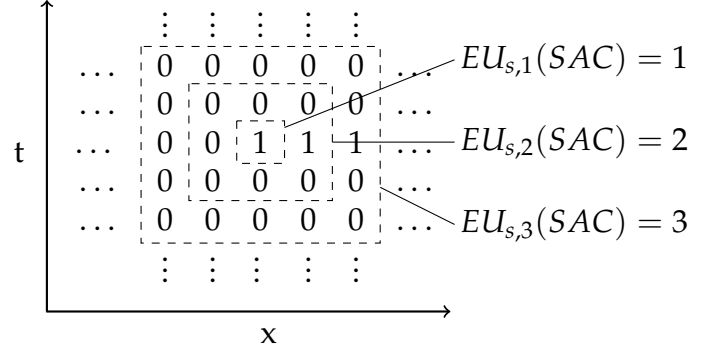


Now consider an acceptable expanding sequence that starts with the first baby, and which expands to include one additional baby at each step. This will yield the following $EU_{s,i}$ values:

³⁰The reason we require the difference to always be at least x , instead just requiring the difference to be positive, is that if the partial sums of these differences converges to zero from above (e.g., the partial sums are $1, 1/2, 1/4, 1/8, \dots$) we don't want to say that the first act is better than the second, since the difference between the two vanishes in the limit.



Do Nothing



Steal All Candy

The difference between these expected utilities at the i th step in this sequence is $EU_{s,i}(DN) - EU_{s,i}(SAC) = i$. It follows that there exists a positive real number x such that at some point i in this sequence the difference in expected utilities is always greater than x (e.g., $x = 1/2$ and $i = 1$). This suggests that Do Nothing is better than Steal All Candy.

To conclusively establish this, we need to show that every acceptable expanding sequence (e.g., ones that start off with different starting regions, ones that expand more or less rapidly, etc.) will yield the same result. In this case it's clear that every acceptable expanding sequence will yield the same result. Thus Arntzenius's proposal yields the desired result that Do Nothing is better than Steal All Candy, and thus that Steal All Candy is impermissible.

Now let's turn to compare the merits of Arntzenius's proposal to the merits of DDT. There are some objections that one could raise to both theories. Since these objections won't distinguish between the two theories, I'll delegate my discussion of these objections to appendix A. In the section 6 I'll consider an objection to DDT that Arntzenius's proposal avoids. In section 7 I'll consider objections to Arntzenius's proposal that DDT avoids.

6 Objections to Direct Difference Taking

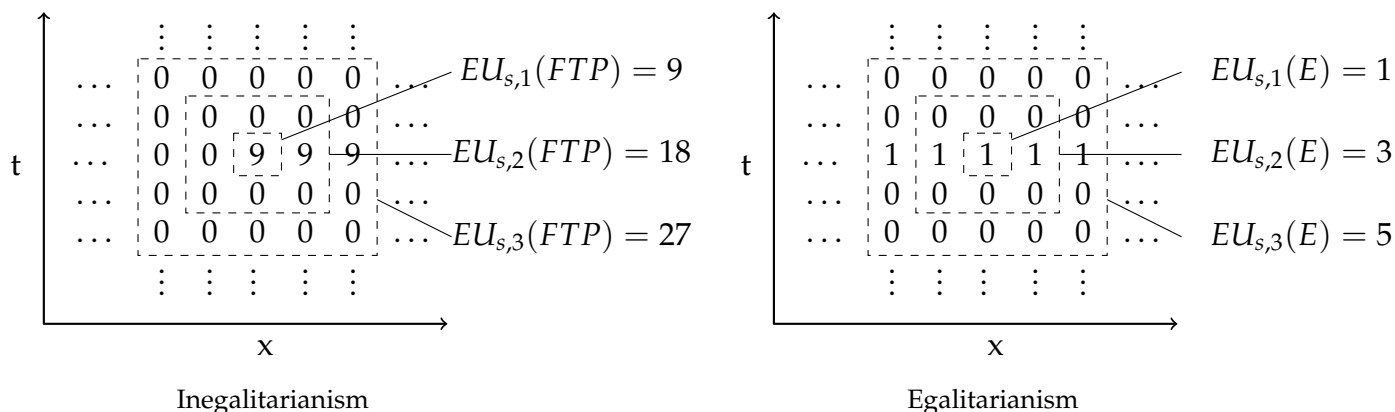
Arntzenius's proposal and DDT will yield similar results in a broad range of cases. But each proposal is subject to some objections that the other avoids. The main objection to DDT that Arntzenius's proposal avoids is that in certain cases DDT will arguably yield prescriptions that are insufficiently discriminating. Unlike DDT, Arntzenius's proposal employs the spatiotemporal structure of the outcomes to provide a privileged ordering and grouping of terms in sums over infinitely many utility contributions. Thus Arntzenius's proposal can provide more fine-grained prescriptions than DDT in cases where the order of these terms matters and the spatiotemporal structure allows us to privilege a particular ordering and grouping.

To get a feel for this, consider the following case:

Favoring the Positive. There are countably infinitely many babies, each named after one of the integers. These babies exist at the same time, and are equally spaced along

some spatial axis, ordered in accordance with the integer they're named after. You have two options. The first (Inegalitarianism) is to give a lot of candy to each baby named after a positive integer, in which case those babies will have a utility of 9, and the other babies a utility of 0. The second (Egalitarianism) is to give a little candy to each baby, in which case each baby will have a utility of 1.

Consider an acceptable expanding sequence that starts with a region containing only baby #1, and which expands uniformly to include two more babies at each step (one on the left, one on the right). This will yield the following $EU_{s,i}$ values:



The difference between these expected utilities at the i th step in this sequence is $EU_{s,i}(FTP) - EU_{s,i}(E) = 7i + 1$. It follows that there exists a positive real number x such that at some point i in this sequence the difference in expected utilities is always greater than x (e.g., $x = 7$ and $i = 1$). This suggests that Inegalitarianism is better than Egalitarianism.

Of course, there are some acceptable expanding sequences, such as ones which start with a region focused on a baby named with a large negative integer (a region far to the left of the one shown above), on which the first batch of expected utility differences will favor Egalitarianism. But as we continue to expand, we'll eventually hit the positive integer-named babies (or "positive babies", for short), and Inegalitarianism will pull into the lead. Since this will be true for any acceptable expanding sequence, Arntzenius's proposal yields the result that Inegalitarianism is better than Egalitarianism, and thus that Egalitarianism is impermissible.

Now suppose that Inegalitarianism gave the positive babies a utility of 1.1 instead of 9. Then, given the same sequence s used before, we'll find that the difference in expected utilities at the i th step is $EU_{s,i}(FTP) - EU_{s,i}(E) = -0.9i + 1$. This suggests that Inegalitarianism is worse than Egalitarianism, and indeed, this result will be borne out by every acceptable expanding sequence. Thus this change in utilities would yield a different verdict – Egalitarianism would now be obligatory, instead of impermissible.

These consequences of Arntzenius's account are appealing in two respects. First, these prescriptions are *prima facie* plausible. In the original case, Egalitarianism raises the utility of the negative babies by 1, but lowers the utility of the positive babies by 8. Since it's natural to think of these two groups of babies as being equally large, it's natural to think that Inegalitarianism is better than Egalitarianism. By contrast, in the altered version

of the case, Egalitarianism raises the utility of the negative babies by 1, and lowers the utility of the positive babies by 0.1, and it's natural to think that Egalitarianism is better than Inegalitarianism.

Second, these prescriptions are nicely fine-grained. That is, they're sensitive to the utilities in a way that results in different verdicts in the original case and the altered version of it. And this is appealing because it's natural to think that the differences in utilities in the original and altered versions of the case *should* make a difference.

Contrast this with how Direct Difference Taking will handle these cases. We can represent the original Favoring the Positive case as follows:

	...	Baby -1	Baby 0	Baby 1	Baby 2	...	Sum
Inegalitarianism	...	0	0	9	9	...	∞
Egalitarianism	...	1	1	1	1	...	∞
Difference	...	-1	-1	8	8	...	undefined

On some orderings, such as ones which alternate between negative and positive babies, the sum over differences will converge to ∞ . On other orderings, however, such as ones which interspace each positive baby with 10 negative babies, the sum over differences will converge to $-\infty$. Since this sum over differences is not permutation invariant, Direct Difference Taking won't deem either act impermissible. Thus according to DDT both options are permissible.

We'll get precisely the same verdicts in the altered version of the case in which Inegalitarianism only gives positive babies a utility of 1.1. For on some orderings (e.g., ones which interspace each negative baby with 11 positive babies) the sum over differences will converge to ∞ , and on other orderings (e.g., ones which alternate between negative and positive babies) it will converge to $-\infty$. Thus, again, both options are permissible.

Does this give us a *prima facie* reason to prefer Arntzenius's proposal to DDT?³¹ It depends on whether we take this apparent advantage of Arntzenius's proposal to actually be an advantage. For example, in the Favoring the Positive case, we might think that the intuition that Inegalitarianism is better than Egalitarianism seems to hang on the thought that the size of the set of babies named after positive integers is the same (in the relevant sense) as the size of the set of babies who are not. After all, if there were ten times as many positive babies as negative babies then our intuitions regarding this case would reverse. And one might reasonably question whether we should rely on the intuition that the positive and negative babies are the same size (in the relevant sense) in these kinds of cases.

7 Objections to Arntzenius's Proposal

Assume that utility contributions aren't directly assigned to spatiotemporal regions; a plausible assumption given that no theory of moral or prudential goods I know of has

³¹It's worth nothing that there are structurally similar cases in which Arntzenius's proposal also fails to give fine-grained prescriptions, such as the Intensity or Duration case discussed in appendix B.3.

proposed such a thing. Then one could raise a number of objections to Arntzenius’s proposal that won’t apply to DDT.³² In what follows, I’ll discuss seven objections to Arntzenius’s proposal, ordered, roughly, by how worrisome they are. And while some of the earlier problems can be avoided or mitigated by modifying the details of Arntzenius’s proposal, the latter problems point to deeper troubles with this approach.³³

7.1 Counterpart Relations Aren’t Identity Relations

Arntzenius’s proposal requires that we be able to construct expanding sequences of spatiotemporal regions in different outcomes, such that the i th member of each sequence is an ordinary counterpart of the i th member of every other sequence. But we can only do this if the ordinary counterpart relations behave like identity relations – e.g., are symmetric and transitive. If these relations aren’t symmetric or transitive, then there may be no expanding sequences where the i th region at each outcome is an ordinary counterpart of the i th region at every other outcome – for while one may be able to construct expanding sequences where the i th region at each outcome is a counterpart of the i th region in *some* other outcome, there may be no sequence where they’re all counterparts of each other. And this is problematic because ordinary counterpart relations over spatiotemporal regions generally *don’t* behave like identity relations – e.g., they’re not symmetric or transitive.³⁴

That said, this is not a deep problem for Arntzenius’s proposal. For we’ve already seen a way to address these problems. In section 3.5.1 we introduced the notion of *acceptable pairings* to deal with these kinds of violations. And by modifying Arntzenius’s proposal so that it appeals to acceptable pairings instead of ordinary counterpart relations, Arntzenius’s proposal can circumvent these problems as well.

³²What if one *does* adopt the implausible stance that utility contributions are directly assigned to spatiotemporal regions? Then many of these objections (e.g., the objections raised in sections 7.4, 7.5, 7.6, and 7.7) will fail to differentiate between DDT and Arntzenius’s proposal, since they’ll apply to both proposals or neither proposal. That said, simply endorsing a picture of utility contributions which directly assigns values to spatiotemporal regions won’t suffice to shield Arntzenius’s proposal from these criticisms. For part of the project Arntzenius and I are engaged in is to provide a way to yield plausible infinite generalizations of moral theories. And if Arntzenius’s proposal has difficulty doing so in a plausible way for preference utilitarians, hedonic utilitarians, and the like, then this is a mark against his proposal for how to generalize moral theories.

³³Arntzenius (2014) briefly considers two other approaches to handling cases with infinitely many utility contributions, which he calls the *Weak Person Condition* and the *Weak Location Condition*. Although these proposals are interesting, it would distract too much from the comparison of Arntzenius’s proposal and DDT to discuss them in the text. So I delegate my discussion of these views to appendix B, in which I describe these proposals, and assess the pros and cons of these proposals relative to Arntzenius’s proposal and DDT

³⁴Here’s a simple case in which ordinary counterpart relations over spatiotemporal regions fail to be symmetric. Consider a pair of spatially finite outcomes, o and o' , where o' is twice the size of o . The first half of o' , r_1 , is a perfect duplicate of the region at o , r . The second half of the o' , r_2 , is an almost perfect duplicate of r . In this case, r_2 ’s counterpart at o is r . But r_2 is not r ’s counterpart at o' , r_1 is. (Similar cases yield counterexamples to transitivity.)

7.2 Island Universes

Arntzenius's proposal runs into problems dealing with *island universes*, worlds with multiple disconnected spatiotemporal regions.³⁵ This is because of the way Arntzenius's proposal works: it starts with some initial region, and then uniformly expands it to include points that are nearby. Thus if the initial region starts in one "island", expansions from that region won't reach any other island, and so won't take their utility contributions into account.

Of course, Arntzenius's proposal doesn't require initial regions to be located at one island. There will be some acceptable expanding sequences with initial regions that aren't contiguous, and which have parts located at every island. And such expanding sequences will take utility contributions at every island into account. But this is of little help. For there will also be some acceptable expanding sequences with initial regions which are located at only one island, and these sequences won't take utility contributions at other islands into account. This is problematic because Arntzenius's proposal only yields the result that an act is impermissible if it's worse than some other act according to *every* acceptable expanding sequence. And thus Arntzenius's proposal will judge some clearly impermissible acts permissible.

For example, consider a version of the Stealing All The Candy case from section 3.2 at an island universe with two islands. In this variant, the situation described in the original case is what happens at the agent's island; the other island consists of empty space regardless of what the agent does. On Arntzenius's proposal, one of the acceptable expanding sequences will start with an initial region located at the second island. And that expanding sequence won't take utility contributions at the agent's island into consideration. Thus there are some acceptable expanding sequences according to which the limiting difference between the expected utilities of the two acts is 0. And so, on Arntzenius's proposal, we won't get the desired result that Steal All Candy is impermissible.

We can avoid these kinds of problems by modifying Arntzenius's proposal again. In particular, we can require that the initial regions of acceptable expanding sequences include some part of every disconnected spatiotemporal region (island) that exists at that world. By adding this requirement, we can ensure that utility contributions at all of the islands at a world are taken into account.³⁶

³⁵Note that one can't dismiss this problem by suggesting that such possibilities are unlikely. For if there are possible agents with non-zero credences in such possibilities, then one's proposal needs to be able to accommodate such possibilities (regardless of how implausible one might personally take them to be). Moreover, it's not entirely clear that we shouldn't take such possibilities seriously; see Bricker (1996) and Bricker (2001) for discussions of such possibilities.

³⁶Of course, this modification won't help with worlds with infinitely many islands. Since the initial region is required to be finite, worlds with infinitely many islands will require the parts of each island in the initial region to get arbitrarily small (e.g., a cubic meter of the first island, half a cubic meter of the second, a quarter of a cubic meter of the third, and so on). But once we perform our first uniform expansion of such a region (e.g., adding all points within a centimeter of the previous region), we'll get a region that's infinite. And since the first stage of Arntzenius's proposal only assigns values to finite regions, such infinite regions won't be assigned values. As a result, the second stage of Arntzenius's proposal (which employs the first stage's assignment of values to regions to evaluate acts) won't get off the ground. Now, one could try to modify the first stage of Arntzenius's proposal so that it also assigns values to some infinite regions as well. But this raises other problems for Arntzenius's proposal, since such regions will generally have infinitely many utility contributions,

7.3 Infinite Regions with Holistic Utility Contributions

Arntzenius's proposal won't take into consideration utility contributions which are holistically determined by infinite regions – contributions which don't show up when we consider any finite region. This leads it to yield implausible verdicts in such cases.

Suppose subjective preference utilitarianism is true – that you should act to maximize expected desire satisfaction. Consider the following case:

Remembrance. The world is infinitely extended, spatially and temporally, and contains infinitely many subjects with various desires. You can choose to bring about one of two outcomes. These outcomes are the same in all respects relevant to desire satisfaction, with one exception: the second outcome will bring about an additional subject with one desire – to be remembered forever – and that desire will be satisfied.

In the second outcome of the Remembrance case, no finite region will entail that the subject's desire to be remembered forever is satisfied. And since every region in an acceptable expanding sequence is either an initial region (which is finite) or a finite expansion of an initial region, every region in an acceptable expanding sequence is finite. It follows that no region in an acceptable expanding sequence will entail that the subject's desire is satisfied. Thus on Arntzenius's proposal the expected utilities of these two options will always be the same; for the difference between their outcomes only shows up when we consider a temporally infinite region, which no acceptable expanding sequence does. As a result, Arntzenius's proposal will take both of these options to be permissible. But this is the wrong verdict – it should be obligatory to bring about the second outcome, as it has all of the desire satisfaction of the first outcome, plus an additional satisfied desire.³⁷

and Arntzenius's proposal is ill-equipped to handle such regions (cf. section 7.6).

A different requirement one might consider is requiring the initial region of acceptable expanding sequences to be located in the island in which the agent is located, leading the proposal to effectively ignore other islands. After all, the agent's acts will generally be probabilistically and causally independent of the utility contributions at other islands. Given this, it seems unproblematic to simply ignore other islands, since they won't bear on what the agent should do. But this modification won't be able to accommodate more exotic possibilities in which utility contributions at other islands aren't probabilistically or causally independent of the agent's acts. (Here are some examples of such cases. EX 1. Suppose one deduces from the laws that there exists an island that's a duplicate of this one. Then one's acts won't be probabilistically independent of what happens at that island. EX 2. Suppose a counterfactual analysis of causation is correct, and that certain counterfactuals hold between your acts and outcomes in other islands. Then one's acts won't be causally independent of what happens at other islands. Of course, one might question whether a *temporally* disconnected region could have any causal relations to this one. But those who are attracted to counterfactual analyses of causation may feel some pressure to admit the possibility of non-temporally mediated causal relations, since the appropriate counterfactual relations can hold despite the lack of temporal relations. See Baron & Miller (2014) and Baron & Miller (2015) for defenses of such a view.)

³⁷One might be tempted to escape this problem by adopting the alternate account of when a utility contribution is in a region discussed in section 5, on which utility contributions due to satisfied desires lie in the region containing only the subject of the desire. This alternate account would allow Arntzenius's proposal to handle Remembrance, since it would now be easy to find finite regions containing the additional subject's utility contribution. But this temptation should be resisted. First, as mentioned earlier, this alternate account is unattractive for other reasons (cf. footnote 26). Second, there are other cases where this alternate account won't help. For example, suppose subjective hedonic utilitarianism is true – you should act to maximize expected happiness – and that some functionalist account of the mind is correct. And consider the following case:

Again, we can modify Arntzenius’s proposal to mitigate this problem, though doing so starts to stray from the spirit of Arntzenius’s original proposal. Let a *holistic utility contribution* at an outcome be a utility contribution that isn’t entailed by any finite region at that outcome, but that is entailed by the outcome as a whole. And let the value of $EU_{s,i}$ now be the expected sum of the utility contributions in the i th region of s (as before) *plus* the expected value of all of the holistic utility contributions at that outcome. With this modification, Arntzenius’s proposal will yield the desired results in cases with finite numbers of holistic utility contributions.

What about cases in which there are infinitely many holistic utility contributions? In these cases, we’ll want to assess the *differences* between these holistic contributions at different outcomes (for the kinds of reasons described in section 3.1), require the resulting sum to be order-independent (for the kinds of reasons described in section 3.3), ensure that we’re taking unpaired contributions into account in the right way (for the kinds of reasons described in sections 3.4), and ensure that we’re pairing contributions in the right way (for the kinds of reasons described in section 3.5). I.e., we’ll want to employ something like a restricted version of Direct Difference Taking.

The resulting account will yield satisfactory verdicts. But this account is implausibly disjunctive, effectively using Arntzenius’s proposal to assess non-holistic contributions and Direct Different Taking to assess holistic contributions. And I take it that few would be attracted to such an *ad hoc* proposal.

7.4 Outcomes with Different Spacetime Structure

Arntzenius’s proposal assumes that all of the outcomes in a decision problem have spacetimes with the same structure. For example, consider the requirement that the initial region in the expanding sequences at different outcomes be the same size. If one outcome has a Newtonian spacetime (in which there are distinct spatial and temporal metrics) and another has a Lorentzian spacetime (in which there’s only one metric, the spacetime interval), then the requirement that the initial regions at these outcomes be the same size – i.e., have the same metric extension – is impossible to satisfy. Likewise, consider the requirement that the expanding sequences at different outcomes expand at the same rate. If one outcome has Newtonian spacetime and the other has Lorentzian spacetime, then this requirement is impossible to satisfy. Thus Arntzenius’s proposal won’t work in such cases.

This is problematic. Agents can be uncertain about what the structure of spacetime is. For example, the physicists engaged in the initial debate regarding Special Relativity were

Big Brain. The world is infinitely extended, spatially and temporally, and contains infinitely many subjects with various degrees of happiness. You can choose to bring about one of two outcomes. These outcomes are the same in almost every respect relevant to happiness, with one exception: at the second outcome, the distribution of subjects in the world is such that there’s a further emergent consciousness – an infinitely extended “big brain” (along the lines of Block’s (1978) China Brain example) – which is happy.

On either account of when a utility contribution is in a region, Arntzenius’s proposal will be indifferent between these two options, since the difference in the utility contributions of their outcomes will be one that no acceptable expanding sequence will take into account. Again, this is the wrong verdict – it’s obligatory to bring about the second outcome, as it has all of the happiness of the first outcome, plus the happiness of the big brain.

uncertain about whether spacetime was Newtonian or Lorentzian. And when agents are uncertain about the structure of spacetime, their decision problems will have outcomes with different spacetime structure. So Arntzenius's proposal won't work for such agents.

The natural way to modify Arntzenius's proposal in light of this problem is to only require coordination between expanding sequences at outcomes with the same spacetime structure. Thus we'd only require initial regions *at outcomes with the same spacetime structure* to be the same size. Likewise, we'd only require expanding sequences *at outcomes with the same spacetime structure* to expand at the same rate.

This modification allows Arntzenius's proposal to apply in cases where there are outcomes with different spacetime structures. And in most cases, this modified proposal will yield plausible prescriptions. But there are some cases in which this proposal will yield implausible verdicts; namely, cases in which the agent believes their actions and the structure of spacetime are probabilistically or causally dependent. (For example, perhaps the agent is considering whether to activate a scientific device that (they believe) will change the nature of spacetime. Or perhaps the agent believes that the different spacetime structures are naturally associated with different laws, and that these laws bear on how agents with cognitive structures like their own are likely to behave. Or perhaps the agent believes they're in a Newcomb-like situation, where God's choice about what spacetime structure to give the world was determined by God's prediction about what the agent's decision in this case would be.) And if we'd like our theory to yield plausible verdicts in general, then we'd like it to yield plausible verdicts in these cases.

For an example of how this modified version of Arntzenius's proposal yields implausible verdicts, consider the following variant of the Stealing All The Candy case from section 3.2:

Stealing All The Candy (ST). There are countably infinitely many babies, each of which has some candy. These babies are equally spaced along a line starting from your current location. You have two options. The first (Do Nothing) is to leave the babies as they are, in which case each will have a utility of 2. The second (Steal All Candy) is to steal candy from all of them, decreasing their utility from 2 to 1. Finally, you're unsure about whether spacetime is Newtonian or Lorentzian. But you think the world is more likely to be Newtonian than Lorentzian if you decide to Do Nothing (0.9/0.1), and equally likely to be Newtonian or Lorentzian if you Steal All Candy (0.5/0.5).

To simplify, let's assume there are only four outcomes we need to consider: a pair of Newtonian outcomes in which you Do Nothing (*N&DN*) or Steal All Candy (*N&SAC*), and a pair of Lorentzian outcomes in which you Do Nothing (*L&DN*) or Steal All Candy (*L&SAC*).

There will be one acceptable expanding sequence, *s*, in which the initial region at all four outcomes includes only the first baby in the line, and where each subsequent region contains one more baby. Given this sequence, the utility in the *i*th region at the Do Nothing outcomes will be twice the number of babies in that region, or $2i$. The utility in the *i*th region at the Steal All Candy outcomes will be equal to the number of babies in that region, or i . Thus the difference in expected utilities between Do Nothing and Steal All

Candy at the i th step in sequence s , s_i , will be:

$$\begin{aligned}
EU_{s,i}(DN) - EU_{s,i}(SAC) &= (cr(N : DN) \cdot u_{ST}(s_i(N\&DN)) + cr(L : DN) \cdot u_{ST}(s_i(L\&DN))) \\
&\quad - (cr(N : SAC) \cdot u_{ST}(s_i(N\&SAC)) + cr(L : SAC) \cdot u_{ST}(s_i(L\&SAC))) \\
&= (0.9 \cdot 2i + 0.1 \cdot 2i) - (0.5 \cdot i + 0.5 \cdot i) \\
&= i,
\end{aligned}$$

which suggests that Do Nothing is better than Steal All Candy, as desired.

But there will also be another acceptable expanding sequence, s^* , in which at the Newtonian outcomes the initial region includes only the first baby, and each subsequent region contains one more baby, but in which at the Lorentzian outcomes the initial region includes the first ten babies, and each subsequent region contains ten more babies. Given this sequence, the utility in the i th region at Newtonian outcomes will be $2i$ if Do Nothing and i if Steal All Candy. And the utility in the i th region at the Lorentzian outcomes will be $20i$ if Do Nothing and $10i$ if Steal All Candy. Thus the difference in expected utilities between Do Nothing and Steal All Candy at the i th step in sequence s^* , s_i^* , will be:

$$\begin{aligned}
EU_{s^*,i}(DN) - EU_{s^*,i}(SAC) &= (cr(N : DN) \cdot u_{ST}(s_i^*(N\&DN)) + cr(L : DN) \cdot u_{ST}(s_i^*(L\&DN))) \\
&\quad - (cr(N : SAC) \cdot u_{ST}(s_i^*(N\&SAC)) + cr(L : SAC) \cdot u_{ST}(s_i^*(L\&SAC))) \\
&= (0.9 \cdot 2i + 0.1 \cdot 20i) - (0.5 \cdot i + 0.5 \cdot 10i) \\
&= -1.7i,
\end{aligned}$$

which suggests that Steal All Candy is better than Do Nothing. Since s and s^* yield different results, neither act will be ranked above the other. Thus we'll get the implausible verdict that Steal All Candy is permissible.

7.5 Outcomes Without Metric Structure, or Without Spacetime

Arntzenius's proposal presupposes that all of the outcomes in a decision problem have spacetime with metric structure. In order to construct an acceptable expanding sequence, we need a way of matching regions of the same size in different outcomes. And if an outcome has a spacetime without metric structure, or no spacetime at all, then we can't construct acceptable expanding sequences. As a result, we can't apply Arntzenius's proposal to cases with such outcomes.

One way of responding to this worry is to suggest that the kinds of cases in which this worry arises are unrealistic or *recherché* cases that agents like us never encounter.³⁸ This response is unsatisfactory because the claim it makes is false. All that's required for this problem to arise is that an agent have a non-zero credence that the world doesn't

³⁸One might be tempted to reply that cases with infinitely many utility contributions are themselves *recherché* cases. Thus if we're bracketing such cases, we should bracket the kinds of cases that motivate adopting Arntzenius's proposal in the first place. But this is too quick; as we saw in section 1, cases with infinitely many utility contributions *are* cases that agents like us encounter. Moreover, many of the debates in decision theory (e.g., the debate between causal and evidential decision theorists) already concern *recherché* cases (e.g., Newcomb's Case) that agents like us never encounter (e.g., see Horwich (1985)). But these debates are nonetheless taken to be important because they bear on what the correct theory of decision-making is.

have metric structure. And there are plausibly reasonable agents who do have non-zero credences in such possibilities.

For example, consider those who are at least somewhat moved by the kind of configuration space realism defended by Albert (2015), in which the three dimensional space of our acquaintance is effectively illusory.³⁹ One might follow Albert and take this view to entail that real space is an extremely high-dimensional space quite unlike the one we think we live in (i.e., the configuration space in which the wave function lives). Or one might take this view to entail that there isn't anything that deserves the name "space", and thus that there isn't anything that corresponds to spatial metric structure.⁴⁰ In a similar vein, one might take certain results in string theory to suggest that, if these theories are correct, then the ordinary spacetime of our acquaintance is effectively illusory.⁴¹ Either way, we would like our proposal to apply to agents who take such possibilities seriously.

A second way of responding to this worry is to modify Arntzenius's proposal along the lines sketched in section 7.4. That is, we might handle outcomes with spacetimes without metric structure by allowing any choice of initial region and expansions at such outcomes to yield acceptable expanding sequences. And we might handle outcomes without spacetime by allowing any set of utility contributions to count as an "initial region" of an acceptable expanding sequence, and any superset of these utility contributions to count as an expansion. But these moves lead to problems similar to the ones we saw in section 7.4. The problem in section 7.4 was that we'll get implausible prescriptions in certain cases – cases in which the agent believes the metric structure of spacetime is probabilistically or causally dependent on their act. The problem here is worse, since we'll get implausible prescriptions even if the agent believes the metric structure of spacetime isn't probabilistically or causally dependent on their act.

To see this, consider a version of the Stealing All The Candy case where you're unsure about whether spacetime has metric structure, and where you take this fact to be probabilistically and causally independent of your act. There will be acceptable expanding sequences such that at outcomes without metric structure the initial region and subsequent expansions at the Do Nothing outcomes include more babies than the initial region and subsequent expansions at the Steal All Candy outcomes. Likewise, there will be acceptable expanding sequences such that at outcomes without metric structure the initial region and subsequent expansions at the Do Nothing outcomes include *fewer* babies than the initial region and subsequent expansions at the Steal All Candy outcomes. These different acceptable expanding sequences won't agree on how to rank acts, and so we'll get the implausible verdict that both acts are permissible.

³⁹"The new picture turns everything elegantly inside out. What had seemed merely abstract and symbolic in the old picture (that is, the wave functions, and the high-dimensional space in which they undulate) becomes real and physical and concrete in the new one, and what had seemed exact and fundamental in the old picture (that is, the talk of particles in a three-dimensional space) becomes vague and approximate and emergent in the new one." (Albert (2015), p146.)

⁴⁰For example, if one rejects metaphysical indeterminacy, then Albert's argument that three-dimensional space is "vague and approximate" (Albert (2015), p146) suggests such a stance.

⁴¹See Allori (2016) and the references therein. As before, one might take these results different ways; one might take them to suggest that the ordinary four dimensional spacetime of our acquaintance still exists, if only as an emergent non-fundamental entity (Huggett & Wüthrich (2013)). Or one might take them to suggest that ordinary spacetime doesn't exist.

7.6 Finite Regions with Infinitely Many Utility Contributions

As we noted in section 5, Arntzenius's proposal for comparing acts can be thought of as a two step procedure. The first step determines the utilities associated with finite spatiotemporal regions at each outcome, while the second step uses these utilities to compare pairs of acts. Focus on the first step. In Arntzenius's paper, the utility associated with a finite region is determined by summing over the individual utility contributions in that region. But this is problematic in cases in which there are finite regions with infinitely many utility contributions. For the question of how to handle finite regions with infinitely many utility contributions is effectively just a version of the question we started with – how to handle outcomes with infinitely many utility contributions. And all of the worries that arise for assigning utilities to outcomes by summing over their utility contributions will also arise for assigning utilities to regions by summing over their utility contributions.

For example, suppose preference utilitarianism is true, and consider a case with a subject with infinitely many satisfied (and no unsatisfied) desires about some finite region they're in. You can either choose to do nothing, or to give them a drug that halves the strength of those desires, thus halving their utility contributions. If we assign utilities to regions by summing over the contributions in that region, we'll assign the same utility (∞) to this region at both outcomes, and both options will be deemed permissible.⁴² But that seems like the wrong answer.

For another example, consider a case in which you can either create no one, or create a subject with infinitely many desires about some finite region they're in, which are satisfied and not satisfied in a way that yields utility contributions corresponding to the alternating harmonic series. Since this sum is order dependent, it's not clear what utility to assign to this region.⁴³

Thus Arntzenius's proposal is incomplete, and incomplete in a way that threatens to undermine the motivation for adopting it. For Arntzenius's answer to the question of how to deal with outcomes with infinitely many contributions presupposes an answer to the question of how to deal with finite regions with infinitely many contributions. But the same kinds of challenges are raised by both questions. So in order to put Arntzenius's proposal on firm footing, we need to find an independent way of dealing with these challenges. Now, one could do this by (say) employing a different account in such cases, such as Direct Difference Taking. But such an account would be implausibly disjunctive. And I think few would find such an *ad hoc* proposal appealing.

7.7 Spreading Utility Cases

Because Arntzenius's proposal compares the utilities associated with spatiotemporal regions instead of comparing utility contributions directly, it takes the spatiotemporal dis-

⁴²Assuming here a notion of convergence to extended reals along the lines sketched in section 3.2.

⁴³Does DDT do better in this case? It does, because DDT tells us how to obtain verdicts in this case. First we pair the utility contributions (in this case, desires) in the two different outcomes. Since there are no desires in the first outcome, this will be a case of zero-point pairing. We then consider the infinite sum over the differences between these utility contributions, which is identical to the sum of the utility contributions in the second outcome. Since the value of this sum is order dependent, we'll get the result that both options are permissible. (I owe an anonymous referee for encouraging me to address this question.)

tribution of utility contributions to bear on what actions we should prefer in implausible ways.

For example, suppose hedonic utilitarianism is true, and consider the following case:

Population Density. There are countably infinitely many people who all exist at the same time, and who need to be assigned a place to live. You have two options. The first (Dense) is to divide space into an infinite grid of 1 meter cubes and cram each person into one of these cubes. In this case, each person will have a cramped and miserable existence; each will have utility of $1/1000$. The second (Sparse) is to divide space into an infinite grid of 1 kilometer cubes and assign each person to one of these cubes. In this case, each person will have a relaxed and idyllic existence; each will have a utility of 1000.

In the Sparse outcome each person is much happier than in the Dense outcome; each person has a million times more utility. As a result, it seems like Sparse should be obligatory. And proposals like Direct Difference Taking, which directly track utility contributions, will yield this result.

But the utility *density* in the Dense outcome is much greater than in the Sparse outcome; there's a thousand times more utility per unit area. And Arntzenius's proposal effectively assesses utility density. On Arntzenius's proposal, every acceptable uniform expansion will favor Dense over Sparse. As a result, Arntzenius's proposal will yield the verdict that Dense is obligatory. This is implausible.⁴⁴

8 Conclusion

Cases with infinitely many utility contributions pose a challenge to expected utility-based theories. And these are cases we need to take seriously, for there are contemporary cosmological theories that suggest that there are infinitely many utility contributions at the actual world. I've motivated and defended a proposal – Direct Difference Taking – for how to deal with such cases. And I've compared it to what I take to be the most plausible alternative on offer, the proposal offered by Arntzenius (2014).

Arntzenius's proposal embodies a natural idea, and has the arguable advantage of being able to offer more fine-grained prescriptions in certain kinds of cases. But Arntzenius's proposal also faces a number of worries. And while some of these worries can be mitigated by revising Arntzenius's proposal, the most serious worries are ones that will

⁴⁴In the comparing outcomes literature, Garcia & Nelson (1994) and Cain (1995) raise the worry that views which appeal to a privileged temporal order are effectively concerned with how quickly utility is produced, rather than how much utility there is. One can see this worry as being in a similar vein.

One might wonder whether Population Density is essentially a Repugnant Conclusion-style worry (see appendix A for discussion), and thus something that can be recreated as an objection to DDT as well. (I owe an anonymous referee for encouraging me to address this worry.) Despite some superficial similarities, Population Density is quite different from Repugnant Conclusion-style worries. Population Density is a same-numbers case, in which the same individuals (and the same number of individuals) exist regardless of what you do. This contrasts with different-numbers cases like those that yield the Repugnant Conclusion. Since both Arntzenius's proposal and DDT are subject to Repugnant Conclusion-style worries, we can bracket those worries for the purposes of comparison. And after bracketing such worries, same-numbers case like Population Density raise further problems for Arntzenius's proposal that do not arise for DDT.

afflict any approach along these lines. Thus I take Direct Difference Taking to be the most attractive way of handling cases with infinitely many utility contributions.⁴⁵

A Objections to Both Proposals

I’ve discussed objections that apply to DDT but not Arntzenius’s proposal in section 6, and objections that apply to Arntzenius’s proposal but not DDT in section 7. In this appendix, I’ll present and respond to two objections that apply to both proposals.⁴⁶

Here is the first objection. As Parfit (1984) notes, aggregative theories like utilitarianism are generally subject to Repugnant Conclusion-style worries. For such theories will generally entail that an outcome with a large number of very high utility people is worse than an outcome with a much larger number of low utility people, as long as the latter outcome has more overall utility. For example, they’ll generally hold that given a choice between a million people with 100 utility and a trillion people with 0.01 utility, we’re obligated to choose the latter. And most find this conclusion repugnant.

Both Arntzenius’s proposal and DDT will yield this verdict. But this is not a demerit of these proposals. For both Arntzenius’s proposal and DDT are, in part, attempts to describe how one should generalize moral theories like utilitarianism so that they yield plausible verdicts in infinite cases. To qualify as infinite generalizations of utilitarianism, these proposals must yield the same verdicts as utilitarianism in finite cases. Since utilitarianism yields Repugnant Conclusion-style worries in finite cases (like the case described above), any infinite generalization of utilitarianism must yield these verdicts as well. Thus Arntzenius’s proposal and DDT must yield these verdicts, given what they’re trying to do.⁴⁷

Here is the second objection. Hamkins & Montero (2000) suggest that a plausible condition on a satisfactory theory is that if there’s a value-preserving bijection between the utility contributions of a pair of outcomes, then these outcomes should be treated interchangeably. Call this condition *Isomorphism*. Now suppose utility contributions are assigned to complete lives, and consider the following case:⁴⁸

	...	Person i	Person j	Person k	Person l	Person m	...	Sum
Act 1	...	-1	0	1	2	3	...	undefined
Act 2	...	-2	-1	0	1	2	...	undefined
Difference	...	1	1	1	1	1	...	∞

⁴⁵I’d like to thank Maya Eddon, Itai Sher, two anonymous referees, participants of the Spring 2018 UMass Brown Bag group, participants of the 2018 Recent Work in Decision Theory and Epistemology conference at Columbia, and the audience of my 2019 colloquium talk at the University of Maryland, College Park, for helpful comments and discussion. I’d also like to thank Sophie Horowitz for suggesting both the title of this paper and the title of section 5. Finally, I’d like to thank Frank Arntzenius for inspiring my interest in these issues and many others.

⁴⁶I owe an anonymous referee for encouraging me to address these objections.

⁴⁷There are, of course, a number of alternatives to utilitarianism that one might consider to try to avoid these worries, such as averaging views, critical levels views, and the like. For an in-depth discussion of these options, see Arrhenius et al. (2017).

⁴⁸This is a variant of case discussed by Vallentyne & Kagan (1997).

In this case DDT says that only Act 1 is permissible, even though there's a value-preserving bijection between them. Thus DDT will violate Isomorphism. And if we assume all of the people in these outcomes are equally spaced along some spatial axis, and occupy the same locations regardless of which act one performs, then Arntzenius's proposal will yield the same prescriptions as DDT in this case. Thus Arntzenius's proposal violates Isomorphism as well.

I concede that there is some force to this objection. But as Hamkins & Montero (2000) and Lauwers & Vallentyne (2004) note, what this shows is that in these kinds of infinite cases we're forced to choose between two plausible but conflicting principles: Isomorphism and dominance. Act 1 and Act 2 are isomorphic, so Isomorphism requires both acts to be permissible. But Act 1 dominates Act 2, so dominance requires Act 2 to be impermissible. So we have to reject either Isomorphism or dominance. Arntzenius and I follow Lauwers & Vallentyne (2004) in taking dominance considerations to be more compelling than Isomorphism.⁴⁹ Thus while both Arntzenius's proposal and DDT reject Isomorphism, I take this rejection to be defensible, as it's required in order to hold on to dominance.

B The Weak Person Condition and the Weak Location Condition

In section 10 of his paper, Arntzenius (2014) briefly mentions two other proposals for how to handle cases with infinitely many utility contributions. In what follows I'll describe these proposals, and compare them to Arntzenius's proposal and DDT.⁵⁰

B.1 The Weak Person Condition

The first proposal is what Arntzenius calls the *Weak Person Condition* (WPC). Like Arntzenius's proposal, WPC has two stages: first it agglomerates utility contributions to assign values to temporally extended individuals (i.e., people's lives), and second it lays out a procedure for using those values to assess the available acts. As with Arntzenius's proposal, the first stage is largely left implicit, though it is presumably determined by something like summing over utility contributions associated with the subject's life. The second stage of assessment is effectively the conjunction of difference-taking (as described in section 3.1.1, but applied to people's lives instead of utility contributions), infinite convergence (section 3.2.1), permutation invariance (section 3.3.1), zero-point pairing (section 3.4.1), and expectation taking (section 3.6.1).

⁴⁹Both Arntzenius's proposal and DDT respect dominance considerations, but the dominance considerations they respect are not identical. Arntzenius's proposal respects dominance with respect to spatiotemporal regions (e.g., if there are two outcomes with the same spatiotemporal regions, and in one outcome every region is at least as good, and sometimes better, than the corresponding region in the other outcome, then the first outcome is better than the second). DDT respects dominance with respect to utility contributions, which in this case we're assuming to be individuals.

⁵⁰I own an anonymous referee for encouraging me to discuss these proposals.

WPC differs from Arntzenius's proposal in that its first stage agglomerates utility contributions associated with people's lives instead of spatiotemporal regions, it employs zero-point pairing, and it does not employ the expanding sequence structure Arntzenius's proposal employs.

WPC differs from DDT in that it has two stages (agglomerating over utility contributions to determine the utility of people's lives, and then using these values to assess acts) instead of one (using the utility contributions to directly assess acts), and does not employ acceptable pairings to overcome the problems that arise due to the failure of counterpart relations to be symmetric, transitive, or one-to-one (as in cases of fission or fusion). It also leaves the order of some of the relevant quantifiers ambiguous in a way that might yield a further deviation from DDT. (For example, does WPC tell us that an act B is permissible *iff* there doesn't exist an act A such that, for every acceptable counterpart pairing, $EU(A) > EU(B)$? Or does WPC tell us that an act B is permissible *iff* for every acceptable counterpart pairing, there doesn't exist an act A such that $EU(A) > EU(B)$? These two disambiguations can yield different verdicts, and only the latter will line up with DDT.)

B.2 The Weak Location Condition

The second proposal is what Arntzenius calls the *Weak Location Condition* (WLC). Roughly, WLC can be seen as a compromise between WPC and Arntzenius's proposal. Like both of these views, WLC has two stages: first it agglomerates utility contributions to assign values to spatiotemporal regions, and second it lays out a procedure for using those values to assess the available acts. And the second stage of assessment is effectively the conjunction of difference-taking (as described in section 3.1.1, but applied to spatiotemporal regions instead of utility contributions), infinite convergence (section 3.2.1), permutation invariance (section 3.3.1), and expectations taking (section 3.6.1).

Since Arntzenius only briefly describes WLC, we're not told what the locations we're summing over and comparing are. To make things concrete, let's fill in these details by assuming that the regions WLC sums over must be mutually exclusive and exhaustive, that these regions must be finite, and that *if* the spacetime structure is the same in a pair of outcomes, then the counterpart relations between such outcomes need to pair regions of the same size.

WPC differs from WLC in assigning values to spatiotemporal regions instead of people's lives, and not appealing to zero-point pairing. WPC differs from Arntzenius's proposal in that it doesn't appeal to Arntzenius's machinery for finding initial regions in the outcomes of the same size, uniformly expanding them, and seeing what the difference between the utility of these regions goes to in the limit.

WLC differs from DDT in that it has two stages (agglomerating over utility contributions to determine the utility of spatiotemporal regions, and then using these values to assess acts) instead of one (using the utility contributions to directly assess acts), and does not employ zero-point pairing or acceptable pairings. And like WPC, WLC leaves the order of some of the relevant quantifiers ambiguous in a way that might yield further deviations from DDT.

B.3 Assessing the Weak Person Condition

In sections 6 and 7 we looked at objections to DDT and Arntzenius's proposal. Now let's look at how WPC compares to these views.

WPC is generally subject to the same worries as DDT regarding coarse-grained prescriptions (cf. section 6).⁵¹ Like DDT, it doesn't employ the the spatiotemporal structure of outcomes to effectively provide a privileged ordering or grouping of terms. So it will yield coarse-grained prescriptions in cases like Favoring the Positive in much the same way as DDT.

WPC is also subject to many of the worries raised for Arntzenius's proposal in section 7, such as the worries raised in section 7.1, and worries similar to those raised in sections 7.4, 7.5, and 7.6.⁵² Finally, WPC will face worries arising from the need to resolve issues like a subject's identity over time in order to make sense of the first stage value assignments the view employs.⁵³

That said, WPC does have an arguable advantage over DDT in certain cases in which agglomerating utility by subject seems natural. Let's suppose hedonic utilitarianism is true, so that the utility contributions are provided by experiences of pleasure or pain. Furthermore, let's assume that given the laws that obtain at the outcomes we'll be concerned with, the time required to experience an episode of pleasure or pain is one minute. Now consider the following case:⁵⁴

Intensity or Duration. There are countably infinitely many babies, each who will live for three minutes. You have two options. The first (Candy Dump) is to give each baby three pieces of candy as soon as they come into existence, in which case each will have a utility of 2 during their first minute, and a utility of 0 in the next two minutes. The second (Candy Rationing) is to give each baby one piece of candy each minute,

⁵¹Though as we'll see below, there are at least some cases of this kind in which WPC arguably yields more plausible verdicts than DDT.

⁵²Let me go through these three worries in reverse order. First, worries similar to those discussed in section 7.6 arise for WPC in cases in which infinitely many utility contributions are associated with a temporally extended individuals. Second, as with Arntzenius's proposal, worries will arise in cases in which some outcomes are spatiotemporal and some are not (cf. section 7.5). WPC works by (1) summing up utility contributions to assigns values to the complete temporal lives of people, and then (2) using those values to assess acts. But in cases in which the outcomes don't have temporal structure there won't be any temporally extended individuals, so WPC will fall silent. Third, as with Arntzenius's proposal, worries will arise in cases in which different outcomes have different temporal structure (cf. section 7.4). In particular, worries will arise here that intersect worries WPC faces regarding how to work out details regarding personal identity. If we adopt a psychological continuity-over-time or physical continuity-over-time account of personal identity, then questions will arise regarding how to pick out temporally extended individuals at temporally discontinuous spacetimes, or spacetimes without a clear spatial/temporal distinction, or spacetimes with multiple temporal dimensions. And even with an answer to these questions, further questions will arise regarding which temporally extended individuals should be identified with each other in cases in which different outcomes have different temporal structures.

⁵³For example, if we adopt a psychological continuity-over-time or physical continuity-over-time account of personal identity, then tricky questions arise regarding how to pick out temporally extended individuals at temporally discontinuous spacetimes, or spacetimes without a clear spatial/temporal distinction, or spacetimes with multiple temporal dimensions.

⁵⁴I thank an anonymous referee for suggesting a version of this case.

in which case each will have a utility of 1 for each minute.

	Baby 1 ($\tau_1/\tau_2/\tau_3$)	Baby 2 ($\tau_1/\tau_2/\tau_3$)	...	Sum
Candy Dump	2/0/0	2/0/0	...	∞
Candy Rationing	1/1/1	1/1/1	...	∞
Difference (per τ)	1/-1/-1	1/-1/-1	...	undefined
Difference (per lifetime)	-1	-1	...	$-\infty$

According to WPC, Candy Dump is impermissible and Candy Rationing is obligatory. For the utility assigned to each subject's life will be 2 given Candy Dump and 3 given Candy Rationing, and the sum of the difference between Candy Dump and Candy Rationing will be $-\infty$.

But according to DDT both Candy Dump and Candy Rationing will be permissible. For the sum over differences in utility contributions will yield infinitely many positive contributions and infinitely many negative contributions. And since different permutations of these terms will yield different verdicts – summing over two positive terms per negative term will yield ∞ , while summing over two negative terms per positive term will yield $-\infty$ – DDT will take both options to be permissible.

(What verdict will Arntzenius's proposal yield? It depends on how the babies are spatiotemporally arranged (and this is true even if we assume the babies occupy the same spatiotemporal locations in both outcomes). For example, suppose that all outcomes are Newtonian, that the babies exist at the same time, and that they are equally spaced along some spatial axis. Then Arntzenius's proposal will yield the verdict that Candy Dump is impermissible and Candy Rationing is obligatory. Suppose instead that the babies aren't equally spaced along the spatial axis, but become exponentially more densely packed (spatially) as we move away from Baby 1, and suppose they don't exist at the same time, but are born later and later as we progress down the axis (at a rate linear with distance). Then there will be some acceptable expanding sequences (e.g., intuitively, ones which expand slowly in the temporal direction and quickly in the spatial direction) which favor Candy Dump (since each expansion will let in more positive utility contributions than negative ones), and some acceptable expanding sequences (e.g., intuitively, ones which expand quickly in the temporal direction and slowly in the spatial direction) which favor Candy Rationing (since each expansion will let in more negative utility contributions than positive ones). So Arntzenius's proposal will yield the verdict that both options are permissible.)

As an objection to DDT, this worry is structurally identical to the Favoring the Positive case discussed in section 6. In both cases, we're presented with cases where differences in utility contributions give us a sum over infinitely many plusses and minuses of non-trivial magnitudes, and since such sums aren't permutation invariant, DDT will take both options to be permissible. And as with Favoring the Positive, how compelling one takes this case to be will depend on how one feels about these kinds of aggregative intuitions. One might argue that the intuition that Candy Rationing is better than Candy Dump seems to hang on the thought that the size of the set of positive τ s is greater (in the relevant sense) than the size of the set of negative τ s. And one might question whether we should rely on such intuitions regarding the sizes of these sets in such cases.⁵⁵

⁵⁵One might be tempted to defend DDT by arguing that those who think DDT yields the wrong verdicts

All said and done, I take WPC to be roughly on a par with Arntzenius’s proposal. For while it can’t provide fine-grained prescriptions in cases like Favoring the Positive, it always (instead of merely sometimes) takes Candy Dump to be impermissible in Intensity or Duration, and it avoids a couple of the worries facing Arntzenius’s account.

Like Arntzenius’s proposal, I take WPC to be less attractive than DDT. While it can arguably yield more plausible verdicts in cases like Candy Rationing, it yields the same verdicts as DDT in cases like Favoring the Positive. And it encounters many of the same problems Arntzenius’s proposal runs into, in addition to worries regarding a subject’s identity over time.

B.4 Assessing the Weak Location Condition

Now let’s look at how WLC compares to these views. Like WPC, WLC is generally subject to the same worries as DDT regarding coarse-grained prescriptions that (cf. section 6 and appendix B.3). Like DDT, it doesn’t employ the the spatiotemporal structure of outcomes (in the case of Favoring the Positive) or the temporal structure of subjects (in the case of Intensity or Duration) to effectively provide a privileged ordering or grouping of terms.⁵⁶ So it will yield coarse-grained prescriptions in cases like Favoring the Positive and Intensity or Duration in much the same way as DDT. WLC is also subject to almost all of the worries raised for Arntzenius’s proposal in section 7, such as the worries raised in sections 7.1, 7.3, 7.4, 7.5, 7.6, and 7.7.⁵⁷

All said and done, I take WLC to be strictly less attractive than the other views we’ve considered. If we take the inability to provide fine-grained prescriptions in cases like Favoring the Positive to be a demerit of a view, then WLC is essentially dominated by Arntzenius’s proposal. For WLC faces virtually all of the same problems as Arntzenius’s proposal, but without the benefit or being able to provide fine-grained prescriptions in cases like Favoring the Positive. On the other hand, if we think that a view shouldn’t be providing fine-grained prescriptions in cases like Favoring the Positive, then WLC is dominated by DDT. For it’s not clear that WLC has any virtues relative to DDT, and it has a lot of vices: it faces virtually all of the same problems as Arntzenius’s proposal.

in Intensity or Duration should find fault with the notion of utility contributions we’re employing, not DDT. After all, if we took utility contributions to be assigned to something like the complete lives of subjects, DDT would not yield those verdicts. I think we should resist this temptation. The goal of proposals like Arntzenius’s proposal and DDT is to provide generalizations of theories like hedonic utilitarianism that yield plausible results in cases with infinitely many utility contributions. And trying to avoid implausible consequences of such proposals by changing the theory we’re trying to generalize doesn’t help the proposal achieve this goal, it just changes the topic (cf. footnote 32).

⁵⁶Since WLC joins Arntzenius’s proposal in taking spatiotemporal regions to be the bearers of utility in the second stage of assessment, it might be surprising that WLC doesn’t do any better than DDT or WPC in cases like Favoring the Positive. But unlike Arntzenius’s proposal, WLC doesn’t work by looking at increasing uniform expansions, and seeing what happens in the limit. Instead, WLC just partitions spacetime and sums over the utility contributions in those regions (in some arbitrary order), depriving it of much of the spatiotemporal information that Arntzenius’s proposal makes use of.

⁵⁷The immediate worries described in section 7.4 and the “no spacetime structure” part of section 7.5 won’t apply to the version of WLC described in appendix B.2, because it effectively builds in the first amendment suggested in response to each of these worries. But it is still subject to the worries raised for these amendments.

References

- Albert, D. Z. (2015). *After Physics*. Harvard University Press.
- Allori, V. (2016). Space, time, and (how they) matter: A discussion about some metaphysical insights provided by our best fundamental physical theories. In G. C. Ghirardi, & S. Wuppuluri (Eds.) *Space, Time, and The Limits of Human Understanding*, (pp. 95–107). Springer.
- Arntzenius, F. (2014). Utilitarianism, decision theory and eternity. *Philosophical Perspectives*, 28(1), 31–58.
- Arrhenius, G., Ryberg, J., & Tnnsj, T. (2017). The repugnant conclusion. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2017 ed.
- Baron, S., & Miller, K. (2014). Causation in a timeless world. *Synthese*, 191(12), 2867–2886.
- Baron, S., & Miller, K. (2015). Causation sans time. *American Philosophical Quarterly*, 52(1), 27–40.
- Bartha, P. (2007). Taking stock of infinite value: Pascal’s wager and relative utilities. *Synthese*, 154(1), 5–52.
- Block, N. (1978). Troubles with functionalism. *Minnesota Studies in the Philosophy of Science*, 9, 261–325.
- Bostrom, N. (2011). Infinite ethics. *Analysis and Metaphysics*, 10, 9–59.
- Bricker, P. (1996). Isolation and unification: The realist analysis of possible worlds. *Philosophical Studies*, 84(2-3), 225–238.
- Bricker, P. (2001). Island universes and the analysis of modality. In G. Preyer, & F. Siebelt (Eds.) *Reality and Humean Supervenience: Essays on the Philosophy of David Lewis*. Rowman and Littlefield.
- Buchak, L. (2013). *Risk and Rationality*. Oxford University Press.
- Cain, J. (1995). Infinite utility. *Australasian Journal of Philosophy*, 73(3), 401–404.
- Carroll, S. M. (forthcoming). Why boltzmann brains are bad. In S. Dasgupta, & B. Weslake (Eds.) *Current Controversies in the Philosophy of Science*. Routledge.
- Chen, E. K., & Rubio, D. (forthcoming). Surreal decisions. *Philosophy and Phenomenological Research*.
- Chichilnisky, G. (1996). An axiomatic approach to sustainable development. *Social Choice and Welfare*, 13, 321–257.
- Chichilnisky, G. (1997). What is sustainable development? *Land Economics*, 73, 467–491.

- Colyvan, M. (2008). Relative expectation theory. *Journal of Philosophy*, 105(1), 37–44.
- Colyvan, M., & Hájek, A. (2016). Making ado without expectations. *Mind*, 125(499), 829–857.
- Divers, J. (2002). *Possible Worlds*. Routledge.
- Easwaran, K. (2014). Principal values and weak expectations. *Mind*, 123(490), 517–531.
- Feldman, F. (2004). *Pleasure and the Good Life: Concerning the Nature, Varieties, and Plausibility of Hedonism*. Clarendon Press.
- Fishkind, D., Hamkins, J. D., & Montero, B. (2002). New inconsistencies in infinite utilitarianism: Is every world good, bad or neutral? *Australasian Journal of Philosophy*, 80(2), 178–190.
- Fletcher, G. (2013). A fresh start for the objective-list theory of well-being. 25, 206–220.
- Garcia, J. L. A., & Nelson, M. T. (1994). The problem of endless joy: Is infinite utility too much for utilitarianism? *Utilitas*, 6(2), 183.
- Gwiazda, J. (2014). Orderly expectations. *Mind*, 123(490), 503–516.
- Hájek, A. (2003a). Waging war on pascal’s wager. *Philosophical Review*, 112(1), 27–56.
- Hájek, A. (2003b). What conditional probability could not be. *Synthese*, 137(3), 273–323.
- Hamkins, J. D., & Montero, B. (2000). With infinite utility, more needn’t be better. *Australasian Journal of Philosophy*, 78(2), 231–240.
- Horwich, P. (1985). Decision theory in light of newcomb’s problem. *Philosophy of Science*, 52(3), 431–450.
- Huggett, N., & Wüthrich, C. (2013). Emergent spacetime and empirical coherence. *Studies in History and Philosophy of Modern Physics*, 44(3), 276–285.
- Joyce, J. M. (1999). *The Foundations of Causal Decision Theory*. Cambridge University Press.
- Koopmans, T. C. (1960). Stationary ordinary utility and impatience. *Econometrica*, 28(2), 287–309.
- Lauwers, L. (2014). The axiomatic approach to the ranking of infinite streams. <http://dx.doi.org/10.2139/ssrn.2409028>.
- Lauwers, L., & Vallentyne, P. (2004). Infinite utilitarianism: More is always better. *Economics and Philosophy*, 20(2), 307–330.
- Lauwers, L., & Vallentyne, P. (forthcoming). Decision theory without finite standard expected value. *Economics and Philosophy*, (pp. 1–25).

- Lewis, D. K. (1968). Counterpart theory and quantified modal logic. *Journal of Philosophy*, 65(5), 113–126.
- Liedekerke, L. V. (1995). Should utilitarians be cautious about an infinite future? *Australasian Journal of Philosophy*, 73(3), 405–407.
- Meacham, C. J. G. (forthcoming). Difference minimizing theory. *Ergo*.
- Melia, J. (2003). *Modality*. McGill-Queen's University Press.
- Nelson, M. T. (1991). Utilitarian eschatology. *American Philosophical Quarterly*, 28(4), 339–47.
- Nover, H., & Hájek, A. (2004). Vexing expectations. *Mind*, 113(450), 237–249.
- Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.
- Ramsey, F. P. (1928). A mathematical theory of saving. *The Economic Journal*, 38(152), 543–559.
- Seegerberg, K. (1976). A neglected family of aggregation problems in ethics. *Noûs*, 10(2), 221–244.
- Vallentyne, P. (1993). Utilitarianism and infinite utility. *Australasian Journal of Philosophy*, 71(2), 212–217.
- Vallentyne, P., & Kagan, S. (1997). Infinite value and finitely additive value theory. *Journal of Philosophy*, 94(1), 5–26.
- vonWeizsacker, C. C. (1965). Existence of optimal programs of accumulation for an infinite time horizon. *The Review of Economic Studies*, 32(2), 85–104.