Suppose there is a medical device that enables doctors to apply electric current to the body in increments so tiny that the patient cannot feel them. The device has 1001 settings: 0 (off) and 1 ... 1000.1 Suppose someone (call him the self-torturer) agrees to have the device, in some conveniently portable form, attached to him in return for the following conditions: The device is initially set at 0. At the start of each week he is allowed a period of free experimentation in which he may try out and compare different settings, after which the dial is returned to its previous position. At any other time, he has only two options — to stay put or to advance the dial one setting. But he may advance only one step each week, and he may never retreat. At each advance he gets $10,000.

Since the self-torturer cannot feel any difference in comfort between adjacent settings, he appears to have a clear and repeatable reason to increase the voltage each week. The trouble is that there are noticeable differences in comfort between settings that are sufficiently far apart. Indeed, if he keeps advancing, he can see that he will eventually reach settings that will be so painful that he would then gladly relinquish his fortune and return to 0.2

The self-torturer is not alone in his predicament. Most of us are like him in one way or another. We like to eat but also care about our appearance. Just one more bite will give us pleasure and won't make us look fatter; but very many bites will. And there may be similar connections between puffs of pleasant smoking and lung cancer, or between pleasurable moments of idleness and wasted lives.

In all these cases, we find a mix of transitive and intransitive preferences. The self-torturer's step-wise preferences are intransitive. All things considered, he prefers 1 to 0, 2 to 1, 3 to 2, etc. ... but certainly not 1000 to 1. This is why he cannot say that any setting is

better than the previous one. “Better than” is, while his step-wise preferences are not, transitive. But when he compares settings so far apart that he prefers the earlier setting, his preferences are transitive. If he prefers 500 to 1000 and 0 to 500, then he prefers 0 to 1000. This seems to permit us to say that 1000 is worse for him than 0.

The self-torturer’s preferences are considered and well-informed. Before forming them, he freely experiments with all the relevant settings. And he is well-informed about the pleasures and advantages that different amounts of money can buy. Despite this, many theorists would condemn his intransitive preferences as irrational. But this response may be too hard on the self-torturer and too easy on the theorist. The self-torturer’s intransitive preferences seem perfectly natural and appropriate given his circumstances. They are the very ones most of us would have in his place (and do have in structurally similar, everyday situations). To insist that he get new, more ‘rational’ preferences might well invite bad faith. He wants to know how he should act on the ones he actually has. Intuitively, this question does not seem to be one that must lack a satisfactory answer. So the theoretical convenience of rejecting the question must be weighed against the failure to address what seems to be a genuine problem. In any case, I shall assume provisionally that the self-torturer has, as he is, a real problem of rational choice: How to take reasonable advantage of what the device offers him without ending up the worse for it.

Let’s begin by considering some objections to my description of the case — objections that would point the way to some familiar way of solving the puzzle or some further grounds for rejecting it.

(1) The self-torturer’s preferences are changing: On this objection, he is like an addict. At 0, the thought of 1000 appalls him, but at 999 it looks good. The changes wrought in him between 0 and 999 affect his outlook. Since he now (at 0) prefers not to change, he should resist any advance that has that effect.

But his preferences do not change. Even at 0 he prefers 1000 to 999. The same holds for plans of sequential choice. He always prefers a