The Cable Guy Paradox

Alan Hájek

Analysis, Vol 65. No. 4, April 2005

The Cable Guy is coming. You have to be home in order for him to install your new cable service, but to your chagrin he cannot tell you exactly when he will come. He will definitely come between 8.a.m. and 4 p.m. tomorrow, but you have no more information than that. I offer to keep you company while you wait. To make things more interesting, we decide now to bet on the Cable Guy's arrival time. We subdivide the relevant part of the day into two 4-hour long intervals, 'morning': (8, 12], and 'afternoon': (12, 4). You nominate an interval on which you will bet. If he arrives during your interval, you win and I will pay you $10; otherwise, I win and you will pay me $10. Notice that we stipulate that if he arrives *exactly* on the stroke of noon, then (8, 12] is the winning interval, since it is closed on the right; but we agree that this event has probability 0 (we have a *very* precise clock!).

At first you think: obviously there is no reason to favour one interval over the other. Your probability distribution of his arrival time is uniform over the 8 a.m. – 4 p.m. period, and thus assigns probability 1/2 to each of the two 4-hour periods at issue. Whichever period you nominate, then, your expected utility is the same. The two choices are equally rational.

But then you reason as follows. Suppose that you choose the morning interval. Then there will certainly be a period during which you will regard the *other* interval as

preferable. For even if the Cable Guy arrives in the morning interval (the best case scenario), *some* portion of that interval will have elapsed *before* his arrival. During that portion, you would be rationally required to swap your bet to the afternoon interval if given the opportunity (although there will be no such opportunity since the bets are irreversible). For example, if he arrives at 10 a.m., then for two hours your interval will have been gradually dwindling, while the afternoon interval remains fully intact. Indeed, an instant before 10 a.m., only two hours (plus that instant) of your interval will remain, and given your uniform distribution for his arrival time, your probability at that time of winning will be (approximately) only half your probability for the afternoon interval winning. Similar reasoning applies however early he might arrive in the morning: at 8:10, or at 8:01, or at a second past 8:00, or …—at least *some* of your interval will have been used up before his arrival, and during that time you will be rationally required to regard the afternoon interval, with its full complement of four hours, as more desirable. For each possible arrival time in the morning, there is an earlier time at which you must prefer the afternoon interval.

Now suppose that you choose the afternoon interval. Then it is *not* the case that there will certainly be a period during which you will regard the other interval as preferable. On the contrary, if at 12:00 the Cable Guy still has not shown up (an event you now assign probability 1/2), then far from wishing that you could exchange your interval for the other one, you will know at that point that you will win. Of course, you do not know *now* that he will arrive during afternoon. You realize that if you nominate this interval, you may be wrong—he may arrive in the morning, and in that case you will (retrospectively) regard that interval as preferable. However, *this* sort of preference-

switching is not guaranteed. And the prospect of *this* sort of preference-switching is the same whichever interval you pick: the risk of being wrong is the same. But nominating the morning interval commits you to a *guaranteed* and *asymmetric* preference-switching caused by watching at least some of your interval erode away before his arrival; nominating the afternoon interval protects you from such preference-switching. So you should nominate the afternoon interval. Where initially you regarded the two choices symmetrically, you have now found a symmetry-breaker .

Note that we have not added any extra details to the original story. We have not, for example, added new considerations about the disutility of the pain of regretting one's choice, forcing a recalculation of the original expected utilities. All the reasoning assumes is that money is the only source of utility in this game, and that you are an expected utility maximizer. You are certain now that a rational future self of yours will assign higher probability to the Cable Guy's arriving in the afternoon than in the morning. Thus, you are certain now that a rational future self of yours will assign higher expected utility to betting on his arriving in the afternoon than in the morning. So you are certain now that a rational future self of yours will prefer betting on the afternoon interval. Why wait, then? It seems rationally required for you to form that preference *now*.

Prompted by this talk of rational future selves, we may put the reasoning in terms of a plausible diachronic rationality principle somewhat in the spirit of van Fraassen's 'Reflection Principle' (1984 and 1995). The idea is that you should not knowingly frustrate a rational future self of yours. I will call it the '*Avoid Certain Frustration* Principle':

> *Suppose you now have a choice between two options. You should not choose one of these options if you are certain that a rational future self of yours will prefer that you had chosen the other one—unless both your options have this property.*

I have carefully worded the principle so as to evade four types of potential counterexample—the first ruled out by the word "certain", the second by "rational", the third by "two options", and the fourth by "unless both your options have this property":

1) Uncertain frustration: Of course, we frequently frustrate our future selves while acting fully rationally, because we are *uncertain* of how events will transpire. Suppose that a fair coin is about to be tossed, and you now have a choice between betting on heads and betting on tails. Quite rationally, you choose to bet on heads, fully aware that the coin may land tails, and that a rational future self of yours—after the coin has landed—will in that case prefer that you had chosen to bet on tails. But this is no counterexample to the *Avoid Certain Frustration* Principle, because the prospect of this frustration is *uncertain*.

2) Certain frustration when there are at least three options: Sometimes it is rational to knowingly frustrate a future self when one is choosing among *three or more* options, and when this does not reduce to a pair-wise choice in which there is such foreseen frustration. This can happen when there is foreseen frustration of some form or other, but it is not foreseen exactly *which* form it will take. (When there are only two options, the frustration over having chosen one option can only take one form: a preference for the other option.) Now suppose, for example, that we add a third choice to the coin tossing case: you may refuse to bet. This is surely a rational choice (even if it is not rationally

required, since the betting options have just as high expected utility). Yet you are certain that a rational future self of yours—after the coin has landed—will prefer that you had chosen to bet on the winning result.[1] But notice that this does not reduce to a two-option choice to which the *Avoid Certain Frustration* Principle would apply, for you are not certain now *which* of the two betting choices will eventually be regarded as preferable to not betting. You are not certain now, for example, that a rational future self of yours will prefer that you had chosen to bet *on heads*. So the Principle does not deliver an incorrect verdict here, because it does not deliver a verdict at all.

3) Rational protection against an irrational future self: In certain special cases, a rational agent may foresee a future lapse in rationality, and may act now knowingly so as to frustrate an *irrational* future self. In the *Odyssey*, as Ulysses approaches the sirens he has two options: tying himself to the mast so that he cannot steer his ship fatally to the cliff where they are irresistibly singing, or not doing so. He rationally chooses to tie himself to the mast, despite being certain that a future self of his (namely, one under the spell of the sirens) will prefer that he had chosen not to do so. He is thus certain to frustrate a future self of his. But Ulysses knows that the future self here is *not* rational, so this is no counterexample to the *Avoid Certain Frustration* Principle.

4) Frustration whatever you choose: The *Avoid Certain Frustration* Principle ends with a caveat: "… unless both your options have this property." There may be pathological cases of choices between two options that activate this caveat: whichever option you

[1] I thank Marc Lange for a closely related example.

choose now, foreseeably a rational future self of yours will prefer that you had chosen the other one. Perhaps the infinite version of the 'two envelope paradox' in which you get to open your envelope is like this: arguably, you know now that whatever you see, you will prefer having the other envelope. (See, e.g., Arntzenius and McCarthy 1997 and Chalmers 2002.) But we need not settle such cases, because the *Avoid Certain Frustration* Principle is suitably weakened by the caveat so as not to address them.

None of these four kinds of excusing conditions for foreseeable future frustration applies in the Cable Guy scenario. The choice to bet on the morning interval falls squarely under the purview of the *Avoid Certain Frustration* Principle. It is thus ruled out. Rationality, then, requires you to bet on the afternoon interval (the only choice that is not so ruled out). This is paradoxical, because your initial reasoning that there is nothing to favour one interval over the other seemed impeccable.

We can equally arrive at the paradox by focusing on the *evidence* that you are certain of receiving. At the moment, your evidence bears symmetrically on both the intervals; but you are certain of getting evidence that will bear asymmetrically on them. The Principle of Indifference seems to apply in your current epistemic state—hence your initial 1/2-1/2 probability distribution over the two intervals. But you are certain that you will soon be in an epistemic state in which the Principle will *not* apply—in which your probability distribution will favour the afternoon interval. All this without any epistemic impairment at any stage—no sirens calling you to choose that interval, or what have you. Maximizing your probability of winning, given your evidence, is all the call you need.

So you are certain now that if you choose the morning interval, you will receive evidence that will *undermine* that choice: having received the evidence, you will

rationally disprefer having made that choice. Call a choice *self-undermining* if, having made it, you are certain that you will receive undermining evidence regarding it. Betting on the morning interval is a self-undermining choice. Not so betting on the afternoon interval: having done so, you are *not* certain that you will receive undermining evidence regarding it. Now let us appeal to an equally plausible[2] and more succinct diachronic rationality principle, which I will call the '*Avoid Self-Undermining Choices* Principle':

> *Whenever you have a choice between two options, you should not make a self-undermining choice if you can avoid doing so.*

You can avoid making a self-undermining choice by betting on the afternoon interval. And so, apparently, you should.

This is so even though the afternoon interval contains neither of its endpoints, so in that sense it appears to be 'shorter'—by a point!—than the morning interval. Remember that the endpoints each have probability zero. Curiously, the topology of the set-up—that the intervals are open on the left—is crucial. Suppose, for example, that the first interval was instead [8, 12], so that the Cable Guy could arrive at 8:00 on the dot. If he did so, then none of the interval would have elapsed before his arrival. Then there would be no period of preference-switching, if that were your chosen interval. The inclusion or exclusion of a probability zero event makes all the difference to your reasoning: the diachronic rationality principles apply in one case, and not in the other. To be sure, we could strengthen those principles so that they would apply either way: by replacing their explicit or implicit reference to what you regard as *certain* with a cagier reference to what you assign *probability 1*. But thanks to the topology of our case, we need not do so.

---

[2] In fact, I believe that it is equivalent to the *Avoid Certain Frustration* Principle, although proving this equivalence is beyond the scope of this paper.

Another curiosity is that despite your apparent preference for the afternoon interval, you should now choose the morning interval given the tiniest incentive. If I paid you a penny up front to choose the morning interval, you should: then, the expected utility of doing so would be a penny better than that of choosing the afternoon interval. If the incentive were a lottery ticket that had any positive probability, however small, of paying a penny, you should still choose the morning interval. You apparently prefer the afternoon interval to the morning interval, but there is no finite amount by which you prefer it.

There is something of the snake eating its own tail in all of this. For the initial reasoning that concluded that the two intervals are exactly on a par was based on a consideration of expected utilities. But this is undercut by the second round of reasoning, whichever way we run it, and it is also essentially a consideration of expected utilities. Expected utility reasoning seems to conflict with itself, and is thus called into question.

This seems absurd. I submit that it is the *Avoid Certain Frustration* and the *Avoid Self-Undermining Choices* Principles, innocent though they appeared, that have to be called into question. Sometimes it is rational to knowingly act against your rationally-formed future preferences, even when you know exactly how to avoid doing so.[3]

[3] I am grateful to Paul Bartha, Adam Elga, Andy Egan, Jim Joyce, Brad Monton, Harris Nover, Neil Thomason, Bas van Fraassen, Manuel Vargas, and especially to James Chase (who suggested a close relative of the first argument for preferring the afternoon interval), Chris Hitchcock, Marc Lange, and Peter Vranas for very useful comments.

REFERENCES

Arntzenius, Frank and David McCarthy. 1997. 'The Two Envelope Paradox and Infinite expectations.' *Analysis*, Vol. 57, No. 1: 42-50.

Chalmers, David. 2002. 'The St. Petersburg Two-Envelope Paradox', *Analysis* 62:155-57,

van Fraassen, Bas. 1984. 'Belief and the Will', *Journal of Philosophy* 81: 235-256

van Fraassen, Bas. 1995, 'Belief and the Problem of Ulysses and the Sirens', *Philosophical Studies* 77: 7-37

MATTHIAS: THERE IS A TIME WHEN YOU'LL WANT TO SELL THE BET – EVENTHOUGH THE DUTCH BOOKIE CAN'T TELL YOU INADVANCE. INFINITELY MANY POISED TO STRIKE? – BUT THEN THEY KNOW MORE THAN YOU!