Philos Stud (2008) 138:317–333
DOI 10.1007/s11098-006-9044-1

ORIGINAL PAPER

Avoiding certain frustration, reflection, and the cable guy paradox

Brian Kierland · Bradley Monton · Samuel Ruhmkorff

Received: 13 July 2005 / Accepted: 17 August 2006 / Published online: 30 September 2006

Abstract We discuss the cable guy paradox, both as an object of interest in its own right and as something which can be used to illuminate certain issues in the theories of rational choice and belief. We argue that a crucial principle—The Avoid Certain Frustration (ACF) principle—which is used in stating the paradox is false, thus resolving the paradox. We also explain how the paradox gives us new insight into issues related to the Reflection principle. Our general thesis is that principles that base your current opinions on your current opinions about your future opinions need not make reference to the particular *times* in the future at which you believe you will have those opinions, but they do need to make reference to the particular *degrees of belief* you believe you will have in the future.

Keywords Bayesianism · Reflection principle · Stopping times

1 The cable guy paradox

Consider the cable guy paradox:

The cable guy will definitely be coming to your house some time in the open interval (8 am, 4 pm). You have no further information about when he will come, so

B. Kierland
Department of Philosophy, University of Missouri-Columbia, 38 General Classroom Bldg., Columbia, MO 65211, USA
e-mail: kierland@missouri.edu

B. Monton (✉)
Department of Philosophy, University of Colorado at Boulder, 169 Hellems, Boulder, CO 80309, USA
e-mail: monton@colorado.edu

S. Ruhmkorff
Division of Social Studies, Simon's Rock College of Bard, 84 Alford Rd., Great Barrington, MA 01230, USA
e-mail: sruhmkor@simons-rock.edu

you are indifferent between the hypothesis that he will come in the interval (8 am, noon) and the hypothesis that he will come in the interval (noon, 4 pm). But then you realize that if you make an even odds bet on the first interval, you will certainly feel regret, since there will surely be a time in the first interval where some time has elapsed and the cable guy hasn't arrived, and so you will wish you had bet on the second interval. So you decide you should bet on the second interval, following the *Avoid Certain Frustration* (ACF) 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 of your options have this property.

The ACF principle dictates that you should favor betting on the second interval, even though you are probabilistically indifferent between the two intervals. This is paradoxical—expected utility reasoning clearly dictates that it's no better to bet on the first interval than the second, and yet the ACF principle dictates otherwise.

Alan Hájek (2005) proposed this paradox, and he also proposed a resolution: reject the ACF principle. We agree with Hájek that the principle should be rejected, but we think that Hájek hasn't said all there is to be said about the paradox. In this paper, we will give a more detailed account of why the ACF principle is false. Furthermore, we will propose and defend a modified version of the ACF principle which avoids paradox. Also, we will show how all of this sheds new light on the Reflection principle. The underlying theme of this paper is as follows. Principles that base your current opinions on your current opinions about your future opinions need not make reference to the particular *times* in the future at which you believe you will have those opinions, but they do need to make reference to the particular *degrees of belief* you believe you will hold in the future.

We'll start, though, with a preliminary point. The cable guy paradox makes a sophisticated assumption about your ability to feel regret, an assumption which is not warranted for actual agents. Specifically, the paradox assumes that, when you bet on the first interval, you can feel regret, even though there is no first time at which you start to feel regret. To see this, suppose that there is a first time at which you come to feel regret. This first time couldn't be before 8 am, or precisely at 8 am, because the regret the paradox has in mind is regret that a certain interval of time where the cable guy could have come has already passed. Thus, the first time at which you come to feel regret would have to be some time t which is after 8 am. But this means that there is an interval of time (8 am, t) in which the cable guy could have come, and you would never have felt regret. So for you to be certain that you will feel regret if you bet on the first interval, you have to be capable of feeling regret even though there is no first time at which you start to feel regret. Specifically, if you bet on the first interval, you would have to feel regret at *every* time in the interval (8 am, T), where T is some time after 8 am, but not beforehand. (T could be the time that the cable guy comes, but it need not; the ACF principle will still apply, even if for some reason you stop feeling regret before the cable guy comes.)

Actual agents are not capable of feeling regret in this way, so the cable guy paradox doesn't show that the ACF principle will dictate problematic judgments for actual agents. But it seems reasonable to take "you" in the ACF principle to range over not just actual agents, but all possible agents. It follows that limitations of actual agents aren't relevant to determining the truth value of the ACF principle.

2 Nearly certain frustration

Hájek points out that the requirement that you be "certain" that a rational future self of yours will prefer the other option is an important one. Hájek explains that we sometimes act in a rational way that ends up frustrating our future self, simply because we are uncertain of how the future will go. For example, if you bet Heads on a future fair coin toss, and the toss comes out Tails, a rational future self of yours will regret your betting on Heads. But beforehand you weren't certain that you would feel regret, since you knew the coin toss could have landed heads.

Hájek does not have anything further to say about the certainty requirement, but we think it's worth emphasizing how crucial it is. Clearly incorrect predictions would result if one weakened the certainty requirement in the ACF principle, by espousing the *Avoid Nearly Certain Frustration* (ANCF) principle:

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

Here the requirement of "near certainty" can be made as stringent as one likes: for any $x < 1$, one can specify that an agent whose opinion is represented by probability function P is nearly certain that A iff $P(A) > x$. Nevertheless, the ANCF principle is clearly false.

To see this, consider a modified version of the cable guy account, where the cable guy will come in the interval (8 am–y, 4 pm), where y is some small but non-zero interval of time. Suppose you are given the option to do an even bet on either the interval (8 am–y, noon] or the interval (noon, 4 pm). Clearly, the first interval should be preferred, regardless of how small y is. But since y is a small interval of time, one is nearly certain to feel regret if one bets on the first interval. As long as the cable guy doesn't come in the interval (8 am–y, 8 am], one will feel regret. It follows that betting on the first interval is nearly certain to lead to regret, even though betting on the first interval is clearly the favored option. Thus, the ANCF principle is false.

Seeing, in this way, that the ANCF principle is false helps to cast suspicion on the ACF principle. Clearly, even being *nearly* certain to feel regret upon choosing an option, you could nonetheless rationally choose that option if your evidence favors it. It's prima facie unclear how the tiniest positive increment of probability that one gets in moving from near certainty to certainty could lead to a difference in the truth values of the two principles.[1]

[1] There are other ways to see that the ANCF principle is false, but we focus on the way in the main text because it is related to the cable guy paradox. One might worry about such a "companions in guilt" argument on the grounds that, e.g., the difference between certainty and near certainty can make all the difference in determining whether it is rational to buy a lottery ticket (so long as the prize is large enough). But in such a situation, there is no special discontinuity at probability 1; instead, there is a threshold at a probability slightly less than 1, determined by the number of tickets, the size of the prize, and the utility one gets out of money. What our example in the main text shows is that, if the ACF principle is true, then, given the falsehood of the ANCF principle, there *is* a special discontinuity at probability 1; note that, in the example, one can make the value of y as small as one likes.

3 Avoiding certain frustration and expecting future gladness

The above discussion may cast suspicion on the ACF principle, but it doesn't explain how it goes wrong. The ACF principle can seem intuitive, because of this apparently uncontroversial phenomenon: present certainty in future rational regret is relevant to present rational choice. We will build up to an explanation of why the ACF principle is false, by considering four points which show this phenomenon to be an instance of a more general phenomenon.

First, just as you can regret having made a choice, you can be glad for having made that choice. For example, suppose that a future fair coin toss will land Heads. Were you now to make an even bet on the toss landing Heads, a rational future self of yours would be glad for your having made this bet. Further, if you now are certain about such future gladness, then it would now clearly be rational for you to make the bet. More generally, just as present certainty in future rational regret is relevant to present rational choice, so also is present certainty in future rational gladness. To be clear, we emphasize this talk of ‘‘gladness'' is not meant to introduce any extra relevant utilities—as it would if (thinking of the coin toss example) it pointed to potential warm, fuzzy feelings as something relevant in addition to potential monetary gains. Rather, it is a way of speaking of how, retrospectively, often in the light of newer information, we find certain options in choice-situations preferable to others. We are retrospectively glad in this way when we prefer the option we in fact chose.[2]

Second, both regret and gladness come in degrees. For example, suppose you make an even bet that at least ten of twenty future fair coin tosses will land Heads. Were the first toss to land Tails, you would rationally regret having made the bet, but only to some minimal amount r_1, since there is still close to a 50% chance of winning the bet. Were the second toss also to land Tails, you would then still rationally regret having made the bet, but now to a higher amount r_2, since your chances of winning the bet would have gone down even further. And so on. Moreover, the amount of regret or gladness one feels depends on the possible outcomes of the bet. For example, the larger the difference in payoffs is between winning and losing the bet, the more regret one would feel after the first toss lands Tails.

How should these amounts of regret and gladness be quantified? Consider a situation where you make a bet, picking between two options. Suppose it turns out that you lose the bet. If you had bet on the other option, you would have gotten some amount of utility y.[3] But in fact you got a lower (perhaps negative) amount of utility x. We specify that the amount of regret you feel is equal to the difference in utilities, $(y–x)$. Suppose instead that it turns out that you win the bet, and you get some amount of utility z. If you had bet on the other option, you would have gotten some lower amount of utility w. We specify that the amount of gladness you feel is equal to the difference in utilities, $(z–w)$. Here we are treating regret as ‘‘negative'' gladness—that is, regret of amount r is the same as gladness of amount $–r$. Of course, you may not at the time in question know whether you've won or lost the bet. In that case, your amount of regret or gladness will be determined in similar fashion by your expected utilities at that

[2] Cf. Hájek (2005, p. 114), who makes the parallel point about his talk of regret.

[3] We are setting aside situations where the outcome of the bet is (causally or counterfactually) dependent on the choice one makes.

time, which may have changed since making the bet as you may have acquired new evidence.

Third, it is not only present certainty in future rational regret and gladness which is relevant to present rational choice—so also are present non-extremal probability assignments to such claims about the future. Just as you can calculate your expected utilities for various outcomes, so you can calculate your expected amounts of regret and gladness for various outcomes. These expected amounts of regret and gladness can guide future action, just as expected utilities do.

Fourth, when it comes to using future regret or gladness to guide present choice, there are many future times that one *can* consider. Return to the series of coin tosses example. You can consider your present probabilities for various amounts of rational regret and gladness immediately after the first toss (but before the second), or immediately after the second toss (but before the third), or, ..., immediately after the twentieth toss. So which of these times *should* one consider? We think it natural to answer: the time after the twentieth toss. This is because it is your future self at *that* time which is (in relevant ways) maximally informed; she is the future self of yours who thus provides the best guide for present choice. Of course, there will usually be many such future times; where the choice concerns a bet which will be settled at some particular point in the future, any point after that time will do.[4] Let's call any such time an *end-of-the-day time*. Given how such a time is defined, your probabilities for various amounts of future rational regret and gladness will be the same, no matter which end-of-the-day time you choose to consider.[5]

With these four details now made explicit, we want to present the *Expected Future Gladness* (EFG) principle. But first, let's define your EFG toward choice of an option O. It is, for some time t, the sum of, for each amount of gladness g: (your probability that you will, at t, rationally feel g units of gladness about having chosen option O) $\times g$.[6] Now, here is the EFG principle:

> Suppose you have a choice between two options. You should not choose one of these options if your EFG at some end-of-the-day time toward choice of this option has some particular negative value.

We endorse the EFG principle, and as we will explain below, this principle can be used to explain why the ACF principle is false.[7] But first, we will make some preliminary points about the EFG principle.

---

[4] That is, any such point will do so long as you are alive then; if not, the time of maximal relevant information can be taken as the last time at which you are alive.

[5] It might be said: what you, in the relevant sense, *should* now consider is that future time such that *you believe*, or *find it probable*, that you will be relevantly maximally informed at that time. Strictly, this is correct. An easy way around the point would just be always to take the last time at which you are alive as your end-of-the-day time. Or at least it would be, if there were no issues about forgetting. We are only interested in rational agents, so we are making the standard Bayesian idealization that an agent never forgets.

[6] If there is a continuum of degrees of gladness, as we think there is, then this definition will strictly have to be modified to make use of integration instead of mere summation. But since the idea should be clear enough, we choose to avoid the complexities involved in such modification.

[7] Although we think it most natural, choosing end-of-the-day times is not necessary; choosing other times (or even aggregating across many different times) leads to principles equivalent to the EFG principle. We leave this for the reader to work out.

Note that the EFG principle is restricted to choice-situations with only two options. This is because the terms 'gladness' and 'regret' are most naturally understood in comparative fashion. For example, you regret having made a choice only if you wish you had made a *different choice*. Restricting ourselves to two-option situations allows us to leave this comparative aspect implicit (since each of two options has only one other to be compared to), and also allows us to avoid complications arising in more-that-two-option situations.[8]

The EFG principle looks a lot like the standard principle of choice in Bayesian theory, namely, the *Expected Utility* principle, which reads as follows when restricted to two-option situations:

> Suppose you have a choice between two options. You should not choose one of these options if choosing it has lower *expected utility* than choice of the other option. Your expected utility for choice of an option O is the sum of, for each degree of utility d: (your probability that you choosing option O will bring you degree of utility d) $\times$ d.[9]

The parallel between the EFG principle and the Expected Utility principle is no accident. The EFG calculation will always yield the same result as the expected utility calculation. When, relative to some specific choice situation, an end-of-the-day time T_e involves knowing (all relevant aspects of) the outcome of having made the choice one way or another, the parallel is fairly direct. This is because, in such a case, your probability now for knowing at T_e that the outcome goes one way upon a certain choice just is your probability for the outcome going that way upon this choice. And probabilities of this latter sort are what determine (together with facts about utilities) your expected utility for this choice. When an end-of-the-day time does not involve knowing (all relevant aspects of) the outcome of having made the choice one way or another, there is a less direct, but equally real, parallel. At such a less-than-fully-informed end-of-the-day time, your gladness for having made this choice will be determined (together with your utilities) by your updated probabilities—your original probabilities (i.e., your probabilities now) updated by whatever relevant information how have learned up to that point. However, these updated probabilities will reflect, not just what you have learned between now and then, but also, in effect, probabilities for learning further relevant information which would take you from where you are then to a (possible imaginary) state of knowing (all the relevant aspects of) the outcome of having made the choice one way or another; this is because your original probabilities can be treated as probabilities for what a relevantly maximally informed individual would know. (And all of this is something

[8] Presumably there is a similar non-comparative principle (one for which there is no reason of simplicity for restricting to two-option situations), but we stick with the language of ''gladness'' and ''regret'' for continuity with Hájek. Note that Hájek (2005, p. 115) himself restricts his ACF principle to two-option situations for a different reason; this reason is inapplicable once we have moved from certain future regret to expected future gladness. Hájek (2005, pp. 115–116) also restricts his ACF principle so that it does not apply when each option has the feature that one is certain to rationally regret having made it in the future. Given the comparative nature of regret, such a situation may seem impossible, but Hájek cites the situation in the two envelope paradox as a possible case. In case such situations are possible (which we doubt), the reader should feel free to read a similar restriction into the EFG principle; but for simplicity, our formulations will leave out such a restriction.

[9] Since expected utility is not comparative in the way that expected future gladness is (see second point above after formulation of the EFG principle), formulation of this version of the Expected Utility principle requires explicitly building in the comparative aspect.

you now know.) So your original probabilities, used in expected utility calculations, are in effect put to the same work in EFG calculations for less-than-fully-informed end-of-the-day times.

What the parallel between the EFG principle and the Expected Utility principle reveals is that, once all the details are made explicit, considerations of future gladness and regret as a guide to present choice do not really offer anything beyond standard expected utility calculations—they simply give us a new way of getting the same result we would get from expected utility calculations. Although in one way disappointing, this revelation does help in treating the cable guy paradox. The paradox consisted in the seeming fact that the Expected Utility principle offered guidance different from that offered by considerations of future gladness and regret. And in fact, there is a difference in guidance, when considerations of future gladness and regret are treated using the ACF principle. But there is no difference in guidance once considerations of future gladness and regret are treated in the correct way, that is, once we have moved from the ACF to the EFG principle. Both the Expected Utility principle and the EFG principle yield the correct result, that you should be indifferent between betting that the cable guy will come in the first interval and betting that he will come in the second.

This might seem puzzling, since it might seem that the ACF principle is a corollary of the EFG principle for those cases where certainty is to be had. This appearance is illusory. The corollary of the EFG principle in the relevant neighborhood is the *Certain Future Gladness* (CFG) principle:

> Suppose you have a choice between two options. You should not choose one of these options if, for some end-of-the-day time, you are *certain* that, at that time, the amount of gladness associated with choice of this option has some particular negative value.

In the cable guy example, it is false that, for some end-of-the-day time (which, in the example, is 4 pm or any time after), you are certain you will feel negative gladness, i.e., regret, at that time. This is because, now attaching equal probabilities to the cable guy coming in the interval (8 am, noon] and his coming in the interval (noon, 4 pm), you now attach equal probabilities to having learned by any end-of-the-day time that he came in (8 am, noon] and to having similarly learned that he came in (noon, 4 pm). Hence, the CFG principle is simply inapplicable to the cable guy example.

It might be objected: we get this result only because the CFG principle focuses upon end-of-the-day times. If, in the cable guy example, we focus on earlier times, maybe the paradox will return; after all, the whole point of the special set-up of the example is so that there is guaranteed to be *some* time before noon when you will regret having bet on the interval (8 am, noon]. We reply: there is no *particular time* at which one will be certain that one feels regret, and hence, no corollary of the EFG principle that substitutes some other particular time for the end-of-the-day time will be applicable. This is simply because for any *particular time* after 8 am in the example, you will, at 8 am, be less than certain that the cable guy won't have come by this time, no matter how soon after 8 am this time is.

So here is our diagnosis of what is wrong with the ACF principle. (We think this diagnosis is correct, but as the reader will see, we have a more fundamental diagnosis to offer below.) We do not think that there is anything wrong with talking about regret, as long as the notion of regret is understood as based on the standard utility framework, where it can come in amounts, and be both positive and negative. As a result, we endorse both the EFG principle and the Expected Utility principle; they always yield the same result. The problem with the ACF principle is that it is not a genuine implication of the EFG principle for cases involving certainty. The correct implication is the CFG principle. The CFG principle, unlike the ACF principle, requires there to be some particular future time at which you are now certain that you will feel negative gladness, and the CFG principle gives the desired result: unlike the ACF principle, it never dictates a choice that violates standard expected utility calculations.

This diagnosis is closely related to Hájek's.[10] It's just that we now see Hájek's diagnosis in a larger context; that is, we offer the diagnosis from a perspective which takes into account more details concerning the rational guidance of considerations of future gladness and regret. Still, the diagnosis as it stands, even in this larger context, may seem puzzling: is it essential that correct principles of this sort consider *particular* times? We maintain that reference to particular times is not essential. Seeing why leads to a more fundamental diagnosis.

Suppose, for a *particular* amount of regret *r*, God, in whom you put absolute trust, tells you that, at *some unspecified* future time *t*, if you make a certain choice, a rational future self of yours will regret, to amount *r*, having made that choice. We think this consideration of future rational regret offers a guide to present choice, despite the fact that there *fails to be* reference to a *particular* time. We take this to be the intuitively correct view and also to connect with why Hájek's and our initial diagnosis can seem puzzling. Now suppose instead, for a *particular* time *T*, God tells you that, at time *T*, a rational future self of yours will regret having made a certain choice to *some unspecified* amount *d* (and none of your prior probabilities allows you to infer anything about the value of *d*). We do not think this consideration of future rational regret offers a guide to present choice, despite the fact that there *is* reference to a *particular* time. This may seem counter-intuitive, and requires explanation.

Suppose you are now wondering how disinclined you should be to make the choice in question. (This disinclination can be objectively measured, by how many units of utility it would take to get you to make the choice that you were initially disinclined to make.) The fact that you will have some unspecified amount of future rational regret offers no guidance, since, for all it tells you, the degree to which your rational future self will regret having made the choice may be *arbitrarily small*. Hence, any amount of disinclination you decide to opt for could go too far in the other direction to an *arbitrarily large extent*. And so the consideration should not disincline you *at all* to making the choice. The sense in which you could go too far to an arbitrarily large extent consists in the fact that the following proportion could be arbitrarily large: $(d_{\text{in-fact}}-d_{\text{warranted}})/(d_{\text{warranted}}-0)$, where $d_{\text{in-fact}}$ is the amount of present disinclination you would in fact opt for if influenced by your knowledge of unspecified future regret and $d_{\text{warranted}}$ is the amount of present disinclination "objectively" warranted by your actual future amount of regret. In other words,

[10] Hájek (2005, 118n) cites Schervish, Seidenfeld, and Kadane (2004) on the idea of *stopping times*, and maintains that it is reference to *particular* times which prevents the cable guy example from being a counter-example to the Reflection principle. Cf. our discussion of the Reflection principle in Section 4.

what you gain by at least getting to $d_{\text{warranted}}$ could be swamped, to an arbitrarily large extent, by how far past $d_{\text{warranted}}$ you go.

So what's really crucial about the cable guy example is *not* that there is no *particular future time* such that, at that time, you are (before 8 am) certain to feel some amount of regret; *rather*, it's that there is no *particular amount of regret* such that you are (before 8 am) certain that, at some future time, a rational future self of yours will feel that amount of regret. In other words, the fundamental diagnosis is not "no reference to particular time", but rather "no reference to particular amount of regret". This is obscured by the fact that, in the cable guy example, later times exactly correspond to higher amounts of regret (as long as the cable guy has not come up to the time in question).

Given this more fundamental diagnosis, we now see that the EFG principle, which considers both particular times and particular amounts of regret, is just a special case of a broader, more fundamental principle which only considers particular amounts of regret. This is the *General Expected Future Gladness* (GEFG) principle:

> Suppose you have a choice between two options. You should not choose one of these options if your EFG at some (perhaps unspecified) future time toward choice of this option has some particular negative value.

We take this to be an interesting result, though, for reasons similar to those above, it does not offer anything in addition to standard expected utility calculations. The importance of the GEFG principle lies in seeing that it matches expected utility calculations and yields the correct result in the cable guy paradox, and does so without making reference to particular future times. The GEFG principle does, however, make reference to particular amounts of regret. We conclude that it is particular amounts of regret that matter, not particular times.

Let's sum up our argument. The ACF principle makes reference to neither particular times nor particular amounts of regret, and yields the incorrect result in the cable guy paradox. The EFG principle makes reference to both particular times and particular amounts of regret, and yields the correct result. So what matters—particular times, or particular amounts of regret, or both? It turns out that knowing the particular amount of regret matters—if you didn't know the particular amount of regret, then you would fall prey to the swamping argument given above: any degree of disinclination you decide to opt for could go too far in the other direction to an arbitrarily large extent. We can't see any reason that knowing the particular time would matter—as long as you know that you will rationally regret your choice to some particular amount in the future, that should influence your decision now. This is why we conclude that it is particular amounts of regret that matter, not particular times.

4 Reflection and quantified reflection

We have explained why we feel there is more to say with respect to Hájek's diagnosis of the problem with the ACF principle. We also feel there is more to say with respect to his explanation of the Reflection principle's success in addressing the cable guy situation; we'll take up this issue now.

To state the Reflection principle, we need some terminology. Let P_{now} denote one's current probability function, P_t denote one's probability function at time t, and

$P_t(X) = x$ denote the proposition that at time t, the probability one assigns to proposition X is x. Now, the Reflection principle (van Fraassen, 1984) is as follows: in situations where agents have precise numerical probabilities for their opinions, for all propositions X, for all future times t, and for all probability assignments x,

$$P_{\text{now}}(X|P_t(X) = x) = x.$$

There is controversy in the literature about to what extent the Reflection principle is applicable to non-ideal agents; we will avoid this controversy by restricting our discussion to ideally rational agents. (Such agents are logically omniscient, the probability functions representing their opinions are coherent, they always update their opinions via conditionalization, and they never suffer information loss.)

Hájek maintains that the Reflection principle, unlike the ACF principle, is not refuted by the cable guy situation. His explanation for this is that the Reflection principle, unlike the ACF principle, mandates reference to a particular time (2005, 118; he references Schervish et al., 2004). One might think that the cable guy scenario shows that the Reflection principle is false, because you know that there will be some time after 8 am but before the cable guy arrives at which you will assign probability lower than 0.5 to the proposition that the cable guy comes in the first interval (call that proposition *MORNING*). But there is no particular time such that you are certain that at that time you will assign a particular probability other than 0.5 to *MORNING*. Therefore it is consistent with Reflection to now (before 8 am) assign 0.5 to *MORNING*. But the ACF principle does not say that, in order for betting on the first interval to be irrational, you must be certain that *at a particular time t*, you will feel regret; rather it says that you must be certain that there is *some time or other* that you will feel regret. Since you are certain that, if you bet on the first interval, at some time or other you will regret having done so, the ACF principle does not permit this choice. Thus, Hájek maintains that the fact that Reflection refers to a specified future time, while the ACF principle refers to an unspecified future time, explains why the former but not the latter is not refuted by the cable guy situation.

Hájek is technically correct here. The Reflection principle does not require that you assign anything other than 0.5 to *MORNING*, because you are not certain that you will make any probability assignment other than 0.5 at any particular future time.[11] But this discussion seems to be missing something. We maintain that Reflection's reference to particular future times is inessential to it, given its underlying motivation. A variant of Reflection can be stated which does not make reference to a particular future time. The variant has the same underlying motivation as Reflection—it ties one's current probabilities to one's opinions about one's future probabilities. Since the variant is not refuted by the cable guy situation, this shows that it is not just the Reflection principle itself, but the underlying motivations for the principle, which are compatible with the cable guy paradox.

[11] See Section 5 for a more detailed discussion of the implications of Reflection for present opinion given uncertainty about future opinion; it turns out in fact that Reflection requires the probability assignment of 0.5 to *MORNING*.

We will call this variant the *Quantified Reflection* principle. Where Ft means that t is a future time, the Quantified Reflection principle is as follows. For all propositions X, and for all probability assignments x,

$$P_{\text{now}}(X|(\exists t)(Ft\&P_t(X) = x)) = x.$$

Reflection and Quantified Reflection universally quantify over conditional probabilities, where the condition in Reflection, $P_t(X) = x$, entails the condition in Quantified Reflection, $(\exists t)(Ft \ \& \ P_t(X) = x)$. (Note that in Reflection, t stands for some particular future time, while in Quantified Reflection, t plays the role of a variable.) By the probability calculus, the ideal agents with which we are concerned will assign $P((\exists t)(Ft \ \& \ P_t(X) = x)) \geq P(P_t(X) = x)$. In cases in which agents are certain that $P_t(X) = x$, they will likewise be certain that $(\exists t)(Ft \ \& \ P_t(X) = x)$.[12] In these cases, the conditional probabilities required by Reflection and Quantified Reflection have the same consequences for $P(X)$. When agents are not certain that $P_t(X) = x$, they may assign $P((\exists t)(Ft \ \& \ P_t(X) = x)) = P(P_t(X) = x)$. In these cases also, the conditional probabilities required by Reflection and Quantified Reflection have identical consequences for $P(X)$. Finally, in some cases in which agents are not certain that $P_t(X) = x$, they assign $P((\exists t)(Ft \ \& \ P_t(X) = x)) > P(P_t(X) = x)$. In these cases, the conditional probabilities required by Quantified Reflection place different demands on an agent's synchronic coherence than those required by Reflection.

We believe that Quantified Reflection is preferable to Reflection because its consequences in cases in which it differs from Reflection are more reasonable than those of Reflection. Here, we will restrict our discussion to cases in which an agent is certain of $(\exists t)(Ft \ \& \ P_t(X) = x$, but not of $P_t(X) = x$ for any particular time t. (In the next section we will take up cases involving uncertain future credences.) Consider the following example. God tells you that at some (unspecified) time in the future, you will assign 0.4 to M, and you fully believe that what God tells you is true. Reflection in combination with the probability calculus does not require you to assign 0.4 to M, because you are not certain of the time in the future at which you will make this probability assignment—there is no t such that you are now certain that $P_t(M) = 0.4$. Quantified Reflection in combination with the probability calculus does require you to assign 0.4 to M, since you are now certain that $(\exists t)(Ft \ \& \ P_t(M) = 0.4)$. We think that, given the motivation underlying the Reflection principle, you ought to be required in this situation to now assign $P(M) = 0.4$. Given that you are an ideally rational agent, and given that you are certain that in the future you will assign probability 0.4 to M, it makes sense that you should assign probability 0.4 to M now.

Here is another way of motivating Quantified Reflection. First, suppose God tells you that at some particular future time t_1 you will assign probability 0.6 to some proposition A. It follows by Reflection and the probability calculus that you should assign 0.6 to A now. Now consider a variation of that scenario: suppose God instead tells you that either at time t_1 or at time t_2 you will assign probability 0.6 to A. Reflection does not demand that you assign 0.6 to A—but it seems that, if in the first case you ought now to assign probability 0.6 to A, so ought you in the second case.

12 Recall that we are restricting our discussion to ideally rational agents. We think it is legitimate to specify that an ideally rational agent is always aware, for any time t, whether or not that time is in the future.

Now suppose that God tells you that either at time t_1 or at time t_2 or at time t_3 you will assign probability 0.6 to A. The same result obtains. It doesn't matter how many times you add—you could have continuum many times, encompassing the interval from now until the infinite future—and it still follows that you should assign probability 0.6 to A now. Quantified Reflection yields this result, but Reflection doesn't.

In conclusion, we maintain that Quantified Reflection more aptly captures the intuitions underlying Reflection than does Reflection itself. It would be strange to be rationally required to keep one's present beliefs in line with what one is certain will be one's future beliefs, *unless* one is not certain precisely when in the future those beliefs are to take place. Knowing the exact time at which one will have those future beliefs does not seem relevant to whether one should take on those future beliefs now. Thus, we accept Quantified Reflection as a more general formulation of the motivations that underlie Reflection.

Now that we have motivated Quantified Reflection, we will show that it is not refuted by the cable guy situation. Even though you are certain that there will be a time when you assign a probability less than 0.5 to *MORNING*, because you are certain there will be a time after 8 am when the cable guy has not yet arrived, you don't know what that time will be; that's why Reflection does not dictate that you should assign anything other than 0.5 to *MORNING*. But Quantified Reflection does not dictate an assignment other than 0.5 either. Quantified Reflection doesn't require that you be certain that there is a particular future time at which you will have a certain probability assignment in order to mandate that you have that probability assignment now, but it *does* require that you be certain that there is some particular probability assignment you will have in the future in order to mandate that you have that probability assignment now. In the cable guy scenario, you know that in the future you will assign probability less than 0.5 to *MORNING*, since you know that there will be a time after 8 am where the cable guy has not yet arrived, but there is no particular probability assignment to *MORNING* for which you fully believe that you will have that probability assignment in the future. This is why Quantified Reflection does not disallow your current probability assignment of 0.5 to *MORNING*. And so we have reached a deeper understanding of why the cable guy situation is puzzling, even though it fails to refute anything essential to Bayesian theory: you should currently assign 0.5 to *MORNING*, even though you know that at some time in the future your probability assignment to *MORNING* will be in the interval [0, 0.5).

One question remains: if reference to specific future times is not essential to the Reflection principle, *why* is reference to specific probabilities? The answer to this can be nicely illustrated in terms of the cable guy situation itself. Suppose, knowing that there is some probability less than 0.5 which you will attach to *MORNING* some time in the future, you decide to lower your current probability below 0.5. Since the cable guy can come arbitrarily close to 8 am, there is no guarantee for any specific probability p which is less than 0.5 that at some time in the future you will attach a probability to *MORNING* which is less than or equal to p. In Section 3, we argued that any amount of disinclination you have for an option you know you will regret, but where you do not know the degree to which you will regret it, risks going too far in the other direction to an arbitrarily large extent. Here, we see that, no matter how little you fix your current probability below 0.5 in an attempt to reflect your future opinion, you risk going too far to an *arbitrarily large extent*.[13] Avoiding this consequence is why reference to specific probabilities is essential to Reflection.

5 Reflection without certain future credences

Here are four ways in which you can have an opinion about your future credence in *X*:

Type 1: You are certain what your future credence in *X* will be, and at what future time you will have it.

Type 2: You are certain what your future credence in *X* will be, but not at what future time you will have it.

Type 3: You are not certain what your future credence in *X* will be—though you do have partial beliefs about your future credence. You know at what future time one will have that different credence.

Type 4: You are not certain what your future credence in *X* will be—though you do have partial beliefs about your future credence. You do not know at what future time you will have that different credence.

The conditional probabilities required by Reflection combined with the probability calculus entail that, in situations of Type 1, your current credence in *X* should match the credence you're certain that you'll adopt at the specified future time. The conditional probabilities required by Quantified Reflection combined with the probability calculus entail that, in situations of Type 1 and Type 2, your current credence should match the credence you're certain that you'll adopt at some future time. Our discussion of Quantified Reflection shows that it's not having beliefs about the particular future time that matters, it's having beliefs about the particular future credence. Thus, one lesson of the cable guy paradox is that it's more important for plausible principles of belief and choice to make reference to specified future credences than to specified future times.

With this in mind, let's discuss cases of Type 3 and 4. For example, consider a variant of a situation of Type 4, where you are certain that your future credence will be lower than the current one, but you are not certain what your future credence will be, nor at what time you will have it. Should we endorse a principle that mandates that in such a situation you should have a lower credence now?

The cable guy paradox shows that we shouldn't. The cable guy paradox presents exactly the sort of situation just described—you currently assign *MORNING* a credence of 0.5, but you know that in the future you will assign it a lower credence, since you know that there will be a time after 8 am where the cable guy hasn't yet arrived. However, there is no particular credence lower than 0.5 such that you fully believe you will assign that credence. Moreover, there is no particular time at which you fully believe that you will assign some lower credence, since for any time after 8 am, the cable guy could have arrived by that time. If you endorsed the principle that in a Type 4 variant where you know your future credence will be lower, you

[13] The sense in which you could go to far to an arbitrarily large extent consists in the fact that the following proportion could be arbitrarily large: $(p_{\text{future}}–p_{\text{current}})/(0.5–p_{\text{future}})$, where p_{future} is the smallest probability you in fact attach to *MORNING* in the future and p_{current} is the present probability you would in fact opt for if influenced by your knowledge of (largely) unspecified future probability. In other words, what you gain by at least getting to p_{future} could be swamped, to an arbitrarily large extent, by how far past p_{future} you go.

should have a lower credence now, then on the basis of that principle you would have to assign a credence lower than 0.5 to *MORNING* now. This is clearly incorrect, and hence the principle in question is false.

Even though the particular principle we just considered is flawed, there is a way to use true principles, namely Reflection, Quantified Reflection, and the Theorem of Total Probability to analyze cases of Type 3 and 4. We have seen that having the conditional probabilities demanded by Reflection and Quantified Reflection can have implications for one's unconditional probability for the relevant proposition in cases of Type 1 and 2. The Theorem of Total Probability can be used in this calculation. This theorem holds that for a partition $Y_1 \ldots Y_n$, $P(X) = \Sigma(i = 1 \text{ to } n)\ P(X \mid Y_i)\ P(Y_i)$. When the certainty in the conditions of Reflection or Quantified Reflection is involved, this reduces to the sum of one term. For example, if $P(P_t(X) = x) = 1$, then $P(X) = P(X \mid P_t(X) = x)\ P(P_t(X) = x) = x$. This explains why, in the special case in which certainty is involved, Reflection entails that your current credence match your certain future credence.

The Theorem of Total Probability also has implications for your unconditional probabilities in cases of Type 3 and Type 4. We'll show how to do this with Reflection for cases of Type 3. Consider the situation where one is not certain that $P_t(X) = a$, but instead is certain that either $P_t(X) = a$ or $P_t(X) = b$, and one assigns particular probabilities to these two possibilities. It follows that:

$$\begin{aligned} P(X) &= P(P_t(X) = a)P(X|P_t(X) = a) + P(P_t(X) = b)P(X|P_t(X) = b) \\ &= aP(P_t(X) = a) + bP(P_t(X) = b). \end{aligned}$$

This application of Reflection and the Theorem of Total Probability is easily generalizable to situations where one believes there are more than two possibilities for what probability one will assign to X at future time t, but we will stick with the two-possibility formulation.[14] We'll do so both for simplicity, and because the two-possibility formulation is directly applicable to the cable guy paradox.

Note that the cable guy paradox presents a situation of Type 3. In the cable guy paradox, you know that for each time after 8 am, your credence for *MORNING* will be different than 0.5, but you don't know what that credence will be. Specifically, you know that at each time after 8 am, there are two possibilities for what your credence for *MORNING* will be. At each time in the interval (8 am, noon], your credence will either be some quantity less than 0.5 (in the case where the cable guy hasn't arrived yet), or 1 (in the case where the cable guy has already arrived). After noon, your credence will be either 0 (in the case where the cable guy didn't arrive in the first interval) or 1 (in the case where he did).

Let's use Reflection and the Theorem of Total Probability to determine the value of $P(MORNING)$. Let's start by defining $r(t)$ to be the proportion of time that has elapsed in the first interval at time t. For any time t in the first interval, where t_{start} is 8 am, and t_{mid} is noon,

$$r = (t - t_{\text{start}})/(t_{\text{mid}} - t_{\text{start}})$$

[14] Note that this reasoning might not generalize to *all* cases of Type 3. For example, it would not apply to a situation where an agent does not have defined expectation values, if such a situation were indeed possible for ideally rational agents.

If the cable guy has come by t, $P_t(MORNING) = 1$. The probability that the cable guy has come by t is simply 0.5 r. Thus, $P(P_t(MORNING) = 1) = 0.5\ r$. If the cable guy hasn't come by t, $P_t(MORNING)$ is determined by the proportion of the interval (t, 4 pm) that is in the interval (t, noon). In other words, where t_{end} is 4 pm,

$$\begin{aligned} P_t(MORNING) &= (t_{\text{mid}} - t)/(t_{\text{end}} - t) \\ &= (1 - r)/(2 - r) \end{aligned}$$

It now follows that

$$\begin{aligned} P(MORNING) &= [(1 - r)/(2 - r)]P(Pt(MORNING) \\ &= (1 - r)/(2 - r)) + 1P(P_t(MORNING) = 1) \\ &= [(1 - r)/(2 - r)](1 - 0.5r) + 0.5r = 0.5. \end{aligned}$$

This is the desired result.

What about if t is a some future time in the second interval? Here,

$$\begin{aligned} P(MORNING) = {} & 0\ P(P_t(MORNING) = 0) \\ & + 1\ P(P_t(MORNING) = 1) \\ = {} & 0(0.5) + 1(0.5) = 0.5. \end{aligned}$$

This is again the desired result. Thus, when you take into account your uncertain future credences using Reflection and the Theorem of Total Probability, you are indifferent between the hypothesis that the cable guy comes in the first interval and the hypothesis that he comes in the second.

We have shown how Reflection applies to cases of Type 3, and we have shown that it gives the correct result in the cable guy scenario. One can similarly apply the Quantified Reflection principle to cases of Type 4. Because we don't know of any particular philosophical problem to which such a calculation would be applicable, we will leave this as an exercise for the reader.

6 General reflection and quantified general reflection

Van Fraassen (1995) distinguishes between the General and Special Reflection principles. What we (and Hájek) have been calling the Reflection principle is also known as the Special Reflection principle. The General Reflection principle is as follows.

> *General Reflection principle*: My current opinion about event E must lie in the range spanned by the possible opinions I may come to have about E at later time t, as far as my present opinion is concerned. (van Fraassen, 1995, p. 16)

The General Reflection principle is stated so as to account for cases in which one is certain that one's future opinion will be in a specified range. The Special Reflection principle is entailed by it, and is the special case in which the range is a single sharp opinion.

General Reflection is not refuted by the cable guy situation. You are certain that at any specified time t in the interval (8:00 am, noon], your probability for *MORNING* will be in the interval [(1–r)/(2–r), 1]. Your current credence, 0.5, will always lie within this interval; therefore, that probability assignment is compatible with General Reflection. Also, you are certain that at any specified time after noon, your probability for *MORNING* will be in the interval [0, 1], since by that point you will know whether the cable guy has arrived. Again, your current probability assignment of 0.5 is within this range, and hence is compatible with General Reflection.

Now, one might think that, just as we endorsed the Quantified Reflection principle as better capturing the motivation behind the Reflection principle, so we will endorse the Quantified General Reflection principle as better capturing the motivation behind the General Reflection principle. But it turns out that one has to be careful here. There are multiple natural principles that one could get by generalizing the General Reflection principle in the way that we did above with the Reflection principle. We will now present two of these principles, and show (via reference to the cable guy scenario) that one of them is true, while the other is false.

First consider the following natural generalization of the General Reflection principle, which quantifies over times in the same way that the Quantified Reflection principle did.

$$\text{CQGR}: P(X|(\exists t)(Ft\&P_t(X) \in [x, y])) \in [x, y]$$

The "CQGR" stands for "Closed Quantified General Reflection"; the "closed" qualifier is there because the interval [x, y] is closed.

We believe that the CQGR principle is true. Consider in contrast the Open Quantified General Reflection principle:

$$\text{OQGR}: P(X|(\exists t)(Ft\&P_t(X) \in [x, y))) \in [x, y)$$

We will now show that the OQGR principle is false.[15]

In the cable guy scenario, you now, prior to 8 am, know that there is some time after 8 am but before the cable guy arrives where your probability assignment for *MORNING* will be less than 0.5. Thus, you know that there is some time at which your probability assignment for *MORNING* will be in the interval [0, 0.5). By the OQGR principle, it follows that your probability assignment now should be in the interval [0, 0.5). But this is inconsistent with the correct probability assignment of P(*MORNING*) = 0.5. Thus, the OQGR principle is false.

The CQGR principle, in contrast, is not refuted by the cable guy scenario. Since it only works with closed intervals, the relevant information about your future credence for *MORNING* is that it will be in the interval [0, 0.5]. Your current credence of 0.5 does fall within that interval, thus satisfying the CQGR principle.[16]

This reasoning gets at something essential to what is puzzling about the cable guy paradox. You are certain that in the future you will assign a probability less than 0.5 to *MORNING*—and yet you currently assign 0.5 to *MORNING*. The key feature

[15] One could also give a principle that deals with the case where the interval is open on the left and closed on the right, and a principle that deals with the case the interval is open on both sides. For reasons similar to the ones we will give in the main text, these other principles that involve open intervals are also false. Similar remarks hold for more complicated ranges, e.g., (x, y] ∪ [w, z).

[16] More generally, whenever the OQGR principle applies, the CQGR also applies. This is because, whenever a certain probability lies in the range [x, y), it also lies in the range [x, y], as the former is a proper subset of the latter.

that allows you to currently have that credence is that your current credence is one endpoint of the interval of possible future credences.

To see that this is a key feature, consider a contrasting example. Suppose you currently assign probability 0.5 to some proposition B, and you are certain that at some unspecified time in the future you will assign some probability in the interval [0, 0.4) to B. You are incoherent; to be rational, your probability assignment to B must be in the interval [0, 0.4].[17] Using considerations analogous to those we gave to motivate the Quantified Reflection principle, we can see that the fact that the future time is unspecified doesn't matter. Thus, in a situation where you are certain that your current credence is not in the interval of possible future credences, the only circumstances in which you are rational are those in which your current credence is one endpoint of that interval.

One thing we have learned from this discussion of the CQGR and OQGR principles is that, in a situation where you are certain that at some (specified or unspecified) future time your probability for some proposition will fall in an open interval, your present probability need not fall within that interval. It should, however, fall within the corresponding closed interval. This is one of the important philosophical lessons inspired by the cable guy paradox.[18]

References

Hájek A. (2005). The cable guy paradox. *Analysis, 65*, 112–119.

Schervish M. J., Seidenfeld T., & Kadane J. B. (2004). Stopping to reflect. *Journal of Philosophy, 101*, 315–322.

van Fraassen B. (1984). Belief and the will. *Journal of Philosophy, 81*, 235–256.

van Fraassen B. (1995). Belief and the problem of Ulysses and the Sirens. *Philosophical Studies, 77*, 7–37.

[17] Why not require that the probability assignment for B be in the interval [0, 0.4)? We think that considerations analogous to those which show that the OQGR principle is false will show that restricting one's probability assignment for B to the interval [0, 0.4) is too stringent.

[18] We thank Alan Hájek for many helpful comments.