ON THE EXPECTED VALUE OF GAMES WITH ABSENTMINDEDNESS

Adam J. Grove
NEC Research Institute
4 Independence Way
Princeton, NJ 08540
grove@research.nj.nec.com

Joseph Y. Halpern*
Computer Science Department
Cornell University
Ithaca, NY 14850
halpern@cs.cornell.edu
http://www.cs.cornell.edu/home/halpern

April 20, 1997

Abstract

Piccione and Rubinstein argue that a seemingly paradoxical form of *time inconsistency* can arise in games of imperfect recall. Their argument depends on calculating the expected value of a game from the standpoint of a player in the middle of play. We claim that this concept is not well defined in games with absentmindedness (where two nodes on a path can be in the same information set) without additional assumptions. We show that, under some reasonable assumptions, no time inconsistency arises. Different assumptions will validate Piccione and Rubinstein's calculations, but these are such as to remove the appearance of paradox.

JEL classification numbers: C70, D80, D84.

*Most of this author's work was performed at the IBM Almaden Research Center. IBM's support is gratefully acknowledged. This paper is essentially identical to one that will appear in *Games and Economic Behavior*.

1 Introduction

In a fascinating paper, Piccione and Rubinstein [1996] (PR from now on) argue that a seemingly paradoxical form of *time inconsistency* can arise in games of imperfect recall. That is, a player may be tempted to change his strategy, despite getting no new information. To illustrate this problem, PR consider what they call the "absentminded driver paradox", which they describe as follows:

Example 1.1: An individual is sitting late at night in a bar planning his midnight trip home. In order to get home he has to take the highway and get off at the second exit. Turning at the first exit leads into a disastrous area (payoff 0). Turning at the second exit yields the highest reward (payoff 4). If he continues beyond the second exit he will reach the end of the highway and find a hotel where he can spend the night (payoff 1). The driver is absentminded and is aware of this fact. When reaching an intersection he cannot tell whether it is the first or the second intersection and he cannot remember how many he has passed. ∎

The situation is described by the game tree in Figure 1, in which E is the action of turning at an exit, B is the action of continuing past an exit, and X_e is the information set consisting of the two exits. Since e_1 and e_2 are in the same information set, the driver must perform the same (possibly randomized) action at each.

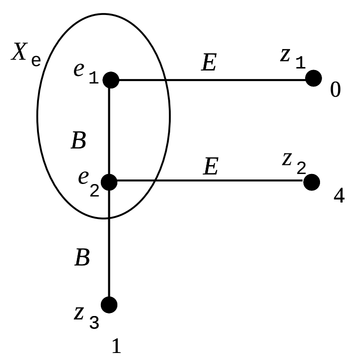

Figure 1: The absentminded driver game.

The only decision the driver has to make is whether to get off when he reaches an exit. The optimal deterministic strategy is clearly never to exit; this gives him a payoff 1. He can actually do better using a behavior strategy. A straightforward computation shows that the driver's optimal strategy *ex ante* is to exit with probability 1/3; this gives him a payoff of 4/3. Throughout this paper, we refer to this strategy as b_{opt}. The interesting question is whether the driver is tempted to depart from b_{opt} when he actually reaches an exit.

The following argument could provide reason for this temptation. When the driver is at an exit, he does not know whether he is at the first or second exit. So he might calculate that his expected payoff following the optimal strategy if he is at the first exit is 4/3, while his expected payoff if he is at the second exit is 2 (because a third of the time he will exit, getting payoff 4, and otherwise he receives payoff 1). Now suppose he ascribes subjective probability α to being at the first exit (and thus probability $1-\alpha$ to being at the second exit). It then seems that his expected payoff when following the strategy b_{opt} is

$$4\alpha/3 + 2(1-\alpha). \tag{1}$$

Notice that this expected payoff is at least 4/3, and as long as $\alpha > 0$, it is greater than 4/3. This seems paradoxical to us. The driver's *ex ante* valuation of the game was 4/3, but as soon as he reaches an exit (which he knew was certain to happen), he apparently thinks the game is worth more. He would refuse an offer of 4/3 to quit the game once it had started, yet at the bar would believe this to be a fair offer. We refer to this divergence in expected utility as the *expectation paradox*.

A calculation similar to that above shows that if the driver uses the strategy of exiting with probability p (with p not necessarily 1/3), then his expected payoff is

$$\alpha((1-p)^2 + 4p(1-p)) + (1-\alpha)((1-p)+4p) = 1 + (3-\alpha)p - 3\alpha p^2. \tag{2}$$

Equation 2 is maximized when $p = \min(1, (3-\alpha)/6\alpha)$ and $(3-\alpha)/6\alpha \geq 1/3$, with equality holding only if $\alpha = 1$. Roughly speaking, what is going on is the following: The driver's expected payoff calculation for the first exit is identical to his calculation while sitting at the bar. So if he is at the first exit the optimal strategy is just b_{opt}, with expected payoff 4/3. But if he is at the second exit, the optimal thing to do is exit (with probability 1), which gives him expected payoff 4. Thus, as long as the driver places positive probability on being at the second exit, it seems reasonable that he would want to use a behavior strategy which places a higher probability on exiting than b_{opt} does. Again, this is paradoxical because he doesn't learn anything new about the game simply because he is at an exit. He can confidently predict that, as soon as he starts his journey (i.e., when he reaches an exit), he will prefer to do something other than follow b_{opt}. We call such an argument a *strategy-change* paradox.

PR discuss their calculations largely in terms of the implications for strategy change. From our perspective, the difficulty with this is that it adds many extra complications that may or may not get to the heart of the matter. In particular, to analyze the strategy-change question carefully, one needs to address several issues that were left somewhat implicit in PR's paper, including the following: How often does the player get to think about changing strategy? If he changes, does he remember his new strategy or is it a one-off deviation? What information does the player have when he chooses? And so on. ([Halpern 1996] addresses the paradox in terms of precisely such questions.)

In contrast, this paper concentrates mostly on the expectation paradox, because it has a much simpler structure. In particular, we hope to avoid clouding the discussion too much with (perhaps controversial) questions about what it really means to choose a new strategy. And yet, as we argued above, the expectation problem is no less a paradox than the strategy-change question. It is also clearly relevant to the latter: after all, if expectations change once one has started the game, then we should not be surprised if b_{opt} (chosen to have the highest *ex ante* expectation) no longer seems to be the best possible strategy. Conversely, just as in the absentminded driver game, strategy-change paradoxes tend to have a correspondingly puzzling expectation version. We discuss the implications of our results to strategy choice at the end of the paper.

However one interprets the paradox, one might suspect that the introduction of subjective probabilities is somehow to blame. After all, in general, the driver's beliefs regarding whether he is at exit 1 or 2 could be arbitrary. PR discuss the question of these beliefs very carefully. Although the beliefs are, in principle, unconstrained, PR are interested in the case where beliefs are "...related in a systematic way with the strategy to be assessed." They discuss two distinct notions of belief which appear, to them, to be reasonable, both of which assign positive probability to being at the second exit. However, as we have just seen, the paradoxical conclusions hold no matter how beliefs are assigned (unless the driver is certain that he never gets to exit 2, which seems quite irrational). So, *provided we accept the calculation of expected utilities described above*, the paradox (in either form) remains.

In contrast to the careful consideration PR give to the question of appropriate choice of beliefs, they appear to take the expected utility calculation for granted. This is indeed a reasonable calculation in games without absentmindedness, that is, games where two distinct nodes in an information set cannot lie in the same path. We show that the notion of expected utility or value of a game at an information set is not well defined in games with absentmindedness, such as the absentminded driver example. To make it well defined, additional assumptions must be made. It is possible to make assumptions that validate PR's calculations. However, the nature of these assumptions is generally such as to remove the appearance of paradox from the expectation calculation.

2 Expected Utility

We focus on single-player games here, just as PR do. Given a game Γ, a behavior strategy b induces a probability distribution that we denote p_b on the space of complete paths in the game tree (or, equivalently, on the terminal nodes in the game tree) in the obvious way. We can associate each node x in the game with an event in this probability space, namely, with the event $\mathit{reach}(x)$ consisting of all paths going through x. Similarly, given a set X of nodes in the tree, we can consider the event of all paths going through some node in X. Thus $p_b(\mathit{reach}(x))$ and $p_b(\mathit{reach}(X))$ are both well defined.

Utility can be viewed as a random variable on this space (where the utility of a path is the utility associated with its terminal node), so we can compute expected utility in the standard way. We can also compute expected utility conditioned on reaching node x or reaching information set X with no difficulty. Let $EU(b)$, $EU(b;\mathit{reach}(x))$, and $EU(b;\mathit{reach}(X))$ denote the expected utility of b, the expected utility of b conditioned on reaching x, and the expected utility of b conditioned on reaching X, respectively.

The expectation $EU(b; reach(X))$ is determined by b, and does not depend on any subjective probabilities. To compare this with Equation 1, which does use subjective probabilities, we need to see how PR choose these probabilities. PR are quite explicit that they are interested in the subjective belief of being *at* a node, not of reaching a node x. To emphasize this issue, given a game Γ, let N_Γ be the set of nodes in Γ, and let $T_\Gamma = \{at(x) : x \in N_\Gamma\}$. (Although T_Γ is isomorphic to N_Γ, we use the "at" notation for events in T_Γ to emphasize the fact that we are considering being at a node x, rather than reaching x.) If $X \subseteq N_\Gamma$, then we take $at(X)$ to be the event $\{at(x) : x \in X\}$ in T_Γ. If b is a (behavior) strategy in the game Γ, let μ_b be the probability distribution on T_Γ defined as

$$\mu_b(at(x)) = p_b(reach(x))/\sum_{x' \in N_\Gamma} p_b(reach(x')). \tag{3}$$

PR observe that $\mu_b(at(x)|at(X))$ is equal to the long-run proportion of times in which one visits $x \in X$ if strategy b is being played repeatedly. More precisely, if the game is played repeatedly using strategy b, and we compute the total amount of time spent at any node in X (counting one time unit per node visited), then the fraction of this time spent at x converges to $\mu_b(at(x)|at(X))$. PR say that a belief assessment μ is *consistent with b* precisely if $\mu(at(x)|at(X)) = \mu_b(at(x)|at(X))$ for each information set X reached with positive probability.[1] If μ is consistent and there is no absentmindedness, PR's expected utility calculation gives exactly what we would expect, as formalized in the following easy proposition:

Proposition 2.1: *In games without absentmindedness, for every strategy b and information set X such that $p_b(reach(X)) > 0$, we have*

$$EU(b; reach(X)) = \sum_{x \in X} \mu_b(at(x)|at(X))\, EU(b; reach(x)).$$

Proof: The proof is straightforward. In games without absentmindedness the events $reach(x)$, $x \in X$, form a partition of the event $reach(X)$, simply because no two nodes in X lie on the same path. So, just from the definition of expectation, we have $EU(b; reach(X)) = \sum_{x \in X} p_b(reach(x)|reach(X)) EU(b; reach(x))$. However, using the disjointness of the $reach(x)$ events again, we have $p_b(reach(X)) = \sum_{x' \in X} p_b(reach(x'))$. Thus

$$p_b(reach(x)|reach(X)) = p_b(reach(x))/p_b(reach(X)) = \mu_b(at(x)|at(X)),$$

from which the result follows. ∎

[1] As is standard in the literature, PR do not assume that there is one belief assessment over all of T_Γ. Rather, they allow a separate assessment $\mu(\cdot|at(X))$ for each X, which (despite the notation) is not necessarily obtained by conditioning on one distribution. Nevertheless, the definition of consistency we give is equivalent to theirs.

Thus, in games without absentmindedness, it really is the case that the expected utility of strategy b conditioned on reaching information set X can be computed by taking the expected utility of the strategy conditioned on reaching each node $x \in X$, and weighting that according to the (consistent) subjective probability of being at x. This computation (i.e., the expression occurring in Proposition 2.1) is our interpretation of the expected utility computation that PR consider in their discussion of time (in-)consistency. In particular, it reduces to Equation 1 in the absentminded driver example. Note that PR's notion of time consistency compares this expression (interpreted as the expected payoff of the continuing with the current strategy) with a different calculation intended to measure the expected payoff of changing to another strategy. This is because, unlike us, they wished to address the strategy-change issue explicitly.

The argument we gave for Proposition 2.1 is not correct in games with absentmindedness, because the events of reaching nodes in an information set no longer define disjoint events on the space of paths. Given an information set X, define the *upper frontier of* X, denoted $\hat{X}$, to consist of all those nodes $x \in X$ such that there is no node $x' \in X$ that strictly precedes x on some path from the root. The notion of upper frontier was introduced in [Halpern 1996], where it was argued that if a player switched strategies at all, the switch had to occur at the upper frontier. The upper frontier also plays a role in the computation of expected value. A similar proof to that of Proposition 2.1 shows:

Proposition 2.2: *For every game, every strategy b, and every information set X such that $p_b(reach(X)) > 0$, we have*

$$EU(b; reach(X)) \quad = \quad \sum_{x \in \hat{X}} \mu_b(at(x)|at(\hat{X}))EU(b; reach(x)).$$

This says that, in order to compute the expected utility of strategy b conditioned on reaching information set X, we can focus on $\hat{X}$, compute the expected utility of b for each node in $x \in \hat{X}$, and weight that according to the subjective probability of being at x conditioned on being at $\hat{X}$. Of course, $X = \hat{X}$ in games without absentmindedness. In the absentminded driver example, if X_e is the information set $\{e_1, e_2\}$, then $\widehat{X_e} = \{e_1\}$, and the expected utility conditional on reaching an exit, according to this calculation, is 4/3, independent of the subjective probability the driver places on being at the first exit and second exit.

Why does this calculation differ from what may seem to be the more intuitive calculation considered in the introduction, which took into account the possibility of being at either exit? The most important difference is that Equation 1 includes terms for both $EU(b; reach(e_1)) = 4/3$ and $EU(b; reach(e_2)) = 2$, even though $reach(e_1)$ and $reach(e_1)$ are not mutually exclusive events. Perhaps our natural inclination is to think of these events as mutually exclusive (which is the case in games without absentmindedness) because we are thinking of being *at* an exit rather than reaching one.

This suggests that we consider a different calculation that uses $EU(b; at(e_i))$ instead of $EU(b; reach(e_i))$; that is,

$$\sum_{x \in X} \mu_b(at(x) | at(X))\, EU(b; at(x)).$$

This keeps the spirit of Equation 1 and Proposition 2.1, but always conditions on disjoint events (i.e., $at(x)$), and does so consistently (rather than mixing $reach(x)$ and $at(x)$). However, $EU(b; at(e_i))$ is not yet well defined. To define it, we need a space in which, not only is "being at an exit" an event, but utility is a random variable.

There are many probability spaces that we might construct in which $EU(b; at(e_i))$ is defined. The space T_Γ is not so convenient for defining utilities; we thus consider a slightly more general construction. Given a game Γ, let S_Γ consist of all pairs (z, x) such that z is a terminal node in Γ (and thus determines a path in the game tree) and x is a node on the path leading to z, which is denoted $x \preceq z$. Intuitively, (z, x) can be thought of as saying "the path taken is (or will be) z, and the player is currently at x". Note that $at(x)$ is a well-defined event in this space for any node x in the tree, consisting of all pairs (z, x) such that $x \preceq z$. Furthermore, note that if x and y are distinct nodes, then $at(x)$ and $at(y)$ are mutually exclusive events. Thus, T_Γ can be embedded in S_Γ in the obvious way.

We now must define a utility function and a probability distribution over S_Γ. There are a number of reasonable choices that can be made for each. We begin by considering a distribution that corresponds to having a belief assessment consistent with the strategy chosen. Given a behavior strategy b in Γ, let $q_b(z, x) = p_b(z) / \sum_{(z', x') \in S_\Gamma} p_b(z')$. It is easy to see that for the strategy b_{opt} in the absentminded driver example, we have $q_{b_{\mathrm{opt}}}(at(e_1) | at(X_e)) = 3/5$ and $q_{b_{\mathrm{opt}}}(at(e_2) | at(X_e)) = 2/5$ (where X_e is the information set $\{e_1, e_2\}$). Note that, according to Equation 3, $\mu_{b_{\mathrm{opt}}}(at(e_1) | at(X_e)) = 3/5$ and $\mu_{b_{\mathrm{opt}}}(at(e_2) | at(X_e)) = 2/5$. In general, for any game Γ, strategy b, and information set X, it is not hard to show that $q_b(\cdot | at(X))$ is a belief assessment consistent with b.

Finally, we must define utility as a random variable over S_Γ. Having the terminal node as one of the components of an element of S_Γ makes this easier to do for S_Γ than for T_Γ. Nevertheless, there are several reasonable ways it can be done. Perhaps one's first thought would be to take $u(z, x)$, the utility of a pair (z, x), to be the payoff at terminal node z; this is the utility function we focus on from here on.

With these definitions we can make formal sense of the expected utility of using strategy b, interpreted here as expected utility in the space $\mathcal{Q}_{\Gamma,b} = (S_\Gamma, q_b, u)$. We denote this expected utility $EU_{\mathcal{Q}}(b)$, to emphasize the fact that we are working in $\mathcal{Q}_{\Gamma,b}$. Similarly, we can compute conditional expected utilities in $\mathcal{Q}_{\Gamma,b}$. It is easy to check that $EU_{\mathcal{Q}}(b_{\mathrm{opt}}; at(e_1)) = 4/3$ and $EU_{\mathcal{Q}}(b_{\mathrm{opt}}; at(e_2)) = 2$. Since $at(e_1)$ and $at(e_2)$ are disjoint events, we can calculate $EU_{\mathcal{Q}}(b_{\mathrm{opt}}; at(X_e)) = 4/3 \times 3/5 + 2 \times 2/5 = 8/5$.

In this space, it seems that the PR's expected value calculations are justified. However, it must be stressed that $\mathcal{Q}_{\Gamma,b_{\mathrm{opt}}}$ is very different from the original space. For example, a straightforward calculation shows that the unconditional expected utility of b_{opt} using

$\mathcal{Q}_{\Gamma,b_{\text{opt}}}$ is 3/2, not 4/3. Thus, we must ask to what extent any expectation in $\mathcal{Q}_{\Gamma,b_{\text{opt}}}$ (and in particular, $EU_{\mathcal{Q}}(b_{\text{opt}}; at(X_e))$) tells us something about the original example. There are other, quite different, spaces that we might have considered instead of $\mathcal{Q}_{\Gamma,b_{\text{opt}}}$. How should we decide among the various possibilities?

A good way of understanding the issues involved is to give a concrete operational interpretation of the new space(s) being considered. One possibility is as follows. Imagine playing the absentminded driver game over and over again. From time to time, an external agent $\mathcal{A}$ stops you in the middle of play, and asks you how much you would pay to receive the payoff of the play being interrupted. To make things even more concrete, we stipulate that the external agent knows how the game will end and, if you do pay him, immediately deposits the amount of the game's payoff in your bank account.[2] It is important to note that this story, interpreted literally, is only relevant to the expectation form of the paradox. The only option the player has is to refuse the offer (which means, if he is never given the offer again in this game, that he has effectively left the game). It would require a much more elaborate model in order to give the player the opportunity to choose to continue playing but with a different strategy.

In this setting, when you are stopped, your situation will be adequately described by some pair (z, x). Thus, S_Γ is a good space over which to analyze this new game. By assumption, if the situation is (z, x), then the payoff you stand to earn is the payoff associated with z. Thus $u(z, x)$ as defined above is also appropriate. The missing ingredient is the probability distribution over S_Γ. In part, this distribution is determined by the strategy one is actually playing. It is worth noting that, in strategy-change problems, this dependence is problematic because it appears somewhat circular. Roughly speaking, this is because in some formulations of the strategy-change paradox, the player must take into account that he might have *already* been in this information set, performed exactly the same calculation he is doing now, and decided to change strategies. However, in our setting this issue is straightforward because the only option is to buy the game's payoff. Even if the driver had done this in the past, he knows that he has also continued to play his initial strategy (and will do so in the future), and so can he can base his beliefs on this knowledge. But even in our setting, the distribution over S_Γ depends on more than the player's initial strategy: it is also determined by the strategy $\mathcal{A}$ uses to decide when to stop you. This is perhaps the more interesting dependence.

For instance, consider the agent $\mathcal{A}_\gamma$ who decides randomly, at each point, whether to stop you (i.e., at each point he tosses a coin of constant bias γ, whose outcome is independent of previous tosses). It is easy to verify that, if you are playing strategy b, then the resulting distribution over places where you are stopped is precisely q_b (independent of γ). We can thus view the new game, involving $\mathcal{A}_\gamma$, as a concrete instantiation of the

[2] The assumption that the external agent can "predict the future" is a rather harmless one. In fact, it is not even necessary for some of the models we discuss and is made only because we want to be extremely concrete about how payoffs are made. Even where the assumption is important, it is easy to simulate it by having all random decisions determined in advance (such as when the driver is still in the bar) and made known to $\mathcal{A}$ at that point.

space $\mathcal{Q}_{\Gamma,b_{\text{opt}}}$ defined earlier. In the new game $EU_{\mathcal{Q}}(b_{\text{opt}}; at(X))$ can be interpreted as the amount the player should be prepared to pay if he has been stopped and all he can tell is that he is in information set X. If he pays this amount, he will break even in the long run.

How should we feel about $EU_{\mathcal{Q}}(b_{\text{opt}}; at(X_e)) = 8/5$, given this interpretation? On the one hand, if one really is playing the new game (involving the external agent $\mathcal{A}_\gamma$), this is undoubtedly the right value to pay. On the other hand, the new game differs in important respects from the original. If γ is high then the agent $\mathcal{A}_\gamma$ may stop you more than once in a game and, according to the rules we have given, will pay you each time you buy into the game. For instance, if you are stopped twice in a play of the original game that ends at z_2, then $\mathcal{A}_\gamma$ will need to reward you with $8 = 4 \times 2$ units altogether. This seems to be against the spirit of the original game (in which, for such a play, exactly 4 units of reward is made available). It is not surprising that the new game is worth more to play.

If we would prefer that the player never be paid twice per evening, there are many ways to ensure this. Some of these involve using a utility function other than u. For example, we could define $u'(z, x)$ to be 0 unless $z = e_1$ (so that $\mathcal{A}$ only has to reward you when you are in fact at the first exit). Or we could expand the space S_Γ to keep track of whether the player has been stopped already that evening (and if so, then the utility at stake in any subsequent stop is 0). Another idea, which does not involve modifying u, is to consider $\mathcal{A}_\gamma$ in the limit as $\gamma \to 0$, in which case the probability of being asked twice in an evening becomes vanishingly small. These possibilities, and many others, are all reasonable. Indeed, the story we have given (concerning an agent $\mathcal{A}$ and repeated plays of the game) is certainly not unique either. What we believe is important is simply that the complete model be defined and interpreted *somehow*. Only then can we form useful opinions as to what (if anything) a certain calculation tells us about the original game.

With respect to understanding the absentminded drivers paradox, it turns out that none of the utility models mentioned in the previous paragraph lead, by themselves, to the *ex ante* expectation of 4/3. But this is not surprising. Consider the case of $\mathcal{A}_\gamma$ in the limit as $\gamma \to 0$. We have already seen that 8/5 continues to be the fair value of this game. But even though you do not stand to earn two rewards per evening, there are some evenings in which you are never stopped at all (and so do not pay anything). Furthermore, if you are stopped it is more likely to be in a game that gets to the second exit (because you are twice as likely to be stopped at some exit in such a game). Because games in which the second exit is reached are preferred, it is not surprising that expected rewards should be biased upward when compared to the original game.

This bias toward longer games shows up in another way. In fact, it allows us to generate a purely probabilistic version of the paradox that does not involve utilities at all. Just as $reach(z_i)$ is an event in the original probability space, it is also an event in S_Γ: $reach(z_i)$ is the set of all pairs (z_i, x) whose first component is z_i. However, $q_b(reach(z_1)|at(X_e)) = 1/4$, whereas $p_b(reach(z_1)|reach(X_e)) = 1/3$; the corresponding probabilities for z_2 and z_3 also differ. Thus, being "at" some node in the information set (which the driver knew all along he was going to reach) will result in the driver changing his initial assessment of the probability of reaching the terminal nodes, so as to favor nodes that terminate longer paths. This change is not just an artifact of our use of q_b; if the agent's beliefs μ_b are consistent with b (in PR's sense), and we define $\mu_b(\mathit{reach}(z_i)|\mathit{at}(X)) = \sum_{x\in X} \mu_b(\mathit{at}(x)|\mathit{at}(X))p_b(\mathit{reach}(z_i)|\mathit{reach}(x))$ (in analogy with PR's expected utility calculation), then we get the same paradoxical result.

What if we are uncomfortable with this implication of q_b? We say that a distribution r over S_Γ is *outcome-uninformative at information X* (under strategy b) if, for all terminal nodes z, we have $r(\mathit{reach}(z)|\mathit{at}(X)) = p_b(\mathit{reach}(z)|\mathit{reach}(X))$. This is not quite as strong as saying that being at a set conveys no more information than reaching the set,[3] but rather that whatever extra information (if any) is conveyed by being at this set, it is not enough to help us predict the game's final outcome. By definition, the purely probabilistic "paradox" vanishes for outcome-uninformative distributions. A little more work shows that if the driver uses an outcome-uninformative distribution, then his *ex ante* expected utility is the same as his expected utility at the information set.

Proposition 2.3: *Suppose r is outcome-uninformative for X under strategy b, and let $\mathcal{R} = (S_\Gamma, r, u)$. Then*

$$EU_{\mathcal{R}}(b; \mathit{at}(X)) = EU(b; \mathit{reach}(X)).$$

Proof: A straightforward calculation gives us

$$\begin{aligned} EU_{\mathcal{R}}(b; \mathit{at}(X)) &= \textstyle\sum_{(z,x)\in \mathit{at}(X)} u(z,x)\, r(z,x)/r(\mathit{at}(X)) \\ &= \textstyle\sum_{z\in Z} \sum_{\{x:(z,x)\in \mathit{at}(X)\}} u(z,x) r(z,x)/r(\mathit{at}(X)) \\ &= \textstyle\sum_{z\in Z} u(z) \ \sum_{\{x:(z,x)\in \mathit{at}(X)\}} r(z,x)/r(\mathit{at}(X)) \\ &= \textstyle\sum_{z\in Z} u(z)\ r(\mathit{reach}(z)|\mathit{at}(X)) \\ &= \textstyle\sum_{z\in Z} u(z)\ r(\mathit{reach}(z)|\mathit{reach}(X)) \\ &= EU(b; \mathit{reach}(X)). \end{aligned}$$

∎

We believe that this proposition provides some resolution to the expectation form of the absentminded driver paradox. Either the driver gains knowledge simply due to the fact that he is contemplating expected utility at an exit (because the probability distribution is not outcome-uninformative), or else he continues to accept the *ex ante* expected utility value.

It turns out that not only does an outcome-uninformative distribution always exist, but that PR have done much of the work of defining one for us. Recall that we said earlier that PR considered two ways of computing subjective beliefs. Although they ended up working mostly with the notion of consistent belief, they also defined a notion called *Z-consistency*. Let Z be the set of terminal nodes in a game Γ. PR say that a belief

[3] This stronger condition might be captured by requiring $r(Y|\mathit{at}(X))$ and $p_b(Y|\mathit{reach}(X))$ to agree for all events Y that are defined in both spaces.

assessment μ is *Z-consistent* with a behavior strategy b in Γ if for every information set X reached with positive probability we have

$$\mu(at(x)|at(X)) = \frac{\sum_{\{z\in Z: x\preceq z\}} \frac{p_b(z)}{|\{x'\in X: x'\preceq z\}|}}{\sum_{\{z\in Z: X\preceq z\}} p_b(z)}, \quad (4)$$

where $X \preceq z$ if $x \preceq z$ for some node $x \in X$.

Does Z-consistency have any analogue in the richer space S_Γ? Given a game Γ and a node x in Γ, let X_x be the information set containing x if x is a decision node (one where the player performs an action), and just $\{x\}$ otherwise. Consider the following distribution:

$$r_b(z,x) = \frac{p_b(z)}{K \mid \{x' \in X_x : x' \preceq z\} \mid}, \quad (5)$$

where K is a normalization constant. It is easy to see that r_b is Z-consistent with b.

We can also understand r_b in terms of an external agent $\mathcal{A}_Z$. Again, we assume that $\mathcal{A}_Z$ knows the outcome of each particular play of the game. At the beginning of a play of the game with outcome z, for each information set X containing a node on the path to z, $\mathcal{A}_Z$ tosses a fair coin to determine whether or not to stop the player in X. If the coin for X comes up heads, $\mathcal{A}_Z$ then chooses the node at which to stop the player by choosing uniformly at random among the nodes on the path to z that are in X (so that, in particular, the player is only stopped once in X). If the coin for X comes up tails, the player is not stopped at X at all. It is easy to see that the long run proportion of times at which the player is stopped at (z,x) is indeed given by r_b. Moreover, r_b is outcome-uninformative. For suppose that the player is stopped somewhere in information set X. What else can the player deduce from the fact that he was stopped? In this game, the only extra information he learns is that $\mathcal{A}_Z$ threw heads on the coin he used for deciding whether to stop in X. But $\mathcal{A}_Z$'s coin toss was independent of all other factors, so this is not especially useful information. This intuition can be formalized to show the following:

Proposition 2.4: *If strategy b is being played, then r_b is outcome-uninformative for every information set.*

To summarize, if we use the distribution r_b (which is very much like PR's Z-consistent distribution, although defined over the space S_Γ), together with the utility function u as defined earlier, and do our calculations in the space S_Γ, then the driver's expected utility at the information set is the same as his *ex ante* expected utility, at the bar.

PR make the interesting observation that a player whose beliefs are Z-consistent is vulnerable to a sort of "money pump". They define a particular scenario in which the player is guaranteed to lose money if he plays the game repeatedly. Their scenario involves a game with absentmindedness, in which the player is offered a certain bet each time he reaches a point in an information set. In particular—and this is crucial to their example—he is offered the bet twice in a play if the information set is reached twice, and if he were to accept twice he must pay twice. Since this is inconsistent with the story we gave to motivate Z-consistency (involving $\mathcal{A}_Z$, who only stops the player once per information set), it is not surprising that Z-consistency leads to the wrong expectation in this game. However, it is easy to construct a similar money pump if the player's belief assessment is consistent (i.e., if the player uses q_b rather than r_b). For example, if the absentminded driver is stopped each time he is in the information set X_e, and pays 8/5 to play the game (which, as we have seen, is what he calculates as his expected utility if he is using strategy $b_{\rm opt}$ and his beliefs are consistent), but is paid only once per play of the game, then he can expect to lose 4/5 on average each time he plays the game.

Such examples can be used to show that *any* notion of expected utility will be appropriate only for certain games, and subjects the player to a money pump if he persists in using it in other cases. Indeed, we doubt that there can be any argument that demonstrates the general superiority of one particular belief assessment approach in games of absentmindedness. This emphasizes the need for operational scenarios corresponding to the choice of any particular belief assessment. While we have provided such scenarios for q_b and r_b (involving external agents $\mathcal{A}_Z$ and $\mathcal{A}_\gamma$, and repeated plays of the original game), as well as giving a more abstract condition (outcome-uninformativeness) that distinguishes r_b from q_b, we do not mean to imply that these are the only possible scenarios (nor that these are the only possible reasonable belief assessments).

We close this discussion by asking how Proposition 2.3 reconciles with the "paradoxical" calculation in the introduction which, in the absentminded drivers example, gives an answer that is greater than 4/3, essentially independent of distribution. The explanation is that the expected utility calculation in the introduction does not apply to r_b because the underlying space is different. In particular, it is easy to verify that $EU_{\mathcal{R}}(b_{\rm opt}; at(e_1)) = 1$, rather than 4/3. If the driver is stopped at e_1, it becomes more likely that he will exit at z_1 (because if the final outcome will be z_2 or z_3, $\mathcal{A}_Z$ has more choices of where to stop him), and so the expectation conditional on being stopped at e_1 is less than 4/3 (and is in fact 1). On the other hand, $EU_{\mathcal{R}}(b_{\rm opt}; at(e_2)) = 2$, as before. Since $r_{b_{\rm opt}}(at(e_1)) = 2/3$ and $r_{b_{\rm opt}}(at(e_2)) = 1/3$, we get that the expected utility is 4/3.

3 Discussion

The message of this paper should be obvious: The notion of expected payoff to a player at a point in the middle of a game is a subtle one, especially in games with absentmindedness. To calculate the expected utility, we must first decide whether to condition on the event of *reaching* the information set or *being at* the information set. In the former case, as we have seen, the correct expectation—$\sum_{x\in\hat{X}} p_b(reach(x)|reach(\hat{X}))EU(b; reach(x))$—is equivalent to that provided by Proposition 2.2, and not to the calculation used in Proposition 2.1 (at least, not in games with absentmindedness).

On the other hand, if we really intend to condition on the event of being *at* the information set, then we must consider carefully how to construct a space that captures this event, and what probability distribution and utility are appropriate for it. Various answers are possible depending on the assumptions made, although the ones that seem best motivated to us do not lead to time inconsistency. Our view of PR's absentminded driver game is that it is an important cautionary example, showing how careful one must be in doing expected utility calculations in games with absentmindedness.

So what does all this tell us about time consistency and the examples given by PR? We have deliberately not discussed strategic concerns here, preferring to focus on the case of a fixed strategy. Nevertheless, we argue that our results are directly relevant to the fairly narrow question of why PR were able to produce paradoxical conclusions in the absentminded driver example. This is because the expected utility calculation that we have been analyzing lies at the heart of PR's formulation of the strategy-change issue (i.e., their notion of time consistency). Our argument is that there are "paradoxes" inherent in this calculation itself, which we have endeavored to resolve.

What about the broader question, of whether and how one should change strategies in the middle of a game? We have not addressed this directly in this paper. However, it should be clear that many of the issues that we have considered in the context of calculating expected utility also arise in the richer context of strategy change. In particular, as shown in [Halpern 1996], the calculations used both by PR and by Aumann, Hart, and Perry [Aumann, Hart, and Perry 1996] can be understood in terms of an external agent that stops the driver at most once in each play of the game. This suggests that we should then consider an expanded game that includes the external agent. As shown in [Halpern 1996], in this expanded game, the paradox disappears.

PR also give examples of apparent time inconsistency in games without absentmindedness, but the issues raised in these examples are quite different from those raised by the absentminded driver example. (See [Halpern 1996] for some discussion of these issues.) The subtleties that arise in calculating expected utility at an information set in games with absentmindedness tend not to arise in games without absentmindedness. The lack of absentmindedness means that there is at most one representative of any given information set per path. Thus, from the standpoint of someone in that information set considering the various possible points at which he might be, there is a one-to-one correspondence between the events of being *at* a point and *reaching* a point (i.e., reaching it sometime during the play of that game). It is thus only natural to conflate the two, and it seems to be relatively harmless to do so. In the case of absentmindedness, there is no one-to-one correspondence between "being at" and "reaching" for us to rely on.

References

Aumann, R. J., S. Hart, and M. Perry (1996). The absent-minded driver. *Games and Economic Behavior*. This issue.

Halpern, J. Y. (1996). On ambiguities in the interpretation of game trees. *Games and Economic Behavior*. This issue.

Piccione, M. and A. Rubinstein (1996). On the interpretation of decision probems with imperfect recall. *Games and Economic Behavior*. This issue.