Diachronic rationality and prediction-based games

David Wallace*

3rd April 2009

Abstract

I explore the debate about causal versus evidential decision theory, and its recent developments in the work of Andy Egan, through the method of some simple games based on agents' predictions of each other's actions. My main focus is on the requirement for rational agents to act in a way which is consistent over time and its implications for such games and their more realistic cousins.

Traditional "evidential" decision theory seemed to work fine, until Newcomb's problem forced philosophers to consider situations where an agent's choices provide good evidence for some fact holding without being causally relevant to that fact holding. In Newcomb's problem, an agent makes a choice between two options: he knows that he will do better in taking the first than the second option no matter what the world is like, but also that taking the second option provides strong evidence that the world is a pretty good place, and that taking the first option provides strong evidence that it's not so good. The majority of philosophers, therefore, converted to "causal" decision theory, which evaluates an act's expected utility with respect to an agent's unconditional credences in the state of the world, not their credences conditional on the actions they in fact perform.

But it's turned out that causal decision theory, too, has its problems. In Newcomb's problem, option one is better than option two whatever the world's like. But what if taking option one provides good evidence that option two would have been the better choice, and vice versa? Andy Egan's recent (2007) discussion of such examples has stirred a controversy as to whether causal decision theory can after all be defended or whether, as he advocates, we need to look for yet another formal system for decision-making.

This paper can be taken as either a partial defence of causal decision theory, or a partial development of an alternative: the decision theory I propose essentially *is* causal decision theory but requires rational agents to conform to

*Balliol College, Oxford OX1 3BJ; and Philosophy Faculty, Oxford University. Email address: david.wallace@balliol.ox.ac.uk

constraints on their credences as well as their preferences. I develop it, in large part, in the context of fairly simple, stylised two-person games.

I should acknowledge up front that much here is not new. The strategy I advocate is a variant on ratificationism (originally advocated by Jeffrey (1983) in defence of *evidential* decision theory, and considered and rejected by Egan). And both the details of my proposal and my diachronic rationale for it are in many ways close to a recent proposal due to Arntzenius (2007). The points of novelty, I think, are: the defence I offer of "diachronic consistency", the intepretation I wish to place on mixed acts, the generality of the analysis I offer of a certain family of games, and the proposal that decision theory may sometimes be silent as to which act an agent should perform not because none of them are rationally acceptable but because several are each acceptable by their own lights.

1 A simple game

Alice and Bob are playing a simple game, which goes like this. There are two opaque boxes, A, and B, and Bob is to choose between them. Alice secretly places either three dollars in box A or two dollars in box B — these aren't her dollars, the gamesmaster has supplied them. Bob then chooses to open either A or B and gets to keep the contents of the box. *If he chooses the empty box*, Alice gets one dollar. (I assume, harmlessly, that utility is linear in dollars for both of them.)

What should each do? We'll come at it in stages. Start by assuming that the game is very heavily loaded in Alice's favour:[1] Bob is compelled to tell her his strategy in advance — before she places the money — and then is bound to keep to it (by the high disutility he places on dishonesty; by getting a third party to execute the strategy; by armed guards with telepathic powers — take your pick).

If Bob's plan is to take box A then Alice will put the money in box B, and vice versa; either way, he gets nothing. But students of game theory will have spotted an alternative: he can decide to take box A with probability p and box B with probability $1-p$ (assume, for the moment, that he carries a handy random-number generator with him for such purposes).

What should Alice do? It's pretty simple: if $p < 0.5$, then Bob is more likely to pick box B than box A, so she should put the money in box A; if $p > 0.5$, she should put the money in box B. And if p is 0.5 exactly, it doesn't matter what she does.[2]

So: what is the optimum choice of p for Bob (assuming that Alice is rational)? He should choose p to be just slightly less than 0.5: this guarantees that Alice puts three dollars in box A, and his expected winnings are just shy of 1.5

[1] As we'll see, it doesn't at all follow that it's loaded in Bob's disfavour. Alice and Bob are not playing against one another.

[2] It's easy to see that Alice doesn't benefit from adopting a probabilistic strategy in this situation.

dollars. (Conversely, if he chose p to be just slightly more than 0.5, she'd put two dollars in box B and his expected winnings would be just under one dollar: the reason for choosing $p < 0.5$ is that Alice puts more money in box A than in box B.[3]

Basically the same thing happens if we reverse the loading, so that Alice rather than Bob has to disclose her strategy in advance: it's easy to show that her best strategy is to put money in box A with probability just under 0.6, and that her expected winnings are then \$0.6 or just under.

In the real game, though,Alice and Bob's possible strategies are unchanged from these simple cases: choose each box with some probability. But now neither has prior knowledge of the other's choice of probability, so each is in the position of having to guess the other's strategy (on whatever grounds come to hand: assuming the other is somewhat rational; prior knowledge of the other; access to a team of well-briefed psychologists, whatever). We will assume that neither is *hopeless* at that task (more accurately, we will assume that neither believes the other to be hopeless), but we will make no assumption that either is perfect, or even preternaturally talented, at the task.

Consider Bob: he wants to choose box A if the money is in A and box B if it's in B, and we might hope to be able to represent his degree of confidence as to which box it's in by some credence: we'll write q for Bob's credence that the money is actually in box B. So if Bob chooses to take box A with probability p (for convenience, say just "if Bob chooses the p-strategy") the expected utility of his choice is

$$U_B(p, q) = 3p(1-q) + 2q(1-p) = 2q + p(3 - 5q). \tag{1}$$

But there is a complication. Bob knows that which box the money is in depends on Alice's predictions of his actions, and if she is any good at predicting his actions, his value of q should increase as p increases: the more likely he is to take box A, the more likely it is that Alice has put the money in box B. Mathematically, this means we should replace q with an increasing function $g(p)$ of p.

Traditionally, decision theory has offered Bob a choice of two theories to decide which p-strategy to select.

- According to *evidential decision theory* (EDT; see, e. g. , Jeffrey (1983)), Bob should choose that value of p which maximises

  $$V_B(p) = U_B(p, g(p)), \tag{2}$$

  the expected utility of Bob choosing the p-strategy evaluated with respect to Bob's credence in the money's location *given that he chooses the p-strategy*. For instance, suppose we make the (unrealistically crude) assumption $g(p) = p$: that is, conditional on Bob choosing the p-strategy,

[3]Strictly, then, Bob has no absolutely *maximal* strategy: whatever his choice of p, there is some slightly better choice available. I take this to be a harmless artifact of the model's simplicity.

Alice has probability p of putting the money in box B. Then

$$V_B(p) = 3p(1-p) + 2p(1-p) = 5p - 5p^2, \tag{3}$$

which is maximised at $p = 0.5$, for an expected return (as calculated by a third-party observer) of 1.25.

- According to *causal decision theory* (CDT; see, e. g. , Lewis (1981)), Bob should choose that value of p which maximises $U(p,q)$ for whatever value of q represents his actual credence in the money being in box B. If $(3-5q) > 0$ — that is, if $q < 0.6$ — then this means he should choose $p = 1$, and take box A with certainty; if $q > 0.6$, he should take box B with certainty.

Both strategies have this odd property: in general they commit Bob to deciding to do something which he will then regret deciding. Start with EDT: in the example given, Bob will choose $p = 0.5$. Conditional on this choice, he now believes there to be a 50% probability that the money is in box A. But on that assumption, his choice no longer looks optimal: choosing box A with certainty gives better rewards.

CDT fares no bettter. Suppose (say) that $q = 0.5$, then CDT suggests — reasonably — that Bob should take box A: the money is equally likely to be there as in box B, and if it is, there's more of it. But in that case, Bob chooses $p = 1$. Conditional on his following this strategy, the money has probability $g(1)$ of being in box B; if Alice is even moderately good at predicting Bob's actions, $g(1) > 0.6$, the threshold at which CDT dictates that choosing B is preferable. So, having chosen his strategy, Bob will (if he regards Alice as reasonably accurate) predict his expected return to be close to zero.

2 Diachronic consistency

Why care if Bob makes a decision he then regrets? Even if he makes a decision that he predicts he'll regret?

Short answer: because if he regrets that decision, nothing stops him from changing it. (It's not like the original case we considered, where the strategy was announced to Alice and Bob had to follow through on it.) And if he changes his decision, what makes it a *decision* at all?

The longer version rests on this principle: that if some time an agent prefers a strategy of doing X at a later time to a strategy of doing Y at that later time, then *at* that later time, the agent prefers doing X to Y. Call this *diachronic consistency*: it is a close cousin of Arntzenius's *desire reflection principle.*

Why accept it? Arntzenius offers two reasons: that diachronically inconsistent agents are susceptible to Dutch books, and that to violate diachronic consistency is to refuse advice from your future self, who is an agent just like you but with more information. That sounds persuasive to me, but I have a different line of defence: if an agent actually chooses Y over X at the later time, what *makes it true* that he has previously decided to choose X-later rather than Y-later? He may *say* "I am deciding to choose $X \cdot Z$", but if he sincerely believes that he will actually choose Y over X at time t_2, can it really count as a decision?

On first sight, this may just be seen as undermining diachronic consistency: is an agent really deciding anything at the earlier time, rather than making a mere prediction about his behaviour when the real decision is made? But I don't think this is a sustainable position. Real-world decisions — applying for a job, going out for a coffee, even crossing the room to turn the light on — take place over a period of time and involve consequences that likewise unfold over time; in general, a realistic action is a composite of smaller actions, and there is no reason to expect a finest grain of individuation for actions (or at any rate for actions *qua* actions; actions *qua* physical processes may be another matter, but that seems to be a matter for particle physics, not decision theory.)

Someone might, I guess, nonetheless suppose that there really are "critical moments" in decision-making when decisions are really taken: it is at *those moments* (one might suppose) that an agent's preferences really have significance, and at other moments an agent's "preferences" are merely aspirations, hopes, predictions, commentaries, or whatever. This might be an attractive position if one regarded actions, and decision-theoretic preferences, as actually instantiated in some concrete psychological processes. However, following (Davidson 1973), (Lewis 1974), and (Dennett 1987), I wish to adopt a more 'interpretative' approach to ascriptions of beliefs and preferences: what makes it true that an agent has a certain credence function and a certain utility function is that those functions best describe their dispositions to action on the assumption that their behaviour is rational, and what makes it true that they are rational agents is that some such functions describe their dispositions to action reasonably well.

From this perspective, it is difficult to see what pattern of behaviour could justify treating a physical system as an agent who is systematically violating diachronic consistency but is otherwise rational, rather than as a system which cannot be treated as rational at all (a system to which the Intentional Stance cannot usefully be adopted, in Dennett's (1987) terminology). To be sure, certain specific behaviours — deciding each morning to give up alcohol, only to succumb that same evening, for instance — might naturally be thought of as violations of diachronic consistency. But (leaving aside the fact that such behaviour is typically considered a paradigm of irrationality) it seems to me plausible that it can only be understood as behaviour at all when seen against an overall pattern of rationality (see (Dennett 1987, 103–116) and Davidson (1973) for further discussion of this point).

I hope that either these arguments, or Arntzenius's, at least make it plausible that diachronic consistency is a reasonable constraint on rational action. At any rate, for the rest of this paper I shall take it as read.

What does a diachronically consistent decision theory look like? It should respect the basic insight of causal decision theory: that an agent should choose that strategy on which his rewards are maximised *given his actual credence function.* But it also had better not be the case that that rule is self-undermining, so that when an agent updates his credence function by conditionalisation on that strategy, he now prefers some other strategy.

What does this entail for Bob? EDT said: choose p to maximise $U_B(p, g(p))$. CDT said: choose p to maximise $U_B(p, q)$ for whatever q is your prior credence in Alice's choice. DDT — *diachronically consistent decision theory*[4] — says: do what CDT says, but make sure that your credence q satisfies $q = g(p)$.

DDT looks somewhat unfamiliar. We are used to thinking of decision theory as being a rule for choosing preferences for given credences; we are less used to simultaneous constraints on credence and preference. But the idea that rationality constrains one's credence function is not new. After all, why require a credence function to satisfy the probability axioms in the first place? A simple answer is that if it doesn't, the agent will be committed to accepting bets which always lose money, and that this is irrational; a better answer comes through decision-theoretic representation theorems like those of Savage (1972), Jeffrey (1983) or Joyce (1999), which establish that agents satisfying certain qualitative constraints of rationality must act as if they maximised some utility function with respect to some credence function conforming the probability axioms. Similarly, diachronic Dutch book arguments Lewis (1997) are widely taken to establish that since having a credence function which updates other than through Bayesian conditionalisation commits one to losing money, it is irrational to have such a credence function. (And it is possible to prove a representation theorem to the same effect analogous to Savage's representation theorem provided one takes diachronic consistency as an assumption; see the appendix of Wallace (2010) for such a theorem.)

In normal circumstances, the constraints that rationality places on credence are fairly mild: any function which synchronically satisfies the probability calculus and diachronically conforms to Bayesian updating will do. Similarly, rationality places essentially no constraints on an agent's utility function. But where an agent regards his own choices as affecting the probabilities of various outcomes, the constraints are more severe: only certain choices of credence and utility function will give rise to diachronically consistent behaviour, so if diachronic consistency is a prerequisite for rationality, only certain choices of credence and utility function will count as rational. And since our evidence for ascribing given credence and utility functions to an agent comes from the assumption that their behaviour is rational or approximately so, it ought to follow that we will never have reason to ascribe non-diachronically-consistent credences and utilities to any agent. Indeed, if rationality is to be understood in interpretative terms, there simply cannot be agents whose credences and utilities are not diachronically consistent.

3 Bob's choice again

Recall Bob's situation: he has available to him a continuous one-parameter family of actions parametrised by p ($0 \leq p \leq 1$): the p-strategy involves choosing

[4] Conveniently, this could also stand for Arntzenius's *deliberational decision theory.*

box A with probability p and box B otherwise. So the expected utility of the strategy, given q, is

$$U_B(p,q) = 3p(1-q) + 2q(1-p). \quad (4)$$

Bob's credence on Alice putting the money in box B, conditional on his choosing strategy p, is $g(p)$. According to DDT, q and p are jointly determined by the requirement that p maximises $U_B(p,q)$ when $q = g(p)$.

$U_B(p,q)$ is linear in p, so there are three possibilities for its maxima:

1. the rate of change of U_B with respect to p is positive, so the maximum value occurs at $p = 1$;

2. the rate of change of U_B with respect to p is negative, so the maximum value occurs at $p = 0$;

3. the rate of change of U_B with respect to p is zero, so that all choices of p give the same expected utility.

Now $\frac{\partial U_B}{\partial p}(p,q) = 3 - 5q$, so the maxima are

1. $p = 1$ (if $q < 0.6$)

2. $p = 0$ (if $q > 0.6$)

3. All values of p (if $q = 0.6$ exactly).

The first is consistent only if $g(1) < 0.6$ (that is, only if Bob believes Alice to be a pretty poor predictor, little better than chance); the second only if $g(0) > 0.6$ (that is, only if Bob believes he's managed to fake out Alice, and she's rather more likely than not to guess wrong). The third, though, will be consistent provided there is some value of p such that $g(p) = 0.6$. And on the assumption that $g(1) < 0.6$ and $g(0) > 0.6$, this will be the case whenever g is continuous — an assumption which seems reasonable if only on physical grounds Arntzenius (2007).

If this third case holds, diachronic consistency commits Bob to choose that value of p such that conditional on his choice he predicts Alice to have probability 0.6 of picking box B. Notice that this doesn't mean Alice is adopting the *strategy* of picking box B with probability 0.6. It just means that Bob, who is predicting Alice's behaviour just as she is predicting his, gives credence 0.6 to her picking box B (conditional on his choosing p).

In any event, we know that the value of $U(p,q)$ in the third case is independent of p: it works out as 1.2, rather lower than the 1.5 that Bob could achieve if he was able to commit to the strategy of picking box B with probability $0.5 + \epsilon$, and slightly below the 1.25 attainable by EDT in the particular case where $g(p) = p$. I take both of these to be symptoms of the same underlying fact: that it is sometimes rational to constrain the choices of your future self.

Applying the same analysis to Alice: she has some credence $h(q)$ that Bob picks box A, conditional on her following a q-strategy. DDT requires that her credence p in Bob's choosing box A, and her choice q of strategy, both (a) maximise $U_A(p,q) = pq + (1-p)(1-q)$ and (b) satisfy $p = h(q)$. Since $\frac{\partial U_A}{\partial q} = 2p - 1$, there are again three possibilities for the maxima of U_A:

1. $q = 0$ (if $p < 0.5$)
2. $q = 0$ (if $q > 0.5$)
3. All values of p (if $q = 0.5$ exactly).

The first is consistent only if $h(1) < 0.5$; the second only if $h(0) > 0.5$; that is, they are consistent only if Alice believes Bob to be worse than chance at predicting her actions. If she is more charitable than this, her optimal strategy is that choice of q for which $h(q) = 0.5$, and again this is guaranteed to exist provided h is continuous. Her expected winnings, again, are independent of h: she expects to win 0.5 dollars.

4 Generalising the game

A more general version of Alice and Bob's game looks (from Bob's perspective) as follows. Bob chooses between options 1 and 2 (boxes A and B in the previous case); Alice predicts which choice he makes (with Alice's prediction and Bob's choice being causally independent). Bob then receives a payoff dependent on his choice and Alice's prediction, so that the general payoff matrix is

	Choose 1	Choose 2
Predict 1	x_{11}	x_{12}
Predict 2	x_{21}	x_{22}

For instance, the particular game we have been considering so far has payoff matrix

	Choose 1	Choose 2
Predict 1	0	2
Predict 2	3	0

(note that the table tracks only Bob's payoffs; from here on, the details of Alice's remuneration don't matter, provided that she has some incentive to make correct predictions).

Our analysis generalises straightforwardly. If Bob has credence q in Alice predicting choice 1, and in fact his strategy is to predict choice 1 with probability p, his expected utility is

$$U(p,q) \quad = \quad pqx_{11} + q(1-p)x_{12} + p(1-q)x_{21} + (1-q)(1-p)x_{22}. \quad (5)$$

If we define $\delta_1 = x_{11} - x_{22}$ and $\delta_2 = x_{21} - x_{22}$ (so that δ_i is the difference in payoff for Bob, conditional on him choosing i, between Alice being right and being wrong in her prediction) we have

$$\frac{\partial U}{\partial p}(p,q) = q\delta_1 - (1-q)\delta_2 = q(\delta_1 + \delta_2) - \delta_2. \quad (6)$$

If this is positive, it is maximised by $p = 1$; if negative, by $p = 0$; if zero, all choices of p give the same reward.

DDT requires that p and q be such that p maximises $U_A(p,q)$ and $q = g(p)$, where $g(p)$ is Bob's credence that Alice picks option 1 given that his strategy is to pick option 1 with probability p. So as before, there are three possibilities:

1. $g(1)(\delta_1 + \delta_2) > \delta_2$ and $p = 1$, in which case the expected utility is

$$U = g(1)x_{11} + (1 - g(1))x_{21} = x_{21} + g(1)(x_{11} - x_{21}) \quad (7)$$

2. $g(0)(\delta_1 + \delta_2) < \delta_2$ and $p = 0$, in which case the expected utility is

$$U = g(0)x_{12} + (1 - g(0))x_{22} = x_{12} + (1 - g(0))(x_{22} - x_{12}) \quad (8)$$

3. p satisfies $g(p) = q_f = \delta_2/(\delta_1 + \delta_2)$, in which case the expected utility is

$$U = \frac{x_{11}x_{22} - x_{21}x_{12}}{\delta_1 + \delta_2} = x_{21} + q_f(x_{11} - x_{21}) = x_{12} + (1-q_f)(x_{22} - x_{12}). \quad (9)$$

At least one of these conditions is guaranteed to hold provided g is continuous: if $(\delta_1 + \delta_2)g(1) < \delta_2 < (\delta_1 + \delta + 2)g(0)$, then $(\delta_1 + \delta_2)g(p) - \delta_2$ (regarded as a function of p) must cross over from positive to negative somewhere between $p = 0$ and $p = 1$.

In nontrivial examples of these prediction games, at least one of δ_1 and δ_2 must be nonzero: if both are zero, Bob can ignore Alice's predictions altogether. We can further distinguish three classes of these games:

1. *Co-operative games*, where δ_1 and δ_2 are both nonnegative. In these games, Alice and Bob both benefit from Alice guessing what Bob is going to do.

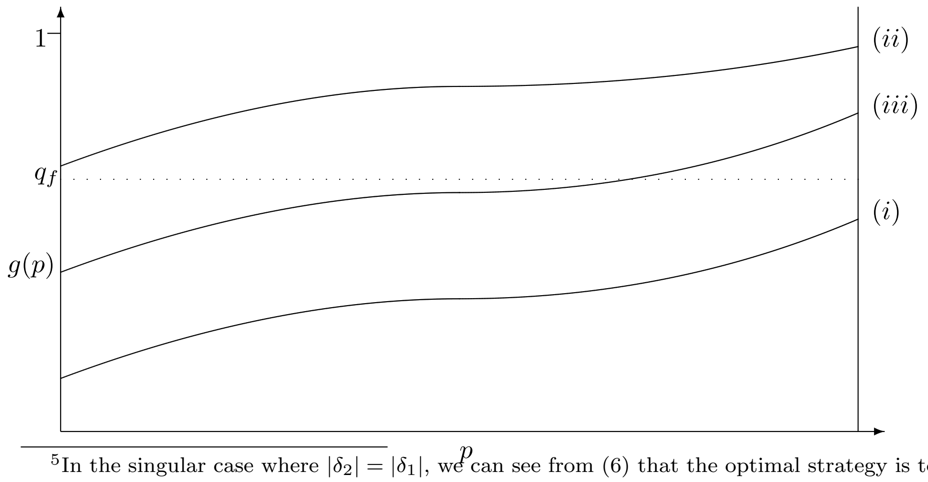
2. *Adversarial games*, where δ_1 and δ_2 are both negative. Here, Bob benefits from faking out Alice: if she guesses wrong, he does better than if she guesses right.

3. *Mixed games*, where one of δ_1 and δ_2 is negative and the other isn't.

We should expect that when $\delta_i > 0$, in general it's diachronically consistent for Bob to make choice i with certainty: he will predict (assuming Alice is any good) that conditional on that choice Alice is fairly likely to predict that he makes that choice, and he will be happy with her making that prediction. Conversely, in general where $\delta_i < 0$, choosing i with certainty will be diachronically inconsistent, since Bob's prediction conditional on that choice is that Alice will predict that he makes that choice, whereupon it becomes rational for him to change it. We will find that this expectation is borne out.

Recall that $q_f = \delta_2/(\delta_1 + \delta_2)$. In a mixed game, either $q_f > 1$ (if $|\delta_2| > |\delta_1|$) or $q_f < 0$ (if $|\delta_2| < |\delta_1|$). So there is exactly one diachronically consistent strategy, and it is either "take choice 1 with certainty" or "take choice 2 with certainty"; furthermore, the optimal strategy does not depend at all on Bob's assessment of Alice's predictive ability.[5]

In a co-operative or adversarial game, $0 \leq q_f \leq 1$. On the assumption that Alice's predictions are correlated at all with Bob's actions, if $p' > p$ then $g(p') > g(p)$. There are then three possibilities for the function $g(p)$.



[5]In the singular case where $|\delta_2| = |\delta_1|$, we can see from (6) that the optimal strategy is to make choice i iff $\delta_i > 0$.

(i) $g(p) < q_f$ for all values of p.

(ii) $g(p) > q_f$ for all values of p.

(iii) $g(0) < q_f$ and $g(1) > q_f$, so that (given that g is continuous) there is some value p_f of p such that $g(p_f) = q_f$.

In case (i), if the game is cooperative then the only diachronically consistent strategy is for Bob to choose option 2 always. This reflects the fact that (a) Bob is willing to choose either 1 or 2 provided he has confidence in Alice predicting that choice with reasonable accuracy — but he is insufficiently confident of her ability to predict his taking option 1 (his threshold is that $g(1)$ must be at least q_f). So he is left with option 2 as the only choice. By similar reasoning, if the game is adversarial then the only diachronically consistent strategy for Bob in case (i) is to choose option 1 always.

Case (ii) follows essentially the same logic: in cooperative games Bob should choose option 1 always, in adversial games he should choose option 2 always.

In case (iii), if the game is adversarial, the only diachronically consistent strategy is for Bob to choose option 1 with probability p_f, and to have credence q_f in Alice predicting that he takes option 1. But if it is cooperative, there are three diachronically consistent strategies: choose 1 always, choose 2 always, or choose 1 with probability p_f.

5 Examples of mixed and adversarial games

- Newcomb's problem is a mixed game. Here (recall), Bob chooses whether to take box A and box B, or only to take box A. Box B always contains a thousand dollars. Alice predicts whether Bob chooses both boxes; if she predicts that he does, she leaves box A empty; if she predicts that he does not, she puts a million dollars into box B. The payoff matrix (assuming utility to be linear in dollars, and choosing 1000$ =1 util) is

	Choose 1	Choose 2
Predict 1	0	1
Predict 2	1000	1001

$\delta_1 = -1$ and $\delta_2 = +1$, so the game is indeed mixed: Bob does better by taking both boxes than by taking one, irrespective of whether or not Alice predicts that he does so. The only diachronically consistent solution is to take both boxes with certainty, in accordance with the predictions of causal decision theory. From the perspective of this paper, the reason that causal decision theory suffices for mixed games is that the preferred strategy for them is independent of any considerations about how accurate the predictor is.

- The game with which I began this article is adversarial: whatever Alice predicts, Bob stands to gain from making the opposite choice. As we found before, Bob's optimal strategy depends upon how much credit he gives Alice as a predictor. If he believes her to be sufficiently inaccurate, his best strategy will be to choose either 1 always or 2 always and take a chance on her getting it wrong; if she is reasonably accurate, he will do better to adopt a mixed strategy.

- Andy Egan's "Newcomb's Firebomb" example Egan (2007) is adversarial. In this example, there are again two boxes, and Bob's choice is to take only box A or to take both. However, unlike Newcomb's original problem, he knows the contents of the boxes: box A definitely contains a million dollars and box B definitely contains a thousand dollars. If Alice predicts that Bob will take both boxes, though, she wires up an undetectable firebomb which will destroy the million dollars iff Bob takes the thousand dollars. So the payoff matrix is

	Choose 1	Choose 2
Predict 1	0	1000
Predict 2	1001	1000

Here $\delta_1 = -1000$ and $\delta_2 = -1$: whatever Alice guesses, Bob does better by confounding her guess than by confirming it. q_f is 1/1001, and so unless Alice is a truly superb predictor, $g(p) > q_f$ for all p, and Bob's best strategy is to take only one box. At the other extreme, $g(p) < q_f$ for all p only if Bob believes Alice will almost certainly predict that he takes one box no matter what he actually does; only in this unlikely circumstance will his best strategy be to take both boxes.

The intermediate case arises only if Bob believes Alice to be an extremely accurate predictor: one who is more than 99.9% likely to have predicted that he takes one box given that he in fact does but who is less likely than this to have made that prediction given that he actually takes both. In this case, both the extreme strategies are diachronically inconsistent:

conditional on him taking both boxes he believes he is likely to be incinderating a fortune; conditional on him taking only one box he believes he is leaving money on the table. The consistent strategy is to choose p such that $g(p) = 1/1001$: in other words, to choose a strategy conditonal on which Alice will almost but not quite certainly put the million dollars in the box.

There is something a little counter-intuitive about this strategy, to be sure, since its expected winnings are (from 9) exactly 1 million dollars, and Bob stood to do just as well as that by adopting the strategy of taking only one box with certainty (this, incidentally, is the strategy which Egan regards as rationally required). It is somewhat reassuring to note that this exact equality depends on the fact that two payoffs in the matrix are exactly equal. From equations (8) and (9), the difference between the expected utilities of the always-take-one-box strategy and the mixed strategy (worked out in both cases with conditional probabilities) is

$$EU_{\text{always 1}} - EU_{\text{mixed}} = (q_f - g(0))(x_{22} - x_{12}). \tag{10}$$

So if the amount of money in box A was just slightly less conditional on Alice predicting that Bob would take one box, the mixed strategy pays better than the alternative: Bob's gamble has a nonzero expected payoff relative to playing it safe. And conversely, if Alice slightly decreases the amount of money in box A when she predicts that Bob takes both boxes, then Bob's expected returns would be better if he played it safe than if he took a gamble — but this is just another instance of the "if you're so smart, why aren't you rich" objection to the two-box strategy in Newcomb's problem. Bob would indeed be better off if he could *commit* to the strategy of playing it safe, but he cannot so commit, because conditional on the assumption that he commits he stands to make a profit by changing his mind. (Because the benefit is only probabilistic, the case is less intuitively obvious than Newcomb's problem, but the basic structure is the same.)

- Egan's "psychopath" case can also be thought of as adversarial. In that case,

  > Paul is deciding whether to press the "kill all psychopaths" button. It would, he thinks, be much better to live in a world with no psychopaths. Unfortunately, Paul is quite confident that only a psychopath would press such a button. Paul very strongly prefers living in a world *with* psychopaths to dying. Should Paul press the button? (Egan 2007, p.97)

  (Egan's position is that (i) CDT advocates pressing the button (assuming that Paul is fairly confident that he's not a psychopath); (ii) it's intuitively obvious that the right strategy is not pressing the button.)

  No actual person is trying to predict whether Paul will press the button, but his credence in being a psychopath is higher conditional on him pressing the button than on him not doing so, and that's all that's needed to apply the adversarial-game analysis. The payoff matrix looks like this (assuming that Paul puts utility, say, +10 on wiping out all the psychopaths (whilst surviving himself) and −1000 on dying):

	Choose 1	Choose 2
Predict 1	-1000	0
Predict 2	10	0

so that $\delta_1 = -1000$, $\delta_2 = -10$, and $q_f = 1/101$.

In this case, though, we're not assuming that Paul thinks whether he's a psychopath depends on what (potentially mixed) strategy he chooses, but only on the end result: what button he presses, that is. Suppose, say, that conditional on his pressing the button he has credence 0.9 in being a psychopath, and that conditional on his not pressing it he has credence 0.1. (After all, merely *taking seriously* the possibility of wiping out all the psychopaths should give Paul pause in his assessment of his own mental state, whether or not he actually goes through with it!) By definition, his credence in pressing the button, conditional on adopting a strategy of pressing the button with probability p, is just p. So his credence in being a psychopath conditional on adopting that strategy is

$$g(p) = 0.9p + 0.1(1-p) = 0.1 + 0.8p \tag{11}$$

so that $g(p) > q_f$ always. The diachronically consistent strategy here, then, accords with Egan's intuitions: definitely don't press the button.

Things change interestingly, though, if we shift the numbers a little. Let's suppose that Paul is quite sure that a psychopath *would* press the button, so that his credence in being a psychopath conditional on not pressing the button is zero. Now we have

$$g(p) = 0.9p \tag{12}$$

so that $g(p) = q_f$ for $p = 10/909 \simeq 0.011$. And now Paul's only diachronically consistent strategy is to press the button with probability 10/909.

Interestingly, having pressed the button, Paul's credence in having pressed it shifts to 1: this being the case, his credence in being a psychopath shifts to 0.9, and he believes that he has very probably signed his own death warrant. Conversely, once the chance to press the button has definitely gone (for simplicity, suppose the device has a time limit for use), Paul's credence in being a psychopath drops to zero, and he believes now that pressing the button would have done him no harm. So he is in the rather unfortunate position of regretting his action, whatever it actually was.

At first sight this might look like a violation of diachronic consistency, but it isn't. Diachronic consistency, as I've formulated it, enforces a consistency between agent's preferences at different times *when those preferences are understood as dispositions to choose one action over another.* Having pressed the button, though, Paul has no further actions to choose. His regret, however tragic from a humanitarian perspective, is decision-theoretically irrelevant. Furthermore, *any* plausible decision theory must predict that Paul will regret his choice, whatever it is. We can see this further by supposing that there is a very small chance (one in a million, say) that Paul will be offered a chance to change his mind after pressing or not pressing the button. Since the chance is so small, its presence won't appreciably affect his original strategy, but we can predict that he will always change his mind.[6]

- "Death at Damascus" is adversarial: recall that Death predicts if his victim is hiding in Damascus (option A) or Samarah (option B), and waits for him there. Assuming that Death is a very good predictor, the optimal strategy is to choose at random. (If Death's powers go beyond prediction into actual prophecy, I'm less sure if the analysis applies.)

6 What to do in cooperative games? Choosing between diachronically consistent alternatives

A simple example of a cooperative game would be: Bob chooses box A or box B; Alice has placed either three dollars in box A or two dollars in box B; unlike our original example, Alice is rewarded whenever she chooses the box that Bob actually selects. For this game, the payoff matrix (for Bob) is

	Choose 1	Choose 2
Predict 1	3	0
Predict 2	0	2

[6] More accurately, we can make this prediction conditional on our believing that Paul is rational. If Paul pressed the button, he's probably a psychopath, so this belief is pretty questionable!

so that $\delta_1 = 3$ and $\delta_2 = 2$: the game is indeed cooperative, with $q_f = 0.4$.

If Bob regards Alice as sufficiently biased towards him taking box A ($g(p) > 0.4$ always) or box B ($g(p) < 0.4$ always) then his best option is to take box A, or box B, with certainty, depending on her bias. But if he regards Alice as at all skilled as a predictor, then there will be some value p_f for p such that $g(p) = 0.4$, and then all three strategies — box A always, box B always, and box A with probability p_f — are diachronically consistent.

This raises the obvious question: what should Bob do in this case? A natural, and intuitively reasonable, rule would be: take that strategy which has the largest expected payoff.[7] (Call this the maxicon strategy: choose that diachronically **con**sistent strategy which **max**imises expected utility.)

But maxicon should be rejected. For one thing, there are straight counterexamples. Suppose we modify the Alice/Bob game a little. Alice still chooses to put either three dollars in box A or two dollars in box B (recall that this isn't her own money), but now she receives three dollars if she puts money in box B and Bob takes it, and two dollars in she puts money in box A and Bob takes it. So *her* payoff matrix is

	Choose 1	Choose 2
Predict 1	2	0
Predict 2	0	3

Let's assume that Alice and Bob both believe the other to be a pretty good predictor of their actions. For Bob, this means $g(1) \sim 1$ and $g(0) \sim 0$. There are three diachronically consistent strategies for him, but the one with the biggest payoff is 'box A always', which has an expected payoff of ~ 3. Alice's situation is the mirror of Bob's: she presumably has some credence function $h(q)$ which is her credence in Bob choosing box A given that she follows a q-strategy, and (given the assumption that she believes Bob to be a good predictor of her actions) $h(0) \sim 0$ and $h(1) \sim 1$. So she too has three diachronically consistent strategies, and the one which pays best is 'box B always'. And so, if Alice and Bob both follow a maxicon principle, she will choose B and he will choose A, and both will get nothing.

Now, we're used to Prisoner's Dilemma situations[8] where both players miss out on a mutually beneficial strategy because each benefits by deviating from the strategy. But this game isn't like that: conditional on Alice choosing box

[7] This rule is espoused by Arntzenius (2007): "[A]n ideally rational person is always in an equilibrium state such that there is no other equilibrium state which has higher utility."

[8] Prisoner's Dilemma is another mixed game, incidentally.

A, both Alice *and* Bob do better by choosing box A. So if Bob really believes that (i) Alice is a good predictor of his actions; (ii) Alice believes him to be a good predictor of her actions; (iii) Alice follows maxicon, then it is irrational for him too to follow maxicon.

This points us towards the positive story as to what's wrong with maxicon. Diachronic consistency says that rational agents cannot but have credences and preferences such that (a) they prefer that action which maximises utility given their credences; (b) their credence in anything conditional on their carrying out their actions are the same as their unconditional credence in it. Maxicon exhorts agents to choose to be that rational agent — that is, to choose their credences and preferences so as to be that rational agent — whose expected utility on performing their preferred strategy is greater than any other such agent.

But if agents really are allowed to adjust their credence functions so as to maximise utility, the sky's the limit. Currently, my credence in Bill Gates giving me a billion dollars, conditional on my writing and asking for it, is (alas) pretty low; that being the case, I'm not planning to write. If I change that credence to something close to one, writing to Gates will have a pretty impressive expected utility — but that doesn't make it rational to make that change, nor to write the letter.

Diachronic decision theory, just like causal decision theory, tells agents to choose that strategy which maximises utility given the agent's actual credence. It additionally requires that agents' credences satisfy constraints somewhat more stringent than just the probability calculus and the Bayesian update rule, but it does not pretend to be a rule for choosing which of the various credence s satisfying those constraints should actually be selected. The story as to which credences an agent has at some time will, as usual, be determined by a combination of their initial beliefs and various ways of updating in the face of evidence.

For instance, suppose (reverting to the original cooperative game) that Alice is a very good predictor and Bob resolves, before she makes her prediction, that he will take box A. Then he will predict that she predicts he takes box A, and so it makes sense for him to follow through on that resolution. Indeed, from the perspective of diachronic consistency it is only because (being rational) he will follow through on the resolution that it *is* a resolution, rather than an empty utterance of "I resolve that …". On the other hand, if for some reason (be it a moment of madness or a lucrative side bet) Bob resolves that he will take box B, it makes sense for him to follow through on *that* resolution too. Indeed, if communication is possible between Alice and Bob, it is in both of their interests to agree a strategy and then stick to that strategy. (This is what separates cooperative games from adversarial or mixed games: in the Prisoner's Dilemma, for instance, it would be in Alice and Bob's interests to agree a strategy, but since it would be irrational for them to follow through on that agreement, they won't be able to make it in the first place.)

In one sense, then, Bob *can* be said to be choosing his strategy by maxicon: any provisional decision he makes as to which box to take will be self-reinforcing, and so it makes sense to provisionally choose that decision which leads to the best expected return. (Again, note the contrast with Newcomb's problem and similar adversarial games: there, such resolutions are self-undermining, and so rational agents will be unable to form them in the first place). But once the prediction has been made, Bob needs to choose that strategy which maximises his utility given his actual credence function. Since he is (ex hypothesi) diachronically consistent, conditionalising on that choice will not change his credence, but he is not licenced to change both credence and preference so as to enter a different diachronically consistent state.

7 What are mixed strategies?

Taking stock: in the two-option games I have been considering so far, we have seen that it is always possible for an agent to be diachronically consistent. And in fact this generalises to games where Bob chooses one of N options, provided we continue to assume that Bob's conditional credences in Alice's choices are continuous functions of his chosen strategy.[9]

However, all of this does assume that agents have access to mixed strategies, and we need to revisit the question of what these are. In the stylised cases we have considered so far, we can at least imagine that Bob is consulting a real random-number generator. But in most circumstances, this doesn't seem realistic.

I wish to defend the claim, though, that mixed strategies really are available to (idealized) agents, even when they don't have some handy source of objective randomness to hand. This suggestion has been considered, and criticized, by Arntzenius (2007), who writes

> The natural alternative view [to the one Arntzenius proposes] is that a mixed decision is a decision to perform certain acts A_i with certain probabilities p_i. But what is it to decide to perform certain acts $A_i$with certain probabilities p_i? A natural suggestion would be that one does this just in case one has a chance device at one's disposal, where one can delegate responsibility of which act is to be performed to this chance device, and one can set the chances with which this

[9]Let X_{ij} be the elements of the payoff matrix; then if Bob has credence q_i in Alice predicting that he takes option i, the strategy of taking option i with probability p_i has utility $U(\mathbf{p},\mathbf{q}) = \mathbf{q}\cdot\mathbf{X}\cdot\mathbf{p}$ (in vector notation). If $g_i(\mathbf{p})$ is Bob's credence in Alice making choice i conditional on his following strategy $\mathbf{p}$, we can define the set-valued function S by

$$S(\mathbf{p}) = \{\mathbf{y} : \mathbf{y}\text{is a maximum of } U(\mathbf{y},\mathbf{g}(\mathbf{p})\}.$$

$\mathbf{p}$ represents a diachronically consistent strategy provided that $\mathbf{p} \in S(\mathbf{p})$. By the Kakutani fixed point theorem (Kakutani 1941), there will be such a point provided that the graph of S is closed and that $S(\mathbf{p})$ is convex and non-empty for every $\mathbf{p}$. Since $U(\mathbf{p},\mathbf{q})$ is linear in p, convexity of $S(p)$ is trivial; since for each $\mathbf{q}$ $U(\mathbf{p},\mathbf{q})$ is a continuous function on a compact set, $S(\mathbf{p})$ is never empty. (Conversely, in the unphysical case where there are infinitely many possibilities, diachronically consistent strategies need not exist.)

Let Y be the complement of the graph of S and suppose that $(\mathbf{p},\mathbf{y}) \in Y$. Then there exists some $\mathbf{z}$ such that $(\mathbf{z}-\mathbf{y})\cdot\mathbf{X}\cdot\mathbf{g}(\mathbf{p}) > 0$. Since this is a continuous function of $\mathbf{z}$ and $\mathbf{p}$, there must be some neighborhood of $(\mathbf{p},\mathbf{y})$ which is contained in Y; hence, Y is open and so the graph of S is closed.

> chance device will act one one's behalf to values p_i of one's choosing. However, in the first place we are hardly ever in a situation in which we can perform such actions. (It is not as if one has such a chance device stored away in some convenient part of one's brain.) In the second place, even if we did it would amount to a different decision situation, namely one in which we have an uncountable infinity of pure acts that we can perform, the acts being the possible ways we have of setting the chance 'dials' of the chance device.

Now, from my perspective Arntzenius's second objection is unproblematic. I agree fully that the mixed strategies, conceived my way, are indeed just "an uncountable infinity of pure acts"; indeed, I have relied on it, and assumed that Bob has credences in Alice's behaviour conditional on each element of that uncountable infinity. Arntzenius's first objection has rather more force.

... And yet, consider: what would you *actually* do if you had, say, Bob's role in Alice and Bob's original, adversarial game? (Notice that there's nothing at all science-fictional about that game: feel free to take time out of the paper, borrow a colleague, and play a few rounds if you don't know the answer). I'm pretty sure I know what I'd do: choose at random. I'm pretty sure I have the capacity to do so: certainly I frequently do something that feels subjectively like choosing at random. If I really want to be careful (if there's a lot riding on the game, say) I might resort to eenie-meenie-minie-moe or some similar prosthetic, but in general, or if there's no time to do otherwise, I'll just pick on a whim.

And what's more, I believe — indeed, have just spent the better part of a paper arguing — that picking at random is the *rational* thing to do in Bob's place (at least given that Alice is a moderately good predictor). So if I've in fact chosen at random, and if choosing at random is the rational thing to do, it seems unreasonable to deny that I *decided* to act randomly, so that my decision really was to take box A with probability 0.5 and box B with probability 0.5.

And now consider what you'd do in Alice's situation. If it were me, I'd *probably* put the money in box B, but I *might* put it in box A. I'd do so because (a) you're reasonably likely to predict what I do, if for no other reason than the assumption that we're both reasonably intelligent and rational people, liable to behave the same way in situations like this, and (b) on that prediction, you'll pick a box at random, which is my best chance of guessing correctly. And it seems to me pretty meaningful to resolve to act in that fashion. Will the actual odds of me choosing box A be 2 in 5? Probably not; but perfect rationality is hard.

One more case, adjusting the settings on the intuition pump: suppose now that you're in Alice's shoes once more, but that now Bob stands to win fifty dollars if he correctly chooses box A, and only one dollar if he correctly chooses box B. In that circumstance, he's going to choose box A if he thinks there's any significant chance there's money in it. But if he's 100% certain there's no money in box A, he'll pick box B. Either way, you lose. So what you ought to do is *almost certainly* put money in box B. Again, it doesn't seem unreasonable to suppose that someone could resolve to act that way.

This understanding of chance acts is not *that* far from Arntzenius's own preferred understanding:

> My tentative view is that to make a certain mixed decision is just to have certain credences in one's acts at the end of a rational deliberation.(Arntzenius 2007)

The difference, though, is that Arntzenius requires that the process which selects which action to perform is "a rational deliberation". I don't see the need for this rider and, indeed, in general I need to reject it: for what is a rational deliberation other than one which results in a rational action? But where mixed acts are the only diachronically consistent actions available, no rational deliberation will lead to a definite decision to make one choice rather than another. I'd like to say, rather, that the internal psychological process by which Bob makes his random choice is not concern of decision theory: it forms what Dennett calls "sub-personal cognitive psychology" (Dennett 1987, chapter 3).

Now, to be sure: it's a little difficult to believe that realistic agents have such a finely graded set of actions available to them that they can, say, choose to take option A with probability 357/589. My position is just that

- *ideal* rational agents must be able to perform mixed acts with any probability, because ideal rational agents must be diachronically consistent, and this requires access to mixed acts.
- Real agents, who actually can decide to pick an option at random, can approximate ideal rational agents at least reasonably well.

No real agent truly has access to the continuous infinity of mixed acts. No real agent is more than approximately rational. So what? Rationality is hard, and perfect rationality is an ideal, not an actually-achievable state. We knew that already.

References

Arntzenius, F. (2007). No regrets; or: Edith Piaf revamps decision theory. Available online at philsci-archive.pitt.edu.

Davidson, D. (1973). Radical interpretation. *Dialectica 27*, 313–328.

Dennett, D. C. (1987). *The intentional stance.* Cambridge, Mass.: MIT Press.

Egan, A. (2007). Some counterexamples to causal decision theory. *Philosophical Review 116*, 93–114.

Jeffrey, R. C. (1983). *The Logic of Decision (2nd edition).* Chicago: University of Chicago Press.

Joyce, J. N. (1999). *The Foundations of Causal Decision Theory.* Cambridge: Cambridge University Press.

Kakutani, S. (1941). A generalization of brouwers fixed point theorem. *Duke Mathematical Journal 8*, 457–459.

Lewis, D. (1974). Radical interpretation. *Synthese 23*, 331–44. Reprinted in David Lewis, *Philosophical Papers*, Volume I (Oxford University Press, Oxford, 1983).

Lewis, D. (1981). Causal decision theory. *Australasian Journal of Philosophy 59*, 5–30. Reprinted in David Lewis, *Philosophical Papers*, Volume II (Oxford University Press, Oxford, 1986).

Lewis, D. (1997). Why conditionalize? In David Lewis, *Papers in Metaphysics and Epistemology* (Cambridge, 1999).

Savage, L. J. (1972). *The foundations of statistics* (2nd ed.). New York: Dover.

Wallace, D. (2010). *Everettian Quantum Mechanics.* Oxford: Oxford University Press.