Solving an Infinite Decision Problem

Brian Weatherson

ABSTRACT: Barrett and Artzenius posed a problem concerning infinite sequences of decisions. It appeared that the strategy of making the rational choice at each stage of the game was, in some circumstances, guaranteed to lead to lower returns than the strategy of making the irrational choice at each stage. This paper shows that there is only the appearance of paradox. The choices that Barrett and Artzenius were calling 'rational' cannot be economically justified, and so it is not surprising that someone who makes them ends up with sub-optimal returns. A solution to the more general problem they pose is also advanced.

KEYWORDS: Decision Theory, Dutch Book, Puzzles.

Barrett and Artzenius (1999) pose the following problem:

"One has an infinite stack of dollar bills with consecutive serial numbers: 1, 2, 3 etc. An agent, who starts with no money, is then offered the following choice, where *n* is equal to the total number of times that the choice has been offered so far:

1. Get one dollar bill off the top of the stack.

2. Get 2^{n+1} dollar bills off the top of the stack, but you must then return the bill with the smallest serial number that you currently have. The returned bill is immediately destroyed.

Suppose that the agent does not know how many times he will be offered this choice." (101)

Several puzzles arise. One puzzle is that option 2 seems to give the agent more money than option 1 at every stage, but if there are countably many stages then choosing option 1 every time leaves the agent with all the bills, and choosing option 2 every time leaves her with nothing. That isn't the puzzle with which Barrett and Artzenius are primarily concerned. Rather, they ponder why it is that option 2 could be the rational choice at every stage, yet a policy of choosing option 2 at every stage yields lower returns than a policy of choosing option 1. "One might worry that irrationality pays so well."[1] (102) Given all this, they ask "Not knowing ahead of time what choices … will be offered, what strategy should the agent follow at each step"? (102).

First Response: It pays to be far-sighted.

Let's assume, for a simplification, that the agent knows the choices will be offered infinitely often. Then it seems, paradoxically, that his best strategy over all choices is to always choose option 1 (that gets all the cash), but his best choice on each individual occasion is to choose option 2 (that gets more cash on

[1] Of course irrationality can always pay well, if the irrational agent is lucky. It isn't a paradox that minimising *expected* returns can sometimes end up maximising *actual* returns.

each occasion). Again the individual and the collective are at odds. However, there is a simple argument that in these circumstances it is not best to choose option 2 on each occasion.

Betty would like to own the new Smashing Pumpkins CD; in fact, she would happily pay $15 for it were this the only way to obtain it. She is walking down a street in here town where there are two strange music shops. At the first shop she will come to she can buy the Smashing Pumpkins CD for $5. At the second shop she can get a copy of the CD for free, provided she has not already bought a copy. (The second store has ways of knowing who buys which CDs.) Betty knows this before she starts down the street.

It would, I take it, be irrational for Betty to buy the CD for $5 at the first store, even though that trade in isolation makes her better off. She would, ceteris paribus, prefer having the CD to having $5. But here other things aren't equal. Making that trade would prevent her from making an even better trade just down the road. As Keynes said, "It is not sensible to pay 25 for an investment of which you believe the prospective yield to justify a value of 30 if you also believe that the market will value it at 20 three months hence." (1936: 155) And it certainly isn't rational to pay $5 for an investment of which you believe the prospective yield to justify a value of $15 if you also believe that the market will value it at 0 three doors hence.

This kind of irrationality is displayed by the person who chooses option 2 at every stage knowing that the game will go on forever. Consider this redescription of the first choice she is offered: "You will be given the bill numbered 1, and you can either keep it, or exchange it for the bills numbered 2, 3 and 4." Option 1 is equivalent to keeping the bill, option 2 to making the exchange. The person who makes the exchange is buying the bills numbered 2, 3 and 4 at a considerable discount: 67% off the cover price. But she does this knowing full well that (a) in a short amount of time she could get a further discount and (b) that further offer is only valid if she refuses the trade now. As we saw in Betty's case, this is an irrational action.

Note that we don't need to appeal directly to the infinite sequence of choices to show that the choice is irrational. We do appeal to it indirectly. If the game isn't slated to continue indefinitely, then the agent does not know that she will receive the higher numbered bills further down the track. Note also that this argument doesn't rely on the agent planning to take option 1 at the later stages. Whether she plans to take option 1 or option 2, she is certain to receive those very bills while either (a) giving up nothing if she chooses option 1 or (b) receiving even more in exchange if she chooses option 2. So, whatever she plans to do in the future, there is a cheaper way to buy bills numbered 2, 3 and 4 than choosing option 2, and that way is foreclosed if she chooses option 2 now. So that option is irrational at the first stage. Because the game is slated to continue indefinitely, this reasoning is perfectly general, and hence shows that it is irrational to choose option 2 at each stage.

If it is not rational to choose option 2 at every stage in this game, it is certainly not rational to pay $1 for the privilege of being allowed to choose it. So the Dutch Book argument which Barrett and Artzenius propose against the agent who always chooses option 1 (p. 103) will fail when it is certain the game goes on forever.

Second Response: Choosing the highest limit

Assume, with some loss of generality, that the agent is given in advance the probability that the game will be stopped after each step. Say r_i is the probability that the game will be continue after the i'th round, given that it has made it that far. (So the probability of the getting to the fourth round is $r_1 r_2 r_3$.) The argument of the previous section showed that when r_n is one for all n, the right choice is option 1. We now suggest a general procedure for choosing a strategy given <r_1, r_2, …>.

Let n be an integer, and τ a strategy for choosing option 1 or option 2 at each stage. Define $S(n, \tau)$ to be the probability of ending up with the dollar bill with serial number n by following strategy τ. So if τ is α, the strategy of always choosing option 1, then $S(n+1, \tau)$ is $r_1 r_2 \ldots r_n$, or the probability that the game will get past the n'th round. Define $T(n, \tau)$ as the sum from 1 to n of $S(i, \tau)$. That is, $T(n, \tau)$ gives the expectation of the number of bills with serial numbers less than or equal to n with which an agent following strategy τ will end up with. The following proposal seems quite plausible: Strategy χ is better than strategy δ iff there is some integer k such that for all $j > k$, $T(\chi, j) > T(\delta, j)$.

Before defending this proposal, it is worthwhile to say a few words about the strategy, call it β, of always choosing option 2. To determine $S(n, \beta)$ we must make two calculations. First, we must work out the probability of bill n getting into the agent's collection in the first place. Second, we must work out the probability of that bill being subsequently destroyed. The first calculation is already done, it is $S(n, \alpha)$. The second calculation is a bit harder. Define a new function f from natural numbers to natural numbers. Intuitively, f measures how long it will take for you to receive bill number n following strategy β. So $f(73) = 7$ means that an agent following strategy β will receive the bill numbered 73 in the seventh round of the game, if it gets that far. In the scenario Barrett and Artzenius describe, we can formalise f as follows[2]:

[2] I have left some scope for generality here, because the precise number of dollars an agent receives under option 2 seems irrelevant to the puzzle. If we replaced '2^{n+1}' in their puzzle with, say, '3', there would still be the oddity that choosing the larger sum of money every time leads to an inferior result.

$$f(n) = \text{Int}(\text{Log}_2(n + 3.5)) - 1$$

The '3.5' here could be replaced by any value in (3, 4) without affecting the formula. ('Int' here is a function which takes reals to the largest integer less than or equal to that real; intuitively it deletes everything after the decimal point.) Given *f*, we can write the formula for S(*n*, β).

$$S(n, \beta) = S(f(n), \alpha) - S(n, \alpha)$$

The first term gives the probability the agent will receive bill *n*, the second term the probability that it will be destroyed. Note that if r_i equals 1 for all i, then S(*n*, β) equals zero for all *n*, and hence T(*n*, β) equals zero for all *n*. This accords with our earlier conclusion that α is a better strategy in this situation. If choosing strategy α has a finite expected return, that return is the limit as *n* tends to infinity of T(*n*, α). (Lemma One in the Appendix proves that for any strategy χ with a finite expected return, that return is the limit as *n* tends to infinity of T(*n*, χ).) In such a case it will always be better to choose β rather than α. Indeed, in any such case the best strategy of all will be β. This is proven as Lemma Two in the Appendix. On the other hand, if the probability that the game will have no end is positive, then α is a better strategy than β, as is shown in Lemma Three. (This of course is not to suggest that α is the best strategy of all in such a situation.) So it seems the proposal gives the right answers in the cases on which we have firm opinions, and gives us a reasonable way of working out the cases on which we have no firm opinions.

At this stage it might seem there is little which is paradoxical. We have a plausible mathematical model which recommends choosing option one in the case where the game will go on forever, and we have two arguments derived from considering finite cases for the legitimacy of this choice. There is, however, some paradox lurking in the background. To see this, consider how the game goes if the r_i are such that the following holds for all *j*:

$$\prod_{i=1}^{j} r_i = \frac{1}{2} + \frac{1}{2^{i+1}}$$

So the probability of the game having no end is ½, and the probability of the game lasting at least j rounds approaches that limit rather quickly as j rises. I leave as an exercise for the interested reader the following theorem. Let χ be any strategy other than β, and let δ be the strategy generated by replacing the first choice of option 1 in χ with a choice of option 2. In this game, δ is a better strategy than χ. The paradox is that α is better than β. So we have, loosely speaking, a kind of infinite intransitivity. Let's try to be rid of some of the looseness.

Let g be any function, partial or complete, from strategies to strategies. As usual, let $g^0(\chi)$ be χ, when it is defined, and $g^{k+1}(\chi)$ be $g(g^k(\chi))$ for all non-negative integers k. Say γ is the limit as k tends to infinity of $g^k(\chi)$ iff for all j there exists a k such that for all $i \geq k$, $g^i(\chi)$ agrees with γ on the first j moves. In the circumstances, this seems a natural enough definition of a limit. The following might look like a plausible transitivity principle: If for all positive k, $g^k(\chi)$ is better than $g^{k-1}(\chi)$, and γ is the limit as k tends to infinity of $g^k(\chi)$, then γ is better than χ.

If we accept my proposal about judging strategies, it turns out this transitivity principle is false. Let g be the function of replacing the first choice of option one with a choice of option two, and let χ be α, the strategy of always choosing option one. Given our definition of the r_i, all the conditions on the transitivity principle are met, but α is better than the limit as k tends to infinity of $g^k(\alpha)$. That is, α is better than β. So perhaps there is something paradoxical about infinite decision puzzles after all.

It might be thought that it is my proposal, not the puzzle case, which is problematic. After all, it might be thought that endorsing this principle of infinite transitivity is a requirement on any rule for evaluating strategies. We can show that thought is mistaken by looking at some much simpler cases. Consider this game, called the pool game[3].

We have a bank, a pool, and an agent. The bank starts with an infinite supply of dollar bills, and the pool starts empty. At each round, the agent has two options. If she chooses option one, one dollar gets added to the pool, she gets everything in the pool, and the game ends. If she chooses option two, one dollar gets added to the pool, and the game goes on. The first round takes ½ of a minute, the second round ¼ of a minute, the third round 1/8 of a minute, and so on, with the game ending after one minute. If the game ends without the agent taking what is in the pool, the pool returns to the bank. Say C_0 is the strategy of always taking option 2, and C_n is the strategy of taking option 1 for the first (and last) time on the *n*'th round. Let *g* be the function which takes C_n to C_{n+1} for all *n*. It is clear that for all positive *k*, $g^k(C_1)$ is better than $g^{k-1}(C_1)$, and C_0 is the limit as *k* tends to infinity of $g^k(C_1)$. And it is clear that C_1, which returns $1, is a better strategy than C_0. So the problem is with the infinite transitivity principle, not with my proposal for evaluating strategies.

Third Response: Don't Always Expect an Answer

As noted earlier, there is a loss of generality in assuming the agent knows the probability that the game will stop after each round. In the problem Barrett and Artzenius describe, the agent simply does not know how likely it is that the game will end at each stage. What should she do? There are two substantive problems to be addressed here before we can help the troubled agent. First, we must provide a theory of how uncertainty is to be represented in such cases. Secondly, we must provide a decision theory based on

[3] This game is a variant on the paradox of Hell discussed in Artzenius and McCarthy (1997).

these representations. Since each of these questions is a matter of some controversy, I will just state the positions which I take to be winning the war.

In the previous section we assumed that the agent's epistemic state could be properly represented by a single probability function. When the agent is not so certain about the conditions under which the game is played, the appropriate move is to represent her epistemic state by a set of probability functions, called her 'representor'. As above, we will equate these functions with sequences $<r_1, r_2, \ldots, r_n, \ldots>$, with $0 \leq r_i \leq 1$ for all i, with each r_i representing the conditional probability that the game would continue after the i'th round. If the agent was completely uncertain about the game conditions, her representor would consist of every possible sequence.

In fact no rational agent is completely this ignorant. Once the choice has been offered the first, say, 100 times, the agent should start to suspect that the game will be long-lasting. So it seems irrational to have any sequence in the representor in which r_{100} is very low. Even within these constraints, however, a representor could contain a wide range of sequences.

This way of representing uncertainty has become quite popular in recent years. (See, for a fair sample, Williams 1976, Levi 1980, Jeffrey 1983, Seidenfeld 1984, van Fraassen 1990 and Walley 1991.) There is considerably less consensus on how we should incorporate these representors into our decision theory. Note that since the elements of representors are probability functions, we still can talk about the expected utility of any choice (or in this case strategy) according to elements of the representor. Most theorists agree that when χ has a higher expected return than δ according to each element of the representor, χ should be chosen over δ. The problem is what to do when this is not the case, when χ does better than δ according to some elements and δ does better than χ according to others.

There are two broad classes of strategy for dealing with this problem. We can either try and write tie-breaker principles to decide whether χ or δ should be chosen, or we can say that in such cases neither option should be preferred to the other. Levi, Williams and Seidenfeld favour the first option; I'm inclined to favour the second. If we take the first option, then what to do will depend on which tie-breaker rule we establish. If we take the second, it will generally turn out that there is no fact of the matter as to which should be chosen. More precisely, choosing either χ or δ will be permissible, and neither choice is obligatory. It is not uncommon for decision theories to make these kinds of recommendations; orthodox expected utility theory says the same thing when the expected return of each choice is identical.

Given how indefinite the probabilities seem to be here, and given this rule for decision making with indefinite probabilities, we will end up not providing a definite answer as to which strategy a player should follow. Some options are ruled out; given what was said above about how we should think the game will keep on going after it has gone on for a while, β is probably not a viable strategy. And unless the player is quite confident the game will have no end, α also seems like a poor strategy. However any strategy which displays a heavy bias towards taking option 2 in the early rounds, but which guarantees that you will only choose it finitely often no matter what happens in the game, seems acceptable. Without any further specification of the problem, it is unwise to expect any more specificity in the answer.

Appendix

Lemma One: Whenever there is a finite expected return from following strategy χ, it is $\lim_{n \to \infty} \mathrm{T}(n,\chi)$.

Define $p(i)$ as $r_1 r_2 \ldots r_{i-1}$, and $e(i)$ as either 1, if the i'th choice in χ is option 1, or $2^{i+1}-1$ if the i'th choice in χ is option 2. That is, $e(i)$ is the net number of dollars the agent collects at stage i following strategy χ if the game gets that far, and $p(i)$ is the probability of the game getting that far. For convenience, let $g(i)$ be the gross number of dollars the agent gets at that stage before any dollars destroyed, that is, either 1 or 2^{i+1} depending on whether the choice is option 1 or option 2.

The assumption that χ has a finite expected return is just the assumption that:

$$\exists l : \lim_{n \to \infty} \left(\sum_{i=1}^{n} e(i)\,p(i)\right) = l$$

To prove: $\lim_{n \to \infty} \mathrm{T}(n,\chi) = l$

Note first that if χ has an expected return, then so must α. So the following is true:

$$\exists a : \lim_{n \to \infty} \left(\sum_{i=1}^{n} p(i)\right) = a$$

By the definition of a limit for any ε > 0 there is some k such that the following inequalities hold:

$$\sum_{i=1}^{k} e(i)\,p(i) > l - \varepsilon \quad \text{and} \sum_{i=1}^{k} p(i) > a - \varepsilon$$

We now construct an m such that $\mathrm{T}(m, \chi) > l - 2\varepsilon$. Define two functions $b(n)$ and $c(n)$ as follows: $b(n)$ is the round at which the agent following strategy χ receives the bill numbered n, if the game gets that far, and $c(n)$ is the round, if any, at which the bill numbered n is destroyed, if the game gets that far. (For convenience, and because it will be only used as a limit, we can say $c(n)$ is infinite if there are fewer than n choices of option 2 in χ.) For any χ, $b(n) \leq 2^{n+2}$-4, with equality if χ is the strategy of taking option 2 at the first n stages, and $c(n) \leq n$, with equality in the same circumstances. Using b, c and p we can give a neat formula for S:

$$\mathrm{S}(n, \chi) = p(b(n)) - p(c(n))$$

So we have the following equation:

$$\mathrm{T}(n,\chi) = \sum_{i=1}^{n} p(b(i)) - \sum_{i=1}^{n} p(c(i))$$

Let m be the largest value such that $b(m) = k$, and apply the above equation to $T(m, \chi)$. Note that the number of values of n such that $b(n) = i$ equals the number of dollar bills the agent is given at the i'th stage in strategy χ. That is, there are 2^{i+1} values of n such that $b(n) = i$ if the i'th move in χ is option 2, and just one such value if the i'th move is option 1. Hence we have:

$$\sum_{i=1}^{m} p(b(i)) = \sum_{i=1}^{k} g(i)\,p(i)$$

If the i'th move in strategy χ is option 2, then there exists a unique n such that $c(n) = i$. Otherwise, there is no such n. Note that $g(i)$ - $e(i)$ equals 1 if the i'th choice is option 2, and 0 if it is option 1, so these two go well together. In general, we have the following:

$$\sum_{i=1}^{n} p(c(i)) = \sum_{i=1}^{c(n)} (g(i) - e(i))\,p(i)$$

Putting the pieces together:

$$\sum_{i=1}^{m} p(b(i)) - \sum_{i=1}^{m} p(c(i)) = \sum_{i=1}^{k} g(i)\,p(i) - \sum_{i=1}^{c(n)} (g(i) - e(i))\,p(i)$$

$$= \sum_{i=1}^{k} e(i)\,p(i) - \sum_{i=k+1}^{c(m)} (g(i) - e(i))\,p(i)$$

To finish this part of the proof, we need only show that the final term here is no greater than ε, which can be seen to be true as follows:

$$\sum_{i=k+1}^{c(m)} (g(i) - e(i))\,p(i) \le \sum_{i=k+1}^{c(m)} p(i) \le \sum_{i=k+1}^{\infty} p(i) < \varepsilon$$

So we have proven that for any arbitrarily small ε there is an *m*, T(*n*, χ) is greater than *l* - 2ε for all *n* > *m*.

To complete the proof we need to show that T(*n*, χ) < *l* for all *n*. Since T(*n*, χ) is monotonic non-decreasing in *n*, if we can show this for every *n* there is an *m* > *n* such that T(*m*, χ) < *l*, this will be sufficient. Let *m* be the largest integer such that *b*(*m*) = *b*(*n*). We call this value, i.e. *b*(*m*), *k*. As was shown above, we have the following equation:

$$\mathrm{T}(m, \chi) = \sum_{i=1}^{k} e(i)\,p(i) - \sum_{i=k+1}^{c(m)} (g(i) - e(i))\,p(i)$$

Since the left-hand term is less than *l*, and the right-hand term is non-negative, this shows T(*m*, χ) < *l*, as required.

Lemma Two: Whenever the expected return from α is finite, β beats α.

Using the terminology of the previous proof, the assumption is that:

$$\exists l : \lim_{n \to \infty} \left(\sum_{i=1}^{n} p(i)\right) = l$$

From Lemma One, it follows that:

$$\lim_{n \to \infty} \mathrm{T}(n, \alpha) = l$$

Let k be some value such that $T(k, \alpha) > l/2$. We aim to show that $T(4k, \beta) > l$, and hence that β is a better strategy than α.

Since $S(n, \alpha)$ is monotonic non-increasing in n, and $f(4n-j) \leq n$ for $j \in \{0, 1, 2, 3\}$, the following inference is sound, for $j \in \{0, 1, 2, 3)$ and any n.

$$S(4n - j, \beta) = S(f(4n - j), \alpha) - S(4n - j, \alpha)$$

$$\geq S(n, \alpha) - S(4n - j, \alpha)$$

Applying this result, we get the following:

$$S(1, \beta) + S(2, \beta) + S(3, \beta) + S(4, \beta) \geq 4S(1, \alpha) - S(1, \alpha) + S(2, \alpha) + S(3, \alpha) + S(4, \alpha)$$

$$S(5, \beta) + S(6, \beta) + S(7, \beta) + S(8, \beta) \geq 4S(2, \alpha) - S(5, \alpha) + S(6, \alpha) + S(7, \alpha) + S(8, \alpha)$$

…

$$S(4k\text{-}3,\beta) + S(4k\text{-}2,\beta) + S(4k\text{-}1,\beta) + S(4k,\beta) \geq 4S(k,\alpha) - S(4k\text{-}3,\alpha) + S(4k\text{-}2,\alpha) + S(4k\text{-}1,\alpha) + S(4k,\alpha)$$

Summing the left and right-hand sides we get:

$$T(4k, \beta) \geq 4T(k, \alpha) - T(4k, \alpha)$$

By hypothesis, $T(k, \alpha) > l/2$ and $T(4k, \alpha) < l$, since S is monotonic non-decreasing, and it approaches l as a limit. Hence we get:

$$T(4k, \beta) > l$$

So $\forall i, i \geq 4k$, we have

$$T(i, \beta) > T(i, \alpha)$$

So β is a better strategy than α.

Lemma Three: Whenever the probability that the game will not end in finitely many steps is positive, α beats β.

Define R_j as $\prod_{i=1}^{j} r_i$. The hypothesis is that $\exists l : \lim_{j \to \infty} R_j = l$. Choose some k such that $R_k < 3l / 2$. Using the function b as in the previous proof, let m be any value such that $b(m) = k$. As was shown above, $S(m, \beta) = S(k, \alpha) - S(m, \alpha)$. Since $S(j, \alpha) = R_j$, we have $S(m, \beta) = R_k - R_m$, and since $R_k < 3l / 2$, and $R_m > l$, this implies $S(m, \beta) < l / 2$. As R_j is non-increasing, and b is non-decreasing, this also shows that for all $n \geq m$, $S(n, \beta) < l / 2$. Let o be any integer such that $o > m$ and $o > m + (T(m, \beta) - T(m, \alpha)) / 2l$. It can be easily shown that $T(o, \alpha) > T(m, \alpha) + (o - m)l$, and that $T(o, \beta) < T(m, \beta) + (o - m)l / 2 < T(o, \alpha)$, which proves that α is a better strategy than β.[4]

References

Arntzenius, Frank and David McCarthy (1997) "The Two-Envelope Paradox and Infinite Expectations" *Analysis* 47: 42-50.

Barrett, Jeffrey and Frank Artzenius (1999) "An Infinite Decision Puzzle" *Theory and Decision* 46: 101-3.

Jeffrey, Richard (1983a) "Bayesianism with a Human Face", in J. Earman (ed.) *Testing Scientific Theories*, Minneapolis: University of Minnesota Press.

Keynes, John Maynard (1936) *The General Theory of Employment, Interest and Money*. London: Macmillan.

Levi, Isaac (1980) *The Enterprise of Knowledge*. Cambridge, MA.: MIT Press.

Seidenfeld, Teddy (1984) Review of Paul Horwich's *Probability and Evidence*. *Philosophical Review* 93 474-483.

van Fraassen, Bas (1990) "Figures in a Probability Landscape" in J. Dunn and A. Gupta (eds) *Truth or Consequences*. Amsterdam: Kluwer, pp 345-356.

Walley, Peter (1991) *Statisical Reasoning with Imprecise Probabilities* London: Chapman & Hall.

Williams, Peter (1976) Indeterminate Probabilities. In M. Przelecki, K. Szaniawski and R. Wojcicki (eds) Formal *Methods in the Methodology of the Empriccal Sciences*. Dordrecht: Reidel, pp 229-246.

Author:

Brian Weatherson

Department of Philosophy

Syracuse University

541 Hall of Languages

Syracuse, NY, 13244

[4] Thanks to John Hawthorne and Ted Sider for helpful comments and suggestions.