Albert and the Dinosaurs
A Solution Concept for Games with Imperfect Recall

by
Steven E. Landsburg
University of Rochester

When agents have imperfect recall, seemingly unambiguous decision problems can allow multiple interpretations. To highlight some of these hidden ambiguities, Piccione and Rubinstein [PR] posed the problem of an absent-minded driver trying to navigate his way home without ever being able to remember how far he's traveled. The apparent paradox is that the driver has clear incentives to deviate from his own optimal strategy.

In response, Aumann, Hart and Perry [AHP] formulated a solution concept that they defended as the unique correct resolution. By contrast, Lipman [L] constructs two apparently isomorphic versions of the problem, only one of which seems amenable to the [AHP] solution. Thus, argues Lipman, the ambiguities remain.

Insofar as the problem is ambiguous, there can of course be no one correct solution. Nevertheless, there are certain properties that any reasonable solution should satisfy. For example, *if a decision maker is unable to distinguish between two nodes, then he must use the same decision procedure at each of them.*

I will argue that if we take this principle seriously, we are led to a new solution concept with significant conceptual advantages over those considered by [PR] and [AHP]. I've labeled this solution concept "Mark 3", to distinguish it from the Aumann/Hart/Perry model ("Mark 2"), the original Piccione/Rubinstein story ("Mark 1"), and a benchmark model with no updating ("Mark 0"). I prove some existence and uniqueness theorems for the Mark 3 model, and compare the predictions of the models under various auxiliary assumptions. (The "Mark X" models specify the procedure by which the driver updates his plans at each intersection; the auxiliary assumptions specify the procedure by which he initializes his plans at the beginning of the journey.)

For ease of exposition, Part I (Sections 1-5) fixes a particularly simple payoff structure. In Part II (Sections 6-9) we repeat the analysis with a more general payoff structure that requires us to allow some randomization in the updating procedure.

Part I

1. The Absent-Minded Driver. Each day, Albert leaves his office (at the bottom of the map below), gets on the Main Highway and attempts to drive home to his house on Second Street. If he turns too soon (onto First Street) or if he overshoots (going all the way to the north end of the Main Highway), he is mauled by dinosaurs.[1]

---

[1] This is essentially the problem posed by Piccione and Rubinfield in [PR], though I've changed the payoffs for expositional reasons. Nothing important depends on this change. In Section 10 and beyond, we'll adopt a more general model that includes the original [PR] setup, with the original payoffs, as a special case.

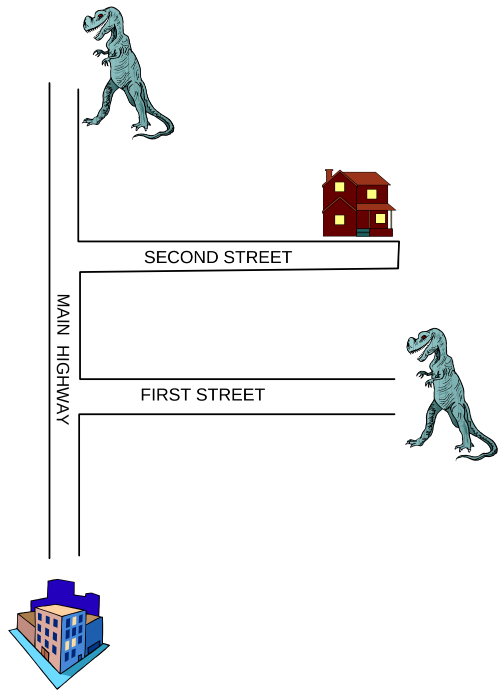

Figure 1

Obviously, Albert's best strategy is to go straight at the first intersection and turn right at the second. Unfortunately, both intersections look identical. Doubly unfortunately, Albert can never remember whether he's already passed the first intersection.

Because Albert can't tell the intersections apart, he needs a single strategy for both of them. Strategy A is to turn right at every intersection. This delivers him directly to the First Street dinosaur mob. Strategy B is to go straight at every intersection, putting him on a direct route to the North Side crew. Neither of these strategies has any chance of getting him home.

Therefore, Albert adopts Strategy C, which is to flip a fair coin at every intersection. This gives him a 1/2 chance of going straight at First Street and a 1/2 chance of turning right at Second, for an overall 1/4 chance of arriving home safely. It's easy to compute that Albert can do no better.

Fortunately, Albert is smart enough to figure this out. Perhaps unfortunately, he's also smart enough to reason a bit further: Under Strategy C, he reaches First Street every day, but Second Street only half the time. Thus when he reaches an intersection, there's a 2/3 chance that he's at First Street. Now if he deviates from Strategy C by going straight with some new probability q, he gets home with probability

$$(2/3)q(1-q)+(1/3)(1-q)$$

To maximize this expression, Albert chooses $q = 1/4$. If he makes this choice at every intersection, he gets home with probability $q(1-q) = 3/16$, which is less than 1/4. Albert would have been better off with the original strategy C.

Piccione and Rubinstein [PR] attribute this paradoxical result to ambiguities in the formulation of the problem. A number of authors have either embraced or attempted to resolve these ambiguities in a variety of ways. The next section incorporates several of these attempts into a general framework.

2. A General Framework. When Albert arrives at an intersection, he chooses a probability of going straight. This choice is made subject to some belief about the probability that he's gone straight at past intersections (if any) and probability that he will go straight at future intersections.

We set:

$$\left.\begin{aligned} p &= \text{The probability of having gone straight in the past.} \\ q &= \text{The probability of going straight in the present.} \\ r &= \text{The probability of going straight in the future.} \end{aligned}\right\} \tag{2.1}$$

We interpret "absent-mindedness" to mean that, beyond knowing the structure of the problem, Albert's mind is capable of storing just one number (in this case p). At the intersection, he updates the stored number to q, goes straight with probability q, and continues to store the number q until he reaches the next intersection.

Moreover, we assume that Albert is an optimizer, so that he chooses q so as to maximize

$$\frac{1}{1+p}q(1-r) + \frac{p}{1+p}(1-q) \tag{2.2}$$

or, equivalently, to maximize

$$q(1-r-p) \tag{2.3}$$

subject to his beliefs about p and r. We assume also that these beliefs are accurate.

This leaves us with two modeling choices:

1. *What determines r?*
2. *At the first intersection, what determines p?*

Although there are surely some assumptions built into this formulation, we will soon see that it is general enough to encompass the solutions contemplated in [PR] and [AHP], as well as the new solution that will be the main focus of the present paper.[2]

We will address these questions separately in the next two sections.

3. Modeling the Future

Recall that Albert chooses q to maximize (2.3) subject to his (accurate) memory of p and his (accurate) forecast of r. In this section we address the question *what determines r?*

[2] Our two modeling choices subsume the two key issues identified by Piccione and Rubinstein in their followup paper [PR2]. First, they ask "Can a decision maker control his future behavior?", i.e. (in our formulation) "Can Albert lock in the value of r?". Our first modeling choice subsumes this question. Second (actually first, because we've reversed the order of the issues here from [PR2]), they ask "Is there a preplay planning stage?", which we interpret as "Can Albert set the initial value of p?". Our second modeling choice subsumes this question.

We offer four possible answers: A baseline model (which we call Mark 0) in which Albert never updates, the model considered by Piccione and Rubinstein (Mark 1), the model endorsed by Aumann, Hart and Perry (Mark 2) and a new model that I will argue is often conceptually preferable (Mark 3). The Mark 3 model is the main contribution of this paper.

Mark 0. Albert never updates. That is, once p is chosen (via some process to be modeled later), he is constrained to set $q = r = p$. With $p = 1/2$, this is the "Strategy C" model of Section 1.

Mark 1. $r = q$. That is, Albert believes (accurately) that he can update once but can never update again. Then to maximize (2.3), Albert sets

$$q = (1 - p)/2 \tag{3.1}$$

. With $p = 1/2$, this is the model from the end of Section 1. There is, of course, something very odd about this model: When Albert reaches the first intersection, he adjusts his going-straight probability and assumes he can *commit* to that new probability for the future, even though he's in the process of violating a commitment that he made in the past. As [PR] observe, this paradox is avoidable if and only if $r = q = p = 1/3$.

Mark 2. $r = p$. This is the assumption on which Aumann, Hart and Perry insist, arguing, in effect, that Albert has no reason to believe that he'll act any differently in the future than he has in the past. Then to maximize (2.3), Albert sets

$$q = \begin{cases} 1 & \text{if } p < 1/2 \\ \text{arbitary} & \text{if } p = 1/2 \\ 0 & \text{if } p > 1/2 \end{cases} \tag{3.2}$$

With $p = 1/2$, this gives $r = q = p = 1/2$.

Mark 3. The Mark 3 model is an alternative which, as far as I am aware, has not previously appeared in the literature. It rests on two principles: First, Albert is sufficiently absent-minded that his state of mind consists of a single number p (which he updates to q upon reaching a decision point) but sufficiently foresighted that he can forecast the future state r. Second, when Albert can't distinguish between two nodes, he must use the same updating procedure at each of them.

It follows, then, that at each intersection, Albert updates the number in his memory according to a rule of thumb that is optimal subject to the assumption that he'll continue to use the same rule of thumb in the future.

More precisely, this implies that $r = f(q)$ and $q = f(p)$ where $f : [0,1] \to [0,1]$ is some function such that $q = f(p)$ maximizes (2.3) subject to the condition $r = f(q)$. In other words, f should be such that the expression

$$q(1 - f(q) - p) \tag{3.3}$$

is maximized at $q = f(p)$. In still other words, for all $p, q \in [0,1]$ we must have

$$f(p)(1 - f(f(p)) - p) \geq q(1 - f(q) - p) \tag{3.4}$$

It turns out that there is a unique function f satisfying (3.4):

Theorem 1. Suppose that $f : [0,1] \to [0,1]$ is a function satisfying (3.4) for all p and q. Then f is the function

$$f(p) = \begin{cases} 1 & \text{if } p < 1 \\ 0 & \text{if } p = 1 \end{cases} \tag{3.5}$$

Proof. We break the proof into two lemmas:

Lemma 1A. $f(1) = 0$.

Proof. Suppose otherwise.

To compute $f(1)$, we maximize

$$q(1 - f(q) - 1) = -qf(q)$$

which is maximized when (and only when) $qf(q) = 0$. Thus $f(1)f(f(1)) = 0$. We've assumed $f(1) \neq 0$, so we get

$$f(f(1)) = 0 \tag{3.6}$$

Now consider (3.4) with $p = q = f(1)$:

$$f(f(1))(1 - f(f(f(1))) - f(1) \geq f(1)(1 - f(f(1)) - f(1))$$

which, in light of (3.6), simplifies to:

$$0 \geq f(1)(1 - f(1)) \tag{3.7}$$

The right-hand side of (3.7) can't be negative, so we have equality. But $f(1) \neq 0$ by assumption, so it follows that

$$f(1) = 1 \tag{3.8}$$

.

Combining (3.6) with (3.8) gives the contradiction:

$$0 = f(f(1)) = f(1) = 1$$

q.e.d.

Lemma 1B. If $p \neq 1$, then $f(p) = 1$.

Proof. $f(p)$ must maximize the expression

$$q(1 - f(q) - p)$$

But $1 - f(q) - p$ is maximized at $q = 1$ by Lemma 1A, and q is *uniquely* maximized at $q = 1$, so the only possibility for the maximand is $q = 1$.

q.e.d.

This completes the proof of the theorem.

Our model therefore requires that Albert employ the function (3.5).

4. Modeling the Initial Condition. In Section 2, we raised two modeling questions: What determines r, and what determines the initial choice of p? The Mark 0, 1, 2, and 3 models of Section 3 are responses to the first of these questions. Here we turn to the second.

There are many ways to model the initial choice of p. Of these, we'll focus on two:

The Consistency Requirement. We require that

$$q = p \tag{4.1}$$

That is, Albert goes straight with the same probability at every intersection. The authors of [AHP] strongly defend this requirement, arguing that because all intersections are identical, Albert must behave identically at each intersection.

The Optimality Requirement. This is the requirement that p be chosen optimally taking as given the rest of the model (the "rest of the model" being, for example, Mark 0, 1, 2, or 3).

In the next section, we'll say a few words about the relative merits of these requirements. Here we'll explore their implications. Combining one or the other requirement with the models of Section 3 gives a menu of options, summarized in Table 1 of the appendix. In each case we evaluate p, q and r at the first intersection, so that Albert's probability of getting home is $q(1-r)$. In what follows, "Mark X.C" means "the Mark X model augmented with the Consistency Requirement" and "Mark X.O" means "the Mark X model augmented with the Optimality Requirement".

Mark 0. Here Albert never updates, so $r = q = p$ and the consistency requirement (3.4) is satisfied automatically. If we further impose the optimality requirement, then, as in Section 1, we have $p = 1/2$ and Albert gets home with probability $1/4$.

Mark 1.C. Here Albert has only one chance to update and must decide not to. From (3.1) and (4.1) we get $p = q = r = 1/3$, so Albert gets home with probability $2/9$. This is one of the models considered in [PR].

Mark 1.O. Let q be the probability that Albert goes straight at the first intersection. From the Mark 1 assumption $r = q$ and the Mark 1 conclusion (3.1), we see that Albert gets home with probability $(1-p)^2/4$, which is maximized at $p = 0$, where Albert gets home with probability $1/4$.

Mark 2.C. This is the model of [AHP]. Here (3.2) and (4.1) imply $p = q = r = 1/2$ and a $1/4$ probability of getting home.

Mark 2.O. Given the Mark 2 condition (3.2), Albert maximizes his chance of getting home by setting $p = 0$, leading to $q = 1$, $r = 0$, and a guaranteed safe arrival at home.

Mark 3.C. The function (3.5) has no fixed point, so there is no solution to the Mark 3.C version of the model.

Mark 3.O. Given (3.5), it is optimal for Albert to choose any $p < 1$, giving $q = 0$, $r = 1$, and a guaranteed safe arrival at thome.

5. Discussion. We've seen that if we augment the Mark 3 model with the Optimality Requirement, Albert achieves the first best outcome, i.e. he gets home for sure, essentially because he manages to coordinate his behavior perfectly across the two intersections.

Perhaps the reader feels we've somehow cheated. After all the spirit of the problem seems to forbid precisely this sort of coordination.

To this there are several responses. First, we did not *contrive* to let Albert coordinate. Instead, we imposed some simple principles and followed where they led.

Second, the outcome is partly illusory, in the sense that it's an artifact of the particular payoff structure in Figure 1. We will see in future sections that if we alter the payoff structure, the Mark 3 version of Albert will be unable to achieve his first best outcome, regardless of the initial choice of p.

Third, and by far most importantly, even with the current parameter values, the first best outcome is *not* an implication of the Mark 3 model. Instead it is an implication of *the Mark 3 model augmented with the Optimality Requirement*. Moreover, the same first best outcome is *also* an implication of the *Mark 2* model augmented by the Optimality Requirement. So it's the Optimality Requirement, not the choice of Mark 3 over Mark 2, that drives the first best outcome. Or, to put this another way, if the possibility of first-best coordination is somehow a strike against the Mark 3 model, then it's a strike against the Mark 2 (Aumann-Hart-Perry model) as well.

Of course, Aumann-Hart-Perry reject the Optimality Requirement in favor of the Consistency Requirement. This raises the question of whether they were right to do so. Indeed, one wonders why the [AHP] version of Albert, in full knowledge of his absent-mindedness and his own updating procedure, would fail to optimize at the outset.[3]

A possible (partial) response is that Albert is so absent-minded that he not only fails to distinguish one intersection from the other, but fails to distintuish the intersections from the starting point. Therefore, there's never a moment when he knows he's choosing p "at the outset".

But there are multiple problems with this response. First, there is of course no choice of directions at the outset; Albert can only go straight. Second, and more fundamentally, the entire setup of the problem assumes that when Albert chooses q, he remembers (and accounts for) the value of p. By contrast, when he chooses p, there is no past value to remember. Surely Albert can recognize that the choice of p is a fundamentally different sort of problem than the choice of q, and this ought to be enough to tell him whether he's at the outset or not.

A second possible response is the one endorsed in the [AHP] paper: If Albert chooses any initial state p other than $1/2$ — say $p = 0$ — then he updates his state to 1 at First Street and to 0 at Second Street. Because he's aware of his state, he can use this information to infer his current location — knowledge that the statement of the problem explicitly disallows.

[3] One wonders also what force could lead Albert to choose precisely that p which happens to satisfy the Consistency Requirement.

One problem with this response is that it's **not true**. According to (5.2), when Albert is in State 0, all he knows is that he is either a) at the first intersection after having started out in a state $p \geq 1/2$ or b) at the second intersection after having started out in a state $p \leq 1/2$. When he's in State 1, all he knows is that he's either at the first intersection after having started out in a state $p \leq 1/2$ or at the second intersection after having started out in a state $p \geq 1/2$.

Indeed, [AHP] suggest a similar mechanism toward the end of their paper, where they imagine that Albert has a handkerchief in which he ties a knot every time he passes an intersection. If the handkerchief has an even number of knots, he goes straight; if it has an odd number of knots, he turns. Being absent-minded, he does not remember whether the handkerchief had an even or an odd number of knots at the outset. ([AHP] impose this condition so as to conform to the proviso that Albert never knows which intersection he's at.) This allows Albert to get home with probability 1/2, which is better than he can do without the handkerchief (or some other correlating device).

The argument here is that [AHP] are mistaken in their belief that they can (in the main part of their paper) simply assume away the handkerchief. Albert — in every model we'll consider, including AHP's — is never so absent-minded as to forget his current state of mind; if he *were* that absent-minded he'd be facing Knightian uncertainty as opposed to Knightian risk, and none of our (or [AHP]'s) calculations would be relevant[4] But Albert's knowledge of his own state can be just as useful as a handkerchief. Our models don't allow us to assume that knowledge away, so we ought to allow Albert to use it. The AHP handkerchief gets Albert home with probability 1/2 only because the authors have (implicitly) assumed that the initial state of the handkerchief is effectively determined by a fair coin flip. If he were smart enough to start with an even number of knots in his handkerchief, he'd make it home for sure.

Part II

6. A More General Payoff Structure. Now we put Albert in a more general environment with the following payoffs (where C and D are non-negative constants):

[4] For example, when Albert optimizes on the assumption that he's reached the first intersection with probability $1/(1+p)$, he surely must know the value of p.

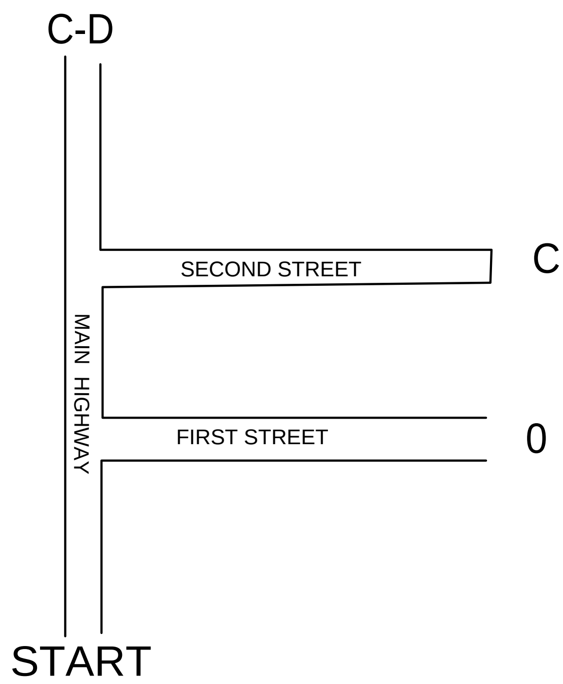

Figure 2

The model of Part I assumes $C = D$. The original Piccione/Rubinstein paper assumes $C = 4$, $D = 3$.

In the general setup of Figure 2, the Mark 3 solution concept requires Albert, taking p as given, to choose $q = f(p)$ so as to maximize

$$0 \cdot (1-q) + 1 \cdot qf(q) + C \cdot q(1-f(q)) + C \cdot p(1-q) + (C-D) \cdot pq$$

or equivalently to maximize

$$q(C - Dp - Df(q))$$

(This is the analogue of (3.3).)

A Mark 3 solution is a function f satisfying

$$f(p)(C - Dp - Df(f(p))) \geq q(C - Dq - Df(q)) \tag{6.1}$$

for all p and q.

Example. Suppose $C \leq D$, and choose $T \in \{0\} \cup (C/D, 1]$. Then the following function sastifies (1):

$$f(p) = \begin{cases} 1 & \text{if } p < C/D \\ T & \text{if } p = C/D \\ 0 & \text{if } p > C/D \end{cases} \tag{6.2}$$

Note in particular that if $C = D$, we must choose $T = 0$, so that (6.2) specifies a unique function, as required by Theorem 1.

In the Piccione-Rubinstein paper, we have $C = 4$, $D = 3$. The following theorem establishes, among other things, that in that case there is no Mark 3 solution to Albert's problem. In Section 7, we will remedy this situation by allowing Albert to use a form of mixed strategies.

Theorem 2:

a) If $C \leq D$, then any f satisfying (6.1) must be of the form (6.2). In particular, if $C/D = 1$, there is a unique such f.

b) If $D < C < 2D$, there is no f satisfying (1).

c) If $2D \leq C$, then the unique function satisfying (6.1) is the constant function $p \mapsto 1$).

Proof. Assume f satisfies (1). We proceed via a sequence of lemmas.

Lemma 2A. f is non-increasing.

Proof. Replacing q with $f(q)$ on the right side of (1), we have, for all p and q:

$$f(p)(C - Dp - Df(f(p))) \geq f(q)(C - Dp - Df(f(q))) \tag{6.3}$$

Interchanging p with q in (6.3) gives

$$f(q)(C - Dq - Df(f(q))) \geq f(p)(C - Dq - Df(f(p))) \tag{6.4}$$

Adding (3) to (4) and simplifying gives

$$(f(p) - f(q))(q - p) \geq 0 \tag{6.5}$$

as needed.

q.e.d.

Lemma 2B:

a) If $p < C/D$, then $f(p) \neq 0$.

b) If $p > C/D$, then $f(p) = 0$.

c) If $p = C/D$ then either $f(p) = 0$ or $f(p) > p$.

d) If $p = C/D = 1$ then $f(p) = 0$.

Proof. (a) Suppose $p < C/D$. Then taking $q = p$ in (6.1) gives

$$f(p)(C - Dp - Df(f(p))) \geq p(C - Dp - Df(p))$$

which simplifies to

$$f(p)(C - Df(f(p)) \geq p(C - Dp) > 0$$

as needed.

(b) Taking $q = 0$ in (6.1) gives

$$f(p)(C - Dp - Df(f(p))) \geq 0 \tag{6.6}$$

If $p > C/D$, this implies $f(p) = 0$.

(c) Suppose $p = C/D$ and $f(p) \neq 0$. Then (6.6) becomes

$$f(p)(-Df(f(p))) \geq 0$$

so

$$f(f(p)) = 0 \tag{6.7}$$

.

Now replace both p and q with $f(p)$ in (1) to get

$$0 = f(f(p))(C - Df(p) - Df(f(f(p)))) \geq f(p)(C - Df(p) - Df(f(p))) = f(p)(C - Df(p))$$

Because $f(p)$ is assumed nonzero, this gives $f(p) \geq C/D = p$. It remains to rule out the possibility that $f(p) = p$. But in view of (6.7), this would yield the contradiction

$$0 = f(f(p)) = f(p) \neq 0$$

(d) This follows immediately from (c).

q.e.d.

Lemma 2C. If $C/D \leq 1$ then $f(1) = 0$.

Proof. For $C/D < 1$, this is Lemma 2B(b) and for $C/D = 1$ this is Lemma 2B(d).

q.e.d.

Lemma 2D.

a) If $p < C/D$ then $C - Dp - Df(f(p)) \geq 0$.

b) If $p < C/D \leq 1$ then $f(p) = 1$.

Proof. (a) follows from (6) and Lemma 2B(a).

For (b), suppose $p < C/D \leq 1$. Now take $q = 1$ in (1) and use Lemma 2C to get to get

$$f(p)(C - Dp - Df(f(p))) \geq 1 \cdot (C - Dp - Df(1)) = C - Dp > 0$$

Thus

$$0 \geq -Df(p)f(f(p)) \geq (1 - f(p))(C - Dp) \geq 0$$

which forces $f(p) = 1$.

q.e.d.

Lemma 2E. If $1 < C/D$, then for all p, q, if $q > f(p)$ then $f(q) > f(f(p))$.

Proof. If $C/D > 1$ then Lemma 2D(a) applies to both p and q. Thus each side of the inequality (1) (reproduced here) is non-negative:

$$f(p)(C - Dp - Df(f(p))) \geq q(C - Dp - Df(q)) \tag{6.8}$$

which suffices.

q.e.d.

Lemma 2F. If $1 < C/D$ then $f \circ f$ is a constant function.

Proof. Let A be the range of f and consider the restricted function $f|_A$. By Lemma 2A, this function is non-increasing, and by Lemma 2E, it's non- decreasing.

q.e.d.

Lemma 2G.

a) If $p < 1 < C/D$, then $f(p) = 1$.

b) If $1 < C/D$, then $f(f(p)) = 1$ for all p.

c) If $1 < C/D$ then $f(1) = 1$.

Proof. By Lemma 2F, we can write $f(f(p)) = K$ for some constant K. Substituting 1 for p in Lemma 2D(a) gives

$$C - D(1 + K) \geq 0 \tag{6.9}$$

a) Taking $p < 1$ and $q = 1$ in (1) and invoking (6.9)gives gives

$$f(p)(C - Dp - DK) \geq (C - Dp - DK) > (C - D(1 + K)) > 0$$

b) We need to show that $K = 1$. If not, then two applications of (a) give

$$K = (f \circ f)(1) = (f \circ f)(f(K)) = f(f \circ f)(K) = f(K) = 1$$

c) By b),

$$f(1) = f(f \circ f(1)) = (f \circ f)(f(1)) = 1$$

q.e.d.

Lemma 2H. If $C/D < 2$, then $f(1) = 0$.

Proof. Taking $p = 1$ and $q = 0$ in (1) and using Lemma 2G(c)gives

$$0 > f(p)(C - 2D) \geq 0$$

q.e.d.

Proof of Theorem 2. (a) is the conjunction of Lemmas 2D(b), 2B(b) and 2B(c). For (b), the existence of f, together with Lemmas 2G(c) and 2H, gives a contradiction. (c) follows from Lemma 2G(a).

q.e.d.

7. Mixed Strategies. According to Theorem 2, there is a considerable range of parameter values for which no Mark 3 solution exists. In those situations, what can we advise Albert to do? If there's no available solution in pure strategies, perhaps he can resort to mixed strategies.

That is, instead of choosing a single $f(p)$ to maximize (7.1), we allow Albert to choose a probability distribution over (possibly multiple) values all of which maximize (7.1):

Definition. A *mixed strategy Mark 3 solution* to the decision problem of Figure 3 is a mapping $\mu : p \mapsto \mu_p$ from $[0, 1]$ to the set of probability measures on $[0, 1]$ such that for any $p \in [0, 1]$ and any probability measure μ on $[0, 1]$, we have

$$\int\int t(C - Dp - Dr)d\mu_t(r)d\mu_p(t) \geq \int\int t(C - Dp - Dr)d\mu_t(r)d\mu(t) \tag{7.1}$$

Example. Suppose that $0 < D < C < 2D$. Then one checks easily that the following satisfies (7.1):

$$\mu_p \begin{cases} \text{is concentrated on } 1 & \text{if } p < 1 \\ \text{assigns probability } C/D - 1 \text{ to } 1 \text{ and } 2 - C/D \text{ to } 0 & \text{if } p = 1 \end{cases} \tag{7.2}$$

Remark. The case $D < C < 2D$ is precisely the case in which Theorem 2 tells us that no pure-strategy Mark 3 solution exists. Thus we are now able to prescribe a solution for any set of parameter values. Moreover, in this case, we have uniqueness:

Theorem 3. When $D < C < 2D$, (7.2) is the unique assignment satisfying (7.1).

Proof. Suppose $p \mapsto \mu_p$ is another such solution. We start with some Lemmas:

Lemma 3A. $\mathcal{E}(\mu_p) = \int t d\mu_p(t)$ is a non-increasing function of p.

Proof. The proof mimics the proof of Lemma 2A. In (7.1), replace μ with μ_q to get an inequality. Then switch p with q to get another. Add those two inequalities and rearrange to get

$$\int\int D(q - p)t[d\mu_p(t) - d\mu_q(t)]d\mu_t(r) \geq 0$$

or

$$\int D(q - p)t[d\mu_p(t) - d\mu_q(t)] \geq 0$$

as needed.

q.e.d.

Lemma 3B. For each p, the measure μ_p is concentrated on 0 and 1.

Proof. If q is in the support of μ_p, then q maximizes $q(C - Dp - D\mathcal{E}(\mu_q))$. If $q \neq 0$, then we must have $C - Dp - D\mathcal{E}(\mu_q) \geq 0$. Therefore, by Lemma 3A, we can increase the value of $D - Dp - D\mathcal{E}(\mu_q)$ if we can increase q. Since the expression is already maximized, we must have $q = 1$.

q.e.d.

Lemma 3C. If μ_p is not concentrated on a single point, we have $p = (C/D) - \mathcal{E}(\mu_1)$.

Proof. Suppose μ_p is not concentrated on a single point. Then by Lemma 3B, it is concentrated on 0 and 1. Therefore we have

$$0 = 0(C - Dp - D\mathcal{E}(\mu_0)) = 1 \cdot (C - Dp - D\mathcal{E}(\mu_1))$$

q.e.d.

Lemma 3D. If μ_p is not concentrated on a single point, we have $p = 1$.

Proof. By Lemma 4C, we have $p = C/D - \mathcal{E}(\mu_1)$. If $p \neq 1$ then we can't also have $1 = (C/D) - \mathcal{E}(\mu_1)$, so μ_1 is concentrated on a single point, which (by Lemma 3B) must be either 0 or 1. If μ_1 is concentrated on 0, then $p = C/D > 1$, which is not possible, so μ_1 is concentrated on 1. But then $E(\mu_1) = 1$, so, by Lemma 4A, $\mathcal{E}(\mu_p) = 1$, whence μ_p is concentrated entirely on 1.

q.e.d.

Proof of Theorem 3. It follows from Lemma 3D that μ_p is concentrated on 1 for every $p \neq 1$. It follows that μ_1 cannot be concentrated on a single point, because then we'd have a solution in pure strategies, contradicting Theorem 2. Thus by Lemma 3C, we must have $1 = (C/D) - (\mu_1)$, or $(\mu_1) = (C/D) - 1$. Together with Lemma 3B, this shows that μ_1 assigns probability $(C/D) - 1$ to 1 and probability $2 - (C/D)$ to 0.

q.e.d.

8. Initial Values.

Theorem 3 establishes the Mark 3 updating procedure in mixed strategies, given an initial value for p. As in Part I, the source of that initial value is a separate modeling question.

For concreteness, we will set $C = 4$ and $D = 3$ in Figure 2. In that case, Theorem 3 tells us that given an initial value for p, Albert updates according to the rule

$$\mu_p = \begin{cases} \text{is concentrated on 1} & \text{if } p < 1 \\ \text{assigns probability 1/3 to 1 and 2/3 to 0} & \text{if } p = 1 \end{cases} \tag{8.1}$$

As in Part I, we face a separate question: How does the initial value of p arise?

Possible answers include analogues of the Consistency and Optimality Requirements in Section 4.

It's easy to check that the Consistency Requirement is satisfied (that is, the probability distribution remains unchanged after updating) if and only if the initial value of p is distributed according to the law

$$p = \begin{cases} 0 & \text{with probability } 2/5 \\ 1 & \text{with probability } 3/5 \end{cases} \tag{8.2}$$

It's also easy to check that with this initial condition, Albert achieves an expected payoff of 1.8.

Alternatively, we can impose the Optimality Requirement so that Albert chooses the initial value of p to maximize his expected payoff. Starting with any $p < 1$, Albert updates to

$$q = 1 \text{ and } r = \begin{cases} 1 & \text{with probability } 1/3 \\ 0 & \text{with probability } 2/3 \end{cases}$$

earning an expected payoff of $E(qr + 4q(1-4)) = 3$. Starting with $p = 1$, he updates to

$$\begin{cases} q = r = 1 & \text{with probability } 1/9 \\ q = 1, r = 0 & \text{with probability } 2/9 \\ q = 0, r = 1 & \text{with probability } 2/3 \end{cases}$$

earning an expected payoff of 1. Thus the Optimality Requirement dictates that Albert choose $p < 1$ and earn a payoff of 3.

9. Comparison of Models. In Part I, where we had $C = D = 1$, Albert was able to achieve first-best optimality with a Mark 3 solution. More generally, he can't — though he can do pretty well.

We continue to fix attention on the original Piccione/Rubinstein payoffs $C = 4$, $D = 3$. Then the Mark 0 optimum is achieved at $p = 2/3$, giving Albert a payoff of 1.33. The Mark 1 (Piccione/Rubinstein) solution with a consistency requirement is $p = q = r = 4/9$, giving him a payoff of 1.19.

The Mark 2 (Aumann/Hart/Perry) solution is

$$q = \begin{cases} 1 & \text{if } p < 2/3 \\ \text{arbitary} & \text{if } p = 2/3 \\ 0 & \text{if } p > 1/2 \end{cases} \tag{9.1}$$

Augmented with the Consistency Requirement, this is the same as the Mark 0 solution.

The Mark 2 solution with an Optimality Requirement — where Albert chooses p with full knowledge that he'll be following the Aumann/Hart/Perry model once he gets underway — is achieved at any $p < 2/3$, so that (according to 8.1) $q = 1$ and then $r = 0$, yielding the maximum payoff of 4.

With the payoff structure of Part 1, Albert was able to match this first-best Mark 2 solution with a Mark 3 solution (provided he chose his initial p wisely). But with the Piccione-Rubinstein payoffs, this is no longer the case. As we saw in the preceding section, the best possible Mark 3 payoff is 3. All of this is summarized in Table 2 of the appendix.

9. Conclusion. If we presume that Albert's state of mind consists of a single number, which he updates at each node subject to an accurate forecast of his own future updates, and if we assume that he must use the same updating procedure at each node, then we are led naturally to the Mark 3 model, which does not seem to have been previously considered.

In order to get solutions for all possible parameter values, we have to extend the Mark 3 model to allow mixed strategies, which means that there is some randomness in the updating process.

The Mark 3 model, like the others we've considered, must be augmented with some assumption about Albert's initial state of mind. [AHP] deal with this issue by imposing the Consistency Requirement. We have argued for imposing the Optimality Requirement instead. However, the main focus of this paper is not on Consistency versus Optimality but on Mark 3 versus Mark 2, and in particular on establishing existence and uniqueness theorems for Mark 3.

For some parameter values, the Mark 3 model augmented with the Optimality Assumption allows Albert to achieve his first best outcome. For those parameter values, the same is true of the Mark 2 ([AHP]) model. For other parameter values, Albert can frequently achieve the first best under Mark 2 but not Mark 3.

The key difference between the models lies not in what Albert can achieve, but in how we think about his maximization problem. The Mark 2 version of Albert makes the (in general) unwarranted assumption that he'll revert to past behavior in the future.[5] The Mark 3 version, by contrast, assumes — correctly! — that he'll optimize in the future exactly as he does in the present. Absent-minded though he may be, our version of Albert is a rational man.

Appendix: Tables

	Assumptions	p	q	r	probability of safe arrival
Mark 0.	r= q=p	p	p	p	p(1-p)
Mark 0.C		p	p	p	p(1-p)
Mark 0.O		1/2	1/2	1/2	0.25
Mark 1.	r= q	p	$\frac{1-p}{2}$	$\frac{1-p}{2}$	$\frac{1-p^2}{4}$
Mark 1.C		1/3	1/3	1/3	0.22
Mark 1.O		0	1/2	1/2	0.25
Mark 2.	r=p	see (3.2)			
Mark 2.C		1/2	1/2	1/2	0.25
Mark 2.O		<1/2	1	0	1.00
Mark 3.	r= f(q) q=f(p)	see (3.5)			
Mark 3.C		no solution			
Mark 3.O		< 1	1	0	1.00

TABLE 1
(for the game in Figure 1)

	Assumptions	p	q	r	expected payoff
Mark 0.	r= q=p	p	p	p	p(4-3p)
Mark 0.C		p	p	p	p(4-3p)
Mark 0.O		2/3	2/3	2/3	1.33
Mark 1.	r= q	p	$\frac{4-3p}{6}$	$\frac{4-3p}{6}$	$\frac{16-9p^2}{12}$
Mark 1.C		4/9	4/9	4/9	1.19
Mark 1.O		0	2/3	2/3	1.33
Mark 2.	r=p	see (3.2)			
Mark 2.C		2/3	2/3	2/3	1.33
Mark 2.O		<2/3	1	0	4.00
Mark 3.	r= f(q) q=f(p)	see (9.1)			
Mark 3.C		0 with prob 2/5, 1 with prob 3/5			1.80
Mark 3.O		< 1	1	0	3.00

TABLE 1
(for the game in Figure 2 with C=4, D=3)

For each version of the model, the charts shows the values of p, q and r at the first intersection, and Albert's probability of getting home. The listed assumptions are the assumptions Albert makes when he chooses q subject to an accurate memory of p and a rational expectation of r. The "C" versions of the models impose the consistency requirement that r= q=p, and the "O" versions impose the requirement that p is chosen optimally, given the rest of the model.

References

[AH] R. Aumann, S. Hart, and M. Perry, "The Absent- Minded Driver", Games and Economic Behavior 20, 102-116 (1997)

[L] B. Lipman, "More absentmindedness", Games and Economic Behavior 20, 97-101 (1997)

[PR] M. Piccione and A. Rubinstein, "On the Interpretation of Decision Problems With Imperfect Recall", Games and Economic Behavior 20, 3-24 (1997)

[5] The assumption becomes warranted in the presence of the Consistency Requirement — which [AHP] impose — but not otherwise. Thus in the absence of this requirement — i.e. unless he happens to choose exactly the right p at the outset — the Mark 2 Albert's belief about his own future behavior is in general incorrect.

[PR2] M. Piccione and A. Rubinstein, "The Absent Minded Driver's Paradox: Synthesis and Responses", Games and Economic Behavior 20, 121-130 (1997).