GAMES AND ECONOMIC BEHAVIOR **20,** 3–24 (1997)
ARTICLE NO. GA970536

On the Interpretation of Decision Problems with Imperfect Recall*

Michele Piccione

Department of Economics, University of British Columbia, Vancouver, Canada

and

Ariel Rubinstein

School of Economics, Tel Aviv University, Tel Aviv, Israel

Received September 18, 1995

We argue that in extensive decision problems (extensive games with a single player) with *imperfect* recall care must be taken in interpreting information sets and strategies. Alternative interpretations allow for different kinds of analysis. We address the following issues: 1. randomization at information sets; 2. consistent beliefs; 3. time consistency of optimal plans; 4. the multiselves approach to decision making. We illustrate our discussion through an example that we call the "paradox of the absentminded driver." *Journal of Economic Literature* Classification Numbers: C7, D0. © 1997 Academic Press

1. INTRODUCTION

This paper is an examination of some modelling problems regarding imperfect recall within the model of extensive games. It is argued that, if the assumption of perfect recall is violated, care must be taken in interpreting the main elements of the model. Interpretations that are inconsequential under perfect recall have important implications in the analysis of games with imperfect recall.

* The first author acknowledges financial support from SSHRC-Canada. The second author acknowledges partial financial support from the United States–Israel Binational Science Foundation, Grant Number 1011-341. We are grateful to the University of Oslo and New York University for their hospitality, and to Cafe Tamar in Tel Aviv for providing an inspiring atmosphere. We thank Bob Aumann, Paolo Battigalli, Sergiu Hart, Dana Heller, Bart Lipman , Avishai Margalit, Roger Myerson, Hugh Neary, Motty Perry, Tim Van Zandt, and Ruth Weintraub for their comments on earlier versions of this paper.

0899-8256/97 $25.00

The distinction between perfect and imperfect recall for extensive games was introduced in Kuhn (1953). Since then, traditional game theory has excluded games with imperfect recall from its scope. In this paper, we wish to readdress this topic. Since the interpretative issues on the agenda appear within the framework of extensive games with a single player, we confine our discussion to *decision problems* with imperfect recall.

Extensive decision problems are a special case of extensive games in that the set of players is a singleton. Our basic understanding of an extensive decision problem includes the following assertions:

1. Informational assumptions are modeled by partitioning the situations (histories) where the decision maker takes an action into "information sets." The interpretation of the "informational structure" is that the decision maker knows the information set he is at but, if an information set is not a singleton, he cannot distinguish among the points in the set he has reached. The decision maker, however, can make inferences or form a belief.

2. A strategy for the decision maker assigns to each information set *one* distinct action which is executed whenever an information set is reached. The decision maker cannot plan to assign different actions to two histories which lie in the same information set.

3. If the decision maker assesses his strategy at an information set including more than one history, he forms beliefs about the histories which led to it. These beliefs are the basis for his considerations.

A decision problem exhibits imperfect recall if, at a point of time, the decision maker holds information which is forgotten later on. Specifically, an information set includes some histories which are incompatible with previously held information. Figures 1–3 are examples that illustrate three main aspects of imperfect recall in decision problems.

The standard motive of imperfect recall appears in Example 2. The decision maker is initially informed about the move of chance and loses

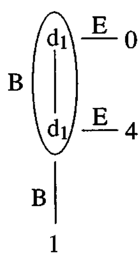

FIG. 1. Example 1.

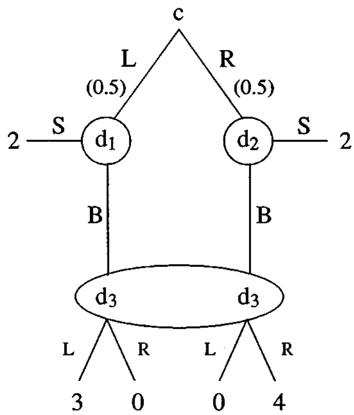

FIG. 2. Example 2.

this information upon reaching the information set d_3. In Example 3, the decision maker forgets his initial move at the information set d_2. Example 1 (which is central in this paper) is rather unconventional; the decision maker does not distinguish between the first and the second nodes. Reaching the second node, he loses the information that he had previously made a choice. The inability of a decision maker to distinguish between two histories on the same path will be referred to as *absentmindedness*.

In the above examples, the information sets determine the ability of the decision maker to recall. In practice, a decision maker can affect what he remembers. In this paper, however, we assume that the decision maker is not allowed to employ an external device to assist him in keeping track of the information which he would otherwise lose. Thus, we sidestep the decision maker's considerations regarding the trade-off between "more memory" and "memory costs."

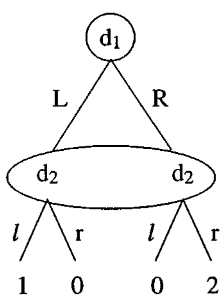

FIG. 3. Example 3.

The common interpretation of a decision problem with imperfect recall refers to a situation in which an individual takes several successive actions and faces memory constraints. Imperfect recall, however, is not necessarily related to the mental phenomenon of memory loss and may also reflect the imperfect ability to make the inferences necessary for distinguishing among different points in the same information set. A decision maker may not realize that he is at the 17th exit along a highway, either because he does not recall whether he has passed the 16th intersection, or because he cannot infer that he is at the 17th intersection, despite the perfect pictures of each intersection in his mind. The latter interpretation of imperfect recall brings our discussion closer to the topic of "bounded rationality."

An alternative interpretation is found in Isbell (1957). The decision maker is an organization consisting of many agents who have the same interests and act at different possible instances. In this case, the decision process of the organization may exhibit imperfect recall either because it must keep the instructions given to agents acting in successive situations simple or because of communication problems between agents. Agents receive instructions on how to behave and the collection of these instructions is equivalent to the notion of a strategy.

From a psychological point of view, imperfect recall is a very important phenomenon as it puts severe constraints on a decision maker's behavior. We believe, however, that the ultimate proof for its relevance in economic analysis can only come from an interesting model which clearly explains an economic phenomenon. Constructing such a model is beyond the confines of this paper.

The structure of the paper is as follows. We begin by presenting an example which we call the "*paradox of the absentminded driver*." This example is used to illustrate many of the points of this paper and will be referred to repeatedly.

After providing the formal definition of the model, we discuss two issues of importance for decision problems with absent-mindedness. First, in Section 4, we review the circumstances in which whether the decision maker is allowed to randomize affect the analysis. Second, in Section 5, we show that the extension of Bayesian updating to decision problems with absentmindedness is not trivial.

The two main topics which will be addressed are the timing of decision and the multiself approaches. The significance of these issues is marginal in decision problems with perfect recall since possible answers are inconsequential for the analysis.

(i) *Timing of decision*. This issue deals with the interpretation of an information set as either a point of decision or a point in which a strategy is executed. In Section 6, we will show that for decision problems with imperfect recall optimal strategies may be time inconsistent and that strategies which are time consistent may not be optimal. The timing of decisions can be an important consideration for a decision maker whereas, with perfect recall, there is no reason to make some decisions at a particular point of time.

(ii) *The multiselves approach to decision making*. Standard dynamic inconsistencies are generally addressed by assuming that the decision maker acts as a collection of distinct "selves" who behave independently. The behavior of the decision maker is analyzed as an equilibrium. In Section 7, we extend the multiself approach to decision problems with imperfect recall and show that in this extension an optimal strategy is dynamically consistent.

2. THE PARADOX OF THE ABSENTMINDED DRIVER

An individual is sitting late at night in a bar planning his midnight trip home. In order to get home he has to take the highway and get off at the second exit. Turning at the first exit leads into a disastrous area (payoff 0). Turning at the second exit yields the highest reward (payoff 4). If he continues beyond the second exit, he cannot go back and at the end of the highway he will find a motel where he can spend the night (payoff 1). The driver is absentminded and is aware of this fact. At an intersection, he cannot tell whether it is the first or the second intersection and he cannot remember how many he has passed (one can make the situation more realistic by referring to the 17th intersection). While sitting at the bar, all he can do is to decide whether or not to exit at an intersection. We exclude at this stage the possibility that the decision maker can include random elements in his strategy. Example 1 describes this situation.

Planning his trip at the bar, the decision maker must conclude that it is impossible for him to get home and that he should not exit when reaching an intersection. Thus, his optimal plan will lead him to spend the night at the motel and yield a payoff of 1. Now, suppose that he reaches an intersection. If he had decided to exit, he would have concluded that he is at the first intersection. Having chosen the strategy to continue, he concludes that he is at the first intersection with probability 1/2. Then, reviewing his plan, he finds that it is optimal for him to leave the highway since it yields an expected payoff of 2. Despite no new information and no change in his preferences, the decision maker would like to change his initial plan once he reaches an intersection! Note that if the decision maker now infers that he would have exited the highway had he passed the first intersection, his reasoning becomes circular; he must conclude that he is at the first intersection and that it is optimal to continue.

We wish to emphasize that this is not a standard example of time inconsistency. Usually, time inconsistency is obtained as a consequence of changes in either preferences (tastes) or information regarding the moves of nature during the execution of the optimal plan. Here, preferences over final outcome are constant and the only factor intervening between planning and execution of the optimal strategy is the occurrence of the situation which calls for execution, that is, reaching the intersection.

We find this example paradoxical as it exhibits a conflict between two ways of reasoning at the intersection. The first is based on a quite minimal principle of rationality; having chosen an optimal strategy, one does not have to verify its optimality at the time of execution unless there is a change in information or in preferences. In our example, the decision maker knew he would reach the intersection with certainty and his preferences are constant. This principle leads to the conclusion that the decision maker should stick to his plan to continue. The second way is based on the principle which calls at each instance to maximize expected payoffs given the relevant beliefs. In our example, this principle leads to the conclusion of exiting. The conflict between these two potential lines of reasoning is at the root of the apparent ambiguity of our example.

3. EXTENSIVE DECISION MODEL

In this section, a formal definition of the extensive decision model is given. The presentation follows that of Osborne and Rubinstein (1994). The reader can easily identify the model with the standard definition in the "tree" language.

A (finite) *decision problem* is a five-tuple $\Gamma = \langle H, u, C, \rho, I \rangle$, where:

(a) H is a finite set of *sequences*. We assume that the empty sequence, ϕ, is an element of H and that if $(a_1, \dots, a_K) \in H$ and $(a_1, \dots, a_K) \neq \phi$ then $(a_1, \dots, a_{K-1}) \in H$.

We interpret a history $(a_1, \dots, a_K) \in H$ as a feasible sequence of actions taken by the decision maker or by chance. The history $(a_1, \dots, a_K) \in H$ is *terminal* if there is no $(a_1, \dots, a_K, a) \in H$. The set of terminal histories is denoted by Z. The set of actions available to the decision maker or chance after a nonterminal history h is defined by $A(h) = \{a: (h,a) \in H\}$. To avoid degenerate cases we assume that $A(h)$ contains at least two elements. When presenting a decision problem diagramatically, we draw H as a tree whose nodes are the set of histories with root ϕ and whose edges combine a node $(a_1, \dots, a_K)$ with a node $(a_1, \dots, a_{K+1})$.

(b) u: $Z \to R$ is a utility function which assigns a number (payoff) to each of the terminal histories. Preferences are defined on the set of all lotteries over terminal histories and satisfy the VNM assumptions.

(c) C is a subset of H. We assume that the chance player moves after histories in C.

(d) ρ is the decision maker's belief about the chance player's behavior. ρ assigns to each history $h \in C$ a probability measure on $A(h)$. To avoid degeneracy, we assume that $\rho(h, a)$ is strictly positive for all $h \in C$ and $a \in A(h)$.

Thus, the set of histories H is partitioned into three subsets:

Z, the set of terminal histories;

C, the set of histories after which chance moves;

$D = H - Z - C$, the set of histories after which the decision maker moves.

(e) The set of information sets, which is denoted by I, is a partition of D. We assume that for all h, h' in the same cell of the partition $A(h) = A(h')$; i.e., the sets of actions available to the decision maker at histories in the same information set are identical. For convenience, with a slight abuse of notation, we will sometimes denote the set of actions which are available at a history in X by $A(X)$.

Note that, in contrast to some authors, we do not exclude from the class of decision problems those which exhibit absentmindedness (see definition below).

If all information sets in I are singletons we say that Γ is a decision problem with *perfect information.*

A (*pure*) *strategy*, f, is a function which assigns to every history $h \in D$ an element of $A(h)$ with the restriction that if h and h' are in the same information set $f(h) = f(h')$. Notice that this definition requires that the decision maker plans an action at histories which he will not reach if he follows the strategy.

We are now ready for the main definitions of this paper. The *experience* of the decision maker at a history h in D, denoted by $\exp(h)$, is the sequence of information sets and actions of the decision maker along the history h. We adopt the convention that the last element in the sequence $\exp(h)$ is the information set which contains h. Thus, in Example 1, $\exp(\phi) = (d_1)$ and $\exp(B) = (d_1, B, d_1)$.

A decision problem has *perfect recall* if for any two histories, $h, h' \in D$, which lie in the same information set, $\exp(h) = \exp(h')$. Thus, in a decision problem with perfect recall, the decision maker "remembers" the succession of the information sets he has faced and the actions he has taken. A decision problem for which the above condition is violated is referred to as a decision problem with *imperfect recall.*

Given a history $h = (a_1, \dots, a_K)$ and $L < K$, the history $h' = (a_1, \dots, a_L)$ is a *subhistory* of h. We say that a decision problem Γ exhibits *absentmindedness* if there are two histories h and h' such that h' is a subhistory of h and both belong to the same information set.

The decision problem illustrated in Example 1 exhibits absentmindedness since history (B) and its subhistory ϕ are in the same information set.

4. THE VALUE OF RANDOMIZATION

In this section, we discuss the implications of enlarging the strategy set of a decision maker to include random strategies. Given that the decision maker behaves as an expected utility maximizer, randomization over pure strategies is redundant for problems of either perfect or imperfect recall. Define a *mixed strategy* to be a probability distribution over the set of pure strategies. It describes a behavior in which randomization occurs only at the outset, before the decision problem unfolds. Each pure strategy induces a lottery over Z. A mixed strategy induces a lottery over Z which is the compound lottery of the lotteries induced by each of the pure strategies in its support. Therefore, no mixed strategy can be strictly preferred to all the pure strategies.

Behavior strategies perform a different method of randomization. A *behavioral strategy*, b, is a function which assigns to every history $h \in D$, a *distribution* $b(h)$ over $A(h)$ such that $b(h) = b(h')$ for any two histories h and h' which lie in the same information set. In decision problems without absentmindedness $b(h)$ is a lottery which is realized when the information set which contains h is reached. For decision problems with absentmindedness we take $b(h)$ to be a random device which is activated independently every time the information set which includes h is reached.

Consider again Example 1. In this problem there are two pure strategies, "B" and "E" which yield payoffs of 1 and 0, respectively. Although the absentminded driver cannot use a pure strategy to reach home with certainty, he can toss a coin and obtain an expected payoff of 1.25. Note that his optimal behavioral strategy is to exit with probability $\frac{1}{3}$ and yields the expected payoff of $\frac{4}{3}$.

It turns out that absentmindedness is necessary for behavioral strategy to be strictly optimal. This was shown in Isbell (1957) and we provide the proof for completeness.

PROPOSITION 1. *Suppose Γ does not exhibit absentmindedness. Then for any behavioral strategy there is a pure strategy which yields a payoff at least as high.*

Conversely, suppose $\Gamma = \langle H, u, C, \rho, I \rangle$ exhibits absentmindedness. Then, there exist a decision problem $\Gamma' = \langle H, u', C, \rho, I \rangle$ and a behavioral strategy which yields a payoff strictly higher than any payoff achieved by a pure strategy.

Proof. See the Appendix.

The Paradox of the Absentminded Driver Revisited. The inconsistency discussed in Section 2 is not a consequence of the restriction that the strategy set includes only pure strategies and persists when the decision maker is allowed to choose random actions. The optimal behavioral strategy is to choose B with probability $\frac{2}{3}$. Reaching the intersection, the driver will form beliefs about where he is. Denote by α the probability he assigns to being at the first intersection. Then, his expected payoff is $\alpha[p^2 + 4(1-p)p] + (1-\alpha)[p + 4(1-p)]$, where p is the probability of not exiting. The optimal p is now $\max\{0, (7\alpha - 3)/6\alpha\}$. This is inconsistent with his original plan unless $\alpha = 1$. In other words, his original plan is time consistent if and only if he believes that there is no chance he has passed the first intersection. We find such a belief unreasonable. Given his strategy it seems natural to assign to the second intersection a probability which is $\frac{2}{3}$ times the probability assigned to the first intersection, which implies $\alpha = 0.6$. The issue of consistent beliefs will be discussed in the next section.

This type of time inconsistency can appear also in decision problems in which the optimal strategy is pure. Consider the following example shown in Fig. 4.

The optimal behavioral strategy is the pure strategy which selects L at d_1 and L at d_2. To verify it, denote by α and β the probabilities of choosing L at d_1 and d_2, respectively, and note that $\alpha^2\beta + 3\alpha(1-\alpha)(1-\beta)$ is strictly less than 1 unless $\alpha = \beta = 1$. Upon reaching the information set d_1, if the decision maker concludes that he is at the two histories with equal probabilities, the strategy R at d_1 and R at d_2 yields the higher expected payoff of 1.5.

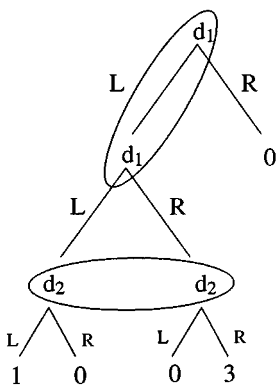

FIG. 4. Example 4.

5. CONSISTENT BELIEFS

If information sets are to be interpreted as points of decision, Examples 1 and 4 suggest that a decision maker who acts on the basis of expected utility maximization may be unable to execute the optimal strategy. The first step in addressing this issue is to specify the decision maker's beliefs at an information set which is not a singleton. As we shall see, finding an appropriate specification for decision problems with absentmindedness is not conceptually trivial.

We define a *belief* system as a function μ which assigns to any information set X and any history $h \in X$ a nonnegative number $\mu(h|X)$ such that $\sum_{h \in X} \mu(h|X) = 1$. The interpretation is that the decision maker, upon reaching X, assigns probability $\mu(h|X)$ to the possibility that he is at h. Let $p(h|h', b)$ be the probability that, conditional on reaching h', the history h will be realized when executing the strategy b. We denote $p(h|\phi,b)$ by $p(h|b)$.

Several alternatives are conceivable for the specification of the decision maker's beliefs. Since our objective is to examine the optimality of a strategy during its execution, we find it natural to assume that the beliefs of the decision maker are related in a systematic way with the strategy to be assessed. The condition that we require a belief system μ to satisfy to be consistent with a behavioral strategy b mirrors the frequency approach to belief formation. Namely, if an information set X is reached with positive probability, $\mu(h|X)$ is assumed to be equal to the long run proportion of times in which "visiting" the information set X involves being in h for a decision maker who plays the decision problem again and again and follows b.

DEFINITION. A belief system μ is *consistent* with the behavioral strategy b if for every information set X which is reached with positive probability and for every $h \in X$, $\mu(h|X) = p(h|b)/\sum_{h' \in X} p(h'|b)$.

Our definition of consistency imposes restrictions only on beliefs at information sets which are reached with positive probability. Notice that, for decision problems without absentmindedness, consistency is equivalent to Bayes' formula. For decision problem with absentmindedness, however, the denominator can be greater than one. The similarity with Bayes' formula is only notational.

To clarify this definition, consider first the absentminded driver example and the strategy that selects B with probability equal to $\frac{1}{2}$. A consistent belief, conditional upon the information set d_1, assigns probability $\frac{2}{3}$ to being at the first intersection. Consider next Example 5 (see Fig. 5) and the strategy which selects B with probability 1.

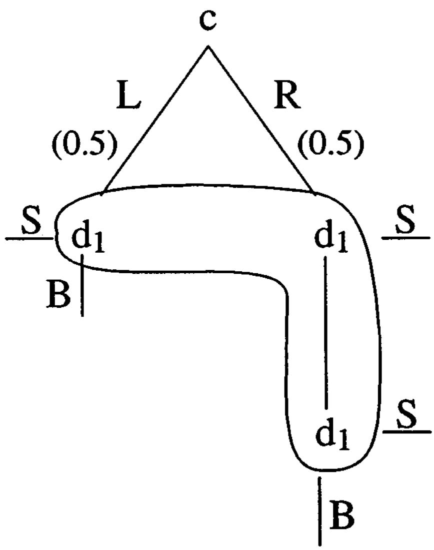

FIG. 5. Example 5.

Consistent beliefs at the information set d_1 assign probability $\frac{1}{3}$ to each history in the information set.

Our definition of consistent beliefs can also be motivated as being derived from a probability space which includes the time at which the decision maker can be. As an illustration consider Example 5 and assume that each action takes one unit of time. The relevant space of instances consists of $(L, 1)$, $(L, 2)$, $(R, 1)$, and $(R, 2)$, where (x, t) is the instance in which the chance player chooses x and time is t. Assuming equal probabilities for each instance is consistent with the description of the chance player in the decision problem. Then, a decision maker who is told that he is at d_1 updates his belief by the Bayesian formula. For example, the unconditional probability of the second node after R is $p/4$, where p is the probability of choosing B at d_1, and the unconditional probability of the node after R is $\frac{1}{4}$. If $p = 1$, the conditional probability of each of the three nodes in d_1 is $\frac{1}{3}$.

In this paper, we adopt the above definition of consistency. However, we find the following definition of consistency reasonable as well.

DEFINITION. A belief system μ is *Z-consistent* with the behavioral strategy b if for every information set X which is reached with positive probability and for every $h \in X$

$$\mu(h|X) = \frac{\sum_{\{z|h \text{ is a subhistory of } z\}} p(z|b)/\#\{h' \mid h' \in X \text{ and is a subhistory of } z\}}{\sum_{\{z|z \text{ has a subhistory in } X\}} p(z|b)}.$$

The rationale behind this definition is that, given the behavioral strategy b, the probability of the event that the random elements which determine the terminal history z are realized is $p(z|b)$; if the history z includes K histories in which the decision maker is asked to act then he assigns to each of them the *ex ante* probability $p(z|b)/K$.

To illustrate the Z-consistency, consider again the absentminded driver example and the strategy that selects B with probability $\frac{1}{2}$. A Z-consistent belief, conditional upon the information set d_1, assigns probability $\frac{3}{4}$ to being at the first intersection. In Example 5, if the behavioral strategy selects B with probability 1, Z-consistent beliefs at the information set d_1 assign probability $\frac{1}{2}$ to the history (L) and probability $\frac{1}{4}$ to each of the histories (R) and (R, B).

Note that the paradoxical flavor of the absentminded driver example is unaffected by the type of consistency we adopt. Also, if the decision problem does not exhibit absentmindedness then consistency and Z-consistency are identical.

We do not have a firm view about the "right" definition of consistent beliefs. The issue deserves further investigation. We wish to point out that if the decision maker adopts Z-consistency then he is exposed to a sort of "money pump." Consider Example 5 and the behavioral strategy that selects B with probability 1. Given Z-consistent beliefs a risk neutral decision maker, upon reaching the information set, will always accept an agreement in which he gains \$1.1 if he is at (L) and loses \$1 otherwise. If such an agreement is offered whenever the information set is reached, the resulting undesirable lottery yields the decision maker \$1.1 with probability 0.5 and $-$\$2 with probability 0.5.

6. TIME CONSISTENCY

When does the decision maker make his decision? Can he decide about the point of time at which a decision is made? To what extent can the decision maker commit to decisions he makes? These questions are superfluous for decision problems with perfect recall since the *ex ante* optimal strategy remains optimal during its execution. In the presence of imperfect recall, the optimal strategy may cease to be optimal along its execution and answers to these questions are significant for the analysis.

We say that a strategy is time consistent if there is not, at any information set which is reached as the decision problem unfolds, a different strategy for the remainder of the decision problem which yields a higher expected payoff. Notice that we require the optimality to be assessed only at the information sets and *not before the initial history*.

The significance of the statement that a strategy is optimal at an information set depends on the beliefs that the decision maker is assumed to hold. Requiring that beliefs be consistent leads to the following definition.

DEFINITION. A behavioral strategy b is *time consistent* if there is a belief μ consistent with b such that for every information set X which is reached with positive probability under b,

$$\sum_{h \in X} \mu(h) \sum_{z \in Z} p(z \mid h, b) u(z) \geq \sum_{h \in X} \mu(h) \sum_{z \in Z} p(z \mid h, b') u(z)$$

for any behavioral strategy b'.

The notion of time consistency that we use is a very strong criterion. We must emphasize that it does not rest on a model of how decisions are made at each information set. The definition presumes that the decision maker, in principle, can commit to a future course of action. Of course, if the decision maker is aware that a revision of the plan is possible at future information sets, expectations of commitments are naive.

Time consistency is an assessment as to whether, conditional upon the realization of information set, there exists an alternative rule of behavior which yields the decision maker a higher payoff. If an optimal strategy satisfies this criterion, the distinction with respect to information sets as points of decision or points of execution is inconsequential for a decision maker who is implementing it and whose beliefs are consistent. If, however, an optimal strategy fails this criterion, this distinction becomes significant and alternative specifications of the domain of choice and the timing of decision of the decision maker can give rise to different analyses.

A well-known result states that for decision problems with perfect recall a strategy is optimal if and only if it is time consistent. This is not the case for problems with imperfect recall. As we have seen, the optimal behavioral strategy for the absentminded driver is not time consistent. It is easy to see that the only time consistent strategy is to exit with probability $\frac{5}{9}$.

The problem of time consistency can also arise without absentmindedness as Example 2 shows. The optimal strategy is to choose S at d_1, B at d_2, and R at d_3. However, upon reaching d_1 the decision maker prefers using B at d_1 and L at d_3. A decision maker who can postpone his decisions until he has reached d_1 or d_2 can make a better decision regarding the choice at d_3 than a decision maker who decides *ex ante*. A complete model should indicate if the domain of choice of the decision maker includes timing of the decision and the extent of his ability to use his knowledge at d_1 and d_2 to influence his choice at d_3.

In the case of absentmindedness, the divergence between optimality and time consistency is two-sided. The only optimal strategy is not time consistent and the only time consistent strategy is not optimal. The next example (Fig. 6) shows that, even without absentmindedness, a strategy may be time consistent and not optimal.

Consider the strategy where the decision maker plays B at both information sets. The belief system consistent with this strategy assigns equal probabilities to each of the histories in each of the information sets. Therefore, this strategy is time consistent although the optimal strategy is to choose S at both information sets.

The next proposition provides conditions under which optimal and time consistent strategies are equivalent. Two histories are said to *split* at h if h is their longest common subhistory.

PROPOSITION 2. *Let* Γ *be a decision problem without absentmindedness for which*

for any information set X *and two histories* $h', h'' \in X$ *which split at* $h \in C$, *the information sets which appear in* $\exp(h')$ *are the same as in* $\exp(h'')$. $(*)$

A behavioral strategy for Γ *is optimal if and only if it is time consistent.*

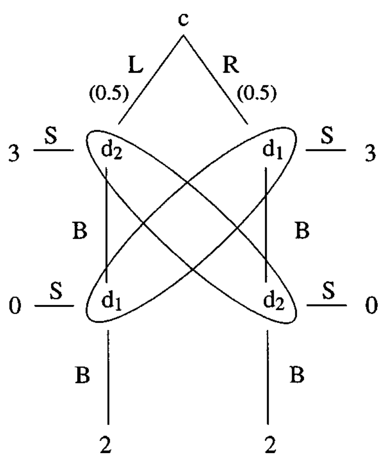

FIG. 6. Example 6.

Proof. See the Appendix.

To clarify the meaning of condition $(*)$ in Proposition 2, consider first Example 2. Both histories (L, B) and (R, B) are in d_3, but $\exp(L, B) = (d_1, B, d_3)$ and $\exp(R, B) = (d_2, B, d_3)$. $(*)$ is violated, since at d_3 the decision maker loses information about the move of chance. In Example 6, even though at no information set the decision maker knows nature's move, $(*)$ is violated since $\exp(L) = (L, d_2)$ and $\exp(R, B) = (R, d_1, B, d_2)$. Notice that, however, if the decision maker had beliefs consistent with the strategy assigning B to d_1 and S to d_2, he would assign probability one to nature's move R immediately after R and probability $\frac{1}{2}$ to both of nature's moves when R is followed by his action B.

7. THE MULTISELF APPROACHES

For standard dynamic inconsistencies, Strotz (1956) provides a framework of analysis in which every information set is assumed to be a point of decision and the decision maker is unable to control his behavior at future information sets. A decision maker acts as a collection of hypothetical agents (selves) whose plans form an equilibrium. More precisely, for any decision problem Γ define $G(\Gamma)$ to be the extensive game in which each information set of Γ is assigned a distinct player, and all players have the same payoff as the decision maker. The behavior of the decision maker in Γ is then analyzed as an equilibrium of $G(\Gamma)$. As is well known, any optimal play for Γ is the play induced by some subgame perfect equilibrium of $G(\Gamma)$, and any subgame perfect equilibrium of $G(\Gamma)$ corresponds to an optimal strategy. This property has the consequence that the backward induction algorithm is a procedure for solving a decision problem with perfect information. An analogous result is valid for decision problems with perfect recall and imperfect information. In this case we use the solution concept of sequential equilibrium which combines sequential rationality with the requirement of consistent beliefs. The set of distributions over the terminal nodes generated by the sequential equilibria of $G(\Gamma)$ is identical to the set of distributions generated by the optimal strategies of Γ (see Hendon, Jacobsen, and Sloth, 1993).

These results are called the "no single improvement" property; a strategy is optimal if, conditional upon reaching an information set, the decision maker cannot improve his expected payoff by changing only the action prescribed by the strategy for that information set. It implies that it makes no difference whether we model a decision problem piecemeal or in its entirety. The equivalence of the single-self and the multiselves approaches for decision problems with perfect recall breaks down when we analyze decision problems with imperfect recall. Consider Example 3. The optimal strategy (R, r) corresponds to a sequential equilibrium in the corresponding multiself game. However, the multiself approach does not rule out the inferior strategy (L, l); the decision maker cannot improve his payoff by simply changing the action prescribed at a single information set. Conditional upon being at d_1, the choice of L is optimal if the decision maker treats his behavior at d_2 as given and unchangeable and believes he will play l.

The Paradox of the Absentminded Driver Re-revisited. Consider again the paradox of the absentminded driver. The formal notion of sequential equilibrium is not applicable to decision problems with absentmindedness. However, the requirement that the behavioral strategy at each information set is optimal, given the consistent beliefs, is meaningful. Despite the fact that there is only one information set, this requirement generates a strong dependence between behavioral strategies and beliefs. As we have noted before, the only strategy which satisfies this requirement is to exit with probability $\frac{5}{9}$, whereas the strategy that maximizes expected payoffs *ex ante* is to exit with probability $\frac{1}{3}$. Nevertheless, one can modify the multiself approach to avoid the possibility that an optimal strategy does not satisfy sequential rationality with respect to consistent beliefs.

Beliefs consistent with the optimal behavioral strategy in the paradox of the absentminded driver assign probability $\frac{3}{5}$ to the first intersection. If the decision maker anticipates that his "twin-self," if asked to move again, will exit with probability $\frac{1}{3}$, then it is optimal for him to exit with probability $\frac{1}{3}$ as he is indifferent between E (yielding the expected payoff $\frac{3}{5}[0] + \frac{2}{5}[4] = \frac{8}{5}$) and B (yielding an expected payoff of $\frac{3}{5}[(\frac{1}{3})4 + (\frac{2}{3})1] + \frac{2}{5}[1] = \frac{8}{5}$). In this modification, when considering a deviation at an information set X from the behavioral strategy assigned to X, the "self" assumes that his "twin-self" will use the equilibrium strategy if asked again to act, *regardless* of the choice he now makes. The requirement that the decision maker uses identical randomization at all histories in the same information set is retained only in equilibrium. In an actual play of the decision problem, if he decides to deviate, he may use different randomizations at different histories in the same information set.

Also notice that, for the optimal strategy to satisfy sequential rationality in this modification of the multiself approach, the decision maker must believe that he is at the first intersection with probability $\frac{3}{5}$. In particular, if one replaces our definition of consistency of beliefs with Z-consistency, sequential rationality fails in this modification as well. We now define the "modified multiself" approach formally and prove that every optimal strategy is modified multiself consistent (discussions with Bob Aumann and Roger Myerson were very helpful in shaping this part of the paper).

DEFINITION. A behavioral strategy b is *modified multiself consistent* if there exists a belief μ consistent with b such that for every information set X which is reached with positive probability and for every action $a \in A(X)$ for which $b(h)(a) > 0$ for $h \in X$, there is no $a' \in A(X)$ such that

$$\sum_{h \in X} \mu(h) \sum_{z \in Z} P(z \mid (h, a'), b) u(z) > \sum_{h \in X} \mu(h) \sum_{z \in Z} P(z \mid (h, a), b) u(z).$$

Note that in contrast with the approach discussed in the first part of this section, the above definition imposes requirements only on the selves which are reached with positive probability.

PROPOSITION 3. *If a behavioral strategy is optimal then it is modified multiself consistent.*

Proof. See the Appendix.

Proposition 3 demonstrates that, within the framework of the "modified" multiself approach, the inconsistency of optimal plans for decision problems with absentmindedness evaporates. The interpretative ambiguities originating in the paradox of the absentminded driver also vanish as the strategy to exit with probability $\frac{2}{3}$ is singled out as the only optimal and consistent strategy. The value of this resolution, however, depends on the appropriateness of the "modified" multiself approach as a theory of decision making under imperfect recall. The behavioral assumption of the multiself approach is that each self maximizes his conditional expected payoff given his own beliefs. In decision problems with perfect recall, beliefs at each information set are independent of the choices of previous selves and each self can determine the optimal decisions of future selves by recursion. In decision problems with imperfect recall, however, the conditional payoff of each self depends on his beliefs about the behavior of earlier selves and the independence of beliefs and actions from previous choices is an additional assumption. For decision problems with absentmindedness, the "modified" multiself approach *assumes* that a decision maker, upon reaching an information set, takes his actions to be immutable at future occurrences of that information set, no matter which course of action he is contemplating now. At the other extreme one finds the opposite axiom for which only one self resides in the information set and expects that, were the information set to occur again, he would adopt whichever behavioral rule he adopts now. In this case, the strategy to exit with probability $\frac{5}{9}$ would be the consistent rule of behavior for Example 1. These two assumptions reflect two alternative ways of reasoning at the information set. We find both of them to have some intuitive appeal and neither to be universally valid.

8. FINAL COMMENTS

Our observations are only a limited exercise intended to suggest some of the issues which a comprehensive theory of imperfect recall must confront. We conclude our discussion with the following three comments:

a. *Other types of imperfect recall.* We model imperfect recall by including in the same information set histories which contain different experiences. Such histories are assumed to be indistinguishable for the decision maker. A different type of imperfect recall is represented by a situation in which "I know that at the first intersection I would be aware of this but at the second intersection I do not know whether I am at the first or the second intersection." One way to model such a scenario is by nonpartitional informational structures (see Geanakopolos, 1990). Rubinstein (1991) discussed the difficulty of modelling considerations such as "At the first intersection I know that I am there but if I reach the second intersection I think that there is a 10% chance that I am at the first intersection" (see also Fluck, 1994). It seems to us that new analytical frameworks are needed to address these issues as well as issues such as partial recall of strategy and imperfect ability to make inferences.

b. *Games with imperfect recall.* This paper focuses on issues related to modeling decision problems with imperfect recall. Obviously, such issues carry over to extensive games. Isbell and Kuhn show that the nonequivalence between mixed and behavioral strategies can cause nonexistence of Nash equilibria in behavioral strategies. Time inconsistency of optimal strategies for some decision problems with imperfect recall makes it possible to construct games with imperfect recall that have no equilibria which satisfy sequential rationality with respect to consistent beliefs.

Ambiguities in the interpretation of games with imperfect recall is probably the reason why only a handful of papers have dealt with the topic. The few exceptions include Isbell (1957) and Alpern (1988), who prove the existence of mixed behavioral strategy equilibrium in a class of games with imperfect recall, and Binmore (1992) who relates imperfect recall and nonpartitional knowledge. The literature on machines playing repeated games (see, for example, Rubinstein, 1986) may also be interpreted as an analysis of a class of repeated games with imperfect recall, where it is costly for the players to distinguish between different histories of the game. Further investigation is needed to adapt solution concepts designed for games with perfect recall to games with imperfect recall.

c. *The paradox of the absentminded driver*. In all its forms the absentminded driver example exhibits a conflict between two types of reasoning. Commitment to the *ex ante* optimal behavioral strategy is obviously the normative rule of behavior.

We do not have a firm view about the resolution of the paradox. We have investigated one resolution which requires dividing a decision maker into multiple independent selves. Another resolution would entail the rejection of expected utility maximization given consistent beliefs when the information set includes histories whose probabilities depend on the decision maker's actions at that information set. Savage's theory views a state as a description of a scenario which is independent of the act. In contrast, "being at the second intersection" is a state which is not independent from the action taken at the first, and, consequently, at the second intersection.

APPENDIX

Proof of Proposition 1. Consider an information set X and $h \in X$. Let b be a behavioral strategy and $b(h)(a)$ denote the probability of choosing $a \in A(h)$ at h. Since Γ does not exhibit absentmindedness, the expected payoff from b can be written as $\sum_{a \in A(h)} b(h)(a) g_a + g$, where each coefficient g_a and g are independent of $b(h)$. Consider $a* \in A(h)$ which achieves the highest g_a. Setting $b(h)(a*) = 1$ yields a payoff at least as high. The claim follows by repeating the argument for every information set.

Suppose that Γ exhibits absentmindedness. Then, there exist a final history $\bar{z}$ and two distinct subhistories of $\bar{z}$, (h, a) and (h', a'), such that h and h' belong to the same information set and $a \neq a'$. Consider a payoff function u' such that $u'(\bar{z}) = 1$ and $u'(z) = 0$ for $z \neq \bar{z}$. The claim follows, since $\bar{z}$ cannot be reached by any pure strategy, and is reached with positive probability by any behavioral strategy which assigns a positive probability to each action. Q.E.D.

Proof of Proposition 2. Let $\pi(b)$ denote the payoff that the decision maker obtains by playing b. Given an information set X, let $Z(X)$ denote the set of final histories having subhistories in X. Then, we can write $\pi(b)$ as

$$\pi(b) = \sum_{z \in Z/Z(X)} p(z \mid b) u(z) + \sum_{z \in Z(X)} p(z|b) u(z).$$

Since Γ does not exhibit absentmindedness, a terminal history z in $Z(X)$ has a unique subhistory h in X. Thus, $p(z|b) = p(h|b) p(z|h, b)$ and

$p(z|h', b) = 0$ for $h' \neq h$, $h' \in X$. Therefore,

$$\pi(b) = \sum_{z \in Z/Z(X)} p(z|b)u(z) + \sum_{h \in X} p(h|b) \sum_{z \in Z(X)} p(z \mid h, b)u(z).$$

Suppose b is optimal and not time consistent. Then, there is an information set X and a behavioral strategy b' such that X is reached with positive probability under b and

$$\sum_{h \in X} \mu(h) \sum_{z \in Z} p(z \mid h, b)u(z) < \sum_{h \in X} \mu(h) \sum_{z \in Z} p(z \mid h, b')u(z),$$

where μ is consistent with b. Then, by the definition of consistent beliefs,

$$\pi(b) < \sum_{z \in Z/Z(X)} p(z \mid b)u(z) + \sum_{h \in X} p(h \mid b) \sum_{z \in Z(X)} p(z \mid h, b')u(z).$$

Denote the right-hand side of the above inequality by π'. Define $H(X)$ to be the union of X and set of histories having subhistories in X. Consider a decision problem Γ' obtained from Γ by "splitting" any information set Y containing histories in both $H(X)$ and $(H - H(X))$ into $Y \cap H(X)$ and $Y \cap (H - H(X))$. Define $\bar{b}$ to be a strategy for Γ' which is identical to b' at X and at any information sets after X and identical to b otherwise. By construction, $\bar{b}$ yields π'. Then, by Proposition 1, there exists a pure strategy f' which yields in Γ' a payoff of at least π'. We now show that, by suitably modifying f', one can obtain a strategy f for Γ which yields the same payoff as f' for Γ'. The construction of f easily follows if we show that, for any information set Y for Γ which has been split into Y' and Y'' in Γ', it is not possible that both Y' and Y'' are reached with positive probability under f'. Suppose that $h' \in Y'$ and $h'' \in Y''$ are both reached with positive probability. By construction, one and only one of $\exp(h')$ and $\exp(h'')$ lists X. Denoting the longest history common subhistory of h' and h'' by $\tilde{h}$, $(*)$ implies that $\tilde{h} \notin C$. A contradiction is then obtained since f' is a pure strategy and assigns positive probability to at most one edge at $\tilde{h}$.

Thus, we can construct a strategy for Γ which yields at least π'. This contradicts the optimality of b.

Now suppose that b is time consistent. First notice that if $\phi \in D$, no absentmindedness implies that b is optimal. Suppose now that $\phi \in C$ and let Δ be the set of information sets which contain some histories with no subhistories in D. We first show that $X \in \Delta$ implies that no history in X has a subhistory in D. Suppose not and let h be a history in $X \in \Delta$ with a subhistory h' in the information set X'. By construction, $\exp(h)$ lists X' and, by no absentmindedness, $X \neq X'$. $(*)$ then yields a contradiction since, by definition, X contains a history with no subhistory in D. Hence,

$$\pi(b) = \sum_{X \in \Delta} \sum_{h \in X} \left[p(h|b) \sum_{z \in Z(X)} p(z \mid h, b) u(z) \right] + \text{constant}$$

and $p(h|b)$ is independent of b since no action of the decision maker precedes information sets in Δ. Since b is time consistent, b maximizes

$$\sum_{h \in X} p(h|b) \left[\sum_{z \in Z(X)} p(z \mid h, b) u(z) \right]$$

for any $X \in \Delta$ which is reached with positive probability (i.e., $\sum_{h \in X} p(h|b) > 0$). The claim follows. Q.E.D.

Proof of Proposition 3. Let b^* be a strategy and suppose that there exist a belief μ^* consistent with b^*, an information set X which is reached with positive probability, $a \in A(X)$ for which $b^*(h)(a) > 0$ for $h \in X$ and $a' \in A(X)$ such that

$$\sum_{h \in X} \mu^*(h) \sum_{z \in Z} p(z \mid (h, a'), b^*) u(z) > \sum_{h \in X} \mu^*(h) \sum_{z \in Z} p(z \mid (h, a), b^*) u(z).$$

We will show that b^* is not optimal. Let b_ϵ be a behavioral strategy which is identical to b^* except for $b_\epsilon(h)(a) = b^*(h)(a) - \epsilon$ and $b_\epsilon(h)(a') = b^*(h)(a') + \epsilon$. Let $\pi(b)$ denote the expected payoff of playing b. We claim that for sufficiently small positive ϵ, $\pi(b_\epsilon) > \pi(b^*)$.

Note that $\pi(b)$ is defined by a polynomial in the probabilities assigned by the strategy b to all possible actions. Fix all the probabilities at the levels assigned by b^*, except for the actions a and a' at the information set X, and denote those probabilities by p and p', respectively. We use the notation $b^*(X)$ to denote $b^*(h)$, for any $h \in X$. We shall assume that $b^*(X)(a') > 0$. The reader will easily see that this assumption keeps the notation simple and does not affect the argument of the proof.

We obtain a polynomial

$$V(p, p') = \sum_{z \in Z} u(z) p^{\delta(z)} p'^{\delta'(z)} p(z \mid b^*) / \left[b^*(X)(a)^{\delta(z)} b^*(X)(a')^{\delta'(z)} \right],$$

where $\delta(z)$ is the number of times that the terminal history z includes a play of a at X (and similarly for $\delta'(z)$).

It is enough to show that at the point $(p, p') = (b^*(X)(a), b^*(X)(a'))$ we have $dV(p, p')/dp < dV(p, p')/dp'$.

Since μ^* is consistent with b^* we have that $\mu^*(h) = p(h \mid b^*)/$

$\sum_{h' \in X} p(h' \mid b^*)$. Thus, it is sufficient to verify that at the point $(b^*(X)(a), b^*(X)(a'))$

$$dV(p, p')/dp' = \sum_{z \in Z} \delta'(z)[p(z \mid b^*)/b^*(X)(a')])u(z)$$
$$= \sum_{h \in X} [p(h \mid b^*)] \sum_{z \in Z} p(z \mid (h, a'), b^*)u(z)$$

and similarly for the partial derivative with respect to p. Q.E.D.

REFERENCES

Alpern, S. (1988). ''Games with Repeated Decisions,'' *SIAM J. Control Optim.* **26**, No. 2, 468–477.

Binmore, K. (1992). *Fun and Games*. Boston: Heath.

Fluck, Z. (1994). ''Games with Forgetfulness,'' mimeo.

Geanakopolos, J. (1990). ''Game Theory without Partitions, and Applications to Speculation and Consensus,'' Santa-Fe Institute, DP. 90-018.

Hendon, E., Jacobsen, J., and Sloth, B. (1993). ''The One-Shot-Deviation Principle for Sequential Rationality,'' *Games Econ. Beh.*, to appear.

Isbell, J. (1957). ''Finitary Games,'' in *Contributions to the Theory of Games III*, pp. 79–96. Princeton, NJ: Princeton Univ. Press.

Kuhn, H. W. (1953). ''Extensive Games and the Problem of Information,'' in *Contributions to the Theory of Games II*. Princeton, NJ: Princeton Univ. Press.

Osborne, M., and Rubinstein, A. (1994). *A Course in Game Theory*, Cambridge, MA: MIT Press.

Rubinstein, A. (1986). ''Finite Automata Play a Repeated Prisioner's Dilemma,'' *J. Econ. Theory* **39,** 83–96.

Rubinstein, A. (1991). ''Comments on the Interpretation of Game Theory,'' *Econometrica* **59,** 909–924.

Strotz, R. H. (1956). ''Myopia and Inconsistency in Dynamic Utility Maximization,'' *Rev. Econ. Stud.* **23,** 165–180.