N 70 28614

NASA CR 110184

PURDUE UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ON BEHAVIOR STRATEGY SOLUTIONS OF FINITE TWO-PERSON CONSTANT-SUM EXTENDED GAMES

V. A. Utgoff
R. L. Kashyap

TR-EE 70-9
February, 1970

Lafayette, Indiana 47907

ON BEHAVIOR STRATEGY SOLUTIONS OF FINITE TWO-PERSON CONSTANT-SUM EXTENDED GAMES

V. A. Utgoff

R. L. Kashyap

TR-EE 70-9

February, 1970

School of Electrical Engineering
Purdue University
Lafayette, Indiana 47907

:his work was partially supported by NASA under Grant NGR 15-005-021,
›y NSF under Grant GK-1970, and by the AFOSR under Grant 69-1776.

ERRATA

"On Behavior Strategy Solutions of Finite Two-Person Constant-Sum Extended Games"

Page No.	Line	Correction
3	36	$P_i \quad A_j \rightarrow P_i \cap A_j$
4	1	$P_o \quad A_j \rightarrow P_o \cap A_j$
5		Add under Assumption 2, "*Assumption 2A:* The vertices contained in any information set U, are of the same rank.
6	5	$U \in p^{\alpha} \rightarrow U \in P^{\alpha}$
6	5	$U \in p^{\beta} \rightarrow U \in P^{\beta}$
10	1	$D \rightarrow \varphi$
12	1	$\Gamma \rightarrow \Gamma^1$
16	13	set → point
17	17	$p(1 \mid \beta(I)) \rightarrow p(2 \mid \beta I(1))$
Figure 1		information set $\beta I(1)$, right-hand vertex → reverse the numbering of the alternatives
33	2	$]] \Sigma \rightarrow]] \div \Sigma$

ABSTRACT

Through the introduction of a concept called recall-sensitivity, it is demonstrated that perfect recall (each player remembers all of its past actions and past knowledge of the other player's and nature's actions) is a sufficient but not a necessary condition for the existence of behavior strategy solutions and ϵ-solutions in finite two-person constant-sum extended games.

A method is presented by which behavior strategies meeting a necessary condition for solutions or ϵ-solutions may be generated.

Comments are made on the practical implications of the material presented.

1.0 Introduction

Consider the three basic types of strategies that may be employed by the players in a finite two-person constant-sum extended game, specifically, pure strategies, mixed strategies, and behavior strategies. If we define the solution of such a game to be any set of strategies such that each player's strategy guarantees it the value of the game, we can make the following statements regarding these strategy types.

(i) In some games, pure strategy solutions, which are the simplest to store and implement, will not exist.

(ii) In every game, a mixed strategy solution will exist, but it may be vastly more complicated to store and implement than a pure strategy.

(iii) In many games, behavior strategy solutions, which are only slightly more complicated to store and implement than pure strategies, will exist.

In reference 1, Kuhn shows that a behavior strategy solution always exists in any game of perfect recall; i.e., in any game in which each player remembers at each move all of its past actions, and its past knowledge of the other player's and nature's actions.

In reference 2, a method is given for finding behavior strategy solutions in games of perfect recall.

In this paper, we intend to show that behavior strategy solutions, or ϵ-solutions (to be defined), also exist in some games in which the players do not have perfect recall. Our argument will take the following form.

(i) For every finite two-person constant-sum extended game there is an associated 2N-person non-cooperative game such that any set of equilibrium behavior strategies in the associated game corresponds to a pair of behavior strategies in the original game, which meets a necessary condition for a solution

(ii) If a player employs its strategy in a pair of behavior strategies which meets the above necessary condition, the difference between the expected return it is guaranteed to receive and the value of the game, is bounded by the sum of both players' recall-sensitivities about the given pair of behavior strategies (to be defined†).

(iii) The recall-sensitivity of any player having perfect recall is zero about any pair of behavior strategies, but a player may also have a recall-sensitivity of small or zero value about a pair of behavior strategies even in cases where it does not have perfect recall.

In the final sections of the paper, we will give a heuristic algorithm for generating equilibrium behavior strategies in the associated game, and we will comment on the practical implications of the above observations.

We will now give a formal description of the type of game that we are going to consider and establish some definitions and notation.

---

†Roughly, a measure of the influence of the forgotten information on the expected return.

1.1 Description of the Class of Games Considered

In reference 1, the author describes a class of games that he calls general n-person games. We are interested in a subclass of this general class, which we shall deliniate by repetition of the general definition, and statement of several restricting assumptions.

Definition 1K: A game tree K is a finite tree with a distinguished vertex 0 which is imbedded in an oriented plane.

Terminology: The alternatives at a vertex $X \in K$ are the edges e incident at X and lying in components of K which do not contain 0 if we cut K at X. If there are j alternatives at X, then they are indexed by the integers $1, \ldots, j$, circling X in the positive sense of the orientation. At the vertex 0, the first alternative may be assigned arbitrarily. If one circles a vertex $X \neq 0$ in the positive sense, the first alternative follows the unique edge at X which is not an alternative. The function thus defined which indexes the alternatives in K will be denoted by ν, thus $\nu(i)$ is the index of alternative i. Those vertices which possess alternatives will be called moves; the remaining vertices will be called plays. The name play will also be used for the unique unicursal path from 0 to a play when no confusion will result. The partition of the moves into sets A_j, $j=1,2,\ldots$, where A_j contains all of the moves with j alternatives will be called the alternative partition. A temporal order on K is defined by $X \leq Y$ if X lies on W_Y, the unicursal path joining 0 to Y; it is a partial order. The rank of a move Y is the number of X such that $X \leq Y$ as, equivalently, the number of moves $X \in W_Y$.

Definition 2K: A general n-person game Γ is a game tree K with the following specifications:

(i) A partition of the moves into n+1 indexed sets $P_o, P_1, \ldots, P_n$ which will be called the player partition. The moves in P_o will be called chance moves; the moves in P_i will be called personal moves of player i for $i=1,2,\ldots,n$.

(ii) A partition of the moves into sets U which is a refinement of the player and alternative partitions (that is, each U is contained in P_i A_j for some i and j) and such that no U contains two moves on the same play. This partition is called the information partition and its sets will be called information sets.

(iii) For each $U \subset P_o$ A_j, a probability distribution on the integers $1, \ldots, j$ which assigns positive probability to each. Such information sets are assumed to be one element sets.

(iv) An n-tuple of real numbers $h(W) = (h_1(W), h_2(W), \ldots, h_n(W))$ for each play W. The function h will be called the payoff function.

To interpret the above formal scheme, we may imagine a number of people called agents isolated from each other and each in possession of the rules of the game. There is one agent for each information set, and they are grouped into players in the natural manner, an agent belonging to the i^{th} player if his information set lies in P_i. This seeming plethora of agents is occasioned by the possibly complicated state of information of our players who may be forced by the rules of the game to forget facts which they knew earlier in a play.

A play begins at the vertex 0. Suppose that it has progressed to the move X. If X is a personal move with j alternatives, then the agent whose information set contains X chooses a positive integer not greater than j, knowing only that he is choosing an alternative at one of the moves in his information set. If X is a chance move, then an alternative is chosen in accordance with the probabilities specified by (iii) for the information set containing X. In this manner, a path with initial point 0 is constructed. It is unicursal and since K is finite, leads to a unique play W. At this point, player i is paid the amount $h_i(W)$ for $i=1, \ldots, n$.

The subclass of games that will be considered is defined with respect to the class of games described above, by means of the following assumptions.

Assumption 1: The number of players is two, and the payoff function satisfies the relations: (a) $h_1(W) + h_2(W) =$ constant; (b) $h_1(W) \geq 0$ and (c) $h_2(W) \geq 0$; for all possible W.

For convenience in what follows, we shall denote player 1 as the α-player and player 2 as the β-player. Any functions or variables associated with the α-player, β-player, or nature, will be superscripted by α, β or η, respectively.

Further, wherever a statement is made regarding the α-player, the corresponding modification needed to make the statement applicable to the β-player will be given in square brackets, which will be employed in the verbal portions of the text only for this purpose.

Assumption 2: If a move of rank k on a play W belongs to the α-player [β-player], a move of rank k on any other play W' also belongs to the α-player [β-player].

With the above assumptions then, we have defined the subclass of games that we shall consider. In order to begin our argument, we shall also need to establish the following definitions and additional notation.

1.2 Preliminary Definitions and Notation

Definition 1: The quality of a given strategy for the α-player [β-player] is the expected payoff to the player, $h^\alpha = \sum_W h_1(W)p(W)$, [$h^\beta = \sum_W h_2(W)p(W)$] assuming that:

(i) the α-player [β-player] uses the given strategy;

(ii) the β-player [α-player] selects one of the strategies open to it which minimizes h^α [h^β] knowing the α-player's [β-player's] chosen strategy.

With respect to the above definition, note that the value of the game is the maximum quality value that any strategy can have for either player.

Denoting the value of the game by v, we formally define solution and ϵ-solution.

Definition 2: Any pair of behavior strategies such that the α-player's strategy and the β-player's strategy are both of quality v, will be called a solution.

<u>Definition 3</u>: Any pair of behavior strategies such that the quality of the α-player's strategy is in the interval $(v-\epsilon, v)$, and such that the quality of the β-player's strategy is in the interval $(v, v+\epsilon)$, will be called an ϵ-solution.

<u>Definition 4</u>: Let $\mathfrak{U}^{\alpha} = \{U \in p^{\alpha}\}$ and $\mathfrak{U}^{\beta} = \{U \in p^{\beta}\}$. A <u>behavior strategy</u> for the α-player [β-player] is a function θ^{α} [θ^{β}] mapping $\mathfrak{U}^{\alpha}$ [$\mathfrak{U}^{\beta}$] into a collection of probability distributions such that for each $U \in \mathfrak{U}^{\alpha}$ [$\mathfrak{U}^{\beta}$], where $U \subset A_j$, $\theta^{\alpha}(U)$ [$\theta^{\beta}(U)$] is a probability distribution on the integers $1, \ldots, j$.

<u>Definition 5</u>: The α-player [β-player] is said to have <u>perfect recall</u> if for any k and any $\ell > k$, no two plays W and W′ pass thru a common information set at the α-player's [β-player's] ℓ^{th} move given that they have been distinguished from each other at the α-player's [β-player's] k^{th} move by virtue of:

(i) different alternative choices at an information set passed thru by both W and W′, or

(ii) passage thru different information sets.

(Note that perfect recall is not the same thing as perfect information.)

Noting that any play W is established as a string of alternative choices by nature and the two players, we establish the following notation. <u>For any play W</u>:

(i) denote by $\alpha E(k)$, $\beta E(k)$, or $\eta E(k)$, the alternative choice on W that is established by the α-player, β-player, or nature at its k^{th} move, and

(ii) denote by $\alpha I(k)$, $\beta I(k)$, or $\eta I(k)$, the information set containing the vertex on W corresponding to the k^{th} move of the α-player, β-player, or nature.

Further, we assume an arbitrary ordering on the information sets for each player at each of its moves. For example, let $\alpha I(k)_i$ denote the i^{th} possible information set for the α-player at its k^{th} move.

With these definitions and notations we can begin our argument; whatever else we need we will introduce as we go along.

2.0 Description of the Associated Game

In describing the associated game, and explaining the algorithm that can be employed to "solve" it, we shall find it convenient to express the expected return to the players in the particular form developed below.

Proceeding, the expected return to the α-player in the original game is first written as:

$$H = \sum_{W \in \Gamma} h_1(W)\, p(W) \tag{1}$$

where p(W) is the probability distribution defining the probabilities of the possible plays $W \in K$.

At this point, we may assume without loss of generality, that each of the players and nature make an equal number of moves N on every play W.† Under this assumption, (1) can be rewritten as:

†If it is not so, we can always imbed the original game in an appropriately defined larger game which has the desired property.

$$H = \sum_{C_W} p(\alpha E(i), \beta E(i), \eta E(i);\ i=1,\ldots,N)\ h_1(\alpha E(i), \beta E(i), \eta E(i);\ i=1,\ldots,N) \tag{2}$$

where C_W is the set of all choice sequences which constitute plays in K. Noting that use of behavior strategies means that the players and nature each make conditionally independent choices at every move, we can apply Bayes rule to (2) to obtain:

$$H = \sum_{C_W} p(\eta E(1)|\eta I(1))\ p(\alpha E(1)|\alpha I(1))\ p(\beta E(1)|\beta I(1)) \ldots p(\eta E(N)|\eta I(N))\ p(\alpha E(N)|\alpha I(N))\ p(\beta E(N)|\beta I(N))\ h_1(\eta E(1), \alpha E(1), \beta E(1), \ldots, \eta E(N), \beta E(N)) \tag{3}$$

Noting again that there are finitely many choices and finitely many information sets, we can construct from $h_1(\alpha E(i), \beta E(i), \eta E(i);\ i=1,\ldots,N)$ a function $g(\alpha E(i), \alpha I(i), \beta E(i), \beta I(i), \eta E(i), \eta I(i);\ i=1,\ldots,N)$ such that H can be written in the following form:

$$H = \sum_{\alpha I(1)} \sum_{\alpha E(1)|\alpha I(1)} p(\alpha E(1)|\alpha I(1)) \ldots \sum_{\alpha I(N)} \sum_{\alpha E(N)|\alpha I(N)}$$

$$\cdot\ p(\alpha E(N)|\alpha I(N)) \sum_{\beta I(1)} \sum_{\beta E(1)|\beta I(1)} p(\beta E(1)|\beta I(1)) \ldots$$

$$\cdot \sum_{\beta I(N)} \sum_{\beta E(N)|\beta I(N)} p(\beta E(N)|\beta I(N)) \Big(\sum_{\eta I(1)} \sum_{\eta E(1)|\eta I(1)}$$

$$\cdot\ p(\eta E(1)|\eta I(1)) \ \cdots \sum_{\eta I(N)} \ \sum_{\eta E(N)|\eta I(N)} p(\eta E(N)|\eta I(N))$$

$$\cdot\ g(\eta E(1), \eta I(1), \alpha E(1), \alpha I(1), \ldots, \beta E(N), \beta I(N))\Big) \qquad (4)$$

where the notation $\sum_{x|y}$ means summation over the set of choices x that are alternatives in information set y.

Carrying out the summations of the term enclosed by parentheses using the behavior strategies given for nature in the definition of the game, we can denote the result as $d(\alpha E(i), \alpha I(i), \beta E(i), \beta I(i); i=1,\ldots,N)$ and rewrite H as:

$$H = \sum_{\alpha I(1)} \ \sum_{\alpha E(1)|\alpha I(1)} \cdots \sum_{\alpha I(N)} \ \sum_{\alpha E(N)|\alpha I(N)} \cdots \sum_{\beta I(N)} \ \sum_{\beta E(N)|\beta I(N)}$$

$$p(\alpha E(1)|\alpha I(1)) \ldots p(\alpha E(N)|\alpha I(N)) \Big(d(\alpha E(i), \alpha I(i), \beta E(i), \beta I(i);$$

$$i=1,\ldots,N)\Big)\ p(\beta E(1)|\beta I(1)) \ldots p(\beta E(N)|\beta I(N)) \qquad (5)$$

Now, under our assigned orderings to each player's possible information sets at each rank k, we define for $k=1,\ldots,N$, partitioned vectors $X^{\alpha k}$, $Y^{\beta k}$ whose ℓ^{th} components, $X^{\alpha k}_{\ell}$ and $Y^{\beta k}_{\ell}$, are defined below.

$X^{\alpha k}_{\ell} = p(\alpha E(k) = j|\alpha I(k)_i)$ where $\ell = j + \sum_{m=1}^{i-1} n^{\alpha}_{mk}$ and n^{α}_{mk} is the total number of alternatives available in information set $\alpha I(k)_m$.

$Y^{\beta k}_{\ell} = p(\beta E(k) = j|\beta I(k)_i)$ where $\ell = j + \sum_{m=1}^{i-1} n^{\beta}_{mk}$ and n^{β}_{mk} is the total number of alternatives available in information set $\beta I(k)_m$.

Denote by ∇ an inner product operation. For example,

$$\sum_i \sum_j \sum_k \alpha_i \beta_j \rho_k A_{ijk} \triangleq \alpha \nabla \beta \nabla \rho \nabla A$$

and let $\prod_{i=1}^{N} X^{\alpha i} \triangleq X^{\alpha 1} \nabla X^{\alpha 2} \nabla \ldots \nabla X^{\alpha N}$.

Using the above notation, we can construct a 2N-dimensional array D from the function $d(\alpha E(i), \alpha I(i), \beta E(i), \beta I(i);\ i=1,\ldots,N)$ such that equation (5) can be rewritten as:

$$H = \left(\prod_{i=1}^{N} X^{\alpha i} \right) \nabla \left(\prod_{i=1}^{N} Y^{\beta i} \right) \nabla D \qquad (6)$$

Note that the set of vectors $(X^{\alpha 1}, \ldots, X^{\alpha N})$, constitutes a behavior strategy for the α-player, with the m^{th} partition of $X^{\alpha k}$ being a probability distribution on the alternatives available in $(\alpha I(k))_m$.

Denoting the original game as Γ, we define the associated game as follows:

Definition 6: The associated game Γ' for a given game Γ, is the 2N-person non-cooperative game defined by considering each rank $k=1,\ldots,2N$, to be under the control of a separate and independent entity whose payoff is identical to that of the player with which it is associated.

The i^{th} entity for the α-player [β-player] will be denoted as the αi-entity [βi-entity] for $i=1,\ldots,N$.

Definition 7: A behavior strategy for the αi-entity [βi-entity], $i=1,\ldots,N$, is any collection of probability distributions such that:

(a) the members of the collection are in one-to-one correspondence with the information sets under the given entity's control, and

(b) each probability distribution is a distribution over the integers 1, ..., j if the information set U with which it is associated is such that $U \subset A_j$.

Note that the set of all admissible values for the vector $X^{\alpha i}$ $[Y^{\beta i}]$ corresponds to the set of all possible behavior strategies for the αi-entity [βi-entity].

<u>Definition 8</u>: A set of entity behavior strategies $(X^{*\alpha i}, Y^{*\beta i}, i=1,\dots,N)$ is an <u>entity behavior equilibrium point</u> if:

$$\max_{X^{\alpha i}} \left[X^{\alpha i} \nabla \left(\prod_{\substack{j=1\\ j\neq i}}^{N} X^{*\alpha j}\right) \nabla \left(\prod_{j=1}^{N} Y^{*\beta j}\right) \nabla D\right] \leq \left(\prod_{j=1}^{N} X^{*\alpha j}\right) \nabla$$

$$\left(\prod_{j=1}^{N} Y^{*\beta j}\right) \nabla D \leq \min_{Y^{\beta k}} \left[Y^{\beta k} \nabla \left(\prod_{\substack{j=1\\ j\neq k}}^{N} Y^{*\beta j}\right) \nabla \left(\prod_{j=1}^{N} X^{*\alpha j}\right) \nabla D\right]$$

for $i,k=1,\dots,N$.

<u>Definition 9</u>: A pair of behavior strategies $(X^{*\alpha i}, i=1,\dots,N)$, $(Y^{*\beta i}, i=1,\dots,N)$ is a <u>player behavior equilibrium point</u> if:

$$\max_{(X^{\alpha i},\, i=1,N)} \left(\prod_{j=1}^{N} X^{\alpha j}\right) \nabla \left(\prod_{j=1}^{N} Y^{*\beta j}\right) \nabla D \leq \left(\prod_{j=1}^{N} X^{*\alpha j}\right) \nabla$$

$$\left(\prod_{j=1}^{N} Y^{*\beta j}\right) \nabla D \leq \min_{(Y^{\beta i},\, i=1,N)} \left(\prod_{j=1}^{N} X^{*\alpha j}\right) \nabla \left(\prod_{j=1}^{N} Y^{\beta j}\right) \nabla D$$

We complete this section by stating the following theorem and lemma.

Theorem 1: For any associated game Γ, there exists at least one entity behavior equilibrium point. (The proof of this theorem is given in the appendix.)

Lemma 1: If a pair of behavior strategies $(X^{*\alpha I}, i=1,\ldots,N)$, $(Y^{*\beta i}, i=1,\ldots,N)$ is to be of maximum possible quality for both players, then the corresponding set of entity behavior strategies must be an entity behavior equilibrium point. (The proof of this lemma follows directly from definitions 8 and 9.)

We will now define the concept of recall-sensitivity. We will then use this concept to establish a bound on the difference between the qualities of behavior strategies which correspond to entity behavior equilibrium points and the value of the game.

3.0 The Concept of Recall-Sensitivity and the Qualities of Behavior Strategies Corresponding to Entity Behavior Equilibrium Points

Denoting $H((X^{\alpha k}, Y^{\beta k}; k=1,\ldots,N)|\alpha I(i))$ as the expected return to the α-player as a function of the behavior strategies employed by the two players, given that a particular information set $\alpha I(i)$ has already been realized, and $H((X^{\alpha k}, Y^{\beta k}; k=1,\ldots,N)|\beta I(i))$ as the expected return given that $\beta I(i)$ has been realized, we define the concept of recall-sensitivity as follows:

Definition 10: The recall sensitivity of the α-player about a given pair of behavior strategies $(X^{*\alpha j}, j=1,\ldots,N)$, $(Y^{*\beta j}, j=1,\ldots,N)$ is R^{α} where:

$$\text{(a)}\quad R^{\alpha} = \sum_{k=2}^{N} \delta^{\alpha k}, \quad \text{and where:}$$

(b) $(\delta^{\alpha 2}, \ldots, \delta^{\alpha N})$ is the smallest set of real numbers such that:

$$\delta^{\alpha i}(X^{*\alpha j}; Y^{*\beta j}, j=1, \ldots, N) = \underset{X^{\alpha i}, \alpha I(i)}{\text{Max}} \Big[2 \Big[\underset{X^{\alpha 1}}{\max} \cdots \underset{X^{\alpha(i-1)}}{\max} H((X^{\alpha 1}, \ldots, X^{\alpha i}, X^{*\alpha(i+1)}, \ldots, X^{*\alpha N}; Y^{*\beta j}, j=1, \ldots, N) | \alpha I(i)) - \underset{X^{\alpha 1}}{\min} \cdots \underset{X^{\alpha(i-1)}}{\min} H((X^{\alpha 1}, \ldots, X^{\alpha i}, X^{*\alpha(i+1)}, \ldots, X^{*\alpha N}; Y^{*\beta j}, j=1, \ldots, N) | \alpha I(i)) \Big] \Big],$$

$i=2, \ldots, N.$

Similarly, the recall-sensitivity of the β-player about $(X^{*\alpha j}, j=1, \ldots, N)$, $(Y^{*\beta j}, j=1, \ldots, N)$ is R^{β} where:

(a) $R^{\beta} = \sum_{k=2}^{N} \delta^{\beta k}$, and where:

(b) $(\delta^{\beta 2}, \ldots, \delta^{\beta N})$ is the smallest set of real numbers such that:

$$\delta^{\beta i}(X^{*\alpha j}; Y^{*\beta j}, j=1, \ldots, N) = \underset{Y^{\beta i}, \beta I(i)}{\text{Max}} \Big[2 \Big[\underset{Y^{\beta 1}}{\max} \cdots \underset{Y^{\beta(i-1)}}{\max} H(X^{*\alpha j}, j=1, \ldots, N; Y^{\beta 1}, \ldots, Y^{\beta i}, Y^{*\beta(i+1)}, \ldots, Y^{*\beta N}) | \beta I(i)) - \underset{Y^{\beta 1}}{\min} \cdots \underset{Y^{\beta(i-1)}}{\min} H((X^{*\alpha j}, j=1, \ldots, N; Y^{\beta 1}, \ldots, Y^{\beta i}, Y^{*\beta(i+1)}, \ldots, Y^{*\beta N}) | \beta I(i)) \Big] \Big]$$

$i=2, \ldots, N.$

We can examine the above definition as follows.

The α-player's expected return, given that it has realized a particular information set $\alpha I(i)$, is dependent upon the functions $(X^{\alpha j}, j=1, \ldots, i-1)$, to the extent that these functions establish a probability distribution on the actions and measurements it knew in the past but cannot recall. If we denote the forgotten information as $\alpha IF(i)$, we can see this fact by expressing $H((X^{\alpha j}, Y^{\beta j},\ j=1,\dots,N|\alpha I(i))$ as follows:

$$H((X^{\alpha j}, Y^{\beta j},\ j=1,\dots,N)|\alpha I(i)) = \sum_{\alpha IF(i)} H((X^{\alpha j}, Y^{\beta j},\ j=1,\dots,N|\alpha I(i),$$

$$\alpha IF(i))\ p(\alpha IF(i)|\alpha I(i))$$

In section 4.0 below, we shall show that $H((X^{\alpha j}, Y^{\beta j},\ j=1,\dots,N)|\alpha I(i), \alpha IF(i))$ is not dependent upon any of the functions $(X^{\alpha j},\ j=1,\dots,i-1)$. Further, by causality, $p(\alpha IF(i)|\alpha I(i))$ is not dependent upon any of the functions $(X^{\alpha j}, Y^{\beta j},\ j=i,\dots,N)$. In the light of these facts, we can see that under either one of the following conditions, the loss of information $\alpha IF(i)$ will have little effect on the recall-sensitivity of the α-player about behavior strategies $(X^{*\alpha j},\ j=1,\dots,N)$, $(Y^{*\beta j},\ j=1,\dots,N)$:

(1) $H((X^{\alpha i}, X^{*\alpha j},\ j=i+1,\dots,N,\ Y^{*\beta k},\ k=1,\dots,N)|\alpha I(i), \alpha IF(i))$ is not strongly dependent upon $\alpha IF(i)$, for any given $X^{\alpha i}$ and $\alpha I(i)$, or,

(2) $p(\alpha IF(i)|\alpha I(i))$ is not strongly dependent upon the functions $(X^{\alpha j},\ j=1,\dots,i-1)$.

Using definition 10, we state the following theorem.

Theorem 2: If the pair of behavior strategies $(X^{*\alpha k},\ k=1,\dots,N)$, $(Y^{*\beta k},\ k=1,\dots,N)$ corresponds to an entity behavior equilibrium point, then:

(a) the quality of $(X^{*\alpha k}, k=1,\ldots,N)$ lies in the interval $(\emptyset, \emptyset - R^{\beta})$,

(b) the quality of $(Y^{*\beta k}, k=1,\ldots,N)$ lies in the interval $(\emptyset + R^{\alpha}, \emptyset)$,

where $\emptyset = \left(\prod_{j=1}^{N} X^{*\alpha j} \right) \nabla \left(\prod_{j=1}^{N} Y^{*\beta 1} \right) \nabla D$

The proof of this theorem is given in the appendix. The following corollary follows from theorem 2 as a direct consequence of the fact that we are dealing with a constant sum game.

Corollary 1: If the pair of behavior strategies $(X^{*\alpha k}, k=1,\ldots,N)$, $(Y^{*\beta k}, k=1,\ldots,N)$ corresponds to an entity behavior equilibrium point, thus,

(a) the quality of $(X^{*\alpha k}, k=1,\ldots,N)$ lies in the interval $(v, v-R^{\alpha}-R^{\beta})$

(b) the quality of $(Y^{*\beta k}, k=1,\ldots,N)$ lies in the interval $(v, v+R^{\alpha}+R^{\beta})$.

By corollary 1, we see that any pair of behavior strategies is an ϵ-solution if:

(i) it corresponds to an entity behavior equilibrium point, and

(ii) the recall-sensitivities of both players about the given pair of behavior strategies sum to no more than ϵ.

The following corollary follows directly from theorem 1 and corollary 1.

Corollary 2: If the α and β players both have recall-sensitivity of value zero, $(R^{\alpha} = R^{\beta} = 0)$ about any pair of behavior strategies, then a behavior strategy solution exists.

In the following section we will give a lemma and an example which together demonstrate that perfect recall for a player is a sufficient but not a necessary condition for that player to have a behavior strategy solution or ϵ-solution.

4.0 The Relationship Between Recall-Sensitivity and Perfect Recall

The following lemma is proven in the appendix.

Lemma 2: If a player has perfect recall, then it has recall-sensitivity of value zero about any pair of behavior strategies.

As a direct consequence of Corollary 1 and Lemma 2 above, we can state the following theorem.

Theorem 3: A behavior strategy solution exists in any game of perfect recall. (Any set of behavior strategies corresponding to an entity behavior equilibrium set is a solution.)

This theorem is very similar to Kuhn's theorem (reference 1) on the equivalence of mixed and behavior strategies.

To see that perfect recall is not necessary for the existence of behavior strategy solutions or ϵ-solutions, we need only show that games exist in which the players do not both have perfect recall, but do have recall-sensitivities of small or zero value about a pair of behavior strategies corresponding to an entity behavior strategy equilibrium point.

The following example describes such a game.

Example: In figure 1 we show a game in which the β-player moves at rank 2, and the α-player moves at ranks 1 and 3. The information sets for the players are appropriately indicated, and the payoffs to the α-player for each play W are indicated at the end of that play. Note that the α-player has forgotten at rank 3, its alternative choice at rank 1.

Since the β-player has perfect recall, its recall-sensitivity has value zero about any pair of behavior strategies. To examine the recall-sensitivity of the α-player, we look at its expected return given that it has realized information set $\alpha I(2)$. This can be expressed as:

$$H((X^{\alpha 1},X^{\alpha 2},Y^{\beta 1})|\alpha I(2)) = \sum_{C_W|\alpha I(2)} H(X^{\alpha 1},X^{\alpha 2},Y^{\beta 1}) \div \sum_{C_{W_2}|\alpha I(2)} p(W_2) \quad (7)$$

where $C_W|\alpha I(2)$ indicates summation over all plays W containing a vertex which is a member of information set $\alpha I(2)$, and $C_{W_2}|\alpha I(2)$ indicates summation over the set of all partial plays leading from 0 to a vertex in information set $\alpha I(2)$.

We can expand equation (7) to obtain:

$$\begin{aligned} H((X^{\alpha 1},X^{\alpha 2},Y^{\beta 1})|\alpha I(2)) &= p(2|\beta I(1)) \Big[p(1|\alpha I(1))p(1|\alpha I(2))h_2 \\ &+ p(1|\alpha I(1))p(2|\alpha I(2))h_3 + p(2|\alpha I(1))p(1|\alpha I(2))h_4 + p(2|\alpha I(1)) \\ &\cdot p(2|\alpha I(2))h_5\Big] \div \Big[p(1|\alpha I(1))p(2|\beta I(1)) + p(2|\alpha I(1))p(1|\beta I(1))\Big] \end{aligned} \quad (8)$$

If we examine equation (8), we see that if $h_2 = h_4$ and $h_3 = h_5$, then $H((X^{\alpha 1},X^{\alpha 2},Y^{\beta 1})|\alpha I(2))$ will not depend on $X^{\alpha 1}$ at all. In this case the α-player will have a recall-sensitivity of value zero about any pair of behavior strategies. In this situation then, both players

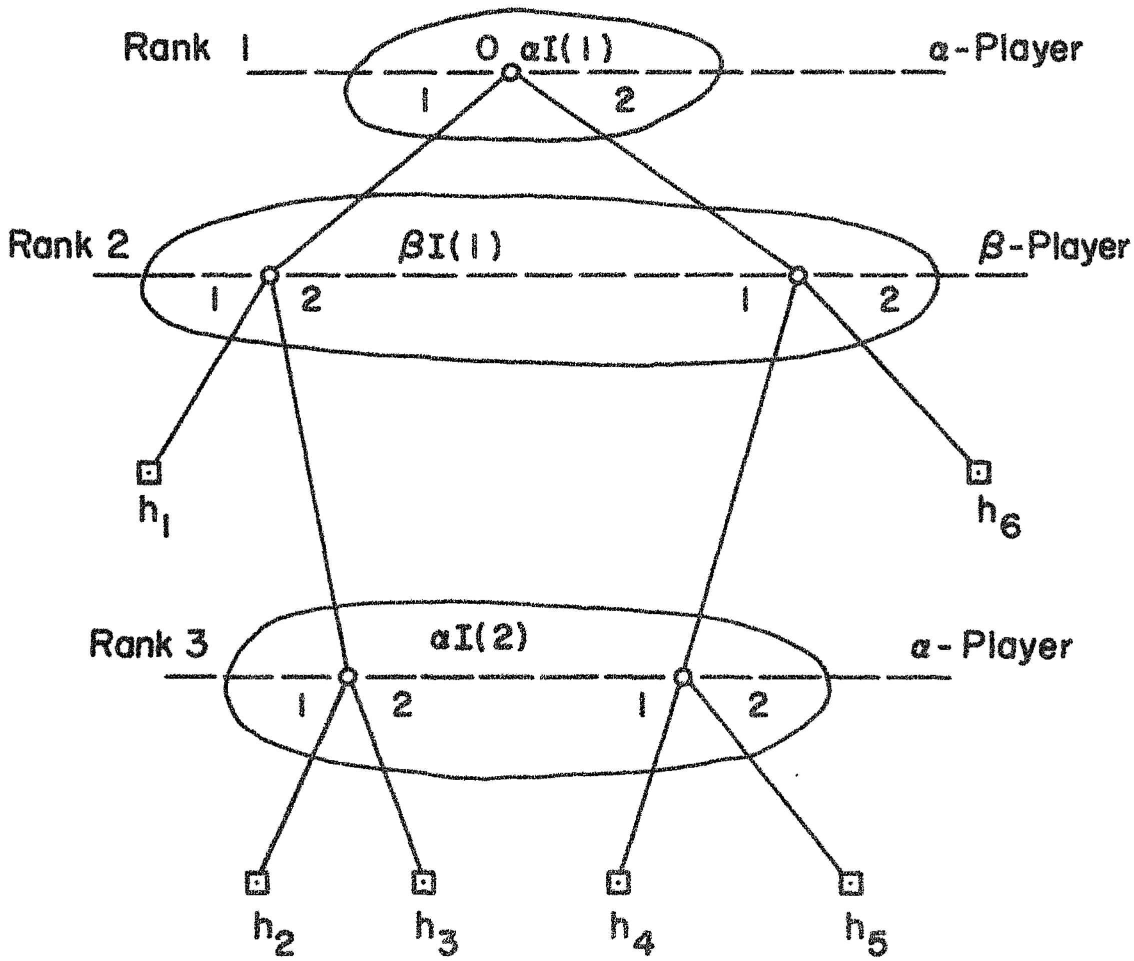

FIGURE 1. AN EXAMPLE IMPERFECT RECALL GAME.

have recall-sensitivities of value zero and thus by Corollary 1, a behavior strategy solution exists even though this is not a game of perfect recall.

If, however, $h_2 = h_4 + \epsilon_1$ and $h_3 = h_5 + \epsilon_2$ where ϵ_1 and ϵ_2 are both small numbers, then it is easy to show that the recall sensitivity of the α-player is bounded by $c = \max(\epsilon_1, \epsilon_2)$ around any pair of behavior strategies. In this solution, a behavior strategy "c-solution" must therefore exist.

Lemma 2 and the above example thus imply (by Corollary 1) that the condition of perfect recall for the players is sufficient but not necessary for the existence of a behavior strategy solution or ϵ-solution.

In the following sections we will describe a heuristic algorithm for generation of entity behavior equilibrium points, and we will comment on the practical implications of our observations.

5.0 Construction of Entity Behavior Equilibrium Points

The following algorithm has been employed by the authors to generate entity behavior equilibrium points in several associated games. We call this algorithm the "extended fictitious play" algorithm because of its similarity to the Brown-Robinson fictitious play algorithm for the solution in mixed strategies of games in normal form (references 3 and 4).

We will describe the algorithm by first establishing some notation, and then defining the sequence of entity behavior strategies it is to generate.

Denote by maxv $X^{\alpha k}$ a vector partitioned in the same manner as $X^{\alpha k}$ and having all zero components except for a single 1.0 entry in each partition in a position corresponding to a maximum element of $X^{\alpha k}$ in that partition.[†]

Correspondingly, denote by minv $Y^{\beta k}$ a vector partitioned in the same manner as $Y^{\beta k}$ and having all zero components except for a single 1.0 entry in each partition in a position corresponding to a minimum element of $Y^{\beta k}$ in that partition.

Definition 11: The set of vector sequences ($V^{\alpha k}(j), X^{\alpha k}(j), Q^{\beta k}(j), Y^{\beta k}(j)$, $k=1,N$ and $j=0,1,\ldots$) will be called an entity behavior strategy sequence for the associated game Γ' which is defined by array D if:

$$\text{(i)} \quad V^{\alpha k}(j) = \left(\prod_{\substack{\ell=1 \\ \ell \neq k}}^{N} X^{\alpha \ell}(j) \right) \nabla \left(\prod_{\ell=1}^{N} Y^{\beta \ell}(j) \right) \nabla D$$

$$\text{(ii)} \quad Q^{\beta k}(j) = \left(\prod_{\ell=1}^{N} X^{\alpha \ell}(j) \right) \nabla \left(\prod_{\substack{\ell=1 \\ \ell \neq k}}^{N} Y^{\beta \ell}(j) \right) \nabla D$$

$$\text{(iii)} \quad X^{\alpha k}(j+1) = \left[j\, X^{\alpha k}(j) + \text{maxv } V^{\alpha k}(j) \right] \left(\frac{1}{j+1} \right)$$

$$\text{(iv)} \quad Y^{\beta k}(j+1) = \left[j\, Y^{\beta k}(j) + \text{minv } Q^{\beta k}(j) \right] \left(\frac{1}{j+1} \right)$$

for $k=1,N$ and $j=0,1,\ldots,$ where ($X^{\alpha k}(0)$, $Y^{\beta k}(0)$, $k=1,N$) are arbitrary initial entity behavior strategies.

[†] In the case of several maximum elements in a given partition, the choice of which one is to correspond to the 1.0 entry is arbitrary.

We have generated entity behavior strategy sequences for a number of games. In each case studied, convergence in the following sense was observed.†

For $\epsilon = 0.05$, and some reasonably small integer j_o, an integer $\xi \leq j_o$ could be found such that for $(X^{\alpha k}(\xi), Y^{\beta k}(\xi),\ k=1,\ldots,N)$, the following expressions are valid.

$$\text{(a)}\quad \max_{X^{\alpha i}} X^{\alpha i} \nabla \Big(\prod_{\substack{\ell=1\\ \ell\neq i}}^{N} X^{\alpha \ell}(\xi)\Big) \nabla \Big(\prod_{\ell=1}^{N} Y^{\beta \ell}(\xi)\Big) \nabla D \leq (1+\epsilon)\ \phi(\xi)$$

$$\text{(b)}\quad \min_{Y^{\beta i}} Y^{\beta i} \nabla \Big(\prod_{\ell=1}^{N} X^{\alpha \ell}(\xi)\Big) \nabla \Big(\prod_{\substack{\ell=1\\ \ell\neq i}}^{N} Y^{\beta \ell}(\xi)\Big) \nabla D \geq (1-\epsilon)\ \phi(\xi)$$

for $i=1,\ldots,N$ and where $\phi(\xi) = \Big(\prod_{\ell=1}^{N} X^{\alpha \ell}(\xi)\Big) \nabla \Big(\prod_{\ell=1}^{N} Y^{\beta \ell}(\xi)\Big) \nabla D.$

To give some idea of the sort of performance obtained using the extended fictitious play algorithm, in the most involved case that we have considered, an entity behavior strategy sequence was generated for a 14-entity associated game. In this associated game, the number of information sets controlled by an entity ranged from 1 for the simplest entity, to 1024 for the most complex, and the number of alternatives was 2 for wvery information set. In this example, conditions (a) and (b) were satisfied for an $\xi \leq j_o = 200$.

We turn now to consideration of the practical implications of the observations we have made in sections 2.0-5.0.

---

†Shapley (reference 7) has given a class of games for which fictitious play will not converge, but this class cannot be transformed into a class of games of the type we are considering.

6.0 Practical Implications

First, we have shown that there are games of imperfect recall in which entity behavior equilibrium points exist which correspond to behavior strategy solutions or ϵ-solutions, and we have described a technique by which such equilibrium points may be generated. As far as we know, no other techniques have yet been developed for determination of behavior strategy solutions and ϵ-solutions in imperfect recall games of sufficient complexity to represent real life situations.

Second, in many games, computation of the value of the game is too complicated to carry out. In such games it may be possible to establish bounds on the value of the game by:

(i) generating an entity behavior equilibrium point,

(ii) establishing bounds on the recall-sensitivities of the players about this equilibrium point, and

(iii) applying appropriately theorem 2 and corollary 1.

Third, there are games in which intuition suggests that a player should possess a behavior strategy guaranteeing an expected return close to the value of the corresponding perfect recall game, using an information partition considerably simpler than that required for perfect recall. For example, in a finite pursuit-evasion game, we might expect that a player can guarantee itself an expected return close to the value of the corresponding perfect recall game, using a simplified information partition; i.e., one which corresponds to recall of only "sufficiently recent" information. Such a simplification is of interest because it implies that less memory will be required to store and implement the corresponding behavior strategy solution or ϵ-solution, but presently no techniques are available to investigate such possibilities.

The above problem is composed of two separate questions. When an information partition is proposed which simplifies the recall requirement of one of the players, a new game is defined. Denoting the value of the new game as w, and the value of the original perfect recall game as v_{pr}, we note that a behavior strategy pair in the new game is an ϵ-solution in the perfect recall game only if:

(i) the behavior strategy pair is an ϵ_1-solution in the new game, and

(ii) $\epsilon_1+\epsilon_2 \leq \epsilon$ where $\epsilon_2 = v_{pr}-w$.

A technique for generation of ϵ-solutions in imperfect recall games is obviously a prerequisite for any search for a simplified information partition. An approach to such a problem can be made by employing the extended fictitious play technique to generate entity behavior equilibrium points for the cases where:

(i) both players have perfect recall, and where

(ii) the player of interest has imperfect recall while its opponent has perfect recall.

By theorem 2, we see that in each case $\emptyset$ is the quality of the generated strategy for the player of interest, and thus if the two values of $\emptyset$ are "close enough" to each other, the tested imperfect recall information partition and correspondingly the behavior strategy generated for the player of interest, is "good enough."

In reference 5, this technique is employed to find a simplified behavior strategy ϵ-solution in a rather involved 5-stage "doctor-patient medical game." A behavior strategy was found for the doctor which guarantees 92 percent of the value of the perfect recall game but requires storage of instructions for only 57 information sets, whereas perfect recall for the doctor would require storage of instructions for 681 information sets.

References

1. Kuhn, H. W., "Extensive Games and the Problem of Information," Contributions to the Theory of Games, Annals of Math Studies, No. 28, Princeton University Press, 1953.

2. Utgoff, V. A., and Kashyap, R. L., "A Pursuit and Evasion Problem with Measurement Uncertainty," TR-EE 69-39, School of Electrical Engineering, Purdue University, October 1969, (to appear in the Journal of Optimization Theory and Applications).

3. Robinson, J., "An Iterative Method of Solving a Game," Annals of Mathematics, Vol. 54, No. 2, September 1951.

4. Danskin, J. M., "Fictitious Play for Continuous Games," Naval Research Logistics Quarterly, Vol. 1, 1954, pp. 313-320.

5. Utgoff, V. A., and Kashyap, R. L., "A Doctor-Patient Problem--A Finite Two-Person Constant-Sum Extended Game with Imperfect Recall," to be published.

6. Nash, J., "Non-Cooperative Games," Annals of Mathematics, Vol. 54, No. 2, September 1951.

7. Shapley, L. S., "Some Topics in Two-Person Games," Annals of Math Studies, Vol. 52, 1964.

APPENDIX

PROOFS

A. Proof of Theorem 1

In order to prove theorem 1, we shall establish the following definitions and notation.

Definition 11: Any entity behavior strategy $X^{\alpha i}$ [$Y^{\beta i}$] which consists of only zeros and a single 1.0 in each information set partition will be called an entity pure strategy, and will be denoted as $X^{p\alpha i}$ [$Y^{p\beta i}$].

Since there are finitely many information sets and alternative choices, there are finitely many entity pure strategies. We shall assign an arbitrary ordering to the pure strategies for each entity, and denote for the αi-entity [βi-entity], $i=1,\ldots,N$:

(a) the kth entity pure strategy as $X_k^{p\alpha i}$ [$Y_k^{p\beta i}$]

(b) the total number of entity pure strategies as $\ell^{\alpha i}$ [$\ell^{\beta i}$]

(c) the set of all entity pure strategies as $C^{p\alpha i}$ [$C^{p\beta i}$].

Definition 12: For any entity αi [βi], $i=1,\ldots,N$, any probability distribution over $C^{p\alpha i}$ [$C^{p\beta i}$] will be called an entity mixed strategy and will be denoted as $X^{m\alpha i}$ [$Y^{m\beta i}$].

We shall denote the probability assigned under $X^{m\alpha i}$ [$Y^{m\beta i}$] to the entity pure strategy $X_k^{p\alpha i}$ [$Y_k^{p\beta i}$] as $X_k^{m\alpha i}$ [$Y^{m\beta i}$].

We shall state definitions 13, 14 and 15 for the αi-entities only. The corresponding definitions for the βi-entities are analogous.

<u>Definition 13</u>: For any entity mixed strategy $X^{m\alpha i}$, the <u>associated entity behavior strategy</u> is the entity behavior strategy defined by:

$$\bar{X}^{\alpha i} = \sum_{k=1}^{\ell^{\alpha i}} X_k^{m\alpha i} \cdot X_k^{p\alpha i} \qquad (1)$$

<u>Definition 14</u>: For any entity behavior strategy $X^{\alpha i}$, the <u>associated entity mixed strategy</u> is the entity mixed strategy defined by:

$$\bar{X}_k^{m\alpha i} = \prod_{\substack{\text{all components} \\ i|j \text{ such that} \\ (X_k^{p\alpha i})_{i|j}=1.0}} (X^{\alpha i})_{i|j} \qquad (2)$$

where the $i|j$ component of $X^{\alpha i}$ (or $X_k^{p\alpha i}$) is the component corresponding to the ith alternative in the partition corresponding to the jth information set.

As a direct consequence of definition 11-14, an entity behavior strategy generates under equation 2 an entity mixed strategy which in turn generates under equation 1 the original entity behavior strategy.

<u>Definition 15</u>: A set of entity mixed strategies ($X^{*m\alpha i}$, $Y^{*m\beta i}$, $i=1,\ldots,N$) is called an <u>entity mixed strategy equilibrium point</u> if it is true that:

$$\max_{X^{m\alpha j}} H(X^{m\alpha j}, Y^{*m\beta j}; X^{*m\alpha i}, Y^{*m\beta i}, i\neq j, \; i=1,\ldots,N) \leq H(X^{*m\alpha i}, Y^{*m\beta i}, i=1,\ldots,N) \leq \min_{Y^{m\beta k}} H(X^{*m\alpha k}, Y^{m\beta k}; X^{*m\alpha i}, Y^{*m\beta i}, i\neq k, \; i=1,\ldots,N)$$

for all $j,k = 1,\ldots,N$.

We now proceed with the proof of theorem 1, by stating and proving the following two lemmas.

Lemma 3: The expected returns under any given set of entity mixed strategies $(X^{m\alpha i}, Y^{m\beta i}, i=1,\ldots,N)$ and the corresponding set of associated entity behavior strategies $(\bar{X}^{\alpha i}, \bar{Y}^{\beta i}, i=1,\ldots,N)$ are equal.

Proof: Under the given set of entity mixed strategies, the expected return is given by:

$$H(X^{m\alpha i}, Y^{m\beta i}, i=1,\ldots,N) = \sum_{C} \Big[(X_{k_1}^{m\alpha 1} \cdot X_{k_1}^{p\alpha 1})^{\dagger} \nabla\, (X_{k_2}^{m\alpha 2} \cdot X_{k_2}^{p\alpha 2}) \nabla \ldots \nabla$$

$$(X_{k_N}^{m\alpha N} \cdot X_{k_N}^{p\alpha N}) \nabla\, (Y_{\ell_1}^{m\beta 1} \cdot Y_{\ell_1}^{p\beta 1}) \nabla \ldots \nabla\, (Y_{\ell_N}^{m\beta N} \cdot Y_{\ell_N}^{p\beta N}) \nabla D \Big] \qquad (3)$$

where C is the set of all possible integer sequences $(k_1, k_2, \ldots, k_N, \ell_1, \ell_2, \ldots, \ell_N)$ such that $1 \leq k_i \leq \ell^{\alpha i}$ and $1 \leq \ell_i \leq \ell^{\beta i}$ for $i=1,\ldots,N$. This expression, however, can be written as:

$$H(X^{m\alpha i}, Y^{m\beta i}, i=1,\ldots,N) = \Big(\sum_{k_1=1}^{\ell^{\alpha 1}} X_{k_1}^{m\alpha 1} \cdot X_{k_1}^{p\alpha 1} \Big) \nabla \Big(\sum_{k_2=1}^{\ell^{\alpha 2}} X_{k_2}^{m\alpha 2} \cdot X_{k_2}^{p\alpha 2} \Big) \nabla$$

$$\ldots \nabla \Big(\sum_{k_N=1}^{\ell^{\alpha N}} X_{k_N}^{m\alpha N} \cdot X_{k_N}^{p\alpha N} \Big) \nabla \Big(\sum_{\ell_1=1}^{\ell^{\beta 1}} Y_{\ell_1}^{m\beta 1} \cdot Y_{\ell_1}^{p\beta 1} \Big) \nabla \ldots \nabla$$

$$\Big(\sum_{\ell_1=1}^{\ell^{\beta 1}} Y_{\ell_N}^{m\beta N} \cdot Y_{\ell_N}^{p\beta N} \Big) \nabla D \qquad (4)$$

By definition 13, however, each of the strategy terms in parentheses above is the associated entity behavior strategy for the given entity

† Note $X_{k_1}^{m\alpha 1}$ is a scalar, while $X_{k_1}^{p\alpha 1}$ is a vector.

mixed strategy which corresponds to it. So we can rewrite (4) as:

$$H(X^{m\alpha i}, Y^{m\beta i}, i=1,\ldots,N) = \left(\prod_{i=1}^{N} \bar{X}^{\alpha i}\right) \nabla \left(\prod_{i=1}^{N} \bar{Y}^{\beta i}\right) \nabla D$$

$$= H(\bar{X}^{\alpha i}, \bar{Y}^{\beta i}, i=1,\ldots,N) \qquad (5)$$

Q.E.D.

Lemma 4: If any set of entity mixed strategies is an entity mixed strategy equilibrium point, then the corresponding set of associated entity behavior strategies is an equilibrium point in entity behavior strategies.

Proof: Suppose $(X^{*m\alpha i}, Y^{*m\alpha i}, i=1,\ldots,N)$ is an entity mixed strategy equilibrium point. By definition 15 then, for any $j,k=1,\ldots,N$ it is true that:

$$\max_{X^{m\alpha j}} H(X^{m\alpha j}, Y^{*m\beta j}; X^{*m\alpha i}, Y^{*m\beta i}, i\neq j,\ i=1,\ldots,N) \leq H(X^{*m\alpha i}, Y^{*m\beta i},$$

$$i=1,\ldots,N) \leq \min_{Y^{m\beta k}} H(X^{*m\alpha k}, Y^{m\beta k}; X^{*m\alpha i}, Y^{*m\beta i}, i\neq k,\ i=1,\ldots,N) \qquad (6)$$

Now, suppose there exists an $X^{\alpha j}$ such that:

$$H(X^{\alpha j}, \bar{Y}^{*\beta j}; \bar{X}^{*\alpha i}, \bar{Y}^{*\beta i}, i\neq j,\ i=1,\ldots,N) > H(\bar{X}^{*\alpha i}, \bar{Y}^{*\beta i}, i=1,\ldots,N) \qquad (7)$$

where $(\bar{X}^{*\alpha i}, \bar{Y}^{*\beta i}, i=1,\ldots,N)$ is the set of associated entity behavior strategies for the given entity mixed strategy equilibrium point.

Applying lemma 3 and definition 14, we can write from equation (7) above, that:

$$H(\bar{X}^{m\alpha j}, Y^{*m\beta j}; X^{*m\alpha i}, Y^{*m\beta i}, i\neq j,\ i=1,\ldots,N) = H(X^{\alpha j}, \bar{Y}^{*\beta j}; \bar{X}^{*\alpha i}, \bar{Y}^{*\beta i},$$

$$i=1,\ldots,N) > H(\bar{X}^{*\alpha i}, \bar{Y}^{*\beta i}, i=1,\ldots,N) = H(X^{*m\alpha i}, Y^{*m\beta i}, i=1,\ldots,N) \qquad (8)$$

This, however, contradicts expression (6), so no such entity behavior strategy can exist, and we can write

$$\max_{X^{\alpha j}} (X^{\alpha j}, \bar{\bar{Y}}^{*\beta j}; \bar{\bar{X}}^{*\alpha i}, \bar{\bar{Y}}^{*\beta i}, i \neq j, \ i=1,\ldots,N) \leq H(\bar{\bar{X}}^{*\alpha i}, \bar{\bar{Y}}^{*\beta i}, i=1,\ldots,N) \quad (9)$$

Similar arguments will establish that:

$$H(X^{*\alpha i}, \bar{\bar{Y}}^{*\beta i}, i=1,\ldots,N) \leq \min_{Y^{\beta k}} (\bar{\bar{X}}^{*\alpha k}, Y^{\beta k}; \bar{\bar{X}}^{*\alpha i}, \bar{\bar{Y}}^{*\beta i}, i=1,\ldots,N) \quad (10)$$

and the arguments can be carried out to obtain equations (9) or (10) for any $j,k=1,\ldots,N$.

Q.E.D.

To complete the proof of theorem 1, we now note that the 2N-entity game in entity mixed strategies is an "n-person non-cooperative finite game" as defined by Nash in reference 6. For such games, Nash gives the following theorem.

Theorem (Nash): Every finite (n-person non-cooperative) game has an equilibrium point (in mixed strategies).

The proof of our theorem 1 now follows trivially from Nash's theorem and lemma 4 above.

B. Proof of Theorem 2

The proof of theorem 2 proceeds as follows.

If the pair of behavior strategies $(X^{*\alpha k}, \ k=1,\ldots,N)$, $(Y^{*\beta k}, k=1,\ldots,N)$ corresponds to an entity behavior strategy equilibrium point, then taking the α-player's part we can write:

$$\max_{X^{\alpha N}} H(X^{\alpha N},Y^{*\beta N};X^{*\alpha i},Y^{*\beta i},\ i=1,\dots,N-1) = \max_{X^{\alpha N}} \sum_{\alpha I(N)} H((X^{\alpha N},Y^{*\beta N};$$

$$X^{*\alpha i},Y^{*\beta i},\ i=1,\dots,N-1)\,|\,\alpha I(N))p(\alpha I(N);(X^{\alpha N},Y^{*\beta N};X^{*\alpha i},Y^{*\beta i},$$

$$i=1,\dots,N-1)) \leq H(X^{*\alpha i},Y^{*\beta i},\ i=1,\dots,N) \tag{11}$$

where $p(\alpha I(N);(X^{\alpha i},Y^{\beta i},\ i=1,\dots,N))$ is the probability that information set $\alpha I(N)$ occurs, as a function of the behavior strategies employed by the α and β players.

By our definition of recall-sensitivity we can write:

$$\max_{X^{\alpha 1}} \dots \max_{X^{\alpha(N-1)}} H((X^{\alpha i},Y^{*\beta i},\ i=1,\dots,N)\,|\,\alpha I(N)) - H((X^{*\alpha 1},\dots,$$

$$X^{*\alpha(N-1)},X^{\alpha N};Y^{*\beta i},\ i=1,\dots,N)\,|\,\alpha I(N)) \leq \delta^{\alpha N},\ \text{for all } \alpha I(N) \text{ and } X^{\alpha N} \tag{12}$$

Combining (11) and (12), and noting that since $X^{\alpha N}$ is a function on $\alpha I(N)$, the summation and maximization operations can be interchanged, we obtain:

$$\sum_{\alpha I(N)} \Big[\max_{X^{\alpha 1}} \dots \max_{X^{\alpha N}} H((X^{\alpha i},Y^{*\beta i},\ i=1,\dots,N)\,|\,\alpha I(N))p(\alpha I(N);$$

$$(X^{*\alpha 1},\dots,X^{*\alpha(N-1)},X^{\alpha N},Y^{*\beta i},\ i=1,\dots,N))\Big] \leq H(X^{*\alpha i},Y^{*\beta i},$$

$$i=1,\dots,N) + \delta^{\alpha N}/2 \tag{13}$$

Now $H(X^{*\alpha i},Y^{*\beta i},\ i=1,\dots,N)$ can also be expressed as a sum over $\alpha I(N)$, and by causality, $p(\alpha I(i);(X^{\alpha k},Y^{\beta k},\ k=1,\dots,N))$ is not dependent upon $(X^{\alpha j},Y^{\beta j},\ j=i,\dots,N)$, so we can rewrite (13) as:

$$\sum_{\alpha I(N)} \Big[\max_{X^{\alpha 1}} \ldots \max_{X^{\alpha N}} H((X^{\alpha i}, Y^{*\beta i},\ i=1,\ldots,N) | \alpha I(N)) - H((X^{*\alpha i},$$

$$Y^{*\beta i},\ i=1,\ldots,N) | \alpha I(N)) \Big]\, p(\alpha I(N); (X^{*\alpha i}, Y^{*\beta i},\ i=1,\ldots,N-1)) \leq \delta^{\alpha N}/2 \qquad (14)$$

Now $p(\alpha I(N)(\ldots))$ takes its values in $[0,1]$, and since the bracketed function is non-negative, we can write from equation (14) that:

$$\max_{X^{\alpha 1}} \ldots \max_{X^{\alpha N}} H((X^{\alpha i}, Y^{*\beta i},\ i=1,\ldots,N) | \alpha I(N)) - H((X^{*\alpha i}, Y^{*\beta i},$$

$$i=1,\ldots,N) | \alpha I(N)) \leq \delta^{\alpha N}/2 \qquad (15)$$

Multiplying (15) by $p(\alpha I(N); (X^{\alpha i}, Y^{*\beta i},\ i=1,\ldots,N-1))$ and summing over $\alpha I(N)$, we obtain:

$$\sum_{\alpha I(N)} \max_{X^{\alpha 1}} \ldots \max_{X^{\alpha N}} H((X^{\alpha i}, Y^{*\beta i},\ i=1,\ldots,N) | \alpha I(N)) p(\alpha I(N); (X^{\alpha i}, Y^{*\beta i},$$

$$i=1,\ldots,N-1)) \leq \sum_{\alpha I(N)} H((X^{*\alpha i}, Y^{*\beta i},\ i=1,\ldots,N) | \alpha I(N)) p(\alpha I(N);$$

$$(X^{\alpha i}, Y^{*\beta i},\ i=1,\ldots,N-1) + \delta^{\alpha N}/2 \qquad (16)$$

Now, reversing the maximization and summation operations in (16) either has no effect on the inequality or strengthens it. Further, by our recall-sensitivity definition we have:

$$|H((X^{*\alpha i}, Y^{*\beta i},\ i=1,\ldots,N) | \alpha I(N)) - H((X^{\alpha 1}, \ldots, X^{\alpha(N-1)}, X^{*\alpha N};$$

$$Y^{*\beta i},\ i=1,\ldots,N) | \alpha I(N))| \leq \delta^{\alpha N}/2 \qquad (17)$$

Combining (16) and (17), we can therefore write:

$$\max_{X^{\alpha 1}} \cdots \max_{X^{\alpha N}} H(X^{\alpha i},Y^{*\beta i},\ i=1,\ldots,N) \leq H(X^{\alpha 1},\ldots,X^{\alpha(N-1)},X^{*\alpha N};$$

$$Y^{*\beta i},\ i=1,\ldots,N)) + \delta^{\alpha N} \tag{18}$$

Starting with the equation corresponding to equation (11), but for maximization over $X^{\alpha(N-1)}$, we can repeat the steps of equations (12) thru (18) to obtain:

$$\max_{X^{\alpha 1}} \cdots \max_{X^{\alpha(N-1)}} H(X^{\alpha 1},\ldots,X^{\alpha(N-1)},X^{*\alpha N};Y^{*\beta i},\ i=1,\ldots,N) \leq$$

$$H(X^{\alpha 1},\ldots,X^{\alpha(N-2)},X^{*\alpha(N-1)},X^{*\alpha N}) + \delta^{\alpha(N-1)} \tag{19}$$

By similar reasoning, we can write for $\ell=2,\ldots,N$ that:

$$\max_{X^{\alpha 1}} \cdots \max_{X^{\alpha \ell}} H(X^{\alpha 1},\ldots,X^{\alpha \ell},X^{*\alpha(\ell+1)},\ldots,X^{*\alpha N};Y^{*\beta i},\ i=1,\ldots,N) \leq$$

$$H(X^{\alpha 1},\ldots,X^{\alpha(\ell-1)},X^{*\alpha \ell},\ldots,X^{*\alpha N};Y^{*\beta i},\ i=1,\ldots,N) + \delta^{\alpha \ell} \tag{20}$$

Now, combining the equation corresponding to equation (11), but for maximization over $X^{\alpha 1}$, with equation (20) for $\ell=2$, we can write:

$$\max_{X^{\alpha 1}} \max_{X^{\alpha 2}} H(X^{\alpha 1},X^{\alpha 2},X^{*\alpha 3},\ldots,X^{*\alpha N};Y^{*\beta i},\ i=1,\ldots,N) \leq H(X^{\alpha 1},X^{*\alpha 2},$$

$$\ldots,X^{*\alpha N};Y^{*\beta i},\ i=1,\ldots,N) + \delta^{\alpha 2} \leq \max_{X^{\alpha 1}} H(X^{\alpha 1},X^{*\alpha 2},\ldots,X^{*\alpha N};$$

$$Y^{*\beta i},\ i=1,\ldots,N) \leq H(X^{*\alpha i},Y^{*\beta i},\ i=1,\ldots,N) + \delta^{\alpha 2} \tag{21}$$

Applying (21) in the right-hand side of (20) for $\ell=3$, we obtain:

$$\max_{X^{\alpha 1}} \max_{X^{\alpha 2}} \max_{X^{\alpha 3}} H(X^{\alpha 1}, X^{\alpha 2}, X^{\alpha 3}, X^{*\alpha 4}, \ldots, X^{*\alpha N}; Y^{*\beta i},\ i=1,\ldots,N) \leq$$

$$H(X^{*\alpha i}, Y^{*\beta i},\ i=1,\ldots,N) + \delta^{\alpha 2} + \delta^{\alpha 3} \qquad (22)$$

Combining for $\ell=4,\ldots,N$, we finally obtain:

$$\max_{X^{\alpha 1}} \ldots \max_{X^{\alpha N}} H(X^{\alpha i}, Y^{*\beta i},\ i=1,\ldots,N) \leq H(X^{*\alpha i}, Y^{*\beta i},\ i=1,\ldots,N) +$$

$$\sum_{i=2}^{N} \delta^{\alpha i} \qquad (23)$$

But, by definition:

$$H(X^{*\alpha i}, Y^{*\beta i},\ i=1,\ldots,N) \leq \max_{X^{\alpha 1}} \ldots \max_{X^{\alpha N}} H(X^{\alpha i}, Y^{*\beta i},\ i=1,\ldots,N) =$$

$$\text{quality of } (Y^{*\beta i},\ i=1,\ldots,N) \qquad (24)$$

From (23) and (24) then, the quality of $(Y^{*\beta i},\ i=1,\ldots,N)$ must lie in the interval $(\emptyset, \emptyset+R^{\alpha})$, and similar arguments show that the quality of $(X^{*\alpha i},\ i=1,\ldots,N)$ must lie in the interval $(\emptyset-R^{\beta}, \emptyset)$.

Q.E.D.

C. Proof of Lemma 2:

The proof of lemma 2 proceeds as follows.

Proof: Suppose for example that the α-player has perfect recall.

Using expression (3) of section 2.0 of the text, we can write:

$$H((X^{\alpha i},Y^{\beta i},\ i=1,\ldots,N)|\alpha I(i)) = \Big[\sum_{C_W|\alpha I(i)} \Big[p(\eta E(k)|\eta I(k))p(\alpha E(k)|$$

$$\alpha I(k))p(\beta E(k)|\beta I(k));k=1,\ldots,N)\Big]\Big] \quad \sum_{C_{W_i}|\alpha I(i)} \Big[p(\eta E(k)|\eta I(k)) \cdot$$

$$p(\alpha E(k)|\alpha I(k))p(\beta E(k)|\beta I(k));\ k=1,i-1)\Big]\Big] \qquad (25)$$

where $C_W|\alpha I(i)$ indicates summation over all plays W under which $\alpha I(i)$ occurs, and $C_{W_i}|\alpha I(i)$ indicates summation over all partial plays to the rank of the α-player's ith move, under which $\alpha I(i)$ occurs.

Now, if the α-player has perfect recall, $\alpha I(i)$ covers the domain of definition of each of the functions defining $p(\alpha E(1)|\alpha I(1))\ldots p(\alpha E(i-1)|\alpha I(i-1))$, so these terms can be factored in (25) to obtain:

$$H((X^{\alpha i},Y^{\beta i},\ i=1,\ldots,N)|\alpha I(i)) = \Big[\Big[\prod_{k=1}^{i-1} p(\alpha E(k)|\alpha I(k))\Big]$$

$$\cdot \sum_{C_W|\alpha I(i)} \Big[p(\eta E(k)|\eta I(k))p(\beta E(k)|\beta I(k));\ k=1,\ldots,i-1\Big]\ (p(\eta E(k)|$$

$$\eta I(k))\ p(\alpha E(k)|\alpha I(k))\ p(\beta E(k)|\beta I(k));\ k=i,\ldots,N\Big]\Big] \div \Big[\Big[\prod_{k=1}^{i-1} \cdot$$

$$p(\alpha E(k)|\alpha I(k))\Big] \cdot \sum_{C_{W_i}|\alpha I(i)} \Big[p(\eta E(k)|\eta I(k))p(\beta I(k)|\beta I(k));$$

$$k=1,\ldots,i-1\Big]\Big] \qquad (26)$$

Noting that all terms defined by $(X^{\alpha 1},\ldots,X^{\alpha(i-1)})$ cancel out, we can write:

$H((X^{\alpha 1},\ldots,X^{\alpha(i-1)},X^{\alpha i},\ldots,X^{\alpha N};\ Y^{\beta k},\ k=1,\ldots,N)|\alpha I(i))$ -

$H((X^{1\alpha 1},\ldots,X^{1\alpha(i-1)},X^{\alpha i},\ldots,X^{\alpha N};\ Y^{\beta k},\ k=1,\ldots,N)|\alpha I(i)) = 0$

for any $\alpha I(i)$, and any values for $(X^{\alpha k},X^{1\alpha k},Y^{\beta k},\ k=1,\ldots,i-1;$

$X^{\alpha j},Y^{\beta j},\ j=1,\ldots,N)$

By definition then $\delta^{\alpha i}=0$, and similar arguments can be carried out for $\delta^{\beta i}$ if the β-player has perfect recall, for $i=1,\ldots,N$.

Q.E.D.