MATHEMATICS OF OPERATIONS RESEARCH
Vol. 7, No. 1, February 1982
Printed in U.S.A.

OPTIMAL BEHAVIORAL STRATEGIES IN 0-SUM GAMES WITH ALMOST PERFECT INFORMATION*†

JEAN-PIERRE PONSSARD‡ AND SYLVAIN SORIN§

This paper provides the general construction of the optimal strategies in a special class of zero sum games with incomplete information, those in which the players move sequentially. It is shown that at any point of the game tree, a player's optimal behavioral strategy may be derived from a state variable involving two components: the first one keeps track of the information he revealed, the second one keeps track of the (vector) payoff he should secure over his opponent's possible position. This construction gives new insights on earlier results obtained in the context of sequential repeated games. Several examples are discussed in detail.

I. Introduction. This paper is concerned with the recursive construction of optimal behavioral strategies for a class of 0-sum extensive games with incomplete information, defined as follows:

(i) Let G be a two-person finite game tree with perfect information.

(ii) Let A_H be the zero-sum payoff associated with a play H of G, A_H is a discrete random variable such that $\operatorname{Prob}(A_H = a_H^{rs}) = p^r q^s$ where $p = (p^r)_{r \in R}$, $q = (q^s)_{s \in S}$ are two independent probability distributions over two given finite sets R and S.

(iii) The game is played the following way: chance selects $r \in R$ and $s \in S$ according to p and q and reveals r to Player I and s to Player II, then the players proceed along the game tree until they reach a play H; at that stage, Player II (the minimizer) pays a_H^{rs} to Player I (the maximizer) and the game ends. All the preceding description is common knowledge (i.e., r is revealed to Player I but not to Player II, similarly for s, but both players know that, etc.)

This class of games called games with almost perfect information was first studied in Ponssard (1975).

It overlaps with the class of repeated games with incomplete information (the study of which was initiated by Aumann and Maschler (1966)), since it includes the sequential finitely repeated games (Ponssard and Zamir, 1973).

Briefly speaking, in all these games, the strategic problem of information usage may be stated in the following terms. First, the players are concerned with using their private information (i.e., for Player I to correlate his moves with the event r selected by chance and only known by himself) but this exposes them to reveal it to their opponent (i.e., through the observation of Player I's past moves, Player II may infer something about the selected r) and possibly reduce its value for the future moves of the game. Second, the players cannot take at its face value the information apparently revealed by their opponent without being subject to bluffing.

These two aspects have been developed in the context of two persons zero-sum infinitely repeated games in order to find "maxmin" strategies (Aumann, Maschler (1966); Kohlberg (1975); Mertens, Zamir (1980)). In this context the main tools to

*Received March 7, 1979; revised November 30, 1979.
AMS 1980 subject classification. Primary: 90D15.
OR/MS Index 1978 subject classification. Primary: 236 Games/group decisions/noncooperative.
Key words. Games, incomplete information.
†This work has been partially supported by Contrat C.O.R.D.E.S. no. 136-77.
‡Ecole Polytechnique, Paris.
§Université Paris VI.

0364-765X/82/0701/0014$01.25

construct an optimal strategy are conditional probabilities (use of private information) and vector payoffs (limiting use of opponent's information).

For games with almost perfect information, it appears that a somewhat similar construction can be explicitly made. However it takes advantage of two special features of such games: the fact that they are finite and the fact that the players move in sequence. In particular, we shall show that at any point of the game tree a player's optimal strategy may be derived from a state variable involving two components: the first one keeps track of the information he revealed, the second one keeps track of the conditional vector payoff he should secure over his opponent's possible positions for the remaining part of the game. This general recursive construction of the optimal strategies is proved in part II and is applied in specific cases in the subsequent sections. The present development uses the linear programming formulation of these games and its implications (Ponssard and Sorin, 1980).

II. The recursive structure of optimal strategies in a compounded game.

1. *Definitions.* Define the following compounded game (Ponssard, 1975):

—Chance selects $r \in R$ and $s \in S$ according to two given probability distributions $p = (p^r)_{r \in R} \in P$, $q = (q^s)_{s \in S} \in Q$, r is revealed to Player I, s is revealed to Player II.

—Player I selects a move $t \in T$ (T is a finite set), then knowing t the players proceed by playing the zero-sum matrix game B_t^{rs}. Denote by $G(p,q)$ this game, and by $V(p,q)$ its value.

Let I and J be the two player's strategy sets in B_t^{rs} (we can always assume that I and J are independent of r, s and t) and let $V^t(p,q)$ be the value of the above compounded game where Player I's set T is restricted to the unique element t. Clearly, $V^t(p,q)$ is the value of the game with incomplete information $(B_t^{rs})_{r \in R, s \in S}$.

2. *Preliminary results.*

LEMMA 1. *$V^t(p,q)$ is the value of the following linear program where the unknowns are $(\alpha_i^r)_{i \in I, r \in R}$ and $(v^s)_{s \in S}$:*

$$\begin{aligned} &\operatorname{Max} \sum_s q^s v^s \\ &\textit{subject to} \quad \sum_{r,i} b_{ij}^{rst} \alpha_i^r \geqslant v^s \quad \forall s, \forall j, \\ &\qquad\qquad\quad \sum_i \alpha_i^r = p^r \qquad \forall r, \\ &\qquad\qquad\quad \alpha_i^r \geqslant 0 \qquad\quad \forall i, \forall r. \end{aligned} \tag{1}$$

This is Lemma 2 in (Ponssard and Sorin, 1980).

Note that if in this program the optimal value of the variables $v = (v^s)_{s \in S}$ are known, the optimal value of the variables $\alpha = (\alpha_i^r)_{r \in R, i \in I}$ may be derived without recourse to the value of q.

Starting with some specific probability distributions p_0, q_0 we shall exhibit Player I's optimal strategy in the compounded game as:

—an optimal move on $T : x_t^r = \text{Prob}(t \mid r)$,

—for each $t \in T$ an associated (p_t, v_t) which can then be used in (1) to derive the optimal moves in $(B_t^{rs})_{r \in R, s \in S}$.

Let us introduce the following notations: For each function $g : P \times Q \to \mathbb{R}$, $\operatorname{Cav}_p g$ is the smallest real function defined on $P \times Q$, concave w.r.t. p and greater than g on $P \times Q$. $\operatorname{Vex}_q g$ is defined in a dual manner.

LEMMA 2.

$$V(p,q) = \operatorname*{Cav}_p \operatorname*{Max}_t V^t(p,q).$$

This is Theorem 2 in (Ponssard, 1975).

DEFINITION. A function $g: P \times Q \to \mathbb{R}$ is "piecewise bilinear" if there is a finite partition of P (resp. Q) into convex polyhedra $(C_k)_{k \in K}$ (resp. $(D_l)_{l \in L}$) such that the restriction of g to each product set $C_k \times D_l$ is bilinear.

LEMMA 3. *V and V^t are concave w.r.t. p, convex w.r.t. q and piecewise bilinear.*

This is Theorem 1 in (Ponssard, Sorin, 1980).

LEMMA 4. *For all p, q in $P \times Q$ there exist $\lambda_t \in [0,1]$ and $p_t \in P$ where $t \in T$ such that:*

$$\sum_t \lambda_t p_t = p, \qquad \sum_t \lambda_t = 1 \quad \textit{and} \quad V(p,q) = \sum_t \lambda_t V^t(p_t, q). \tag{2}$$

PROOF. Since V is continuous, Caratheodory's Theorem and Lemma 2 imply the existence of $\lambda_r \in [0,1]$, $p_r \in P$ with $r \in R$ such that:

$$\sum_r \lambda_r p_r = p, \qquad \sum_r \lambda_r = 1,$$

and

$$V(p,q) = \sum_r \lambda_r \max_t V^t(p_r, q)$$

$$= \sum_r \lambda_r V^{t_r}(p_r, q).$$

Now since each V^t is concave w.r.t. p we can assume that if $\lambda_r \cdot \lambda_{r'} \neq 0$ then $t_r \neq t_{r'}$, so that we can redefine the indices and this gives the result.

Since V^t is a concave convex function we can define the following sets, for all $p \in P$, $q \in Q$, $t \in T$:

$$L^t(p) = \{ l \in \mathbb{R}^S;\ \ l \cdot q \leqslant V^t(p,q) \text{ on } Q \},$$

$$L^t(p,q) = \{ l \in \mathbb{R}^S;\ \ l \in L^t(p) \text{ and } l \cdot q = V^t(p,q) \},$$

$$M^t(q) = \{ m \in \mathbb{R}^R;\ \ m \cdot p \geqslant V^t(p,q) \text{ on } P \},$$

$$M^t(p,q) = \{ m \in \mathbb{R}^R;\ \ m \in M^t(q) \text{ and } m \cdot p = V^t(p,q) \}.$$

We define similarly $L(p)$, $L(p,q)$, $M(q)$, $M(p,q)$ for the function V.

3. *The optimal strategies.*

THEOREM 1. *For all $l \in L(p_0, q_0)$ there exist (λ_t, p_t) $t \in T$ which satisfy (2) at p_0, q_0, and $l_t \in L^t(p_t, q_0)$ for all $t \in T$ such that:*

$$\left(\sum_t \lambda_t l_t - l \right) \cdot q \geqslant 0 \quad \textit{for all } q \in Q.$$

PROOF. Let us write the formula (2) in the following way

$$V(p_0, q_0) = \sum_i \lambda_i V^{t_i}(p_i, q_0)$$

such that

$$V(p_i, q_0) = V^{t_i}(p_i, q_0).$$

Now since V is piecewise linear and convex with respect to q there is a neighborhood Q' of q_0 and a finite number of $R \times S$ matrices A_k, $k \in K$ such that:

$$V(p_0, q) = \max_k p_0 A_k q \quad \text{on } Q'.$$

It follows that if l belongs to $L(p_0, q_0)$ there exist μ_k, $\mu_k \in [0,1]$, $k \in K$, such that $l = \sum_k p_0 \mu_k A_k$ with $\sum_k \mu_k = 1$.

Now since V is piecewise bilinear, this implies that, for all i,

$$l_i = \sum_k p_i \mu_k A_k \quad \text{belongs to } L(p_i, q_0).$$

It remains to see that each $p_i A_k$ belongs to $L^{t_{ik}}(p_i, q_0)$ for some t_{ik} in T, since $V = \operatorname{Cav} \max_t V^t$.

Finally due to the concavity of each V^t the t_{ik} are different for different values of i.

Now starting with formula (2) we have

$$V(p, q_0) = \sum_i \lambda_i V^{t_{ik}}(p_i, q_0)$$

with $\sum_i \lambda_i = 1$ and $\sum_i \lambda_i p_i = p$.

Let us introduce $\rho_{ik} = \lambda_i \mu_k$, $p_{ik} = p_i$ and $l_{ik} = p_i A_k$ for all i and k, then we have:

$$\sum_{i,k} p_{ik} = 1, \quad \sum_{i,k} \rho_{ik} p_{ik} = p,$$

$$V(p, q_0) = \sum_{i,k} \rho_{ik} V^{t_{ik}}(p_{ik}, q_0)$$

and

$$\sum_{i,k} \rho_{ik} l_{ik} = l$$

which gives the result.

Theorem 2. *An optimal strategy for Player* I *in* $G(p_0, q_0)$ *is defined as follows*:
(a) *choose* $l \in L(p_0, q_0)$ *and then* $\lambda_t, p_t, l_t, t \in T$ *such that*:

$$\begin{aligned} &\text{(i)} \quad \lambda_t \in [0,1], \quad \textstyle\sum_t \lambda_t = 1, \quad \sum_t \lambda_t p_t = p_0; \\ &\text{(ii)} \quad l_t \in L^t(p_t) \quad \textit{for all } t \in T; \\ &\text{(iii)} \quad (\textstyle\sum_t \lambda_t l_t - l) q \geqslant 0 \quad \textit{for all } q \in Q. \end{aligned} \tag{3}$$

(b) *Denoting by* x_t^r *the probability of playing* t *given* r, *Player* I*'s intital move is given by*:

$$x_t^r = \lambda_t \cdot \frac{p_t^r}{p_0^r} \quad \textit{for all } t, r \textit{ with } p_0^r > 0.$$

(c) *If Player* I*'s first move is* t *then he selects his move in* B_t^{rs} *according to* (p_t, l_t) *in* (1).

Proof. From Theorem 1, it follows that the admissible set of λ_t, p_t, l_t in (a) is not empty, since $L^t(p_t, q_0) \subset L^t(p_t)$, hence the strategy is well defined.

We shall now prove that this strategy, denoted by σ, guarantees $V(p_0, q_0)$ to Player I.

Assume that Player II knows σ. Learning t he can compute a posterior probability on R, given σ and p which is precisely p_t since we have:

$$\operatorname{Prob}(r/t) = \frac{\operatorname{Prob}(r \text{ and } t)}{\operatorname{Prob}(t)} = \frac{x_t^r}{\sum_r p_0^r x_t^r} = \frac{\lambda_t \cdot p_t^r}{\lambda_t \sum_r p_t^r} = p_t^r.$$

Now for each $q \in Q$, Player I can guarantee at (p_r, q) the amount $l_t \cdot q$ since l_t belongs to $L^t(p_t)$.

It remains to compute the total probability of playing t which is $\sum_r p_0^r x_t^r = \lambda_t$ so that Player I can assure himself $\sum_t \lambda_t l_t \cdot q$. Hence from (a)(iii) this amount is greater or equal than $l \cdot q$ so that at q_0 we obtain $l \cdot q_0$ which is precisely $V(p_0, q_0)$.

Note that this strategy rely only on the information Player I gives to Player II (i.e., the posteriors p_t) and not on the information he can get (Player I will play in G^t without computing posteriors on S but by using p_t, l_t in (1)).

III. Applications to special cases. The implications of Theorem 2 will now be developed in the context of 0-sum games with almost perfect information. First, we shall consider games with lack of information on one side. This will have the advantage to explicitly dissociate the two problems mentioned in the introduction (using one's own information, limiting use of opponent's information) thus giving rise to simpler structure for optimal strategies. After some general considerations (part A) we shall study two examples of sequential repeated games (part B). In the first one, the informed player is playing in the second position. The second example where the informed player in playing first was studied in Ponssard and Zamir (1973). Then in part C we shall illustrate the case of lack of information on both sides through poker-type games (Example 3).

A. *Games with lack of information on one side: General considerations.* Let us give additional notation:

(i) the finite game tree G is assumed to consist in a succession of moves $i_1, j_1, \ldots, i_t, j_t, \ldots, i_n, j_n$ in which i_t belongs to some fixed set I for Player I and j_t to some fixed set J for Player II; a partial history of G, noted h, consists of a sequence of moves starting from the beginning of the game tree; (it will be noted by h' if the number of moves is even, h'' otherwise; moreover $h'i$ denote the history h' followed by i and similarly for $h''j$).

(ii) Player I is informed. Let us denote this game by $G(p)$. Recall the general formula that gives the value of such games (Ponssard, 1975): For all complete histories H of G let $V^H(p)$ be $V^H(p) = \sum_{r \in R} p^r a_H^r$ then define recursively $V^h(p)$ by:

$$\begin{aligned} V^{h'}(p) &= \operatorname*{Cav}_{p} \operatorname*{Max}_{i \in I} V^{h'i}(p), \\ V^{h''}(p) &= \operatorname*{Min}_{j \in J} V^{h''j}(p). \end{aligned} \tag{4}$$

Eventually one obtains $V(p)$ which gives the value of the game for any $p \in P$. Note that for all h, $V^h(p)$ is the value of the game in which the players are restricted to begin the game G by playing the dummy moves indicated in h and then effectively play over the rest of the game tree of G.

COROLLARY 1. *The two players' optimal behavioral strategies in $G(p_0)$ are the following*:

For Player I, *for all h define $p_h \in P$ starting with the initial probability distribution p_0, such that*:

(i) $p_{h''j} = p_{h''} \quad \forall j \in J$;

(ii) *the $p_{h'i}$'s, for all $i \in I$, are obtained by solving the following system*:

$$\sum_{i \in I} \lambda_{h'}^i V^{h'i}(p_{h'i}) \geqslant V^{h'}(p_{h'}),$$

$$\sum_{i \in I} \lambda_{h'}^i p_{h'i} = p_{h'},$$

$$\sum_{i \in I} \lambda_{h'}^i = 1$$

then player I'*s optimal strategy after the history* h' *is given by*

$$\operatorname{Prob}(i \mid h' \text{ and } r) = \lambda^i_{h'} p^r_{h'i} / p^r_{h'}.$$

For Player II, *for all* h, *define* $m_h \in \mathbb{R}^R$ *starting with some* m_0 *which is a supporting hyperplane to* $V(p)$ *at* p_0 (*i.e.* $m_0 \in M(p_0)$), *and such that*:

(iii) $m_{h'} = m_{h'i} \quad \forall i \in I$.

(iv) *starting from* $m_{h''}$ *the* $m_{h''j}$ *are obtained by solving the following system*:

$$m_{h''j} \cdot p \geqslant V^{h''j}(p) \quad \forall p \in P,$$

$$\left(\sum_j \mu^j_{h''} m_{h''j} - m_{h''} \right) \cdot p \leqslant 0 \quad \forall p \in P,$$

$$\sum_j \mu^j_{h''} = 1.$$

Then Player II'*s optimal strategy after the history* h'' *is given by* $\operatorname{Prob}(j \mid h'') = \mu^j_{h''}$.

PROOF. These strategies are merely the counterpart of Theorem 2 and its development all along the game tree G since any restricted game may be viewed as a compounded game through appropriate normalization.

For Player I, the informed player, the notation of supporting hyperplane is reduced to its simplest form (the cardinality of S is one). Then, condition (3) may be directly expressed in terms of the points $V^{h'i}(p_{h'i})$ and this is precisely condition (ii).

For Player II, the noninformed player, it is the notion of conditional probabilities that becomes trivial. Condition (iv) is exactly condition (3) but taken from Player II's point of view.

COROLLARY 2. *The set of conditional probabilities* p_h *is a martingale over* P, *the set of conditional hyperplanes* m_h *is decreasing in expectation over all* P.

Furthermore, for all partial histories h *that may occur with a positive probability when both players use optimal strategies we have*:

$$\sum_{r \in R} m^r_h p^r_h = V^h(p_h).$$

PROOF. The first two statements follow from Corollary 1, as for the last statement it merely means that if Player II were to be revealed p_h, then he could not obtain more than the value of the restricted game initiated by h and with a starting probability distribution $p_0 = p_h$.

It may be interesting to give a graphical interpretation of the construction of the optimal strategies in the simplest case $|R| = 2$.

Starting at some $p_{h'}$, Player I's strategy will generate conditional probabilities on P as depicted in Figure 1. It should be observed that Player I's optimal strategy is conditioned over the sequence of probabilities which he generates himself: whereas he knows the r selected by chance, the martingale on P is a state variable that keeps track of what he revealed through his past moves thus generating a new starting point for the future. This idea has been present in the literature from the very beginning (Aumann and Maschler, 1966).

Player II's optimal strategy does not rely on posterior probabilities on P such as guessing some p_h and then maximizing an expected payoff at this point. Although by Corollary 2 it is consistent with Bayerian learning, it appears to be best interpreted in a vector framework. Consider the value function $V^{h''}$. By construction $V^{h''} = \operatorname{Min}(V^{h''1}, V^{h''2})$. The recursive construction of an optimal strategy for Player II has the following graphical counterpart. For any supporting hyperplane to $V^{h''}$, say $m^{h''}$, find

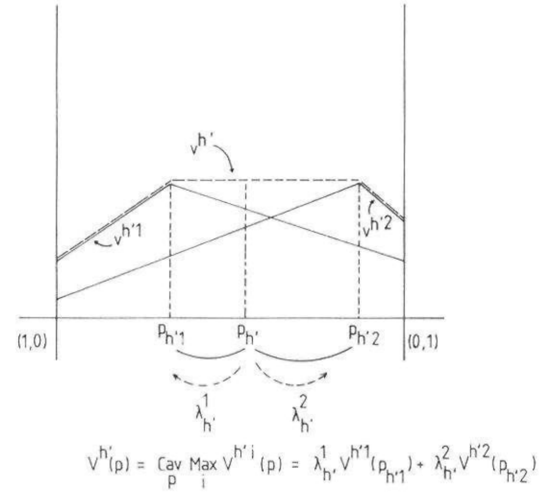


FIGURE 1. The informed player's optimal strategy.

two supporting hyperplanes to $V^{h''1}$ and $V^{h''2}$, respectively, such that a linear combination of these two is everywhere below $m^{h''}$. The coefficients of this combination give the probabilities of the respective moves. Once Player II's move is made the hyperplane for future reference becomes either $m^{h''1}$ or $m^{h''2}$. Thus, Player II's optimal strategy is conditioned at each stage by a set of successive hyperplanes each of which appears as a vector payoff that should be secured over the remaining stages of the game.

This construction has a similar feature with the construction of an optimal strategy for the noninformed player in infinitely repeated games, called the "Blackwell Strategy" (Aumann, Maschler, 1966), The initial probability distribution p_0 plays a role only through the specification of some supporting hyperplane. Then this hyperplane l_0 defines a set

$$A = \{ l \in \mathbb{R}^R \quad l^r \leqslant l_0^r \quad \forall r \in R \}$$

and Player II is now playing a game with vector-payoffs, each component corresponding to a state r, as defined by Blackwell (1956). An optimal strategy for Player II will then approach A (that is, with probability one, the distance between the average n

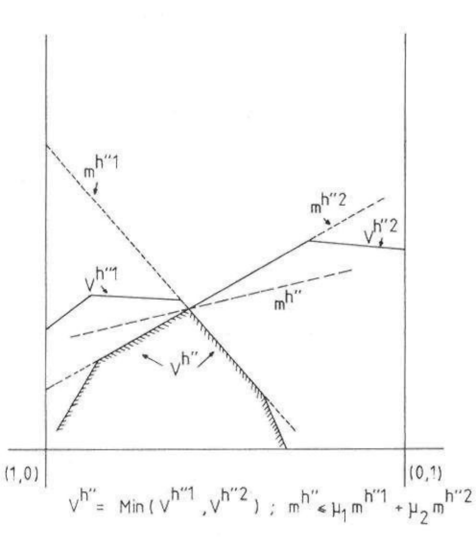


FIGURE 2. The noninformed player's optimal strategy.

stage payoff and A will be close to zero for some large n on). Hence l_0 is also a kind of security level vector.

B. *Examples*. The two following examples fall in the class of sequential repeated games. Since in a repeated game the total payoff is the sum of the stage payoffs (a_{ij}^{rs}), the value of the n repeated game with lack of information on both sides, V_n is given by:

$$V_n(p,q) = \operatorname*{Cav}_p \operatorname*{Max}_{i \in I} \operatorname*{Vex}_q \operatorname*{Min}_{j \in J}\Big(\sum_{r,s} a_{ij}^{rs} p^r q^s + V_{n-1}(p,q)\Big) \tag{5}$$

that is

$$V_n^i(p,q) = \operatorname*{Vex}_q \operatorname*{Min}_{j \in J}\Big(\sum_{r,s} a_{ij}^{rs} p^r q^s + V_{n-1}(p,q)\Big),$$

$$V_n(p,q) = \operatorname*{Cav}_p \operatorname*{Max}_{i \in I} V_n^i(p,q) \quad \text{and} \quad V_0(p,q) \equiv 0$$

under the following assumptions: Player I is playing first, maximizes, is informed about r and Player II, who plays second and minimizes, is informed about s.

EXAMPLE 1. In this example $R = \{1\}$ and $S = \{1,2\}$ that is Player I is uninformed and Player II is informed. The stage game tree is described below:

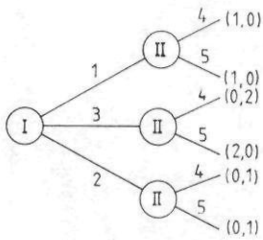


FIGURE 3

Hence the strategies are $i = 1, 2, 3$ for Player I and $j = 4, 5$ for Player II, and, for instance, the payoff $(2,0)$ stands for (a_{35}^1, a_{35}^2).

Note that if Player I's move is 1 or 2, Player II is a dummy. Hence, Player I can make a relevant observation on Player II's strategy only by playing 3, but this is clearly a bad move in the one stage game. Indeed if $q = (q_1, q_2)$ then $V_1(q) = \max\{q_1, q_2\}$.

Using formula (5) which gives here

$$V_n^i(q) = \operatorname*{Vex}_q \operatorname*{Min}_j \Big[\sum a_{ij}^s q^s + V_{n-1}(q)\Big], \tag{α}$$

$$V_n(q) = \max_i V_n^i(q). \tag{β}$$

The successive $V_n(q)$ are easily derived and it can be seen that $V_n(q)/n$ is constant for $n \geqslant 3$.

Figure 4 gives the construction of V_n up to V_4, where V_n^{ij} stands for

$$V_n^{ij}(q) = \sum_s a_{ij}^s q^s + V_{n-1}(q) \tag{γ}$$

It should also be mentioned that the (a) figures are made only for $i = 3$ since for $i = 1$ and 2 we have

$$V_n^1 = q^1 + V_{n-1} \quad \text{and} \quad V_n^2 = q^2 + V_{n-1}.$$

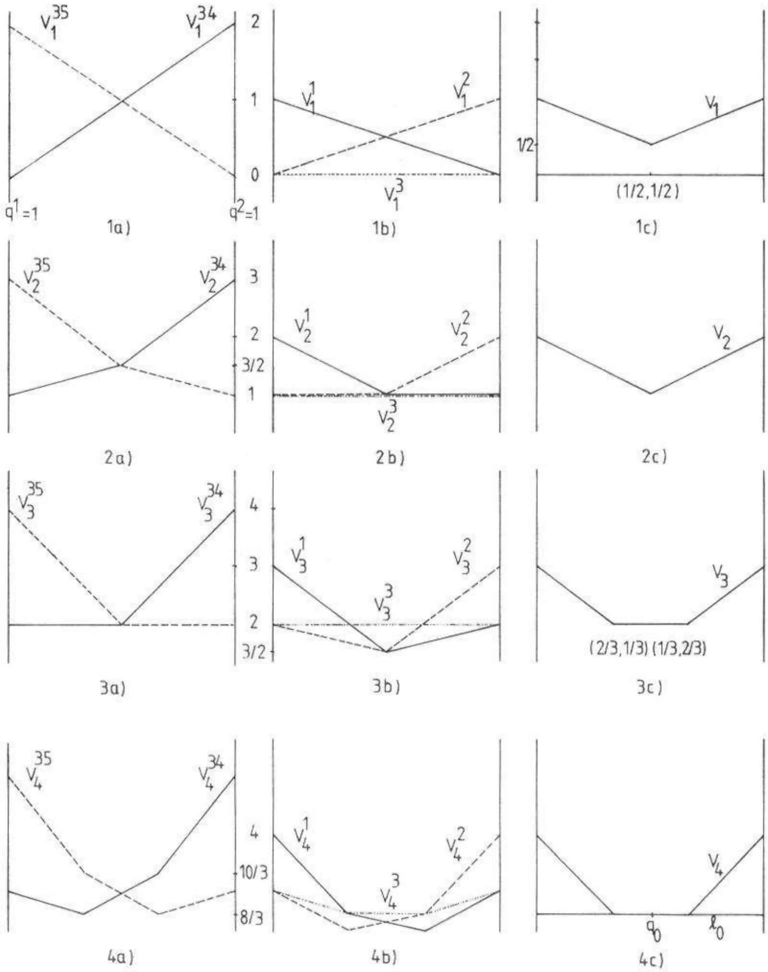


FIGURE 4. Construction of V_n $n \leqslant 4$.

(α) is used to construct Figure K(b) from Figure K(a) then (β) is used to get Figure K(c) from Figure K(b), then Figure $K+1$(a) comes from Figure K(c) through (γ).

Now we shall compute optimal strategies in the 4-stage game. Assume that $q_0 = (1/2, 1/2)$ so that the supporting hyperplane to V_4 at q_0 is $l_0 = (8/3, 8/3)$ (see 4(c)).

Note that a behavioral strategy for each player is constructed in four steps since it gives the move at each stage given the past history.

(1) *First step for Player* I. Choose $\lambda^i \in [0,1]$ and $l_i = (l_i^1, l_i^2)$, $i \in I$ such that $\sum_i \lambda^i = 1$, $l_i \cdot q \leqslant V_4^i(q)$ and $(\sum_i \lambda_i l_i - l_0) \cdot q \geqslant 0$.

From Figure 4(b) this implies $\lambda^3 = 1$ and $l_3 = l_0$. Hence Player I has to play 3 and to keep the same reference hyperplane $l_3 = l_0$.

(2) *First step for Player* II. Assuming Player I's move being 3, Player II has to find $\mu_3^j \in [0,1]$ and $q_{3j} \in Q$, $j \in J$ such that: $\sum_j \mu_3^j q_{3j} = q_0$ and $\sum \mu_3^j V_4^{3j}(q_{3j}) \leqslant V_4^3(q_0)$.

From Figure 4(a) and 4(b) this implies $\mu_3^4 = \mu_3^5 = 1/2$; $q_{34} = (2/3, 1/3)$ $q_{35} = (1/3, 2/3)$.

Hence an optimal strategy for Player II is to play 4 with probability 2/3 if $s = 1$ ($\text{Prob}(j = 4 \mid s = 1) = \mu_3^4 q_{34}^1 / q_0^1 = 2/3$) and with probability 1/3 if $s = 2$, keeping in mind q_{3j} if his move is actually j.

Assuming that the history is 34 we now obtain:

(3) *Second step for Player* I. Choose $\lambda_{34}^i \in [0,1]$ and l_{34i} $i \in I$ such that $\sum_i \lambda_{34}^i = 1$, $l_{34i} \cdot q \leqslant V_3^i(q)$ and $(\sum_i \lambda_{34}^i l_{34i} + a_{34} - l_3) \cdot q \geqslant 0$ (obviously we have to deflate a_{34} from the initial vector security level).

This gives $\lambda_{34}^1 = 1/3$, $\lambda_{34}^2 = 0$, $\lambda_{34}^3 = 2/3$, $l_{341} = (3,0)$, $l_{343} = (2,2)$ (from Figure 3(b)).

Hence Player I's optimal strategy at this step is to play 1 with probability $1/3$ and 3 with probability $2/3$, the vector security level being, according to the actual move $(3,0)$ or $(2,2)$.

(4) *Second step for Player* II. If the history is 341, Player II's move is not relevant. If $h = 343$ Player II can either play 4 with probability 1 and then $q_{3434} = q_{34} = (2/3, 1/3)$ or choose some q_{3435} with $q_{3435}^2 \geqslant 1/2$ and then λ_{3434}, λ_{3435}, q_{3434} such that:

$$\lambda_{3434} q_{3434} + \lambda_{3435} q_{3435} = q_{34},$$

$$\lambda_{3434} + \lambda_{3435} = 1$$

(for example $\lambda_{3434} = \lambda_{3435} = 1/2$, $q_{3434} = (1,0)$ $q_{3435} = (1/3, 2/3)$ and then play 4 with probability: $\lambda_{3434} q_{3434}^1 / \frac{2}{3}$ if $s = 1$ and with probability $\lambda_{3434} \cdot q_{3434}^2 / \frac{1}{3}$ if $s = 2$).

(5) *Third step for Player* I.

If $h = 3414$ or 3415 the relevant vector security level is $(2,0)$. Hence Player I has to play 1 with probability one and $l_{34141} = l_{34151} = (2,0)$ (see Fig. 2(b)).

If $h = 3434$ the relevant vector security level is $(2,0)$. Hence Player I has to play 1 with probability one and $l_{34341} = (2,0)$.

If $h = 3435$ we start with $l_{3435} - a_{35} = (0,2)$. Hence Player I plays 2 with probability one and $l_{34351} = (0,2)$.

(6) *Third step for Player* II. If Player II's move at this step is relevant (i.e., after a move 3 by Player I, which is nonoptimal given the past history) then the optimal move is to play completely revealing, i.e., 4 if $s = 1$ and 5 if $s = 2$.

(7) *Fourth step for Player* I.

–if $h = 341x1y$ (recall that Player II's move after move 1 of the Player I is irrelevant) Player I's optimal move is to play 1 with probability one;

–if $h = 34341x$ same result as above;

–if $h = 34352x$ Player I's optimal move is to play 2 with probability one.

(8) *Fourth step for Player* II. Here again if Player I plays optimally given the past history, Player II's move is irrelevant. Otherwise (i.e., if Player I chooses 3), then Player II plays completely revealing, i.e., 4 if $s = 1$ and 5 if $s = 2$.

Let us discuss the optimal strategies along the following remarks.

REMARK 1. Let $u(q) = \max_i \min_j \sum_s a_{ij}^s q^s$. Then the value of the infinitely repeated game is given by $V(q) = \operatorname{Vex} u(q)$. (Aumann, Maschler, 1966).

Hence from Figure 5 it follows that the optimal first move of the informed player in this four-stage game is his first optimal move in the infinitely repeated game. Whereas in that case he should retain the conditional probabilities $(2/3, 1/3)$ or $(1/3, 2/3)$ for the rest of the game, here we observe the end play effect: the informed player does not care about revealing information at the end and it is quite advantageous if the opportunity arises. It is not known whether this kind of structure for an optimal strategy is true in general.

Observe that, in spite of the fact that $V_3/3 = V_4/4$ the optimal strategies for Player II at the first step in both games are not the same.

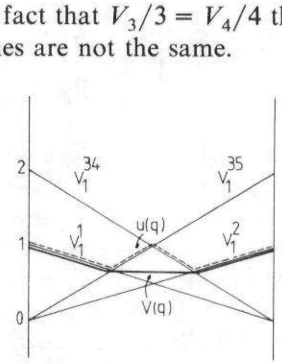


FIGURE 5

REMARK 2. The "Blackwell strategy" for Player I (the uninformed player) in the infinitely repeated game would approach the set $A = \{(x, y) \in R^2;\ x \geqslant 2/3, y \geqslant 2/3\}$. With our notations this means that the "security level" vector would be $(2/3, 2/3)$.

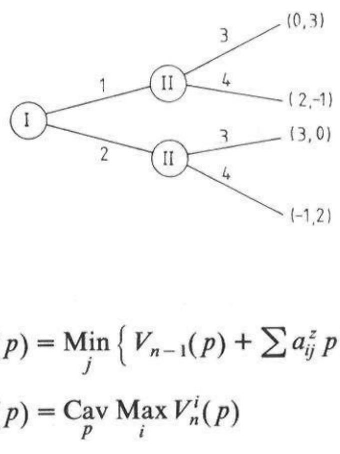
Note that in our construction it is $(2n/3, 2n/3)$ at stage n, for $n \geqslant 3$ (the average vector payoff being $(2/3, 2/3)$), then it becomes $(2n/3, 2(n-3)/3)$ if Player II's first move is 4 for example, etc.

Moreover Player I needs to randomize in order to achieve some security level vectors whereas in the Blackwell strategy for sequential games he does not have to randomize in order to approach the set A. Again, we observe similarities and differences.

Note that the optimal strategy of the noninformed player makes intuitive sense about information usage. The observation of a move 4 by Player I is a sign that $s = 1$ is more likely thus triggering a move 1 in response, whereas a move 5 triggers a move 2. But in order to avoid being bluffed, Player I has to come back from time to time to move 3 which may be interpreted as an investment for information. As we said earlier, this "story" need not be true; Player I's strategy is a maximin strategy.

EXAMPLE 2. (the trap phenomenon revisited). We shall show here that there are many more optimal strategies for the noninformed player than the one given in a previous paper (Ponssard and Zamir, 1973), some of them exploiting mistakes more than others. The idea of supporting hyperplane will enable us to formulate more precisely the "trap phenomenon" pointed out in Ponssard (1976).

Let us consider the following sequential game where Player I, the maximizer, is informed and plays first. The game is as follows:



Formula (5) gives here:

$$V_n^i(p) = \operatorname*{Min}_j \left\{ V_{n-1}(p) + \sum a_{ij}^z p^r \right\},$$

$$V_n(p) = \operatorname*{Cav}_p \operatorname*{Max}_i V_n^i(p)$$

and it is easy to see that:

$$V_n(p)/n = V_1(p) = \begin{cases} 3p^1 & \text{if } p^1 \in [0, 1/3], \\ 1 & \text{if } p^1 \in [1/3, 2/3], \\ 3p^2 & \text{if } p^1 \in [2/3, 1]. \end{cases}$$

Starting from $p_0 = (1/2, 1/2)$ the optimal strategies in the one-stage game $G_1(1/2)$ follow from Figure 6. For player I this gives $\lambda_1 = \lambda_2 = 1/2$ and $p_1 = (2/3, 1/3)$, $p_2 = (1/3, 2/3)$. Hence an optimal strategy is to play 1 with probability 2/3 if $r = 1$ and with probability 1/3 if $r = 2$.

Now for Player II we start with $m_0 = (1, 1)$. Let us suppose that Player I's move is 1, then we have: $m_{13} = (0, 3)$, $m_{14} = (2, -1)$ and $\mu_{13} = \mu_{14} = 1/2$; hence Player II plays 3 with probability 1/2 and similarly if Player I's move is 2.

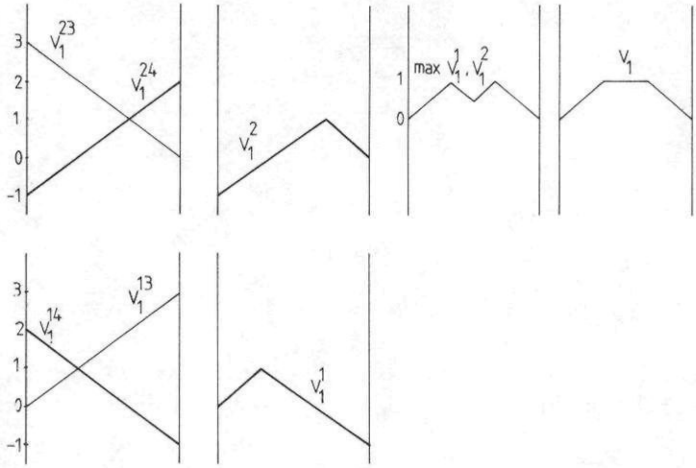


FIGURE 6

Let us now consider the two stage game $G_2(1/2)$ and compute optimal strategies.

(a) *First step for Player* I. From Figure 7 below it follows that the strategy at this step is the same as in $G_1(1/2)$.

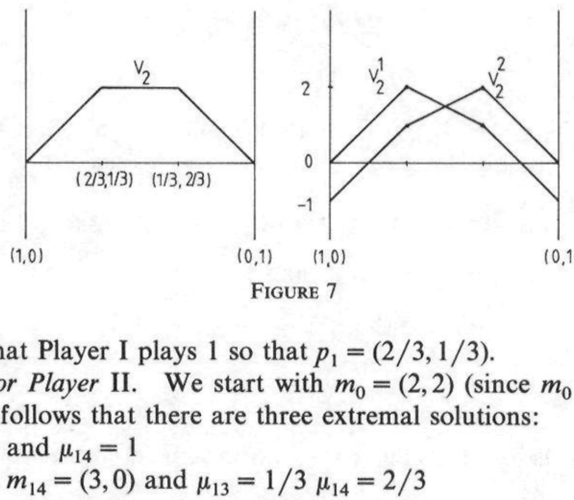


FIGURE 7

Assume then that Player I plays 1 so that $p_1 = (2/3, 1/3)$.

(b) *First step for Player* II. We start with $m_0 = (2, 2)$ (since $m_0 \in M(1/2)$); hence from Figure 8, it follows that there are three extremal solutions:

(1) $m_{14} = (2, 2)$ and $\mu_{14} = 1$

(2) $m_{13} = (0, 6)$ $m_{14} = (3, 0)$ and $\mu_{13} = 1/3$ $\mu_{14} = 2/3$

(3) $m_{13} = (1, 4)$ $m_{14} = (3, 0)$ and $\mu_{13} = 1/2$ $\mu_{14} = 1/2$.

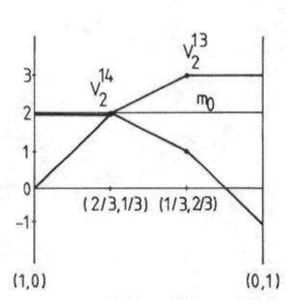


FIGURE 8

Let us suppose that Player II plays 4.

(c) *Second step for Player* I *given the history* 14. The construction of V_2^{14} is given in Figure 9. Since $p_{14} = (2/3, 1/3)$ the only optimal move is 1.

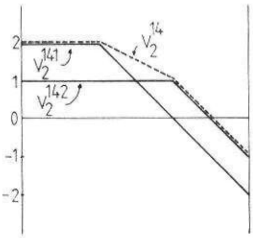


FIGURE 9

(d) *Second step for Player* II *given the history* 141. Using Figure 10 it follows that we have:

If $m_{14} = (2,2)$ then the only solution is $\mu_{1413} = 1$ (i.e., play 3) with $m_{1413} = (2,2)$,

If $m_{14} = (3,0)$ the only solution is $\mu_{1413} = \mu_{1414} = 1/2$ with $m_{1413} = (2,2)$ and $m_{1414} = (4,-2)$.

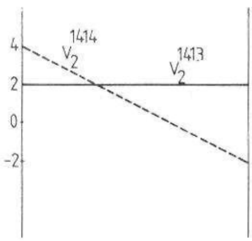


FIGURE 10

(e) *Second step for Player* II *given the history* 142. In this case Player I's strategy involves an observable mistake. From Figure 11 we can make the following observations:

(1) If $m_{14} = (2,2)$ then Player II can use two extremal possibilities:

$$m_{1424} = (1,1) \quad \text{and} \quad \mu_{1424} = 1$$

$$m_{1423} = (5,-1), \qquad m_{1424} = (1,1)$$

and

$$\mu_{1423} = 1/4, \qquad \mu_{1424} = 3/4.$$

Note that as long as he selects the first one with some nonzero probabilities we have:

$$\sum_j \mu_{142j} m_{142j} < m_{14}.$$

Player II's vector security level is strictly decreasing, that is, he is exploiting the mistake.

(2) If $m_{14} = (3,0)$ the only solution is now

$$m_{1423} = (5,-1), \qquad m_{1424} = (1,1)$$

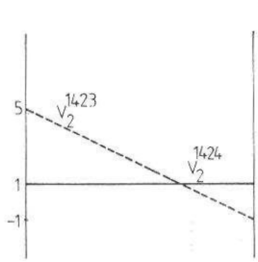


FIGURE 11

and

$$\mu_{1423} = 1/2, \qquad \mu_{1424} = 1/2;$$

but in this case we have

$$\sum_j \mu_{142j} m_{142j} = m_{14},$$

that is, Player II cannot exploit the mistake.

Let us now suppose that Player II plays 3 at the first stage, that is, uses with positive probability strategy (b)(2) or strategy (b)(3). Similar computation shows that if $m_{13} = (0,6)$ then the exploitation of a mistake is feasible whereas if $m_{13} = (1,4)$ it is not.

Let us now conclude with some remarks.

REMARK 1. (Comparison with the optimal strategies exhibited in Ponssard and Zamir, 1973). Player II's set of optimal strategies is enlarged. In the earlier paper, only strategy (b)(3) was used. In terms of exploiting the opponents' mistake we just showed that this is possible only by selecting (b)(1) or (b)(2) with some positive probability. Of course any optimal strategy is good enough to secure the value of the game, but in a game in extensive form one may ask for more.

REMARK 2. The fact that the exploitation of a mistake at the second stage need be prepared is made explicit by the fact that it crucially depends on which hyperplane is selected at the first stage. But, we fail to see intuitively why (b)(1) or (b)(2) should be preferred to (b)(3) by only considering V_2^{13} and V_2^{14}.

REMARK 3. The interested reader will note that the algorithm proposed in Ponssard (1975) to recursively obtain the optimal strategies breaks down in such an example because of the following degeneracy: V_2^{13} and V_2^{14} precisely intersects at the point at which they are kinked.

C. *Games with lack of information on both sides*: *a poker type game*. Just as chess, poker has always been a reference for game theory. From an economic point of view, poker may be considered as a rough model of escalation with incomplete information, and many extensions, involving nonzero sum elements, are certainly worth studying. In this respect, it is interesting to note that the development of games with incomplete information provides new tools for the analysis of such situations.

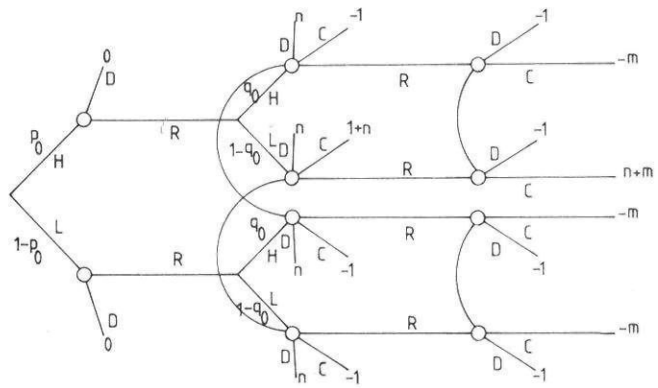
The following model is simply used as an illustration of the recursive computation of optimal strategies and certainly not as a contribution to poker itself. The most striking simplifications are the number of players, the number of cards, the independence of the chance moves, the lack of interaction between successive plays which in practice gives rise to a survival problem.

1. *Description of the game*. The following "poker game" is a generalization of Friedman's "simple bluffing situations with possible reraise" (1971). Player I has a low card up and one card down, Player II has a high card up and one card down. If both players have either a high or a low card down, then Player II wins; otherwise the player with the high card down wins. There are n units in the pot. Player I may either drop or raise 1 unit. Then Player II may either drop, call or reraise $(m-1)$ units. Finally Player I may either drop or call.

Let p_0 and q_0 be the respective probabilities that Player I and Player II have a high card down. Of course each player knows his own card down.

The game has the following extensive form and Player I is the maximizer, where the letters D, C, R stand for drop, call or raise, respectively.

2. *Computation of the value*. We shall make the computation stage by stage using the recursive formula and represent the successive functions on the unit square $(0 \leqslant p \leqslant 1, 0 \leqslant q \leqslant 1)$.



Here V^{RRD}, V^{RRC}, V^{RC}, V^{RD} and V^{D} are directly obtained from the game tree. Now using Lemma 2 we get:

$$V^{RR} = \operatorname*{Cav}_{p} \operatorname{Max}(V^{RRD}, V^{RRC}),$$

$$V^{R} = \operatorname*{Vex}_{q} \operatorname{Min}(V^{RR}, V^{RC}, V^{RD}),$$

$$V = \operatorname*{Cav}_{p} \operatorname{Max}(V^{R}, V^{D}).$$

For the computation of the optimal strategies, it will be helpful to keep track of how the Cav and Vex are constructed. This will be done by labeling the corresponding sides of the rectangles. For instance for $q \in [0, (m+n+1)/(2m+n)]$, V^{RR} is a linear combination of V^{RRD} at $p = 0$ and V^{RRC} at $p = 1$, thus the respective labels D and C; for $q \geqslant (m+n+1)/(2m+n)$, $V^{RR} = V^{RRD}$ and we give the label D to both sides of the rectangle, and so on.

3. *Computation of the optimal strategies.* As an example we shall explicitly make the computation for the case

$$p_0 \in \left]0, \frac{n+1}{n+2}\right[, \qquad q_0 \in \left]0, \frac{n(m+n+1)}{(n+1)(2m+n)}\right[$$

(this case is the most interesting one because if $q \geqslant n(m+n+1)/(n+1)(2m+n)$, the first move of Player I is obviously D, and it is R in the last case, see Figure 12 and the labels corresponding to each case).

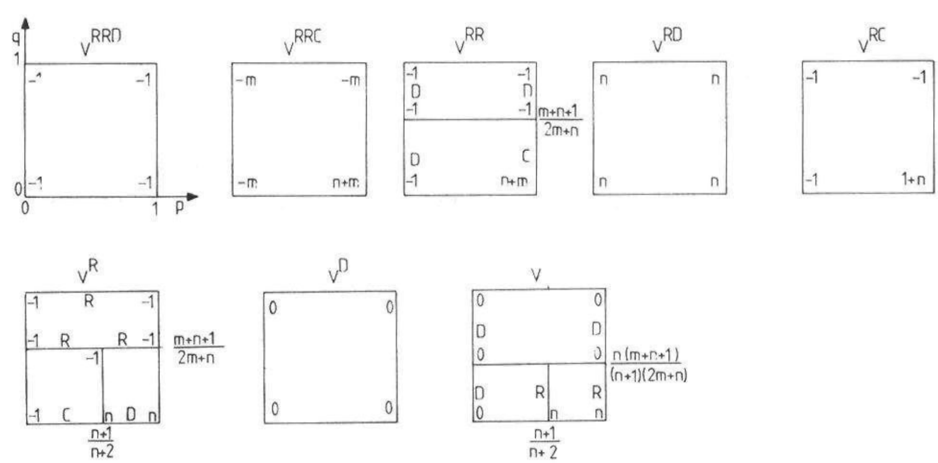


FIGURE 12. Computation of the value. (The numbers inside the squares stand for the value of the respective functions, the numbers outside give the coordinates in p and q.)

(a) *Player* I *first optimal move*. Since (p_0, q_0) is an interior point of the left lower rectangle, the construction is quite easy. The security level vector at (p_0, q_0) is a linear combination of the respective security level vectors on each side of the rectangle. Thus to satisfy (3) of Theorem 2 we proceed as follows: take

$$P_D = 0, \qquad l_D = (0,0),$$
$$P_R = (n+1)/(n+2), \qquad l_R = (n, -(n+2)m/(n+m+1))$$

(because $l_R \cdot (q, 1-q)$ equals n for $q = 0$ and 0 for $q = n(m+n+1)/(n+1) \cdot (2m+n)$).

Now, find λ_R and λ_D such that

$$\lambda_R p_R + \lambda_D p_D = p_0.$$

We obtain

$$\lambda_R = (n+2)p_0/(n+1), \qquad \lambda_D = 1 - \lambda_R.$$

By construction $\lambda_R l_R + \lambda_D l_D$ is the supporting hyperplane to $V(p_0, \cdot)$ at q_0 and all conditions are satisfied.

Player I's first optimal move is derived as:
–if H play R,
–if L play R with probability $p_0/(n+1)(1-p_0)$
play D with probability $1 - p_0/(n+1)(1-p_0)$.

(b) *Player* II'*s optimal move*. We may assume that R is played by Player I since otherwise the game ends. The supporting hyperplane to $V(\cdot, q_0)$ at p_0 is:

$$m = (0, n(n+2)/(n+1) - (n+2)(2m+n)q_0/(n+m+1)).$$

By construction m is tangent to V^R at $\tilde{p} = (n+1)/(n+2)$ so that to satisfy (3) we are interested in the rectangles of V^R that contain the point $(\tilde{p}, q_0)$. Then knowing the labels of these two rectangles, we proceed as follows: Take

$$q_{RR} = (m+n+1)/(2m+n), \qquad m_{RR} = (-1, -1),$$
$$q_{RC} = 0, \qquad m_{RC} = (-1, n+1),$$
$$q_{RD} = 0, \qquad m_{RD} = (n, n).$$

Now, find $\lambda_{RR}, \lambda_{RC}, \lambda_{RD}$ such that

$$\lambda_{RR} q_{RR} + \lambda_{RC} q_{RC} + \lambda_{RD} q_{RD} = q_0,$$
$$\lambda_{RR} m_{RR} + \lambda_{RC} m_{RC} + \lambda_{RD} m_{RD} = m.$$

The only solution appears to be

$$\lambda_{RR} = q_0(2m+n)/(m+n+1), \qquad \lambda_{RC} = 1 - \lambda_{RR} - \lambda_{RD}, \qquad \lambda_{RD} = 1/(n+1).$$

Player II's optimal move is derived as:
–if H play R,
–if L play R with probability $(m-1)q_0/(m+n+1)(1-q_0)$,
play C with probability $(n/(n+1) - q_0(2m+n)/(m+n+1))/(1-q_0)$,
play D with probability $1/(n+1)(1-q_0)$.

(c) *Player* I'*s optimal second move*. Again, we may assume that R is played by Player II since otherwise the game ends. We start with $p_{RR} = p_R = (n+1)/(n+2)$ and $l_{RR} = (n, -(n+2)m/(n+m+1))$. l_{RR} is tangent to V^{RR} at $\tilde{q} = (m+n+1)/(2m+n)$. This gives the relevant rectangles of V^{RR} (in fact all of them at this stage).

To satisfy (3) one may proceed as follows: Take

$$p_{RRC} = 1, \qquad l_{RRC} = (n+m, -m),$$

$$p_{RRD} = x, \qquad l_{RRD} = (-1, -1).$$

Find λ_{RRC}, λ_{RRD} and x such that

$$\lambda_{RRC}p_{RRC} + \lambda_{RRD}p_{RRD} = 1,$$

$$\lambda_{RRC}l_{RRC} + \lambda_{RRD}l_{RRD} = l_{RR}.$$

The only solution appears to be

$$\lambda_{RRC} = (n+1)/(n+m+1), \qquad \lambda_{RRD} = m/(n+m+1),$$

$$p_{RRD} = (n+1)(m-1)/(n+2)m.$$

Player I's optimal second move is derived as:
–if H play C with probability $(n+2)/(n+m+1)$
 play D with probability $(m-1)/(m+n+1)$
–if L play D.

(4) *Remarks.* The structure of the optimal strategies may be summarized through the values they generate on the spaces of the state variables. As a numerical example, let $m = n = 2$ and $p_0 = q_0 = 1/2$. In terms of conditional probabilities the optimal strategies would, for instance, reveal the following information: Player I's first move is R; then the probability that he has a high card goes from $1/2$ to $3/4$. Player II's move is R; then his probability of having a high card goes from $1/2$ to $5/6$. Player I's second move is a drop; his probability of having a high card goes from $3/4$ to $3/8$ (see Figure 12). But the optimal strategies do not use the information "revealed" by the opponent. They are based on a sequence of vector security levels (see Figures 13 and 14). For instance at the first stage Player I secures a vector payoff $m_0 = (4/3, -16/15)$.

In this example Corollary 2 says that for any history which occurs with a positive probability given the optimal strategies, we have:

$$V^h(p_h, q_h) = (p_h, 1-p_h) \cdot m_h = (q_h, 1-q_h) \cdot l_h$$

for example at the beginning:

$$V(p_0, q_0) = V(\tfrac{1}{2}, \tfrac{1}{2}) = \tfrac{2}{15} = \tfrac{1}{2} \times \tfrac{4}{15} + \tfrac{1}{2} \times 0 = \tfrac{1}{2} \times \tfrac{4}{3} - \tfrac{1}{2} \cdot \tfrac{16}{15}$$

and if $h = RC$, $V^{RC}(p_{RC}, q_{RC}) = 2$ with

$$p_{RC} = 3/4, \qquad m_{RC} = (3, -1),$$

$$q_{RC} = 0, \qquad l_{RC} = (-8/5, 2).$$

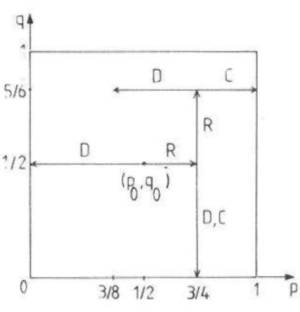


FIGURE 13. The conditional probabilities.

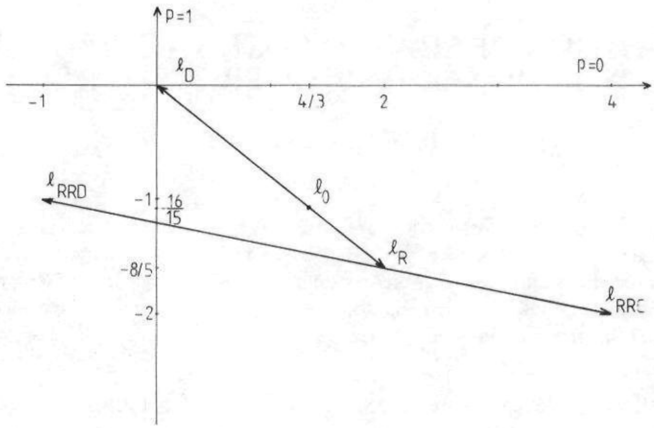


FIGURE 14. Player I's vector security levels.

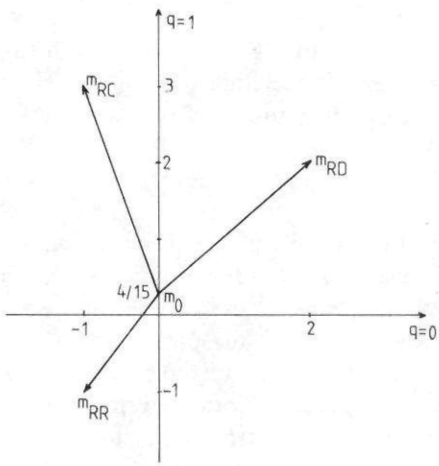


FIGURE 15. Player II's vector security levels.

References

[1] Aumann, R. J. and Maschler, M. (1966). Game Theoretic Aspects of Gradual Disarmament. Report to the U.S. A.C.D.A. ST. 80, Chapter V prepared by Mathematica, Princeton, N.J.

[2] Blackwell, D. (1956). An Analog of the Minimax Theorem for Vector Payoffs. *Pacific J. Math.* **6** 1–8.

[3] Friedman, L. (1971). Optimal Bluffing Strategies in Poker. *Management Sci.* **17** B764–B771.

[4] Kohlberg, E. (1975). Optimal Strategies in Repeated Games of Incomplete Information. *Internat. J. Game Theory* **4** 7–24.

[5] Mertens, J. F. and Zamir, S. (1980). Minmax and Maxmin of Repeted Games with Incomplete Information. *Internat. J. Game Theory* **9** 201–215.

[6] Ponssard, J. P. (1975). Zero-Sum Games with Almost Perfect Information. *Management Sci.* **21** 794–805.

[7] ———. (1976). On the Subject of Non-Optimal Play in Zero-Sum Extensive Games: the Trap Phenomenon. *Internat. J. Game Theory* **5** 107–115.

[8] ——— and Sorin, S. (1980). The L.P. Formulation of Finite Zero-Sum Games with Incomplete Information. *Internat. J. Game Theory* **9** 99–105.

[9] ——— and Zamir, S. (1973). Zero-Sum Sequential Games with Incomplete Information. *Internat. J. Game Theory*. **2** 99–107.

PONSSARD: CENTRE DE RECHERCHE EN GESTION, ECOLE POLYTECHNIQUE, 17 RUE DESCARTES, 75005 PARIS, FRANCE

SORIN: LABORATOIRE D'ECONOMETRIE, UNIVERSITE PARIS VI, 4, PLACE JUSSIEU, 75230 PARIS CEDEX 05, FRANCE