Zero-Sum Sequential Games with Incomplete Information

By Jean-Pierre Ponssard [1]) and Shmuel Zamir [2])

Abstract: Repeated zero-sum two-person games of incomplete information on one side are considered. If the one-shot game is played sequentially, the informed player moving first, it is proved that the value of the *n*-shot game is constant in *n* and is equal to the concavification of the game in which the informed player disregards his extra information. This is a strengthening of Aumann and Maschler's results for simultaneous games. Optimal strategies for both players are constructed explicitly.

Introduction

Consider two players, say player I and player II who are playing repeatedly the same zero-sum two-person game. The two players are in an asymmetric position due to the fact that one of them, say player I, knows the payoff matrix of the game while the other is uncertain about it. More precisely, he knows only the probability distribution according to which "chance" has chosen this matrix. Games of this type were studied by Aumann and Maschler (AM). Their main result can be described as follows: Let $v_n(p)$ be the (minimax) value per play of the game with n plays, as a function of the apriori probability distribution p. Then $v_n(p)$ converges to Cav $u(p)$, the concavification of $u(p)$, where $u(p)$ is the value of the game in which player I disregards completely his extra information about the game. Cav $u(p)$ is also the value of the game with infinitely many plays.

Intuitively the operation of concavification reflects in this context a behavior which is very typical to these games namely a partial revealing of information. It turns out that the best thing for player I to do is generally to reveal *only part* of his information to player II. This portion of the information is to be released during the *n* plays of the game. Another question is therefore what is the optimal rate of releasing information. This optimal rate determines the speed of convergence $v_n(p) \to$ Cav $u(p)$ and it was studied by AM and by Zamir.

AM's result is for zero-sum two-person games in which the two players move simultaneously. This includes of course games in which the moves are made sequentially: Player I moves first, player II is informed of player I's move and then makes his own move. (Such a game can be looked at as a simultaneous game after redefining the pure strategies of player II.)

[1]) J. P. Ponssard, École polytechnique, Groupe de recherche en gestion des organisations 17, rue Descartes, Paris V⁰, France.
[2]) S. Zamir, University of California, Los Angeles, Department of Mathematics, 405 Hilgard Ave. Los Angeles, Cal. 90024, U.S.A.

For sequential games AM's result can be strengthened considerably namely: $v_n(p) = \text{Cav } u(p)$ for all n, and hence also $v_\infty(p) = \text{Cav } u(p)$. There is therefore no question about the speed of convergence. This result reflects the disadvantage of player I in the sequential game compared to the simultaneous game in which $v_n(p) \geq \text{Cav } u(p)$. In sequential games, the player who is first to move has a clear disadvantage. These results were obtained by PONSSARD using convex analysis. In this paper we present a slightly different proof which has the advantage of leading to a simpler construction for the optimal strategies.

These optimal strategies are constructed for both players in the n-shot game, a task which is difficult in the general AM case where one could have only asymptotic optimal strategies. The optimal strategy for player I involves some lotteries the probability distribution of which depends on the actual game chosen by chance. The optimal strategy of player II is obtained by solving certain ordinary matrix games determined by the supporting hyperplane to the function $\text{Cav } u(p)$. The calculation of the value and the optimal strategies are demonstrated by an example.

An interesting thing about the optimal strategies, which is worth some attention, is the fact that player I has usually to use a randomized strategy although for him the game is of perfect information. When he is called to make his move his information sets are all singletones. The incomplete information of player II, not about player I's move but about the chance move, makes the sequential game very much like the simultaneous game as far as the information structure is concerned. We may also interpret this fact in terms of bluffing; that is: Whenever player I is in a weak situation, with some probability he will play as if he were in a strong situation and vice versa. Actually, games with incomplete information have closed connections with poker games. The interested reader is referred to the original work of PONSSARD.

1. The Model

Let $A^1 \ldots A^r$ be $m \times n$ matrices viewed as payoff matrices of two-person zero-sum games in which the two players are player I and player II and the sets of pure strategies are $I = \{1, \ldots, m\}$ and $J = \{1, \ldots, n\}$ respectively. Let $K = \{1, \ldots, r\}$ and denote by a^r_{ij}, $i \in I$, $j \in J$, the elements of A^k, $k \in K$. Let P be the simplex:

$$P = \{p = (p_1 \ldots p_r) \mid p_k \geq 0 \ \forall k \in K \quad \text{and} \quad \sum_{k \in K} = 1\}.$$

For each $p \in P$ let $\Gamma_1(p)$ be the two-person zero-sum game played as follows:

Step 0.   A chance move chooses an element $k \in K$ according to the probability vector p. Player I is informed which k was chosen but player II is not.

Step 1_I.   Player I (knowing k) chooses $i \in I$. Player II is told which i was chosen.

Step 1_II.   Player II (knowing i but not k) chooses $j \in J$.

After step 1_{II}, player II pays player I a_{ij}^k. The game $\Gamma_1(p)$ will be referred to as *one-shot* sequential game. Later we will discuss the *n*-shot game $\Gamma_n(p)$ in which the play consisting of steps 1_I and 1_{II} is repeated *n* times and the payoff is made after each play according to the matrix A^k chosen by chance at stage 0. However we assume that the *payoffs are not announced* before the end of the *n* plays. We also redefine the payoff to be the average payoff per stage $\rho_n = \frac{1}{n} \sum_{m=1}^{n} h_m$, where h_m is the payoff for the *m*-th play. This is a mere normalization which enables us to compare the values of $\Gamma_n(p)$ for different values of *n*.

Clearly $\Gamma_n(p)$ has a value according to the classical minimax theorem. Denote this value by $v_n(p)$.

2. The Main Theorem

The main result of the paper is concerning the value $v_1(p)$ of the one-shot game $\Gamma_1(p)$. To state this result let us define for each $p \in P : A(p) = \sum_{k \in K} p_k A^k$. Let $a_{ij}(p)$, $i \in I$, $j \in J$, be the elements of $A(p)$. One easily realizes that $A(p)$ is the payoff matrix for a modified one-shot game, $\Delta_1(p)$, in which none of the players is informed about the choice of chance k, (or equivalently, the informed player, player I, disregards his extra information). The value of $\Delta_1(p)$ is readily found to be: $u(p) = \max_{i \in I} \min_{j \in J} a_{ij}(p)$.

Let $\text{Cav}\,u(p)$ be the concavification of $u(p)$, i.e. the smallest concave function $f(p)$ on P which satisfies $f(p) \geq u(p) \;\forall\, p \in P$.

Theorem 1:

$$v_1(p) = \operatorname*{Cav}_{p} u(p).$$

For the proof of this theorem we first prove:

Lemma 2.1:

$v_1(p)$ is a concave function on P.

Proof:

By KUHN's theorem, since $\Gamma_n(p)$ is a game or perfect recall, we may think of the strategies of the players as behavioral strategies. The sets of behavioral strategies in $\Gamma_1(p)$ are X for player I and Y for player II defined by:

$$X = \{x = (x_1^1, \dots, x_m^1; \dots; x_1^r, \dots, x_m^r) \,|\, \sum_{i \in I} x_i^k = 1 \;\forall\, k \in K; x_i^k \geq 0 \;\forall\, i \;\forall\, k\}$$

$$Y = \{y = (y_1^1, \dots, y_n^1; \dots; y_1^m, \dots, y_n^m) \,|\, \sum_{j \in J} y_j^i = 1 \;\forall\, i \in I; y_j^i \geq 0 \;\forall\, i \;\forall\, j\}.$$

In words the strategy x for player I is to play the mixed strategy $(x_1^k, \dots, x_m^k)$ if the chance choice is *k*. The strategy y for player II is to play the mixed strategy

$(y_1^i, \ldots, y_n^i)$ if player I annonces choice i. The expected payoff as x and y are used is $H(p, x, y)$ given by:

$$H(p, x, y) = \sum_k \sum_i \sum_j p_k x_i^k y_j^i a_{ij}^k. \qquad (2.1)$$

By the minimax theorem:

$$v_1(p) = \min_y \max_x H(p, x, y) = \min_y \max_x \sum_k p^k \sum_i x_i^k \sum_j y_j^i a_{ij}^k$$

$$= \min_y \sum_k p^k (\max_i \sum_j y_j^i a_{ij}^k)$$

and this is a concave function being the minimum of linear functions in p.

Proof of Theorem 1:

Since the only difference between the games $\Gamma_1(p)$ and $\Delta_1(p)$ is that the set of strategies for player I in $\Delta_1(p)$ is a subset of X (namely $X^1 = \{x \in X \mid x_i^k = x_i^{k'} \ \forall i, k, k'\}$), one has the following relation between the values:

Hence by Lemma (2.1): $v_1(p) \geq u(p)$.

$$v_1(p) \geq \operatorname{Cav} u(p). \qquad (2.2)$$

To prove the opposite inequality, for any $x \in X$ and $y \in Y$ let:

$$\bar{x}_i = \sum_k p_k x_i^k; p_k^i = \begin{cases} p_k x_i^k / \bar{x}_i & \text{if } \bar{x}_i \neq 0 \\ \text{arbitrary} & \text{if } \bar{x}_i = 0 \end{cases}; \quad \forall i, k. \qquad (2.3)$$

It follows directly from (2.3) that $p^i \in P$ whenever $\bar{x}_i \neq 0$ and

$$E(p^i) = \sum_i \bar{x}_i p^i = p. \qquad (2.4)$$

Combining (2.1) and (2.3) we get

$$H(p, x, y) = \sum_i \sum_j \sum_k \bar{x}_i p_k^i y_k^i a_{ij}^k = \sum_i \sum_j \bar{x}_i y_j^i a_{ij}(p^i).$$

Hence by the minimax theorem:

$$v_1(p) = \max_x \min_y H(p, x, y) = \max_x \sum_i \bar{x}_i \min_j a_{ij}(p^i)$$

$$\leq \max_x \sum_i \bar{x}_i \max_v \min_j a_{v,j}(p^i) = \max_x \sum_i x_i u(p^i).$$

But by (2.4) $\sum_i x_i u(p^i) \leq \operatorname{Cav}_p u(p)$. Hence $v_1(p) \leq \operatorname{Cav}_p u(p)$.

3. Optimal Strategies

In this section we give explicit detailed description of optimal strategies for both players. As is often the case in such games the optimal strategies are essentially different for the two players. This is very much expected in view of the asymmetric roles of the two players.

The optimal strategy for player I is actually the strategy described by AUMANN and MASCHLER and by MERTENS-ZAMIR in a slightly different context. It is based on the following observations:

Lemma 3.1:

Let $p, p^1, \ldots, p^r$ be $r + 1$ points in P such that $p = \sum_k \lambda_k p^k$, $\lambda_k \geq 0$ $\sum_k \lambda_k = 1$, then there are r lotteries with the same set of possible outcomes $(O_1 \ldots O_r)$ with the following property: If k is chosen according to p and then the k-th lottery is performed then:

(a) The total probability of outcome O_v is λ_v, $v = 1, \ldots, r$.

(b) The conditional probability distribution on the chosen r given the outcome O_v is p^v, $v = 1, \ldots, r$.

The proof of this lemma which is straight forward is given in MERTENS-ZAMIR (Lemma 2 p. 46).

Lemma 3.2:

If player I announces his strategy σ and then does a move i after hearing k, let $p(\sigma, i)$ be the conditional distribution on K given σ and i. Then the value of the game will not change if we modify the rules to let the payoff be made according to a payoff matrix newly chosen according to $p(\sigma, i)$.

This was proved in MERTENS-ZAMIR (Theorem 3.1, p. 45).

Now for any $p \in P$ let $v_1(p) = \underset{p}{\mathrm{Cav}}\, u(p) = \sum_{v=1}^{r} \lambda_v u(p^v)$, where $p^v \in P$, $\lambda_v \geq 0$, $\sum_v \lambda_v = 1$, $\sum_v \lambda_v p^v = p$. Let $L_1, \ldots, L_r$ be the lotteries defined by lemma 3.1 for $\{\lambda_v, p_v\}$. The strategy for player I can now be described as follows:

After hearing k chosen by chance perform the lottery L_k if the outcome is O_v, play optimally in the game $\Delta_1(p^v)$.

To see that this strategy is optimal we notice that the conditional probability distribution on K given the strategy and given O_v is p^v. By Lemma 3.2 we may assume that the payoff matrix is newly chosen according to p^v. Hence by playing optimally in $\Delta_1(p^v)$ player I guarantees $u(p^v)$. Since the total probability of O_v is λ_v, player I guarantees at least $\sum_v \lambda_v u(p^v) = \mathrm{Cav}\, u(p)$.

For the *optimal strategy of player II* consider $p^\circ = (p_1^\circ, \ldots, p_r^\circ) \in P$ and let $\alpha_1 p_1 + \cdots + \alpha_r p_r = z$ be a supporting hyperplane to $\mathrm{Cav}\, u(p)$ at p°. i.e.

$$\sum_k \alpha_k p_k \geq \mathrm{Cav}\, u(p) \quad \forall p \in P$$

$$\sum_k \alpha_k p_k^\circ = \mathrm{Cav}\, u(p)\big|_{p=p^\circ} \tag{3.1}$$

For each $i \in I$ consider the 2-person zero-sum game G_i defined by the $r \times n$ matrix $(a_{ij}^k - \alpha_k)$, $k = 1, \ldots, r, j = 1, \ldots, n$.

Lemma 3.3:

$\forall i \in I$ the value of G_i is ≤ 0.

Proof:

For any $p \in P$, if player I plays i independently of k, his payoff in $\Gamma_1(p)$ will be $\min_j \sum_j p_k a_{ij}^k$. But this is evidently not more than the value of $\Gamma_1(p)$ which is $\text{Cav} u(p)$.

Hence:

$$\forall i \in I \ \forall p \in P \ \exists j \in J \text{ s.t. } \sum_k p_k a_{ij}^k \le \text{Cav} u(p) \le \alpha_1 p_1 + \cdots + \alpha_k p_k,$$

or:

$$\sum_k p_k(a_{ij}^k - \alpha_k) \le 0.$$

Interpreting p as a mixed strategy in G_i this means that the value of G_i is ≤ 0 for all $i \in I$.

Corollary (3.4):

For each $i \in I$ there is a mixed strategy $y^i = (y_1^i, \ldots, y_n^i)$ such that $\sum_j a_{ij}^k y_j^i \le \alpha_k$ for all $k \in K$.

Proof:

Such a strategy is any optimal strategy for player II in G_i since it guarantees: $\forall k \in K: \sum_j y_j^i(a_{ij}^k - \alpha_k) \le 0$ which is $\sum_j a_{ij}^k y_j^i \le \alpha_k$.

The behavioral strategy $y = (y^1, \ldots, y^m)$, where y^i is the mixed strategy determined by Corollary (3.4), is an optimal strategy for player II in $\Gamma_1(p^\circ)$. To see that let $x \in X$ then:

$$H(p^\circ, x, y) = \sum_k \sum_i \sum_j p_k^\circ x_i^k a_{ij}^k y_j^i = \sum_k p_k^\circ \sum_i x_i^k \sum_j a_{ij}^k y_j^i$$

$$\le \sum_k p_k^\circ \sum_i x_i^k \alpha_k = \sum_k p_k^\circ \alpha_k = \text{Cav} u(p)|_{p = p^\circ}.$$

4. The *n*-Shot Game

We turn now to the *n*-shot game $\Gamma_n(p)$ in which the game chosen by chance is played repeatedly n times by the same players. Recall that the payoff in $\Gamma_n(p)$ is defined as the average payoff per stage, hence we expect $v_n(p)$ to be of the same order of magnitude. However a much stronger statement can be made:

Theorem 2:

$$v_n(p) = v_1(p) \quad \text{for} \quad n = 2, 3, \ldots, . \tag{4.1}$$

Proof:

An optimal strategy for player I in $\Gamma_1(p)$ described in the previous section with the understanding that the lottery L_k is made only once in the beginning and if the outcome is O_v player I plays in all stages optimally in $\Delta_1(p^v)$. By exactly the same arguments that we had for $\Gamma_1(p)$ we get that this strategies guarantee $\text{Cav} u(p)$ per play and hence $v_n(p) \ge \text{Cav} u(p)$. Similarly if player II plays in each stage the optimal strategy for $\Gamma_1(p)$ he guarantees $\text{Cav} u(p)$ per play and hence $v_n(p) \le \text{Cav} u(p)$.

The relation of our results to those of AUMANN and MASCHLER about simultaneous games now becomes clear. Since any sequential game can be written as a simultaneous game by redefining the sets of strategies we have by AUMANN and MASCHLER:

$$\lim_{n \to \infty} v_n(p) = \text{Cav}\, u(p). \qquad (4.2)$$

However we see that for sequential repeated games (4.2) can be replaced by a much stronger statement, namely $v_n(p) = \text{Cav}\, u(p)$. We remark that if we define the infinite stage game $\Gamma_\infty(p)$ as it was defined in AM we have readily that the optimal strategies that we described for $\Gamma_n(p)$ are also optimal in $\Gamma_\infty(p)$ and hence $\Gamma_\infty(p) = \text{Cav}\, u(p)$, a fact that also follows from AM results.

5. An Example

In our example $r = m = n = 2$,

$$A^1 = \begin{pmatrix} \dfrac{5}{2} & -2 \\ 12 & 0 \end{pmatrix} \qquad A^2 = \begin{pmatrix} 1 & 10 \\ 0 & 4 \end{pmatrix}$$

Solution:

Let p' denote $1 - p$ for $0 \le p \le 1$. In $\Gamma_n(p)$ the probability distribution according to which A^k is chosen is (p,p'). Clearly, the scalar $p(0 \le p \le 1)$ can replace the vector (p,p') as a variable of $\Gamma_n(p)$, $\Delta_1(p)$, $u(p)$ and $v(p)$.
$\Delta_1(p)$ corresponds to the matrix

$$\begin{pmatrix} \dfrac{5}{2} p + p' & -2p + 10 p' \\ 12 p & 4 p' \end{pmatrix}, \quad 0 \le p \le 1.$$

Hence:

$$u(p) = \max \left\{ \min \left(\frac{5}{2} p + p', -2p + 10 p' \right), \min(12 p, 4 p') \right\} \qquad (5.5)$$

The function $u(p)$ and its concavification $v(p)$ are shown in Fig. 1. The only values of p for which $v(p) = u(p)$ are 0, $\frac{1}{4}$, $\frac{2}{3}$ and 1.

Let us demonstrate the computation of optimal strategies say for $0 \le p \le \frac{1}{4}$. For these values of p:

$$v(p) = (1 - 4p)u(0) + 4p u(1/4).$$

Hence $P^1 = 0$, $P^2 = \frac{1}{4}$, $\lambda_1 = 1 - 4p$, $\lambda_2 = 4p$, the corresponding lotteries of Lemma (3.1) for player I are: $L_1 = (0,1)$, $L_2 = \left(1 - \dfrac{3p}{p'}, \dfrac{3p}{p'} \right)$. If the outcome is O_1 he plays optimally in $\Delta_1(0)$ which is $(1,0)$, if the outcome is O_2 he plays optimally in $\Delta_1(\frac{1}{4})$ which is $(0,1)$.
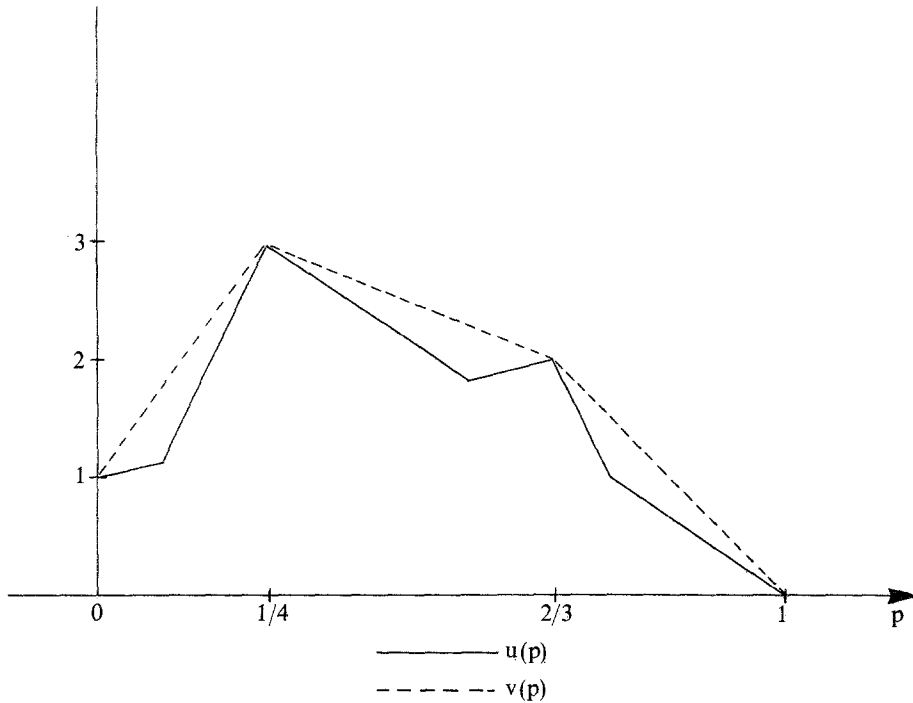
Fig. 1

For player II we need the supporting hyperplane to $v(p)$ which is $9p + p' = z$. From this we construct G_1 and G_2 for finding optimal mixed strategies for player II:

$$G_1 = \begin{pmatrix} \dfrac{5}{2} - 9 & -2 - 9 \\ 1 - 9 & 10 - 1 \end{pmatrix} = \begin{pmatrix} \dfrac{-11}{2} & -11 \\ 0 & 9 \end{pmatrix}, \text{ optimal strategy for II is } (1,0).$$

Similarly: $G_2 = \begin{pmatrix} 3 & -9 \\ -1 & 3 \end{pmatrix}$, optimal strategy for II is $\left(\dfrac{3}{4}, \dfrac{1}{4} \right)$.

By similar computations for $\frac{1}{4} \le p \le \frac{2}{3}$ and $\frac{2}{3} \le p \le 1$ we get finally the optimal strategies listed in the following tables:

Optimal strategies for player I

p	L_1 (prob O_1)	L_2 (prob O_1)	strategy if O_1	strategy if O_2
$0 \le p \le \dfrac{1}{4}$	0	$1 - \dfrac{3p}{p'}$	$(1,0)$	$(0,1)$
$\dfrac{1}{4} \le p \le \dfrac{2}{3}$	$\dfrac{2 - 3p}{5p}$	$\dfrac{6 - 9p}{5p'}$	$(0,1)$	$(1,0)$
$\dfrac{2}{3} \le p \le 1$	$\dfrac{2p'}{p}$	1	$(1,0)$	$(0,1)$

Optimal strategy for player II

p	y^1 mixed st. if $i = 1$	y^2 mixed st. if $i = 2$
$0 \leq p \leq \dfrac{1}{4}$	$(1,0)$	$\left(\dfrac{3}{4}, \dfrac{1}{4}\right)$
$\dfrac{1}{4} \leq p \leq \dfrac{2}{3}$	$\left(\dfrac{32}{45}, \dfrac{13}{45}\right)$	$\left(\dfrac{1}{10}, \dfrac{9}{10}\right)$
$\dfrac{2}{3} \leq p \leq 1$	$\left(\dfrac{4}{9}, \dfrac{5}{9}\right)$	$(0,1)$

Acknowledgement

References

AUMANN, R. J., and M. MASCHLER: Repeated Games with Incomplete Information. A survey of recent results. Report to the U.S. Arms control and Disarmament Agency, Washington, D.C. Final report on contract ACDA/ST-116 prepared by Mathematics, Princeton, N.J. September 1967, Chapter III, 287−403.

MERTENS, J. F., and S. ZAMIR: The value of Two-Person Zero-Sum Repeated Games with lack of Information on Both Sides. The Int. J. of Game Theory 1, No. 1, 39−64, 1971.

PONSSARD, J. P.: "Information Usage in Non-Cooperative Game Theory", Ph. D. Dissertation, E.E.S. Department, Stanford University, California, Sept. 1971.

ZAMIR, S.: On the Relation Between Finitely and Infinitely Repeated Gan w   plete Information. The Int. J. of Game Theory 1, No. 3, 1973.

−: On Repeated Games with General Information Function. Research pr   am in Game Theory and Mathematical Economics, Research Memorandum No. 46, Ma   969. Department of Mathematics, The Hebrew University of Jerusalem, Jerusalem Israel.   appear in the Int. J. of Game Theory.