Universität Bielefeld/IMW

Working Papers
Institute of Mathematical Economics

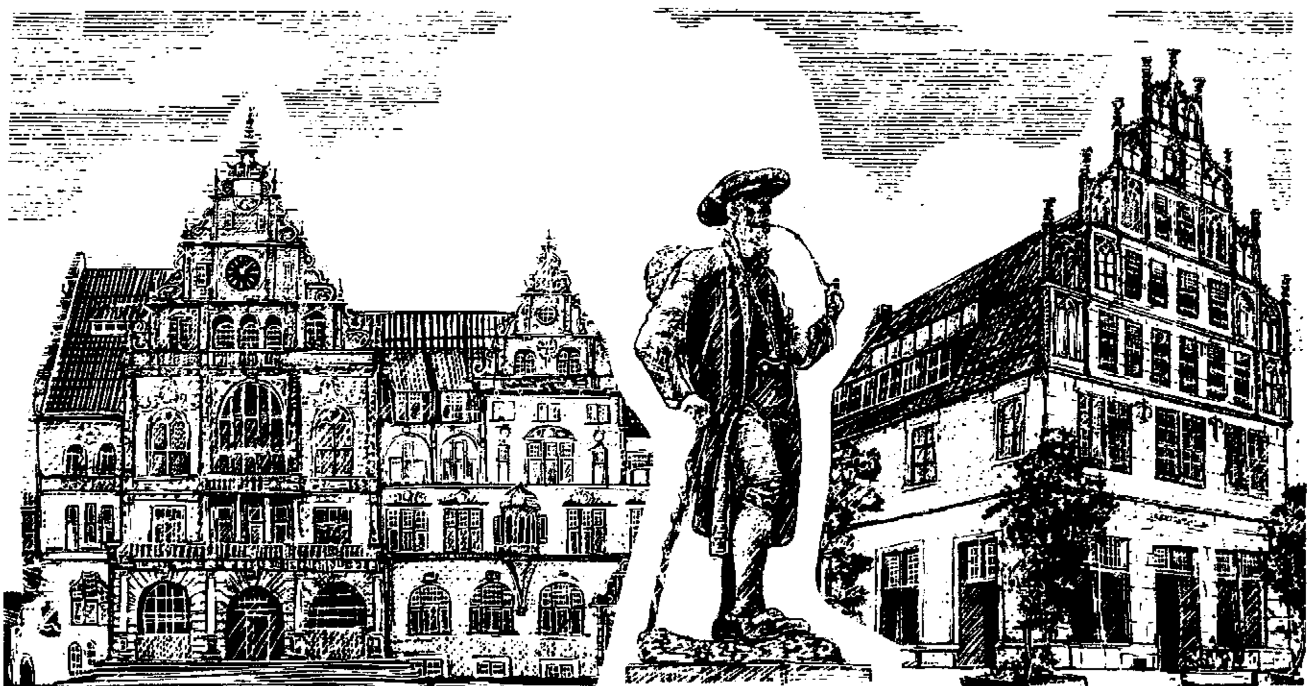
Arbeiten aus dem
Institut für Mathematische Wirtschaftsforschung

Nr. 154

An information-theoretic approach to infinitely repeated games with lack of information on both sides - the independent case -

Hans-Martin Wallmeier

March 1987



H. G. Bergenthal

Institut für Mathematische Wirtschaftsforschung
an der
Universität Bielefeld
Adresse/Address:
Universitätsstraße
4800 Bielefeld 1
Bundesrepublik Deutschland
Federal Republic of Germany

1. Introduction

The investigation of two-person zero-sum repeated games started with a series of papers by R.AUMANN and M.MASCHLER in 1966. Since then those games attracted a lot of interest, in particular because it was felt that the way of collecting information as described in those games is a central aspect of using information strategically.

The common core of the class of games to be investigated ("repeated games") here is characterized by the presupposition of a finite set of "states of nature" together with a probability distribution given thereupon. Once been chosen the state is assumed to be fixed along the duration of a play - an assumption which contrasts e.g. to the model of stochastic games. The general model of repeated games allows for - and enforces - the treatment of a series of specific cases. However, the yet available (positive) results are mainly related to the two-person zero-sum case. This case may further be classified according to the amount of information the players may acquire in the course of a play. In this exposition we shall confine ourselves to the assumption of lack of information (on the state) on both sides and of knowledge acquisition by the observation of the opponent's actions in the past. Finally we shall assume the game to be of infinite duration.

The class of games to be considered in the sequel has been dealt with before e.g. by R.AUMANN, M.MASCHLER, and R.STEARNS [67], J.F.MERTENS and S.ZAMIR [77] and, within a survey, by S.SORIN [79]. In those papers formulae for the maxmin and the minmax of the attainable payoff were given, and, using those formulae, the existence of games without a value was shown.

The reasons for re-investigation of the models are twofold.

Firstly the author estimates the available proofs to be not convincing even though the later versions are by part elaborate versions of earlier attempts. The main gap seems to be related to the separation theorem, which provides a further game with an (hopefully) existing and identical value. This two-step game may be thereafter used iteratively to give a basis for elaborating a single-stage formula for the payoff achievable in the repeated game.

Secondly, the author proposes a new method of proving one-stage formulae for both, maxmin and minmax. Its tools are provided by information-theory.

The main observation is that, given any pair of strategies, the sequence of the joint conditional probabilities on the states, given "histories" of increasing length, becomes stationary. This frees the analysis from paying attention to the dynamics contained in the optimal strategies of the repeated game, at least from some stage on.

2. The Model

Let us formally describe the games to be investigated.

Assume to be given finite sets of states of nature $\mathcal{R}$ and $\mathcal{S}$. Let $\mathcal{X}$ and $\mathcal{Y}$ denote the finite sets of actions available to the players and assume the payoff to depend on the actions and the states via

$$u: (\mathcal{R} \times \mathcal{S}) \times (\mathcal{X} \times \mathcal{Y}) \longrightarrow \mathbb{R}.$$

$u\ (r,s;\ x,y)$ is the amount given to player 1 and to be paid by player 2.

Let R and S denote random variables with values in $\mathcal{R}$ and $\mathcal{S}$, their joint distribution is given by $\mu \in \Delta\ (\mathcal{R} \times \mathcal{S})$.

We shall later assume R and S to be independent, which means

$$\bigwedge_{r,s} \mu(r,s) = \mu_1(r) \cdot \mu_2(s)$$

for some $\mu_1 \in \Delta(\mathcal{R})$, $\mu_2 \in \Delta(\mathcal{S})$.

Having provided the parameters of the game we may now define the rules according to which the game is played.

(1) At stage 0 a pair (r,s) is chosen according to μ.

(2) The strategies of the players are represented as infinite sequences of conditional probabilities $(X_t)_{t\geq 1}$ and $(Y_t)_{t\geq 1}$, respectively. Using the notation $U \mid \mathcal{M} \Longrightarrow \mathcal{U}$ for a conditional probability on $\mathcal{U}$, given $m \in \mathcal{M}$, we denote the strategy sets as

$$\Sigma_1 = \{(X_t)_{t\geq 1} / \; X_{t+1} \mid \mathcal{R} \times (\mathcal{X} \times \mathcal{Y})^t \Longrightarrow \mathcal{X}\}$$

and

$$\Sigma_2 = \{(Y_t)_{t\geq 1} / \; Y_{t+1} \mid \mathcal{S} \times (\mathcal{X} \times \mathcal{Y})^t \Longrightarrow \mathcal{Y}\}.$$

Verbally: at each stage t+1 the players may use all the actions x^t and y^t chosen previously and the knowledge on the states of nature r (player 1) and s (player 2) to find the action to be used at stage t+1.

For convenience we prefer to use random variables instead of their (conditional) distributions. Observe that the random-mechanism μ and the strategies of the players give rise to the definition of random variables on the set infinite sequences on actions $(\mathcal{X} \times \mathcal{Y})^{\infty}$. By a slight abuse of notation we may alternatively describe the strategies by means of random variables $(X_t)_{t\geq 1}$ and $(Y_t)_{t\geq 1}$ satisfying the markov-chain condition

$$\bigwedge_{t\in\mathbb{N}} (S,Y_{t+1}) \Leftrightarrow (R,X^t,Y^t) \Leftrightarrow X_{t+1}$$

(3)

$$\bigwedge_{t\in\mathbb{N}} (R,X_{t+1}) \Leftrightarrow (S,X^t,Y^t) \Leftrightarrow Y_{t+1},$$

respectively.

The symbol "$\Leftrightarrow$" is meant to indicate markov-chain property:

$$A \Leftrightarrow B \Leftrightarrow C$$

is defined to express that A is independent of C given B, or

$$\Pr\{A=a\,|B=b,\ C=c\} = \Pr\{A=a\,|B=b\}.$$

Observe that upper subscripts denote sequences, where lower subscripts denote their components, e.g.

$$x^t = (x_1,\ldots,x_t).$$

It should be noted that the restriction on behavior strategies as given above causes no loss of generality since Kuhn's theorem applies.

Continuing the description of the game we shall now provide the players' incentives to act in some way or other.

(4) Assuming states r and s been chosen and actions x and y being selected at some stage, the payoff $u(r,s;\ x,y)$ results.

Observe that this payoff is not known to the players, compare (2).

The players are assumed to be rational in as much as they are interested in the maximization of their payoff. Due to the infinite duration of the game one has to be careful about the definition of the players' aims. Observe first that the aggregate payoffs are quite generally not defined. Also the near at hand approach of average payoffs up to stages $\ldots, \tau, \tau+1, \ldots.$ is not available since its limit

$$\lim_{\tau \to \infty} \tau^{-1} E\left[\sum_{t=1}^{\tau} u(R,S;\ X_t, Y_t)\right]$$

does not exist in general. But the following definition may be used:

Definition

Player 1 can guarantee u, if

$$\bigwedge_{\epsilon > 0} \bigvee_{(X_t)} \bigvee_{T} \bigwedge_{\tau > T} \bigwedge_{(Y_t)} E[\tau^{-1} \sum_{t=1}^{\tau} u(R,S;X_t . Y_t)] \geq u-\epsilon.$$

Player 2 can guarantee u, if

$$\bigwedge_{\epsilon > 0} \bigvee_{(Y_t)} \bigvee_{T} \bigwedge_{\tau > T} \bigwedge_{(X_t)} E[\tau^{-1} \sum_{t=1}^{\tau} u(R,S;X_t , Y_t)] \leq u+\epsilon$$

Let u_* (u^*) denote the maximum (minimum) player 1 (2) can guarantee. Quite obviously $u_* \leq u^*$.

The game is said to have a value, if $u_* = u^*$.

For easier reference and for obvious reasons u_* is called the "maxmin" whereas u^* is denotes as "minmax".

The payoff function $u(\cdot)$ and the probability distribution μ give rise to a normal form game $\Gamma(\mu)$ via definition of the payoff

$$\tilde{u}(x,y) = \sum_{r,s} \mu(r,s)\ u(r,s;\ x,y).$$

The value of this game ($\mathrm{val}(\Gamma(\mu))$, as well as the value of our repeated game both heavily rely on the probability distribution μ; - but more than this is true. It comes out that a characterization of the functional dependency between the initial distribution μ and the value of the repeated game can be given in terms of the value-function for the one-shot games given above.

In fact, the following theorem may be proved:

Theorem

For a repeated game with initial distribution $\mu = \mu_1 * \mu_2$ of product form the identities

$$\max\min = (\operatorname*{cav}_{\tilde{\mu}_1} \operatorname*{vex}_{\tilde{\mu}_2} \operatorname{val} \Gamma(\tilde{\mu}_1, \tilde{\mu}_2))\ (\mu_1, \mu_2)$$

and

$$\min\max = (\operatorname*{vex}_{\tilde{\mu}_2} \operatorname*{cav}_{\tilde{\mu}_1} \operatorname{val} \Gamma(\tilde{\mu}_1, \tilde{\mu}_2))\ (\mu_1, \mu_2)$$

hold.

Already in 1967 R.Aumann and M.Maschler gave an example for which cav vex val and vex cav val did not coincide. Thereby they showed the existence of a repeated game without a value. For this and other examples the article of S.Sorin may be consulted.

The above result suggests that the players can make only limited use of their information concerning the true state of nature. It serves only as far as concavification and convexification is concerned. The heuristics behind this phenomenon may be given as follows:

Making use of private information, i.e. using a strategy depending on the true state to the ignorant player. In fact, using the knowledge on the actions chosen by the opponent makes posterior probabilities on the space of unknown states accessible. The more one player uses his information, the more information is revealed and thereby acquired by the opponent. The theorem shows that the use of information in some way or another changes the game to one in which no player makes use of his private information, since the formula apparently suggests computation of the value as an averaged one-shot game.

3. A Fundamental Result

As indicated in the previous section we shall now prove that in the course of time the repeated game is converted in such a way that neither player uses his private information any more. In fact, we shall prove that, given any pair of strategies (X,Y) the whole private information ever used by the players is revealed prior to some stage T(X,Y).

First we give some basic definitions and information-theoretical results needed in the sequel.

For random variables U,V and W we list:

(1) H(U) denotes the <u>entropy</u> of U, defined as

$$H(U) = -\sum_{u\in\mathcal{U}} \Pr\{U=u\} \log \Pr\{U=u\}$$

(2) H(V|U) denotes the <u>conditional entropy</u> of V given U, defined as

$$H(V|U) = -\sum_{u\in\mathcal{U}} \Pr\{U=u\} \sum_{v\in\mathcal{V}} \Pr\{V=v|U=u\} \log \Pr\{V=v|U=u\}$$

(3) I(U∧V) denotes the <u>mutual information</u> of U and V, defined by

$$I(U\wedge V) = H(V) - H(V|U).$$

(4) I(U∧V|W) denotes the <u>conditional mutual information</u> of U and V, given W, defined as

$$I(U\wedge V|W) = H(U|W) - H(U|V,W).$$

As basic properties we mention:

(1) $0 \leq \begin{array}{l} H(U) \\ H(U|V) \\ I(U \wedge V) \\ I(U \wedge V|W) \end{array}$

(2) $H(U|V,W) \leq \begin{array}{l} H(U|V) \\ H(U|W) \end{array}$

(3) $H(U,V|W) \geq \begin{array}{l} H(U|W) \\ H(V|W) \end{array}$

(4) $I(U \wedge V,W) = I(U \wedge V) + I(U \wedge W|V)$
(Kolmogorov-identity)

(5) U are V independent, if and only if $H(U|V) = H(U)$ and, equivalently if and only if $I(U \wedge V) = 0$.

(6) U and V are conditional independent given W, if and only if $I(U \wedge V|W) = 0$. We may equivalently use the notation: U,W and V for a markov-chain and denote this fact by U⊖W⊖V.

(7) The (conditional) mutual information and the entropy are continuous real-valued functions with respect to the probability distributions involved. No difficulties may occur since the sets $\mathcal{U}$,s $\mathcal{V}$ and $\mathcal{W}$ will all be assumed to be finite.

3.1. Definition

A sequence $(\mu_t(\cdot|z^t))_{t=1,2,...}$ of conditional probabilities is said to become (ϕ_T)-stable, if

$$\bigwedge_{\epsilon,\delta>0} \bigvee_{T} \bigwedge_{t>T} \phi_T \left(\{z^T / \|\mu_t(\cdot|z^t) - \mu_T(\cdot|z^T)\| > \delta\}\right) < \epsilon,$$

where $z^T = \text{proj}_T(z^t)$ denotes the projection of the sequence z^t on its first T components.

3.2. Proposition

Given any pair of strategies (X,Y) the sequences $(\mu(\cdot,\cdot|x^t,y^t))_{t=1,2,...}$ of conditional probabilities on $\mathcal{X} \times \mathcal{Y}$ becomes stable with respect to the marginal distributions on $\mathcal{X}^t \times \mathcal{Y}^t$ induced by $\mu \in \Delta(\mathcal{X} \times \mathcal{Y})$ and (X,Y).

Proof

Observing the conditional entropy $H(R,S|X^t,Y^t)$ to be antitonic in t and bounded from below (by zero) the sequence converges. As a Cauchy-sequence, given any $\rho > 0$ there exists $T \in \mathbb{N}$ such that for all $\tau > T$ the inequality

$$\rho > H(R,S|X^T,Y^T) - H(R,S|X^\tau,Y^\tau)$$
$$= I(R,S \wedge (X_{T+1},Y_{\tau+1},\dots X_\tau,T_\tau)|X^T,Y^T)$$

holds.

Due to the continuity of the conditional mutual information and because of (6) we get the desired result. #

From our assumption on the strategies available to the players we may easily infer a central identity:

Since $\bigwedge_{t\in\mathbb{N}} I(X_{t+1} \wedge S,Y_{t+1} \mid X^t,Y^t,R)$

$$= 0$$
$$= I(Y_{t+1} \wedge R,X_{t+1} \mid X^t,Y^t,S),$$

for any pair of strategies (X,Y) and concluding by antitonicity,

$$0 = I(Y_{t+1} \wedge R|X^t,Y^t,S)$$

we infer the central

3.3 Identity

$$\bigwedge_{t\in\mathbb{N}} I(X_{t+1}, R \wedge Y_{t+1},S \mid X^t,Y^t)$$
$$= I(R \wedge Y_{t+1}, S \mid X^t,Y^t)$$
$$= I(R \wedge S \mid X^t,Y^t).$$

We may now investigate properties of the conditional probabilities on the set of states induced in the course of a play. Those conditional probabilities are shown to be of product-type if the initial probability $\mu \in \Delta(\mathcal{R} \times \mathcal{Y})$ is a product of $\mu_1 \in \Delta(\mathcal{R})$ and $\mu_2 \in \Delta(\mathcal{Y})$.

Assume from now on R and S to be independent.

3.4 Lemma

Let (X,Y) denote any pair of strategies. Then for all $\tau \in \mathbb{N}$ the conditional probability on $\mathcal{R} \times \mathcal{S}$ under the condition (X^τ, Y^τ), is of product-type.

Proof

It is sufficient to show

$$I(R \wedge S \mid X^\tau, Y^\tau) = 0.$$

However, by recursive reasoning and using the central identity we find

$$\begin{aligned}
0 &= I(R \wedge S) \\
&= I(R, X_1 \wedge S, Y_1) \\
&\geq I(R \wedge S \mid X_1, Y_1) \\
&= I(R, X_2 \wedge S, Y_2 \mid X_1, Y_1) \\
&\geq I(R \wedge S \mid X^2, Y^2) \\
&\vdots \\
&= I(R, X_{t+1} \wedge S, Y_{t+1} \mid X^t, Y^t) \\
&\geq (R \wedge S \mid X^{t+1}, Y^{t+1}) \\
&\vdots \\
&\geq I(R \wedge S \mid X^\tau, Y^\tau). \qquad \#
\end{aligned}$$

Re-investigating the central identity 3.3 we find all the terms to be zero, provided the independency of R and S holds. Thus we find the

3.5 Condition

$$\bigwedge_{t \in \mathbb{N}} (X_{t+1}, R) \multimap (X^t, Y^t) \multimap (Y_{t+1}, S)$$

to be a surrogate for the two markovity-conditions given in (3), section 2. (condition 3.5 obviously implies both parts of (3).)

By proposition 3.2 we found that the conditional probabilities in the set of states become stable. We shall now investigate the assumptions under which the approximate result "stable" becomes true in a strict sense. To put it another way: What is the assumption on the strategies that guarantees the sequence of conditional probabilities to become stationary?

3.6 Lemma

For strategies (X,Y) the following conditions are equivalent:

(i) $$\bigwedge_{\tau > T} (R,S) \mathrel{-\!\!\circ\!\!-} (X^T,Y^T) \mathrel{-\!\!\circ\!\!-} (X_{\tau+1},Y_{T+1},\ldots X_\tau,Y_\tau)$$

and

(ii) $$\bigwedge_{t \geq T} R \mathrel{-\!\!\circ\!\!-} (X^t,Y^t) \mathrel{-\!\!\circ\!\!-} X_{t+1}$$

and

$$\bigwedge_{t \geq T} S \mathrel{-\!\!\circ\!\!-} (X^t,Y^t) \mathrel{-\!\!\circ\!\!-} Y_{t+1}$$

Proof

The markovity

$$\bigwedge_{\tau > T} (R,S) \mathrel{-\!\!\circ\!\!-} (X^T,Y^T) \mathrel{-\!\!\circ\!\!-} (X_{t+1},Y_{T+1},\ldots X_\tau,Y_\tau)$$

holds, if and only if

$$0 = I(R,S) \wedge (X_{T+1},Y_{T+1},\ldots X_\tau,Y_\tau) \mid X^T,Y^T)$$

- which is, using the Kolmogorov-identity extensively, equivalent to

$$\bigwedge_{t \geq T} (R,S) \mathrel{-\!\!\circ\!\!-} (X^t,Y^t) \mathrel{-\!\!\circ\!\!-} (X_{t+1},Y_{t+1}).$$

On the other hand, $I(R,S \wedge X_{T+1},Y_{T+1},\ldots X_\tau,Y_\tau \mid X^T,Y^T)$ may be split up to give

$$0 = H(R,S|X^T,Y^T) - H(R,S|X^\tau,Y^\tau)$$

$$= H(R|X^T,Y^T) - H(R|X^\tau,Y^\tau)$$

$$+ H(S|X^T,Y^T) - H(S|X^\tau,Y^\tau)$$

$$= I(R \wedge X_{T+1},Y_{T+1},\ldots X_{\tau}, Y_{\tau} \mid X^{T},^{T})$$

$$+ I(S \wedge X_{T+1},Y_{T+1},\ldots X_{\tau}, Y_{\tau} \mid X^{T},Y^{T}),$$

where we used lemma 3.4 to ensure the second equality.

For symmetry reasons it is sufficient to show the equivalence of

$$I(R \wedge X_{T+1}, Y_{T+1},\ldots X_{\tau}, Y_{\tau} \mid X^{T},Y^{T}) = 0$$

and $\bigwedge_{t>T} R \leftrightarrow (X^t,Y^t) \leftrightarrow X_{t+1}$. Using the Kolmogorov identity we infer

$$0 = I(R \wedge X_{T+1},Y_{T+1},\ldots X_{\tau},Y_{\tau} \mid X^{T},Y^{T})$$

$$= \sum_{t=T}^{\tau-1} I(R \wedge X_{t+1}, Y_{t+1} \mid X^{t},Y^{t})$$

$$= \sum_{t=T}^{\tau-1} I(R \wedge X_{t+1} \mid X^{t},X^{t}) + I(R \wedge Y_{t+1} \mid X^{t+1},Y^{t}).$$

As a consequence it remains to prove the equality

$$0 = I(R \wedge Y_{t+1} \mid X^{t+1},Y^{t})$$

quite generally for all strategies X,T and $t \in \mathbb{N}$.

Observe however,

$$I(R \wedge Y_{t+1} \mid X^{t+1}, Y^{t})$$

$$\leq I(R,X_{t+1} \wedge Y_{t+1} \mid X^{t},Y^{t})$$

$$= 0,$$

the last identity following by assumption (3), section 2. #

4. An alternative Representation of the Payoff

Having been occupied with the dynamic structure of plays of the game, we shall now investigate the payoff resulting from an arbitrary pair (X,Y) of strategies. It will emerge that the payoff may be divided into two parts. The first one results from the choice of actions up to some stage τ. Since this interval is finite it may be neglected from the payoff--centered point of view. The remaining portion may be viewed as being derived from playing a supergame. This is due to the fact that the conditional probabilities on later stages may be approximated with arbitrary exactness by the conditional probability already obtained at some stage T. Thus we shall argue that the true sequence may be replaced by the constant sequence of conditional probabilities on $\mathcal{R} \times \mathcal{S}$, computable.as early as T.

4.1 Lemma

Let $\epsilon > 0$ and a pair of strategies (X,Y) be given. Then there exists $T \in \mathbb{N}$ such that for all $\tau > T$

$$\epsilon > \Big| E_{RSX^\tau Y^\tau} [\tau^{-1} \sum_{t=1}^{\tau} u(R,S;\, X_t, Y_t)]$$
$$- \tau^{-1} \sum_{t=1}^{\tau} E_{X^{t-1},Y^{t-1}} [E_{X_t Y_t R_T S_T} [u(R_T, S_T; X_t, Y_t) \,|X^{t-1}, Y^{t-1}]] \Big|$$

Here the notation is as follows:

The joint distribution on $(\mathcal{R} \times \mathcal{S}) \times (\mathcal{X}^\tau \times \mathcal{Y}^\tau)$ induced by R,S and X,Y is rewritten as marginal distribution on $(\mathcal{X}^{\tau-1} \times \mathcal{Y}^{\tau-1})$ times the conditional distribution on $(\mathcal{R} \times \mathcal{S}) \times (\mathcal{X} \times \mathcal{Y})$. (R_τ, S_τ) denotes the conditional probability on $\mathcal{R} \times \mathcal{S}$ given X^τ, Y^τ.

PROOF

$$E_{RSX^{\tau}Y^{\tau}}[\tau^{-1} \sum_{t=1}^{\tau} u\ (R,S;\ X_t,Y_t)]$$

$$= \tau^{-1}\ E_{X_1Y_1}[E_{R_1S_1}[u(R_1S_1;X_1Y_1)$$

$$+E_{X_2Y_2\ldots X_{\tau}Y_{\tau}}[\sum_{t=2}^{\tau} u[R_1,S_1;X_tY_t)\ |X_1,Y_1,R_1,S_1]]]$$

$$= \tau^{-1}\ (E_{X_1Y_1}[E_{R_1,S_1}[u(R_1,S_1;X_1,Y_1)]]$$

$$+\ E_{X_1Y_1}\ [E_{R_1S_1X_2Y_2\ldots X_{\tau}Y_{\tau}}[\sum_{t=2}^{\tau} u(R_1,S_1;X_t,Y_t)(X_1,Y_t]])$$

By iteration we obtain the coincidence of the payoff with

$$\tau^{-1}\ (E_{X^1Y^1}[E_{R_1S_1}[u(R_1,S_1;\ X_1,Y_1)]]$$

$$+\ E_{X^2,Y^2}[E_{R_2S_2}[u(R_2S_2;\ X_2Y_2)\ |X^2,Y^2]]$$

$$+$$

$$\vdots$$

$$+\ E_{X^{\tau-1}Y^{\tau-1}}[E_{X_{\tau}Y_{\tau}R_{\tau}S_{\tau}}[u(R_{\tau},S_{\tau},X_{\tau}Y_{\tau}\,|X^{\tau-1},Y^{\tau-1}]])$$

$$= \tau^{-1} \sum_{t=1}^{\tau}\ E_{X^{t-1}Y^{t-1}}\ [E_{X_tY_tR_tS_t}[u(R_t,S_t;\ X_tY_t)\ |X^{t-1},Y^{t-1}]]$$

Now proposition 3.2 may be used to replace the conditional distribution (R_t,S_t) by (R_T,S_T) for all $t > T$. In fact, the terms

$$E_{X^tY^t}E_{R_tS_t}[u(R_t,S_t,X_tY_t)\ |X^tY^t]]$$

are approximately equal to

$$E_{X^tY^t}[E_{R_TS_T}[u(R_T,S_T,X_t,Y_t)\ |X^tY^t]]$$

for $t > T$, T sufficiently large and thus even the weighted sum may be replaced by a sum in which the expressions within the tail are replaced.

#

The lemma may be summarized verbally as:

Given any pair of strategies (X,Y), there exists $T \in \mathbb{N}$ such that from T on the conditional probabilities $\mu(\cdot,\cdot|x^t,y^t)$, $t \geq T$ are changing so little that, as far as the payoff is concerned, the true probabilities may each be replaced by the conditional probabilities $\mu(\cdot,\cdot|x^T,y^T)$ induced at stage T.

Assume a strategy Y of player 2 to be fixed.

A strategy $\tilde{X}$ is called <u>T-compatible (with X)</u> if

(i) $\bigwedge_{t=1,\dots,T} \tilde{X}_t = X_t$

(ii) $\bigwedge_{t>T} R \multimap (\tilde{X}^t, Y^t) \multimap \tilde{X}_{t+1}$

Looking back to the proof of lemma 3.6 we find that T-compatibility is equivalent to

$$\bigwedge_{\tau>T} R \multimap (\tilde{X}^T, Y^T) \multimap (\tilde{X}_{T+1}, Y_{T+1}, \dots \tilde{X}_\tau, Y_\tau)$$

or

$$\bigwedge_{\tau>T} H(R|\tilde{X}^\tau, Y^\tau) = H(R|\tilde{X}^T, Y^T)$$
$$(= H(R|X^T, Y^T))$$

This means that the conditional probabilities $\mu_1(\geq|x^t,y^t)$ on R become stationary and equal to $\mu_1(\geq|x^T,y^T)$ for $t > T$, where, as usual we assume ("by notation") $proj_T\ (x^t,y^t) = (x^T,y^T)$.

As a consequence the conditional distribution on R is also not affected by a variation of the strategy Y of player 2 which takes place on stages $t > T$. Thus we may denote $\tilde{X}$ to be T compatible with $(\mu_1(\geq|x^T,y^T))$, where we tacitly assume (X^T,Y^T) to be given.

We shall now prove that for any pair of strategies (X,Y) there exists T and T-compatible strategies $(\tilde{X},\tilde{Y})$ which are almost payoff equivalent to (X,Y).

4.2 Corollary

Let $\epsilon > 0$ and X,Y be given. Then there exists $T \in \mathbb{N}$ and T-compatible strategies $\tilde{X}$ (and $\tilde{Y}$, respectively) such that

$$\epsilon > |E_{RSX^{\tau}Y^{\tau}}[\tau^{-1} \sum_{t=1}^{\tau} u(R,;X_t,Y_t)]$$

$$- \tau^{-1} \sum_{t=1}^{\tau} E_{\tilde{X}^{t-1}Y^{t-1}}[E_{\tilde{X}_t,Y_t,R_T,S_T}[u(R_T,S_T;\tilde{X}_t,Y_t)\,|\tilde{X}^{t-1},Y^{t-1}]]\,|$$

(and

$$\epsilon > |E_{R,X^{\tau}Y^{\tau}}[\tau^{-1} \sum_{t=1}^{\tau} u(R,S;X_t,Y_t)]$$

$$- \tau^{-1} \sum_{t=1}^{\tau} E_{X^{t-1}\tilde{Y}^{t-1}}[E_{X_t\tilde{Y}_tR_TS_T}[u(R_T,S_T;X_t,\tilde{Y}_t)\,|X^{t-1},\tilde{Y}^{t-1}]])\,|)$$

Proof

As a consequence of the preceding lemma it is only to be shown that for sufficiently large T and t > T the expectation

$$E_{\tilde{X}^tY^t}[E_{R_TS_T}[u(R_T,S_T;\tilde{X}_t,Y_t)\,|\tilde{X}^{t-1},Y^{t-1}]]$$

is approximately equal to

$$E_{X^tY^t}[E_{R_TS_T}[u(R_TS_T;X_t,Y_t)\,|X^{t-1},Y^{t-1}]].$$

This however, follows from proposition 3.2, lemma 3.6, the continuity of $I(R,S\wedge X_{t+1}Y_{t+1}|X^{t-1},Y^{t-1})$ and its coincidence with zero, if and only if $(X_t,Y_t) \leftrightarrow (X^{t-1},Y^{t-1}) \leftrightarrow (R,S)$, in as much as thereby the nearby independency of (X_t,Y_t) and (R,S) given (X^{t-1},Y^{t-1}) is insured. Continuity of the expected payoff as function of the underlying probabilities given the desired result. #

Now, suppose (X^t,Y^t) to be given. The sets $\tilde{\Sigma}_1^T$. $\tilde{\Sigma}_2^T$ of strategies T-compatible with (X^T,Y^T) (or, with $(\mu(\cdot,\cdot|x^T,y^T))$ may be viewed as sets of strategies for supergames for which pre-play correlation takes place by observation of randomly generated sequences (x^T,y^T) preceding stage T.

Regarding (X^T,Y^T) to be fixed for a while, we may denote those supergames as a correlated supergames.
Observe that all the T-compatible strategies trivially give rise to an identical correlated supergame and are thus indistinguishable at stage T. The most efficient pursuit for the players, once x^T,y^T and $\mu(\cdot,\cdot|x^T,y^T)$ being common knowledge, is consequently the one guaranteeing an optimum payoff within the correlated supergame.

5. The value of a correlated supergame

The value of a correlated supergame will be shown to be easily derivable from a particular case of the famous "folk theorem". This result, attributed to R.AUMANN and L.SHAPLEY, establishes the coincidence of the set of equilibria of an infinitely iterated normal-form game with the set of individually rational payoffs. This characterization allows for the computation of equilibrium payoffs in the supergame, - at least in principal.

Let us give the formal definitions.

Suppose $\mathcal{N} = \{1,\dots,N\}$ to denote the set of players. For each $n \in \mathcal{N}$ a finite set of actions $\mathcal{X}_n$ available to player n is given and assume the payoff functions

$$u_n : \prod_n \mathcal{X}_n \to \mathbb{R}, \qquad n \in \mathcal{N}$$

to relate the payoffs to all players with the selected actions.

A supergame Γ^{∞} is defined as the infinite iteration of a normal-form game, such that its strategies are given as vectors $(X_t)_{t\in\mathbb{N}}$ and $(Y_t)_{t\in\mathbb{N}}$, where $X_{t+1} | \mathcal{X}^t \times \mathcal{Y}^t \Longrightarrow \mathcal{X}$ and $Y_{t+1} \,|\, \mathcal{X}^t \times \mathcal{Y}^t \Longrightarrow \mathcal{Y}$.

The concept of an equilibrium (value) is burdened with the same difficulties as for infinitely repeated games, thus the same solution is used here. We omit its formulation.

For the one-shot game we shall use the following notation: We shall denote the convex hull of the (achievable) payoff-vectors by C, i.e.

$$C = \text{conv}\ (\{u^N(x^N)\,|x^N \in \mathcal{X}^N\}).$$

Define the maximum payoff player n can guarantee for himself to be r_n,

$$r_n = \max_{X_n} \min_{X^{N-n}} \{E[u_n(X_n, X^{N-n})]\}$$

$$(= \min_{X^{N-n}} \max_{X_n} \{E[u_n(X_n, X^{N-n})]\}$$

Then the set of individually rational payoff-vectors is defined as

$$C_R = \{a^N \in C \,/\, \bigwedge_{n\in N} a_n \geq r_n\}.$$

R.AUMANN and L.SHAPLEY found

5.1 Theorem (folk theorem)

> The set of payoff-vectors induced by Nash-equilibrium strategies of the supergame Γ^{∞} coincides with the set of individually rational payoff-vectors C_R of Γ.

Whereas the set of equilibria generally is greatly enlarged by the transition from the one-shot to the infinitely iterated game, this is not the case for the class of two-person zero-sum games.

In fact, in this case the above sets and numbers reduce to

$$C = \{(a_1,a_2)/\min_{x_1,x_2} \{u(x_1,x_2)\} \leq a_1 \leq \max_{x_1,x_2} \{u(x_1,x_2)\}\ a_2 = -a_1\}$$

$$\begin{aligned} r_1 &= \max_{X_1} \min_{X_2} E[u(X_1,X_2)] \\ &= \text{val } (\Gamma) \\ &= -r_2 \end{aligned}$$

and consequently

$$C_R = \{(a_1,a_2) \ / \ a_1 \geq \text{val}(\Gamma),\ a_2 \geq -\text{val}(\Gamma),\ a_2 = -a_1\}$$

yielding

5.2 Corollary $\qquad C_R = \{(\text{val}(\Gamma), -\text{val}(\Gamma))\}.$

Verbally:

The payoff for player 1 resulting of equilibrium strategies of a zero--sum supergame is given by val(Γ), or

$$\text{val}(\Gamma^{\infty}) = \text{val}(\Gamma).$$

Suppose now correlation of strategies of a supergame to be available by joint observation of sequences $x^T \in \mathcal{X}^T$ and $y^T \in \mathcal{Y}^T$, randomly chosen. (The mechanism according to which x^T and y^T may be specified arbitrarily, known or unknown to the players.) Denote this game by Γ^t_{corr}. It is observed that this type of correlation does not change the value of the supergame.

5.3 Lemma

$$\wedge_T \text{ val } (\Gamma^T_{corr}) = \text{val } (\Gamma^\infty) \ (= \text{val}(\Gamma)).$$

Proof

A correlation svector x^T, y^T and the realization of a strategy within Γ^T_{corr} may be viewed as realization of strategies within Γ^∞. The total difference in payoff obtained from interpretation in either case is

$$\sum_{t=1}^{T} u(x_t, y_t)$$

which is negligible in as much as it is only the asymptotic behavior of the payoff, that counts.

Suppose now a strategy $X(Y)$ within Γ^∞ to guarantee some payoff a. Then its cut version $\bar{X}(\bar{Y})$ of the available strategies of Γ^T_{corr} equally guarantees a. Consequently val (Γ^∞) = val (Γ^T_{cor}). #

6. Generation of Supergames and Attainable Information

The particular supergame which is generated in the course of the play depends on a stochastic and a strategic component. The first one is due to a particular realization of the random strategies whereas the second depends on the strategies which the players decided to use. It remains to investigate the latter component.

Consider now the set of privately observable states of nature of any one of the players. Our first observation will be that the players may generate conditional probabilities on this set on their own, by playing appropriately. More precisely, for some strategies of player I the conditional probabilities on R only depends on the realization of X, but not on those of the opponent's strategy.

6.1 Lemma

Let X denote a strategy of player I additionally satisfying

$$\bigwedge_{t<\tau} Y^t \multimap (X^t, R) \multimap X_{t+1},$$

$$\text{then} \quad I(Y^\tau \wedge R \,|X^\tau) = 0.$$

Proof

By Kolmogorov's identity

$$I(Y^\tau \wedge R\,|X^\tau)$$

$$= I(Y^{\tau-1} \wedge R\,|X^\tau) + I(Y_\tau \wedge R\,|X^\tau, Y^{\tau-1}).$$

Now $\quad I(Y_\tau \wedge R\,|X^\tau, Y^{\tau-1}) \leq I(Y_\tau \wedge R, X_\tau\,|X^{\tau-1}, Y^{\tau-1})$

$$= 0$$

since strategies satisfy (3), section 2. Consequently it is sufficient to prove $0 = I(Y^{\tau-1} \wedge R\,|X^\tau)$.

The latter term is bounded from above by

$$I(Y^{\tau-1} \wedge R, X_\tau | sX^{\tau-1})$$

$$= I(Y^{\tau-1} \wedge R | X^{\tau-1}) + I(Y^{\tau-1} \wedge X_\tau | R, X^{\tau-1}).$$

By assumption the second term is equal to zero, whereby we reduced the claim to be proved from stage τ to stage τ-1. Iterating we come out with the upper bound

$I(Y_1 \wedge R | X_1) \leq I(Y_1 \wedge R, X_1) = 0$ to the term $I(Y^\tau \wedge R | X^\tau)$, proving the claim. #

Recall that $I(Y^\tau \wedge R | X^\tau)$ is equivalent to the independency of the conditional distribution on $\mathcal{R}$ form the realizations of Y^τ. Thus we find

$$\bigwedge_{y^\tau} \mu(\cdot | x^\tau, y^\tau) = \mu(\cdot | x^\tau)$$

under the above presupposition.

6.2 Lemma

The set of conditional probability vectors $(\mu_1(\cdot | x^T)_{x^T \in \mathcal{X}^T}$ asymptotically coincides with

$\{(\nu(\cdot | k)_{k \in K}) \;/\; K$ finite set, $\nu(\cdot|k)$ prob.distr.
on R for each k, $\sum_k \eta_k \; \nu(\cdot | k) = \mu_1(\cdot)$
for some probability-vector $(\eta_k)_{k \in K}\}$.

Proof

Trivially the set of conditional probabilities induced by strategies $X \in \Sigma_1$ is contained in the above set, therefore it remains to prove the converse relationship.

Assume $(\nu(\cdot|k))_{k\in K}$ satisfyiang $\sum_k \eta_k\ \nu(\cdot|k)\ \mu_1(\cdot)$ to be given. Choose T such that $|\mathfrak{X}|^T \geq |K|$ and identify somehow elements $k \in K$ with sequences $x^T \in \mathfrak{X}^T$.

Now define

$$\Pr\{X^T = x^T | R = e\} = \eta_{x^T} \frac{\nu\ (r\,|\ x^T)}{\sum_{\overline{x}} \eta_{\overline{x}^T}\ \nu\ (r\,|\overline{x}^T)} = \eta_{x^T} \frac{\nu\ (r\,|x)}{\mu(r)}$$

Given $r \in R$ this defines a probability distribution on $\mathfrak{X}^T$, the distribution of X^T.

We obtain for the conditional distribution on R induced by X^T:

$$\bigwedge_{r\in R} \mu_1(r\,|x^T) = \frac{\mu(r)\cdot\Pr\{X^T{=}x^T\,|R = r\}}{\sum_{\overline{r}} \mu(\overline{r})\ \Pr\{X^T{=}x^T\,|R = \overline{r}\}}$$

$$= \frac{\mu(r)\cdot\mu_{x^T}\ \frac{\nu(r\,|x^T)}{\mu(r)}}{\sum_{\overline{r}} \mu(\overline{r})\ \ \eta_{x^T}\ \frac{\nu(\overline{r}\,|x^T)}{\mu(\overline{r})}}$$

$$= \frac{\eta_{x^T}\ \nu(r\,|x^T)}{\eta_{x^T}\ \sum_{\overline{r}} \nu(r\,|x)}$$

$$= \nu(r\,|x^T) \qquad \#$$

Some strategic freedom of the players only remain in generating vectors of conditional probabilities. We observed above that only those vectors $(\mu_1(\cdot|k))_{k\in\mathcal{K}}$ and $(\mu_2(\cdot|l))_{l\in\mathcal{L}}$ are obtainable which satisfy the condition

$$\mathbb{E}\ [\mu_1(\cdot|K)] = \mu_1(\cdot)$$

and

$$\mathbb{E}\ [\mu_2(\cdot|L)] = \mu_2(\cdot),$$

where K and L denote random variables with values in $\mathcal{K}$ and $\mathcal{L}$. These properties give rise to additional concavifications and convexifications of the payoff-functions.

6.3 Lemma

Let $\epsilon > 0$ be given.

Let $f : \Omega \to \mathbb{R}$ denote a function defined on a compact, convex set Ω. Then there exists a finite set $\mathcal{Z} = \mathcal{Z}(\epsilon)$ such that for

$$\Sigma = \Sigma(\mathcal{Z}) = \{(Z,(g_z)_{z\in\mathcal{Z}} \mid Z \text{ random variable with value in } \mathcal{Z} \;\; \bigwedge_z g_z : \Omega \to \Omega \text{ satisfying } \bigwedge_\omega E[g_Z(\omega)] = \omega\}$$

the following inequalities hold:

(i)
$$\begin{aligned} & (\underset{\tilde\omega}{\text{vex}}\ f\ (\tilde\omega))\ (E[g_Z(\omega)]) + \epsilon \\ \geq\ & \min_{(Z,(g_z)_z)\in\Sigma} \{E[f \circ g_Z(\omega)]\} \\ \geq\ & (\text{vex}\ f(\tilde\omega))\ (E[g_Z(\omega)]) \\ =\ & (\underset{\tilde\omega}{\text{vex}}\ f(\tilde\omega))\ (\omega) \end{aligned}$$

and

(ii)
$$\begin{aligned} & (\text{cav}\ f(\tilde\omega))\ (\tilde\omega \\ =\ & (\underset{\tilde\omega}{\text{cav}}\ f(\tilde\omega))\ (E[g_Z(\omega)]) \\ \geq\ & \max_{(Z(g_z))\in\Sigma} \{E[f \circ g_Z(\omega)]\} \\ \geq\ & (\underset{\tilde\omega}{\text{cav}}\ f(\tilde\omega))\ (\omega). \end{aligned}$$

Proof

For symmetry reasons it is sufficient to consider only (i).

Since $(\operatorname{vex}_{\tilde{\omega}} f(\tilde{\omega}))(\cdot)$ is convex per definition, the inequality

$$(\operatorname{vex}_{\tilde{\omega}} f(\tilde{\omega}))(\omega) \leq \sum_z \Pr\{Z=z\}\ (f \circ g_z)(\omega) = E[(f \circ g_z)(\omega)]$$

holds for any $(Z, (g_z)) \in \Sigma$.

On the other hand, a result of Caratheodory ensures the existence of a pair $(Z, (g_z))$ such that

$$E[(f \circ g_z)(\omega)] = \sum_z \Pr\{Z=z\} \cdot (f \circ g_z)(\omega) \leq (\operatorname{vex}_{\tilde{\omega}} f(\tilde{\omega}))(\omega) + \epsilon.$$

By now we found that any pair of strategies eventually makes the interpretation of the repeated game as a correlated supergame accessible. Yet we did not investigate the strategic considerations of the players concerning the correlation in as much as different portions (X^T, Y^T) yield different supergames (with respect to their stationary distribution on states).

However observe that in section 5 we analyzed the payoff from an overlooking person's point of view. In fact we assumed that at some stage T a conditional propability $\mu(\cdot,\cdot|x^T,y^T)$ and its stationarity are common knowledge. We shall investigate this assumption, its implications and thereby shall investigate the strategial aspect of information to be acquired.

Our central question now is:

> What type of information do the players have on supergame being eventually achieved by the available strategies?

From this point on it is important to discriminate the succession in which the players select their strategies. We shall investigate the case of "max min", that is, we assume player I to be the first to announce the strategy that he will use in the play. The minimizing player II thus may use his knowledge on the strategy of player I in order to find an appropriate strategy of reaction. Of course, the case with reversed rôle, "min max", follows by duality.

Suppose the strategy X of player I been selected and, - assuming player I to be preplaying, - to be common knowledge. Looking at X as a vector $(\mathfrak{X}_r)_{r\in\mathcal{R}}$ with components $\mathfrak{X}_r$ to be used whenever state r prevails, both players may compute the conditional distribution $\mu_1(\cdot|x^t,y^t)$ on the basis of their common observation (x^t,y^t).

6.4 Proposition

Suppose player I to be the first to select a strategy. Then player II may guarantee a maximum payoff not exceeding

$$\text{cav vex val } (\Gamma(\tilde{\mu}_1,\tilde{\mu}_2))\ (\mu_1,\mu_2)$$

Proof

It is sufficient to provide a strategy for player II using this payoff for him.

The definition of player II's strategy may depend on the strategy selected by player I, as we are in the case of investigating max min.

Observe that according to the lemmata 3.2 and 3.6 any strategy of player I almost satisfies the markov chain condition

$$R \multimap (X^t,y^t) \multimap X_{t+1}$$

for t sufficiently large. Thus, in view of lemma 4.2, we may assume X to be T_X-compatible with $(\mu_1(\cdot|x^{T_X},y^{T_X}))$.

Let $\epsilon>0$ be given and suppose $(\eta_l)_{\in\mathcal{L}}$ to be such that

$$|E_{X^T,Y^T}[\text{vex val } (\Gamma(\mu_1(\cdot|x^{T_X},{}^{T_X})\cdot\tilde{\mu}_2))\ (\mu_2)]$$

$$- E_{X^T,Y^T}[\sum_{l\in\mathcal{L}}\eta_l \text{ val } (\Gamma(\mu_1(\cdot|x^{T_X},y^{T_X})\cdot\nu_2(\cdot|l)))]| < \epsilon$$

where $\mathcal{L}$ is some finite set.

According to lemma 6.2 player II can find an arbitrarily good approximation to $(\nu_2(\cdot|1))$ by conditional probabilities $(\mu_2(\cdot|y^{T_y}))$, choosing $T_y \in \mathbb{N}$ sufficiently large.

In fact, assuming Y to satisfy the conditions

(i) $$\bigwedge_{t \leq T_x} (X^t, S) \multimap (Y^t) \multimap Y_{t+1}$$

and

(ii) $$\bigwedge_{T_y \geq t \geq T_x} X^t \multimap (S, Y^t) \multimap Y_{t+1}$$

admits for selecting Y such that the conditional probabilities satisfy

$$\bigwedge_{t \leq T_x} \mu(\cdot,\cdot|x^t, y^t) = \mu_1(\cdot|x^t y^t) \cdot \mu_2(\cdot)$$

and

$$\bigwedge_{T_y \geq t \geq T_x} \mu(\cdot,\cdot|x^t, y^t) = \mu_1(\cdot|x^{T_x}, y^{T_x}) \cdot \mu_2(\cdot|y^t).$$

Consequently, assuming Y to satisfy

$$\bigwedge_{t > T_y} S \multimap (X^t, Y^t) \multimap Y_{t+1}$$

we obtain

$$\bigwedge_{t > T_y} \mu(\cdot,\cdot|x^t y^t) = \mu_1(\cdot|x^{T_x}, y^{T_x}) \cdot \mu_2(\cdot|y^{T_y}).$$

To put it other way, subsequent to stage T_y the conditional probabili-ties on $\mathcal{X}$ and $\mathcal{Y}$ remain fixed.

Let us now assume in favor of player I that he is aware of the strategy used by player II, in particular on the conditional probabilities arising. Then the assumption on T_y compatibility of both strategies yields that a commonly known correlated supergame is played.

Observe that, since the strategies do not refer to the (fictitious) variation of states, the average payoff-function

$$\sum_{r,s} \mu(r,s\,|x^{T_y},y^{T_y}) \cdot u\,(r,s;\cdot,\cdot) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$$

has to be considered.

For this payoff-function the value may be evaluated by lemma 5.3, giving at most $\operatorname{val}\,(\Gamma(\mu(\cdot,\cdot\,|x^{T_y},y^{T_y})))$ for fixed x^{T_y}, y^{T_y}, assuming $(Y_t)_{t>T_y}$ to be optimal under the above constraints. Properties (i) and (ii) of player II's strategies allow for approximation of any vector of conditional probabilities satisfying lemma 6.2, provided T_y is sufficiently large. Thus lemma 5.3 becomes applicable showing the payoff to be upperbound by

$$E_{X^{T_x},Y^{T_x}}\,[\operatorname{vex}\operatorname{val}\,(\Gamma(\mu_1(\cdot\,|X^{T_x},Y^{T_y})\ \tilde{\mu}_2))\ (\mu_2)] + \epsilon$$

the ϵ results from the quality of approximation of the optimal vector of conditional probabilities. (Remind the continuity of the value for correlated supergames with respect to the payoff-function, which in turn follows from the continuity of the value for normal form games with respect to the payoff-function.)

Observing now that by lemma 6.2 $E_{X^T,Y^t}\,[\mu_1(\cdot\,|X^T,Y^T)] = \mu_1(\cdot)$ we find as an upper bound to the payoff any number exceeding

$$\operatorname{cav}\operatorname{vex}\operatorname{val}\,(\Gamma(\tilde{\mu}_1,\tilde{\mu}_2))\ (\mu_1,\mu_2)$$

whence the claim follows. #

To obtain an upper bound for the asymptotic per stage payoff we put the preplaying player in a better position than he really is. We assumed him to know the conditional distribution $\mu_2(\cdot\,|y^T)$ and thereby the determinants of the correlated supergame. In order to give a lower bound, however, we have to stick to the actual information available to the player I and, according to this, computation of conditional probabilities is not feasible, since he is ignorant of the strategy used by player II. The usual approach to overcome this difficulty, namely upperbounding "max min" and lowerbounding "min max" in an analogous manner does not give results, since equality of "min max" and "max min", i.e. the existence of a value is not guaranteed in advance (and even is false in general). This illustrates the difference to finitely repeated games, for which the existence of a value is ensured "exogeneously".

Thus in order to find bounds to the payoff the roles of the players as to be preplaying or postplaying come out to be decisive.

An expedient for computation of payoff achievable for player I comes from D.Blackwell's theorem on approachability of sets of payoff-vectors. In order to provide Blackwell's theorem we shall need the following notation:

Assume a vector payoff-function

$$u : \mathcal{A} \times \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}^M$$

been given.

Now the definition of a vector-payoff supergame is a trivial extension of the definition of supergames given earlier.

The notion of a value does not generalize in this case. Mostly there cannot be given a partition into sets in which - using optimal strategies - the asymptotic payoff is sure to be found. This contrast jthe case of payoffs given by real numbers where the set $[\mathrm{val}(\Gamma),\infty)$ is approachable by player I, whereas $(-\infty, s\ \mathrm{val}\ \Gamma]$ is approachable by player II.

The concrete definition and results are:

Definition

A set $\mathcal{A} \subset \mathbb{R}^M$ is approachable (by player I), if

$$\bigwedge_{\epsilon>0} \bigvee_{(X_t)} \bigwedge_{\tau} \bigwedge_{(Y_t)} \epsilon > Pr_{X_t}((Y_t)\{ \bigvee_{t_o>\tau} d\,(t_o^{-1} \sum_{t=1}^{t_o} u(a,x_t y_t),A) \geq \epsilon\},$$

where $d(\cdot,\cdot)$ denotes the euclidean distance on $\mathbb{R}^M$.

Additionally we define for random variables X with values in $\mathcal{X}$

$$T(X) = \underset{y \in \mathcal{Y}}{\text{conv}} \{(E[u(a,X,y)])_{a \in \mathcal{A}}\}.$$

D.Blackwell [56] proved

Theorem

Assume $\mathcal{A}$ to be a closed subset of $\mathbb{R}^M$. If for every $\alpha \notin \mathcal{A}$ there exists a random variable X with values in $\mathcal{X}$, such that the hyperplane through $a(\alpha)$, the nearest point to α within $\mathcal{A}$, perpendicular to the line segment $\overline{\alpha a(\alpha)}$, separates α from $T(X)$, then S is approachable with $X = (X_t)$ satisfying

$$X_{t+1} = \begin{cases} X(\alpha_t) & \text{if} \quad t \geq 1 \text{ and } \\ & \alpha_t = t^{-1} \sum_{\tilde{t}=1}^{t} \alpha_{\tilde{t}} \notin A \\ \text{arbitrary} & \text{otherwise.} \end{cases}$$

Comment

Observe that this result immediately generalizes to vector-payoff supergames in which correlation may be performed by mutual observation of payoff-irrelevant actions in stages 1 to T:

Approachability is not affected by bringing forward the beginning of the payoff-relevant stages, thereby transforming the supergame endowed with preplay correlation to an ordinary supergame, compare section 5.

Using Blackwell's theorem we are now able to prove

6.5 Proposition

Among the strategies $X \in \Sigma_1$ of player I T-compatible with the vector $(\mu_1(\cdot|x^T,y^T))$ of conditional probabilities, there is one ensuring asymptotically a per stage payoff exceeding $\underset{\tilde{\mu}_2}{\text{vex}}\, (\text{val}\, \Gamma\, (\mu_1(\cdot|x^T,y^T)\,,\, \tilde{\mu}_2(\cdot)))\, (\mu_2)$.

Proof

In view of the preceding comment we may assure that the first T stages were used to guarantee $(\mu_1(\cdot|x^T,y^T))_{x^T}$. Subsequent to stage T player I is assumed to use the Blackwell strategy with initial payoff already obtained. Observe, that he may use the Blackwell strategy since the vector-payoff

$$\left(\sum_{t=1}^{T} u(r,s;x_t,y_t) + \sum_s \mu_1(r|x^T,y^T) \cdot \sum_{t=t+1}^{\tau} u(r,s;x_t,y_t)\right)_{s\in\mathscr{S}}$$

obtained from actions (x^τ, y^τ) selected up to stage τ is computable. Moreover recall that the Blackwell strategy is independent of s and thus T-compatible with $(\mu_1(\cdot|x^T,y^T))$.

We shall now define a set of payoff-vectors B such that each vector herein, if obtained as asymptotic per stage payoff, exceeds

$$\underset{\tilde{\mu}_2}{\text{vex}} \ (\text{val}\ \Gamma(\mu_1(\cdot|x^T,y^T) * \tilde{\mu}_2(\cdot))\ (\mu_2)$$

in the mean and prove that B is approachable.

Let H denote a supporting hyperplane to

$$\underset{\tilde{\mu}_2}{\text{vex}} \ (\text{val}\ \Gamma(\mu_1(\cdot|x^T,y^T) * \tilde{\mu}_2(\cdot)) \quad \text{at}\ 1\mu_2(\cdot).$$

It may be characterized by $\alpha \in \mathbb{R}^{(\mathscr{S})}$ via

$$\underset{\tilde{\mu}_2}{\text{vex}} \ (\text{val}\ \Gamma(\mu_1(\cdot|x^T,y^T) * \tilde{\mu}_2(\cdot)))\ (\mu_2) = \alpha \cdot \mu_2(\cdot)$$

and

$$\bigwedge_{\tilde{\mu}_2(\cdot)} \text{val}\ \ (\Gamma(\mu_1(\cdot|x^T,y^T) \cdot \tilde{\mu}_2(\cdot)) \le \alpha \cdot \tilde{\mu}_2(\cdot).$$

Suppose we are able to show that player I has a strategy guaranteeing him an "average" payoff exceeding α_s whenever state s prevails. Then the strategy ensures him an expected payoff at least equal to

$$\underset{\tilde{\mu}_2}{\text{vex}} \ (\text{val}\ (\Gamma(\mu_1(\cdot|x^T,y^T) * \mu_2(\cdot))))\ (\mu_2),$$

as desired. Consequently it is sufficient to prove approachability of

$B = \{(\beta_s) \mid \wedge_s \beta_s \geq \alpha_s\}$. We shall verify the presuppositions of Blackwell's theorem.

Let $\delta \notin B$ be given. Define σ to be its projection on B. The specific shape of B gives rise to some observations:
For $\delta_s > \alpha_s$ we find $\sigma_s = \alpha_s$, whereas otherwise $\sigma_s = \delta_s$. Thus $\sigma - \delta \geq 0$.
Define λ such that

$$\tilde{\mu}_2 = \lambda \cdot (\sigma - \delta) \in \Delta(\mathcal{f}).$$

Then $\tilde{\mu}_2(s) > 0$ enforces $\sigma_s - \delta_s > 0$ which in turn yields $\sigma_s = \alpha_s$ by the above observation.

Now set

$$\tilde{H} = \{\beta / \tilde{\mu}_2 \ (\beta - \sigma) = 0\},$$

this hyperplane contains σ.

We find, according to the non-negativity of σ-δ and $\tilde{\mu}$:

$$\wedge_s \tilde{\mu}_2(s) \cdot (\delta_s - \sigma_s) \leq 0,$$

and, since $\delta \notin B$, using the shape of B again we are sure of the existence of $s_o \in S$ such that $\delta_{s_o} < \sigma_{s_o}$ thereby finding $\tilde{\mu}_2(s_o) > 0$ and

$$\tilde{\mu}_2 \cdot (\delta - \sigma) < 0.$$

We shall verify at least that $\tilde{H}$ separates δ from T(X) for some appropriate X.
Let $\tilde{X}$ be an optimal strategy of player I in the one-shot game $\Gamma(\mu_1(\cdot|x^T, y^T) * \tilde{\mu}_2(\cdot))$. This strategy may be computed by player I. In fact he is aware of δ, the intermediate per stage payoff-vector obtained from the actions up to the stage in question and, since α is known as depending on $\mu_2(\cdot)$, he may additionally find the projection of δ on B.

Now for the strategy $\hat{X}$ we find

$$\bigwedge_{Y} \sum_{r,s} \mu_1(r|x^T,y^T) \cdot \tilde{\mu}_2(s) \cdot E\,[u(r,s;\hat{X},Y)]$$
$$\geq \text{val}\ (\Gamma(\mu_1(\cdot|x^T,y^T) * \tilde{\mu}_2(\cdot)))$$
$$\geq \alpha \cdot \tilde{\mu}_2(\cdot).$$

Consequently $\hat{H}$ separates δ from

$$\text{conv}\ \{\sum_{y\in\mathcal{Y}}\ \sum_{r,s} \mu_1(r|x^T,y^T)\ \tilde{\mu}_2(s)\ E\,[u(r,s,\hat{X},y)]\}$$

which was to be shown. #

Lemma 6.5 guarantees some vector-payoff for player I depending on the stationary conditional distribution obtained at some stage T. we shall now have to investigate optimization with respect to the stationary condition distribution.

6.5 Corollary

Suppose player I to be the first to select a strategy. Then he may guarantee a minimum payoff at least equal to

$$\text{cav vex (val } (\Gamma(\tilde{\mu}_1,\tilde{\mu}_2)))\quad (\mu_1,\mu_2).$$

Proof

Assume $(\sigma_k)_{k\in\mathcal{K}}$ to be such that

$$\sum_{k\in\mathcal{K}} \sigma_k \ \underset{\tilde{\mu}_2}{\text{vex}}\ \text{val}\ \Gamma\ (\nu_1(\cdot|k),\ \tilde{\mu}_2)\quad (\mu_2)$$
$$= (\text{cav vex val } (\Gamma(\tilde{\mu}_1,\tilde{\mu}_2)))\quad (\mu_1,\mu_2),$$

$\mathcal{K}$ finite.

By lemmata 6.1, 6.2 $(\nu_1(\cdot|k))$ may be approximated with arbitrary precision by a vector of conditional probabilities.Now using lemma 6.3 and the continuity of the value of normal form games with respect to the payoff functions, we may find for any $\epsilon > 0$ some $T \in \mathbb{N}$, such that

$$\sum_{x^T \in I^T} \Pr\{X^T = x^T\} \cdot \underset{\tilde{\mu}_2}{\text{vex}} \text{ val } (\Gamma(\mu_1(\cdot|x^T), \tilde{\mu}_2))) \ (\mu_2)$$

$$\geq (\text{cav vex val } (\Gamma(\tilde{\mu}_1, \tilde{\mu}_2))) \ (\mu_1, \mu_2) - \epsilon.$$

The preceeding lemma shows attainability of

$$\underset{\tilde{\mu}_2}{\text{vex}} \text{ val } (\Gamma(\mu_1(\cdot|x^T), \tilde{\mu}_2)) \ (\mu_1) - \epsilon$$

if x^T prevails. The expected payoff therefor emerges to be as claimed. #

REFERENCES

[66] AUMANN, R. and M.MASCHLER: Reports to the U.S. Arms Control and Disarmament Agency, prepaed by Mathematica Inc., Princeton, N.J.

[81] CSISZAR, I. and J.KÖRNER: Information Theory, Akademiai Kiadc, Budapest, (1981).

[71] MERTENS, J.-F. and S.ZAMIR: The Value of Two-Person Zero-Sum Repeated Games with Lack of Information on Both Sides, IJGT 1, 39-64, (1971/72).

[77] MERTENS, J.-F. and S.ZAMIR: Minmax and Maxmin of Repeated Games with Incomplete Information, Core DP, 7742, (1977).

[79] SORIN, S.: Une Introduction aux Jeux Repetes a Deux Joneurs, a Somme Nulle et a Information Incomplete, Cartiers du Groupe de Mathematiques Economiques (1), (1979).