Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of APPLIED MATHEMATICS AND MECHANICS
www.elsevier.com/locate/jappmathmech
Journal of Applied Mathematics and Mechanics 68 (2004) 597–605

MULTISTAGE GAMES†

L. V. GRAUER and L. A. PETROSYAN

St. Petersburg

e-mail: grauer_lidia@pochta.ru; spbuoasis7@peterlink.ru

(*Received* 14 *January* 2004)

Infinite stage and finite stage games are considered in a tree-like graph in which a certain simultaneous game corresponds to each vertex. A definition of a strong Nash transferable equilibrium is given. In the case of infinite stage games, a regularization procedure is introduced which enables a strong transferable equilibrium to be constructed. A strong transferable equilibrium is found in explicit form for the specific case o the n-person, repeated, infinite stage "Prisoner's dilemma" game. A new class of Nash equilibria, based on the use of penalty strategies, is defined in the case of finite stage games. Explicit analytical formulae are obtained for the number of stages required for the penalty. It is shown that the payoffs in a given equilibrium exceed the payoffs in the classical absolute equilibrium.

These are theorems in the literature [1–3] on the theory of multistage and repeated games, from which it follows that it is possible to construct Pareto optimal equilibria using penalty strategies. Since the authorship of these theorems has not been determined, they have been called public theorems. However, in the above-mentioned publications, these public theorems have been proved in the case of infinite stage games. The transfer of these results to the case of finite stage games came up against the impossibility of realizing penalizing strategies in the last stages of the game. A new approach was developed in [4] which uses both penalty strategies as well as the usual Nash equilibrium for repeated bimatrix games. The problem of constructing the corresponding Nash equilibria for finite stage, n-person games is completely solved in this paper.

The problem of constructing strong equilibria, that is, of equilibria which are stable with respect to the deviations of coalitions of players, is important in the game theory. It does not have a specific meaning in the classical steady-state case since, as a rule, such equilibria do not exist. Therefore, in the classical game theory, equilibria which are stable with respect to the deviations of individual players were actually considered [5, 6]. Light has been thrown on this question in the greatest detail in [7]. In this paper, we have succeeded, using the ideology of the public theorems, in constructing a strong equilibrium, making fairly broad assumptions and in obtaining the conditions for its existence in the case of infinite stage games. A multistage game, at each stage of which the n-person "Prisoner's dilemma" game [8] is played, is taken as the basic illustrative example.

1. AN INFINITE STAGE GAME

Consider an infinite, tree-like graph $G = (Z, L)$, where z is the set of vertices and $L : Z \to 2^Z$, $L(z) = L_z \subset Z$, $z \in Z$ and L_z is the set of vertices which follow after z. We shall assume that the sets L_z, $z \in Z$ are finite. A simultaneous n-person game

$$\Gamma(z) = \langle N; X_1^z, \ldots, X_n^z; K_1^z, \ldots, K_n^z \rangle$$

corresponds to each vertex $z \in Z$, where $N = \{1, \ldots, n\}$ is the set of players, which is the same for all $z \in Z$, X_i^z is the set of strategies of a player $i \in N$ (the sets X_i^z, $z \in Z$ are finite) and $K_i^z\,(x_1^z, \ldots, x_n^z)$ is the payoff function of player i ($i \in N$, $x_i^z \in X_i^z$). The game $\Gamma(z)$ is called a single-stage game.

†*Prikl. Mat. Mekh.* Vol. 68, No. 4, pp. 667–677, 2004.

0021–8928/$—see front matter.
doi: 10.1016/j.jappmathmech.2004.07.012

A transition function

$$T(z; x_1^z, \dots, x_n^z) = T(z; x^z) \in L_z \quad (T(z; x^z) \neq \varnothing \Leftrightarrow L_z \neq \varnothing)$$

is defined for each $z \in Z$. For each vertex z and situation x^z in the game $\Gamma(z)$, the function T determines the vertex z' and, correspondingly, the game $\Gamma(z')$ $(z' = T(z; x^z) \in L_z)$ at the following stage.

We define a multistage game $\bar{G}(z_0)$ in the tree $G = (Z; L)$ using the simultaneous games $\Gamma(z)$ and the transition function T as follows. A simultaneous game $\Gamma(z_0)$ occurs at the initial vertex $z_0 \in Z$ at the first stage. If a situation x^{z_0} occurs in this game, then a game $\Gamma(z_1)$, where $z_1 = T(z_0; x^{z_0})$, will take place at the following stage. If a simultaneous game $\Gamma(z_{k-1})$ has occurred at step k and a situation $x^{z_{k-1}}$ has been realized in $\Gamma(z_{k-1})$, then the game $\Gamma(z_k)$ $(z_k = T(z_{k-1}; x^{z_{k-1}}))$ will take place at the following stage.

The multistage game $\bar{G}(z_0)$ is terminated if we have $L_{z_l} = \varnothing$ at a certain stage l (in this case, $T(z_l; x^{z_l}) = \varnothing$). We call the sequence of situations which has been obtained a trajectory and the corresponding sequence of vertices $z_0, z_1, \dots, z_k, \dots$ a path in the graph.

The pure strategy of the conduct $\pi_i(y), y \in Z$ of a player $i \in N$ in the multistage game $\bar{G}$ is a function which places the pure strategy of player i in the single-stage game $\Gamma(y) : \pi_i(y) = x_i^y \in X_i^y$ in correspondence to each vertex $y \in Z$. A mixed strategy of conduct $q_i(y), y \in Z$ of a player $i \in N$ in the game $\bar{G}$ is defined as the mapping, which places the mixed strategy of player i in the single-stage game $\Gamma(y)$ in correspondence to each vertex $y \in Z$.

We define the payoff function in the game $\bar{G}(z_0)$ as the discounted sum of the gains in the single-stage games along the path which has been realized. It will include the discounting factor δ, $\delta \in (0; 1)$, since infinite paths $z_0, z_1, \dots, z_m, \dots$ can appear in the game $\bar{G}(z_0)$. Hence, the payoff is equal to

$$K_i = \sum_{m=0}^{\infty} K_i^{z_m}(x^{z_m})\delta^m, \quad i \in N \tag{1.1}$$

In order to guarantee the existence of the sum (1.1), we shall assume that all the payoffs in the single-stage games are uniformly bounded $K_i^z(x^z) < K, z \in Z$.

In the game $\bar{G}(z_0)$, the players possess complete information in the sense that, at each vertex $z \in Z$, they known the simultaneous game $\Gamma(z)$ in which they are playing and each player remembers all of the strategies chosen at the preceding vertices by all the players.

For each vertex $y \in Z$, we consider the subgames $\bar{G}(y)$ of the game $\bar{G}(z_0)$ which begin from y and are played in the subgraph $G(y) = (Z^y, L)$. Here, Z^y is the set of vertices of the subgraph $G(y)$. The payoff function of player i in the game $\bar{G}(y)$ is defined as $\sum_{l=m}^{\infty} K_i^{z_l}(x^{z_l})\delta^{l-m}$.

We will now consider the trajectory $\bar{x}^{\bar{z}_0}, \bar{x}^{\bar{z}_1}, \dots, \bar{x}^{\bar{z}_m}, \dots$ with the corresponding path $\bar{z}_0 = z_0, \bar{z}_1, \bar{z}_2, \dots$, where $\bar{z}_k = T(\bar{z}_{k-1}; \bar{x}^{\bar{z}_{k-1}})$ in which the sum of the payoffs of all the players is a maximum, that is,

$$\sum_{i \in N} \sum_{m=0}^{\infty} K_i^{\bar{z}_m}(\bar{x}^{\bar{z}_m})\delta^m = \max_{x_1^{z_0}, \dots, x_m^{z_m}, \dots} \sum_{i \in N} \sum_{m=0}^{\infty} K_i^{z_m}(x^{z_m})\delta^m = V(z_0; N) \tag{1.2}$$

We shall call this trajectory a *cooperative* trajectory (it is assumed that the maximum is attained).

For each subgame $\bar{G}(z)$, $z \in Z$, we consider a corresponding game $\bar{\bar{G}}(z) = \langle N, V(z, S)\rangle$ in the form of a characteristic function. The characteristic function $V(z; S)$, $S \subset N$ is defined as the value of the antagonistic game $\bar{G}_{S, N\setminus S}(z)$, constructed using the structure of the game $\bar{G}(z)$ between a coalition S, which plays the role of a maximizing player, and a coalition $N\setminus S$, which plays the role of a minimizing player. The payoff of player 1 (the coalition S) is defined as the sum of the payoffs of its members. It is additionally assumed that the values $V(z; S)$ exist for every $z \in Z$ and $S \subset N$. We denote the pair of optimal strategies of conduct in the game $\bar{G}_{S, N\setminus S}(z)$ by $(\bar{q}_S^z(\cdot), \bar{q}_{N\setminus S}^z(\cdot))$. Here,

$$\bar{q}_S^z(\cdot) = \{\bar{q}_i^z(\cdot); i \in S\}, \quad \bar{q}_{N\setminus S}^z(\cdot) = \{\bar{q}_i^z(\cdot); i \in N\setminus S\} \tag{1.3}$$

Hence, the pair of strategies $(\bar{q}_S^z(\cdot), \bar{q}_{N\setminus S}^z(\cdot))$ in the game $\bar{G}_{S, N\setminus S}(z)$ forms a certain situation in the game $\bar{G}(z) : \bar{q}^z(\cdot) = (\bar{q}_1^z(\cdot), \dots, \bar{q}_n^z(\cdot))$.

We consider a sequence of subgames $\bar{\bar{G}}(\bar{z}_m)$ along $z_0, \bar{z}_1, \dots, \bar{z}_m, \dots$. Bellman's equation is satisfied in the case of the characteristic function $V(\bar{z}_m, N)$ and

$$V(\bar{z}_m; N) = \max_{x^{\bar{z}_m}} \left\{ \sum_{i \in N} K_i^{\bar{z}_m}(x^{\bar{z}_m}) + \delta V(T(\bar{z}_m; x^{\bar{z}_m}); N) \right\}$$

$$V(\bar{z}_m; N) = \sum_{i \in N} K_i^{\bar{z}_m}(\bar{x}^{\bar{z}_m}) + \delta V(\bar{z}_{m+1}; N)$$

In each subgame $\bar{\bar{G}}(\bar{z}_m)$, we consider the C-kernel $C(\bar{z}_m)$ and assume that $C(\bar{z}_m) \neq \varnothing$. In each subgame $\bar{G}(\bar{z}_m)$, we construct a new characteristic function $\hat{V}(\bar{z}_m, S)$ as follows:

$$\hat{V}(\bar{z}_m; S) = \max_{x_S} \left\{ \sum_{i \in S} K_i^{\bar{z}_m}(\bar{x}^{\bar{z}_m} \| x_S) + \delta V(T(\bar{z}_m; \bar{x}^{\bar{z}_m} \| x_S); S) \right\}, \quad x_S = \{x_j, j \in S\} \tag{1.4}$$

Then

$$\hat{V}(\bar{z}_m; N) = V(\bar{z}_m; N), \quad \begin{cases} \hat{V}(\bar{z}_m; S) \geq V(\bar{z}_m; S), & S \subset N \\ \hat{V}(\bar{z}_m; S_1) \geq \hat{V}(\bar{z}_m; S_2), & S_2 \subset S_1 \end{cases} \tag{1.5}$$

The characteristic function $\hat{V}$ cannot be superadditive. In the game $\hat{G}(\bar{z}_m) = \langle N, \hat{V}(\bar{z}_m; S) \rangle$, we construct the C-kernel $\hat{C}(\bar{z}_m)$ using the new characteristic function.

It follows from relations (1.5) that $\hat{C}(\bar{z}_m) \subset C(\bar{z}_m)$ $(m = 0, 1, \ldots)$.

2. REGULARIZATION OF THE GAME $\bar{G}(z_0)$

We will assume the $\hat{C}(\bar{z}_m) \neq \varnothing$ $(m = 0, 1, \ldots)$. Suppose the sharing $\alpha = (\alpha_1, \ldots, \alpha_n) \in \hat{C}(z_0)$. The sequence of vectors $\beta = \beta_1, \ldots, \beta_l, \ldots$ $(\beta_l = (\beta_{1l}, \ldots, \beta_{nl}))$ is called the *procedure of the distribution of shares in time* (PDS) if the following conditions are satisfied

$$\alpha_i = \sum_{l=0}^{\infty} \beta_{il}\delta^l, \quad \alpha_i^m = \sum_{l=m}^{\infty} \beta_{il}\delta^{l-m}, \quad i \in N(\alpha = \alpha^0), \quad \alpha^m = (\alpha_1^m, \ldots, \alpha_n^m) \in \hat{C}(\bar{z}_m) \tag{2.1}$$

A PDS β exists for each sharing $\alpha \in \hat{C}(z_0)$ for an arbitrary sequence of shares $\alpha = \alpha^0, \alpha^1, \ldots, \alpha^m, \ldots$ $(\alpha^m \in \hat{C}(\bar{z}_m))$ and it can be defined as follows: $\beta_{im} = \alpha_i^m - \delta\alpha_i^{m+1}, i \in N$ $(m = 0, 1, \ldots)$.

We will now consider a multistage game $\hat{G}_\beta(z_0)$ which differs from the game $\bar{G}(z_0)$ solely in the values of the payoff functions along the cooperative trajectory. We will assume that, in the game $\hat{G}_\beta(z_0)$, in each single-stage game $\Gamma(\bar{z}_m)$ $(m = 0, 1, \ldots)$ the payoff in the situation $(\bar{x}_1^{\bar{z}_m}, \ldots, \bar{x}_n^{\bar{z}_m}) = \bar{x}^{\bar{z}_m}$ is defined as

$$\bar{K}_i(\bar{x}_1^{\bar{z}_m}, \ldots, \bar{x}_n^{\bar{z}_m}) = \beta_{im}, \quad i \in N, \quad m = 0, 1, \ldots$$

instead of $K_i^{\bar{z}_m}(\bar{x}^{\bar{z}_m})$ (the payoff in the game $\Gamma(\bar{z}_m)$). For all other situation $(x_1^{\bar{z}_m}, \ldots, x_n^{\bar{z}_m})$, the payoff remains unchanged: $\bar{K}_i(x_1^{\bar{z}_m}, \ldots, x_n^{\bar{z}_m}) = K_i^{\bar{z}_m}(x_1^{\bar{z}_m}, \ldots, x_n^{\bar{z}_m})$. As above, $q_i^{z_0}(\cdot), i \in N$, is the conduct strategy in the game $\hat{G}_\beta(z_0)$ and Q_i is the set of conduct strategies.

The situation in the conduct strategies $q^*(\cdot) = (q_1^*(\cdot), \ldots, q_n^*(\cdot))$ is called a *strong transferable Nash equilibrium* in the game $\hat{G}_\beta(z_0)$ if

$$\sum_{i \in S} K_i(q^*(\cdot)) \geq \sum_{i \in S} K_i(q^*(\cdot) \| q_S(\cdot)), \quad \forall S \subset N, \quad q_S(\cdot) \in \prod_{j \in S} Q_j \tag{2.2}$$

Theorem 1. The strong transferable Nash equilibrium (2.2) in the conduct strategies exists in the game $\hat{G}_\beta(z_0)$.

Proof. We consider the following situation $\hat{q}(\cdot) = (\hat{q}_1(\cdot), \ldots, \hat{q}_n(\cdot))$ in the game $\hat{G}_\beta(z_0)$

$$\hat{q}_i(\cdot) = \bar{x}_i^{\bar{z}_m} \text{ for } z = \bar{z}_m; \quad \hat{q}_i(\cdot) = \bar{q}_i^{\bar{z}_p}(z) \text{ when } z \in Z^{\bar{z}_p}; \tag{2.3}$$

the component $\hat{q}_i(\cdot)$ is arbitrary in other cases where $\Gamma(\bar{z}_p)$ is the first single-stage game from the sequence of games $\Gamma(\bar{z}_0), \dots, \Gamma(\bar{z}_m), \dots$, in which a coalition $S \subset N$ exists such that $i \notin S$ and the players $j \in S$ deviate from $\bar{x}_j^{\bar{z}_p}$; $\bar{q}_i^{\bar{z}_p}(z)$ is the ith component of the strategy $\bar{q}_{N\setminus S}^{\bar{z}_p}(\cdot)$ in the game $\bar{G}_{S,N\setminus S}(\bar{z}^p)$.

We will prove that the situation $\hat{q}(\cdot) = (\hat{q}_1(\cdot), \dots, \hat{q}_n(\cdot))$ is a strong transferable Nash equilibrium. It follows from the definition of the strategies $\hat{q}_i(\cdot)$ $(i = 1, 2, \dots, n)$ that

$$K_S(\hat{q}(\cdot)) = \sum_{i \in S} K_i(\hat{q}(\cdot)) = \sum_{i \in S} \sum_{m=0}^{\infty} \beta_{im}\delta^m = \sum_{i \in S} \alpha_i = \sum_{i \in S} \alpha_i^0$$

We will now consider the situations $(\hat{q}(\cdot)\|q_S(\cdot))$, $S \subset N$ and the payoffs in these situations. If $q_S(\cdot)$ is identical with $\hat{q}_S(\cdot)$ along $\bar{z}_0, \dots, \bar{z}_m, \dots$, then it is clear that

$$K_S(\hat{q}(\cdot)) = K_S(\hat{q}(\cdot)\|q_S(\cdot))$$

We will assume that the strategy $q_S(\cdot)$ prescribes the conduct in one of the single-stage games $\Gamma(\bar{z}_m)$ $(m = 0, 1, \dots)$, which differs from the conduct which is prescribed by the strategy $\hat{q}_S(\cdot)$. We will denote the first vertex of the path $\bar{z}_0, \dots, \bar{z}_m, \dots$, in which $q_i(\bar{z}_m) \neq \bar{x}_j^{\bar{z}_p}, j \in S$ by $\bar{z}_p$.

In this case, in the situation $(\hat{q}(\cdot)\|q_S(\cdot))$ the deviating coalition S cannot obtain more than

$$\delta^p \hat{V}(\bar{z}_p; S) = \delta^p \max_{x_S}\left\{\sum_{j \in S} K_j^{\bar{z}_p}(\bar{x}^{\bar{z}_p}\|x_S) + \delta V(T(\bar{z}_p; \bar{x}^{\bar{z}_p}\|x_S); S)\right\}$$

since, after deviating from $\bar{x}_j^{\bar{z}_p}$, $j \in S$, the coalition of players $N\setminus S$ will play against the coalition S: in accordance with the property (2.3) the players from the coalition $N\setminus S$ will play against the coalition S in the antagonistic game $\bar{G}_{S,N\setminus S}(z')$, where $z' = T(\bar{z}_p; \bar{x}^{\bar{z}_p}\|x_S)$ with a value of the game $V(T(\bar{z}_p; \bar{x}^{\bar{z}_p}\|x_S); S$.

From the definition of the PDS β (see condition (2.1)), we then obtain

$$\begin{aligned}\sum_{i \in S} K_i(\hat{q}(\cdot)) &= \sum_{i \in S}\sum_{m=0}^{\infty} \beta_{im}\delta^m = \sum_{i \in S}\left[\sum_{m=0}^{p-1} \beta_{im}\delta^m + \delta^p\alpha_i^p\right] \geq \sum_{i \in S}\left[\sum_{m=0}^{p-1} \beta_{im}\delta^m\right] + \\ &+ \delta^p \sum_{i \in S}\alpha_i^p \geq \sum_{i \in S}\sum_{m=0}^{p-1} \beta_{im}\delta^m + \delta^p \hat{V}(\bar{z}_p; S) \geq \sum_{i \in S} K_i(\hat{q}(\cdot)\|q_S(\cdot))\end{aligned} \tag{2.4}$$

The inclusion $\alpha^p \in \hat{C}(\bar{z}_p)$ has been used here and, also the fact that, in the penultimate link in the chain of inequalities (2.4), the sum over $i \in S$ is equal to the gain which the coalition S receives in the first p stages when the players from S do not deviate from the cooperative trajectory and the last term is equal to the upper boundary of the payoffs which the coalition S can obtain after having deviated. This completes the proof.

We will consider the case when $\bar{G}(z_0)$ is the infinitely repeated game $\Gamma = \Gamma(z)$. Then, $K_i^z = H_i$, $\bar{x}^z = \bar{x}$ at each stage.

In this case, it is necessary to consider a cooperative modification $\bar{\Gamma}$ of the game Γ, the characteristic function of which $V(S)$ is defined as the value of the antagonistic game $\Gamma_{S,N\setminus S}$ between a coalition S as player 1 (maximizing) and a coalition $N\setminus S$ as player 2 (minimizing) which is generated by the game Γ. We assume that the kernel in the game Γ is not empty. It is denoted by $C(\Gamma)$. The value of the game $\bar{G}_{S,N\setminus S}(z_m)$ will be equal to

$$V(\bar{z}_m; S) = \sum_{l=m}^{\infty} V(S)\delta^{l-m} = \frac{V(S)}{1-\delta} \tag{2.5}$$

and the characteristic function $\hat{V}(\bar{z}_m; S)$ can be calculated using the formula

$$\hat{V}(\bar{z}_m; S) = \max_{x_S}\left\{\sum_{j \in S} H_j(\bar{x}\|x_S) + \frac{\delta V(S)}{1-\delta}\right\}$$

(see expression (1.4) in which it is necessary to put $K_j^z = H_j, \bar{x}^z = \bar{x}$).

For any sharing $\gamma = \gamma_0 \in C(\Gamma)$, defining the PDS β as follows: $\beta_{im} = \gamma_i$ (m = 0, 1, ...), we consider the regularized game $\hat{G}_\gamma(z_0)$.

The existence of the PDS is equivalent to the non-emptiness of the kernel $\hat{C}(\bar{z}_m)$ with a characteristic function $\hat{V}(\bar{z}_m; S)$ which implies the existence of a solution of the following inequalities

$$\sum_{i \in S} \gamma_i \geq \hat{V}(S), \quad S \subset N; \quad \hat{V}(S) = (1-\delta)\max_{x_S}\left\{\sum_{j \in S} H_j(\bar{x}\|x_S)\right\} + \delta V(S) \tag{2.6}$$

Here, account has been taken of the fact that

$$\alpha_i^m = \sum_{l=m}^{\infty} \beta_{il}\delta^{l-m} = \sum_{l=m}^{\infty} \gamma_i \delta^{l-m} = \frac{\gamma_i}{1-\delta}$$

It follows from inequalities (2.6) that the sharing $\gamma \in C(\Gamma)$ must always also belong to $\hat{C}(\Gamma)$, the C-kernel, generated by the characteristic function $\hat{V}(S)$ (2.6) in the single-stage game Γ. It is clear that $\hat{V}(S) \geq V(S)$, $S \subset N$ and $\hat{C}(\Gamma) \neq \varnothing$ means that $C(\Gamma) \neq \varnothing$ also.

The characteristic function $\hat{V}(S)$ can be interpreted as a mathematical description of the maximum payoff of the coalition S in the game Γ if the coalition S deviates from cooperation with a probability δ ($\delta \in (0; 1)$).

The following theorem can be formulated.

Theorem 2. If the kernel $\hat{C}(\Gamma)$ in the single-stage game Γ, which is defined by the characteristic function (2.6), is not empty, then a strong transferable equilibrium exists for any sharing $\gamma \in \hat{C}(\Gamma)$ in the regularization $\hat{G}_\gamma$ of the game $\bar{G}(z_0)$.

3. THE REPEATED n-PERSON "PRISONER'S DILEMMA" GAME

In the n-person "Prisoner's dilemma" game, each of the n players has two strategies, which we call C and D, such that

(1) for each player, D is the predominant strategy ($D \succ C$).
(2) if all the players choose strategy D, their payoffs will be smaller than if all the players choose strategy C.

We will consider a repeated game $\bar{G}$ in which the n-person "Prisoner's dilemma" game $\Gamma = \langle N, \{X_i\}_{i \in N}, \{H_i\}_{i \in N}\rangle$ is played at each stage.

For each player i, the set of strategies X_i consists of two strategies: C and D ($X_i = \{C, D\}$); $H_i(x_1, \dots, x_n)$ is the payoff function of the player i ($i \in N, x_i \in X_i$). The payoff H_i of player i depends on which strategy (C or D) the player chooses as well as on the number of players k ($k = 0, 1, 2, \dots, n-1$) who have chosen strategy C, and it is defined as follows:

$H_i(\cdot) = c_k$ when player i chooses strategy C,

$H_i(\cdot) = d_k$ when player i chooses strategy D.

Here k is the number of players from $N\backslash\{i\}$ who have chosen strategy C (the remaining $(n-1-k)$ players have chosen strategy D) The parameters c_k, d_k satisfy the conditions

(1) $d_k > c_k$ for all $k = 0, 1, 2, \dots, n-1$, since $D \succ C$;
(2) $c_{n-1} > d_0$;
(3) $c_k \geq c_{k-1}$ and $d_k \geq d_{k-1}$ for $k = 0, 1, 2, \dots, n-1$.

Since $D \succ C$, the situation $(D, D, \dots, D)$ is a Nash equilibrium in the game Γ but it is not Pareto optimal since the situation $(C, C, \dots, C)$ is better for all players.

We will now consider a cooperative modification of the game Γ with a characteristic function $V(S)$. We have

$$V(S) = \max_{0 \leq l \leq s}(lc_{l-1} + (s-l)d_l), \quad s = |S| \tag{3.1}$$

since strategy D for each player from the coalition $N\backslash S$ is the optimal strategy of the coalition $N\backslash S$ in the game $\Gamma_{S, N\backslash S}$. Then

$$V(\{i\}) = d_0, \quad i \in N$$

$$V(N) = \max_{x \in \Pi_{i \in N} X_i} \sum_{i \in N} H_i(x) = \max_{0 < l \leq n}(lc_{l-1} + (n-l)d_l)$$

It can be shown that the C-kernel, $C(\Gamma)$, in the game Γ is not empty. To do this, it is sufficient to verify that the sharing $\gamma = (\gamma_1, \dots, \gamma_n)$ is such that $\gamma_i = V(N)/n$, $(i = 1, 2, \dots, n)$ belongs to the C-kernel $C(\Gamma)$.

Suppose $\bar{x}$ is an arbitrary situation in the game Γ for which the sum of the payoffs of the players is a maximum, that is,

$$\sum_{i \in N} H_i(\bar{x}) = \max_{x \in \Pi_{i \in N} X_i} \sum_{i \in N} H_i(x) = V(N) \tag{3.2}$$

We will define the cooperative trajectory as follows: $\bar{x}^z = \bar{x}$. We consider the characteristic function $\hat{V}(S)$ (see (2.6)) and the corresponding C-kernel $\hat{C}(\Gamma)$. In the case of an arbitrary sharing $\gamma \in \hat{C}(\Gamma)$, we consider the regularized game $\hat{G}_\gamma(z_0)$. At each step m $(m = 0, 1, \dots)$ of the game $\hat{G}_\gamma(z_0)$, in the situation $\bar{x}^z$ player i receives a payoff $\beta_{im} = \gamma$. For convenience in the presentation, the following notation is introduced

$$H(S) = \sum_{i \in S} \gamma_i, \quad K(S) = \max_{x_S \in \Pi_{i \in x_S} X_i} H_i(\bar{x} \| x_S)$$

We can formulate the following theorem.

Theorem 3. Suppose $H(S) > V(S)$ for all $S \subset N$. If the discounting factor $\delta \in (0, 1)$ is such that

$$1 > \delta \geq \max_{S \subset N \,:\, K(S) > H(S)} \frac{K(S) - H(S)}{K(S) - V(S)} \tag{3.3}$$

then the situation of a strong transferable equilibrium exists in the regularization $\hat{G}_\gamma$ of the repeated "Prisoner's dilemma" game.

To prove this, it suffices to note that, in the repeated "Prisoner's dilemma" game, the penalty is the same for all coalitions $S \subset N$ who have deviated from the cooperative trajectory (each player from the coalition $N \backslash S$ chooses strategy D).

It follows from the definition of the C-kernel $C(\Gamma)$ that $H(S) \geq V(S)$. If there is just one coalition S_0, $|S_0| < n$, for which $H(S_0) = V(S_0)$, it is impossible to punish the coalition S_0.

4. THE FINITE STAGE GAME

Consider the finite tree of the game $G = (Z, L)$. As earlier, a simultaneous (single-stage), n-person game

$$\Gamma(z) = \langle N; X_1^z, \dots, X_n^z; K_1^z, \dots, K_n^z \rangle$$

is set in correspondence to each vertex $z \in Z$.

Using the single-stage games $\Gamma(z)$ and the transition function T, we define the multistage game $\bar{G}(z_0)$ in the graph $G = (Z, L)$ in a similar way as in the case of infinitestage games. The multistage game $\bar{G}(z_0)$ is terminated if $lL_{z_l} = \varnothing$ at a certain stage (in this case, $T(z_l; x^{z_l}) = \varnothing$).

The payoff function in the game $\bar{G}(z_0)$ is defined as the sum of the payoffs in the single-stage games along the path $z_0, \dots, z_k, \dots, z_l$ which has been realized. Hence, the payoff of player i in the game $\bar{G}(z_0)$ is equal to

$$K_i = \sum_{m=0}^{l} K_i^{z_m}(x^{z_m}), \quad i \in N$$

In the game $\bar{G}(z_0)$, the players possess complete information in the sense that, at each vertex $z \in G$, they know the game $\Gamma(z)$ in which they are simultaneously playing and each player remembers all the strategies chosen at the preceding vertices by all the players.

For each vertex $y \in Z$, we consider the subgames $\bar{G}(y)$ of the game $\bar{G}(z_0)$ which start at the vertex y and are played in the subgraph $G(y) = (Z^y, L)$.

We shall assume that all the single-stage games are finite (they have finite sets of strategies). We now consider the antagonistic game $\bar{\bar{G}}_i(z)$, which is played using the structure of the game $\bar{G}(y)$ between a player i (maximizing) and the coalition $(N \backslash \{i\})$, which plays as a minimizing player. A value of this game exists in the strategy of the conduct, and we denote it by $V(y; \{i\})$. We denote the corresponding equilibrium situation in the game $\overline{\overline{G}}_i(y), y \in Z$ by $(\hat{q}_i^y(z), q_{N\setminus\{i\}}^y(z)) = (\hat{q}_1^y(z), \ldots, q_n^y(z))$.

A certain situation $\bar{x}^z = (\bar{x}_1^z, \ldots, \bar{x}_n^z)$ in the game $\Gamma(z)$ is specified, and, for each vertex $z \in Z$, we determine the following functions

$$w_i(z) = \max_{x_i^z \in X_i^z} \{K_i^z(\bar{x}^z \| x_i^z) + V[T(z; \bar{x}^z \| x_i^z); \{i\}]\} = \max_{x_i^z \in X_i^z} \{K_i^z(\bar{x}^z \| x_i^z) + V(z'; \{i\})\}$$

$w_i(z)$ is the maximum payoff which a player i can be guaranteed in a subgame $\overline{G}(z)$ if he deviates at the first stage the subgame $\Gamma(z)$ from the specified situation $\bar{x}^z = (\bar{x}_1^z, \ldots, \bar{x}_n^z)$.

We denote a certain Nash equilibrium in the strategies of the conduct in the subgame $\overline{G}(y)$ by $\bar{\pi}^y(z) = (\bar{\pi}_1^y(z), \ldots, \bar{\pi}_n^y(z))$ and the corresponding payoffs of the players in the subgame $\overline{G}(y)$ $(y \in Z)$ by $\lambda_i(y), i \in N$.

We now consider the expression

$$H_i^s(z_0, \ldots, z_l) = \sum_{m=0}^{s} K_i^{z_m}(x^{z_m}) + \lambda_i(z_{s+1}) \tag{4.1}$$

for each path $z_0, z_1, \ldots, z_l$ in the game $\overline{G}(z_0)$ and s, $0 \le s \le l$.

Theorem 4. We will assume that an s $(0 \le s \le l)$ exists such that the following conditions are satisfied for each $s' < s$

$$H_i^s(z_0, \ldots, z_l) = \sum_{m=0}^{s} K_i^{z_m}(x^{z_m}) + \lambda_i(z_{s+1}) \ge \sum_{m=0}^{s'-1} K_i^{z_m}(x^{z_m}) + w_i(z_{s'}) \tag{4.2}$$

or

$$\sum_{m=s'}^{s} K_i^{z_m}(x^{z_m}) + \lambda_i(z_{s+1}) \ge w_i(z_{s'}) \tag{4.3}$$

Then, a Nash equilibrium exists in the game $G(z_0)$ with the payoffs

$$H_i^s(z_0, \ldots, z) = \sum_{m=0}^{s} K_i^{z_m}(x^{z_m}) + \lambda_i(z_{s+1}), \quad i \in N \tag{4.4}$$

Proof. We will denote the strategy of the conduct of player i in the game $\overline{G}(z_0)$ by $u_i(z), z \in Z, i \in N$ and the corresponding payoffs by $\overline{K}_i(u_1(z), \ldots, u_n(z)) = \overline{K}_i(u(z))$. We will define the situation in the strategies of the conduct $u^*(z) = (u_1^*(z), \ldots, u_n^*(z))$ in the game $\overline{G}(z_0)$ as follows:

$$u_i^*(z) = \begin{cases} x_i^{z_m} & \text{for} \quad z = z_m, \quad m \le s \\ \bar{\pi}_i^{z_{s+1}} & \text{for} \quad z \in Z^{z_{s+1}} \\ \hat{q}_i^{z_p}(z) & \text{for} \quad z \in Z^{z_p} \end{cases}$$

where $\Gamma(z_p)$ is the first single-stage game in the sequence of games $\Gamma(z_0), \ldots, \Gamma(z_l)$ in which there is a player $j \ne i$ who has deviated from $x_j^{z_p}$. We will prove that this situation $u^*(z) = (u_1^*(z), \ldots, u_n^*(z))$ is a Nash equilibrium in the game $\overline{G}(z_0)$.

It follows from the definition of the strategies $u_i^*(z), i \in N$ and expression (4.1) that

$$K_i(u^*(\cdot)) = H_i^s(z_0, \ldots, z_l) = \sum_{m=0}^{s} K_i^{z_m}(x^{z_m}) + \lambda_i(z_{s+1})$$

We now consider the situation $(u^*(z) \| u_j(z)), j \in N$ and the payoffs in this situation.

If $u_j(z)$ coincides with $u_j^*(z)$ along $z_0, \dots, z_s$ and, for $z \in Z^{z_s+1}$, the situation in the strategies in the conduct of $u^*(z)$ and $(u^*(z) \| u_j(z))$ generate the same probability distributions in the pencil of paths with a common root $z_0, z_1, \dots, z_s$, then $\bar{K}_j(u^*(z)) = \bar{K}_j(u^*(z) \| u_j(z))$. We now assume that $u_j(z)$ is identical to $u_j^*(z)$ for $z = z_0, z_1, \dots, z_s$ but prescribes a behaviour which differs from $\bar{\pi}_j^{z_s+1}(z)$ in the subgame $\bar{G}(z_{s+1})$. Since $\bar{\pi}^{z_s+1}(z) = (\bar{\pi}_1^{z_s+1}(z), \dots, \bar{\pi}_n^{z_s+1}(z))$ is a Nash equilibrium in the subgame $\bar{G}(z_{s+1})$, the payoff of player j in this subgame cannot exceed $\lambda_i(z_{s+1})$ and, in the steps $z_0, \dots, z_s$ (in the single-stage games $\Gamma(z_0), \dots, \Gamma(z_s)$), the payoffs of player j are the same in the situations $u^*(z)$ and $(u^*(z) \| u_j(z))$. Then, in this case

$$\bar{K}_j(u^*(z)) \geq \bar{K}_j(u^*(z) \| u_j(z))$$

We will assume that $u_j(z)$ prescribes a strategy differing from $u_j^*(z)$ in one of the games $\Gamma(z_m)$, $0 \leq m \leq s$. The first vertex of the path $z_0, \dots, z_s$, at which $u_j(z_p) \neq x_j^{z_p}$, is denoted by z_p. In this case, player j in the situation $(u^*(z) \| u_j(z))$ cannot receive more than

$$\sum_{m=0}^{p-1} K_i^{z_m}(x^{z_m}) + \max_{\tilde{x}_j^{z_p} \in X_j^{z_p}} [K_i^{z_p}(x^{z_p} \| \tilde{x}_j^{z_p}) + V(T(z^p; x^{z_p} \| \tilde{x}_j^{z_p}); \{j\})] =$$
$$= \sum_{m=0}^{p-1} K_i^{z_m}(x^{z_m}) + w_j(z_p)$$

since he will be penalised by the coalition $N \backslash \{j\}$ after deviating from $x_j^{z_p}$ at stage p and since the players from $N \backslash \{j\}$, in accordance with the definition $u^*(\cdot)$, will play against him in the antagonistic game $\bar{\bar{G}}(z')$ $(z' = T(z^p; x^{z_p} \| \tilde{x}_j^{z_p} = u_j(z_p))$ with the value $V(T(z_p; x^{z_p} \| \tilde{x}_j^{z_p} = u_j(z_p)); \{j\}$. From conditions (4.2) and (4.3) of the theorem, we obtain the inequality

$$K_j(u^*(z)) \geq K_j(u^*(z) \| u_j(z)), \quad j \in N$$

which it was required to prove.

We will now consider the case of a repeated game $\bar{G}(z_0)$ when the same game Γ takes place in each stage. In this case, we introduce the value of the single-stage antagonistic game Γ_i in which player i comes out as a maximizing player and the coalition $N \backslash \{i\}$ as the minimizing player. The game proceeds according to the structure of the single-stage game Γ. The value of the game Γ_i is denoted by V_i, $i \in N$. All of the paths in the graph G have the same length M. The value of the game $\bar{\bar{G}}_i(z_k)$ is equal to $V(z_k; \{i\}) = (M-k)V_i$ and depends solely on the number of stages in the game $\bar{\bar{G}}_i(z_k)$ and is independent of the vertex z_k. Also, $\lambda_i(z_k) = (M-k)\lambda_i$, where λ_i is the payoff in a certain specified Nash equilibrium in the game Γ.

Theorem 5. We now consider a situation $x = (x_1, \dots, x_n)$ in a game and use the notation

$$H_i(x_1, \dots, x_n) = \bar{\lambda}_i, \quad i \in N$$

It is assumed that an s, $0 \leq s \leq M$ exists such that the following conditions are satisfied

$$\bar{\lambda}_i + (M-s)\lambda_i \geq \max_{\tilde{x}_i} H_i(x \| \tilde{x}_i) + (M-s)V_i, \quad i \in N \tag{4.5}$$

Then, a Nash equilibrium exists in the game $\bar{G}(z_0)$ with the payoffs.

$$s\bar{\lambda}_i + (M-s)\lambda_i, \quad i \in N$$

Theorem 2 is a corollary of Theorem 1 for repeated games. Note that, in repeated games, if conditions (4.5) are satisfied for a certain s, $0 \leq s \leq M$, they are also satisfied for all $s' < s$ (since $\lambda_i \geq V_i$) which is not true in the general case considered in Theorem 1.

5. EXAMPLES

The 3 × 3 "Prisoner's dilemma" game. We consider a 40-stage, 2-person, repeated game in which the bimatrix game

$$\Gamma : I = \left\| \begin{matrix} (8;\,8) & (0;\,0) & (0;\,20) \\ (0;\,0) & (0;\,0) & (0;\,0) \\ (20;\,0) & (0;\,0) & (1;\,1) \end{matrix} \right\|$$

is played at each stage.

Here,

$$\bar{\lambda}_1 = \bar{\lambda}_2 = 8, \quad \lambda_1 = \lambda_2 = 1, \quad V_1 = V_2 = 0$$

From conditions (4.5), we obtain that $s \le 28$. This means that a Nash equilibrium exists in the game being considered in which the players choose the situation (1.1) at the first s ($s \le 28$) stages and situation (3.3) at each stage in the last $40 - s$ ($40 - s \ge 12$) stages. In the Nash equilibrium proposed in this paper, by selecting $s = 8$, the players can obtain $8 \times 28 + 12 \times 1 = 236$ while, in a repeated Nash equilibrium, they only get $40 \times 1 = 40$, which is substantially less.

The repeated M-stage-person "Prisoner's dilemma" game $\bar{G}$ (see Section 3). For each player i the payoff in the Nash equilibrium situation in the game Γ will be $\lambda_i = H_i(D, \ldots, D) = d_0$ and is equal to the maximum payoff V_i which a player i can be guaranteed in a game Γ_i so that, in this game, the method of constructing a new class of equilibria proposed in Section 4 cannot be used.

For this reason, it is proposed that a regularization of the game be constructed. Suppose $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)$ is a situation in the game Γ such that $H_i(\bar{x}) \ge d_0$ for $i \in N$. Such a situation always exists, for example, $(C, C, \ldots, C)$. We now consider the cooperative trajectory with the corresponding path $\bar{x}^{\bar{z}_0}, \bar{x}^{\bar{z}_1}, \ldots, \bar{x}^{\bar{z}_M}$. Here, $\bar{x}^{z} = \bar{x}$ and $\bar{z}_k = T(\bar{z}_{k-1}; \bar{x}^{\bar{z}_{k-1}})$.

We construct a new tree-like graph $G_m = (Z_m, L^m)$ using the tree-like graph G, as follows: the path $z_0, z_1, \ldots, z_M$ in the graph G is such that $k \le M$: $z_j = \bar{z}_j$ exists for all $j < k$ and $z_k = T(\bar{z}_{k-1}; \bar{x}\,\|x_i) \ne \bar{z}_k$ for a certain $i \in N$, then, in the graph G_m, we have $L_{z_M} \ne \varnothing$ and any path which passes through the vertex z_M has a length $(M + m)$. In the remaining cases, $L_{z_M} \ne \varnothing$. Using the single-stage game Γ and the transition function T, we determine the regularized game $\bar{G}_m$ in the tree-like graph G_m. Hence, in certain cases, the single-stage game Γ is repeated $(M + m)$ times in the game $\bar{G}_m$, that is, if a player i deviates from the cooperative trajectory, the game will be extended by m additional stages and, in the case of the players from $N\backslash\{i\}$, it will be possible to penalize player i.

Theorem 6. Assume that $d_0 < 0$. For

$$m \ge \max_{i \in N}\{[H_i(\bar{x}) - \max_{x_i \in X_i} H_i(\bar{x}\|x_i)]/V_i\}$$

in the game $\bar{G}_m$, a Nash equilibrium situation with the payoffs

$$H_i^m(z_0, z_1, \ldots, z_l) = MH_i(\bar{x}), \quad i \in N.$$

exists.

REFERENCES

1. FUDENBERG, D. and TIROLE, J., *Game Theory*. MIT Press, Cambridge; 1991.
2. OSBORNE, M. J. and RUBINSTEIN, A., *A Course in Game Theory*. MIT Press, Cambridge; 1996.
3. PETROSYAN, L. A., ZENKEVICH, N. A. and SEMINA, Ye. A., *Game Theory*. Vyssh. Shkola, Moscow, 1998.
4. PETROSJAN, L. A. and EGOROVA, A. A., New class of solutions for repeated bimatrix games. *Proc. 11th IFAC Workshop on Control Application of Optimization*. St Petersburg, Russia, 2000, **2**, 617–622.
5. NASH, J. F., Equilibrium points in n-person games. *Proc. Nat. Acad. Sci. U.S.A.*, 1950, **36**, 48–90.
6. KUHN, H. W., Extensive games and the problem of information. *Ann. Math. Studies*, 1953, **28**, 193–216.
7. VAN DAMME, E. E. C., *Stability and Perfection of Nash Equilibria*. Springer, Berlin: New York; 1991.
8. STRAFFIN, P. D., *Game Theory and Strategy*. Math. Associat. America, Washington, 1993.

Translated by E.L.S.