

COOPERATIVE DYNAMICS IN A REPEATED THREE-PERSON GAME WITH FINITE NUMBER OF STRATEGIES

Anatolii F. Kleimenov^{*.1}, Maxim A. Schneider^{**}

** Inst. of Mathematics and Mechanics of RAS, Dynamical
Systems Department, S.Kovalevskaya Str. 16,
Ekaterinburg, 620066 Russia, phone: +7-343-375-3456,
fax: +7-343-374-2581, e-mail: kleimenov@imm.uran.ru*

*** Ural State Economic University, Chair of Applied
Informatics in Economics, 8 Marta Str., 62, Ekaterinburg,
620219 Russia, phone:+7-343-375-3028, fax:
+7-343-375-3062, e-mail: agency@ural.ru*

Abstract: In this game two players act in the class of mixed strategies and the third player acts in the class of pure strategies. The suggested approach for building dynamics uses the principle of non-decrease of players' payoffs, various behavior types for players and some special procedure of using Nash equilibria in auxiliary games. Cooperative variant of this procedure is proposed. As an example, three-person repeated game of Prisoner's dilemma type is considered.

Copyright©2005 IFAC

Keywords: Dynamic models, Game theory, Cooperation, Behavior

1. INTRODUCTION

There are various approach to constructing dynamics in repeated games (see e.g. (Maynard Smith, 1982); (Hofbauer and Sigmund, 1988), (Friedman, 1991), (Tarasyev, 1994)). The suggested approach is based on the theory of positional differential games ((Krasovskii and Subbotin, 1988), (Krasovskii, 1985)) and, more specifically, on nonzero-sum branch of this theory (Kleimenov, 1993). For repeated bimatrix 2x2 games this approach was proposed in (Kleimenov, 1998), (Kleimenov and Kryazhimskii, 1998), (Kleimenov, 2000). Using the principle of non-decrease of players' payoffs, various behavior types for players and some special procedure based on Nash equilibria in auxiliary bimatrix games is characteristic for the approach.

Noncooperative variant of a repeated three-person game with two pure strategies for each player was studied in (Kleimenov, 2004). It was assumed there that two players use mixed strategies and the third player uses pure strategies only.

In present paper cooperative variant of repeated three-person game with finite number pure strategies for each player is considered. As in (Kleimenov, 2004) it is assumed that two players use mixed strategies and the third player uses pure strategies only.

The paper is organized as follows. Section 2 contains notations. Cooperative variant of dynamics is described in Section 3. Description of various behavior types of players is given in Section 4. Trajectories generated by the proposed algorithm are described in Section 5 for a game of Prisoner's dilemma type. Numerical illustrations are given in Section 6.

¹ Supported by the Russian Fund of Fundamental Researches, grant 03-01-00228

2. REPEATED THREE-PERSON GAME

The following static three-person game lies in the basis of the considered repeated game. The player 1 (P1) has l strategies numbered $1, \dots, l$, player 2 (P2) has m strategies numbered $1, \dots, m$, and player 3 (P3) has n strategies numbered $1, \dots, n$. The values f_{ijk} , g_{ijk} and h_{ijk} are payoffs to P1, P2 and P3, respectively, for a strategy triple (i, j, k) , where $i \in L = \{1, \dots, l\}$, $j \in M = \{1, \dots, m\}$, and $k \in N = \{1, \dots, n\}$.

Let the players choose their strategies sequentially in rounds $1, 2, \dots$. Assume, that in each round P1 and P2 act in the class of mixed strategies. A collection $\mathbf{p} = (p_1, p_2, \dots, p_l)$, where $p_i \geq 0$, $\sum_{i=1}^l p_i = 1$ is a mixed strategy of P1, and a collection $\mathbf{q} = (q_1, q_2, \dots, q_m)$, where $q_j \geq 0$, $\sum_{j=1}^m q_j = 1$ is a mixed strategy of P2. So, the $(l-1)$ -dimensional simplex S_{l-1} and $(m-1)$ -dimensional simplex S_{m-1} are the sets of mixed strategies of P1 and P2, respectively. These mixed strategies can be realized physically, if P1 and P2 represent large groups of identical agents or populations. Indeed, p_1, p_2, \dots, p_l can be interpreted as shares of all agents in the first group, who play the first strategy, the second one, ..., the l -th one, respectively. Assume also that P3 uses only pure strategies from the set N .

Let the set $S = S_{l-1} \times S_{m-1} \times N$ be called the state space of the repeated game. Elements of S will be called states. The expected payoffs (briefly, the payoffs) to the P1, P2 and P3 at a state $(\mathbf{p}_t, \mathbf{q}_t, k_t) \in S$ are defined by

$$\begin{aligned} f(\mathbf{p}_t, \mathbf{q}_t, k_t) &= \sum_{i=1}^l \sum_{j=1}^m p_{i,t} q_{j,t} f_{ijk_t}, \\ g(\mathbf{p}_t, \mathbf{q}_t, k_t) &= \sum_{i=1}^l \sum_{j=1}^m p_{i,t} q_{j,t} g_{ijk_t}, \\ h(\mathbf{p}_t, \mathbf{q}_t, k_t) &= \sum_{i=1}^l \sum_{j=1}^m p_{i,t} q_{j,t} h_{ijk_t}, \end{aligned} \quad (1)$$

where $p_{i,t}$, $q_{j,t}$ are the components of the vectors $\mathbf{p}_t, \mathbf{q}_t$.

3. COOPERATIVE DYNAMICS

The suggested approach to building dynamics in repeated three-person games is characterized by the following two features. Firstly, it is based on the ideology of guaranteed approach used in the positional differential game theory ((Krasovskii and Subbotin, 1988), (Krasovskii, 1985)). This ideology is expressed here in the principle of non-decrease of guaranteed results for all players along solution trajectories. Secondly, some special procedure based on Nash equilibria in some auxiliary static games (Kleimenov, 1998) is used.

Assume that in each round t the players observe the current state $(\mathbf{p}_t, \mathbf{q}_t, k_t)$ and choose a state $(\mathbf{p}_{t+1}, \mathbf{q}_{t+1}, k_{t+1})$ for the next round taking into account the following restriction

$$\begin{aligned} &(\mathbf{p}_{t+1}, \mathbf{q}_{t+1}) \in U_{\alpha, \beta}(\mathbf{p}_t, \mathbf{q}_t) = \\ &\{(\mathbf{p}, \mathbf{q}) \in S_{l-1} \times S_{m-1} : |p_{i,t} - p_i| \leq \alpha, \\ &|q_{j,t} - q_j| \leq \beta, i = 1, \dots, l, j = 1, \dots, m\} \end{aligned} \quad (2)$$

Here α and β are positive, sufficiently small numbers. The smallness of α and β means that the inner structure of the groups of interacting agents evolves slowly enough. It should be noted that the strategy k_{t+1} can be chosen from the set N arbitrarily.

In the paper the transition $(\mathbf{p}_t, \mathbf{q}_t, k_t) \Rightarrow (\mathbf{p}_{t+1}, \mathbf{q}_{t+1}, k_{t+1})$ will be realized following the principle of non-decrease of players' payoffs during the game. There are many variants of such a transition. The proposed approach is based on using Nash equilibria in auxiliary bimatrix game (Kleimenov, 1998, Kleimenov and Kryazhinskiy, 1998).

Let $(\mathbf{p}_t, \mathbf{q}_t, k_t)$ be a current state in the repeated game. By fixing $k_{t+1}^* \in N$ the following problems are formulated.

Problem 1. Find a pair $(\mathbf{p}^1, \mathbf{q}^1)$ maximizing the function $f(\mathbf{p}_t, \mathbf{q}_t, k_{t+1}^*)$ (1) over the set $U_{\alpha, \beta}(\mathbf{p}_t, \mathbf{q}_t)$ (2) under the condition

$$g(\mathbf{p}, \mathbf{q}, k_{t+1}^*) \geq g(\mathbf{p}_t, \mathbf{q}_t, k_{t+1}^*) \quad (3)$$

Problem 2. Find a pair $(\mathbf{p}^2, \mathbf{q}^2)$ maximizing the function $g(\mathbf{p}_t, \mathbf{q}_t, k_{t+1}^*)$ (1) over the set $U_{\alpha, \beta}(\mathbf{p}_t, \mathbf{q}_t)$ (2) under the condition

$$f(\mathbf{p}, \mathbf{q}, k_{t+1}^*) \geq f(\mathbf{p}_t, \mathbf{q}_t, k_{t+1}^*) \quad (4)$$

Consider an auxiliary bimatrix game (A^*, B^*) with the matrices

$$\begin{aligned} A^* &= \begin{pmatrix} f(\mathbf{p}^1, \mathbf{q}^1, k_{t+1}^*) & f(\mathbf{p}^1, \mathbf{q}^2, k_{t+1}^*) \\ f(\mathbf{p}^2, \mathbf{q}^1, k_{t+1}^*) & f(\mathbf{p}^2, \mathbf{q}^2, k_{t+1}^*) \end{pmatrix}, \\ B^* &= \begin{pmatrix} g(\mathbf{p}^1, \mathbf{q}^1, k_{t+1}^*) & g(\mathbf{p}^1, \mathbf{q}^2, k_{t+1}^*) \\ g(\mathbf{p}^2, \mathbf{q}^1, k_{t+1}^*) & g(\mathbf{p}^2, \mathbf{q}^2, k_{t+1}^*) \end{pmatrix} \end{aligned} \quad (5)$$

In this bimatrix game, i -th strategy of P1 is "to choose \mathbf{p}^i " and the j -th strategy of P2 is "to choose \mathbf{q}^j " ($i = 1, 2$; $j = 1, 2$). To obtain $(\mathbf{p}_{t+1}, \mathbf{q}_{t+1})$ the players find Nash equilibria. It is easily proved that the game (A^*, B^*) has at least one Nash equilibrium in the class of pure strategies. Two cases are possible.

- (1) The game has an unique Nash equilibrium $(\mathbf{p}^N, \mathbf{q}^N)$. Then $(\mathbf{p}_{t+1}, \mathbf{q}_{t+1}) = (\mathbf{p}^N, \mathbf{q}^N)$.
- (2) The game has two Nash equilibria $(\mathbf{p}^{N1}, \mathbf{q}^{N1})$ and $(\mathbf{p}^{N2}, \mathbf{q}^{N2})$. Then $(\mathbf{p}_{t+1}, \mathbf{q}_{t+1}) = \frac{1}{2}(\mathbf{p}^{N1} + \mathbf{p}^{N2}, \mathbf{q}^{N1} + \mathbf{q}^{N2})$.

Thus, the pair $(\mathbf{p}_{t+1}, \mathbf{q}_{t+1})$ is determined; it depends on k_{t+1}^* .

After that the strategy k_{t+1} of P3 is chosen from the condition

$$h(\mathbf{p}_{t+1}, \mathbf{q}_{t+1}, k_{t+1}) - h(\mathbf{p}_t, \mathbf{q}_t, k_t) \geq 0 \quad (6)$$

If such strategy k_{t+1} is nonunique, then one chooses strategy giving the maximum to the left-hand side of (6).

So, the dynamics of the considered repeated game is completely determined.

Besides local criteria of players, given by (1), quite often one gives also global criteria evaluating the quality of the process wholly. And dynamics constructed above not always leads to a state optimizing global criteria. It appears that the using so-called behavior types for players can be effective in this case.

4. BEHAVIOR TYPES

Following the paper (Kleimenov, 1998) various behavior types for players will be introduced in this paper.

Until now it was assumed that each player is *normal* (*nr*) in the sense that his (or her) behavior is aimed at maximizing his (or her) own payoff. However, there might be other behavior types such as *altruistic* (*al*) ("the better to my rival, the better to me"), *aggressive* (*ag*) ("the worse to my rival, the better to me"), and *paradoxical* (*pr*) (the worse to me, the better to me"). These three behavior types can be formalized in the following way.

Let us say that some player is:

- i) altruistic with respect to other player whenever he (or she) identifies his (or her) payoff with the payoff of other player;
- ii) aggressive with respect to other player whenever he (or she) identifies his (or her) payoff with the payoff of other player taken with opposite sign;
- iii) paradoxical whenever he (or she) identifies his (or her) payoff with the same payoff taken with opposite sign.

These definitions indicate extremes in players' behaviors. Real individuals behave partially normal, partially altruistic, partially aggressive. and partially paradoxical. In other words, mixtures of behavior types would better agree with real dynamics.

If we restrict each player to "pure" behavior types then, for example, in two-person game with pay-

off functions f and g there exist 16 admissible combinations shown in Table 1. In 4 combinations players' interests coincide, and they solve problems of optimal choice. In other 4 combinations the players have opposite interests and, thus, play zero-sum matrix games. The rest 8 pairs determine nonzero-sum bimatrix games.

Table 1

	<i>nr</i>	<i>al</i>	<i>ag</i>	<i>pr</i>
<i>nr</i>	(f,g)	(f,f)	(f,-f)	(f,-g)
<i>al</i>	(g,g)	(g,f)	(g,-f)	(g,-g)
<i>ag</i>	(-g,g)	(-g,f)	(-g,-f)	(-g,-g)
<i>pr</i>	(-f,g)	(-f,f)	(-f,-f)	(-f,-g)

In three-person game the number of admissible combinations of "pure" behavior types for players is much more.

5. THREE-PERSON GAME OF PRISONER'S DILEMMA TYPE

Now consider a game G with three players P1, P2 and P3 each of which has two strategies: the first one is denoted by C (cooperate) and the second one is denoted by D (defect).

Payoffs for P1, P2 and P3 in the game G are given as follows (see e.g. (Straffin, 1993)):

$$\begin{aligned} A^C &= \begin{pmatrix} a_4 & a_3 \\ a_6 & a_5 \end{pmatrix}, B^C = \begin{pmatrix} a_4 & a_6 \\ a_3 & a_5 \end{pmatrix}, \\ H^C &= \begin{pmatrix} a_4 & a_3 \\ a_3 & a_1 \end{pmatrix}, \end{aligned} \quad (7)$$

if P3 chooses strategy C , and

$$\begin{aligned} A^D &= \begin{pmatrix} a_3 & a_1 \\ a_5 & a_2 \end{pmatrix}, B^D = \begin{pmatrix} a_3 & a_5 \\ a_1 & a_2 \end{pmatrix}, \\ H^D &= \begin{pmatrix} a_6 & a_5 \\ a_5 & a_2 \end{pmatrix}, \end{aligned} \quad (8)$$

if P3 chooses strategy D .

Here A , B , H denote the payoff matrices of P1, P2 and P3, correspondingly. The following inequalities are assumed

$$a_6 > a_5 > a_4 > a_3 > a_2 > a_1 \quad (9)$$

Remind that P1 and P2 act in the class of mixed strategies and P3 acts in the class of pure strategies. The pair $(p, 1-p)$, $0 \leq p \leq 1$ is a mixed strategy of P1, and the pair $(q, 1-q)$, $0 \leq q \leq 1$ is a mixed strategy of P2. The parameter r belonging to $\{C, D\}$ is a strategy of P3.

Let this game be repeated. Assuming that P1, P2 and P3 construct the dynamics in the same manner as in Section 3, we conclude that trajectories generated by this dynamics are placed either in the unit square

$$E^D = \{(p, q, r) : 0 \leq p \leq 1, 0 \leq q \leq 1, r = D\}$$

or in the unit square

$$E^C = \{(p, q, r) : 0 \leq p \leq 1, 0 \leq q \leq 1, r = C\},$$

Suppose that P3 can make momentary switchings from strategy C to strategy D and back.

Thus, P1 and P2 construct dynamics in repeated bimatrix games (A^C, B^C) and (A^D, B^D) according to the algorithm from Section 3 and taking into account the condition that the payoff of P3 can not decrease along trajectories. In turn, P3 realizes switchings from the square E^C to the square E^D and back as required.

This algorithm of constructing dynamics of the repeated game was analyzed in completely analytical form. Typical trajectories generated by the algorithm for values $r = D$ and $r = C$ are represented in Fig.1 and Fig.2, correspondingly, for the case under conditions

$$a_2 + a_3 - a_1 - a_5 > 0, 2a_2 > a_1 + a_5. \quad (10)$$

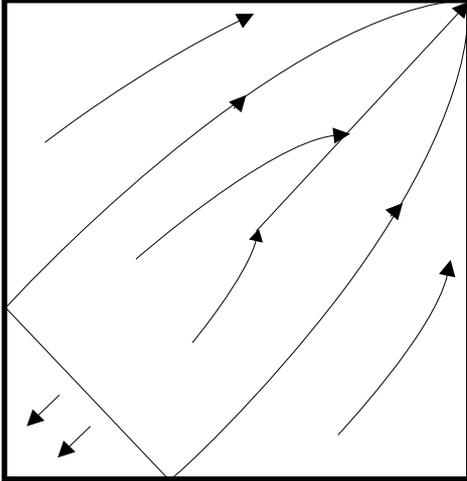


Fig.1

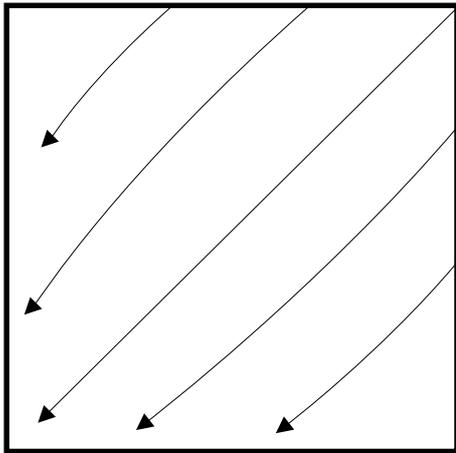


Fig.2

Arrows in Fig. 1,2 show directions of motion for a current state (p, q) .

Note that all the players increase their payoffs along the trajectories shown in Fig.1. At the same

time along the trajectories shown in Fig.2 P1 and P2 increase their payoffs while P3 decrease his payoff, i.e. actually all the trajectories should stay in initial points.

Now assume in addition that during the process of forming dynamics in repeated games (A^C, B^C) and (A^D, B^D) , P1 and P2 can use as regards each to other various behavior types described in Section 4. Finally, P3 can use paradoxical or normal behavior types.

Let the condition of leading a state to the state (C, C, C) be taken as a global criterion. Analyzing Fig.1 and Fig.2 one can conclude that this global condition can fail to be fulfilled if all the players use only normal behavior type, meanwhile the using of abnormal behavior types makes possible to solve the problem (see (Kleimenov and Volegova, 1999), where a repeated bimatrix 2×2 game is considered).

Various optimization problems arise, for example, a problem of minimizing time of using abnormal behavior types. The proposed approach of constructing dynamics in repeated three-person game gives a tool for statement and solving similar problems

6. NUMERICAL ILLUSTRATIONS

Let the parameters a_1, \dots, a_6 be as follows

$$a_1 = -2, a_2 = 2, a_3 = 3, a_4 = 4, a_5 = 5, a_6 = 6.$$

Then the conditions (10) are fulfilled. Trajectories on the square E^D look like in Fig.1, and trajectories on the square E^C stay in initial points.

The collection (D, D, D) is NE with players' payoffs $f = g = h = 2$. However, the collection (C, C, C) gives to each player the payoff equal to 4.

Let the global criterion be to lead states to the state (C, C, C) . Denote

$$\begin{aligned} G_1 &= \left\{ (p, q) \in E^D : p + q \leq \frac{1}{3} \right\}, \\ G_2 &= \left\{ (p, q) \in E^D : 2pq - 4p + 3q \geq 1 \right\}, \\ G_3 &= \left\{ (p, q) \in E^D : 2pq + 3p - 4q \geq 1 \right\}, \\ G_4 &= E^D \setminus (G_1 \cup G_2 \cup G_3). \end{aligned}$$

These sets are represented in Fig.3. Analyzing generated trajectories one can obtain that:

- i) For initial states $(p^0, q^0) \in G_4$, $r = D$ the proposed dynamics leads to the state (C, C, D) after that P3 switches this state to the state (C, C, C) .
- ii) For initial states $(p^0, q^0) \in G_1 \cup G_2 \cup G_3$, $r = C$ or $r = D$ it is impossible to reach the state (C, C, D) without using abnormal behavior types.

- iii) For initial states $(p^0, q^0) \in G_2 \cup G_3$, $r = D$ one can pass to the state (C, C, D) by using normal behavior type for P1 (or P2), and aggressive type for P2 (or P1), and paradoxical type for P3 up to the reaching of the boundary of G_4 , after that all the players use normal behavior type.
- iv) For initial states $(p^0, q^0) \in G_1$, $r = D$ one can pass to the state (C, C, D) by using aggressive behavior type for P1 and P2 and normal type for P3 up to the reaching the set G_4 , after that all the players use normal behavior type.

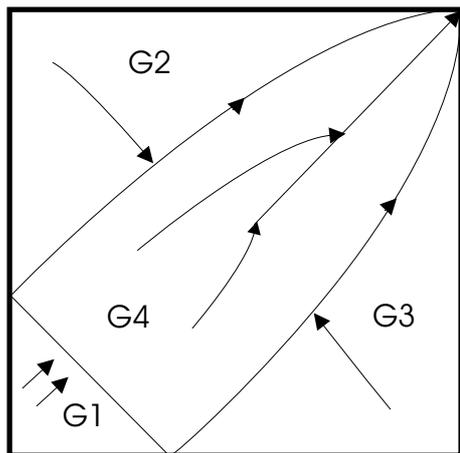


Fig.3

7. CONCLUSION

The proposed approach of constructing dynamics in the repeated three-person game gives a tool for statement and solving various optimization problems, for example, the problem of minimizing time of using abnormal behavior types.

REFERENCES

- Friedman, D. (1991). Evolutionary games in economics,. *Econometrica* **59(3)**, 637–666.
- Hofbauer, J. and K. Sigmund (1988). *The Theory of Evolution and Dynamic Systems*. Cambridge Univ. Press. Cambridge.
- Kleimenov, A.F. (1993). *Nonantagonistic Positional Differential Games*. Nauka. Ekaterinburg. (in Russian).
- Kleimenov, A.F. (1998). An approach to building dynamics for repeated bimatrix 2x2 games involving various behavior types. In: *Dynamic and Control*,. pp. 195–204. Gordon & Breach Sci. Publ.. London.
- Kleimenov, A.F. (2000). Problem of constructing dynamics for nonantagonistic positional differential games. In: *Proceedings of the Steklov Institute of Mathematics*. Vol. Suppl. 2. pp. S67–S78.

- Kleimenov, A.F. (2004). Construction of dynamics in repeated three-person game with two strategies for players. In: *Proceedings of the IFAC Workshop GSCP-04*. pp. 132–137.
- Kleimenov, A.F. and A.V. Kryazhimskii (1998). Normal behaviour, altruism and aggression in cooperative game dynamics. Interim Report IR 98-076. IIASA, Laxenburg.
- Kleimenov, A.F. and E.I. Volegova (1999). Problems of control by dynamics for repeated bimatrix 2x2 games with switching of local criteria for the players. In: *Nonsmooth and Discontinuous Problems of Control and Optimization* (Batukhtin V.D. and Kirillova F.M., Eds.). Pergamon. New York. pp. 127–132.
- Krasovskii, N.N. (1985). *Control of a Dynamical System*. Nauka. Moscow (in Russian).
- Krasovskii, N.N. and A.I. Subbotin (1988). *Game-Theoretical Control Problems*. Springer, NY. Berlin.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge Univ. Press.. Cambridge.
- Straffin, P. (1993). *Game theory and strategy*. Math. Associat. of America .
- Tarashev, A.M. (1994). A differential model for a 2x2 evolutionary game dynamics. Working Paper WP-94-63. IIASA, Laxenburg.