Nash Equilibria Conditions for Stochastic Positional Games

Dmitrii Lozovanu[1] **and Stefan Pickl**[2]

[1] *Institute of Mathematics and Computer Science,*
Academy of Sciences of Moldova,
Academy str., 5, Chisinau, MD–2028, Moldova
E-mail: lozovanu@math.md
http://www.math.md/structure/applied-mathematics/math-modeling-optimization/
[2] *Institute for Theoretical Computer Science, Mathematics and*
Operations Research, Universität der Bundeswehr, München
85577 Neubiberg-München, Germany
E-mail: stefan.pickl@unibw.de

Abstract We formulate and study a class of stochastic positional games using a game-theoretical concept to finite state space Markov decision processes with an average and expected total discounted costs optimization criteria. Nash equilibria conditions for the considered class of games are proven and some approaches for determining the optimal strategies of the players are analyzed. The obtained results extend Nash equilibria conditions for deterministic positional games and can be used for studying Shapley stochastic games with average payoffs.

Keywords: Markov decision processes, stochastic positional games, Nash equilibria, Shapley stochastic games, optimal stationary strategies.

1. Introduction

In this paper we consider a class of stochastic positional games that extends deterministic positional games studied by Moulin,1976, Ehrenfeucht and Mycielski, 1979, Gurvich at al., 1988, Condon, 1992, Lozovanu and Pick, 2006, 2009. The considered class of games we formulate and study applying the concept of positional games to finite state space Markov decision processes with average and expected total discounted costs optimization criteria. We assume that the Markov process is controlled by several actors (players) as follows: The set of states of the system is divided into several disjoint subsets which represent the corresponding position sets of the players. Additionally the cost of system's transition from one state to another is given for each player separately. Each player has to determine which action should be taken in each state of his position set of the Markov process in order to minimize his own average cost per transition or the expected total discounted cost. In these games we are seeking for a Nash equilibrium.

The main results of the paper are concerned with the existence of Nash equilibria for the considered class of games and determining the optimal strategies of the players. Necessary and sufficient conditions for the existence of Nash equilibria in stochastic positional games that extend Nash equilibria conditions for deterministic positional games are proven. Based on the constructive proof of these results we propose some approaches for determining the optimal strategies of the players. Additionally we show that the stochastic positional games are tightly connected with Shapley stochastic games (Shapley, 1953) and the obtained results can be used for studying a special class of Shapley stochastic games with average payoffs.

2. Formulation of the Basic Game Models and Some Preliminary Results

We consider two game-theoretic models. We formulate the first game model for Markov decision processes with average cost optimization criterion and call it *the stochastic positional game with average payoffs*. We formulate the second one for Markov decision processes with discounted cost optimization criterion and call it *stochastic positional game with discounted payoffs*. Then we show the relationship of these games with Shapley stochastic games.

2.1. Stochastic Positional Games with Average Payoffs

To formulate the stochastic positional game with average payoffs we shall use the framework of a Markov decision process (X, A, p, c) with a finite set of states X, a finite set of actions A, a transition probability function $p : X \times X \times A \to [0, 1]$ that satisfies the condition

$$\sum_{y \in X} p^a_{x,y} = 1, \quad \forall x \in X, \quad \forall a \in A$$

and a transition cost function $c : X \times X \to R$ which gives the costs $c_{x,y}$ of states transitions of the dynamical system from an arbitrary $x \in X$ to another state $y \in X$ (see Howard, 1960; Puterman, 2005). For the noncooperative game model with m players we assume that m transition cost functions

$$c^i : X \times X \to R, \quad i = 1, 2, \ldots, m$$

are given, where $c^i_{x,y}$ expresses the cost of the system's transition from the state $x \in X$ to the state $y \in X$ for the player $i \in \{1, 2, \ldots, m\}$. In addition we assume that the set of states X is divided into m disjoint subsets $X_1, X_2, \ldots, X_m$

$$X = X_1 \cup X_2 \cup \cdots \cup X_m \quad (X_i \cap X_j = \emptyset, \quad \forall i \neq j),$$

where X_i represents the positions set of the player $i \in \{1, 2, \ldots, m\}$. So, the Markov process is controlled by m players, where each player $i \in \{1, 2, \ldots, m\}$ fixes actions in his positions $x \in X_i$. We assume that each player fixes actions in the states from his positions set using stationary strategies, i.e. we define the stationary strategies of the players as m maps:

$$s^i : x \to a \in A^i(x) \;\; \text{for} \;\; x \in X_i, \; i = 1, 2, \ldots, m,$$

where $A^i(x)$ is the set of actions of the player i in the state $x \in X_i$. Without loss of generality we may consider $|A^i(x)| = |A^i| = |A|$, $\forall x \in X_i$, $i = 1, 2, \ldots, m$. In order to simplify the notation we denote the set of possible actions in a state $x \in X$ for an arbitrary player by $A(x)$. A stationary strategy s^i, $i \in \{1, 2, \ldots, m\}$ in the state $x \in X_i$ means that at every discrete moment of time $t = 0, 1, 2, \ldots$ the player i uses the action $a = s^i(x)$. Players fix their strategy independently and do not inform each other which strategies they use in the decision process.

If the players $1, 2, \ldots, m$ fix their stationary strategies $s^1, s^2, \ldots, s^m$, respectively, then we obtain a situation $s = (s^1, s^2, \ldots, s^m)$. This situation corresponds to a simple Markov process determined by the probability distributions $p^{s^i(x)}_{x,y}$ in the states $x \in X_i$ for $i = 1, 2, \ldots, m$. We denote by $P^s = (p^s_{x,y})$ the matrix of probability transitions of this Markov process. If the starting state x_0 is given, then for the Markov process with the matrix of probability transitions P^s we can determine the average cost per transition $\omega^i_{x_0}(s^1, s^2, \dots, s^m)$ with respect to each player $i \in \{1, 2, \dots, m\}$ taking into account the corresponding matrix of transition costs $C^i = (c^i_{x,y})$. So, on the set of situations we can define the payoff functions of the players as follows:

$$F^i_{x_0}(s^1, s^2, \dots, s^m) = \omega^i_{x_0}(s^1, s^2, \dots, s^m), \quad i = 1, 2, \dots, m.$$

In such a way we obtain a discrete noncooperative game in normal form which is determined by a finite set of strategies $S^1, S^2, \dots, S^m$ of m players and the payoff functions defined above. In this game we are seeking for a *Nash equilibrium* (see Nash, 1951), i.e., we consider the problem of determining the stationary strategies

$$s^{1^*}, s^{2^*}, \dots, s^{i-1^*}, s^{i^*}, s^{i+1^*}, \dots, s^{m*}$$

such that

$$F^i_{x_0}(s^{1^*}, s^{2^*}, \dots, s^{i-1^*}, s^{i^*}, s^{i+1^*}, \dots, s^{m*}) \\ \le F^i_{x_0}(s^{1^*}, s^{2^*}, \dots, s^{i-1^*}, s^{i}, s^{i+1^*}, \dots, s^{m*}), \ \ \forall s^i \in S^i, \ \ i = 1, 2, \dots, m.$$

The game defined above is determined uniquely by the set of states X, the position sets $X_1, X_2, \dots, X_m$, the set of actions A, the cost functions $c^i : X \times X \to R$, $i = 1, 2, \dots, m$, the probability function $p : X \times X \times A \to [0, 1]$ and the starting position x_0. Therefore, we denote this game by $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ x_0)$. In the case $m = 2$ and $c^2 = -c^1$ we obtain an antagonistic stochastic positional game. If $p^a_{x,y} = 0 \vee 1$, $\forall x, y \in X$, $\forall a \in A$ the stochastic positional game $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ x_0)$ is transformed into the cyclic game (Ehrenfeucht and Mycielski, 1979, Gurvich at al., 1988, Condon, 1992, Lozovanu and Pick, 2006). Some results concerned with the existence of Nash equilibria for stochastic positional games with average payoffs have been derived by Lozovanu at al., 2011. In particular the following theorem has been proven.

Theorem 1. *If for an arbitrary situation $s = (s^1, s^2, \dots, s^m)$ of the stochastic positional game with average payoffs the matrix of probability transitions $P^s = (p^s_{x,y})$ induces an ergodic Markov chain then for the game there exists a Nash equilibrium.*

If the matrix P^s for some situations do not correspond to an ergodic Markov chain then for the stochastic positional game with average payoffs a Nash equilibrium may not exist. This follow from the constructive proof of this theorem (see Lozovanu at al., 2011). An example of a deterministic positional game with average payoffs for which Nash equilibrium does not exist has been constructed by Gurvich at al., 1988. However, in the case of antagonistic stochastic positional games saddle points always exist (Lozovanu and Pickl, 2014), i.e. in this case the following theorem holds.

Theorem 2. *For an arbitrary antagonistic positional game there exists a saddle point.*

The existence of saddle points for deterministic positional games with average payoffs have been proven by Ehrenfeucht and Mycielski, 1979, Gurvich at al., 1988.

2.2. Stochastic Positional Games with Discounted Payoffs

We formulate the stochastic positional game with discounted payoffs in a similar way as the game from Section 2.. We assume that for the Markov process m transition cost functions $c^i : X \times X \to R$, $i = 1, 2, \ldots, m$, are given and the set of states X is divided into m disjoint subsets $X_1, X_2, \ldots, X_m$, where X_i represents the positions set of the player $i \in \{1, 2, \ldots, m\}$. The Markov process is controlled by m players, where each player $i \in \{1, 2, \ldots, m\}$ fixes actions in his positions $x \in X_i$ using stationary strategies, i.e. the stationary strategies of the players in this game are defined as m maps:

$$s^i : x \to a \in A(x) \quad \text{for} \quad x \in X_i; \quad i = 1, 2, \ldots, m.$$

Let $s^1, s^2, \ldots, s^m$ be a set of stationary strategies of the players that determine the situation $s = (s^1, s^2, \ldots, s^m)$. Consider the matrix of probability transitions $P^s = (p^s_{x,y})$ which is induced by the situation s, i.e., each row of this matrix corresponds to a probability distribution $p^{s^i(x)}_{x,y}$ in the state x where $x \in X_i$. If the starting state x_0 is given, then for the Markov process with the matrix of probability transitions P^s we can determine the discounted expected total cost $\sigma^i_{x_0}(s^1, s^2, \ldots, s^m)$ with respect to each player $i \in \{1, 2, \ldots, m\}$ taking into account the corresponding matrix of transition costs $C^i = (c^i_{x,y})$. So, on the set of situations we can define the payoff functions of the players as follows:

$$\widehat{F}^i_{x_0}(s^1, s^2, \ldots, s^m) = \sigma^i_{x_0}(s^1, s^2, \ldots, s^m), \quad i = 1, 2, \ldots, m.$$

In such a way we obtain a new discrete noncooperative game in normal form which is determined by the sets of strategies $S^1, S^1, \ldots, S^m$ of m players and the payoff functions defined above. In this game we are seeking for a Nash equilibrium. We denote the stochastic positional game with discounted payoffs by $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ \gamma,\ x_0)$.

For this game the following result has been proven (Lozovanu, 2011).

Theorem 3. *For an arbitrary stochastic positional game* $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ \gamma,\ x_0)$ *with given discount factor* $0 < \gamma < 1$ *there exists a Nash equilibrium.*

Based on a constructive proof of Theorems 1,3 some iterative procedures for determining Nash equilibria in the considered positional games have been proposed (see Lozovanu at al., 2011).

2.3. The Relationship of Stochastic Positional Games with Shapley Stochastic Games

A stochastic game in the sense of Shapley (see Shapley, 1953) is a dynamic game with probabilistic transitions played by several players in a sequence of stages, where the beginning of each stage corresponds to a state of the dynamical system. The game starts at a given state from the set of states of the system. At each stage players select actions from their feasible sets of actions and each player receives a stage payoff that depends on the current state and the chosen actions. The game then moves to a new random state the distribution of which depends on the previous state and the actions chosen by the players. The procedure is repeated at a new state and the play continues for a finite or infinite number of stages. The total payoff of a player is either the limit inferior of the average of the stage payoffs or the discounted sum of the stage payoffs.

So, an average Shapley stochastic game with m players consists of the following elements:

1. A state space X (which we assume to be finite);
2. A finite set $A^i(x)$ of actions with respect to each player $i \in \{1,2,\dots,m\}$ for an arbitrary state $x \in X$;
3. A stage payoff $f^i(x,a)$ with respect to each player $i \in \{1,2,\dots,m\}$ for each state $x \in X$ and for an arbitrary action vector $a \in \prod_i A^i(x)$;
4. A transition probability function $p : X \times \prod_{x\in X}\prod_i A^i(x) \times X \to [0,1]$ that gives the probability transitions $p^a_{x,y}$ from an arbitrary $x \in X$ to an arbitrary $y \in Y$ for a fixed action vector $a \in \prod_i A^i(x)$, where $\sum_{y\in X} p^a_{x,y} = 1,\ \ \forall x \in X,\ a \in \prod_i A^i(x)$;
5. A starting state $x_0 \in X$.

The stochastic game starts in state x_0. At stage t players observe state x_t and simultaneously choose actions $a^i_t \in A^i(x_t),\ \ i = 1,2,\dots,m$. Then nature selects a state x_{t+1} according to probability transitions $p^{a_t}_{x_t,y}$ for fixed action vector $a_t = (a^1_t, a^2_t, \dots, a^m_t)$. A play of the stochastic game $x_0, a_0, x_1, a_1, \dots, x_t, a_t, \dots$ defines a stream of payoffs $f^i_0, f^i_1, f^i_2, \dots,$ where $f^i_t = f^i(x_t, a_t),\ t = 0,1,2,\dots$. The t-stage average stochastic game is the game where the payoff of player $i \in \{1,2,\dots,m\}$ is

$$F^i_t = \frac{1}{t}\sum_{\tau=1}^{t-1} f^i_\tau .$$

The infinite average stochastic game is the game where the payoff of player $i \in \{1,2,\dots,m\}$ is

$$\overline{F}^i = \lim_{t\to\infty} F^i_t .$$

In a similar a Shapley stochastic game with expected discounted payoffs of the players is defined. In such a game along to the elements described above also a discount factor λ $(0 < \lambda < 1)$ is given and the total payoff of a player represents the expected discounted sum of the stage payoffs.

By comparison for Shapley stochastic games with stochastic positional games we can observe the following. The probability transitions from a state to another state as well as the stage payoffs of the players in a Shapley stochastic game depend on the actions chosen by all players, while the probability transitions from a state to another state as well as the stage payoffs (the immediate costs of the players) in a stochastic positional game depend only on the action of the player that controls the state in his position set. This means that a stochastic positional game can be regarded as a special case of the Shapley stochastic game. Nevertheless we can see that stochastic positional games can be used for studying some classes of Shapley stochastic games.

The main results concerned with determining Nash equilibria in Shapley stochastic games have been obtained by Gillette, 1957, Mertens and Neyman, 1981, Filar and Vrieze, 1997, Lal and Sinha, 1992, Neyman and Sorin, 2003. Existence of Nash equilibria for such games are proven in the case of stochastic games with a finite set of stages and in the case of the games with infinite stages if the total payoff of each player is the discounted sum of stage payoffs. If the total payoff of a player represents the limit inferior of the average of the stage payoffs then the existence of a Nash equilibrium in Shapley stochastic games is an open question. Based on the results mentioned in previous sections we can show that in the case of the average non-antagonistic stochastic games a Nash equilibrium may not exist. In order to prove this we can use the average stochastic positional game $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ x_0)$ from section 2. It is easy to observe that this game can be regarded as a Shapley stochastic game with average payoff functions of the players, where for a fixed situation $s = (s^1, s^2, \dots, s^m)$ the probability transition $p^s_{x,y}$ from a state $x = x(t) \in X_i$ to a state $y = x(t+1) \in X$ depends only on a strategy s^i of player i and the corresponding stage payoff in the state x of player $i \in \{1, 2, \dots, m\}$ is equal to $\sum_{y\in X} p^s_{x,y} c^i_{x,y}$. Taking into account that the cyclic game represents a particular case of the average stochastic positional game and for the cyclic game Nash equilibrium may not exist (see Gurvich at al., 1988) we obtain that for the average non-antagonistic Shapley stochastic game a Nash equilibrium may not exist. However in the case of average payoffs Theorem 1 can be extended for Shapley stochastic games.

3. Nash Equilibria Conditions for Stochastic Positional Games with Average Payoffs

In this section we formulate Nash equilibria conditions for stochastic positional games in terms of bias equations for Markov decision processes. We can see that Nash equilibria conditions in such terms may be more useful for determining the optimal strategies of the players.

Theorem 4. *Let* $(X, A, \{X_i\}_{i=\overline{1,m}}, \{c^i\}_{i=\overline{1,m}}, p, \overline{x})$ *be a stochastic positional game with a given starting position* $\overline{x} \in X$ *and average payoff functions*

$$F^1_{\overline{x}}(s^1,\, s^2,\, \dots,\, s^m),\ F^2_{\overline{x}}(s^1,\, s^2,\, \dots,\, s^m),\ \dots,\ F^m_{\overline{x}}(s^1,\, s^2,\, \dots,\, s^m)$$

of the players $1, 2, \dots, m$*, respectively. Assume that for an arbitrary situation* $s = (s^1, s^2, \dots, s^m)$ *of the game the transition probability matrix* $P^s = (p^s_{x,y})$ *corresponds to an ergodic Markov chain. Then there exist the functions*

$$\varepsilon^i : X \to R, \quad i = 1, 2, \dots, m$$

and the values $\omega^1, \omega^2, \dots, \omega^m$ *that satisfy the following conditions:*

1) $\mu^i_{x,a} + \sum\limits_{y\in X} p^a_{x,y}\varepsilon^i_y - \varepsilon^i_x - \omega^i \geq 0, \quad \forall x \in X_i, \quad \forall a \in A(x),\ i = 1, 2, \dots, m,$
   where $\mu^i_{x,a} = \sum\limits_{y\in X} p^a_{x,y} c^i_{x,y}$;

2) $\min\limits_{a\in A(x)} \{\mu^i_{x,a} + \sum\limits_{y\in X} p^a_{x,y}\varepsilon^i_y - \varepsilon^i_x - \omega^i\} = 0, \quad \forall x \in X_i,\ i = 1, 2, \dots, m;$

3) *on each position set* X_i, $i \in \{1, 2, \dots, m\}$ *there exists a map* $s^{i^*} : X_i \to A$ *such that*
$$s^{i^*}(x) = a^* \in Arg \min_{a\in A(x)} \left\{\mu^i_{x,a} + \sum_{y\in X} p^a_{x,y}\varepsilon^i_y - \varepsilon^i_x - \omega^i\right\}$$

and

$$\mu^j_{x,a^*} + \sum_{y\in X} p^{a^*}_{x,y}\varepsilon^j_y - \varepsilon^j_x - \omega^j = 0, \quad \forall x \in X_i, \quad j = 1, 2, \ldots, m.$$

The set of maps ${s^1}^*, {s^2}^*, \ldots, s^{m*}$ *determines a Nash equilibrium situation* $s^* = ({s^1}^*, {s^2}^*, \ldots, s^{m*})$ *for the stochastic positional game* $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ \overline{x})$ *and*

$$F^i_{\overline{x}}({s^1}^*, {s^2}^*, \ldots, s^{m*}) = \omega^i, \quad \forall \overline{x} \in X,\ i = 1, 2, \ldots, m.$$

Moreover, the situation $s^* = ({s^1}^*, {s^2}^*, \ldots, s^{m*})$ *is a Nash equilibrium for an arbitrary starting position* $\overline{x} \in X$.

Proof. Let a stochastic positional game with average payoffs be given and assume that for an arbitrary situation s of the game the transition probability matrix $P^s = (p^s_{x,y})$ corresponds to an ergodic Markov chain. Then according to Theorem 1 for this game there exists a Nash equilibrium $s^* = ({s^1}^*, {s^2}^*, \ldots, s^{m*})$ and we can set

$$\omega^i = F^i_{\overline{x}}({s^1}^*, {s^2}^*, \ldots, s^{m*}), \quad \forall \overline{x} \in X,\ i = 1, 2, \ldots, m.$$

Let us fix the strategies ${s^1}^*, {s^2}^*, \ldots, {s^{i-1}}^*, {s^{i+1}}^*, \ldots, s^{m*}$ of the players $1, 2, \ldots, i-1, i+1, \ldots, m$ and consider the problem of determining the minimal average cost per transition with respect to player i. Obviously, if we solve this decision problem then we obtain the strategy ${s^i}^*$. We can determine the optimal strategy of this decision problem with an average cost optimization criterion using the bias equations with respect to player i. This means that there exist the functions $\epsilon^i : X \to R$ and the values ω^i, $i = 1, 2, \ldots, m$ that satisfy the conditions:

1) $\mu^i_{x,a} + \sum\limits_{y\in X} p^a_{x,y}\varepsilon^i_y - \varepsilon^i_x - \omega^i \geq 0, \quad \forall x \in X_i,\ \forall a \in A(x);$

2) $\min\limits_{a\in A(x)} \left\{\mu^i_{x,a} + \sum\limits_{y\in X} p^a_{x,y}\varepsilon^i_y - \varepsilon^i_x - \omega^i\right\} = 0, \quad \forall x \in X_i.$

Moreover, for fixed strategies ${s^1}^*, {s^2}^*, \ldots, {s^{i-1}}^*, {s^{i+1}}^*, \ldots, s^{m*}$ of the corresponding players $1, 2, \ldots, i-1, i+1, \ldots, m$ we can select the strategy ${s^i}^*$ of player i where

$${s^i}^*(x) \in Arg \min_{a\in A(x)} \left\{\mu^i_{x,a} + \sum_{y\in X} p^a_{x,y}\varepsilon^i_y - \varepsilon^i_x - \omega^i\right\}$$

and $\omega^i = F^i_x({s^1}^*, {s^2}^*, \ldots, s^{m*})$, $\forall x \in X,\ i = 1, 2, \ldots, m$. This means that conditions 1)–3) of the theorem hold.

Corollary 1. *If for a stochastic positional game* $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ x)$ *with average payoffs there exist a Nash equilibrium* $s^* = ({s^1}^*, {s^2}^*, \ldots, s^{m*})$ *which is a Nash equilibrium for an arbitrary starting position of the game* $x \in X$ *and for arbitrary two different starting positions* $x, y \in X$ *holds* $F^i_x({s^1}^*, {s^2}^*, \ldots, s^{m*}) = F^i_y({s^1}^*, {s^2}^*, \ldots, s^{m*})$ *then there exists the functions*

$$\varepsilon^i : X \to R, \quad i = 1, 2, \ldots, m$$

and the values $\omega^1, \omega^2, \dots, \omega^m$ *that satisfy the conditions* $1) - 3)$ *from Theorem 4. So,* $\omega^i = F_x^i(s^{1^*}, s^{2^*}, \dots, s^{m*}),\ \forall x \in X,\ i = 1, 2, \dots, m$ *and an arbitrary Nash equilibrium can be found by fixing*

$$s^{i^*}(x) = a^* \in Arg \min_{a \in A(x)} \Big\{ \mu^i_{x,a} + \sum_{y \in X} p^a_{x,y} \varepsilon^i_y - \varepsilon^i_x - \omega^i \Big\}.$$

Using the elementary properties of non ergodic Markov decision processes with average cost optimization criterion the following lemma can be gained.

Lemma 1. *Let* $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ x)$ *be an average stochastic positional game for which there exists a Nash equilibrium* $s^* = (s^{1^*}, s^{2^*}, \dots, s^{m*})$, *which is a Nash equilibrium for an arbitrary starting position of the game with* $\omega^i_x = F^i_x(s^{1^*}, s^{2^*}, \dots, s^{m*})$. *Then* $s^* = (s^{1^*}, s^{2^*}, \dots, s^{m*})$ *is a Nash equilibrium for the average stochastic positional game* $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{\overline{c}^i\}_{i=\overline{1,m}},\ p,\ x)$, *where*

$$\overline{c}^i_{x,y} = c^i_{x,y} - \omega^i_x,\ \forall x, y \in X,\quad i = 1, 2, \dots, m$$

and

$$\overline{F}^i_x(s^{1^*}, s^{2^*}, \dots, s^{m*}) = 0,\ \forall x \in X,\quad i = 1, 2, \dots, m.$$

Now using Corollary 1 and Lemma 1 we can prove the following results.

Theorem 5. *Let* $(X, A, \{X_i\}_{i=\overline{1,m}}, \{c^i\}_{i=\overline{1,m}}, p, x)$ *be an average stochastic positional game. Then in this game there exists a Nash equilibrium for an arbitrary starting position* $x \in X$ *if and only if there exist the functions*

$$\varepsilon^i : X \to R,\ i = 1, 2, \dots, m$$

and the values $\omega^1_x, \omega^2_x, \dots, \omega^m_x$ *for* $x \in X$ *that satisfy the following conditions:*

1) $\mu^i_{x,a} + \sum_{y \in X} p^a_{x,y}\varepsilon^i_y - \varepsilon^i_x - \omega^i_x \geq 0, \quad \forall x \in X_i,\ \forall a \in A(x),\ i = 1, 2, \dots, m,$ *where* $\mu^i_{x,a} = \sum_{y \in X} p^a_{x,y} c^i_{x,y}$;

2) $\min_{a \in A(x)} \{\mu^i_{x,a} + \sum_{y \in X} p^a_{x,y}\varepsilon^i_y - \varepsilon^i_x - \omega^i_x\} = 0, \quad \forall x \in X_i,\ i = 1, 2, \dots, m;$

3) *on each position set* X_i, $i \in \{1, 2, \dots, m\}$ *there exists a map* $s^{i^*} : X_i \to A$ *such that*
$$s^{i^*}(x) = a^* \in Arg \min_{a \in A(x)} \Big\{ \mu^i_{x,a} + \sum_{y \in X} p^a_{x,y} \varepsilon^i_y - \varepsilon^i_x - \omega^i \Big\}$$
and
$$\mu^j_{x,a^*} + \sum_{y \in X} p^{a^*}_{x,y} \varepsilon^j_y - \varepsilon^j_x - \omega^j = 0, \quad \forall x \in X_i, \quad j = 1, 2, \dots, m.$$

If such conditions hold then the set of maps $s^{1^*}, s^{2^*}, \dots, s^{m*}$ *determines a Nash equilibrium of the game for an arbitrary starting position* $x \in X$ *and*

$$F^i_x(s^{1^*}, s^{2^*}, \dots, s^{m*}) = \omega^i_x,\ i = 1, 2, \dots, m.$$

Proof. The sufficiency condition of the theorem is evident. Let us prove the necessity one. Assume that for the considered average stochastic positional game there exists a Nash equilibrium $s^* = (s^{1^*}, s^{2^*}, \dots, s^{m*})$ which is a Nash equilibrium for an arbitrary starting position of the game. Denote

$$\sigma_x^i = \widehat{F}_x^i(s^{1^*}, s^{2^*}, \dots, s^{m*}), \quad \forall x \in X, \quad i = 1, 2, \dots, m$$

and consider the following auxiliary game $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{\overline{c}^i\}_{i=\overline{1,m}},\ p,\ x)$, where

$$\overline{c}_{x,y}^i = c_{x,y}^i - \omega_x^i, \ \forall x, y \in X, \quad i = 1, 2, \dots, m.$$

Then according to Lemma 1 the auxiliary game has the same Nash equilibrium $s^* = (s^{1^*}, s^{2^*}, \dots, s^{m*})$ as initial one. Moreover, this equilibrium is a Nash equilibrium for an arbitrary starting position of the game and

$$\overline{F}_x^i(s^{1^*}, s^{2^*}, \dots, s^{m*}) = 0, \ \forall x \in X, \quad i = 1, 2, \dots, m.$$

Therefore, according to Corollary 1, for the auxiliary game there exist the functions

$$\varepsilon^i : X \to R, \quad i = 1, 2, \dots, m$$

and the values $\overline{\omega}^1, \overline{\omega}^2, \dots, \overline{\omega}^m$ ($\overline{\omega}^i = 0,\ i = 1, 2, \dots, m$), that satisfy the conditions of Theorem 4, i.e.

1) $\overline{\mu}_{x,a}^i + \sum\limits_{y \in X} p_{x,y}^a \varepsilon_y^i - \varepsilon_x^i - \overline{\omega}_x^i \geq 0, \quad \forall x \in X_i,\ \forall a \in A(x),\ i = 1, 2, \dots, m,$
where $\overline{\mu}_{x,a}^i = \sum\limits_{y \in X} p_{x,y}^a \overline{c}_{x,y}^i$;

2) $\min\limits_{a \in A(x)} \{\overline{\mu}_{x,a}^i + \sum\limits_{y \in X} p_{x,y}^a \varepsilon_y^i - \varepsilon_x^i - \overline{\omega}_x^i\} = 0, \quad \forall x \in X_i,\ i = 1, 2, \dots, m;$

3) on each position set $X_i,\ i \in \{1, 2, \dots, m\}$ there exists a map $s^{i^*} : X_i \to A$ such that

$$s^{i^*}(x) = a^* \in Arg \min_{a \in A(x)} \left\{\overline{\mu}_{x,a}^i + \sum_{y \in X} p_{x,y}^a \varepsilon_y^i - \varepsilon_x^i - \overline{\omega}^i\right\}$$

and

$$\overline{\mu}_{x,a^*}^j + \sum_{y \in X} p_{x,y}^{a^*} \varepsilon_y^j - \varepsilon_x^j - \overline{\omega}^j = 0, \quad \forall x \in X_i, \quad j = 1, 2, \dots, m.$$

Taking into account that $\overline{\omega}_x^i = 0$, and $\overline{\mu}_{x,a}^i = \mu_{x,a}^i - \omega_x^i$ (because $\overline{c}_{x,y}^i = c_{x,y} - \omega_x^i$) we obtain conditions $1 - 3$ of the theorem.

4. Nash Equilibria Conditions for Stochastic Positional Games with Discounted Payoffs

Now we formulate Nash equilibria conditions in the terms of bias equations for stochastic positional games with discounted payoffs.

Theorem 6. *Let a stochastic positional game* $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ \gamma,\ \overline{x})$ *with a discount factor* $0 < \gamma < 1$ *be given. Then there exist the values* $\sigma_x^i,\ \ i = 1, 2, \dots, m$, *for* $x \in X$ *that satisfy the following conditions:*

1) $\mu^i_{x,a} + \gamma \sum_{y\in X} p^a_{x,y}\sigma^i_y - \sigma^i_x \geq 0, \quad \forall x \in X_i, \quad \forall a \in A(x), \quad i = 1,2,\ldots,m,$
where $\mu^i_{x,a} = \sum_{y\in X} p^a_{x,y} c^i_{x,y}$.

2) $\min_{a\in A(x)} \left\{\mu^i_{x,a} + \gamma \sum_{y\in X} p^a_{x,y}\sigma^i_y - \sigma^i_x\right\} = 0, \quad \forall x \in X_i, \quad i = 1,2,\ldots,m;$

3) *on each position set* X_i, $i \in \{1,2,\ldots,m\}$ *there exists a map* $s^{i^*} : X_i \to A$ *such that*

$$s^{i^*}(x) = a^* \in Arg \min_{a\in A(x)} \left\{\mu^i_{x,a} + \gamma \sum_{y\in X} p^a_{x,y}\sigma^i_y - \sigma^i_x\right\}, \quad \forall x \in X_i$$

and

$$\mu^j_{x,a^*} + \gamma \sum_{y\in X} p^{a^*}_{x,y}\sigma^j_y - \sigma^j_x = 0, \quad \forall x \in X_i, \quad j = 1,2,\ldots,m.$$

The set of maps $s^{1^*}, s^{2^*}, \ldots, s^{m*}$ *determines a Nash equilibrium situation* $s^* = (s^{1^*}, s^{2^*}, \ldots, s^{m*})$ *for the stochastic positional game with discounted payoffs, where*

$$\widehat{F}^i_{\overline{x}}(s^{1^*}, s^{2^*}, \ldots, s^{m*}) = \sigma^i_{\overline{x}}, \ \forall \overline{x} \in X, \quad i = 1,2,\ldots,m.$$

Moreover, the situation $s^* = (s^{1^*}, s^{2^*}, \ldots, s^{m*})$ *is a Nash equilibrium for an arbitrary starting position* $\overline{x} \in X$.

Proof. According to Theorem 3 for the discounted stochastic positional game $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ \gamma,\ \overline{x})$ there exists a Nash equilibrium $s^* = (s^{1^*}, s^{2^*}, \ldots, s^{m*})$ which is a Nash equilibrium for an arbitrary starting position $\overline{x} \in X$ of the game. Denote

$$\sigma^i_x = \widehat{F}^i_x(s^{1^*}, s^{2^*}, \ldots, s^{m*}), \quad \forall x \in X, \quad i = 1,2,\ldots,m.$$

Let us fix the strategies $s^{1^*}, s^{2^*}, \ldots, s^{i-1^*}, s^{i+1^*}, \ldots, s^{m*}$ of the players $1,2,\ldots,i-1, i+1,\ldots,m$ and consider the problem of determining the expected total discounted cost with respect to player i. Obviously, the optimal stationary strategy for this problem is s^{i^*}. Then according to the properties of the bias equations for this Markov decision problem with discounted costs there exist the values σ^i_x, $i = 1,2,\ldots,m$, for $x \in X$ that satisfy the conditions:

1) $\mu^i_{x,a} + \gamma \sum_{y\in X} p^a_{x,y}\sigma^i_y - \sigma^i_x \geq 0, \quad \forall x \in X_i, \ \forall a \in A(x), \quad i = 1,2,\ldots,m;$

2) $\min_{a\in A(x)} \left\{\mu^i_{x,a} + \gamma \sum_{y\in X} p^a_{x,y}\sigma^i_y - \sigma^i_x\right\} = 0, \qquad \forall x \in X_i \quad i = 1,2,\ldots,m.$

Moreover, for fixed strategies $s^{1^*}, s^{2^*}, \ldots, s^{i-1^*}, s^{i^*}, s^{i+1^*}, \ldots, s^{m*}$ of the corresponding players $1,2,\ldots,i-1,i+1,\ldots,m$ we can select the strategy s^{i^*} of the player i where

$$s^{i^*}(x) \in Arg \min_{a\in A(x)} \left\{\mu^i_{x,a} + \gamma \sum_{y\in X} p^a_{x,y}\sigma^i_y - \sigma^i_x\right\}$$

and

$$\widehat{F}^i_{\overline{x}}(s^{1^*}, s^{2^*}, \ldots, s^{m*}) = \sigma^i_{\overline{x}}, \quad \forall \overline{x} \in X, \quad i = 1,2,\ldots,m.$$

This means that the conditions 1)–3) of the theorem hold.

5. Saddle Point Conditions for Antagonistic Stochastic Positional Games

The antagonistic stochastic positional game with the average payoff corresponds to the case of the game from Section 2 in the case $m = 2$ when $c = c^1 = -c^2$. So, we have a game $(X, A, X_1, X_2, c, p, \overline{x})$ where the stationary strategies s^1 and s^2 of the players are defined as two maps

$$s^1 : x \to a \in A^1(x) \ \text{ for } \ x \in X_1; \quad s^2 : x \to a \in A^1(x) \ \text{ for } \ x \in X_2.$$

and the payoff function $F_x(s^1, s^2)$ of the players is determined by the values of average costs ω_x^s in the Markov processes with the corresponding probability matrices P^s induced by the situations $s = (s^1, s^2) \in S$. For this game saddle points ${s^1}^*, {s^2}^*$ always exists (Lozovanu and Pickl, 2014) , i.e. for a given starting position $\overline{x} \in X$ holds

$$F_{\overline{x}}({s^1}^*, {s^2}^*) = \min_{s^1 \in S^1} \max_{s^2 \in S^2} F_{\overline{x}}(s^1, s^2) = \max_{s^2 \in S^2} \min_{s^1 \in S^1} F_{\overline{x}}(s^1, s^2).$$

Theorem 7. *Let* $(X,\ A,\ X_1, X_2, c,\ p, \overline{x})$ *be an arbitrary antagonistic stochastic positional game with an average payoff function* $F_{\overline{x}}(s_1, s_2)$*. Then the system of equations*

$$\begin{cases} \varepsilon_x + \omega_x = \max\limits_{a \in A(x)} \left\{ \mu_{x,a} + \sum\limits_{y \in X} p_{x,y}^a \varepsilon_y \right\}, & \forall x \in X_1; \\ \\ \varepsilon_x + \omega_x = \min\limits_{a \in A(x)} \left\{ \mu_{x,a} + \sum\limits_{y \in X} p_{x,y}^a \varepsilon_y \right\}, & \forall x \in X_2; \end{cases}$$

has solution under the set of solutions of the system of equations

$$\begin{cases} \omega_x = \max\limits_{a \in A(x)} \left\{ \sum\limits_{y \in X} p_{x,y}^a \omega_x \right\}, & \forall x \in X_1; \\ \\ \omega_x = \min\limits_{a \in A(x)} \left\{ \sum\limits_{y \in X} p_{x,y}^a \omega_x \right\}, & \forall x \in X_2, \end{cases}$$

i.e. the last system of equations has such a solution $\omega_x^*,\ x \in X$ *for which there exists a solution* $\varepsilon_x^*,\ x \in X$ *of the system of equations*

$$\begin{cases} \varepsilon_x + \omega_x^* = \max\limits_{a \in A(x)} \left\{ \mu_{x,a} + \sum\limits_{y \in X} p_{x,y}^a \varepsilon_y \right\}, & \forall x \in X_1; \\ \\ \varepsilon_x + \omega_x^* = \min\limits_{a \in A(x)} \left\{ \mu_{x,a} + \sum\limits_{y \in X} p_{x,y}^a \varepsilon_y \right\}, & \forall x \in X_2. \end{cases}$$

The optimal stationary strategies of the players

$$ {s_1}^* : x \ \to \ a^1 \in A(x) \quad for \quad x \in X_1;$$

$$ {s_2}^* : x \ \to \ a^2 \in A(x) \quad for \quad x \in X_2$$

in the antagonistic stochastic positional game can be found by fixing arbitrary maps ${s_1}^*(x) \in A(x)$ *for* $x \in X_1$ *and* ${s_2}^*(x) \in A(x)$ *for* $x \in X_2$ *such that*

$$s_1{}^*(x) \in \left(Arg \max_{a\in A(x)} \left\{ \sum_{y\in X} p^a_{x,y}\omega^*_x \right\}\right) \bigcap \left(Arg \max_{a\in A(x)} \left\{ \mu_{x,a} + \sum_{y\in X} p^a_{x,y}\varepsilon^*_y \right\}\right),$$

$$\forall x \in X_1$$

and

$$s_2{}^*(x) \in \left(Arg \min_{a\in A(x)} \left\{ \sum_{y\in X} p^a_{x,y}\omega^*_x \right\}\right) \bigcap \left(Arg \min_{a\in A(x)} \left\{ \mu_{x,a} + \sum_{y\in X} p^a_{x,y}\varepsilon^*_y \right\}\right).$$

$$\forall x \in X_2$$

For the strategies ${s^1}^*, {s^2}^*$ *the corresponding values of the payoff function* $F_{\overline{x}}({s^1}^*, {s^2}^*)$ *coincides with the values* $\omega^*_{\overline{x}}$ *for* $\overline{x} \in X$ *and*

$$F_x({s^1}^*, {s^2}^*) = \min_{s^1\in S^1} \max_{s^2\in S^2} F_x(s^1, s^2) = \max_{s^2\in S^2} \min_{s^1\in S^1} F_x(s^1, s^2) \quad \forall x \in X.$$

.

Based on the constructive proof of this theorem (see Lozovanu and Pickl, 2014) an algorithm for determining the saddle points in antagonistic stochastic positional games has been elaborated. The saddle point conditions for antagonistic stochastic positional games with a discounted payoff can be derived from Theorem 6.

6. Conclusion

Stochastic positional games with average and discounted payoffs represent a special class of Shapley stochastic games that extends deterministic positional games. For the considered class of games Nash equilibria conditions have been formulated and proven. Based on these results new algorithms for determining the optimal stationary strategies of the players can be elaborated.

References

Condon, A. (1992). *The complexity of stochastic games.* Informations and Computation. **96(2)**, 203–224.

Ehrenfeucht, A., Mycielski, J. (1979). *Positional strategies for mean payoff games.* International Journal of Game Theory, **8**, 109–113. **8**–113.

Filar, J.A., Vrieze, K. (1997). *Competitive Markov Decision Processes.* Springer, 1997.

Gillette, D. (1957) *Stochastic games with zero stop probabilities.* Contribution to the Theory of Games, **vol. III**, Princeton, 179-187.

Gurvich, V.A., Karzanov, A.V., Khachian, L.G. (1988). *Cyclic games and an algorithm to find minimax cycle means in directed graphs.* USSR, Computational Mathematics and Mathematical Physics, **28**, 85–91.

Howard, R.A. (1960). *Dynamic Programming and Markov Processes.* Wiley.

Lal, A.K., Sinha S. (1992) *Zero-sum two person semi-Markov games*, J. Appl. Prob., **29**, 56-72.

Lozovanu, D. (2011) *The game-theoretical approach to Markov decision problems and determining Nash equilibria for stochastic positional games.* Int. J. Mathematical Modelling and Numerical Optimization, **2(2)**, 162–164.

Lozovanu, D., Pickl, S. (2006) *Nash equilibria conditions for Cyclic Games with p players.* Eletronic Notes in Discrete Mathematics, **25**, 117–124.

Lozovanu, D., Pickl, S. (2009) *Optimization and Multiobjective Control of Time-Discrete Systems*. Springer.

Lozovanu, D., Pickl, S., Kropat, E. (2011) *Markov decision processes and determining Nash equilibria for stochastic positional games*, Proceedings of 18th World Congress IFAC-2011, 13398–13493.

Lozovanu, D., Pickl, S. (2014) *Antagonistic Positional Games in Markov Decision Processes and Algorithms for Determining the Saddle Points*. Discrete Applied Mathematics, 2014 (Accepted for publication).

Mertens, J.F., Neyman, A. (1981) *Stochastic games*. International Journal of Game Theory, **10**, 53-66.

Moulin, H. (1976) *Prolongement des jeux a deux joueurs de somme nulle*. Bull. Soc. Math. France., Mem **45**.

Nash, J.F. (1951) *Non cooperative games*. Annals of Mathematics, **2**, 286–295.

Neyman, A., Sorin, S. (2003). *Stochastic games and applications*. NATO ASI series, Kluver Academic press.

Puterman, M. (2005). *Markov Decision Processes: Stochastic Dynamic Programming*. John Wiley, New Jersey.

Shapley L. (1953). *Stochastic games*. Proc. Natl. Acad. Sci. U.S.A. **39**, 1095-1100.