Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

Markov Decision Processes and Determining Nash Equilibria for Stochastic Positional Games

Dmitrii Lozovanu * **Stefan Pickl** ** **Erik Kropat** ***

* *Academy of Sciences of Moldova, Institute of Mathematics and Computer Science, 5 Academiei str., Chişinău, MD−2028, Moldova (e-mail: lozovanu@math.md).*
** *Universität der Bundeswehr München, Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany (e-mail: stefan.pickl@unibw.de)*
*** *Universität der Bundeswehr München, Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany (e-mail: erik.kropat@unibw.de)*

Abstract: A class of stochastic positional games which extend the cyclic games and Markov decision problems with average and discounted optimization costs criteria is formulated and studied. Nash equilibria conditions for considered class of stochastic positional games are derived and some approaches for determining Nash equilibria are described.

Keywords: Markov decision processes; Noncooperative Games, Stochastic positional games, Nash equilibria; Optimal stationary strategies, Cyclic games

1. INTRODUCTION

In this paper we formulate and study a class of stochastic positional games applying the game-theoretical concept to Markov decision problems with average and discounted costs optimization criteria. We consider Markov decision processes that may be controlled by several actors (players). The set of states of the system in such processes is divided into several disjoint subsets which represent the corresponding positions sets of the players. Each player has to determine which action should be taken in each state of his positions set in order to minimize his own average cost per transition or discounted expected total cost. The cost of system's transition from one state to another in the Markov process is given for each player separately. In addition the set of actions, the transition probability functions and the starting state are known. We assume that players use only stationary strategies, i.e. each player in an arbitrary his position uses the same action for an arbitrary discrete moment of time. In the considered stochastic positional games we are seeking for a Nash equilibrium.

The main results we describe in this paper are concerned with existence of Nash equilibria in the considered games and elaboration of algorithms for determining the optimal stationary strategies of players. We show that Nash equilibria for the game model with average cost payoff functions of the players exists if an arbitrary situation generated by the strategies of players induces a Markov unichain. For the game model with discounted payoff function we show that Nash equilibria always exists. The obtained results can be easy extended for antagonistic game models of Markov decision problems and the corresponding conditions for existence of saddle points in such games can be derived.

The proposed approach for Markov decision processes can be extended for multi-objective decision problems with Stackelberg and Pareto optimization principles and the corresponding algorithms for determining the optimal solutions of problems in the sense of Stackelberg and Pareto can be developed.

2. STOCHASTIC POSITIONAL GAMES WITH AVERAGE PAYOFF FUNCTIONS OF PLAYERS

We consider a class of stochastic positional games that extends and generalizes cyclic games (Gurvich [1988], Lozovanu [2006]) and Markov decision problems with average and discounted optimization costs criteria (Puterman [2005], White [1993]). The considered class of games we formulate using the framework of 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 $c : X \times X \to \mathbb{R}$ which gives the costs $c_{x,y}$ of states transitions for the dynamical system when it makes a transition from the state $x \in X$ to another state $y \in X$.

We consider the noncooperative game model with m players in which m transition cost functions are given

$$c^i : X \times X \to \mathbb{R}, \quad i = 1, 2, \dots, m,$$

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

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

where X_i represents the positions set of 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 consider the stationary game model, i.e. we assume that each player fixes actions in the states from his positions set using stationary strategies. The stationary strategies of players we define as m maps:

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

where $A^i(x)$ is the set of actions of 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 $p^{s^i}_{x,y}(x)$ in the states $x \in X_i$ for $i = 1, 2, \ldots, m$. We denote $P^s = (p^s_{x,y})$ the matrix of probability transitions of this Markov process. If the starting state x_{i_0} is given, then for the Markov process with the matrix of probability transitions P^s we can determine the average cost per transition $M^i_{x_0}(s^1, s^2, \ldots, s^m)$ 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 players as follows:

$$F^i_{x_{i_0}}(s^1, s^2, \ldots, s^m) = M^i_{x_{i_0}}(s^1, s^2, \ldots, s^m),\ i = 1, 2, \ldots, m.$$

In such a way we obtain a discrete noncooperative game in normal form which is determined by finite sets of strategies $\mathbb{S}^1, \mathbb{S}^2, \ldots, \mathbb{S}^m$ of m players and the payoff functions defined above. In this game we are seeking for a Nash equilibrium (Nash [2050]), i.e. we consider the problem of determining the stationary strategies

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

such that

$$F^i_{x_{i_0}}(s^{1^*}, s^{2^*}, \ldots, s^{i-1^*}, s^{i^*}, s^{i+1^*} \ldots, s^{m*}) \leq$$
$$\leq F^i_{x_{i_0}}(s^{1^*}, s^{2^*}, \ldots, s^{i-1^*}, s^{i}, s^{i+1^*} \ldots, s^{m*}),$$

$(\forall s^i \in S^i, \quad i = 1, 2, \ldots, m)$. The game defined above is determined uniquely by the set of states X, the positions sets $X_1, X_2, \ldots, X_m$, the set of actions A, the cost functions $c^i : X \times X \to \mathbb{R}$, $i = 1, 2, \ldots, m$, the probability function $p : X \times X \times A \to [0, 1]$ and the starting position x_{i_0}. Therefore we denote it $(X,\ A,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ p,\ x_{i_0})$. We call this game *stochastic positional game with average payoff functions* . In the case $p^a_{x,y} = 0 \vee 1$, $\forall x, y \in X$, $\forall a \in A$ the stochastic positional game is transformed into the cyclic game studied by Gurvich [1988], Lozovanu [2009].

3. DETERMINING NASH EQUILIBRIA FOR STOCHASTIC POSITIONAL GAMES WITH AVERAGE PAYOFF FUNCTIONS

To provide the existence of Nash equilibria for the considered stochastic positional game we shall use the following condition. We assume that an arbitrary situation $s = (s^1, s^2, \ldots, s^m)$ of the game generates a Markov unichain with the corresponding matrix of probability transitions $P^s = (p^s_{x,y})$. The Markov process with such property with respect to the situations $s = (s^1, s^2, \ldots, s^m) \in S$ of the game we call *perfect Markov decision process*. We show that in this case the problem of determining Nash equilibria for a stochastic positional game can be formulated as continuous model that represents the game variant of the following optimization problem:

Minimize

$$\psi(s, q) = \sum_{x \in X} \sum_{a \in A(x)} \mu_{x,a} s_{x,a}\, q_x \tag{1}$$

subject to

$$\begin{cases} \displaystyle\sum_{x \in X} \sum_{a \in A(x)} p^a_{x,y} s_{x,a} q_x = q_y, \quad \forall y \in X; \\ \displaystyle\sum_{x \in X} q_x = 1; \\ \displaystyle\sum_{a \in A(x)} s_{x,a} = 1, \quad \forall x \in X; \\ s_{x,a} \geq 0, \quad \forall x \in X,\ a \in A(x), \end{cases} \tag{2}$$

where

$$\mu_{x,a} = \sum_{y \in X^+(x)} c_{x,y}\, p^a_{x,y}$$

is the immediate cost in the state $x \in X$ for a fixed action $a \in A(x)$.

It is easy to observe that the problem (1), (2) represents the continuous model for Markov decision problem with average cost criterion. Indeed, an arbitrary stationary strategy $s : X \to A$ can be identified with the set of boolean variables $s_{x,a} \in \{0, 1\}$, $x \in X$, $a \in A(x)$ that satisfy the conditions

$$\sum_{a \in A(x)} s_{x,a} = 1, \quad \forall x \in X; \quad s_{x,a} \geq 0,\ \forall x \in X, a \in A.$$

These conditions determine all feasible solutions of the system (2). The rest restrictions in (2) correspond to the system of linear equations with respect to q_x for $x \in X$. This system of linear equations reflects the ergodicity condition for the limiting probability $q_x, x \in X$ in the Markov unichain, where $q_x, x \in X$ are determined uniquely for given $s_{x,a}$, $\forall x \in X, a \in A(x)$. Thus, the value of the objective function (1) expresses the average cost per transition in this Markov unichain and an arbitrary optimal solution $s^*_{x,a}, q^*_x$ $(x \in X,\ a \in A)$ of problem (1), (2) with $s^*_{x,a} \in \{0, 1\}$ represents an optimal stationary strategy for Markov decision problem with average cost criterion. If such an optimal solution is known, then an optimal action for Markov decision problem can be found by fixing $a^* = s^*(x)$ for $x \in X$ if $s^*_{x,a} = 1$.

The problem (1), (2) can be transformed into a linear programming problem using the notations $\alpha_{x,a} =$

$s_{x,a}q_x$, $\forall x \in X, a \in A(x)$. Based on such transformation of the problem we will describe some additionally properties of the optimal stationary strategies in Markov decision processes.

Lemma 1. Let a Markov decision process (X, A, p, c) be given and consider the function

$$\psi(s) = \sum_{x\in X} \sum_{a\in A(x)} \mu_{x,a} s_{x,a}\, q_x,$$

where q_x for $x \in X$ satisfy the condition

$$\begin{cases} \displaystyle\sum_{x\in X} \sum_{a\in A(x)} p^a_{x,y} s_{x,a} q_x = q_y, \quad \forall y \in X; \\ \displaystyle\sum_{x\in X} q_x = 1. \end{cases} \tag{3}$$

Assume that an arbitrary stationary strategy s in the Markov decision process generates a Markov unichain, i.e we have a perfect Markov decision process. Then the function $\psi(s)$ depends only on $s_{x,a}$ for $x \in X, a \in A(x)$, and on the set S of solutions of the system

$$\begin{cases} \displaystyle\sum_{a\in A(x)} s_{x,a} = 1, \quad \forall x \in X; \\ s_{x,a} \geq 0, \quad \forall x \in X,\ a \in A(x), \end{cases} \tag{4}$$

the function $\psi(s)$ is monotone.

Proof. In the perfect Markov decision processes an arbitrary basic solution of the system (4) corresponds to a stationary strategy that generates a Markov unichain. For such an arbitrary strategy the rank of system (3) is equal to $|X|$ and (3) has a unique solution with respect to q_x ($x \in X$) (see Puterman [2005], White [1993]). Moreover, in the mentioned references is shown that for Markov unichain the system of linear equations (3) uniquely determines $q_x, \forall x \in X$ for an arbitrary solution of system (4).

Now let us prove the second part of the lemma. We show that on the set of solutions of system (4) the function $\psi(s)$ is monotone. For this reason it is sufficient to show that for arbitrary $s', s'' \in S$ with $\psi(s') \neq \psi(s'')$ the following relation holds

$$\min\{\psi(s'), \psi(s'')\} < \psi(\overline{s}) < \max\{\psi(s'), \psi(s'')\}. \tag{5}$$

if $\overline{s} = \theta s' + (1-\theta)s''$, $0 < \theta < 1$.

We show that the relation (5) holds for an arbitrary $\overline{s} \in S(s', s'')$, where

$$S(s', s'') = \{\ s \mid \min\{s'_{x,a}, s''_{x,a}\} < s_{x,a} < \max\{s'_{x,a}, s''_{x,a}\},$$
$$\forall x \in X, a \in A(x)\}$$

and the equations

$$\psi(s) = \psi(s'), \quad \psi(s) = \psi(s'')$$

on the set

$$\overline{S}(s', s'') = \{\ s \mid \min\{s'_{x,a}, s''_{x,a}\} \leq s_{x,a} \leq \max\{s'_{x,a}, s''_{x,a}\},$$
$$\forall x \in X, a \in A(x)\}$$

have the unique solutions $s = s'$ and $s = s''$, respectively. The correctness of this property we prove using the relationship of the problem (1), (2) with the following linear programming problem:

Minimize

$$\overline{\psi}(\alpha) = \sum_{x\in X} \sum_{a\in A(x)} \mu_{x,a}\, \alpha_{x,a} \tag{6}$$

subject to

$$\begin{cases} \displaystyle\sum_{x\in X} \sum_{a\in A(x)} p^a_{x,y}\, \alpha_{x,a} = q_y, \quad \forall y \in X; \\ \displaystyle\sum_{x\in X} q_x = 1; \\ \displaystyle\sum_{a\in A(x)} \alpha_{x,a} = q_x, \quad \forall x \in X; \\ \alpha_{x,a} \geq 0, \ \ \forall x \in X,\ a \in A(x). \end{cases} \tag{7}$$

The problem (6), (7) is obtained from (1), (2) introducing the substitutions $\alpha_{x,a} = s_{x,y}q_x$ for $x \in X$, $a \in A(x)$. These substitutions allow us to establish a bijective mapping between the set of feasible solutions of the problem (1), (2) and the set of feasible solutions of the linear programming problem (6), (7). So, if $\alpha_{x,a}$ for $x \in X$, $a \in A(x)$ and $\overline{\psi}(\alpha)$ are known then we can uniquely determine

$$s_{x,a} = \frac{\alpha_{x,a}}{q_x}, \quad \forall x \in X,\ a \in A(x) \tag{8}$$

for which $\psi(s) = \overline{\psi}(\alpha)$. In particular, if an optimal basic solution α^*, q^* of the linear programming problem (6), (7) is found, then the optimal stationary strategy for Markov decision problem can be found fixing

$$s^*_{x,a} = \begin{cases} 1, & \text{if} \quad \alpha^*_{x,a} > 0; \\ 0, & \text{if} \quad \alpha^*_{x,a} = 0. \end{cases}$$

Let s', s'' be arbitrary solutions of the system (4) where $\psi(s') < \psi(s'')$. Then there exist the corresponding feasible solutions α', α'' of the linear programming problem (6), (7) for which

$$\psi(s') = \overline{\psi}(\alpha'), \quad \psi(s'') = \overline{\psi}(\alpha''),$$
$$\alpha'_{x,a} = s'_{x,a}q'_x, \quad \alpha''_{x,y} = s''_{x,a}q''_x \ \ \forall x \in X,\ a \in A(x),$$

where q'_x, q''_x are determined uniquely from the system of linear equations (3) for $s = s'$ and $s = s''$, respectively. The function $\overline{\psi}(\alpha)$ is linear and therefore for an arbitrary $\overline{\alpha} = \theta\alpha' + (1-\theta)\alpha''$, $0 \leq \theta \leq 1$ the following equality holds

$$\overline{\psi}(\overline{\alpha}) = \theta\overline{\psi}(\alpha') + (1-\theta)\overline{\psi}(\alpha''),$$

where $\overline{\alpha}$ is a feasible solution of the problem (6), (7), that in initial problem (1), (2) corresponds to a feasible solution $\overline{s}$ for which

$$\psi(\overline{s}) = \overline{\psi}(\overline{\alpha}); \quad \overline{q}_x = \theta q'_x + (1-\theta)q_x, \ \forall x \in X.$$

Using (8) we have

$$\overline{s}_{x,a} = \frac{\overline{\alpha}_{x,a}}{\overline{q}_x}, \quad \forall x \in X,\ a \in A(x),$$

i.e.

$$\overline{s}_{x,a} = \frac{\theta\alpha'_{x,a} + (1-\theta)\alpha''_{x,a}}{\theta q'_x + (1-\theta)q''_x} = \frac{\theta s'_{x,a}q'_x + (1-\theta)s''_{x,a}q''_x}{\theta q'_x + (1-\theta)q''_x} =$$
$$= \frac{\theta q'_x}{\theta q'_x + (1-\theta)q''_x} s'_{x,a} + \frac{(1-\theta)q''_x}{\theta q'_x + (1-\theta)q''_x} s''_{x,a}.$$

So, we obtain

$$\overline{s}_{x,a} = \overline{\theta}_x s'_{x,a} + (1-\overline{\theta}_x)s''_{x,a},$$

where

$$\overline{\theta}_x = \frac{\theta q'_x}{\theta q'_x + (1-\theta)q''_x}, \quad 0 \leq \theta \leq 1.$$

It is easy to observe that $0 \le \overline{\theta}_x \le 1$, were $\overline{\theta}_x = 0$, $\forall x \in X$ if and only if $\theta = 0$ and $\overline{\theta}_x = 1$, $\forall x \in X$ if and only if $\theta = 1$. This means that for an arbitrary $\overline{s} \in S(s', s'')$ the condition (5) holds and the equations

$$\psi(s) = \psi(s'), \quad \psi(s) = \psi(s'')$$

on the set $\overline{S}(s', s'')$ have the unique solutions $s = s'$ and $s = s''$, respectively. Thus the function $\psi(s)$ on the set of solutions of system (4) is monotone.

Now we extend the results described above for the continuous model of stochastic positional game with average payoff functions. We consider the game model for perfect Markov decision processes.

Let denote by S^i, $i \in \{1, 2, \dots m\}$ the set of solutions of the system

$$\begin{cases} \displaystyle\sum_{a \in A(x)} s^i_{x,a} = 1, \quad \forall x \in X^i; \\ s^i_{x,a} \ge 0, \quad \forall x \in X^i, \ a \in A(x). \end{cases} \tag{9}$$

So, S^i is a convex compact set and its arbitrary extreme point corresponds to a basic solution s' of the system (9), where $s'_{x,a} \in \{0, 1\}, \forall x \in X_i, a \in A(x)$. Thus, if s' is an arbitrary basic solution of system (9), then $s' \in \mathbb{S}^i$.

On the set $S = S^1 \times S^2 \times \cdots \times S^m$ we define m payoff functions

$$\psi^i(s^1, s^2, \dots, s^m) = \sum_{i=1}^{m} \sum_{x \in X_i} \sum_{a \in A(x)} \mu^i_{x,a} s^i_{x,a} q_x, \quad i = \overline{1, m}, \tag{10}$$

where

$$\mu^i_{x,a} = \sum_{y \in X} c^i_{x,y} p^a_{x,y}$$

is the immediate cost of player $i \in \{1, 2, \dots, m\}$ in the state $x \in X$ for a fixed action $a \in A(x)$; q_x for $x \in X$ are determined uniquely from the following system of linear equations

$$\begin{cases} \displaystyle\sum_{i=1}^{m} \sum_{x \in X_i} \sum_{a \in A(x)} p^a_{x,y} s^i_{x,a} q_x = q_y, \quad \forall y \in X; \\ \displaystyle\sum_{x \in X} q_x = 1 \end{cases} \tag{11}$$

when $s^1, s^2, \dots, s^m$ are given.

The main results we prove for our game model represent the following properties:

- The set of Nash equilibria situations of the continuous model is non empty if and only if the set of Nash equilibria situations of the game in positional form is not empty;

- If $(s^1, s^2, \dots, s^m)$ is an extreme point of S then $F^i_x(s^1, s^2, \dots, s^m) = \psi(s^1, s^2, \dots, s^m)$, $\forall x \in X$, $i = 1, 2. \dots, m$ and all Nash equilibria situations for the continuous game model that correspond to extreme points in S represent Nash equilibria situations for the game in positional form.

From Lemma 1 as a corollary we obtain the following result.

Lemma 2. For perfect Markov processes each payoff function $\psi^i(s^1, s^2, \dots, s^m)$, $i \in \{i, 2, \dots, m\}$ possesses the property that $\psi^i(\overline{s}^1, \overline{s}^2, \dots, \overline{s}^{i-1}, s^i, \overline{s}^{i+1}, \dots, \overline{s}^m)$ is monotone with respect to $s^i \in S^i$ for arbitrary fixed $\overline{s}^k \in S^k$, $k = 1, 2, \dots, i-1, i+1, \dots, m$.

Using this lemma we can prove the following theorem.

Theorem 3. Let $(X, A, \{X_i\}_{i=\overline{1,m}}, \{c^i\}_{i=\overline{1,m}}, p, x)$ be a stochastic positional game with a given starting position $x \in X$ and average payoff functions

$$F^1_x(s^1, s^2, \dots, s^m), \ F^2_x(s^1, s^2, \dots, s^m), \dots,$$
$$F^m_x(s^1, s^2, \dots, s^m),$$

of players $1, 2, \dots, m$, respectively. If 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 a Markov uni-chain then for the stochastic positional game $(X, A, \{X_i\}_{i=\overline{1,m}}, \{c^i\}_{i=\overline{1,m}}, p, x)$ there exists Nash equilibrium $s^* = (s^{1*}, s^{2*}, \dots, s^{m*})$. Moreover, for this game there exists a situation $s^* = (s^{1*}, s^{2*}, \dots, s^{m*})$ which is a Nash-equilibrium for an arbitrary starting position $x \in X$.

Proof. According to Lemma 2 each function

$$\psi^i(s^1, s^2, \dots, s^m), i \in \{i, 2, \dots, m\}$$

satisfies the condition that

$$\psi^i(\overline{s}^1, \overline{s}^2, \dots, \overline{s}^{i-1}, s^i, \overline{s}^{i+1}, \dots, \overline{s}^m)$$

is monotone with respect to $s^i \in S^i$ for arbitrary fixed $\overline{s}^k \in S^k$, $k = 1, 2, \dots, i-1, i+1, \dots, m$. In the considered game each subset S^i is convex and compact. Therefore these conditions (see Debreu [1952], Dasgupta [1986], Simon [1987] and Reny [1999]) provide the existence of Nash equilibrium for the functions $\psi^i(s^1, s^2, \dots, s^m)$, $i \in \{i, 2, \dots, m\}$ on $S^1 \times S^2 \times \cdots \times S^m$. Taking into account that S is a polyhedron set and the functions $\psi^i(\overline{s}^1, \overline{s}^2, \dots, \overline{s}^{i-1}, s^i, \overline{s}^{i+1}, \dots, \overline{s}^m)$ are monotone we obtain that there exists a Nash equilibrium $s^{1*}, s^{2*}, \dots, s^{m*}$ that corresponds to a basic solution of the system (9). This means that $(s^{1*}, s^{2*}, \dots, s^{m*})$ is Nash equilibrium for the functions

$$F^1_x(s^1, s^2, \dots, s^m), \ F^2_x(s^1, s^2, \dots, s^m), \dots,$$
$$F^m_x(s^1, s^2, \dots, s^m)$$

on the set of situations $\mathbb{S} = \mathbb{S}^1 \times \mathbb{S}^2 \times \cdots \times \mathbb{S}^m$.

Using the results described above we may conclude that in the case of perfect Markov decision processes Nash equilibrium for stochastic positional games can be determined by using classical iterative methods for the continuous game models with payoff functions $\psi^i(s^1, s^2, \dots, s^m)$, $i \in \{i, 2, \dots, m\}$ on the set $S^1 \times S^2 \times \cdots \times S^m$.

In general, for stochastic positional games with average payoff functions of players, Nash equilibrium may not exists if the stationary strategies do not generate Markov uni-chain. Moreover, Nash equilibrium may not exists even for deterministic positional games (see Gurvich [1988], Lozovanu [2009]). So, the Theorem 3 in the case $p^a_{x,y} \in \{0, 1\}$, gives conditions for existence of Nash equilibria in cyclic games with average payoff functions.

4. STOCHASTIC POSITIONAL GAMES WITH DISCOUNTED PAYOFF FUNCTIONS OF PLAYERS

The stochastic positional game model for discounted Markov decision problem we formulate in a similar way as the game model from Section 2. We apply the game-theoretical concept to discounted Markov decision process (X, A, p, c) with given discounted factor γ, $0 < \gamma < 1$ (see Puterman [2005], White [1993]). So, in for our game model we assume that m transition cost functions $c^i : X \times X \to \mathbb{R}$, $i = 1, 2, \ldots, m$, 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 player $i \in \{1, 2, \ldots, m\}$. Thus, 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. The stationary strategies of players in this game we define 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 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 probability distributions $p^{s^i(x)}_{x,y}$ in the state x were $x \in X_i$. 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:

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

In such a way we obtain a new discrete noncooperative game in normal form which is determined by the sets of strategies $\mathbb{S}^1, \mathbb{S}^2, \ \ldots, \ \mathbb{S}^m$ of m players and the payoff functions defined above. In this game we are seeking for a Nash equilibrium.

This game is determined uniquely by the set of states X, the positions sets $X_1, X_2, \ldots, X_m$, the set of actions A, the cost functions $c^i : X \times X \to \mathbb{R}$, $i = 1, 2, \ldots, m,$, the probability function $p : X \times X \times A \to [0, 1]$ the discounted factor γ and the starting position x_0. Therefore we denote it $(X, \ A, \ \{X_i\}_{i=\overline{1,m}}, \ \{c^i\}_{i=\overline{1,m}}, \ p, \ \gamma, \ x_0)$. We call this game *stochastic positional game with discounted payoff functions.*

5. DETERMINING NASH EQUILIBRIA FOR STOCHASTIC POSITIONAL GAMES WITH DISCOUNTED PAYOFF FUNCTIONS

In this section we show that Nash equilibrium exists for an arbitrary stochastic positional game with discounted payoff functions of the players and given discounted factor γ, $0 < \gamma < 1$. To prove this result we shall use a continuous game which represent the game model for the following continuous optimization problem:
Maximize

$$\varphi_{x_0}(\sigma, s) = \sigma_{x_0} \tag{12}$$

subject to

$$\begin{cases} \sigma_x - \gamma \sum\limits_{y \in X} \sum\limits_{a \in A(x)} s_{x,a}\, p^a_{x,y}\, \sigma_y = \sum\limits_{a \in A(x)} s_{x,a}\, \mu_{x,a}, \forall x \in X; \\ \sum\limits_{a \in A(x)} s_{x,a} = 1, \quad \forall x \in X; \\ s_{x,a} \geq 0, \quad \forall x \in X, \, a \in A(x), \end{cases} \tag{13}$$

where

$$\mu_{x,a} = \sum_{y \in X} p^a_{x,y}\, c^a_{x,y}.$$

This problem represents the continuous model for discounted Markov decision problems. Based on this model we can determine the optimal stationary strategy of the discounted Markov decision problem for an arbitrary starting state $x \in X$. In (13) the system of linear equations with respect to σ_x has a unique solution and therefore the objective function (12) on the set of feasible solutions depends only on s. It is easy to observe that these equations in (13) can be changed by inequalities ($\leq$). If after that we dualize (12), (13) with respect to σ_x for fixed s then we obtain the following problem:
Minimize

$$\overline{\varphi}(s, \beta) = \sum_{x \in X} \sum_{a \in A(x)} \mu_{x,a}\, s_{x,a} \beta_x \tag{14}$$

subject to

$$\begin{cases} \beta_y - \gamma \sum\limits_{x \in X} \sum\limits_{a \in A(x)} p^a_{x,y}\, s_{x,a} \beta_x \geq 0, \quad \forall y \in X \setminus \{x_0\}; \\ \beta_y - \gamma \sum\limits_{x \in X} \sum\limits_{a \in A(x)} p^a_{x,y}\, s_{x,a} \beta_x \geq 1 \quad for \quad y = x_0; \\ \sum\limits_{a \in A(x)} s_{x,a} = 1, \quad \forall x \in X; \\ \beta_y \geq 0 \quad \forall y \in X; \quad s_{x,a} \geq 0, \quad \forall x \in X, \, a \in A(x). \end{cases} \tag{15}$$

Using elementary transformations in this problem and introducing the notations $\alpha_{x,a} = s_{x,s} \beta_x, \forall x \in X, a \in A(x)$ we obtain the following linear programming problem:
Minimize

$$\phi(s, \beta) = \sum_{x \in X} \sum_{a \in A(x)} \mu_{x,a}\, \alpha_{x,a} \tag{16}$$

subject to

$$\begin{cases} \beta_y - \gamma \sum\limits_{x \in X} \sum\limits_{a \in A(x)} p^a_{x,y}\, \alpha_{x,a} \geq 0, \quad \forall y \in X \setminus \{x_0\}; \\ \beta_y - \gamma \sum\limits_{x \in X} \sum\limits_{a \in A(x)} p^a_{x,y}\, \alpha_{x,a} \geq 1 \quad for \quad y = x_0; \\ \sum\limits_{a \in A(x)} \alpha_{x,a} = \beta_x, \quad \forall x \in X; \\ \beta_y \geq 0, \quad \forall y \in X; \quad \alpha_{x,a} \geq 0, \quad \forall x \in X, \, a \in A(x). \end{cases} \tag{17}$$

If (α^*, β^*) is an optimal basic solution of problem (16), (17) then the optimal stationary strategy s^* for the discounted Markov decision problem is determined as follows:

$$s^*_{x,a} = \begin{cases} 1, & \text{if} \quad \alpha^*_{x,a} \neq 0; \\ 0, & \text{if} \quad \alpha^*_{x,a} = 0. \end{cases} \tag{18}$$

and $\alpha^*_{x,a} = s^*_{x,a}\beta^*_x, \ \forall x \in X, a \in A(x)$.

It is easy to observe that $\beta_x > 0, \ \forall x \in X$ if for the considered Markov decision process there exists an action $a \in A(x_0)$ such that $p_{x_0,y} > 0, \forall x \in X$. Without loss of generality we may assume that such condition for our problem holds; otherwise we can add a fictive action a' in the state x_0 for which $p^{a'}_{x,y} > 0, \ \forall y \in X$ ($\sum\limits_{y\in X} p^a_{x,y} = 1$) and $c^{a'}_{x_0,y} = K, \ \forall y \in X$, where K is a suitable big value.

For continuous model of discounted Markov decision problem we prove a similar properties as for average Markov decision model.

Lemma 4. Let a Markov decision process (X, A, p, c) with discounted factor $\gamma, \ 0 < \gamma < 1$ be given. Consider the function

$$\varphi_{x_0}(s) = \sigma_{x_0},$$

where σ_x for $x \in X$ satisfy the condition

$$\sigma_x - \gamma \sum_{y\in X} \sum_{a\in A(x)} s_{x,a}\, p^a_{x,y}\, \sigma_y = \sum_{a\in A(x)} s_{x,a}\, \mu_{x,a}, \ \ \forall x \in X. \tag{19}$$

Then the function $\varphi_{x_0}(s)$ depends only on $s_{x,a}$ for $x \in X, a \in A(x)$, and on the set S of solutions of the system

$$\begin{cases} \sum\limits_{a\in A(x)} s_{x,a} = 1, \quad \forall x \in X; \\ \\ s_{x,a} \geq 0, \quad \forall x \in X, \ a \in A(x) \end{cases}$$

the function $\varphi(s)$ is monotone.

The proof of this lemma is similar to the proof of Lemma 1.

The continuous game model with m players for discounted Markov decision problem we formulate as follow: on the set $S = S^1 \times S^2 \times \cdots \times S^m$ we consider m payoff functions functions

$$\varphi^i_{x_0}(s^1, s^2, \ \dots \ s^m) = \sigma^i_{x_0}, \ \ i = 1, 2, \dots, \ m,$$

where σ^i_x for $x \in X$ satisfy the conditions

$$\sigma^i_x - \gamma \sum_{y\in X} \sum_{a\in A(x)} s^k_{x,a}\, p^a_{x,y} \sigma^i_y = \sum_{a\in A(x)} s^k_{x,a}\, \mu^i_{x,a},$$
$$\forall x \in X_k; \ i, k = 1, 2, \dots, m;$$

This game model possesses the same property as the previous continuous model:

- The set of Nash equilibria situations of the continuous model is non empty if and only if the set of Nash equilibria situations of the game in positional form is not empty;

- If $(s^1, s^2, \ \dots, \ s^m)$ is an extreme point of S then $\overline{F}^i_x(s^1, s^2, \ \dots, \ s^m) = \varphi(s^1, s^2, \ \dots, \ s^m), \ \ \forall x \in X, \ \ i = 1, 2. \ \dots, \ m$ and all Nash equilibria situations for the continuous game model that correspond to extreme points in S represent Nash equilibria situations for the game in positional form.

From Lemma 4 as a corollary we obtain the following result.

Lemma 5. For an arbitrary discounted Markov decision process each payoff function $\varphi^i_{x_0}(s^1, s^2, \ \dots, \ s^m), \ \ i \in \{i, 2, \ \dots, \ m\}$ possesses the property that $\varphi^i_{x_0}(\overline{s}^1, \overline{s}^2, \ \dots, \ \overline{s}^{i-1}, s^i, \overline{s}^{i+1}, \ \dots, \ \overline{s}^m)$ is monotone with respect to $s^i \in S^i$ for arbitrary fixed $\overline{s}^k \in S^k, \ \ k = 1, 2, \ \dots, \ i-1, \ i+1, \ \dots, \ m$..

Using this lemma we can prove the following theorem.

Theorem 6. Let $(X, \ A, \ \{X_i\}_{i=\overline{1,m}}, \ \{c^i\}_{i=\overline{1,m}}, \ p, \ \gamma, \ x)$ be a stochastic positional game with a given starting position $x \in X$ and discounted payoff functions

$$\overline{F}^1_x(s^1, \, s^2, \, \dots, \, s^m), \ \overline{F}^2_x(s^1, \, s^2, \, \dots, \, s^m), \dots,$$
$$\overline{F}^m_x(s^1, \, s^2, \, \dots, \, s^m)$$

of players $1, 2, \ \dots, \ m$, respectively. Then in the considered game there exists Nash equilibrium $s^* = (s^{1^*}, \, s^{2^*}, \, \dots, \, s^{m*})$. Moreover, in this game there exists a situation $s^* = (s^{1^*}, \, s^{2^*}, \, \dots, \, s^{m*})$ which is a Nash-equilibrium for an arbitrary starting position $x \in X$.

6. CONCLUSION

In this paper a new class of stochastic positional games that extend the well known deterministic and stochastic positional games is studied. A new results concerned with existence of Nash equilibria for the game models of Markov decision problems with average and discounted costs optimization criteria are obtained. Based on these results the problem of determining the optimal stationary strategies of players in the considered games can be reduced to continuous similar problems for which classical numerical methods can be applied. The described results may be useful for elaboration of suitable iteration procedures of determining the optimal stationary strategies in positional games, furthermore we extend the results to general network topological problems. This will be presented in the second part of the presentation.

REFERENCES

P. Dasgupta P, E. Maskin. The existence of Equilibrium in Discontinuous Ecomomic Games. *Review of Economic Studies*, 53, 1986, pp. 1-26.

G. Debreu. A Social Equilibrium Existence Theorem. *Proceedings of the National Academy of Aciences*, pp. 386-393, 1952.

V.A. Gurvich, A.V. Karzanov, L.G. Khachian Cyclic Games and an Algorithm to Find Minimax Cycle Means in Directed Graphs. *USSR, Computational Mathematics and Mathematical Phisics*, 28, pp. 85-91, 1988.

D. Lozovanu, S. Pickl Nash Equilibria Conditions for Cicle Games with p players. *Eletronic Notes in Discrete Mathemathics*, 2006, 25, 2006, pp. 117-124.

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

J.F. Nash Non Cooperative Games. *Annals of Mathematics*, 2, 1950, pp.286-295.

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

F.P. Reny On the existence of Pure and Mixed Strategy Nash Equilibria In Discontinuous Games. *Econometrica*, 67, 1999, pp. 1029-1056.

L. Simon Games with Discontinuous Payoffs. *Review of Economic Studies*, 54, 1987, pp. 569-597.

D.J. White Markov Decision Processes. Wiley, New York, 1993.