Contributions to Game Theory and Management, VIII, 187–198

On Nash Equilibria for Stochastic Games and Determining the Optimal Strategies of the Players

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 consider n -person stochastic games in the sense of Shapley. The main results of the paper are related to the existence of Nash equilibria and determining the optimal stationary strategies of the players in the considered games. We show that a Nash equilibrium for the stochastic game with average payoff functions of the players exists if an arbitrary situation induces an ergodic Markov chain. For the stochastic game with discounted payoff functions we show that a Nash equilibrium always exists. Some approaches for determining Nash equilibria in the considered games are proposed.

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

1. Introduction

In this paper we consider the infinite n-person stochastic games. An n-person stochastic game (Owen, 1982; Neyman and Sorin, 2003; Mertens and Neyman, 1981) is a dynamic game with probabilistic transitions played by players in a sequence of stages, where the beginning of each stage corresponds to a state from a given finite set of states of the game. The game starts at a given state. 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 an infinite number of stages. The total payoff of a player is either the average of the stage payoffs or the discounted sum of the stage payoffs. The considered stochastic games have been studied by Gillette, 1957; Mertens and Neyman, 1981; Filar and Vrieze, 1997; Lal and Sinha, 1992; Neyman and Sorin, 2003. Existence of Nash equilibria for n-person games are proven in the case of stochastic games when the total payoff of each player is the discounted sum of stage payoffs and for some special cases of the games with average payoffs. In the general case, for the game with total payoffs that represents the average of the stage payoffs a Nash equilibrium may not exist (Lozovanu and Pickl, 2014).

The main results we describe in this paper are concerned with the existence of Nash equilibria in the considered games and elaboration of algorithms for determining the optimal stationary strategies of the players. We consider the stationary strategies in the sense of Shapley. The stationary strategy weWe show that a Nash equilibrium for the stochastic game with average payoff functions of the players exists if an arbitrary situation generated by the strategies of the players induces a Markov unichain. For the stochastic game with discounted payoff functions we show that a Nash equilibrim always exists. The obtained results can be easily extended for antagonistic stochastic games and the corresponding conditions for the existence of saddle points can be derived.

2. Formulation of the basic game models

A stochastic game with n 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 $i \in \{1, 2, \dots, n\}$ for an arbitrary state $x \in X$;
3. A stage payoff $f^i(x, a)$ with respect to each player $i \in \{1, 2, \dots, n\}$ 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 $a^i_t \in A^i(x_t), \ \ i = 1, 2, \dots, n$. Then nature selects state $y = x_{t+1}$ according to probability transitions $p^{a_t}_{x_t,y}$ for a fixed action vector $a_t = (a^1_t, a^2_t, \dots, a^n_t)$. 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 $i \in \{1, 2, \dots, n\}$ 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 $i \in \{1, 2, \dots, n\}$ is

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

In a similar the stochastic game with discounted sum payoffs of the players is defined. In such a game along to elements described above also a discount factor λ $(0 < \lambda < 1)$ is given and the t-stage stochastic game with discounted sum payoffs is the game where the payoff of player $i \in \{1, 2, \dots, n\}$ is

$$\sigma^i_t = \sum_{\tau=1}^{t-1} \lambda^\tau f^i_\tau.$$

The infinite stochastic game with discounted payoffs is the game where the payoff of player $i \in \{1, 2, \dots, n\}$ is

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

The considered games can be formulated in terms of *stationary strategies* that correspond to pure strategies of the players. In this case the stationary strategies of the players we define as n maps:

$$s^i : x \to a^i \in A^i(x) \ \text{ for } \ x \in X, \ i = 1, 2, \dots, n.$$

Obviously, the corresponding sets of stationary strategies $S^i, S^2, \dots, S^n$ of the players are finite sets.

Let $s = (s^1, s^2, \dots, s^n)$ be a situation determined by a set of stationary strategies $s^1, s^2, \dots, s^n$ of the players $1, 2, \dots, n$. This situation induces a Markov process with the probability distributions $p_{x,y}^{s(x)}$ in the states $x \in X$, i.e. we obtain the matrix of probability transitions $P^s = (p_{x,y}^s)$. For this process we can determine the matrix of limiting probabilities $Q^s = (q_{x,y}^s)$ that correspond to P^s. Therefore, if the starting state x_0 is given, then for the Markov process with the matrix of probability transitions P^s we can calculate the corresponding average costs per transition $F_{x_0}^1(s^1, s^2, \dots, s^n)$, $F_{x_0}^2(s^1, s^2, \dots, s^n), \dots, F_{x_0}^n(s^1, s^2, \dots, s^n)$ for the players as follows:

$$F_{x_0}^i(s^1, s^2, \dots, s^n) = \sum_{x \in X} q_{x_0,x}^s f^i(x, s^1(x), s^2(x), \dots, s^n(x)), \quad i = 1, 2, \dots, n.$$

In such a way on the set of situations $S = S^1 \times S^2 \times, \dots, \times S^n$ we obtain the functions $F_{x_0}^i(s^1, s^2, \dots, s^n)$, $i = 1, 2, \dots, n$ that define the stochastic game with average payoffs in pure strategies. This game is determined by the set of states X, the sets of actions of the players $\{A^i\}_{i=\overline{1,n}}$, the probability function p, the set of stage payoffs $\{f^i(x,a\}_{i=\overline{1,n}}$, where $a \in A$, $A = \prod_{i=1}^n A^i$, and the starting position of the game x_0. Therefore we denote this game $(X, \{A^i\}_{i=\overline{1,n}}, \{f^i(x,a\}_{i=\overline{1,n}}, p, x_0)$.

We define the stochastic game with a discounted sum of stage payoffs in pure strategies in analogues way if for the Markov process with the matrix of probability transitions $P^s = (p_{x,y}^s)$ we consider the matrix $W^s(\lambda) = (w_{x,y}^s(\lambda))$ where $W^s(\lambda) = (I - \lambda P^s)^{-1}$. Then for a situation $s = (s^1, s^2, \dots, s^n)$ the total discounted sum of stage payoffs $\sigma_{x_0}^i(s^1, s^2, \dots, s^n)$ with given discount factor λ $(0 < \lambda < 1)$ for the players can be calculated as follows:

$$\sigma_{x_0}^i(s^1, s^2, \dots, s^n) = \sum_{y \in X} w_{x_0,x}^s(\lambda) f^i(x, s^1(x), s^2(x), \dots, s^n(x)), \quad i = 1, 2, \dots, n.$$

So, on the set of situations $S = S^1 \times S^2 \times, \dots, \times S^n$ we obtain the functions $\sigma_{x_0}^i(s^1, s^2, \dots, s^n)$, $i = 1, 2, \dots, n$ that define the stochastic game with discounted payoffs in pure strategies. In a similar way as the previous game we can denote the discounted stochastic game in pure strategies $(X, \ \{A^i\}_{i=\overline{1,n}}, \{f^i(x,a\}_{i=\overline{1,n}}, p, \gamma, \ x_0)$.

For these games Nash equilibria in pure strategies may not exist . Therefore in this paper we study the stochastic game using *stationary strategies in the sense of Shapley* (Shapley, 1953) that correspond to mixed strategies. For such games we formulate conditions for the existence of Nash equilibria and describe some approaches for determining the optimal strategies of the players.

3. Determining Nash Equilibria for Stochastic Games with Average Payoffs

We shall use a continuous model for studying the average stochastic games. We construct such a model as follows: At first we identify an arbitrary stationary strategy $s^i : x \to a^i \in A^i(x)$ with the set of boolean variables $s^i_{x,a^i} \in \{0,1\}, x \in X, a^i \in A^i(x)$, where $s^i_{x,a^i} = 1$ if and only if player i fixes the action $a^i \in A^i(x)$ in the state x. So, the set of stationary strategies of player i we regard as the set of solutions of the following system:

$$\sum_{a^i \in A(x)} s^i_{x,a^i} = 1, \quad \forall x \in X; \quad s^i_{x,a^i} \in \{0,1\}, \ \forall x \in X, \forall a^i \in A^i(x).$$

Then in this system we change the $s^i_{x,a^i} \in \{0,1\}$ by the condition $0 \le s^i_{x,a^i} \le 1$ and we obtain the set of stationary strategies in the sense of Shapley (Shapley, 1953), where s^i_{x,a^i} is treated as probability of the choices of the action a^i by player i every time when the state x is reached by any route in the dynamic stochastic game. Additionally, we shall use the following condition for the average stochastic games. We assume that an arbitrary situation $s = (s^1, s^2, \dots, s^n) \in S$ generates a Markov unichain with the corresponding matrix of probability transitions $P^s = (p^s_{x,y})$. We call a game with such a property with respect to the situations $s = (s^1, s^2, \dots, s^n) \in S$ *perfect game* (Lozovanu, 2011). We show that in this case the problem of determining Nash equilibria for a stochastic game can be formulated as a continuous model that represents the game variant of the following optimization problem:
Minimize

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

subject to

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

This problem represents a continuous model for an average Markov decision problem with immediate costs $f_{(x,a)}$ in the states $x \in X$ for given actions $a \in A(x)$ and probability transitions $p^a_{x,y}$, where $\sum_{y \in X} p^a_{x,y} = 1, \forall x \in X, \forall a \in A$. More precisely, problem (1), (2) corresponds to a Markov decision problem where each strategy induces a Markov unichain (see Lozovanu and Pickl, 2015). This is easy to show, if we identify an arbitrary stationary strategy with the set of boolean variables $s_{x,a} \in \{0,1\}, \ x \in X, \ a \in A(x)$ that satisfy the conditions

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

These conditions determine all feasible strategies in (2). The remaining 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 a Markov decision problem with an average cost criterion. If such an optimal solution is known, then an optimal action for a 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 $\alpha_{x,a} = s_{x,a}q_x,\ \forall x \in X, a \in A(x)$ (see Lozovanu and Pickl, 2015). 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 an average Markov decision problem be given, where an arbitrary stationary strategy* s *generates a Markov unichain, and consider the function*

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

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

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

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

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

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

Proof. If an arbitrary strategy s for a Markov decision problem induces a Markov unichain then for such an arbitrary a 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). Moreover, the system of linear equations (3) uniquely determines $q_x, \forall x \in X$ for an arbitrary solution of system (4). So, the function $\psi(s)$ depends only on $s_{x,a}$ for $x \in X, a \in A(x)$,

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

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

We proof the correctness of this property using the relationship of the problem (1),(2) with the following linear programming problem:

Minimize

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

subject to

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

The problem (6),(7) is obtained from (1),(2) introducing the substitutions $\alpha_{x,a} = s_{x,a} q_x$ 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 a 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 $\alpha',\ \alpha''$ 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 $\overline{\alpha} = \theta\alpha' + (1-\theta)\alpha''$, $0 \le \theta \le 1$ 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

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

It is easy to observe that $0 \le \overline{\theta}_x \le 1$, where $\overline{\theta}_x = 0,\ \forall x \in X$ if and only if $\theta = 0$ and $\overline{\theta}_x = 1,\ \forall x \in X$ if and only if $\theta = 1$. Moreover, it can be easily seen from the following proof that $\psi(\overline{s}) = \psi(s')$ in the case $\psi(s') = \psi(s'')$. 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 a stochastic game with average payoffs. We consider the continuous model for perfect stochastic games.

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

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

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

On the set $\overline{S} = \overline{S}^1 \times \overline{S}^2 \times \cdots \times \overline{S}^n$ we define n payoff functions

$$\psi^i(s^1, s^2, \ldots, s^n) = \sum_{x\in X} \sum_{(a^1,a^2,\ldots,a^n)\in A(x)} \prod_{k=1}^{n} s^k_{x,a^k} f^i_{(x,a^1,a^2\ldots a^n)} q_x, \quad i = 1, 2, \ldots, n, \tag{10}$$

where q_x for $x \in X$ are determined uniquely from the following system of linear equations

$$\begin{cases} \sum\limits_{x\in X} \sum\limits_{(a^1,a^2,\ldots,a^n)\in A(x)} \prod\limits_{k=1}^{n} s^k_{x,a^k} p^{(a^1,a^2,\ldots,a^n)}_{x,y} q_x = q_y, \quad \forall y \in X; \\ \sum\limits_{x\in X} q_x = 1 \end{cases} \tag{11}$$

when $s^1, s^2, \ldots, s^m$ 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 stochastic game in pure strategies is not empty;
- If $(s^1, s^2, \ldots, s^m)$ is an extreme point of $\overline{S}$ then $F^i_x(s^1, s^2, \ldots, s^n) = \psi(s^1, s^2, \ldots, s^n)$, $\forall x \in X$, $i = 1, 2. \ldots, n$ and all Nash equilibria situations for the continuous game model that correspond to extreme points in $\overline{S}$ represent Nash equilibria situations for the stochastic game in pure strategies.

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

Lemma 2. *For a perfect stochastic game each payoff function* $\psi^i(s^1, s^2, \ldots, s^n)$, $i \in \{1, 2, \ldots, n\}$ *possesses the property that* $\psi^i(\overline{s}^1, \overline{s}^2, \ldots, \overline{s}^{i-1}, s^i, \overline{s}^{i+1}, \ldots, \overline{s}^n)$ *is monotone with respect to* $s^i \in \overline{S}^i$ *for arbitrary fixed* $\overline{s}^k \in \overline{S}^k$, $k = 1, 2, \ldots, i-1, i+1, \ldots, n$.

Using this lemma we can prove the following theorem.

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

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

of players $1, 2, \ldots, n$, *respectively. If for an arbitrary situation* $s = (s^1, s^2, \ldots, s^n) \in S$ *of the game the transition probability matrix* $P^s = (p^s_{x,y})$ *corresponds to a Markov unichain then for the continuous game on* $\overline{S}$ *there exists a Nash equilibrium* $s^* = (s^{1*}, s^{2*}, \ldots, s^{n*})$ *which is a Nash equilibrium for an arbitrary starting position* $x \in X$ *of the game.*

Proof. According to Lemma 2 each function $\psi^i(s^1, s^2, \ldots, s^n)$, $i \in \{1, 2, \ldots, n\}$ satisfies the condition that $\psi^i(\overline{s}^1, \overline{s}^2, \ldots, \overline{s}^{i-1}, s^i, \overline{s}^{i+1}, \ldots, \overline{s}^n)$ is monotone with respect to $s^i \in \overline{S}^i$ for an arbitrary fixed $\overline{s}^k \in \overline{S}^k$, $k = 1, 2, \ldots, i-1, i+1, \ldots, n$. In the considered game each subset $\overline{S}^i$ is convex and compact. Therefore, these conditions (see Debreu, 1952, Dasgupta and Maskin, 1986, Simon, 1987 and

Reny, 1999) provide the existence of a Nash equilibrium $s^* = (s^{1^*}, s^{2^*}, \dots, s^{n*})$ for the functions $\psi^i(s^1, s^2, \dots, s^n),\ i \in \{i, 2, \dots, n\}$ on $\overline{S}^1 \times \overline{S}^2 \times \cdots \times \overline{S}^n$. This Nash equilibrium is a Nash equilibrium for an arbitrary starting position x of the game.

Corollary 1. *For the average stochastic game there exists a Nash equilibrium in pure strategies if and only if the continuous game has a Nash equilibrium in pure strategies.*

Using the results described above we may conclude that in the case of perfect games a Nash equilibrium for stochastic games with average payoffs can be determined by using classical iterative methods for the continuous game models with payoff functions $\psi^i(s^1, s^2, \dots, s^n),\ i \in \{i, 2, \dots, n\}$ on the set $\overline{S}^1 \times \overline{S}^2 \times \cdots \times \overline{S}^n$. If we refer these iterative methods to a discrete game model with payoff functions $F^1_x(s^1, s^2, \dots, s^n),\ F^2_x(s^1, s^2, \dots, s^n),\ \dots,\ F^m_x(s^1, s^2, \dots, s^n)$ on $S^1 \times S2 \times \cdots \times S^n$, then we obtain the iterative procedures where players fix successively their strategies in order to minimize their payoff functions, respectively, and finally to reach Nash equilibrium (if such an equilibrium exists).

Note that if a stochastic game is not perfect, then Nah equilibrium may not exist (Lozovanu and Pickl, 2014, 2015)

4. Determining Nash Equilibria for Stochastic Games with Discounted Payoffs

In this section we show that a Nash equilibrium in mixed strategies exists for an arbitrary stochastic game with discounted payoff functions of the players and given discount factor $\gamma,\ 0 < \gamma < 1$. To prove this we shall use the continuous model for the stochastic game. We will formulate such a model using the following auxiliary optimization problem:
Maximize

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

subject to

$$\sigma_x - \gamma \sum_{y \in X} \sum_{a \in A(x)} s_{x,a}\, p^a_{x,y}\, \sigma_y = \sum_{a \in A(x)} s_{x,a}\, f_{(x,a)}, \quad \forall x \in X, \tag{13}$$

where $s_{x,a}, x \in X, a \in A(x)$ correspond to a fixed strategy that satisfy (1).

This problem represents a continuous model for the discounted Markov decision problem (see Lozovanu and Pickl, 2015). The system of linear equations (13) with respect to σ_x has a unique solution for a fixed s and we can find all σ_x for $x \in X$ that represent the discounted sum of immediate costs in the decision problem with the corresponding starting positions $x \in X$. It is easy to observe that if we consider the optimization problem (12), (13) with respect to σ then the equations in (13) can be changed by inequalities ($\leq$) and the values of the optimal solutions of the problem (12), (13) will correspond to the same σ_x for $x \in X$. Therefore if after that we dualize (12), (13) with respect to σ_x for fixed s then we obtain the following linear programming problem :
Minimize

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

subject to

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

If we add to this system the condition (1) and will minimize with respect to s then we obtain the following optimization problem:
Minimize

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

subject to

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

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

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

subject to

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

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 $\alpha^*_{x,a} = s^*_{x,a}\beta^*_x, \ \forall x \in X, a \in A(x)$ (see Lozovanu and Pickl, 2015).

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

Lemma 3. *Let a discounted Markov decision problem with the discount 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*

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

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* $\overline{S}$ *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 $\varphi_{x_0}(s)$ *is monotone.*

The proof of this lemma is similar to the proof of Lemma 1 if instead of the linear programming formulation (6), (7) we shall use the linear programming formulation (16), (17).

We formulate the continuous model for the stochastic game with discounted payoffs as follows: On the set $\overline{S}=\overline{S}^1\times\overline{S}^2\times\cdots\times\overline{S}^n$ we consider n payoff functions

$$\varphi^i_{x_0}(s^1,s^2,\ \dots\ s^n)=\sigma^i_{x_0},\quad i=1,2,\dots,\ n, \tag{20}$$

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

$$\begin{aligned}\sigma^i_x-\gamma\sum_{y\in X}\ \sum_{(a^1,a^2,\dots,a^n)\in A(x)}\ \prod_{k=1}^n\ s^k_{x,a^k}p^{(a^1,a^2,\dots,a^n)}_{x,y}\sigma^i_y=\\ =\sum_{(a^1,a^2,\dots,a^n)\in A(x)}\ \prod_{k=1}^n\ s^k_{x,a^k}f^i_{(x,a^1,a^2\dots a^n)},\quad \forall x\in X;\ i=1,2,\dots,n.\end{aligned} \tag{21}$$

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 stochastic game in pure strategies is not empty;

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

From Lemma 3 as a corollary we obtain the following result.

Lemma 4. *For an arbitrary stochastic game with discounted payoffs each payoff function* $\varphi^i_{x_0}(s^1,s^2,\ \dots,\ s^n),\quad i\in\{1,2,\ \dots,\ n\}$ *possesses the property that* $\varphi^i_{x_0}(\overline{s}^1,\overline{s}^2,\ \dots,\ \overline{s}^{i-1},s^i,\overline{s}^{i+1},\ \dots,\ \overline{s}^n)$ *is monotone with respect to* $s^i\in\overline{S}^i$ *for arbitrary fixed* $\overline{s}^k\in\overline{S}^k,\quad k=1,2,\ \dots,\ i-1,\ i+1,\ \dots,\ n.$

Using this lemma we can prove the following theorem.

Theorem 2. *Let a stochastic game* $(X, \{A^i\}_{i=\overline{1,n}}, \{f^i(x,a\}_{i=\overline{1,n}}, p, \gamma, x_0)$ *with the starting position* $x \in X$ *and discounted payoff functions*

$$\sigma_x^1(s^1, s^2, \dots, s^n),\ \sigma_x^2(s^1, s^2, \dots, s^m),\ \dots,\ \sigma_x^n(s^1, s^2, \dots, s^n)$$

of the players $1, 2, \dots, n$, *be given. Then in the considered game there exists a Nash equilibrium* $s^* = (s^{1^*}, s^{2^*}, \dots, s^{n*})$ *on* $\overline{S}$ *which is a Nash equilibrium for an arbitrary starting position* $x \in X$.

The proof of this theorem is similar to the proof of Theorem 1, i.e. the existence of Nash equilibria for the continuous game with payoff functions $\varphi_{x_0}^i(s^1, s^2, \dots, s^n)$, $i \in \{1, 2, \dots, n\}$ on $\overline{S}$ can be gained in a analogues way as for the game with average payoffs if we apply Lemma 4 and the corresponding results from (Debreu, 1952, Dasgupta and Maskin, 1986, Simon, 1987 and Reny, 1999).

5. Conclusion

The considered n-person stochastic games can be studied using the continuous game models. Based on the proposed approach new Nash equilibria conditions for the games with average and discounted payoffs have been derived and some approaches for determining the optimal stationary strategies of the players are proposed. The obtained results can be extended for the antagonistic stochastic games.

References

Dasgupta, P, Maskin, E. (1986). *The existence of Equilibrium in Discontinuous Economic Games.* Review of Economic Studies, **53**, 1–26.

Debreu, G. (1952). *A Social Equilibrium Existence Theorem*, Proceedings of the National Academy of Aciences, 386–393.

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.

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. (2014). *Nash equilibria conditions for stochastic positional games.* Contribution to Game Theory and Management, **VII**, Saint. Petersburg State University, **10**, 201–213.

Lozovanu, D., Pickl, S. (2015). *Optimization of Stochastic Discrete Systems and Control on Complex Networks.* Springer.

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

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

Owen, G. (1982). *Game Theory, 2nd edition*, Academic Press, New York.

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

Reny, F. (1999). *On the existence of Pure and Mixed Strategy Nash Equilibria In Discontinuous Games.* Econometrica, **67**, 1029–1056.

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

Simon, L. (1987). *Games with Discontinuous Payoffs.* Review of Economic Studies, **54**, 569–597.