Saddle Point Conditions for Antagonistic Positional Games in Complex Markov Decision Processes

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

[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
A class of stochastic antagonistic positional games for Markov decision processes with average and expected total discounted costs optimization criteria are formulated and studied. Saddle point conditions in the considered class of games that extend saddle point conditions for deterministic parity games are derived. Furthermore algorithms for determining the optimal stationary strategies of the players are proposed and grounded.

Keywords : Markov Decision Processes, Stochastic Positional Games, Antagonistic Positional Games, Saddle Point, Optimal Stationary Strategies
MSC: 65K05, 68W25 .

1 Introduction and Problem Formulation

Markov decision processes have a prominent role within the theory of dynamic games and anticipatory systems. In this paper we study a class of antagonistic positional games for finite state space Markov decision processes with average and expected total cost optimization criteria. We formulate saddle point conditions for this class of games that extend saddle point conditions for deterministic parity games from [1, 4, 5, 7, 9]. Thus, the considered class of games and the presented results generalize deterministic antagonistic positional games and saddle point conditions for such games. Moreover we show that for the considered class of games saddle points always exists and the optimal stationary strategies of the players can be found using an efficient finite iterative procedure.

International Journal of Computing Anticipatory Systems, Volume 30, 2014
Edited by D. M. Dubois, CHAOS, Liège, Belgium, ISSN 1373-5411 ISBN 2-930396-19-9

The formulated basic game models are using 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 for the dynamical system when it makes a transition from a state $x \in X$ to another state $y \in X$.

We assume that the Markov decision process is controlled by two players as follows: The set of states X is divided into two disjoint subsets X_1 and X_2, $X = X_1 \cup X_2$ $(X_1 \cap X_2 = \emptyset)$, where X_1 represents the positions set of the first player and X_2 represents the position set of the second player. Each player fixes actions in his positions, i.e. if the dynamic system at the given moment of time is in the state x which belongs to the position set of first player then the action $a \in A$ is fixed by the first player; otherwise the action is fixed by the second player. The player fixes actions in their position sets using stationary strategies. The stationary strategies of the players are defined as two maps:

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

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

where $A^1(x)$ is the set of actions of the first player in the state $x \in X_1$ and $A^2(x)$ is the set of actions of the second player in the state $x \in X_2$. Without loss of generality we may consider $|A^i(x)| = |A^i| = |A|$, $\forall x \in X_i$, $i = 1, 2$. In order to simplify the notation we denote the set of possible actions in a state $x \in X$ for an arbitrary player by $A(x)$.

If players fix their stationary strategies s_1 and s_2, respectively, then we obtain a situation $s = (s_1, s_2)$. 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 $i = 1, 2$. 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 and the matrix of transition costs $C = (c_{x,y})$ we can determine the average cost per transition $\omega_{x_0}(s_1, s_2)$ that corresponds to the situation $s = (s_1, s_2)$.

So, on the set of situations $S^1 \times S^2$ we can define the payoff function

$$F_{x_0}(s_1, s_2) = \omega_{x_0}(s_1, s_2).$$

In such a way we obtain an antagonistic positional game which is determined by the corresponding finite sets of strategies S_1, S_2 of the players and the payoff function $F_{x_0}(s_1, s_2)$ defined on $S = S_1 \times S_2$.

This game is determined uniquely by the set of states X, the positions sets X_1, X_2, the set of actions A, the cost function $c : X \times X \to R$, the probability function $p : X \times X \times A \to [0,1]$ and the starting position x_0. Therefore we denote it by $(X,\ A,\ X_1,\ X_2,\ c,\ p,\ x_0)$ and call this game *stochastic antagonistic positional game with average payoff function.* We show that for the players in the considered game there exist the optimal stationary strategies, i.e. we show that there exist the strategies $s_1{}^* \in S_1, s_2{}^* \in S_2$ that satisfy the condition

$$F_x(s_1{}^*, s_2{}^*) = \max_{s_1 \in S_1} \min_{s_2 \in S_2} F_x(s_1, s_2) = \min_{s_2 \in S_2} \max_{s_1 \in S_1} F_x(s_1, s_2),\ \forall x \in X. \tag{1}$$

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 parity game studied in [1, 4, 5, 9]. In [2, 3] the stochastic positional game of m players where some sufficient conditions for the existence of Nash equilibria are derived have been formulated. However, for a stochastic positional game with m players in the general case Nash equilibria may not exist. The main result we present in this paper shows that there always exists a saddle point for stochastic antagonistic positional games with average payoff functions.

In this paper we consider additionally the *stochastic antagonistic positional game with a discounted payoff function.* We define this special game in a similar way as the game above. We consider a Markov decision processes (X, A, c, p) that may be controlled by two players with the corresponding position sets X_1 and X_2, where the players fix actions in their position sets using stationary strategies. Here, we assume that for the Markov decision process the discount factor λ, $0 < \lambda < 1$ is given, where the cost of system's transition from a state $x \in X$ to a state y at the moment of time t is discounted with the rate λ^t, i.e. the cost of system's transition from the state x at the moment of time t to the state y at the moment of time $t+1$ is equal to $\lambda^t c_{x,y}$. For fixed stationary strategies s_1, s_2 of the players we obtain a situation $s = (s_1, s_2)$ that determines a simple Markov process with the transition probability matric $P^s = (p^s_{x,y})$ and the matrix of transition costs $C = (c_{x,y})$. Therefore if the starting state x_0 is known then we can determine the expected total discounted cost $\sigma^\lambda_{x_0}(s_1, s_2)$ that corresponds to the situation $s = (s_1, s_2)$. So, on the set of situations $S_1 \times S_2$ we can define the payoff function

$$\overline{F}_{x_0}(s_1, s_2) = \sigma^\lambda_{x_0}(s_1, s_2).$$

We denote the stochastic antagonistic positional game with a discounted payoff function by $(X,\ A,\ X_1,\ X_2,\ c,\ p, \lambda,\ x_0)$. We show that for the players in this game there exist the optimal stationary strategies, i.e. there exist the strategies $s_1{}^* \in S_1, s_2{}^* \in S_2$ that satisfy the condition

$$\overline{F}_{x_0}(s_1{}^*,\ s_2{}^*) = \max_{s_1 \in S_1} \min_{s_2 \in S_2} \overline{F}_{x_0}(s_1,\ s_2) = \min_{s_2 \in S_2} \max_{s_1 \in S_1} \overline{F}_{x_0}(s_1,\ s_2).$$

Some approaches of determining the saddle points in the considered games are described.

2 Determining the Saddle Points for Stochastic Antagonistic Positional Games with an Average Payoff Function

In this section we show for an arbitrary stochastic antagonistic positional game $(X,\ A,\ X_1,\ X_2,\ c, p,\ x)$ with an average payoff function $F_x(s_1, s_2)$ that there exists saddle points, i.e we show that there exists the stationary strategies $s_1{}^*, s_2{}^*$ for which condition (1) holds. This fact is proved using the properties of the bias equations for so called Markov multi-chains [6].

For an arbitrary state $x \in X$ and a fixed action $a \in A(x)$ we denote by

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

i.e. $\mu_{x,a}$ is the immediate cost in the state $x \in X$ for a given action $a \in A(x)$.

Theeorem 1. Let $(X,\ A,\ X_1, X_2, c,\ p, \overline{x})$ be an arbitrary stochastic positional game with an average payoff function $F_{\overline{x}}(s_1, s_2)$. Then the system of equations

$$\begin{cases} \varepsilon_x + \omega_x = \max\limits_{a \in A(x)} \left\{ \mu_{x,a} + \sum\limits_{y \in X} p^a_{x,y} \varepsilon_y \right\}, & \forall x \in X_1; \\ \\ \varepsilon_x + \omega_x = \min\limits_{a \in A(x)} \left\{ \mu_{x,a} + \sum\limits_{y \in X} p^a_{x,y} \varepsilon_y \right\}, & \forall x \in X_2; \end{cases} \tag{2}$$

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

$$\begin{cases} \omega_x = \max\limits_{a \in A(x)} \left\{ \sum\limits_{y \in X} p^a_{x,y} \omega_x \right\}, & \forall x \in X_1; \\ \\ \omega_x = \min\limits_{a \in A(x)} \left\{ \sum\limits_{y \in X} p^a_{x,y} \omega_x \right\}, & \forall x \in X_2, \end{cases} \tag{3}$$

i.e. the system of equations *(3)* 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

$$\begin{cases} \varepsilon_x + \omega^*_x = \max\limits_{a \in A(x)} \left\{ \mu_{x,a} + \sum\limits_{y \in X} p^a_{x,y} \varepsilon_y \right\}, & \forall x \in X_1; \\ \\ \varepsilon_x + \omega^*_x = \min\limits_{a \in A(x)} \left\{ \mu_{x,a} + \sum\limits_{y \in X} p^a_{x,y} \varepsilon_y \right\}, & \forall x \in X_2. \end{cases} \tag{4}$$

The optimal stationary strategies of the players

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

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

in the 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

$$s_1{}^*(x) \in \Big(\arg\max_{a\in A(x)} \Big\{ \sum_{y\in X} p^a_{x,y}\omega^*_x \Big\}\Big) \cap \Big(\arg\max_{a\in A(x)} \Big\{ \mu_{x,a} + \sum_{y\in X} p^a_{x,y}\varepsilon^*_y \Big\}\Big), \quad \forall x \in X_1 \tag{5}$$

and

$$s_2{}^*(x) \in \Big(\arg\min_{a\in A(x)} \Big\{ \sum_{y\in X} p^a_{x,y}\omega^*_x \Big\}\Big) \cap \Big(\arg\min_{a\in A(x)} \Big\{ \mu_{x,a} + \sum_{y\in X} p^a_{x,y}\varepsilon^*_y \Big\}\Big). \quad \forall x \in X_2 \tag{6}$$

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 *(1)* holds.

Proof. Let $x \in X$ be an arbitrary state and consider the stationary strategies $\overline{s}_1 \in S_1$, $\overline{s}_2 \in S_2$ for which

$$F_x(\overline{s}_1, \overline{s}_2) = \min_{s_2\in S_2} \max_{s_1\in S_1} F_x(s_1, s_2).$$

We show that

$$F_x(\overline{s}_1, \overline{s}_2) = \max_{s_1\in S_1} \min_{s_2\in S_2} F_x(s_1, s_2),$$

i.e we show that (1) holds and $\overline{s}_1 = s_1{}^*$, $\overline{s}_2 = s_2{}^*$.

According to the properties of the bias equations from [6] for the situation $\overline{s} = (\overline{s}_1, \overline{s}_2)$ the system of linear equations

$$\begin{cases} \varepsilon_x + \omega_x = \mu_{x,a} + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y, & \forall x \in X_1,\ a = \overline{s}^1(x); \\ \varepsilon_x + \omega_x = \mu_{x,a} + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y, & \forall x \in X_2,\ a = \overline{s}^2(x); \\ \omega_x = \sum\limits_{y\in X} p^a_{x,y}\omega_x, & \forall x \in X_1,\ a = \overline{s}^1(x); \\ \omega_x = \sum\limits_{y\in X} p^a_{x,y}\omega_x, & \forall x \in X_2,\ a = \overline{s}^2(x) \end{cases} \tag{7}$$

has the solution ε^*_x, ω^*_x $(x \in X)$ which for a fixed strategy $\overline{s}_2 \in S_2$ satisfies the condition

$$\begin{cases} \varepsilon_x^* + \omega_x^* \geq \mu_{x,a} + \sum\limits_{y \in X} p_{x,y}^a \varepsilon_y^*, & \forall x \in X_1,\ a \in A(x); \\ \varepsilon_x^* + \omega_x^* = \mu_{x,a} + \sum\limits_{y \in X} p_{x,y}^a \varepsilon_y^*, & \forall x \in X_2,\ a = \overline{s}_2(x); \\ \omega_x^* \geq \sum\limits_{y \in X} p_{x,y}^a \omega_x^*, & \forall x \in X_1,\ a \in A(x); \\ \omega_x^* = \sum\limits_{y \in X} p_{x,y}^a \omega_x^*, & \forall x \in X_2,\ a = \overline{s}_2(x) \end{cases}$$

and $F_x(\overline{s}_1, \overline{s}_2) = \omega_x^*,\ \forall\ x \in X$.

Taking into account that $F_x(\overline{s}_1, \overline{s}_2) = \min_{s^2 \in S^2} F_x(\overline{s}_1, s_2)$ then for a fixed strategy $\overline{s}_1 \in S_1$ the solution $\varepsilon_x^*,\ \omega_x^*\ (x \in X)$ satisfies the condition

$$\begin{cases} \varepsilon_x^* + \omega_x^* = \mu_{x,a} + \sum\limits_{y \in X} p_{x,y}^a \varepsilon_y^*, & \forall x \in X_1,\ a = \overline{s}^1(x); \\ \varepsilon_x^* + \omega_x^* \leq \mu_{x,a} + \sum\limits_{y \in X} p_{x,y}^a \varepsilon_y^*, & \forall x \in X_2,\ a \in A(x); \\ \omega_x^* = \sum\limits_{y \in X} p_{x,y}^a \omega_x^*, & \forall x \in X_1,\ a = \overline{s}^1(x); \\ \omega_x^* \leq \sum\limits_{y \in X} p_{x,y}^a \omega_x^*, & \forall x \in X_2,\ a \in A(x). \end{cases}$$

So, the following system

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

has a solution, which satisfies the condition (7). This means that $s_1^* = \overline{s}_1,\ s_1^* = \overline{s}_1$ and

$$\max_{s_1 \in S_1} \min_{s_2 \in S_2} F_x(s_1, s_2) = \min_{s_2 \in S^2} \max_{s_1 \in S_1} F_x(s_1, s_2),\ \forall x \in X,$$

i.e. the theorem holds.

Thus, the optimal stationary strategies s_1^*, s_2^* of the players for an average antagonistic positional game can be found using the solutions of the system of equations (2)-(4) and conditions (5),(6). Below we describe an iterative algorithm for determining the optimal strategies s_1^*, s_2^*.

3 An Algorithm for Determining the Optimal Stationary Strategies for Stochastic Positional Games with an Average Payoff Function

Preliminary step (Step 0): Fix the arbitrary stationary strategies

$$s_1^0 : x \to y \in X(x) \text{ for } x \in X_1.$$

$$s_2^0 : x \to y \in X(x) \text{ for } x \in X_2.$$

that determine the situation $s^0 = (s_1^0, s_2^0)$.

General step (Step k, $k \geq 1$): Determine the matrix $P^{s^{k-1}}$ that corresponds to the situation $s = (s_1^{k-1}, s_2^{k-1})$ and find $\omega^{s^{k-1}}$ and $\varepsilon^{s^{k-1}}$ which satisfy the conditions

$$\begin{cases} (P^{s^{k-1}} - I)\omega^{s^{k-1}} = 0; \\ \mu^{s^{k-1}} + (P^{s^{k-1}} - I)\varepsilon^{s^{k-1}} - \omega^{s^{k-1}} = 0. \end{cases}$$

Then find a situation $s^k = (s_1^k, s_2^k)$ such that

$$s_1^k(x) \in \arg\max_{a \in A(x)} \Big\{ \sum_{y \in X} p_{x,y}^a \omega_x^{s_1^{k-1}} \Big\}, \qquad \forall x \in X_1;$$

$$s_2^k(x) \in \arg\max_{a \in A(x)} \Big\{ \sum_{y \in X} p_{x,y}^a \omega_x^{s_2^{k-1}} \Big\}, \qquad \forall x \in X_2$$

and set $s^k = s^{k-1}$ if

$$s_1^{k-1}(x) \in \arg\max_{a \in A(x)} \Big\{ \sum_{y \in X} p_{x,y}^a \omega_x^{s_1^{k-1}} \Big\}, \qquad \forall x \in X_1;$$

$$s_2^{k-1}(x) \in \arg\max_{a \in A(x)} \Big\{ \sum_{y \in X} p_{x,y}^a \omega_x^{s_2^{k-1}} \Big\}, \qquad \forall x \in X_2.$$

After that check if $s^k = s^{k-1}$? If $s^k = s^{k-1}$ then go to next step $k+1$; otherwise choose a situation $s^k = (s_1^k, s_2^k)$ such that

$$s_1^k(x) \in \arg\max_{a \in A(x)} \Big\{ \mu_{x,a} + \sum_{y \in X} p_{x,y}^a \varepsilon_y^{s_1^{k-1}(x)} \Big\} \qquad \forall x \in X_1;$$

$$s_2^k(x) \in \arg\max_{a \in A(x)} \Big\{ \mu_{x,a} + \sum_{y \in X} p_{x,y}^a \varepsilon_y^{s_1^{k-1}(x)} \Big\} \qquad \forall x \in X_2$$

and set $s^k = s^{k-1}$ if

$$s_1^{k-1}(x) \in \arg\max_{a \in A(x)} \Big\{ \mu_{x,a} + \sum_{y \in X} p_{x,y}^a \varepsilon_y^{s_1^{k-1}(x)} \Big\} \qquad \forall x \in X_1;$$

$$s_2^{k-1}(x) \in \arg\max_{a\in A(x)} \left\{ \mu_{x,a} + \sum_{y\in X} p_{x,y}^a \varepsilon_y^{s_1^{k-1}(x)} \right\} \qquad \forall x \in X_2.$$

After that check if $s^k = s^{k-1}$? If $s^k = s^{k-1}$ then STOP and set $s^* = s^{k-1}$; otherwise go to next step $k+1$.

Remark. If $p_{x,y} \in \{0.1\}, \forall x, y \in X$ then the algorithm is transformed for determining the optimal stationary strategies of the players in the deterministic parity games.

The convergence of the algorithm described above can be determined in a similar way as the convergence of the iterative algorithm for determining the optimal solution of the Markov decision problem with an average cost optimization criteria (see [6, 8]).

4 Determining the Saddle Points for Stochastic Antagonistic Games with a Discounted Payoff Function

We consider an arbitrary stochastic antagonistic positional game $(X,\ A,\ X_1,\ X_2,\ c,\ p, \lambda,\ x_0)$ and show that saddle points always exist. This result follows as particulary case from results from [2, 3] where the stochastic positional game of m players with discounted payoff functions have been formulated and studied. In [2, 3] it is shown that in such games Nash equilibria always exists. Therefore on the bases of these results we can be proven the following theorem.

Theorem 2. Let $(X,\ A,\ X_1,\ X_2,\ c\, p,\ \lambda,\ \overline{x})$ be an arbitrary stochastic antagonistic positional game with a discounted payoff function $\widehat{F}_{\overline{x}}(s_1, s_2)$. Then there exist the values σ_x for $x \in X$ that satisfy the conditions:

1) $\max\limits_{a\in A(x)} \left\{ \mu_{x,a} + \lambda \sum\limits_{y\in X} p_{x,y}^a \sigma_y - \sigma_x \right\} = 0, \qquad \forall x \in X_1;$

2) $\min\limits_{a\in A(x)} \left\{ \mu_{x,a} + \lambda \sum\limits_{y\in X} p_{x,y}^a \sigma_y - \sigma_x \right\} = 0, \qquad \forall x \in X_2.$

The optimal stationary strategies s^{1^*}, s^{2^*} of the players in the game can be found by fixing the maps

$$s^{1^*}(x) = a^* \in \arg\max_{a\in A(x)} \left\{ \mu_{x,a} + \lambda \sum_{y\in X} p_{x,y}^a \sigma_y - \sigma_x \right\}, \ \ \forall x \in X_1;$$

$$s^{2^*}(x) = a^* \in \arg\min_{a\in A(x)} \left\{ \mu_{x,a} + \lambda \sum_{y\in X} p_{x,y}^a \sigma_y - \sigma_x \right\}, \ \forall x \in X_2,$$

where

$$\widehat{F}_{\overline{x}}(s_1^*, s_2^*) = \sigma_{\overline{x}}, \ \forall \overline{x} \in X.$$

Based on this theorem we can propose the following algorithm for determining the optimal stationary strategies of the players.

5 An Algorithm for Determining the Optimal Stationary Strategies for Stochastic Positional Games with a Discounted Payoff Function

Preliminary step (Step 0): Fix the arbitrary stationary strategies

$$s_1^0 : x_i \to a \in A(x_i) \text{ for } x_i \in X_1;$$

$$s_1^0 : x_i \to a \in A(x_i) \text{ for } x_i \in X_1.$$

and determine the situation $s^0 = (s_1^0, s_2^0)$.

General step (Step k, $k > 0$): Calculate

$$\mu_{x_i, s^{k-1}} = \sum_{y \in X(x_i)} p_{x_i,y}^{s^{k-1}} c_{x_i,y}^{s^{k-1}}$$

for every $x_i \in X$. Then solve the system of linear equations

$$\sigma_{x_i} = \mu_{x_i, s^{k-1}(x_i)} + \lambda \sum_{x_j \in X} p_{x_i,x_j}^{s^{k-1}(x_i)} \sigma_{x_j}, \quad i = 1, 2, \dots, n$$

and find the solution $\sigma_{x_1}^{k-1}, \sigma_{x_2}^{k-1}, \dots, \sigma_{x_n}^{k-1}$. After that determine the new strategies

$$s_1^k : x_i \to a \in A(x_i) \text{ for } x_i \in X_1,$$

$$s_2^k : x_i \to a \in A(x_i) \text{ for } x_i \in X_2,$$

and the corresponding situation $s^k = (s_1^k, s_2^k)$, where

$$s_1^k(x_i) = \arg \max_{a \in A(x_i)} \left[\mu_{x_i,a} + \lambda \sum_{x_j \in X} p_{x_i,x_j}^a \sigma_{x_i}^{k-1} \right] \text{ for } x_i \in X_1;$$

$$s_2^k(x_i) = \arg \min_{a \in A(x_i)} \left[\mu_{x_i,a} + \lambda \sum_{x_j \in X} p_{x_i,x_j}^a \sigma_{x_i}^{k-1} \right] \text{ for } x_i \in X_2.$$

Check if the following conditions hold

$$\begin{aligned} &s_1^k(x_i) = s_1^{k-1}(x_i), \quad \forall x_i \in X_1; \\ &s_1^k(x_i) = s_1^{k-1}(x_i), \quad \forall x_i \in X_2? \end{aligned} \tag{8}$$

If the condition (8) holds then fix

$$s_1^* = s_1^k; \quad \sigma_{x_i}^* = \sigma_{x_i}^k, \quad \forall x_i \in X_1;$$

$$s_2^* = s_1^k; \quad \sigma_{x_i}^* = \sigma_{x_i}^k, \quad \forall x_i \in X_2;$$

otherwise go to the next step $k+1$. The strategies s_1^* and s_2^* represent the optimal stationary strategies of the players in the game.

The convergence of this algorithm can be grounded in a similar way as the convergence of the iterative algorithm for determining the optimal solution of the Markov decision problem with a discounted optimization criteria (see [6, 8]).

6 Determining Saddle Points for Stochastic Antagonistic Positional Games with Stopping States

The stochastic antagonistic positional game model with a discounted payoff function can be modified if we assume that the dynamical system in the Markov process stops transitions as soon as a given state $z \in X$ is reached. Thus, we may assume that z is an absorbing state and the cost $c_{z,z}$ is equal to zero. It is easy to observe that if λ satisfies the condition $0 < \lambda < 1$ then for Markov processes with an absorbing state $z \in X$ with $c_{z,z} = 0$ the saddle point condition and the algorithm for determining the optimal stationary strategies of the players from Section 3 are valid for the considered game. If $\lambda = 1$ then the results from Section 3 can be used for determining the optimal stationary strategies of the players in the antagonistic positional games for Markov decision processes with stopping states only for a special class of games. Note that in the case $p_{x,y} \in \{0, 1\}, \forall x, y \in X$ and $\lambda = 1$ the considered class of games is transformed into finite dynamic c-games studied in [4, 5]. Therefore a saddle point condition for finite dynamic c-games can be derived from the results from Section 3. Some algorithms for determining the optimal stationary strategies in dynamic c-games are proposed and grounded in [4].

7 Conclusions

As Markov decision processes have a prominent role within the theory of dynamic games and anticipatory systems we have developed a new characterization which might support modern decision support systems. The considered stochastic positional games generalize deterministic parity cames and dynamic c-games. For antagonistic positional games in Markov decision processes the saddle points always exists and the optimal stationary strategies in such games can be found using efficient iterative calculation procedures.

Acknowledgements

The authors are grateful to the referees for their useful remarks that improve the quality of the paper.

References

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

[2] 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.

[3] Lozovanu D., Pickl S., Kropat E. (2011) Markov decision processes and determining Nash equilibria for stochastic positional games, Proceedings of 18th World Cogress IFAC-2011, 320–327.

[4] Lozovanu D., Pickl S. (2009) Optimization and Multiobjective Control of Time-Discrete Systems. Springer Verlag.

[5] Lozivanu D., Pickl S. (2009) Discrete control and algorithms for solving antagonistic dynamic games on networks. Optimization, 58(6), 665–683.

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

[7] Voge J., Jurdzinsky M. (2000) A discrete strategy improvement algorithm for solving parity games. In Emerson EA and Sistla AP (eds), Computer Aided Verification, 12th International Conference, volume 1855 of LNCS, Springer, 202–215.

[8] White D.J. (1993) Markov Decision Processes. Wiley, New York.

[9] Zwik U., Paterson M. (1996) The complexity of mean payoff games on graphs. TCS, 158, 344–359.