Stochastic and Differential Games
Theory and Numerical Methods

Martino Bardi
T.E.S. Raghavan
T. Parthasarathy
Editors

Foreword by Tamer Başar

Birkhäuser
Boston • Basel • Berlin
Annals of the International Society of Dynamic Games
Volume 4

Series Editor
Tamer Başar

Editorial Board

Leonard D. Berkovitz, Purdue University
Pierre Bernhard, INRIA, Sophia-Antipolis
Raimo P. Hämäläinen, Helsinki University of Technology
Alain Haurie, University of Geneva
N.N. Krasovskii, Academy of Sciences, Ekaterinburg
George Leitmann, University of California, Berkeley
Geert Jan Olsder, Delft University of Technology
T.E.S. Raghavan, University of Illinois, Chicago
Josef Shinar, Technion—Israel Institute of Technology
B. Tolwinski, Operations Research Experts, Black Hawk, Colorado
Klaus H. Well, Stuttgart University
Stochastic and differential games: theory and numerical methods /
[edited by] Martino Bardi ... [et al.].
   p. cm. — (Annals of the International Society of Dynamic
   Games ; v. 4)
   Includes bibliographical references.
   ISBN 0-8176-4029-0 (hc : Alk. paper)
   (Martino) II. Series.
   QA272. S76 1998
   519.3—dc21
   98-28498
   CIP

AMS Subject Classifications: 49-02

Printed on acid-free paper.
© 1999 Birkhäuser Boston

All rights reserved. This work may not be translated or copied in whole or in part without the written
permission of the publisher (Birkhäuser Boston, c/o Springer-Verlag New York, Inc., 175 Fifth Avenue,
New York, NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis.
Use in connection with any form of information storage and retrieval, electronic adaptation, computer
software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the
former are not especially identified, is not to be taken as a sign that such names, as understood by the
Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

ISBN 0-8176-4029-0
ISBN 3-7643-4029-0

Formatted from the authors' disks by TpXniques, Inc., Cambridge, MA.
Printed and bound by Braun-Brunmfeld, Inc., Ann Arbor, MI.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1
Contents

Foreword and Dedication ........................................... vii

Tamer Başar

Preface ................................................................. ix

Contributors ......................................................... xv

Part I. Zero-Sum Differential Games and
Numerical Methods ................................................. 1

1 Constructive Theory of Positional Differential Games
and Generalized Solutions to Hamilton–Jacobi Equations 3
Andrei I. Subbotin

2 Two-Player, Zero-Sum Differential Games and
Viscosity Solutions ............................................... 69
Panagiotis E. Souganidis

3 Numerical Methods for Pursuit-Evasion Games
via Viscosity Solutions .......................................... 105
Martino Bardi, Maurizio Falcone, and Pierpaolo Soravia

4 Set-Valued Numerical Analysis for Optimal Control
and Differential Games ........................................... 177
Pierre Cardaliaguet, Marc Quincampoix, and Patrick Saint-Pierre

Part II. Stochastic and Nonzero-Sum Games
and Applications .................................................... 249

5 An Introduction to Gambling Theory
and Its Applications to Stochastic Games .................... 251
A. Maitra and W. Sudderth

6 Discounted Stochastic Games: A Complex
Analytic Perspective ............................................... 271
S.A. Connell, W.W. Szczechla, Jerzy A. Filar, and O.J. Vrieze
7 Nonzero-Sum Stochastic Games  ....  297
   Andrzej S. Nowak and Krzysztof Szajowski

8 The Power of Threats in Stochastic Games  ....  343
   Frank Thuijsman and O.J. Vrieze

9 A Markov Game Approach for Optimal Routing
   Into a Queuing Network  ....  359
   Eitan Altman

10 On Linear Complementarity and
   A Discounted Polystochastic Game  ....  377
   S.R. Mohan, S.K. Neogy, and T. Parthasarathy
Nonzero-Sum Stochastic Games

Andrzej S. Nowak and Krzysztof Szajowski

Abstract

This chapter discusses stochastic games. We focus on nonzero-sum games and provide a detailed survey of selected recent results. In Section 1, we consider stochastic Markov games. A correlation of strategies of the players, involving "public signals," is described, and a correlated equilibrium theorem proved recently by Nowak and Raghavan for discounted stochastic games with general state space is presented. We also report an extension of this result to a class of undiscounted stochastic games, satisfying some uniform ergodicity condition. Stopping games are related to stochastic Markov games. In Section 2, we describe a version of Dynkin's game related to observation of a Markov process with random assignment mechanism of states to the players. Some recent contributions of the second author in this area are reported. The chapter also provides a brief overview of the theory of nonzero-sum stochastic games and stopping games.

Stochastic Markov Games

Nonzero-sum versions of Shapley's stochastic games [87] with the discounted payoff criterion were first studied by Fink [88] and Takahashi [89]. The theory of nonzero-sum stochastic games with the average payoffs per unit time for the players started with the papers by Rogers [1] and Sobel [2]. They considered finite state spaces only and assumed that the transition probability matrices induced by any stationary strategies of the players are unichain. Until now only special classes of nonzero-sum average payoff stochastic games were shown to possess Nash equilibria or ϵ-equilibria. Parthasarathy and Raghavan [3] considered games in which one player is able to control transition probabilities and proved the existence of stationary equilibria in such a case. A constructive proof of their results is given in Nowak and Raghavan [90], and other results

Key words and phrases: Average payoff stochastic games, correlated stationary equilibria, nonzero-sum games, stopping time, stopping games.
concerning algorithms for stochastic games are reported in a survey paper by Raghavan and Filar [91] and in the book by Filar and Vrieze [92]. Non-stationary $\varepsilon$-equilibria were shown to exist in games with state independent transitions by Thuijsman [4] and in games with absorbing states by Vrieze and Thuijsman [5]. Thuijsman and Raghavan [93] constructed $\varepsilon$-equilibria for stochastic games with additive transition and reward structure and in perfect information games. Existence of $\varepsilon$-equilibria for 2-person nonzero-sum stochastic games with finite state and action spaces was established by Vieille [94–96], and $\varepsilon$-correlated equilibria were constructed for $n$-person stochastic games by Solan and Vieille [97]. The theory of zero-sum stochastic games with the limiting average criterion is rather complete in view of the contributions by Mertens and Neyman [98] and Maitra and Sudderth [99], [100]. Parthasarathy [6] first considered nonzero-sum stochastic games with countable state spaces and proved that every discounted stochastic game always has a stationary Nash equilibrium solution. Some extensions of Parthasarathy's result [6] to nonzero-sum stochastic games with general (continuous) utility functions defined on the space of infinite histories of the play were given by Nowak [101] and Rieder [102]. Federgruen [7] extended the works of Rogers and Sobel to average payoff nonzero-sum stochastic games with countably many states, satisfying a natural uniform geometric ergodicity condition. Federgruen's results [7] was considerably generalized by Altman, Hordijk, and Sipesma [103] who assumed a $\mu$-recurrence condition implying a weaker version of geometric ergodicity property of the Markov chains induced by stationary strategies of the players. Sennott [104] considered some conditions described in terms of the optimality equations and obtained related results. Borkar and Ghosh [8] and Spiessma and Passchier [105] (see also [106]) studied stochastic games with countably many states under a stability assumption of Lyapunov type and proved some results on almost sure optimality in such games. In a recent paper by Nowak [107] some sensitive optimality criteria are considered in a class of ergodic stochastic games with countable state spaces. The payoff (cost) functions are assumed in [103, 104, 105] to be unbounded in the state variable. Such an approach is motivated by possible applications of stochastic games to queueing theory [108, 103, 106].

In many applications of stochastic games, especially in economic theory, it is desirable to assume that the state spaces are not discrete; see, for example, Duffie et al. [10], Dutta [11], Karatzas et al. [12], or Majumdar and Sundaram [13]. The mentioned papers deal with dynamic program-
ng or discounted stochastic games only. There are also some papers
evoted to nonzero-sum average payoff stochastic games with uncount-
le state spaces. Dutta and Sundaram [14] studied a class of dynamic
onomic games. They proved the existence of stationary Nash equilibria
a class of games satisfying a number of specific conditions and a con-
vergence condition imposed on discounted Nash equilibria as the discount-
tor tends to one. A related result was given for discounted games by
urtat [109]. Ghosh and Bagchi [15] studied games under some separa-
bility assumptions and a recurrence condition which is stronger than
aim geometric ergodicity.

Our main objective in this section is to describe the idea of correlated
ilibrium notion and report a correlated equilibrium theorem proved
or discounted stochastic games by Nowak and Raghavan [16]. We will
so report an extension of this result to undiscounted stochastic games
ained by Nowak [17].

To describe the model, we need the following definition. Let $X$ be
metric space, $(S, \Sigma)$ a measurable space. A set-valued map or a cor-
respondence $F$ from $S$ into a family of subsets of $X$ is said to be lower
measurable if for any open subset $G$ of $X$ the set $\{ s \in S : F(s) \cap G \neq \emptyset \}$
ells to $\Sigma$. For a broad discussion of lower measurable correspond-
ces with some applications to control and optimization theory consult
astaing and Valadier [18] or Himmelberg [19].

An $N$-person nonzero-sum stochastic game is defined by the following
objects:

$$((S, \Sigma), X_k, A_k, r_k, q)$$

with the interpretation that

(i) $(S, \Sigma)$ is a measurable space, where $S$ is the set of states for the
game, and $\Sigma$ is a countably generated $\sigma$-algebra of subsets of $S$.

(ii) $X_k$ is a non-empty compact metric space of actions for player $k$.
We put $X = X_1 \times X_2 \times \cdots \times X_N$.

(iii) $A_k$s are lower measurable correspondences from $S$ into non-empty
ompact subsets of $X_k$. For each $s \in S$, $A_k(s)$ represents the set of
ctions available to player $k$ in state $s$. We put

$$A(s) = A_1(s) \times A_2(s) \times \cdots \times A_N(s), \quad s \in S.$$ 

(iv) $r_k : S \times X \to R$ is a bounded product measurable payoff function
or player $k$. It is assumed that $r_k(s, \cdot)$ is continuous on $X$, for every
$S \in S$. 


(v) \( q \) is a product measurable transition probability from \( S \times X \) to \( S \), called the law of motion among states. If \( s \) is a state at some stage of the game and the players select an \( x \in A(s) \), then \( q(\cdot \mid s, x) \) is the probability distribution of the next state of the game. We assume that the transition probability \( q \) has a density function, say \( z \), with respect to a fixed probability measure \( \mu \) on \( (S, \Sigma) \), satisfying the following \( L_1 \) continuity condition:

For any sequence of joint action tuples \( x^n \to x^0 \),

\[
\int_S \left| z(s, t, x^n) - z(s, t, x^0) \right| \mu(dt) \to 0 \quad \text{as} \quad n \to \infty.
\]

The \( L_1 \) continuity above is satisfied via Scheffe's theorem when \( z(s, t, \cdot) \) is continuous on \( X \). It implies the norm continuity of the transition probability \( q(\cdot \mid s, x) \) with respect to \( x \in X \).

The game is played in discrete time with past history as common knowledge for all the players. An individual strategy for a player is a map which associates with each given history a probability distribution on the set of available actions. A stationary strategy for player \( k \) is a map which associates with each state \( s \in S \) a probability distribution on the set \( A_k(s) \) of actions available to him at \( s \), independent of the history that lead to the state \( s \). A stationary strategy for player \( k \) can thus be identified with a measurable transition probability \( f \) from \( S \) to \( X_k \) such that \( f(A_k(s) \mid s) = 1 \), for every \( s \in S \).

Let \( H = S \times X \times S \times \cdots \) be the space of all infinite histories of the game, endowed with the product \( \sigma \)-algebra. For any profile of strategies \( \pi = (\pi_1, \ldots, \pi_N) \) of the players and every initial state \( s_1 = s \in S \), a probability measure \( P_{s_1}^\pi \) and a stochastic process \( \{\sigma_n, \alpha_n\} \) are defined on \( H \) in a canonical way, where the random variables \( \sigma_n \) and \( \alpha_n \) describe the state and the actions chosen by the players, respectively, on the \( n \)th stage of the game (cf. Chapter 7 in Bertsekas and Shreve [20] or Neveu [32]). Thus, for each profile of strategies \( \pi = (\pi_1, \ldots, \pi_N) \), any finite horizon \( T \), and every initial state \( s \in S \), the expected \( T \)-stage payoff to player \( k \) is

\[
\Phi_k^T(\pi)(s) = E_s^\pi \left( \sum_{n=1}^T r_k(\sigma_n, \alpha_n) \right).
\]

Here \( E_s^\pi \) means the expectation operator with respect to the probability measure \( P_{s_1}^\pi \). If \( \beta \) is a fixed real number in \((0, 1)\), called the discount factor, then we can also consider the \( \beta \)-discounted expected payoff to player \( k \) defined as

\[
\Phi_k^\beta(\pi)(s) = E_s^\pi \left( \sum_{n=1}^\infty \beta^{n-1} r_k(\sigma_n, \alpha_n) \right).
\]
The average payoff per unit time for player \( k \) is defined as

\[
\Phi_k(\pi)(s) = \limsup_{T} \frac{1}{T} \Phi_k^T(\pi)(s).
\]

Let \( \pi^* = (\pi_1^*, \ldots, \pi_N^*) \) be a fixed profile of strategies of the players. For any strategy \( \pi_k \) of player \( k \), we write \( (\pi_k^*, \pi_k) \) to denote the strategy profile obtained from \( \pi^* \) by replacing \( \pi_k^* \) with \( \pi_k \).

A strategy profile \( \pi^* = (\pi_1^*, \ldots, \pi_N^*) \) is called a Nash equilibrium for the average payoff stochastic game if no unilateral deviations from it are profitable, that is, for each \( s \in S \),

\[
\Phi_k(\pi^*)(s) \geq \Phi_k(\pi_{-k}^*, \pi_k)(s),
\]

for every player \( k \) and any strategy \( \pi_k \). Of course, Nash equilibria are analogously defined for the \( \beta \)-discounted stochastic games.

It is still an open problem whether the \( \beta \)-discounted stochastic games with uncountable state space have stationary equilibrium solutions. A positive answer to this problem is known only for some special classes of games, where the transition probabilities satisfy certain additional separability assumptions (cf. Himmelberg et al. [21]), or some other specific conditions (cf. Majumdar and Sundaram [13], Dutta and Sundaram [14], Karatzas et al., [12] or if the game is of perfect information [110]). Whitt [22] and Nowak [23] proved the existence of stationary \( \epsilon \)-equilibrium strategies in discounted stochastic games using some (different) approximations by games with countably many states. The assumptions on the model in Nowak [23] are as in (i) through (v) above plus some extra integrability condition on the transition probability density. Whitt [22] assumed that the state spaces are separable metric and imposed some uniform continuity conditions on the payoffs and transition probabilities. Breton and L'Ecuyer [24] extended Whitt's result to games with a weaker form of discounting. Considerable extensions of Nowak's result [111] to some classes of nonzero-sum stochastic games with unbounded cost functions were recently given by Nowak and Altman in [112] where the expected average cost criterion is also considered under some stochastic stability condition inspired by a recent paper by Meyn and Tweedie [113]. Mertens and Parthasarathy [25] proved the existence of non-stationary Nash equilibria for discounted stochastic games with arbitrary state spaces. Finally, Nowak and Raghavan [16] obtained stationary equilibrium solutions in the class of correlated strategies of the players with symmetric information or "public signals" (see Theorem 1). A related result is reported in Duffie et al. [10]. They used some stronger
assumptions about the primitive data of the game, but showed that there exists a stationary correlated equilibrium which induces an ergodic process. Nonstationary correlated equilibria in a class of dynamic games with weakly continuous transition probabilities were studied by Harris [26]. As already mentioned, Dutta and Sundaram [14] proved an existence theorem for stationary Nash equilibria in some undiscounted dynamic economic games. Equilibria for classes of strongly ergodic stochastic games of perfection formation were shown to exist by Küenle [110].

1.1 Correlated Equilibria

In this subsection we extend the sets of strategies available to the players in the sense that we allow them to correlate their choices in a natural way described below. The resulting solution is a kind of extensive-form correlated equilibrium (cf. Forges [27]).

Suppose that \( \{\xi_n : n \geq 1\} \) is a sequence of so-called signals, drawn independently from \([0, 1]\) according to the uniform distribution. Suppose that at the beginning of each period \( n \) of the game the players are informed not only of the outcome of the preceding period and the current state \( s_n \), but also of \( \xi_n \). Then the information available to them is a vector \( h^n = (s_1, \xi_1, x_1, \ldots, s_{n-1}, \xi_{n-1}, x_{n-1}, s_n, \xi_n) \), where \( s_i \in S, x_i \in A(s_i), i = 1, \ldots, n - 1 \). We denote the set of such vectors by \( H^n \).

An extended strategy for player \( k \) is a sequence \( \pi_k = (\pi^1_k, \pi^2_k, \ldots) \), where every \( \pi^n_k \) is a (product) measurable transition probability from \( H^n \) to \( X_k \) such that \( \pi^n_k(A_k(s_n) \mid h^n) = 1 \) for any history \( h^n \in H^n \). (Here \( s_n \) is the last state in \( h^n \).) An extended stationary strategy for player \( k \) is a strategy \( \pi_k = (\pi^1_k, \pi^2_k, \ldots) \) such that each \( \pi^n_k \) depends on the current state \( s_n \) and the last signal \( \xi_n \) only. In other words, a strategy \( \pi_k \) of player \( k \) is called stationary if there exists a transition probability \( f \) from \( S \times [0, 1] \) to \( X_k \) such that for every period \( n \) of the game and each history \( h^n \in H^n \), we have \( \pi^n_k(\cdot \mid h^n) = f(\cdot \mid s_n, \xi_n) \). Assuming that the players use extended strategies we actually assume that they play a stochastic game in the sense of Section , but with the extended state space \( S \times [0, 1] \). The law of motion, say \( \bar{q} \), in the extended state space model is obviously the product of the original law of motion \( q \) and the uniform distribution \( \eta \) on \([0, 1]\).

More precisely, for any \( s \in S, \xi \in [0, 1], a \in A(s) \), any set \( C \in \Sigma \) and any Borel set \( D \subseteq [0, 1], \bar{q}(C \times D \mid s, \xi, a) = q(C \mid s, a)\eta(D) \).

For any profile of extended strategies \( \pi = (\pi_1, \ldots, \pi_N) \) of the players, the undiscounted \([\beta\text{-discounted}]\) payoff to player \( k \) is a function of the initial state \( s_1 \) and the first signal \( \xi_1 \) and is denoted by \( E_k(\pi)(s_1, \xi_1) \) \( [E^\beta_k(\pi)(s_1, \xi_1)] \).
We say that \( f^* = (f_1^*, \ldots, f_N^*) \) is a Nash equilibrium in the average payoff stochastic game in the class of extended strategies if for each initial state \( s_1 \in S \),

\[
\int_0^1 \Phi_k(f^*)(s_1, \xi_1)\eta(d\xi_1) \geq \int_0^1 \Phi_k(f_{-k}^*(\pi), \xi_1)\eta(d\xi_1),
\]

for every player \( k \) and any extended strategy \( \pi_k \).

A Nash equilibrium in extended strategies is also called a correlated equilibrium with public signals. The reason is that after the outcome of any period of the game, the players can coordinate their next choices by exploiting the next (known to all of them, i.e., public) signal and using some coordination mechanism telling which (pure or mixed) action is to be played by each of them. In many applications, we are particularly interested in stationary equilibria. In such a case the coordination mechanism can be represented by a family of \( N + 1 \) measurable functions \( \lambda^1, \ldots, \lambda^{N+1} : S \to [0, 1] \) such that \( \sum_{i=1}^{N+1} \lambda^i(s) = 1 \) for every \( s \in S \).

(We remind the reader that \( N \) is the number of players. The number \( N + 1 \) appears in our definition because Carathéodory's theorem is applied in the proofs of the main results in [16] and [17].) A stationary Nash equilibrium in the class of extended strategies can be constructed then by using a family of \( N + 1 \) stationary strategies \( f_1^k, \ldots, f_k^{N+1} \), given for each player \( k \), and the following coordination rule. If the game is at a state \( s \) on the \( n \)th stage and a random number \( \xi_n \) is selected, then each player \( k \) is suggested to use \( f_k^m(\cdot \mid s) \), where \( m \) is the least index for which \( \sum_{i=1}^{m} \lambda^i(s) \geq \xi_n \). The \( \lambda^i \)'s and \( f_k^m \)'s fixed above induce an extended stationary strategy \( f_k^* \) for each player \( k \) as follows

\[
f_k^*(\cdot \mid s, \xi) = f_k^l(\cdot \mid s) \quad \text{if} \quad \xi \leq \lambda^1(s), \quad s \in S,
\]

and

\[
f_k^*(\cdot \mid s, \xi) = f_k^m(\cdot \mid s) \quad \text{if} \quad \sum_{i=1}^{m-1} \lambda^i(s) < \xi \leq \sum_{i=1}^{m} \lambda^i(s),
\]

for \( s \in S, \ 2 \leq m \leq N + 1 \). Because the signals are independent and uniformly distributed in \([0, 1]\), it follows that at any period of the game and for any current state \( s \), the number \( \lambda^i(s) \) can be interpreted as the probability that player \( k \) is suggested to use \( f_k^l(\cdot \mid s) \) as his mixed action. Now it is quite obvious that a strategy profile \((f_1^*, \ldots, f_N^*)\) obtained by the above construction is a stationary Nash equilibrium in the class of extended strategies of the players in a game iff no player \( k \) can unilaterally improve upon his expected payoff by changing any of his strategies \( f_k^i \), \( i = 1, \ldots, N + 1 \).
The following result was proved by Nowak and Raghavan [16] by a fixed point argument.

**Theorem 1** Every nonzero-sum discounted stochastic game satisfying (i) through (v) has a stationary correlated equilibrium with public signals.

To report an extension of this result to undiscounted stochastic games obtained in Nowak [17], we need some additional assumptions on the transition probability $q$. For any stationary strategy profile $f$ and $n \geq 1$, let $q^n(\cdot \mid s, f)$ denote the $n$-step transition probability determined by $q$ and $f$. The following condition is used in the theory of Markov decision processes (cf. Tweedie [28], Hernández-Lerma et al. [29, 30], and their references):

**C1 (Uniform Geometric Ergodicity):** There exist scalars $\alpha \in (0, 1)$ and $\gamma > 0$ for which the following holds: For any profile $f$ of stationary strategies of the players, there exists a probability measure $p_f$ on $S$ such that

$$\|q^n(\cdot \mid s, f) - p_f(\cdot)\|_\nu \leq \gamma \alpha^n$$

for each $n \geq 1$.

Here $\| \cdot \|_\nu$ denotes the total variation norm in the space of finite signed measures on $S$.

It is well known that C1 follows from the following assumption (cf. Theorem 6.15 and Remark 6.1 in Nummelin [31] or page 185 in Neveu [32]):

**M (Minorization Property):** There exists a positive integer $p$, a constant $\theta > 0$, and a probability measure $\delta$ on $S$, such that

$$q^p(D \mid s, f) \geq \theta \delta(D),$$

for every stationary strategy profile $f, s \in S$, and for each measurable subset $D$ of $S$.

Condition M was used in stochastic dynamic programming (one-person stochastic game) by Kurano [33] who proved only the existence of "p-periodic" optimal strategies in his model. It is satisfied and easy to verify in some inventory models (cf. Yamada [34]) and some control of water resources problems (cf. Yakovitz [35]). Küenle used condition M (with $p = 1$) to study equilibria in stochastic games of perfect formations (see [110]) and the references therein.

Condition C1 has often been used (even in some stronger versions) in control theory of Markov chains (cf. Georgin [36], Hernández-Lerma et
al. [29, 30] and the references therein). We mention here some conditions which are known to be equivalent to C1. By \( F \) we denote the set of all stationary strategy \( N \)-tuples of the players.

\textbf{C2 (Uniform Ergodicity):} For each \( f \in F \), there exists a probability measure \( p_f \) on \( S \) such that, as \( n \to \infty \),

\[
\| q^n(\cdot \mid s, f) - p_f(\cdot) \|_{\nu} \to 0, \text{ uniformly in } s \in S \text{ and } f \in F.
\]

\textbf{C3:} There exist a positive integer \( r \) and a positive number \( \delta < 1 \) such that

\[
\| q^r(\cdot \mid s, f) - q^r(\cdot \mid t, f) \|_{\nu} \leq 2\delta, \text{ for all } s, t \in S \text{ and } f \in F.
\]

Obviously C1 implies C2 and C3 follows immediately from C2 and the triangle inequality for the norm \( \| \cdot \|_{\nu} \). Finally, C3 implies C1 by Ueno's lemma [37]. For details consult pages 275–276 in [36].

Another equivalent version of C1, called the \textit{simultaneous Doeblin condition}, was used by Hordijk [38] in control theory and Federgruen [7] in stochastic games with countably many states. It can also be formulated for general state space stochastic games following pages 192 and 221 in Doob [39].

\textbf{C4:} There is a probability measure \( \psi \) on \( S \), a positive integer \( r \), and a positive \( \epsilon \) such that

\[
q^r(B \mid s, f) \leq 1 - \epsilon \quad \text{for each } s \in S \text{ and } f \in F \text{ if } \psi(B) \leq \epsilon.
\]

Moreover, for each \( f \in F \), the Markov chain induced by \( q \) and \( f \) has a single ergodic set and this set contains no cyclically moving subsets.

It turns out that C1 is equivalent to C4; see Chapter V in Doob [39] for details. For a further discussion of several recurrence and ergodicity conditions which have been used in the theory of Markov decision processes in a general state space, consult Hernández-Lerma et al. [30]. Now the main result of Nowak [17] can be formulated.

\textbf{Theorem 2} Every nonzero-sum undiscounted stochastic game satisfying (i) through (v) and C1 has a stationary correlated equilibrium with public signals.

We now mention some special classes of nonzero-sum undiscounted stochastic games for which there exist Nash equilibria without public
signals. First, we consider games satisfying the following separability conditions:

\textbf{SC1:} For each player \( k \) and any \( s \in S, x = (x_1, \ldots, x_N) \in A(s), \)

\[ r_k(s, x) = \sum_{j=1}^{N} r_{kj}(s, x_j), \]

where each \( r_{kj} \) is bounded and \( r_{kj}(s, \cdot) \) is continuous on \( X_j. \)

\textbf{SC2:} For any \( s \in S, x = (x_1, \ldots, x_N) \in A(s), \)

\[ q(\cdot \mid s, x) = \frac{\sum_{j=1}^{N} q_j(\cdot \mid s, x_j)}{N}, \]

where \( q(\cdot \mid s, x_j) \) is a transition probability from \( S \times X_j \) to \( S, \) norm continuous with respect to \( x_j \in X_j. \)

Himmelberg et al. [21] and Parthasarathy [40] already showed that nonzero-sum \( \beta \)-discounted stochastic games satisfying SC1 and SC2 possess stationary Nash equilibria. Their theorem was extended to undiscounted stochastic games in Nowak [17].

\textbf{Theorem 3} Every nonzero-sum undiscounted stochastic game satisfying (i) through (v), C1 and separability conditions SC1 and SC2 has a stationary Nash equilibrium without public signals.

By Theorem 2, the game has a stationary correlated equilibrium, say \( f^\Lambda. \) For each player \( k \) and any \( s \in S, \) we define \( f_k^*(\cdot \mid s) \) to be the marginal of \( f^\Lambda(\cdot \mid s) \) on \( X_k \) and put \( f^* = (f_1^*, \ldots, f_N^*). \) It turns out that \( (f_1^*, \ldots, f_N^*) \) is a Nash equilibrium point for the stochastic game, satisfying SC1 and SC2.

A version of Theorem 3 with a recurrence assumption, which is much stronger than the uniform geometric ergodicity, was independently proved (by a different method) in Ghosh and Bagchi [15].

Parthasarathy and Sinha [41] showed that \( \beta \)-discounted stochastic games with state independent transitions and finite action spaces have stationary Nash equilibria. An extension of their result to the average payoff stochastic games, obtained in Nowak [17] is as follows.

\textbf{Theorem 4} Assume that the action spaces \( X_k \) are finite sets and \( A_k(s) = X_k \) for each \( s \in S. \) Assume also that the transition probability \( q(\cdot \mid s, x) \) depends on \( x \) only and is non-atomic for each \( x \in X. \) If (i), (iv), (v), and C1 are also satisfied, then the nonzero-sum average payoff stochastic game has a stationary Nash equilibrium without public signals.
We do not know if condition C1 can be dropped from Theorem 4. When we deal with zero-sum average payoff stochastic games with state independent transitions, then no ergodicity properties of the transition probability \( q \) are relevant (cf. Thuysius [4] for the finite state space case or Theorem 2 in Nowak [42] for general state space games).

The basic idea of the proof of Theorem 2 is rather simple. Let \( L \) be any positive number such that \( |r_k| \leq L \) for every player \( k \). Then, for every discount factor \( \beta \), and any stationary correlated equilibrium \( f_0^\lambda \) obtained in Theorem 1, \( (1 - \beta)\Phi_k(f_0^\lambda)(\cdot) \) is in a compact ball \( B(L) \) with radius \( L \) in \( L^\infty(S, \Sigma, \mu) \) space, endowed with the weak-star topology \( \sigma(L^\infty, L_1) \). Therefore, it is possible to find a sequence \( \{\beta_n\} \) of discount factors which converges to one and \( (1 - \beta_n)\Phi_k(f_{\beta_n}^\lambda) \) converges to some function \( J_k \in B(L) \). Using C1, it is shown that \( J_k \) are constant equilibrium functions of the players, and \( f_{\beta_n}^\lambda \) converges (in some sense) to a stationary correlated equilibrium for the undiscounted game.

As far as two-person zero-sum games are concerned, it is possible to drop the assumption that the transition probability is dominated by some probability measure \( \mu \). To prove the existence of stationary optimal strategies of the players, one can use the following assumption (see Nowak [42]).

**B:** Assume (i) through (iv) and that \( q(D|s, \cdot) \) is continuous on \( X = X_1 \times X_2 \) for each \( D \in \Sigma \). Let \( v_\beta(\cdot) \) be the value function of the \( \beta \)-discounted game, \( \beta \in (0, 1) \). We assume that there exists a positive constant \( L \) such that

\[
|v_\beta(s) - v_\beta(t)| \leq L \text{ for all } s, t \in S \text{ and } \beta \in (0, 1).
\]

It is easy to see that C1 implies B. Moreover, B holds if the transition probability \( q \) is independent of the state variable. The main tool in the proof given in [42] is Fatou’s lemma for varying probabilities (see Dutta [11] or Schäl [43] for a related approach in dynamic programming). That is the main difference between the proofs contained in [17] and [42]. Zero-sum expected average payoff stochastic games with general state spaces were also studied by Rieder [114] where a span fixed point argument is used under an ergodicity assumption, which is essentially stronger than C1. An interesting result on computable bounds for geometric convergence rates of Markov chains proven by Meyn and Tweedie [115] made it possible to formulate, and relatively simple to verify, general stochastic stability assumptions on the transition structure of the stochastic game implying a weaker version of ergodicity of Markov chains induced by stationary strategies of the players, called the \( V \)-uniform geometric er-
2. Stopping Games

The theory of stopping games started with the paper by Dynkin [44]. He considered a zero-sum game based on the optimal stopping problem for discrete time stochastic processes. Let \( \{X_n\}_{n=0}^{\infty} \) be a stochastic sequence defined on some fixed probability space \((\Omega, \mathcal{F}, P)\). Define \( \mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n) \) to be the \( \sigma \)-field generated by the random variables \(X_0, X_1, \ldots, X_n\). The random variable \( \lambda = \lambda(\omega) \) with values in \( \bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\} \) is said to be Markov time with respect to the system \( F = \{\mathcal{F}_n\}, n \in \bar{\mathbb{N}} \) if for each \( n \in \bar{\mathbb{N}} \)

\[ \{\omega : \lambda(\omega) = n\} \in \mathcal{F}_n. \]

The Markov time \( \lambda = \lambda(\omega) \) is said to be a stopping time or a finite Markov time if \( P\{\lambda(\omega) < \infty\} = 1 \). Strategies for the players are Markov times with respect to \( F = \{\mathcal{F}_n\}, n \in \bar{\mathbb{N}} \). If Players 1, 2 choose strategies \( \lambda, \mu \), respectively, then Player 2 pays Player 1 the amount \( R(\lambda, \mu) = X_{\lambda(\mu)}. \)

We will be interested in the expectation of \( ER(\lambda, \mu) \). For this we must assume that \( R(\lambda, \mu) \) is integrable. The aim of Player 1 (resp., Player 2) is to make \( ER(\lambda, \mu) \) as large (resp., as small) as possible. Dynkin [44] assumed a restriction on the moves of the game. Namely, the strategies of the players are such that Player 1 can stop on odd moments only and Player 2 can choose even moments. Under this assumption, Dynkin proved the existence of value and optimal strategies for the players. Kifer [45] provided another sufficient condition. Neveu [46] modified Dynkin's game as follows. There are two preassigned stochastic sequences \( \{X_n\}_{n=0}^{\infty}, \{Y_n\}_{n=0}^{\infty} \) measurable with respect to some increasing sequence of \( \sigma \)-fields \( \mathcal{F}_n \). The players' strategies are stopping times with respect to \( \{F_n\}_{n=0}^{\infty} \). The payoff equals

\[
R(\lambda, \mu) = \begin{cases} 
X_\lambda & \text{on } \{\lambda \leq \mu\}, \\
X_\mu & \text{on } \{\lambda > \mu\}, 
\end{cases}
\]  

(3)
with the condition
\[ X_n \leq Y_n \text{ for each } n. \] (4)

Under some regularity condition Neveu proved the existence of the game value and c-optimal strategies.

The restriction (4) has been suppressed in some cases by Yasuda [47]. He considers the zero-sum stopping game with payoff equals
\[ R(\lambda, \mu) = X_\lambda I_{(\lambda \leq \mu)} + W_\lambda I_{(\lambda = \mu)} + Y_\mu I_{(\lambda > \mu)}, \]
where I is an indicator function. To solve the game the set of strategies has been extended to a class of randomize strategies.

A version of Dynkin's game for Markov chains was considered by Fried [48]. More general version of the stopping game for the discrete time Markov processes was solved by Elbakidze [49]. Let \((X_n, F_n, P_x)_{n=0}^\infty\) be a homogeneous Markov chain with state space \((E, \mathcal{B})\), while \(g, G, e\) and \(C\) are certain \(\mathcal{B}\)-measurable real valued functions. There are two players. The process can be stopped at any instant \(n \geq 0\). If the process is stopped by the first, second or simultaneously by the two players, then the payoffs of the player are \(g(X_n), G(X_n)\) and \(e(X_n)\), respectively. For an unlimited duration of the game the payoff of the first player equals \(\limsup_{n \to \infty} C(X_n)\). The strategies of the first and second player are given by Markov moments relative to \(\{F_n\}_{n=0}^\infty\). Let \(\mathcal{L}\) denote a class of \(\mathcal{B}\)-measurable functions \(f\) such that \(\mathbb{E}_x\{\sup_{n} |f(X_n)|\} < \infty\). It is assumed that
\[ g(x) \leq e(x) \leq G(x), \quad g(x) \leq C(x) \leq G(x), \quad x \in E \text{ and } g, G \in \mathcal{L}. \]

Under these assumptions the value of the game and c-optimal strategies are constructed.

Two-person nonzero-sum stopping games is investigated, among others, by Ohtsubo [50]. Let \(\{X_n^i\}_{n=0}^\infty, \{Y_n^i\}_{n=0}^\infty\) and \(\{W_n^i\}_{n=0}^\infty, i = 1, 2\), be six sequences of real-valued random variables defined on fixed probability space and adapted to \(\{F_n\}_{n=0}^\infty\). It is assumed that

(i) \(\min(X_n^i, Y_n^i) \leq W_n^i \leq \max(X_n^i, Y_n^i)\) for each \(i = 1, 2\).

(ii) \(\mathbb{E}[\sup_{n} |X_n^i|] < \infty\) and \(\mathbb{E}[\sup_{n} |Y_n^i|] < \infty\) for each \(i = 1, 2\).

The strategies of the players are stopping times with respect to \(\{F_n\}_{n=0}^\infty\). If the first and the second players choose stopping times \(\tau_1\) and \(\tau_2\), respectively, as their controls, then the \(i\)-th player gets the reward
\[ g_i(\tau_1, \tau_2) = X_{\tau_1}^i I_{(\tau_1 < \tau_2)} + Y_{\tau_2}^i I_{(\tau_2 < \tau_1)} + W_{\tau_1}^i I_{(\tau_1 = \tau_2 < \infty)} + \limsup_{n} W_n^i I_{(\tau_i = \tau_j < \infty)}, i, j = 1, 2, j \neq i. \]
Under the above assumption the Nash equilibrium for the game is constructed. Ohtsubo [50] gave the solution for the version of the game for the Markov processes. Recently, Ferenstein [51] solved the version of the nonzero-sum Dynkin's game with a different, special, payoff structure.

Continuous time version of stopping games were studied by Bensoussan and Friedman [52], [53], Krylov [54], Bismut [55], Stettner [56], Lepeltier and Maingueneau [57], and many others.

In this chapter we focus on a version of the stopping games called the random priority stopping game.

2.1 Zero-Sum Random Priority Stopping Game

Let \((X_n, \mathcal{F}_n, P_n)_{n=0}^{N}\) be a homogeneous Markov process defined on probability space \((\Omega, \mathcal{F}, P)\) with fixed state space \((E, B)\). The decision makers, henceforth called Player 1 and Player 2, observe the process sequentially. They want to accept the most profitable state of the process from their point of view.

We adopt the zero-sum game model for the problem. In view of this approach, the preferences of each player are described by gain function \(f : E \times E \rightarrow \mathbb{R}\). The function depends on the state chosen by both players. It would be natural to consider the stopping times with respect to \((\mathcal{F}_n)_{n=0}^{N}\) as the strategies of the player if the players could obtain the state which they want. Since there is only one random sequence \((X_n)_{n=0}^{N}\) on a trial, therefore at each moment \(n\) only one player can obtain realization \(x_n\) of \(X_n\). The problem of assigning an object to the players when both want to accept the same one at the same moment is solved by adopting the random mechanism i.e. a lottery chooses the player who benefits. The player chosen by the lottery obtains realization \(x_n\) and the player thus deprived of the acceptance of \(x_n\) at \(n < N\) may select any later realization. The realization can only be accepted when it appears. No recall is allowed. We can think about the decision process as an investigation of objects with characteristics described by the Markov process. Both players together can accept at most two objects.

In the next section a formal model of the random priority zero-sum stopping game is constructed. This very interesting question concerns the influence of the priority on the value of the game or the probability of obtaining the required state of the process (or, in other words, the required object). The tip of the problem is shown by the example related to the secretary problem in Section 2.1.2. The simplest problem with asymmetric aims of the players is considered. The first player's aim is to choose the best applicant (BA), and the second player wants to accept
the best or the second best (BOS) but a better one than the applicant chosen by the opponent. The numerical solution provides that the game is fair when Player 1 has priority \( p \approx 0.7579 \) (in the limiting case when \( N \to \infty \)).

2.1.1 Random Priority and Stopping the Markov Process

Let a homogeneous Markov chain \( (X_n, \mathcal{F}_n, P_x)_{n=0}^{\infty} \) be defined on a probability space \((\Omega, \mathcal{F}, P)\) with a state space \((E, B)\) and let \( f : E \times E \to \mathbb{R} \) be a \( B \times B \) real valued measurable function. Horizon \( N \) is finite. The players observe the Markov chain and they try to accept the "best realization" according to function \( f \) and a possible selection of another player. Each realization \( x_n \) of \( X_n \) can be accepted by only one player and each player can accept at most one realization. If the players have not accepted previous realizations, they evaluate the state of the Markov chain at instant \( n \) and they have two options, either to accept the observed state of the process at moment \( n \) or to reject it. If both players want to accept the same realization, the following random priority mechanism is applied.

Let \( \xi_1, \xi_2, \ldots, \xi_N \) be a sequence of i.i.d. r.v. with the uniform distribution on \([0, 1]\) and \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N) \) be a given vector of real numbers, \( \alpha_i \in [0, 1] \). When both players want to accept realization \( x_n \) of \( X_n \), then Player 1 obtains \( x_n \) if \( \xi_n \leq \alpha_n \); otherwise Player 2 benefits. If Player 1 rejects the state, then Player 2 has to make a decision. If one of the players accepts realization \( x_n \) of \( X_n \), then the other one is informed about it and he continues to play alone. If, in the above decision process, Player 1 and Player 2 have accepted states \( x \) and \( y \), respectively, then Player 2 pays \( f(x, y) \) to Player 1. When only Player 1 (Player 2) accepts state \( x \) (\( y \)) then Player 1 obtains \( f_1(x) = \sup_{y \in E} f(x, y) \) (\( f_2(y) = \inf_{x \in E} f(x, y) \)) by assumption. If both players finish the decision process without any accepted state, then they gain 0. The detail construction of the model is given in [74].

The decision model described above is a generalization of the problems considered by Szajowski [58] and Radzik and Szajowski [59]. Related questions, when Player 1 has permanent priority, have been considered by any authors. One can mention, for example, the papers of Ano [60], Enns and Ferenstein [117], Ferenstein [62], and Sakaguchi [63]. Many papers on the subject were inspired by the secretary problem (see the papers by Enns and Ferenstein [64], Fushimi [65], Majumdar [66], Sakaguchi [63, 67], Ravindran and Szajowski [68], and Szajowski [69] where nonzero sum versions of the games have also been investigated). Sakaguchi [63] considered nonzero-sum two-person games related to the full information
best choice problem with random priority. A review of these problems is contained in Ravindran and Szajkowski [68]. For the original secretary problem and its extension, the reader is referred to Gilbert and Mosteller [70], Freeman [71], Rose [72], and Ferguson [73]. We recall the best choice problem in Section 2.1.2.

A brief construction of the model of stopping game with random priority is presented now. Let $S^N$ be the aggregation of Markov times with respect to $(F_N)^N_{n=0}$. We admit that $\mathbb{P}(\tau \leq N) < 1$ for some $\tau \in S^N$ (i.e., there is a positive probability that the Markov chain will not be stopped). The elements of $S^N$ are possible strategies for the players with the restriction that Player 2 cannot stop at the moment at which Player 1 stops. If both players declare willingness to accept the same object, the random device decides who is endowed. Let us formalize these in this game model. Denote $S^N_k = \{\tau \in S^N : \tau \geq k\}$. Let $\Lambda^N_k$ and $M^N_k$ be copies of $S^N_k$ ($S^N = S^N_0$). The sets of strategies for Players 1 and 2 are: $\Lambda^N = \{\lambda, \{\sigma^1_n\} : \lambda \in \Lambda^N, \sigma^1_n \in \Lambda^N_{n+1} \text{ for every } n\}$ and $\tilde{M}^N = \{\mu, \{\sigma^2_n\} : \mu \in M^N, \sigma^2_n \in M^N_{n+1} \text{ for every } n\}$, respectively. Denote $\tilde{F}_n = \sigma(F_n, \xi_1, \xi_2, \ldots, \xi_n)$ and let $\tilde{S}^N$ be the set of stopping times with respect to $(\tilde{F}_n)^N_{n=0}$. Define

$$\tau_1 = \lambda I_{\{\lambda < \mu\}} + (\lambda I_{\{\xi_1 < \alpha_1\}} + \sigma^1_\mu I_{\{\xi_1 > \alpha_1\}})I_{\{\lambda = \mu\}} + \sigma^1_\mu I_{\{\lambda > \mu\}}$$

and

$$\tau_2 = \mu I_{\{\lambda > \mu\}} + (\mu I_{\{\xi_1 > \alpha_1\}} + \sigma^2_\lambda I_{\{\xi_1 > \alpha_1\}})I_{\{\lambda = \mu\}} + \sigma^2_\lambda I_{\{\lambda < \mu\}}.$$

Lemma 1 Random variables $\tau_1$ and $\tau_2$ are Markov times with respect to $(\tilde{F}_n)^N_{n=0}$ and $\tau_1 \neq \tau_2$.

Let $E_x f^+_n(X_n) < \infty$ and $E_x f^-_m(X_m) < \infty$ for $n, m = 0, 1, \ldots, N$ and $x \in \mathbb{E}$. Let $s \in \Lambda^N$ and $t \in \tilde{M}^N$. Define $\tilde{R}(x, s, t) = E_x f(X_{\tau_1}, X_{\tau_2})$ as the expected gain of Player 1. In this way the normal form of the game $(\Lambda^N, \tilde{M}^N, \tilde{R}(x, s, t))$ is defined. This game is denoted by $\mathcal{G}$. The game $\mathcal{G}$ is a model of the considered bilateral stopping problem for the Markov process.

Definition 1 Pair $(s^*, t^*)$, $s^* \in \Lambda^N$, $t^* \in \tilde{M}^N$ is an equilibrium point in the game $\mathcal{G}$ if for every $x \in \mathbb{E}$, $s \in \Lambda^N$, and $t \in \tilde{M}^N$ we have

$$\tilde{R}(x, s, t^*) \leq \tilde{R}(x, s^*, t^*) \leq \tilde{R}(x, s^*, t).$$

The aim is to construct the equilibrium pair $(s^*, t^*)$. To this end, the following auxiliary game $\mathcal{G}_a$ is considered.
Define \( s_0(x, y) = S_0(x, y) = f(x, y) \) and
\[
\begin{align*}
    s_n(x, y) &= \inf_{\tau \in \mathcal{S}^n} E_y f(x, X_\tau), \\
    S_n(x, y) &= \sup_{\tau \in \mathcal{S}^n} E_x f(X_\tau, y)
\end{align*}
\]
for all \( x, y \in \mathcal{E}, n = 1, 2, \ldots, N \). By the theory of optimal stopping for the Markov processes [75], the function \( s_n(x, y) (S_n(x, y)) \) can be constructed by the recursive procedure as \( s_n(x, y) = Q_{\min}^n f(x, y) \) (\( S_n(x, y) = Q_{\max}^n f(x, y) \)), where \( Q_{\min}^n f(x, y) = f(x, y) \land T_2 f(x, y) \) and \( Q_{\max}^n f(x, y) = f(x, y) \lor T_1 f(x, y) \). The operations \( \land \) and \( \lor \) preserve measurability. This can be proved in a standard way. Hence, \( s_n(x, y) (S_n(x, y)) \) are \( \mathcal{B} \otimes \mathcal{B} \) measurable (cf. [76]). If Player 1 is the first to accept \( x \) at moment \( n \), then his expected gain is
\[
h(n, x) = E_x s_{N-n-1}(x, X_1), \quad (5)
\]
for \( n = 0, 1, \ldots, N - 1 \) and \( h(N, x) = f_1(x) \). When Player 2 is the first, then the expected gain of Player 1 is
\[
H(n, x) = E_x S_{N-n-1}(X_1, x), \quad (6)
\]
for \( n = 0, 1, \ldots, N - 1 \) and \( H(N, x) = f_2(x) \). Functions \( h(n, x) \) and \( H(n, x) \) are well defined. They are \( \mathcal{B} \)-measurable of the second variable, \( h(n, X_1) \) and \( H(n, X_1) \) are integrable with respect to \( P_x \). Let \( \Lambda^N \) and \( M^N \) be sets of strategies in \( \mathcal{G}_a \) for Player 1 and Player 2, respectively. For \( \lambda \in \Lambda^N \) and \( \mu \in M^N \), define payoff function
\[
r(\lambda, \mu) = \begin{cases} 
    h(\lambda, X_\lambda)(\mathbf{1}_{\{\lambda \leq \mu\}} + \mathbf{1}_{\{\lambda = \mu, \xi_\lambda \leq \alpha_\lambda\}}) & \text{if } \lambda \leq N \text{ or } \mu \leq N, \\
    +H(\mu, X_\mu)(\mathbf{1}_{\{\mu > \lambda\}} + \mathbf{1}_{\{\lambda = \mu, \xi_\mu > \alpha_\mu\}}) & \text{otherwise}, \\
    0
\end{cases} \quad (7)
\]
where \( \mathbf{1}_A \) is a characteristic function of set \( A \). As a solution of the game we search for equilibrium pair \((\lambda^*, \mu^*)\) such that
\[
R(x, \lambda, \mu^*) \leq R(x, \lambda^*, \mu) \leq R(x, \lambda^*, \mu) \quad \text{for all } x \in \mathcal{E}, \quad (8)
\]
where \( R(x, \lambda, \mu) = E_x r(\lambda, \mu) \). By (7) we can observe that \( \mathcal{G}_a \) with the sets of strategies \( \Lambda^N \) and \( M^N \) is equivalent to the Neveu’s stopping problem [46] considered by Yasuda [47] if the sets of strategies are extended to the set of stopping times not greater than \( N + 1 \) and the payoff function is (7). The monotonicity of gains are not fulfilled here, but the solution is still
in pure strategies. Because the Markov process is observed here, one can define a sequence \( v_n(x), \ n = 0, 1, \ldots, N + 1 \) on \( \mathbb{E} \) by setting \( v_{N+1}(x) = 0 \) and

\[
v_n(x) = \operatorname{val} \left[ \begin{array}{c}
h(n, x)\alpha_n + (1 - \alpha_n)H(n, x) \\
H(n, x)
\end{array} \right]
\]

for \( n = 0, 1, \ldots, N \), where \( Tv(x) = E_z v(X_1) \) and \( \operatorname{val} A \) denotes a value of the two-person zero-sum game with payoff matrix \( A \) (see [77], [47]).

To prove the correctness of the construction, let us observe that the payoff matrix in (9) has the form

\[
A = \begin{bmatrix}
s & f \\
(s - b)\alpha + b & a \\
(b & c)
\end{bmatrix},
\]

(10)

where \( a, b, c, \alpha \) are real numbers and \( \alpha \in [0, 1] \). By direct checking we have the following:

**Lemma 2** The two-person zero-sum game with payoff matrix \( A \) given by (10) has an equilibrium point \((\epsilon, \delta)\) in pure strategies, where

\[
(\epsilon, \delta) = \begin{cases}
(s, s) & \text{if } a \geq b, \\
(s, f) & \text{if } c \leq a < b, \\
(f, s) & \text{if } a < b \leq c, \\
(f, f) & \text{if } a < c < b.
\end{cases}
\]

Notice that \( v_{N+1} \) is measurable. Let us assume that functions \( v_i, i = N, N - 1, \ldots, n + 1 \) are measurable. Denote

\[
A_n^{ss} = \{ x \in \mathbb{E} : h(n, x) \geq H(n, x) \}
\]

\[
A_n^{sf} = \{ x \in \mathbb{E} : h(n, x) < H(n, x), h(n, x) \geq Tv_{n+1}(x) \}
\]

\[
A_n^{fs} = \{ x \in \mathbb{E} : h(n, x) < H(n, x), H(n, x) \leq Tv_{n+1}(x) \}
\]

and

\[
A_n^{sf} = \mathbb{E} \setminus (A_n^{ss} \cup A_n^{sf} \cup A_n^{fs}).
\]

By sets \( A_n^{ss}, A_n^{sf}, A_n^{fs} \in \mathcal{B} \) and Lemma 2 we have

\[
v_n(x) = [(h(n, x) - H(n, x))\alpha_n + H(n, x)]I_{A_n^{ss}}(x) + h(n, x)I_{A_n^{sf}}(x) + H(n, x)I_{A_n^{fs}}(x) + Tv_{n+1}(x)I_{A_n^{sf}}(x);
\]

hence, \( v_n(x) \) is \( \mathcal{B} \)-measurable.

Define \( \lambda^* = \inf_n \{ X_n \in A_n^{ss} \cup A_n^{sf} \} \) and \( \mu^* = \inf_n \{ X_n \in A_n^{ss} \cup A_n^{fs} \} \).
Theorem 5 ([74]) Game $G_\alpha$ with payoff function (7) and sets of strategies $\Lambda^N$ and $M^N$ for Player 1 and 2, respectively, have a solution. Pair $(\lambda^*, \mu^*)$ is the equilibrium point and $v_0(x)$ is the value of the game.

Let us construct an equilibrium pair for game $G$. Define (see [76])

$$
\sigma_{n}^{1*} = \inf\{m > n : S_{N-m}(X_m, X_n) = f(X_m, X_n)\}, \quad (11)
$$
$$
\sigma_{n}^{2*} = \inf\{m > n : S_{N-m}(X_m, X_n) = f(X_n, X_n)\}. \quad (12)
$$

Let $(\lambda^*, \mu^*)$ be an equilibrium point in $G_\alpha$.

Theorem 6 ([74]) Game $G$ has a solution. Pair $(s^*, t^*)$ such that $s^* = (\lambda^*, \{\sigma_{n}^{1*}\})$ and $t^* = (\mu^*, \{\sigma_{n}^{2*}\})$ is the equilibrium point. The value of the game is $v_0(x)$.

Proof. Let

$$
\tau_1^* = \lambda^* I_{(\lambda^* < \mu^*)} + (\lambda^* I_{(\lambda^* \leq \mu^*)} + \sigma_{\mu}^{1*} I_{(\lambda^* > \mu^*)}) I_{(\lambda^* = \mu^*)} + \sigma_{\lambda}^{1*} I_{(\lambda^* > \mu^*)}
$$

and

$$
\tau_2^* = \mu^* I_{(\lambda^* > \mu^*)} + (\mu^* I_{(\lambda^* > \alpha^*)} + \sigma_{\lambda}^{2*} I_{(\lambda^* \leq \alpha^*)}) I_{(\lambda^* = \mu^*)} + \sigma_{\lambda}^{2*} I_{(\lambda^* < \mu^*)}.
$$

We obtain by the properties of conditional expectation and by (11) and (12)

$$
\tilde{R}(x, s^*, t^*) = E_x f(X_{\tau_1^*}, X_{\tau_2^*}) = E_x [I_{(\lambda^* < \mu^*)} + (\lambda^* I_{(\lambda^* \leq \mu^*)} + \sigma_{\mu}^{1*} I_{(\lambda^* > \mu^*)}) I_{(\lambda^* = \mu^*)} + \sigma_{\lambda}^{1*} I_{(\lambda^* > \mu^*)}] f(X_{\tau_1^*}, X_{\tau_2^*})
$$

$$
= E_x I_{(\lambda^* < \mu^*)} + (\lambda^* I_{(\lambda^* \leq \mu^*)} + \sigma_{\mu}^{1*} I_{(\lambda^* > \mu^*)}) E_{X_{\lambda^*}} f(X_{\tau_1^*}, X_{\tau_2^*})
$$

$$
+ E_x I_{(\lambda^* > \mu^*)} + (\lambda^* I_{(\lambda^* > \alpha^*)} + \sigma_{\lambda}^{2*} I_{(\lambda^* < \mu^*)}) E_{X_{\mu^*}} f(X_{\tau_1^*}, X_{\tau_2^*})
$$

$$
= R(x, \lambda^*, \mu^*).
$$

Let $t = (\mu, \{\sigma_{n}^{2*}\}) \in \hat{M}^N$. We obtain

$$
\tilde{R}(x, s^*, t^*) = R(x, \lambda^*, \mu^*) \leq R(x, \lambda^*, \mu)
$$

$$
= E_x [I_{(\lambda^* < \mu^*)} + (\lambda^* I_{(\lambda^* \leq \mu^*)} + \sigma_{\mu}^{1*} I_{(\lambda^* > \mu^*)}) H(\lambda^*, X_{\lambda^*})
$$

$$
+ I_{(\lambda^* > \mu^*)} + (\lambda^* I_{(\lambda^* > \alpha^*)} + \sigma_{\lambda}^{2*} I_{(\lambda^* < \mu^*)}) H(\mu, X_{\mu})]
$$

$$
= E_x [I_{(\lambda^* < \mu^*)} + (\lambda^* I_{(\lambda^* \leq \mu^*)} + \sigma_{\mu}^{1*} I_{(\lambda^* > \mu^*)}) E_{X_{\lambda^*}} f(X_{\tau_1^*}, X_{\tau_2^*})
$$

$$
+ I_{(\lambda^* > \mu^*)} + (\lambda^* I_{(\lambda^* > \alpha^*)} + \sigma_{\lambda}^{2*} I_{(\lambda^* < \mu^*)}) E_{X_{\mu}} f(X_{\tau_1^*}, X_{\tau_2^*})]
$$

$$
= E_x f(X_{s^*}, X_{t}) = \tilde{R}(x, s^*, t)
$$
Similarly one can show that for every \( s \in \tilde{A}^N \) we have \( \tilde{R}(x, s, t^*) \leq \tilde{R}(x, s^*, t^*) \). Hence, \((s^*, t^*)\) is the equilibrium pair for \( G \).

\[ \Box \]

2.1.2 The Best vs. the Best or the Second Best Game

Two employers, Player 1 and Player 2, are to view a group of \( N \) applicants for a vacancies in their enterprises sequentially. Each of the applicants has some characteristic unknown to the employer. Let \( \mathcal{K} = \{x_1, x_2, \ldots, x_N\} \) be the set of characteristics, assuming that the values are different. The employer observes a permutation \( \eta_1, \eta_2, \ldots, \eta_N \) of the elements of \( \mathcal{K} \) sequentially. We assume that all permutations are equally likely. Let \( Z_k \) denote the absolute rank of the object with the characteristics \( \eta_k \), i.e.,

\[ Z_k = \min \{ r : \eta_k = \bigwedge_{1 \leq i_1 < \ldots < i_r \leq N} \bigvee_{1 \leq j \leq r} \eta_{i_j} \}, \]

(\( \bigwedge, \bigvee \) denote minimum and maximum, respectively). The object with the smallest characteristics has the rank 1. The decisions of the employer at each time \( n \) are based on the relative ranks \( Y_1, Y_2, \ldots, Y_N \) of the applicants and the previous decisions of the opponent, where

\[ Y_k = \min \{ r : \eta_k = \bigwedge_{1 \leq i_1 < \ldots < i_r \leq k} \bigvee_{1 \leq j \leq r} \eta_{i_j} \}. \]

The random priority assignment model is applied when both players want to accept the same applicant. We assume that \( \alpha_n = p, p \in [0, 1] \) for every \( n \). If the applicant is viewed the employer must either accept or reject her. Once accepted the applicant cannot be rejected, once rejected cannot be reconsidered. Each employer can accept at most one applicant. The aim of Player 1 is to accept BA and Player 2 is to accept BOS but a better one than that chosen by the opponent. Both players together can accept at most two objects. It makes the problem resembling to the optimal double stop of Markov process (cf. [78], [79], [76]). It is a generalization of the optimal choice problem. We adopt the following payoff function here. The player obtains +1 from another if he has chosen the required applicant, −1 when the opponent chosen, and 0 otherwise.

Let us describe the mathematical model of the problem. With sequential observation of the applicants we connect the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). The elementary events are a permutation of the elements of \( \mathcal{K} \) and the probability measure \( \mathbb{P} \) is the uniform probability on \( \Omega \). The observable sequence of relative ranks \( Y_k, k = 1, 2, \ldots, N \) defines a sequence
of $\sigma$-fields $\mathcal{F}_k = \sigma(Y_1, \ldots, Y_k)$, $k = 1, 2, \ldots, N$. The random variables
$Y_k$ are independent and $P(Y_k = i) = 1/k$. Denote by $S^N$ the set of all
Markov times $\tau$ with respect to the $\sigma$-fields $\{\mathcal{F}_k\}_{k=1}^N$. The problem
considered can be formulated as follows. For $s \in \bar{\Lambda}^N$ and $t \in \bar{M}^N$ denote
$A_1 = \{\omega : X_{\tau_1} = 1\}$ and $A_2 = \{\omega : X_{\tau_2} = 1\} \cup \{\omega : X_{\tau_2} = 2, X_{\tau_1} \neq 1\}$. Define
the payoff function $g(s, t) = I_{A_1} - I_{A_2}$ and the expected payoff
$G(s, t) = E g(s, t)$. We are looking for $(s^*, t^*)$ such that for every $s \in \bar{\Lambda}^N$
and $t \in \bar{M}^N$

$$G(s, t^*) \leq G(s^*, t^*) \leq G(s^*, t).$$

It is obvious that the essential decisions of the players can be taken
when applicants with relative rank 1 or 2 have appeared. We will call
them candidates. For further consideration it is convenient to define the
following random sequence $(W_k)_{k=1}^N$. Let $W_1 = (1, Y_1) = (1, 1)$, $\rho_1 = 1$. Define

$$\rho_t = \inf \{r > \rho_{t-1} : Y_r \in \{1, 2\}, \ t > 1, \$$

$$(\inf \emptyset = \infty) \text{ and } W_t = (\rho_t, Y_{\rho_t}).$$. If $\rho_t = \infty$, then we put $W_t = (\infty, \infty)$.
Markov chain $(W_t, \mathcal{G}_t, \mathcal{P}_{(1,1)})_{t=1}^N$ with state space $\mathcal{E} = \{(s, l) : l \in \{1, 2\}, s = 1, 2, \ldots, N\} \cup \{(\infty, \infty)\}$ and $\mathcal{G}_t = \sigma(W_1, W_2, \ldots, W_t)$ is homogeneous.
One step transition probabilities are following.

$$p(r, s) = \mathcal{P}\{W_{t+1} = (s, l_s) \mid W_t = (r, l_r)\}$$

$$= \begin{cases}
\frac{1}{2} & \text{if } r = 1, s = 2, \\
\frac{r(r-1)}{s(s-1)(s-2)} & \text{if } 2 \leq r < s, \\
0 & \text{if } r \geq s \text{ or } (r = 1, s \neq 2),
\end{cases}$$

$p(\infty, \infty) = 1$, $p(r, \infty) = 1 - 2 \sum_{s=r+1}^N p(r, s)$ for $l_s, l_r \in \{1, 2\}$ and $1 \leq r \leq s \leq N$. We will call this Markov chain the auxiliary Markov chain
(AMC).

The solution of the two decision makers problem will use partially
the solution of the problem of choosing BOS (see [70], [80], [68]). The
problem can be treated as an optimal stopping problem for AMC with
the following payoff function:

$$f_{\text{BOS}}(r, l_r) = \begin{cases}
\frac{r(2N-r-1)}{N(N-1)} & \text{if } l_r = 1, \\
\frac{r(r-1)}{N(N-1)} & \text{if } l_r = 2.
\end{cases}$$

Let $T^N = \{\tau \in S^N : \tau = \tau \Rightarrow Y_{\tau} \in \{1, 2\}\}$. It is a set of stopping times
with respect to $\mathcal{G}_t$, $t = 1, 2, \ldots, N$. We search $\tau^* \in S^N$ such that

$$\mathcal{P}\{Z_{\tau^*} \in \{1, 2\}\} = \sup_{\tau \in S^N} \mathcal{P}\{Z_\tau \in \{1, 2\}\} = \sup_{\sigma \in T^N} E_{(1,1)} f_{\text{BOS}}(W_\sigma).$$
Denote $\Gamma(r, s) = \{(t, l_t) : t > r, l_t = 1\} \cup \{(t, l_t) : t > s, l_t = 2\}$. Let $r < s$ and $c(r, s) = E_{(r, l_r)} f_{BOS}(W_\sigma)$, where $\sigma = \inf\{t : W_t \in \Gamma(r, s)\}$. Denote $c(r) = E_{(r, l_r)} f_{BOS}(W_{\sigma_1}) = \frac{2^r \binom{N-r}{N-1}}{N}$, where $\sigma_1 = \inf\{t : W_t \in \Gamma(r, r)\}$. We have

$$c(r, s) = \frac{r}{N(N-1)} \sum_{i=r+1}^{s-1} \frac{2N-i-1}{i-1} + \frac{r}{s-1} c(s-1)$$

(16)

for $r < s$, $s = 1, 2, \ldots, N$ ($\sum_r = 0$ if $s < r$). Define $r_a = \inf\{1 \leq r \leq N : f_{BOS}(r, 2) \geq c(r, r)\}$ and $r_b = \inf\{1 \leq r \leq r_a : f_{BOS}(r, 1) \geq c(r, r_a)\}$. Denote

$$\tilde{c}_{BOS}(r, l_r) = \sup_{\tau \in S_{N+r+1}} P\{Z_r \in \{1, 2\} \mid Y_r = l_r\}.$$ 

We have

$$\tilde{c}_{BOS}(r, l_r) = \tilde{c}_{BOS}(r) = \begin{cases} 
  c(r) & \text{if } r_a \leq r \leq N, \\
  c(r, r_a) & \text{if } r_b \leq r < r_a, \\
  c(r_b-1, r_a) & \text{if } 1 \leq r < r_b.
\end{cases}$$

(17)

The optimal stopping time for the one decision maker problem of choosing BOS is $\sigma^* = \inf\{t : W_t \in \Gamma(r_b, r_a)\} \in T^N$ or $r^* = \inf\{r : (r, Y_r) \in \Gamma(r_b, r_a)\} \in S^N$. We have $a = \lim_{N \to \infty} \frac{r_a}{N} = \frac{2}{3}$, $b = \lim_{N \to \infty} \frac{r_b}{N} \approx 0.3470$ and $\lim_{N \to \infty} \tilde{c}_{BOS}(1) \approx 0.5736$ (cf. [81], [80], [82]).

To solve the two-person competitive stopping problem described at the beginning of the section, let us perform a strategy of the players when one of them accepts some observation at moment $r$ with relative rank $Y_r = l_r$. Since the aims of the players are different we have to consider independently the situation when Player 1 has stopped as the first and when Player 2 has done it. We introduce useful notation

$$h_{ik}(r, l_r) = P(A_k | \tau_i = r, \tau_j > \tau_i, Y_r = l_r)$$

for $k, i, j = 1, 2$, $i \neq j$, $r = 1, 2, \ldots, N$, $l_r = 1, 2$.

Let Player 1 stop the process as the first at the moment $r$ on the object with $Y_r = l_r$. As he wants to accept the object with the absolute rank 1, it is obvious that he will stop on the relatively first object. He will also accept probably, in some circumstances, the relatively second object to disturb Player 2 in the realization of his aims. We will see that this supposition is true. Player 2 staying alone will use a strategy $\sigma^2_r = \varsigma^*(r, l_r)$ defined by

$$h(r, l_r) = E_{(r, l_r)} g((r, \sigma^1_\mu), (\mu, \varsigma^*(r, l_r))) = \inf_{\sigma \in S_{N+r+1}} E_{(r, l_r)} g((r, \sigma^1_\mu), (\mu, \sigma)).$$

(18)
where the expectation is taken with respect to $P_{(r, l_r)}$ of AMC. To perform strategy $\zeta^*(r, l_r)$ let us consider the possible essential situations. Let $W_t = (r, l_r)$. Since Player 2 minimizes his expected loss (cf. (18)), he can do it by stopping on some object with relative rank 1 or 2. If $l_r = 1$, then he cannot change the payoff stopping on the objects with relative rank 2 before another one having relative rank 1 has appeared. Let $W_u = (m, 1)$ and $W_u = (n, 2)$ for $u = t + 1, t + 2, \ldots, s - 1$. Player 1 can be the winner in this case if $W_{s+1} = (\infty, \infty)$ and Player 2 does not accept the $m$th object. We see that it is the first moment after the accepting decision of Player 1 when Player 2 can change the gain of Player 1. We want to know if it is optimal to stop at $(m, 1)$ for Player 2. If he stops, he has $-1$ with the probability $f_{BOS}(m, 1)$ (see (15)). When he passes over and he will behave optimally in future, he has $-1$ with probability $\bar{c}_{BOS}(m)$. Since he minimizes his loss, his optimal strategy in $(m, 1)$ is the same as in the mentioned one player problem. If it happens that $n < m < r_b$, then according to the optimal strategy in the one player choosing BOS problem, $m$th object will not be accepted and Player 2 will behave according to $\sigma^*$. It means Player 1 will have $+1$ if $n$th object is the best or it happens that his candidate is the absolutely second and the best one will not be chosen by Player 2 (because she has appeared before $r_b$). Hence, by (14), (17), and (15) we have

$$h_{12}(r, l_r) = P \{ A_2 \mid \tau_1 = r, \tau_2 > \tau_1, Y_r = l_r \}$$

$$= \left\{ \begin{array}{ll} \sum_{s=r+1}^{N} \frac{r}{s(s-1)} \max \{ \frac{s(2N-s-1)}{N(N-1)}, \bar{c}_{BOS}(s) \} & \text{if } l_r = 1, \\
\bar{c}_{BOS}(r) & \text{if } l_r = 2, \end{array} \right.$$  

where $\sigma^*_n = \zeta^*(n, Y_n)$, $s = (\lambda, \{\sigma^*_n\})$ and $t = (\mu, \{\sigma^*_n\})$. The optimal strategy $\zeta^*$, after the first acceptance has been done at moment $r$ on $Y_r = l_r$ has the form

$$\zeta^*(r, l_r) = \left\{ \begin{array}{ll} \vartheta_r & \text{if } \vartheta_r \geq r_b, \\
\sigma^*_r & \text{if } \vartheta_r < r_b \end{array} \right. \quad \text{for } l_r = 1,$$  

$$\sigma^*_r \quad \text{for } l_r = 2,$$  

(19)

where $\vartheta_r = \inf \{ s > r : Y_s = 1 \}$ and $\sigma^*_r = \inf \{ s > r : (s, Y_s) \in \Gamma(r_b, r_a) \}$. Consequently,

$$h(r, l_r) = h_{11}(r, l_r) - h_{12}(r, l_r),$$

where we have $h_{11}(r, l_r) = \frac{r}{N}$ for $l_r = 1$ and 0 otherwise.

Let us assume that Player 2 has stopped the process as the first on some object at moment $r$ with relative rank $Y_r = l_r$. Player 1 will use a
strategy \( \sigma^*_t = \delta^*(r, l_r) \). The strategy \( \delta^*(r, l_r) \) is such that

\[
H(r, l_r) = E_{(r, l_r)} g_1(\langle \lambda, \delta^*(r, l_r) \rangle, (r, \sigma^2_\lambda)) = \sup_{\sigma \in S^N_{r + 1}} E_{(r, l_r)} g_1(\langle \lambda, \sigma \rangle, (r, \sigma^2_\lambda)).
\]

Let \( W_t = (r, l_r) \). Since Player 1 maximizes his expected gain and he would like to choose the best object, he can do it by stopping on some object with relative rank 1. Denote \( \tilde{c}_{BA}(r) = \sup_{\tau \in S^N_{r + 1}} P\{Z_\tau = 1 | Y_r = l_r\} \), \( r_c = \inf\{1 \leq r \leq N : \sum_{i=r+1}^N \frac{1}{i-1} \leq 1\} \) and \( r^*_c = \inf\{s > r : Y_s = 1, s \geq r_c\} \). The optimal strategy \( \delta^* \) of Player 1, after the first acceptance done at the moment \( r \) on \( Y_r = l_r \) by Player 2, has the form

\[
\delta^*(r, l_r) = \begin{cases} 
\vartheta_r & \text{if } \vartheta_r \geq r_c, \\
\tau^*_r & \text{if } \vartheta_r < r_c
\end{cases} \quad \text{for } l_r = 1, \quad (20)
\]

where \( \vartheta_r \) is the first moment after \( r \) when \( Y_r = 1 \). We have

\[
H(r, l_r) = h_{21}(r, l_r) - h_{22}(r, l_r),
\]

where

\[
h_{21}(r, l_r) = \sum_{s=r+1}^N p(r, s)\max\{s, \tilde{c}_{BA}(r)\} + \tilde{c}_{BA}(r) = \tilde{c}_{BA}(r)
\]

and

\[
h_{22}(r, l_r) = P\{A_2 \mid \tau_2 = r, \tau_1 > \tau_2, Y_r = l_r\}
\]

\[
= \begin{cases} 
\frac{r}{N} + \left( \sum_{s=r+1}^{r-1} \frac{r(s-1)}{N(N-1)} \right) & \text{if } r \geq r_c \\
\frac{r(r-1)}{N(N-1)} & \text{if } r < r_c
\end{cases} \quad \text{for } l_r = 1,
\]

\[
= \begin{cases} 
\frac{r}{N} + \left( \sum_{s=r+1}^{r-1} \frac{r(s-1)}{N(N-1)} \right) & \text{if } r \geq r_c \\
\frac{r(r-1)}{N(N-1)} & \text{if } r < r_c
\end{cases} \quad \text{for } l_r = 2
\]

Denote \( h_p(r, l_r) = ph(r, l_r) + (1-p)H(r, l_r) \). Define \( r_d = \min\{1 \leq r \leq N : h(r, 2) \geq H(r, 2)\} \) and \( r_\ell = \min\{1 \leq r \leq r_d : h(r, 1) \geq H(r, 1)\} \). During the recursive construction of \( \tilde{v}(r, l_r; p) \) and the strategy according to Theorem 5 and 6 (see also (9)) for a large \( N \) we get that there exist \( r_{\nu(p)} = \min\{r < r_d : H(r, 2) \leq \tilde{v}(r; p)\} \) and \( \tilde{p}_1 = \min\{0 \leq p \leq 1 : h(1, 1) < \tilde{v}(\ell; p)\} \). For \( p \geq \tilde{p}_1 \) there exists \( r_{\kappa(p)} = \min\{r \leq r_\ell : H(r, 1) \leq \tilde{v}(r; p)\} \) and for \( p < \tilde{p}_1 \) there exists \( r_{\kappa(p)} = \min\{r \leq r_\ell : h(r, 1) \geq \tilde{v}(r; p)\} \). These points \( r_d, r_\ell, r_{\nu(p)}, r_{\kappa(p)} \) are such that

\[
v(r, l_r; p) = \begin{cases} 
h_p(r, l_r) & \text{if } (r, l_r) \in B_{r\ell}N(1) \cup B_{rd}N(2), \\
H(r, l_r) & \text{if } (r, l_r) \in B_{r\nu(p)}r_{d-1}(2), \\
H(r, l_r)1_{\{p \geq \tilde{p}_1\}} & \text{if } (r, l_r) \in B_{r\kappa(p)}r_{\ell-1}(1), \\
\tilde{v}(r; p) & \text{if } (r, l_r) \in B_{1r_{\kappa(p)}-1}(1) \cup B_{1r_{\nu(p)}-1}(2)
\end{cases} \quad (21)
\]
where

\[ \bar{v}(r; p) = Tv(r, l_r; p) = \begin{cases} 
  w(r, r + 1, r + 1, r + 1; p) & \text{if } r_d \leq r \leq N, \\
  w(r, r + 1, r + 1, r_d; p) & \text{if } r_{\nu(p)} \leq r < r_d, \\
  w(r, r + 1, r_{\nu(p)}, r_d; p) & \text{if } r_t \leq r < r_{\nu(p)}, \\
  w(r_t, r_{\nu(p)}, r_d; p) & \text{if } r_{\kappa(p)} \leq r < r_t, \\
  w(r_{\kappa(p)}, r_t, r_{\nu(p)}, r_d; p) & \text{if } 1 \leq r < r_{\kappa(p)} 
\end{cases} \]

and

\[ w(r, s, t, u; p) = \sum_{j=r+1}^{s} \frac{r}{j(j-1)} [H(j, 1)I_{[p \geq \tilde{p}_1]} + h(j, 1)I_{[p < \tilde{p}_1]}] \\
+ \sum_{j=s}^{t-1} \frac{r}{j(j-1)} h_p(j, 1) \\
+ \sum_{j=t}^{u-1} \frac{r(t-2)}{j(j-1)(j-2)} [h_p(j, 1) + H(j, 2)] \\
+ \sum_{j=u}^{N} \frac{r(t-2)}{j(j-1)(j-2)} [h_p(j, 1) + h_p(j, 2)] \]

for \( r \leq s \leq t \leq u \). The optimal first stop strategy is given by sets

\[ A_t^{ss} = B_{r_tN}(1) \cup B_{r_dN}(2), A_t^{sf} = (I_{[p \geq \tilde{p}_1]}B_{r_{\nu(p)}r_{t-1}(1)})) \cup B_{r_{\nu(p)}r_d-1}(2), A_t^{sf} = I_{[p < \tilde{p}_1]}B_{r_{\nu(p)}r_{t-1}(1)}, A_t^{ff} = E \backslash (A_t^{ss} \cup A_t^{sf} \cup A_t^{sf}), t = 1, 2, \ldots, N. \]

Here we adopt convention that for every set \( A \) we have \( 1 \cdot A = A \) and \( 0 \cdot A = \emptyset \), where \( \emptyset \) the empty set.

The function \( w(r, s, t, u; p) \) depends also on \( r_b \) and \( r_c \). Let \( r \leq s \leq t \leq u \). When \( N \to \infty \) and \( \frac{r}{N} \to x_1, \frac{s}{N} \to x_2, \frac{t}{N} \to y_1, \frac{u}{N} \to y_2 \) we get

\[ \bar{w}(x_1, x_2, y_1, y_2; p) = \lim_{N \to \infty} w(r, s, t, u; p) \]

\[ = \bar{w}_{21}(x_1, x_2, y_1, y_2; p)I_{[p \geq \tilde{p}_1]} \\
+ \bar{w}_{22}(x_1, x_2, y_1, y_2; p)I_{[p < \tilde{p}_1]}, \]

where

\[ \bar{w}_{21}(x_1, x_2, y_1, y_2; p) = x_1[(\ln \frac{x_1}{x_2} - \frac{1}{2}((\ln x_2)^2 - (\ln x_1)^2)I_{[x_1 > c]}) \\
- (\frac{1}{2} + \ln x_2 + \frac{1}{2}(\ln x_2)^2)I_{[x_1 \leq c]}] \\
+ \frac{x_1I_{[x_1 > c]} + cI_{[x_1 \leq c]}}{x_2}\bar{w}_1(x_2, y_1, y_2; p), \]
\[
\hat{w}_{22}(x_1, x_2, y_1, y_2; p) = x_1 \left( (\ln x_2)^2 - (\ln x_1)^2 + x_1 - x_2 + 2 \ln \frac{x_2}{x_1} I_{\{x_1 > b\}} \right)
+ ((4 - 2b + 2 \ln b) \ln \frac{b}{x_1} + (2 - b)(x_1 - b)
+ (\ln x_2)^2 - (\ln b)^2 - x_2 + b + 2 \ln \frac{x_2}{b} I_{\{x_1 \leq b\}}) \right]
+ \frac{x_1 I_{\{x_1 > b\}} + b I_{\{x_1 \leq b\}}}{x_2} w_1(x_2, y_1, y_2; p)
\]

and

\[
\hat{w}_1(x_2, y_1, y_2; p) = x_2 [(3p - 1) \ln \frac{y_1}{x_2} + \frac{3p - 1}{2} ((\ln y_1)^2 - (\ln x_2)^2) - p(y_1 - x_2)]
+ x_2 y_1 [3(2p - 1)(\frac{1}{y_1} - \frac{1}{y_2}) + (3p - 2)(\frac{\ln y_1}{y_1} - \frac{\ln y_2}{y_2})]
+ (1 + p) \ln \frac{y_1}{y_2} + y_1^2 [(p - 1)(\frac{1}{y_2} - 1)]
+ (4p - 2)(\frac{\ln y_2}{y_2} + \frac{1}{y_2} - 1) - (2p - 1) \ln y_2].
\]

Parameter \( p_1 \) is asymptotic equivalent of \( \hat{p}_1 \). The value of \( p_1 \) can be determined as the solution of some equation which will be given later.

Let \( d = \lim N \rightarrow \infty \frac{r_d}{N} \approx 0.7587 \) and \( \ell = \lim N \rightarrow \infty \frac{r_\ell}{N} \approx 0.4237 \). We have \( \nu(p) = \lim N \rightarrow \infty \frac{r_\nu(p)}{N} \) is the solution of the equation \( \hat{w}_1(\nu, \nu, d; p) = \hat{H}(\nu, 2) \) in \([\ell, d]\). Now we can determine \( p_1 \) as the solution of the equation \( \hat{w}_1(\ell, \nu(p), d; p) = \hat{H}(\ell, 1) \) with respect to \( p \) in \([0, 1]\). Such solution exists since \( \hat{w}_1(\ell, \nu(p), d; p) \) is non-decreasing function of \( p \) and \( \hat{H}(\ell, \nu(1), d; 1) < \hat{w}_1(\ell, \nu(1), d; 1) \). We have \( p_1 \approx 0.5659 \).

Determine \( \kappa(p) = \lim N \rightarrow \infty \frac{r_\kappa(p)}{N} \). The decision point \( \kappa(p) \) is the solution of the equation \( \hat{w}(\kappa, \ell, \nu(p); p) = \hat{h}(\kappa, 1) I_{\{p < p_1\}} + \hat{H}(\kappa, 1) I_{\{p \geq p_1\}} \):

\[
\hat{v}(x; p) = \lim N \rightarrow \infty \hat{v}(r; p) = \begin{cases}
\hat{w}(x, x, x, x; p) & \text{if } d \leq x \leq 1,
\hat{w}(x, x, x, d; p) & \text{if } \nu(p) \leq x < d,
\hat{w}(x, x, \nu(p), d; p) & \text{if } \ell \leq x < \nu(p),
\hat{w}(x, \ell, \nu(p), d; p) & \text{if } \kappa(p) \leq x < \ell,
\hat{w}(\kappa(p), \ell, \nu(p), d; p) & \text{if } 0 \leq x < \kappa(p).
\end{cases}
\]

We can formulate the following:

**Theorem 7** For \( N \) large enough in the competitive two-person problem of choosing the best vs the best or the second best applicant but a better than the opponent, the asymptotically optimal strategy of the first stop is
described by the sets $A_{*}^{s}$, $A_{t}^{s}$, $A_{t}^{f}$ and $A_{t}^{ff}$. The second step is according
to $\zeta^{*}$ given by (19) for Player 2 and $\delta^{*}$ given by (20) for Player 1. The
value function of the problem is given by (21), the expected value with
respect to $P_{(r,\zeta)}$ of AMC by (22) and its limit by (23).

More examples of zero-sum random priority games and further consi-
erations can be found in [74].

2.2 Nonzero-Sum Random Priority Stopping Game

A construction of Nash equilibria for a random priority finite horizon two-
person nonzero-sum game with stopping of Markov process is considered
in this section. Let $(X_{n}, F_{n}, P_{x})_{n=0}^{N}$ be a homogeneous Markov process
defined on a probability space $(\Omega, F, P)$ with a state space $(E, B)$. At
each moment $n = 1, 2, ... , N$ the decision makers are able to observe the
Markov chain. Each player has his utility function $g_{i} : E \to \mathbb{R}, i = 1, 2,$
and at each moment $n$ each decides separately if he accepts or rejects
the realization $x_{n}$ of $X_{n}$. We admit $g_{i}$ are measurable and bounded. If it
happens that both players have selected the same moment $n$ to accept $x_{n}$,
then the similar random assignment mechanism, as in the zero-sum game
model described in Section 2.1, is applied. If a player has not chosen any
realization of Markov process, he gets $g_{i}^{*} = \inf_{x \in E} g_{i}(x)$. The aim of each
player is to choose a realization which maximizes his expected utility.

The problem with permanent priority for Player 1 (i.e. $\alpha_{n} = 1, n =
1, 2, \ldots$) had been solved by Ferenstein [62]. This game is also strictly
connected with optimal stopping of stochastic processes. The Bellman
principle (see Kuhn [83], Rieder [84]) as well as the ideas of Yasuda [47]
and Ohtsubo [50] will be adopted to this random priority game model.
On the basis of this approach, we solve an example in which we deal with
two-person time sequential nonzero-sum game version of the best choice
problem (the secretary problem). This example is a generalization of a
game model of the problem considered by Fushimi [65].

A tip from the two-person nonzero-sum generalized secretary problem
with fixed priority has been given in Szajowski [69] and from the random
priority game in Sakaguchi [63]. A very interesting illustration of the
problem in stopping games are models of two-person best choice problems
considered by Fushimi [65]. One of them is generalized in this paper and
can be described as follows.

Two companies (Player 1 and Player 2) interview a sequence of applicants
one by one (as in the best choice problem which has been recalled
above) every morning independently of the other company, and the results
of the interviews are communicated to the applicant in the afternoon. If
only one of the companies decides to accept the applicant, she agrees to
this offer at once, the other company is informed of this fact and continues
the interviewing process. If, on the other hand, both companies decide
to accept the applicant, she selects one of them with equal probabilities
and the other company can continue interviewing and employ another
applicant. In Fushimi (1981) the threshold strategies for the players were
admitted. It was shown that equilibrium strategies for players in the
model are different. One of the players should behave more hastily than
in the original secretary problem and he should start solicitation at .2865
for the limiting version of the problem. There are two Nash equilibria in
the considered set of strategies for this game with values (.2865,.2963)
and (.2963,.2865), respectively.

The following generalization of the above problem has been considered
in Section 2.2.2. It is assumed that if both companies want to accept the
same applicant, Player 1 is selected with fixed probability \( \alpha \), Player 2
with probability \( 1 - \alpha \), \( \alpha \in [0,1] \), and the player who has not been
chosen continues interviewing and employs another applicant. Also a
more general set of strategies is admitted. This particular game problem
is presented as interesting per se. The mathematical model of the above
formulated problem will be presented and equilibria for each \( \alpha \) will be
derived. The problem need modified set of strategies with respect to
those applied in the zero-sum random priority game (see Section 2.1).
More details are in [86].

2.2.1 The Payoff Functions and Strategies

In the problem of optimal stopping, the basic class of strategies \( T^N \) are
Markov times with respect to \( \sigma \)-fields \( \{ \mathcal{F}_n \}_{n=1}^N \). We admit that \( \mathbb{P}(\tau \leq N) < 1 \) for some \( \tau \in T^N \). The class of strategies described in Section 2.1
is not sufficient in the nonzero-sum stopping game. To extend the class
of strategies we consider a class of randomized stopping times. It is
assumed that the probability space is rich enough to admit the following
constructions.

**Definition 2** (see Yasuda [47]) A strategy for each player is a random
sequence \( p = (p_n) \in \mathcal{P}^N \) or \( q = (q_n) \in \mathcal{Q}^N \) such that, for each \( n \),

(i) \( p_n, q_n \) are adapted to \( \mathcal{F}_n \);

(ii) \( 0 \leq p_n, q_n \leq 1 \) a.s. .

If each random variables equals either 0 or 1 we call it a pure strategy.
Let $A_1, A_2, \ldots, A_N$ and $B_1, B_2, \ldots, B_N$ be i.i.d.r.v. of the uniform distribution on $[0, 1]$ and independent of Markov process $(X_n, F_n, P_x)^{N}_{n=0}$. Let $\mathcal{H}_n$ be the $\sigma$-field generated by $F_n$, $\{A_1, A_2, \ldots, A_n\}$ and $\{B_1, B_2, \ldots, B_n\}$. A randomized Markov time $\lambda(p)$ for strategy $p = (p_n) \in \mathcal{P}^N$ and $\mu(q)$ for strategy $q = (q_n) \in \mathcal{Q}^N$ are defined by $\lambda(p) = \inf\{N \geq n \geq 1 : A_n \leq p_n\}$ and $\mu(q) = \inf\{N \geq n \geq 1 : B_n \leq q_n\}$, respectively. We denote by $\Lambda^N$ and $\hat{M}^N$ the sets of all randomized strategies of Player 1 and Player 2. Clearly, if each $p_n$ is either zero or one, then the strategy is pure and $\lambda(p)$ is in fact an $\{F_n\}$-Markov time. In particular an $\mathcal{F}_n$-Markov time $\lambda$ corresponds to the strategy $p = (p_n)$ with $p_n = I(\lambda_n=1)$, where $I_A$ is an indicator function of the set $A$.

Denote $T_k^N = \{\tau \in T^N : \tau \geq k\}$. One can define the set of strategies $\Lambda^N = \{(p, \{\sigma_n^1\}) : p \in \mathcal{P}^N, \{\sigma_n^1\} \in T_{n+1}^N \text{ for every } n\}$ and let $\hat{M}^N = \{(q, \{\sigma_n^2\}) : q \in \mathcal{Q}^N, \{\sigma_n^2\} \in T_{n+1}^N \text{ for every } n\}$ for Player 1 and Player 2, respectively.

Let $\xi_1, \xi_2, \ldots$ be i.i.d.r.v. uniformly distributed on $[0, 1]$ and independent of $\bigvee_{n=1}^N \mathcal{H}_n$ and the lottery is given by $\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_N)$. Denote $\mathcal{H}_n = \sigma(\mathcal{H}_n, \xi_1, \xi_2, \ldots, \xi_n)$ and let $\hat{T}^N$ be the set of Markov times with respect to $\{\mathcal{H}_n\}^N_{n=0}$. For every pair $(s, t)$ such that $s \in \Lambda^N$, $t \in \hat{M}^N$ we define $\tau_1(s, t) = \lambda(p)I_{(\lambda(p)<\mu(q))} + (\lambda(p)I_{(\lambda(p)\leq \sigma_{\lambda(p)})} + \sigma_{\lambda(p)}^1I_{(\lambda(p)>\sigma_{\lambda(p)})})I_{(\lambda(p)=\mu(q))} + \sigma_{\mu(q)}^1I_{(\lambda(p)>\mu(q))}$ and $\tau_2(s, t) = \mu(q)I_{(\lambda(p)<\mu(q))} + (\mu(q)I_{(\lambda(p)\leq \sigma_{\mu(q)})} + \sigma_{\mu(q)}^2I_{(\lambda(p)>\sigma_{\mu(q)})})I_{(\lambda(p)=\mu(q))} + \sigma_{\lambda(p)}^2I_{(\lambda(p)<\mu(q))}$. The random variables $\tau_1(s, t), \tau_2(s, t) \in \hat{T}^N$ for every $s \in \Lambda$ and $t \in \hat{M}$.

**Definition 3** The Markov times $\tau_1(s, t)$ and $\tau_2(s, t)$ are selection times of Player 1 and Player 2 when they use strategies $s \in \Lambda$ and $t \in \hat{M}$, respectively, and the lottery is $\bar{\alpha}$.

For each $(s, t) \in \Lambda^N \times \hat{M}^N$ and given $\bar{\alpha}$ the payoff function for the $i$th player is defined as $f_i(s, t) = g_i(X_{\tau_i(s, t)})$. Let $\tilde{R}_i(x, s, t) = E_x f_i(s, t) = E_x g_i(X_{\tau_i(s, t)})$ be the expected gain of $i$th player if the players use $(s, t)$. We have defined the game in normal form $(\Lambda^N, \hat{M}^N, \tilde{R}_1, \tilde{R}_2)$. This random priority game will be denoted $G_{rp}$.

**Definition 4** A pair $(s^*, t^*)$ of strategies such that $s^* \in \Lambda^N$ and $t^* \in \hat{M}^N$ is called a Nash equilibrium in $G_{rp}$ if for all $x \in E$

\[
\begin{align*}
v_1(x) & = \tilde{R}_1(x, s^*, t^*) \geq \tilde{R}_1(x, s, t^*) \text{ for every } s \in \Lambda^N, \\
v_2(x) & = \tilde{R}_2(x, s^*, t^*) \geq \tilde{R}_2(x, s^*, t) \text{ for every } t \in \hat{M}^N.
\end{align*}
\]
Denote \( h_i(n, X_n) = \text{esssup}_{r \in T^i_n} \mathbb{E}_{X_r} g_i(X_r) \) and \( \sigma^{*i} \) a stopping time such that \( h_i(0, x) = \mathbb{E}_x g_i(X_{r^i}) \) for every \( x \in \mathbb{E}, i = 1, 2 \). Let \( \Gamma_n^i = \{ x \in \mathbb{E} : h_i(n, x) = g_i(x) \} \). We have \( \sigma^{*i} = \inf \{ n : X_n \in \Gamma_n^i \} \) (cf. Shiryaev (1978)). Denote \( \sigma^{*k}_i = \inf \{ n > k : X_n \in \Gamma_n^i \} \). Taking into account the above definition of \( G_{trp} \) one can conclude that the Nash values of this game are the same as in the auxiliary game \( G_{wp} \) with the sets of strategies of the players \( \mathcal{P}^N, \mathcal{Q}^N \) and payoff functions (cf. Yasuda (1985))

\[
\varphi_1(p, q) = g_1(X_{\lambda(p)})I_{\lambda(p) < \mu(q)} + \tilde{h}_1(\mu(q), X_{\mu(q)})I_{\lambda(p) > \mu(q)} \\
+ \left[ g_1(X_{\lambda(p)})\alpha_{\lambda(p)} + \tilde{h}_1(\lambda(p), X_{\lambda(p)})(1 - \alpha_{\lambda(p)}) \right] I_{\lambda(p) = \mu(q)},
\]

\[
\varphi_2(p, q) = g_2(X_{\mu(q)})I_{\mu(q) < \lambda(p)} + \tilde{h}_2(\lambda(p), X_{\lambda(p)})I_{\mu(q) > \lambda(p)} \\
+ \left[ g_2(X_{\mu(q)})(1 - \alpha_{\lambda(p)}) + \tilde{h}_2(\lambda(p), X_{\lambda(p)})\alpha_{\lambda(p)} \right] I_{\lambda(p) = \mu(q)},
\]

for each \( p \in \mathcal{P}, q \in \mathcal{Q}, \) where \( \tilde{h}_i(n, X_n) = \text{esssup}_{r \in T^i_{n+1}} \mathbb{E}_{X_r} g_i(X_r) = \mathbb{E}_{X_n} h_i(n + 1, X_{n+1}) \). Denote \( R_i(x, p, q) = \mathbb{E}_x \varphi_i(p, q) \) for every \( x \in \mathbb{E}, i = 1, 2 \).

Let \( \mathcal{P}^N_n = \{ p = (p_n) \in \mathcal{P} : p_1 = \ldots = p_{n-1} = 0, p_n = 1 \} \) and \( \mathcal{Q}^N_n = \{ q = (q_n) \in \mathcal{Q} : q_1 = \ldots = q_{n-1} = 0, q_n = 1 \} \). We will use the following convention: if \( p \in \mathcal{P}^N \), then \( (p_n, p) \) is the strategy belonging to \( \mathcal{P}^N \) in which the \( n \)th coordinate is changed to \( p_n \).

**Definition 5** A pair \( (p^*, q^*) \in \mathcal{P}^N_n \times \mathcal{Q}^N_n \) is called an equilibrium point of \( G_{wp} \) at \( n \) if

\[
v_1(n, X_n) = \mathbb{E}_{X_n} \varphi_1(p^*, q^*) \geq \mathbb{E}_{X_n} \varphi_1(p, q^*) \text{ for every } p \in \mathcal{P}^N_n, \text{ P}_x\text{-a.s.,}
\]

\[
v_2(n, X_n) = \mathbb{E}_{X_n} \varphi_2(p^*, q^*) \geq \mathbb{E}_{X_n} \varphi_2(p^*, q) \text{ for every } q \in \mathcal{Q}^N_n, \text{ P}_x\text{-a.s.}
\]

A Nash equilibrium point at \( n = 0 \) is a solution of \( G_{wp} \).

**Theorem 8** ([86]) *There exists a Nash equilibrium \((p^*, q^*)\) in the game \( G_{wp} \). The Nash value and an equilibrium point can be calculated recursively.*

**Proof.** At moment \( N \) the players play the following bimatrix game

\[
\begin{pmatrix}
(\tilde{g}_1(N, X_N), \tilde{g}_2(N, X_N)) & (g_1(X_N), g_2^*) \\
(g_1^*, g_2(X_N)) & (g_1^*, g_2^*)
\end{pmatrix}
\]

where \( \tilde{g}_1(n, x) = \alpha_n g_1(x) + (1 - \alpha_n)\tilde{h}_1(n, x) \) and \( \tilde{g}_2(n, x) = (1 - \alpha_n)g_2(x) + \alpha_n\tilde{h}_2(n, x) \). This game always has an equilibrium in pure or randomized strategies on \( \{ \omega : X_N = x \} \) for every \( x \in \mathbb{E} \). We denote a Nash...
equilibrium in $\mathcal{P}_N \times \mathcal{Q}_N$ by \((p_N^*, q_N^*)\) and the corresponding Nash value by \((v_1(N, x), v_2(N, x))\). Let us assume that an equilibrium \((p^*, q^*) \in \mathcal{P}_{n+1} \times \mathcal{Q}_{n+1}\) has been constructed and \((v_1(n + 1, x), v_2(n + 1, x))\) is the Nash value corresponding to this strategy on \(\{\omega : X_n = x\}\). We consider the following bimatrix game

\[
\begin{pmatrix}
(g_1(n, X_n), \tilde{g}_2(n, X_n)) & (g_1(X_n), \tilde{h}_2(n, X_n)) \\
(\tilde{h}_1(n, X_n), g_2(X_n)) & (\tilde{v}_1(n, X_n), \tilde{v}_2(n, X_n))
\end{pmatrix}
\]

(26)

where \(\tilde{v}_j(n, x)\) is such that \(\tilde{v}_j(n, X_n) = \mathbb{E}_{X_n} v_j(n + 1, X_{n+1}), j = 1, 2\).

On the set \(\{\omega : X_n = x\}\) there is at least one equilibrium point in pure or randomized strategies in this bimatrix game. By measurability of \(g_i(x)\) there exists \((p_n^*, q_n^*)\) such that \(p_n^*, q_n^* \in \mathcal{F}_n\) and \((p_n^*, q_n^*)\) is a Nash equilibrium in the above bimatrix game. We are now in a position to show that \(((p_n^*, p^*), (q_n^*, q^*))\) is an equilibrium of \(G_{wp}\) in \(\mathcal{P}_n \times \mathcal{Q}_n\). Let \((p_n, p) \in \mathcal{P}_n\), where \(p \in \mathcal{P}_{n+1}\). By properties of conditional expectation and induction assumption we have \(P_x\)-a.s.

\[
\mathbb{E}_{X_n} \phi_1((p_n, p), (q_n^*, q^*)) = p_n q_n^* g_1(n, X_n) + p_n (1 - q_n^*) g_1(X_n) \\
+ (1 - p_n) q_n^* \tilde{h}_1(n, X_n) \\
+ (1 - p_n) (1 - q_n^*) \mathbb{E}_{X_n} \mathbb{E}_{X_{n+1}} \phi_1(p, q^*) \\
\leq p_n^* q_n^* g_1(n, X_n) + p_n^* (1 - q_n^*) g_1(X_n) \\
+ (1 - p_n^*) q_n^* \tilde{h}_1(n, X_n) \\
+ (1 - p_n^*) (1 - q_n^*) \mathbb{E}_{X_n} v_1(n + 1, X_{n+1}) \\
= v_1(n, X_n).
\]

for each \(x \in \mathbb{E}\). The same is valid for Player 2. This proves the theorem.

\[\square\]

The solution of the game \(G_{rp}\) can be constructed based on the solution \((p^*, q^*)\) of the corresponding game \(G_{wp}\).

**Theorem 9** ([86]) Game \(G_{rp}\) has a solution. The pair \((s^*, t^*)\), where \(s^* = (p^*, \{s_{1n}^*\}) \in \bar{\Lambda}^N\) and \(t^* = (q^*, \{s_{2n}^*\}) \in \bar{M}^N\), is an equilibrium point. The value of the game is \((v_1(0, x), v_2(0, x))\).

In fact, the players play optimally \(G_{rp}\) using a Nash equilibrium strategy from \(G_{wp}\). If the strategy of both players indicates stopping at moment \(n\) and neither player has stopped earlier, then the lottery chooses one of them. The player who has not been selected will accept any future realization according to the adequate optimal strategy in the optimization problem.
2.2.2 Two-Person Best Choice Problem with Random Priority

The solution of the best choice problem (one player game), described in Section 2.1.1, is auxiliary in the solution of the two-person game with random priority. It is used in further consideration. Let us consider the two-person nonzero-sum game with random priority described in Section 2.2 related to the secretary problem. We admit that both players observe Markov chain $W_t$, $t = 1, 2, ...$ and their utility functions $g_j(r) = f(r)$, $j = 1, 2$, $r \in \mathbb{E}$. Let lottery $\bar{\alpha}$ be constant, i.e., $\alpha_i = \alpha$, $i = 1, 2, ..., N$. Denote $\bar{c}(r) = \bar{c}_{BA}(r)$ defined in Section 2.1.1 and $r_a = \inf\{1 \leq r \leq N : \sum_{i=r+1}^{N} \frac{1}{i-1} \leq 1\}$ and $r_a^* = \inf\{s > r : Y_s = 1, s \geq r_a\}$. We have $\tilde{g}_1(r) = \alpha f(r) + (1 - \alpha)\bar{c}(r)$, $\tilde{g}_2(r) = (1 - \alpha)f(r) + \alpha\bar{c}(r)$ and $g_i^* = 0$. Our aim is to determine the equilibria which give the highest and lowest value for Player 1. At first, we construct the highest value Nash equilibrium for Player 1. By analysis of the matrices (26) we have that $p_r^* = q_r^* = 1$ is an equilibrium point for $r \geq r_a$. We have then

$$
\tilde{v}_1(r) = \sum_{i=r+1}^{N} p(r, i)\tilde{g}_1(i)
$$

$$
\tilde{v}_2(r) = \sum_{i=r+1}^{N} p(r, i)\tilde{g}_2(i)
$$

for $j = 1, 2$. For $r = r_a - 1$ we have two pure equilibria in (26) in this case: $(1, 0)$ and $(0, 1)$ and one in randomized strategies. Since for $r < r_a$ we have $f(r) < \bar{c}(r)$, henceforth we can choose $(1, 0)$ at $r = r_a - 1$ and assume for induction that the same strategy is optimal for $r < r_a$. Under this assumption

$$
\tilde{v}_1(r) = \sum_{i=r+1}^{r_a-1} p(r, i)g_1(i) + \sum_{i=r_a}^{N} p(r, i)\tilde{g}_1(i)
$$

$$
\tilde{v}_2(r) = \sum_{i=r+1}^{r_a-1} p(r, i)\bar{c}(i) + \sum_{i=r_a}^{N} p(r, i)\tilde{g}_2(i).
$$

Since $f(r)$ is increasing and $\bar{c}(r)$ is constant for $r < r_a$, the strategy $(1, 0)$ can be used as equilibrium in $r_b \leq r \leq r_a$, where $r_b = \inf\{r < r_a : \tilde{v}_1(r) \leq g_1(r)\}$. Denote $r_u = \inf\{r < r_a : \tilde{v}_2(r) \leq g_2(r)\}$. For large $N$ we have $r_b < r_u$ if $\alpha < \alpha_0 = \min\{\alpha \in [0, 1] : \frac{2}{2 + \alpha} \geq e^{-\frac{1-\alpha}{2}}\} \approx 0.5299$. Denote

$$
w_1(r, s, \alpha) = \sum_{i=r+1}^{s-1} p(r, i)f(i) + \sum_{i=s}^{N} p(r, i)\tilde{g}_1(i)
$$
7. Nonzero-sum stochastic games

\[ w_2(r, s, \alpha) = \sum_{i=r+1}^{s-1} p(r, i) \tilde{c}(i) + \sum_{i=s}^{N} p(r, i) \tilde{g}_2(i). \]

For \( \alpha < \alpha_0 \) we have

\[ (p^*_r, q^*_r) = \begin{cases} 
(1, 1) & \text{if } r \geq r_a, \\
(1, 0) & \text{if } r_b \leq r < r_a, \\
(0, 0) & \text{if } 1 \leq r < r_b, 
\end{cases} \quad (27) \]

and

\[ v_j(r) = \begin{cases} 
\tilde{w}_j(r, r + 1, \alpha) & \text{if } r \geq r_a, \\
\tilde{w}_j(r, r_a, \alpha) & \text{if } r_b \leq r < r_a, \\
\tilde{w}_j(r_b - 1, r_a, \alpha) & \text{if } 1 \leq r < r_b, 
\end{cases} \quad (28) \]

for \( j = 1, 2 \). The value of the game is \((v_1, v_2) = (v_1(1), v_2(1))\). When \( N \to \infty \) such that \( \frac{r}{N} \to x \) we obtain

\[ \hat{v}_j(x) = \lim_{N \to \infty} v_j(r) = \begin{cases} 
\hat{\tilde{w}}_j(x, x, \alpha) & \text{if } x \geq a, \\
\hat{\tilde{w}}_j(x, a, \alpha) & \text{if } b \leq x < a, \\
\hat{\tilde{w}}_j(b, a, \alpha) & \text{if } 0 < x < b, 
\end{cases} \]

where

\[ \hat{\tilde{w}}_1(x, y, \alpha) = -x \ln x + (1 - \alpha)x(\ln y + \frac{(\ln y)^2}{2}), \]

\[ \hat{\tilde{w}}_2(x, y, \alpha) = -x - (1 - \alpha)x \ln y + \alpha x \frac{(\ln y)^2}{2} \]

and \( b = \lim_{N \to \infty} \frac{r_b}{N} = e^{-\frac{3a}{4}} \). The asymptotic value of the game in this equilibrium is

\[ (\hat{v}_1, \hat{v}_2) = (e^{-\frac{3a}{4}}, e^{-1} - \frac{\alpha}{2} e^{-\frac{3a}{4}}). \quad (29) \]

Let \( \alpha \geq \alpha_0 \). Denote

\[ u_1(r, s, t, \alpha) = \sum_{i=r+1}^{s-1} p(r, i) \tilde{c}(i) + \sum_{i=s}^{t-1} p(r, i) f(i) + \sum_{i=t}^{N} p(r, i) \tilde{g}_1(i), \]

\[ u_2(r, s, t, \alpha) = \sum_{i=r+1}^{s-1} p(r, i) f(i) + \sum_{i=s}^{t-1} p(r, i) \tilde{c}(i) + \sum_{i=t}^{N} p(r, i) \tilde{g}_2(i). \]

Similar analysis as above leads to conclusion that

\[ (p^*_r, q^*_r) = \begin{cases} 
(1, 1) & \text{if } r \geq r_a, \\
(1, 0) & \text{if } r_b \leq r < r_a, \\
(0, 1) & \text{if } r_c \leq r < r_b, \\
(0, 0) & \text{if } 1 \leq r < r_c, 
\end{cases} \quad (30) \]
and

\[ v_j(r) = \begin{cases} 
  u_j(r, r + 1, r + 1, \alpha) & \text{if } r \geq r_a, \\
  u_j(r, r + 1, r, \alpha) & \text{if } r_b \leq r < r_a, \\
  u_j(r, r_b, r, \alpha) & \text{if } r_c \leq r < r_b, \\
  u_j(r_c - 1, r, \alpha) & \text{if } 1 \leq r < r_c,
\end{cases} \]

(31)

\( j = 1, 2 \), where \( r_c = \inf \{ r < r_b : \hat{v}_2(r) \leq g_2(r) \} \). When \( N \to \infty \) such that \( \frac{r}{N} \to x \) we have

\[ \hat{v}_j(x) = \lim_{N \to \infty} v_j(r) = \begin{cases} 
  \hat{u}_j(x, x, x, \alpha) & \text{if } x \geq a, \\
  \hat{u}_j(x, x, a, \alpha) & \text{if } b \leq r < a, \\
  \hat{u}_j(x, b, a, \alpha) & \text{if } c \leq r < b, \\
  \hat{u}_j(c, b, a, \alpha) & \text{if } 0 \leq r < c,
\end{cases} \]

where \( \hat{u}_1(x, y, z, \alpha) = z - x \frac{\varepsilon}{y} + \frac{\varepsilon}{y} \hat{w}_1(y, z, \alpha) \) and \( \hat{u}_2(x, y, z, \alpha) = x \ln \frac{y}{x} + \frac{\varepsilon}{y} \hat{w}_2(y, z, \alpha) \). The asymptotic value of the game for this equilibrium point is

\[ (\hat{v}_1, \hat{v}_2) = (e^{-1} + e^{-\frac{\varepsilon}{2} + e^{\frac{1-\alpha}{2}}} (1 - e^{\frac{1-n}{2}}), e^{-\frac{\varepsilon}{2} + e^{\frac{1-\alpha}{2}}}). \]

(32)

**Theorem 10** In the random priority two-person nonzero-sum game of choosing the best applicant, the Nash equilibrium which gives the maximal probability of success for Player 1 is given by (27) for \( \alpha < \alpha_0 \) and by (30) for \( \alpha \geq \alpha_0 \). The Nash equilibrium payoffs are given by (28) and (31), respectively, and in the asymptotic case they are given by (29) and (32), respectively.

Now, we construct the Nash equilibrium with the lowest probability of success for Player 1. The same arguments as above suggest that one can choose \( (0, 1) \) in \( r_a - 1 \). Using backward induction procedure as long as possible we minimize the equilibrium strategy. For \( \alpha \geq 1 - \alpha_0 \)

\[ (p_r^*, q_r^*) = \begin{cases} 
  (1, 1) & \text{if } r \geq r_a, \\
  (0, 1) & \text{if } r_d \leq r < r_a, \\
  (0, 0) & \text{if } 1 \leq r < r_d,
\end{cases} \]

(33)

and the Nash value

\[ v_j^*(r) = \begin{cases} 
  w_j^*(r, r + 1, \alpha) & \text{if } r \geq r_a, \\
  w_j^*(r, r_a, \alpha) & \text{if } r_d \leq r < r_a, \\
  w_j^*(r_a - 1, r_a, \alpha) & \text{if } 1 \leq r < r_d,
\end{cases} \]

(34)

where \( w_j^*(r, s, \alpha) = w_2(r, s, 1 - \alpha) \), \( w_j^*(r, s, \alpha) = w_1(r, s, 1 - \alpha) \) and \( r_d = \inf \{ r < r_a : \hat{v}_1^*(r) \leq g_1(r) \} \).
For $\alpha < 1 - \alpha_0$ we have
\[
(p^*_r, q^*_r) = \begin{cases} 
(1,1) & \text{if } r \geq r_a, \\
(0,1) & \text{if } r_d \leq r < r_a, \\
(1,0) & \text{if } r_f \leq r < r_d, \\
(0,0) & \text{if } 1 \leq r < r_f,
\end{cases}
\tag{35}
\]
and
\[
v^*_j(r) = \begin{cases} 
u^*_j(r, r + 1, r + 1, \alpha) & \text{if } r \geq r_a, \\
u^*_j(r, r + 1, r_a, \alpha) & \text{if } r_d \leq r < r_a, \\
u^*_j(r, r_d, r_a, \alpha) & \text{if } r_f \leq r < r_d, \\
u^*_j(r_f - 1, r_d, r_a, \alpha) & \text{if } 1 \leq r < r_f,
\end{cases}
\tag{36}
\]
where $u^*_1(r, s, t, \alpha) = u^*_2(r, s, t, 1 - \alpha)$, $u^*_2(r, s, t, \alpha) = u^*_1(r, s, t, 1 - \alpha)$ and $r_f = \inf\{ r < r_d : \tilde{v}_1(r) \leq g_1(r) \}$. When $N \to \infty$ such that $\frac{r}{N} \to x$ we obtain $\frac{r_d}{N} \to d = e^{-\frac{2}{1+\alpha}}$ and $\frac{r_f}{N} \to f = e^{-\frac{2}{1+\alpha}} + e^{\frac{\alpha}{2}}$. The asymptotic value of the game in this equilibrium is
\[
(\tilde{v}^*_1, \tilde{v}^*_2) = \begin{cases} 
(e^{-\frac{\alpha}{2}} + e^\frac{\alpha}{2}, e^{-1} + e^{-\frac{\alpha}{2}} + e^\frac{\alpha}{2} (1 - e^{\frac{\alpha}{2}})) & \text{if } \alpha < 1 - \alpha_0, \\
e^{-\frac{1-\alpha}{2}} e^{-\frac{2}{1+\alpha}}, e^{-\frac{2}{1+\alpha}}) & \text{if } \alpha \geq 1 - \alpha_0.
\end{cases}
\tag{37}
\]

**Theorem 11** In the random priority two-person nonzero-sum game of choosing the best applicant the Nash equilibrium which gives the lowest probability of success for Player 1 is given by (33) for $\alpha \geq 1 - \alpha_0$ and by (35) for $\alpha < 1 - \alpha_0$. The Nash equilibrium payoffs are given by (34) and (36), respectively, and in the limiting case, they are given by (37).

**Remark 1** These solutions do not exhaust all Nash points in considered game. The other pure Nash equilibria can be obtained, roughly speaking, by more often "switches" between $(1, 0)$ and $(0, 1)$ strategy (when both strategies are the Nash equilibria in bimatrix game (26)). This idea is used in Remark 3 to construct Nash equilibria, for $\alpha \in [1 - \alpha_0, \alpha_0]$, with equal Nash values for both players.

**Remark 2** For $\alpha \in (0, 1)$ and for $r$ such that $\tilde{v}_1(r) \leq f(r) < \tilde{c}(r)$ and $\tilde{v}_2(r) \leq f(r) < \tilde{c}(r)$ we can use the randomized Nash equilibrium (see Moulin (1986) for details)
\[
(p^*_r, q^*_r) = \left( \frac{\tilde{v}_2(r) - f(r)}{\tilde{v}_2(r) - f(r) + (1 - \alpha)(f(r) - \tilde{c}(r))}, \frac{\tilde{v}_1(r) - f(r)}{\tilde{v}_1(r) - f(r) + \alpha(f(r) - \tilde{c}(r))} \right).
\]
Table 1. Decision Points for Selected $\alpha$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$d$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>.3677</td>
<td>.2908</td>
</tr>
<tr>
<td>.525</td>
<td>.3432</td>
<td>.2936</td>
</tr>
<tr>
<td>.520</td>
<td>.3335</td>
<td>.2966</td>
</tr>
<tr>
<td>.515</td>
<td>.3264</td>
<td>.2998</td>
</tr>
<tr>
<td>.510</td>
<td>.3205</td>
<td>.3032</td>
</tr>
<tr>
<td>.505</td>
<td>.3154</td>
<td>.3069</td>
</tr>
<tr>
<td>.500</td>
<td>.3109</td>
<td>.3109</td>
</tr>
</tbody>
</table>

Remark 3 Let $\alpha \in (1 - \alpha_0, \alpha_0)$. We have at least two Nash equilibria with the same Nash values for both players equal $\exp(-3 - \alpha_0/2)$ (in the limiting case). The first pair of strategies is (30) and the second pair is (35) with $c = f \approx .2908$ and $b$, $d$, chosen in an appropriate way. The values of $b$ and $c$ one can obtain as solution of the system of equation $\hat{u}_1(c, b, a, \alpha) = \hat{u}_2(c, b, a, \alpha) = \exp(-(3 - \alpha_0)/2)$. Similarly, $d$ and $f$ is solution of the system of equation $\hat{u}_1^*(f, d, a, \alpha) = \hat{u}_2^*(f, d, a, \alpha) = \exp(-(3 - \alpha_0)/2)$. The values of $b$ and $d$ for selected $\alpha$ are given in Table 1.

References


[28] R. L. Tweedie. Criteria for rates of convergence of Markov chains, with an application to queueing and storage theory; In *Papers in*


7. Nonzero-sum stochastic games


7. Nonzero-sum stochastic games


Institute of Mathematics
Technical University of Wrocław
Wybrzeże Wyspińskiego 27,
PL-50-370 Wrocław, Poland