CONTRIBUTIONS TO GAME THEORY AND MANAGEMENT

Volume III

The Third International Conference
Game Theory and Management
June 24-26, 2009, St. Petersburg, Russia

Collected papers
Edited by Leon A. Petrosjan and Nikolay A. Zenkevich

Graduate School of Management
St. Petersburg University
St. Petersburg
2010
Comparison among Some Optimal Policies in Rank-Based Selection Problems

Krzysztof Szajowski

1 Institute of Mathematics and Computer Science, Wroclaw University of Technology, Wybrzeże Wyspiańskiego 27, PL-50-370 Wroclaw, Poland
E-mail: Krzysztof.Szajowski@pwr.wroc.pl
WWW home page: http://neyman.im.pwr.wroc.pl/~szajow

2 Institute of Mathematics, Polish Academy of Science, ul. Śniadeckich 8, PL-00-956 Warszawa, Poland

Abstract The paper deals with the sequential selection problem of the best object. Interviewers observe applicants or items and may decide to stop and hire the current applicant. He has some knowledge about the total number of applicant available. No recall of previously observed candidates is allowed. Knowledge about current applicant is restricted to his relative rank among interviewed so far. The graders have to select, each of them, exactly one item, when it appears, and receives a payoff which is a function of the unobserved realization of random variable assigned to the item or its rank. When there is only one grader the optimal strategy for wide class of payoff functions has a threshold form. It means that in optimal behavior the decision maker should observe the fixed number of items \( k^* \), a learning sample, and to choose the first one after which is better than all those previously observed. The optimality of the strategy is shown by optimal stopping methods for the Markov sequences. The experimental results have shown that the decision makers in problems like choice of partner, the best real investment, try to accept the reasonable option earlier than the optimal strategy of mathematical models suggest. The main aim of the research is to investigate the assumptions of the mathematical model to show their influence on the optimal threshold.

Keywords: rank-based selection, mathematical models of choice, secretary problem, optimal stopping

1. Introduction

In the applied mathematics or the operations research is known the secretary problem (the beauty contest problem, the dowry problem or the marriage problem). The classical secretary problem CSP in its simplest form can be formulated following (Ferguson, 1989):

(i) There is only one secretarial position available.
(ii) The number of applicants, \( N \), is known in advance.
(iii) The applicants are interviewed sequentially in a random order.
(iv) All the applicants can be ranked from the best to the worst without any ties.

Further, the decision to accept or to reject an applicant must be based solely on the relative ranks of the interviewed applicants.

* Dedicated to memory of my friend and colleague Professor Minoru Sakaguchi
(v) An applicant once rejected cannot be recalled later. The employer is satisfied with nothing but the very best. The solicited applicant has no option to reject the offer.

(vi) The payoff is 1 if the best of the \( N \) applicants is chosen and 0 otherwise.

This model can be used as a mathematical description of choice in many decisions in everyday life, such as buying a car, hiring an employee, or finding an apartment (Corbin, 1980). It has appeared as an exercise or funny question in Martin Gardner column of Scientific American (Gardner, 1960). The part of research on the problem has been devoted to modified version of the problem where some important assumption of the model has been changed to fit it to the real life context. There are analysis of decision maker’s aims, decision in a competitive case, an uncertain employment, an unknown or random number of candidates. It was shown that the optimal strategy in many of these problems has very simple threshold form. The items are observed and rejected up to some moments \( k^* \) (thresholds) after which it is optimal to accept the first candidate with reasonable relative rank which depends on the payoff function. This strategy is rather intuitive. When the candidates run low we admit acceptance the lowest rank of chosen item. If the aim is to choose the second best item then the form of the optimal strategy is not so intuitively obvious [see (Szajowski, 1982), (Rose, 1982), (Móri, 1988)].

There are also experimental research with subjects confronted with the classical secretary problem [see (Seale and Rappaport, 1997) and Seale and Rappaport, 2000)]. The optimal strategy of the grader in the classical secretary problem is to pass \( k^*_N - 1 \) applicants, where \( k^*_N \equiv \lceil Ne^{-1} \rceil \) and stop at the first \( j \geq k^*_N \) which is better that those seen so far. If none exists nothing is chosen. The experimental study by (Seale and Rappaport, 1997) of this problem shows that subjects under study have tendency to terminate their search earlier than in the optimal strategy. These effects are not unexpected if we take into account that it is difficult to check the analysis scheme of decision makers. When the above specified assumption are not taken into account by interviewer then the optimal strategy will change. Some small departures will not vary the form of the optimal strategy but the effect will be seen in threshold value. The violation of assumption will be subject of further investigation.

2. Mathematical formulation of the model

Let us assume that the grader observes a sequence of up to \( N \) applicants whose values are i.i.d. random variables \( \{X_1, X_2, \ldots, X_N\} \) with uniform distribution on \( E = [0, 1] \). The values of the applicants are not observed. Let us define

\[
R_k = \#\{1 \leq i \leq k : X_i \leq X_k\}.
\]

The random variable \( R_k \) is called relative rank of \( k \)-th candidate with respect of items investigated to the moment \( k \). The grader can see the relative ranks instead of the true values. All random variables are defined on a fixed probability space \((\Omega, \mathcal{F}, \mathbb{P})\). The observations of random variables \( R_k, k = 1, 2, \ldots, N \), generate the sequence of \( \sigma \)-fields \( \mathcal{F}_k = \sigma\{R_1, R_2, \ldots, R_k\}, k \in \mathbb{T} = \{1, 2, \ldots, N\} \). The random variables \( R_k \) are independent and \( \mathbb{P}\{R_k = i\} = \frac{1}{k} \).
Denote by \( \mathfrak{M}^N \) the set of all Markov moments \( \tau \) with respect to \( \sigma \)-fields \( \{F_k\}_{k=1}^N \).

Let \( q : T \times \mathcal{B} \times \mathbb{D} \to \mathbb{R}^+ \) be the gain function. Define

\[
v_N = \sup_{\tau \in \mathfrak{M}^N} \mathbb{E}q(\tau, R_\tau, X_\tau).
\]  

(1)

We are looking for \( \tau^* \in \mathfrak{M}^N \) such that \( \mathbb{E}q(\tau^*, R_{\tau^*}, X_{\tau^*}) = v_N \).

**Remark 2.1** In CSP there are no assumption that the labels \( \{X_i\}_{i=1}^N \) come from known distribution. It is important that they are different. The function \( q() \) depends on relative ranks only. The optimal stopping time \( \tau^* = \inf\{k_{N} < n \leq N : R_n = 1\} \).

It is shown that \( \lim_{N \to \infty} k_{N} = \exp(-1) \approx 0.367879441 \)

Since \( \{q(n, R_n, X_n)\}_{n=1}^N \) is not adapted to the filtration \( \{\mathcal{F}_n\}_{n=1}^N \), the gain function can be substituted by the conditional expectation of the sequence with respect to the filtration given. By property of the conditional expectation we have \( \mathbb{E}q(\tau, R_{\tau}, X_\tau) = \mathbb{E}\tilde{g}(\tau, R_{\tau}) \), where

\[
\tilde{g}(r, R_r) = \mathbb{E}[q(r, R_r, X_r)|\mathcal{F}_r]
\]

(2)

for \( r = 1, 2, \ldots, N \). On the event \( \{\omega : R_r = s\} \) we have \( \tilde{g}(r, s) = \mathbb{E}[q(r, R_r, X_r)|R_r = s] \).

### 3. A rank-based selection with cardinal payoffs and a cost of choice

(Bearden, 2006) has considered application the best choice problem to the model of choice for the trader who makes her selling decision at each point in time solely on the basis of the rank of the current price with respect to the previous prices, but, ultimately, derive utility from the true value of the selected observation and not from its rank. The assumption (vi) is not fulfilled in this case. He shows that if the true values \( X_j \) are i.i.d. uniformly distributed on \([0, 1]\) then for every \( N \) the optimal strategy is to pass \( c-1 \) applicants, and stop with the first \( j \geq c \) with rank 1. If none exists, stop at time \( N \). The optimal value of \( c \) is either \( \sqrt[2]{N} \) or \( \sqrt[3]{N} \)

This payoff scheme when the i.i.d. \( X_j \)'s come from other than the uniform distribution has been studied by (Samuel-Cahn, 2007). Three different families of distributions, belonging to the three different domains of attraction for the maximum, have been considered and the dependence of the optimal strategy and the optimal expected payoff has been investigated. The different distributions can model various tendency in perception of the searched items.

For the author it is not obvious if the observed behavior of decision makers in experimental research allows to expect that the threshold of the optimal strategy should tend to 0 when number of the applicants tend to infinity. This point of view has been taken into account in search for model with cost of choice [see (Szajowski, 2009)].

The risk is connected with each decision of the grader. The personal feelings of the risk are different. When the decision process is dynamic we can assume that the feeling of risk appears randomly at some moment \( \xi \). Its distribution is a model of concern for correct choice of applicant. It is assumed that \( \xi \) has uniform distribution on \([0, 1, \ldots, N]\).

**Remark 3.1** Let us assume that the cost of choice or the measure of stress related to the decision of acceptance of the applicant is \( c \). It appears when the decision is
after $\xi$ and its measure will be random process $C(t) = c\{\xi \geq t\}$. Based on the observed process of relative ranks and assuming that there are no acceptance before $k$ we have
\[
c(k, t) = \mathbb{E}[C(t) | \mathcal{F}_k] = c \frac{N - t + 1}{N - k + 1}.
\]

The applied model is a consequence of observation that the fear of the wrong decision today is highest than the concern for the consequence of the future decision.

The aim of the grader is to maximize the expected value of applicant chosen and at the same time to minimize the cost of choice.

In this case the function
\[
q(t, R_t, X_t) = \begin{cases} 
(X_t - C(t))I_{\{R_t=1\}}(R_t) & \text{if } t < N, \\
X_N - c & \text{otherwise.}
\end{cases}
\]

We have for $t \geq r$ that $\tilde{g}_c(r, t, R_t) = \mathbb{E}[g_c(t, R_t, X_t) | \mathcal{F}_r] = (\frac{1}{r+1} - c \frac{N - t + 1}{N - r + 1})I_{\{R_t=1\}}(R_t)$. The optimal strategy has the threshold form with $k_N^* \cong \lfloor N \alpha \rfloor$ where $\alpha$ is solution of the equation $\log(x) = (1 + \frac{1}{2}c)(x - 1)$ in $(0, 1)$. For example for $c = 0.1$ the limit $\alpha \cong 0.002516$.

4. Imprecise number of objects vs. uncertain selection

The proposed in the subject of this section violation of assumption (ii) and (v) are logically connected. When the candidates can reject offers the decision maker has random number of available objects. Such modification of CSP was proposed and solved by (Smith, 1975). He admitted that the interviewer can solicit more than ones. Let us assume that the probability of rejection of an offer is $p$. The optimal strategy is threshold one. The grader should interview some number $k^*$ of object which the number depends on the horizon length and the rejection probability $p$. It was shown by (Smith, 1975) that $\lim_{N \to \infty} k_N^* = p^{1-p}$.

Let us assume that grader has no precise knowledge about the number of candidates. One of the natural case is when the number of available candidates is bounded by some number $M$, but the exact value of them $N$ has uniform distribution. The analysis of such modification was made by (Rasmussen and Robbins, 1975) and they showed that the optimal threshold has approximate value $k_M^* \cong \lfloor 2Me^{-2} \rfloor$, i.e. $\lim_{M \to \infty} \frac{k_M^*}{M} \cong 0.270670566$. The uncertainty related to the real number of candidates makes that the decision maker has tendency to accept earlier than in CSP.

Remark 4.1 One can expect that the problem with uncertain selection when $p = \frac{1}{2}$ has as level of difficulty to select the best applicant as the Rasmussen and Robbins problem. However, the chance for success in both problem is the same when the probability of rejection of an offer is $p^* \cong 0.57039$ (i.e. fulfills the equation $p^{1-p} = 2e^{-2}$).

5. Competitive best choice models

When there are $m$ decision makers and each of them has his own stream of $N$ candidates [see (Sonin, 1976)] the definition of optimality is understood as behavior which leads to stability. Let us assume that the each grader observes a sequence of up to $N$ applicants whose values are i.i.d. random variables $\{X_1^i, X_2^i, \ldots, X_N^i\}_{i=1}^m$ with
uniform distribution on $\mathbb{E} = [0,1]$. The values of the applicants are not observed. Let us $R^i_k = \#\{1 \leq j \leq k : X^i_j \leq X_k\}$. The player $i$ observes the random variables $R^i_k, k = 1,2,\ldots,N$. It generates the sequence of $\sigma$-fields $\mathcal{F}^i_k = \sigma\{R^i_1,R^i_2,\ldots,R^i_k\}$, $k \in \mathbb{T} = \{1,2,\ldots,N\}$. The random variables $R^i_k$ are independent and $P\{R^i_k = j\} = \frac{1}{k}$ for $j = 1,2,\ldots,k$.

Denote by $\mathfrak{M}^{i,N}$ the set of all Markov moments $\tau^i$ with respect to $\sigma$-fields $\{\mathcal{F}^i_k\}_{k=1}^N$. Let $q: \mathbb{T} \times \mathbb{R} \times \mathbb{E} \rightarrow \mathbb{R}^+$ be the gain function. Define $f_N(\tau^1,\tau^2,\ldots,\tau^m) = \mathbb{E}q(\tau^1 \wedge \tau^2 \wedge \ldots \wedge \tau^m, R^i_1, X^i_1)$. We are looking for $\tau^i \in \mathfrak{M}^{i,N}$ such that

$$f_N(\tau^1,\tau^2,\ldots,\tau^i,\ldots,\tau^m) \geq f_N(\tau^1,\tau^2,\ldots,\tau^m). \quad (5)$$

In other words the strategy will give the equilibrium. Taking into account this definition of optimality (the Nash equilibrium approach) and assuming that each player has the same aim it is proved that all players have the same threshold strategy defined by $k^*_N \cong [MZ^*_N]$ where $z^*_m$ fulfills the equation: $mz = 1 - (1 + z \log(z))^m$. For two players $z^*_2 \cong 0.295330$.

However, if there is $m$ players and only one stream of candidates in the mathematical model and in fact the real life rules of assignment should be determine. Mainly, when more than one player wants to accept the same candidate the rules of assignment are needed. The reasonable solution is to determine the priority of the players in deterministic or random manner. Let us report the two person game model with random assignment which gives the priority to the first player with probability $p$ [see (Szajowski, 1993)]. The problem is transformed to the Dynkin’s game with players payoff similar to those in the uncertain employment version of the CSP. Both decision makers’ strategies are pairs of stopping times $(\tau_i, \sigma_j)$, $i,j = 1,2,j \neq i$. For $p \in (0,1)$ there is more than one equilibrium. The stopping times $\tau^*_i(p)$ have threshold form at the equilibrium. For $p = 0.5$ there are two equilibria with stopping times $\tau^*_i\left(\frac{1}{2}\right)$ defined by the threshold for the first equilibrium $k_2^i \cong [MX^*_i]$ where $x^*_i$ fulfills the equation $mx = 1 - (1 + z \log(z))^m$. This special case of the game has been posed and solved by (Fushimi, 1981). When the first player has priority the thresholds are $[Mexp(1), Mexp(-1)]$. One of the players should behave more hastily than in the original secretary problem and should start solicitation at [0.2865N] approximately.

Various game models with waives from the presented in this section problems one can find in review by (Sakaguchi, 1995) who investigated the topic widely. In some what similar model has been investigate by (Sakaguchi and Mazalov, 2004) with the aim to choose the one candidate acceptable for all players. There are two players that jointly interview a given number $n$ of applicants. The applicants present in a random order, and each such an order is equally likely. If both players accept a candidate then she is accepted, if both reject, she is rejected. If their choices are different, then an arbitration comes in and forces them to agree with the first player or the second one with probability $p$ and $1-p$, respectively. The aim of the players is to minimize his expected payoff. These model of choice is not analyzed here as well as the voting stopping games [see (Kurano et al., 1980), (Szajowski and Yasuda, 1996), (Mazalov and Banin, 2003), (Ferguson, 2005)] and the games where players have more than one effective chance to choose the item [see e.g. (Szajowski, 2002) ].
The next section is devoted to the games with given priorities. Two cases are taken into account which differs on knowledge collection during the game. The priority and information change rapidly perspectives of the decision makers.

6. Privileges vs. information in two person CSP

In this section we present the solution of two kinds of two person nonzero-sum games, which are based on CSP. The solution of the best choice problem (one player game) is auxiliary in the solution of these games. With sequential observation of the applicants some natural probability space \((\Omega, \mathcal{F}, P)\) is connected. The elementary events are the permutation of all applicants and the probability measure \(P\) is the uniform distribution on \(\Omega\). The observable sequence of relative ranks \(Y_k, k = 1, 2, \ldots, N\) defines the sequence of the \(\sigma\)-fields \(\mathcal{F}_k = \sigma(Y_1, \ldots, Y_k), k = 1, 2, \ldots, N\). The random variables \(Y_k\) are independent and \(P(Y_k = i) = \frac{1}{k}\). Denote \(S^N\) the set of all Markov times \(\tau\) with respect to the \(\sigma\)-fields \(\{\mathcal{F}_k\}_{k=1}^N\) bounded by \(N\). The secretary problem can be formulated as follows: we are searching \(\tau^* \in S^N\) that \(P\{Z_{\tau^*} = 1\} = \sup_{\tau \in S^N} P\{Z_{\tau} = 1\}\). In the two person games each player wants to get the best applicant. There is only one stream of candidates and it is the reason why the assignment system should be defined. It will be formulated as the priority rules [see (Ravindran and Szajowski, 1992), (Szajowski, 1992)].

The permanent priority scheme is assumed, i.e. Player 1 has highest priority and Player 2 has lower priority than Player 1. In the section 6.3. the model of priority properties from (Ramsey and Szajowski, 2001) is used. It means that all players have the common knowledge about the random sequence observed sequentially. The \(n\)th applicant seen has relative rank \(i\), where \((i \in \{1, 2, \ldots, n\})\), if the \(n\)th applicant seen is the \(i\)-th best applicant among seen so far. The object of each player is to obtain the best applicant. Thus, it is assumed that each player obtains a reward of 1 if he obtains the best applicant and 0 in all other cases. In such cases a player should only accept an applicant, who has relative rank 1. These applicants will be referred to as candidates. If more than one player wishes to accept a candidate, then a priority scheme decides which player has priority and hence he employs that candidate, as previously described. If one player obtains a candidate, then he/she stops searching. The other players are informed of this and then they are allowed to continue searching.

The game of the second type assumes the dynamic assignment of knowledge about the random sequence to the players. It means that the items are presented to the players according their priority and they make decision about acceptance or rejection before the item is presented to next player. In case of acceptance of the observation (an item) by player \(i\)-th the player with lower priority has no knowledge about the accepted state. In a consequence the relative ranks observed after the first acceptance are not taken into account the “hidden observation” and a new, based on restricted history, relative rank are observed. It is assumed, as in the first example, that the aim is to choose the best applicant. If the new relative rank is 1, then the it is actually the relatively the first or the second and consequently, there is only 50% chance that the relatively best after the first acceptance is the best when we compare with the common knowledge case. In each case it is assumed that the number of objects (hereon, called applicants) presented to the players is very large. It means that the asymptotics of the Nash values and the equilibrium strategy when \(N \to \infty\) are obtained. This game is the subject of consideration in the section 6.4.
Comparison among Some Optimal Policies in Rank-Based Selection Problems

6.1. The Markov chain related to the best choice problem

The mathematical model of the secretary problem which is useful for these games is a some version of the Markov chain with the two dimensional state space [see (Dynkin and Yushkevich, 1969), (Presman and Sonin, 1975), (Szajowski, 1982)].

Let us define \( W_0 = (1, Y_1) = (1, 1) \) and put \( a = \max(A) \), where \( A \subset \{1, 2, \ldots, N\} \). \( \gamma_t = \inf\{r > \gamma_{t-1} : Y_r \leq \min(a, r)\} \) (\( \inf \emptyset = \infty \)) and \( W_t = (\gamma_t, Y_{\gamma_t}) \). If \( \gamma_t = \infty \) then define \( W_t = (\infty, \infty) \). \( W_t \) is the Markov chain with following one step transition probabilities [see (Szajowski, 1982)]

\[
p(r, s) = \mathbb{P}\{W_{t+1} = (s, l_s)|W_t = (r, l_r)\} = \frac{1}{\pi}, \quad \text{if } r < a, s = r + 1,
\]

\[
\frac{(r)_{a-1}}{(r)_{a-1}}, \quad \text{if } a \leq r < s,
\]

\[
0, \quad \text{if } r \geq s \text{ or } r < a, s \neq r + 1,
\]

with \( p(\infty, \infty) = 1 \), \( p(r, \infty) = 1 - a \sum_{s=r+1}^{N} p(r, s) \), where \((s)_a = s(s - 1)(s - 2) \ldots (s - a + 1), (s)_0 = 1\). Let \( \mathcal{G}_t = \sigma(W_1, W_2, \ldots, W_t) \) and \( \mathfrak{M}_N \) be the set of stopping times with respect to \( \{\mathcal{G}_t\}_{t=1}^{N} \). Since \( \gamma_t \) is increasing, then we can define \( \mathfrak{M}_N^{\infty} = \{\sigma \in \mathfrak{M}_N : \gamma_{\sigma} > r\} \).

Let \( \mathbb{P}_{(r,l)}(\cdot) \) be probability measure related to the Markov chain \( W_t \), with trajectory starting in state \((r, l)\) and \( \mathbb{E}_{(r,l)}(\cdot) \) the expected value with respect to \( \mathbb{P}_{(r,l)}(\cdot) \). From (6) we can see that the transition probabilities do not depend on relative ranks, but only on moments \( r \) where items with relative rank \( l \leq \min(a, r) \) appear. Based on the following lemma we can solve the problem (1) with gain function \( (4) \) using the embedded Markov chain \( W_t \) and the gain function given by (2).

Lemma 1 (see (Szajowski, 1982)).

\[
\mathbb{E}_{W_N}(s + 1, Y_{s+1}) = \mathbb{E}_{(s,l)} w_N(W_t) \text{ for every } l \leq \min(a, r).
\]

Let us define \( W_1 = 1 \) and . Define \( W_t = \inf\{r > W_{t-1} : Y_r = 1\}, \; t > 1 \), \( \inf \emptyset = \infty \). \( (W_t, \mathcal{F}_t, \mathbb{P}^{(1,1)})_{t=1}^{N} \) is the homogeneous Markov chain with the state space \( \mathcal{E} = \{1, 2, ..., N\} \cup \{\infty\} \), \( \mathcal{F}_t = \sigma(W_1, W_2, ..., W_t) \) and the following one-step transition probabilities: \( p(r, s) = \mathbb{P}\{W_{t+1} = s\mid W_t = r\} = \frac{r}{(s-1)} \) if \( 1 \leq r < s \leq N \), \( p(r, \infty) = 1 - \sum_{s=r+1}^{N} p(r, s) \), \( p(\infty, \infty) = 1 \) and 0 otherwise. The payoff function for the problem defined on \( \mathcal{E} \) has a form \( f(r) = \frac{r}{N} \).

In the games considered in the section 6.3. and the section 6.4. both players stop at most two times. It will be interesting to compare the values of the games and the strategies with the solution of the best choice problem with two stops.

6.2. Choosing the best with two stops

The solution of the best choice problems (one player game) with one and two stops are auxiliary in the solution of the two person game with fixed priority of the player.

Classical best choice problem with one stop Let \( W_1 = 1 \). Define \( W_t = \inf\{r > W_{t-1} : Y_r = 1\}, \; t > 1 \), \( \inf \emptyset = \infty \). \( (W_t, \mathcal{F}_t, \mathbb{P}^{(1,1)})_{t=1}^{N} \) is the homogeneous Markov chain with the state space \( \mathcal{E} = \{1, 2, ..., N\} \cup \{\infty\} \), \( \mathcal{F}_t = \sigma(W_1, W_2, ..., W_t) \) and the following one-step transition probabilities: \( p(r, s) = \mathbb{P}\{W_{t+1} = s\mid W_t = r\} = \frac{r}{(s-1)} \) if \( 1 \leq r < s \leq N \), \( p(r, \infty) = 1 - \sum_{s=r+1}^{N} p(r, s) \), \( p(\infty, \infty) = 1 \) and 0 otherwise. The payoff function for the problem defined on \( \mathcal{E} \) has a form \( f(r) = \frac{r}{N} \).
6.3. The fixed scheme of priorities and common knowledge

Let us denote by $\tau_r$, the strategy of the first player at equilibrium is $a$ such that $a = \lim_{N \to \infty} \frac{r}{N} = e^{-1} \approx 0.3679$.

Two stopping

Let us solve the best choice problem with double stopping. Denote $\tilde{c}^{(2)}(r) = \sup_{\tau_1, \tau_2 \in [0,1]} \mathbb{P}_r(\min\{X_{\tau_1}, X_{\tau_2}\} = 1)$.

The general method of solving the double stopping problem leads to conclusion that $\tilde{c}^{(2)}(r) = E_r c^{(2)}(W_1)$, where $c^{(2)}(r) = \max\{f(r) + \alpha(r), E_r c^{(2)}(W_1)\}$. Solving the equation recursively we get $c^{(2)}(r)$. Define $r_b = \inf\{1 \leq r \leq N : c^{(2)}(r) = f(r) + \alpha(r)\}$. We have $\tau_1^* = \inf\{1 \leq r \leq N : r \geq r_b, Y_r = 1\}$ and $\tau_2^* = \inf\{r \geq r_a : r > \tau_1^*, Y_r = 1\}$. When $N \to \infty$ such that $\frac{r}{N} \to x$ we obtain

$$\lim_{N \to \infty} c^{(2)}_N(x) = \begin{cases} \frac{x - x \log(x)}{e^{-1} - x} & \text{if } e^{-1} \leq x \leq 1, \\ x + e^{-1} & \text{if } e^{-1} - \frac{1}{2} \leq x < e^{-1}, \\ e^{-1} + e^{-\frac{1}{2}} & \text{if } 0 < x \leq e^{-\frac{1}{2}} \end{cases}$$

and

$$\lim_{N \to \infty} \tilde{c}^{(2)}_N(x) = \begin{cases} -x \log(x) + \frac{1}{2} \log^2(x) & \text{if } e^{-1} \leq x \leq 1, \\ e^{-1} - \frac{1}{2} - x \log(x) & \text{if } e^{-1} - \frac{1}{2} \leq x < e^{-1}, \\ e^{-1} + e^{-\frac{1}{2}} & \text{if } 0 < x \leq e^{-\frac{1}{2}}. \end{cases}$$

The asymptotic behavior of thresholds are following $b = \lim_{N \to \infty} \frac{r_b}{N} = e^{-\frac{1}{2}}$ and $a = \exp(-1)$.

6.3. The fixed scheme of priorities and common knowledge

Let us denote by $v_i(t)$, $i = 1, 2$ and $t \in (0,1]$, the $i$-th Player value of the game which start at moment $t$. The player 1 has the highest priority. It means, that he will use the strategy for the optimal stopping problem for Markov chain presented in Section 6.1. We have $v_1(r) = \tilde{c}(r)$ given by (8) and when $N \to \infty$ such that $\frac{r}{N} \to x$ we obtain $v_1(t) = \tilde{c}_1(t) = -t \log t$, $a = \lim_{N \to \infty} \frac{r}{N} = e^{-1} \approx 0.3679$. The strategy of the first player at equilibrium is $\tau_1^* = \inf\{r_a \leq r \leq N : Y_r = 1\}$.

If at moment $r \geq r_a$ the relatively best appears and no player has accepted the candidate than, according to the strategy of the first player, he accepts it and the second player will take the next relatively best. At the left hand side neighborhood of $r_a$, taking into account the highest priority of the second vs third player and the highest immediate payoff than the expected value of future decision, $v_2(r)$ can be get by determining the equilibrium in suitable two person matrix game. Since $v_1(r_a - 1) > f(r_a - 1)$ and $v_2(r_a - 1) \equiv \frac{1}{(r_a - 1)^2} < f(r_a - 1)$ the strategy $(f, s)$ is in equilibrium. For close to $r_a$ moment we have

$$v_2(r) = \sum_{i=r+1}^{r_a-1} \frac{r}{i(i-1)} \frac{r}{N} + \sum_{i=r_a}^{N} \frac{r}{i(i-1)} \tilde{c}(i) = \sum_{i=r+1}^{r_a-1} \frac{r}{i(i-1)} \frac{r}{N} + \frac{r}{r_a - 1} v_2(r_a - 1)$$
Define \( r_\text{b} = \inf \{ 1 \leq r \leq N : v_2(r) \leq f(r) \} \). We have \( \tau_2^* = \inf \{ 1 \leq r \leq N : r \geq r_\text{b}, Y_r = 1, \tau_1^* \neq \emptyset \} \). When \( N \to \infty \) such that \( \frac{N}{\tau_2^*} \to x \) we obtain

\[
\lim_{N \to \infty} v_2(r) = \begin{cases} \frac{2}{x} \log^2(x) & \text{if } e^{-1} \leq x \leq 1, \\ -x \log(x) - \frac{2}{x} & e^{-\frac{2}{x}} \leq x < e^{-1}, \\ e^{-\frac{2}{x}} & 0 < x \leq e^{-\frac{2}{x}}. \end{cases}
\]

The asymptotic behavior of thresholds are following \( b = \lim_{N \to \infty} \frac{r_\text{b}}{N} = e^{-\frac{2}{x}} \cong 0.22313 \).

### 6.4. Two person best choice game with hidden state

In the second type game introduced at the beginning of the section 6. the items are presented to the players according their priority and they make decision about acceptance or rejection before the item is presented to next player. In case of acceptance of the observation (an item) by player \( i \)-th the player with lower priority has no knowledge about the accepted state. As a consequence the relative ranks observed after the first acceptance do not take into account the “hidden observation” and new relative rank is observed. It is assumed, as in the first example, that the aim is to choose the best applicant. If the new relative rank is 1, then it is the relatively the first or the second if the rank is based on all observation and consequently, there is only 50\% chance that the relatively best after the first acceptance is the best when we compare with the common knowledge case. Let us assume that the number of objects (hereon, called applicants) presented to the players is very large and the asymptotic of the Nash values and the equilibrium strategy when \( N \to \infty \) are obtained.

The construction of the Nash value and the equilibrium is made by backward induction. When game starts at moment \( r \) (i.e. no player has chosen the item), observation \( Y_r = 1 \) and the first player has highest priority then the observation \( r \) is accepted by the first player iff \( r \geq r_\text{a} \). The value for the first player \( v_1(r) \) is given by the same formula as the value for the first player in the game with common knowledge described in Section 6.3.

If at moment \( r \geq r_\text{a} \) the relatively best appears and no player has accepted the candidate than, according to the strategy of the first player, he accepts it and the second player will take the next relatively best. However, the observation chosen by the first player is hidden and unknown for the second player. Player 2 can assume that the Player 1 is rational and he is not accepting observation which is not potentially the best. Based on observation \( Y_1, Y_2, \ldots, Y_{r-1}, Y_r, \ldots, Y_s \) and taking into account the rationality of Player 1 the relative ranks \( S_{r+1}, S_{r+2}, \ldots, S_s \) means that the unobserved true relative ranks \( Y_{r+1}, \ldots, Y_s \) take wider set of values e.g. \( S_{r+1} = 1 \) implies that \( Y_{r+1} \in \{1, 2\} \). If the aim is to find the best item the expected payoff by choosing \( S_{r+1} = 1 \) is equal \( g(r) = \frac{1}{2} \).

Let us define \( \mathcal{M}_r = (1, Y_1) = (r, 1) \). Let us define \( \delta_t = \inf \{ r > \delta_{t-1} : S_r = 1 \} \) (\( \inf \emptyset = \infty \)) and \( \mathcal{M}_t = (\delta_t, Y_{\delta_t}) \). If \( \delta_t = \infty \) then define \( \mathcal{M}_t = (\infty, \infty) \). \( \mathcal{M}_t \) is the Markov chain with following one step transition probabilities

\[
p(i, j) = P(\delta_{t+1} = j | \delta_t = i) = \frac{i(i-1)}{j(j-1)(j-2)}
\]

[see (Szajowski, 1982)]. For \( r \geq r_\text{a} \) the value function \( w_2(r) = \sum_{j=r_\text{a}+1}^N \frac{r(r-1)}{j(j-1)(j-2)} \frac{1}{2} \). The equilibrium strategy of Player 2 suggests to
choose the item at moment \( r > r_o \) with \( S_r = 1 \) iff \( w_2(r) \leq \frac{1}{2} \frac{\alpha}{N} \). Let \( r_o = \inf\{ r \leq N : w_2(r) \leq \frac{1}{2} \frac{\alpha}{N} \} \).

Let us consider the asymptotic case. When \( N \to \infty \) such that \( \frac{\alpha}{N} \to x \) we obtain

\[
 w(t) = \lim_{\frac{\alpha}{N} \to t} w(r) = \int_{t}^{1} \frac{2 r^2}{s^3} ds + t(1 - t) \quad \text{and} \quad \alpha = \lim_{N \to \infty} \frac{r_o}{N} = \frac{1}{2}.
\]

At the left hand side neighborhood of \( r_o \), taking into account the low priority of the second player, the expected value of future decision, \( w_2(t) \) can be get by determining the equilibrium in suitable two person matrix game. The careful analysis shows that the equilibrium value of the second player for \( a < t < \alpha \) has to fulfill the integral equation

\[
 w(t) = \int_{t}^{\alpha} \frac{2 t^2}{s^3} w_2(s) ds + \frac{t^2}{\alpha^2} w_2(\alpha).
\]

The solution of this equation in \( (a, \alpha] \), taking into account the continuity property at \( \alpha \) gives \( w_2(t) = \frac{1}{t} \). For \( t < a \) we have any restriction for the decision and knowledge for the second player. If at moment \( r < r_o \) we have \( Y_r = 1 \) and \( r \) is close to \( r_o \) he can get \( \frac{\alpha}{N} \cong e^1 \) by accepting the item which appears at \( r \). If he chooses option to reject this item his expected gain is approximately \( w_2(\alpha) \cong \frac{1}{t} < e^{-1} \). Then, he will accept the relatively best in some interval \( (\beta, a] \) and he will play with the equilibrium strategy after \( a \). His behavior after \( r_o \) is following: he is waiting for the choice of item by the first player and after the moment of acceptance the item by Player 1, he will accept the first item with \( S_r = 1 \) iff \( r \geq r_o \). Such behavior leads to gain

\[
 w_2(t) = \int_{t}^{\alpha} \frac{t}{s^3} ds + \int_{\alpha}^{1} \frac{t}{s^3} w_2(s) ds
 = t(\frac{e}{4} - 2 + \log(2) - \log(t)).
\]

If we compare the expected value of optimal behavior after the rejection of the relatively best at \( r \) such that \( \frac{\alpha}{N} \cong t \) with the immediate payoff by accepting the item we get that \( w_2(t) \leq t \) for \( t < \beta \) where \( \beta = e^{\frac{1}{2} - 3 + \log(2)} = 2 e^{\frac{1}{2} - 3} \cong 0.196463 \).

The asymptotic value of the game, when \( N \to \infty \) such that \( \frac{\alpha}{N} \to t \) has the form

\[
 w_2(x) = \lim_{\frac{\alpha}{N} \to t} w_2(r) = \begin{cases} 
 t^2 - t \log(t) & \text{if } \alpha \leq t \leq 1, \\
 \frac{1}{4} + (\log(2) - 1)t & \text{if } \alpha \leq t < \alpha, \\
 (\frac{1}{4} - 2 + \log(2))t - t \log(t) & \text{if } \beta \leq t < \alpha, \\
 \beta t & \text{if } 0 \leq t < \beta,
\end{cases}
\]

where \( \alpha = 0.5 \), \( a = e^{-1} \) and \( \beta = 2 e^{\frac{1}{2} - 3} \cong 0.196463 \).

**Remark 6.1** It is worth to emphasize the form of the strategy of the second player. If he reaches the interval \( [a, \alpha] \) without acceptance of any candidate he has to be silent in this interval and do not accept the candidates if the first player has done it before. His activity comes back when the game reaches the moment asymptotically later than \( a \).

7. Conclusion

The conclusion from the presented cases is that in the simple best choice problems the decision maker can imagine various threats. In many real problems one can
observe that the decision maker hesitates to long and postpones the final decision [see e.g. (Bearden, 2006), (Seale and Rappaport, 1997)]. He rejects relatively best option too long. It looks that he fears to loss the potential options. The level of fear can be dependent on the value of the item or independent. The mathematical models are able to explain the behavior of the decision maker by specifying more adequate assumptions. The assumption, which makes the model more adequate, are the consequences of the hidden supposition of the decision maker or his creation of circumstances. The presented examples do not exhaust all such deviation of the problem known in literature.

References


