

A TIME SEQUENTIAL GAME RELATED TO THE SEQUENTIAL STOCHASTIC ASSIGNMENT PROBLEM

Tōru Nakai
University of Osaka Prefecture

(Received September 26, 1980; Final October 14, 1981)

Abstract A two-person zero-sum time sequential game related to the sequential stochastic assignment problem is considered. The i.i.d. random variables appear one by one at a time. At each stage, after observing the value x of the random variable, both players must decide whether to accept or reject, and if player I who has n men to assign decides to accept, then he selects one of the n men in order to assign. When both players accept, this random variable is selected and assigned to the selected man with the ability p . The immediate payoff to player I is px . After the man is assigned, he is unavailable for the future stages. If otherwise, they get a new observation and the selected man is assigned in the future stages. The objective of player I (resp. player II) is to maximize (resp. minimize) the total expected reward. In this paper we obtain the recursive equation from which we get the optimal strategies of both players and the value of the game. The optimal strategies of the players are similar to that of the stochastic assignment problem. We observe the property of the special case of this game in the asymptotic case.

1. Introduction

In their paper [4], Derman, Lieberman and Ross consider the following sequential stochastic assignment problem: There are n men available to perform n jobs. The n jobs arrive in sequential order, i.e., job 1 appears first, followed by job 2, etc. Let $X_j, j=1, \dots, n$, be independently and identically distributed random variables, each assuming a real value, interpreted as the value of the j -th job. Upon arrival of the job, the decision maker has to select one of the n men to assign. If the random variable X_j takes on the value x_j and the i -th man p_i is assigned, the expected reward is given by $p_i x_j$ where $0 \leq p_i \leq 1, i=1, \dots, n$, are the known constants. After a man is assigned to a job, he is unavailable for the future stages. The problem is to assign n men to the n jobs so as to maximize the total expected reward. Here we consider that the p_i is an ability of the i -th man, i.e., if $p_i = 1$, this man is "perfect" and obtains the reward x_j for the realized

value x , and if $p_i < 1$, he has a less ability than the "perfect" man (for example, he takes more time or makes more mistakes than the "perfect" man).

As for the above problem the decision maker can attack all appeared n jobs and get their values. In this paper we consider the above sequential assignment problem in the competitive situation and we formulate as the two-person zero-sum game with the rejections as follows. Here we assume that player I desires to maximize the total expected reward and on the other hand player II desires to minimize the total expected cost, where we assume that the cost is in proportion to the value px of the result. Hence player II desires to minimize the total expected reward.

Prosecution and defence (player I and player II, respectively) must choose n jobs and assign the n men of player I to the jobs. The value of the arriving jobs are assumed to be chosen randomly and come up one at a time for decisions. Let $X_j, j=1, \dots, n$, be the independently and identically distributed random variables, each assuming a real value, interpreted as the value of the j -th arriving job. Upon arrival of the job, player I and II must decide, simultaneously and independently of the other player's choice, whether to accept or to reject it according to the observed value. At the same time, if player I decides to accept, he must select one of the remaining n p_i 's in order to assign to this job, where $p_i, i=1, \dots, n$, are known constants. If both players accept and player I selects the i -th man p_i , this i -th man is assigned to this job and the immediate payoff to player I is $p_i x$ (x is the realized value of the random variable). After a man is assigned to a job, he is unavailable for the future stages. If either player rejects, a new job comes up for the next stage. In this case the selected p_i is not assigned and is used for the future stages. The number of the preemptory rejections for each side is restricted, and player I and II can reject r and s times, respectively. The process continues until n men are assigned to the n selected jobs. The object of this game is to find the strategy for player I and II which maximize for player I, and minimize for player II, the total expected payoff of the game $E \sum_{j=1}^n p_{\sigma(j)} X(j)$, where $\sigma \in \mathfrak{S}_n$ and j 's represent the sequence of the jobs selected (\mathfrak{S}_n is the n -dimensional symmetric group). This game is the game theoretic extension of the above sequential stochastic assignment problem. We assume that the above game has perfect recall. On the other hand if we consider the special case where $p_i = 1 (i=1, \dots, n)$, this game is the one considered in Sakaguchi [8].

There is another interpretation of the above problem as follows. We would like to employ the private teachers (or tutors) for the n students. The private teachers arrive one by one in sequential order and his capacity

" x " as a teacher is obtained as the realized value of the i.i.d. random variable X . If the student with the ability p studies under a private teacher with the capacity x , he makes good grades in proportion to the value px . Here we consider that the tuition fees of the teachers are in proportion to the results. Player I desires that the students get good grades in their studies, and player II wants to minimize the total of the tuition fees.

The purpose of this paper is to present the solution of this game and get the optimal strategy of the players. Concerning the Lemma 3 of Section 3, we obtain the result similar to the fact mentioned in Sakaguchi [8], where we get the result in the simpler manner. Finally we consider the asymptotic property of the game for the special case where $p_i = 1$ ($i=1, \dots, n$). At this time we get the solution for the open problem discussed in the Remark 3 of the paper by Sakaguchi [8]. This case is considered as a generalization of the work by Gilbert and Mosteller [5], which is a classical stopping problem, and a game theoretic extension of the study of Albright and Derman [2].

2. Formulation of the time sequential game

Let $X_i, i=1, 2, \dots$, be the independently and identically distributed random variables that can be observed one by one sequentially. The common cdf (cumulative distribution function) $F(x)$ of each X_i is assumed to be known by both players. Now we assume that $\mu \equiv E(X) < \infty$. Player I and player II are asked to select n observations sequentially and further, player I is asked to assign one of his p 's for these n selected observations (where the p 's are known constants). Let $X_{(j)}, j=1, \dots, n$, be the sequence of selected observations, and we suppose that the object of player I (resp. player II)'s sequential decisions is to maximize (resp. minimize) the total expected payoff $E(\sum_{j=1}^n p_{\sigma(j)} X_{(j)})$, where $\sigma \in \mathfrak{S}_n$ and $\sigma(j)$ is defined to be the man (identified by numbers) assigned to the j -th observations selected by these two players.

Let $G_n(p_1, \dots, p_n, r, s)$ ($p_1 \leq \dots \leq p_n$) denote the game in which n random variables remain to be selected and assigned, and the player I and II can reject another r and s times, respectively. This game is played as follows. Step 1. (chance move) Both players observe the realized value x of the random variable X .

Step 2. After observing the realized value x , player I decides whether to accept or to reject, and if he decides to accept, then he also decides which p he selects to assign to this job. Simultaneously player II decides whether to accept or to reject, independently of the player I's choice.

Step 3. If either player rejects, the immediate payoff to player I is 0 and the players observe the new realized value of X where n random variables remain to be selected and the number of rejections of the player who rejected decreases by 1. If both players accept and player I selects p_i , the immediate payoff to player I is $p_i x$ and both players face the game

$$G_{n-1}(p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_n, r, s).$$

Now these three steps are repeated again and again. After n men (or p 's) are assigned the game stops.

The following notations will now be introduced: Let

$$v_n(p_1, \dots, p_n, r, s) = \text{the value of the game } G_n(p_1, \dots, p_n, r, s),$$

$$v_n(p_1, \dots, p_n, r, s | x) = \text{the conditional value of the game } G_n(p_1, \dots, p_n, r, s) \text{ given that the observation } x \text{ is currently observed.}$$

From the dynamic programming formulation of this game and the fact that this game has perfect recall, we have the following recursive equation,

$$(1) \quad v_n(p_1, \dots, p_n, r, s) = E v_n(p_1, \dots, p_n, r, s | X),$$

$$(2) \quad v_n(p_1, \dots, p_n, r, s | x)$$

$$= \text{val} \left[\begin{array}{c} \begin{array}{cc} \text{rej.} & \text{acc.} \\ \text{assign } p_1 & p_1 x + v_{n-1}(p_2, \dots, p_n, r, s) \\ \cdot & \cdot \\ \cdot & \cdot \\ \text{assign } p_n & p_n x + v_{n-1}(p_1, \dots, p_{n-1}, r, s) \end{array} \\ \begin{array}{cc} v_n(p_1, \dots, p_n, r-1, s-1) & v_n(p_1, \dots, p_n, r-1, s) \\ v_n(p_1, \dots, p_n, r, s-1) & \\ \dots & \dots \\ \dots & \dots \\ v_n(p_1, \dots, p_n, r, s-1) & p_n x + v_{n-1}(p_1, \dots, p_{n-1}, r, s) \end{array} \end{array} \right]$$

where $\text{val } A$ denote the value of the two-person zero-sum matrix game A . The interpretation of $p_1 x + v_{n-1}(p_2, \dots, p_n, r, s)$ is that, in this stage, the realized value of the observation is x , both players accept and player I selects the first man p_1 for this observation x , and player I gets the immediate payoff $p_1 x$ and then faces the game $G_{n-1}(p_2, \dots, p_n, r, s)$. The interpretation of the others are considered similarly. The boundary condition of the above recursive equation is $v_0(r, s) = 0$.

In proving the result of this paper, a well-known theorem due to Hardy [6] will be used (as in [4]).

Lemma 1. If $x_1 \leq \dots \leq x_n$ and $y_1 \leq \dots \leq y_n$ are the sequences of numbers, then

$$(3) \quad \max_{\sigma \in S_n} \sum_{j=1}^n x_{\sigma(j)} y_j = \sum_{j=1}^n x_j y_j .$$

3. Optimal strategy and the value of the game

In this section we give the main theorem which embodies the optimal strategies of both players and the value of the game, and some related lemmas.

First of all we define the sequence $\{a_{i,N}\}_{0 \leq i \leq N}$ by the following recursive formula.

$$a_{0,0} = \infty ,$$

$$a_{0,N} = -\infty , a_{N,N} = \infty \text{ for } N \geq 1 \text{ and}$$

$$(5) \quad a_{i,N} = \int_{a_{i-1,N-1}}^{a_{i,N-1}} x dF(x) + a_{i-1,N-1} F(a_{i-1,N-1}) + a_{i,N-1} (1-F(a_{i,N-1}))$$

for $N \geq 2$,

where $-\infty \cdot 0$ and $0 \cdot \infty$ are defined to be 0.

Then we have

Lemma 2. For each $N \geq 1$ and $i=1, \dots, N$, we have

$$a_{i-1,N} \leq a_{i-1,N-1} \leq a_{i,N} .$$

Proof: The proof is by induction on N . When $N = 1$, the fact that $a_{0,1} = -\infty$ and $a_{1,1} = \infty$ yields this lemma. From the induction assumption, we have $a_{i-1,N-1} \leq a_{i-1,N-2} \leq a_{i,N-1} \leq a_{i,N-2}$, and the calculation yields $a_{i-1,N} - a_{i-1,N-1} \leq 0$ and $a_{i-1,N-1} - a_{i,N} \leq 0$. Then we obtain this Lemma.

For simplicity, we call the game $G_n(p_1, \dots, p_n, r, s)$ where $n+r+s=N$ as N-game. We prove the following theorem by induction on N . When $N = 1$, this game is $G_1(p_1, 0, 0)$ and the following theorem is valid, i.e., $a_{0,1} = -\infty$, $a_{1,1} = \infty$ and $v_0(0, 0, 0) = 0$, $v_1(p_1, 0, 0) = p_1 a_{1,2} = p_1 \mu$.

Theorem 1. For each $N \geq 1$, concerning the sequence $\{a_{i,N}\}_{0 \leq i \leq N}$ defined above, we get the following facts for the game $G_n(p_1, \dots, p_n, r, s)$ ($n+r+s=N$).

i) The optimal strategy of player I is:

$$\text{if } \begin{cases} x \leq a_{r,N} \\ a_{r+i-1,N} < x \leq a_{r+i,N} \\ a_{r+n-1,N} < x \end{cases} \text{ then } \begin{cases} \text{reject} \\ \text{assign } p_i \quad (i=1, \dots, n-1) \\ \text{assign } p_n \end{cases}$$

ii) The optimal strategy of player II is:

$$\begin{aligned} \text{if } x \leq a_{n+r,N} & \text{ then accept} \\ \text{otherwise} & \text{ reject.} \end{aligned}$$

iii) The value of the N-game $G_n(p_1, \dots, p_n, r, s)$ where $n+r+s=N$ is as follows.

$$(4) \quad v_n(p_1, \dots, p_n, r, s) = \sum_{i=1}^n p_i a_{r+i, N+1} \cdot$$

Proof: We employ the induction on N and we assume **ii)** for $N=0$ concerning the induction assumption. We assume that **iii)** for $N-1$ and $N-2$ are valid. Now we point the fact from Lemma 2 that

$$-\infty = a_{0, N} \leq a_{1, N} \leq \dots \leq a_{N-1, N} \leq a_{N, N} = \infty \cdot$$

Here we use the following notation. In the N-game $G_n(p_1, \dots, p_n, r, s)$ ($n+r+s=N$), let

$$\alpha = v_n(p_1, \dots, p_n, r, s-1),$$

$$(6) \quad \beta = v_n(p_1, \dots, p_n, r-1, s),$$

$$\gamma = \max_{k=1, \dots, n} \{p_k x + v_{n-1}(p_1, \dots, p_{k-1}, p_{k+1}, \dots, p_n, r, s)\},$$

then the equation (2) can be expressed as

$$v_n(p_1, \dots, p_n, r, s) = E \text{ val} \left(\begin{array}{cc} v_n(p_1, \dots, p_n, r-1, s-1) & \beta \\ \alpha & \gamma \end{array} \right)$$

From the inductive hypothesis, we get

$$\alpha = \sum_{i=1}^n p_i a_{r+i, N} \text{ and } \beta = \sum_{i=1}^n p_i a_{r+i-1, N} \cdot$$

On the other hand the game $G_n(p_1, \dots, p_n, r-1, s-1)$ is the $(N-2)$ -game and we obtain

$$v_n(p_1, \dots, p_n, r-1, s-1) = \sum_{i=1}^n p_i a_{r+i-1, N-1} \cdot$$

Using Lemma 2, $a_{i-1, N} \leq a_{i-1, N-1} \leq a_{i, N}$ ($i=1, \dots, N$), yields

$$\beta \leq v_n(p_1, \dots, p_n, r-1, s-1) \leq \alpha.$$

This means that player I stands advantageous if one of player II's available rejection is transferred to player I for his own use. We obtain the following Lemma 3.

Lemma 3. If both players play optimally, they will never both reject the same observation values.

We remark here: Lemma 3 contains the lemma stated in Sakaguchi [8] and it says there as follows. "This fact, when $p_i=1$ ($i=1, \dots, n$), is the conjecture that can be proven by induction but is very roundabout to be ascertained." This time we can prove Lemma 3 in proving Theorem 1. The similar cases are seen in Sakaguchi [7] and Brams and Davis [3], etc.

Proof of i) and ii) of Theorem 1: First we prove for the case $rs \neq 0$. From the inductive hypothesis, $v_{n-1}(p'_1, \dots, p'_{n-1}, r, s) = \sum_{i=1}^{n-1} p'_i a_{i+r, N}$ where

the p'_1, \dots, p'_{n-1} represent the remaining $(n-1)$ p 's of the original n p 's after the first, and we get from Lemma 3,

$$(7) \quad v_n(p_1, \dots, p_n, r, s | x) = \min \{ \alpha, \max \{ \beta, \gamma \} \} .$$

1) When $a_{i+r-1, N} < x \leq a_{i+r, N}$ ($i=1, \dots, n$). We get from Lemma 2 and Hardy's lemma

(Lemma 1),

$$\begin{aligned} \gamma &= \max_{k=1, \dots, n} \{ p_k x + \sum_{j=1}^{k-1} p_j a_{j+r, N} + \sum_{j=k+1}^n p_j a_{j+r-1, N} \} \\ &= p_i x + \sum_{j=1}^{i-1} p_j a_{j+r, N} + \sum_{j=i+1}^n p_j a_{j+r-1, N} . \end{aligned}$$

From the comparison of γ and β, α , we get $\beta < \gamma \leq \alpha$. Then we have

$v_n(p_1, \dots, p_n, r, s | x) = \gamma$ and the optimal strategy of player I (resp. player II) is to accept and to select p_i (resp. to accept).

2) When $x \leq a_{r, N}$. We have $\gamma \leq \beta \leq \alpha$, and $v_n(p_1, \dots, p_n, r, s | x) = \beta$. The optimal strategy of player I (resp. player II) is to reject (resp. to accept) the observation.

3) $x > a_{r+n, N}$. We get $\beta \leq \alpha \leq \gamma$ and $v_n(p_1, \dots, p_n, r, s | x) = \alpha$. The optimal strategy of player I (resp. player II) is to select the n -th man p_n (resp. to reject) for the observed value of the random variable.

Next we consider the boundary cases where $rs=0$. When $r=0$, we get the recursive equation as $v_n(p_1, \dots, p_n, 0, s | x) = \min\{\alpha, \gamma\}$ and when $s=0$, we have $v_n(p_1, \dots, p_n, r, 0 | x) = \max\{\beta, \gamma\}$. So we can obtain the optimal strategy and the value of the N -game $G_n(p_1, \dots, p_n, r, s)$ ($rs=0$ and $N=r+s+n$) similarly. This completes the proof of i) and ii) of Theorem 1.

Proof of the part iii) of Theorem 1: From the inductive hypothesis, i) and ii) are valid for the N -game $G_n(p_1, \dots, p_n, r, s)$ ($n+r+s=N$). Suppose that $\{a_{i, N}\}_{0 \leq i \leq N}$ be the sequence of the numbers used in i) and ii).

From i) and ii), concerning the N -game $G_n(p_1, \dots, p_n, r, s)$, under the optimal strategy of the players, the i -th man p_i is assigned in the initial stage whenever the random variable X is contained in the interval

$(a_{r+i-1, N}, a_{r+i, N}]$. Hence we can consider the value of $a_{i+r, N+1}$ as the expected value of the quantity to which the i -th smallest p is assigned in the N -game $G_n(p_1, \dots, p_n, r, s)$ under the optimal strategy. (See Remark 1)

The total expected payoff of this N -game is given by

$$v_n(p_1, \dots, p_n, r, s) = \sum_{i=1}^n p_i a_{r+i, N+1} . \quad (n+r+s=N)$$

Hence we have the fact iii) of Theorem 1. This completes the proof of Theorem 1.

Remark 1. In the N-game $G_n(p_1, \dots, p_n, r, s)$, under the optimal strategy, the quantity to which the i -th smallest p is assigned is as follows. If x is contained in the interval $(a_{r+i-1, N}, a_{r+i, N}]$, this p is assigned and the quantity is x . If the value x is less than or equal to $a_{r+i-1, N}$, then this p is not assigned in this stage and the expected quantity to which this p is assigned in the following (N-1)-game is $a_{r+i-1, N}$ because in this (N-1)-game this p is the $(i-1)$ -th smallest p or the remaining number of rejections of player I decreases by 1. If x is greater than $a_{r+i, N}$, the expected quantity to which this p is assigned in the next (N-1)-game is $a_{r+i, N}$, similar to the above case.

Remark 2. Our problem is a generalization of the sequential stochastic assignment problem discussed in Derman, Lieberman and Ross [4], mentioned in the initial section, i.e., there is a single decision maker (i.e. player I) and there are only n jobs to assign. The objective of the decision maker of their problem is to maximize the total expected reward from assigning n men to n jobs sequentially.

Here we point the fact that the result of Theorem 1 is the same style as that of the result of their paper [4]. Theorem 1 includes the result of Derman, Lieberman and Ross, and when $s=0$, the equation (5) of this paper is equal to the equation (8) of their paper [4].

Remark 3. The special case where the p 's are equal to 1 is considered. This game is studied as the time sequential game for sums of the random variables discussed in Sakaguchi [8]. Theorem 1 of this paper is the same as Theorem 1 of Sakaguchi [8] in this special case. If we write the game $G_n(1, \dots, 1, r, s)$ as $G_n(r, s)$ and let the value of the game $G_n(r, s)$ be $v_n(r, s)$, we point the following fact from Theorem 1.

$$(8) \quad v_n(r, s) = \sum_{i=1}^n a_{r+i, N+1}, \quad (N=r+s+n)$$

and $v_n(r-1, s) - v_{n-1}(r, s) = a_{r, N}$, $v_n(r, s-1) - v_{n-1}(r, s) = a_{r+n, N}$. The recursive relation of the paper [8] can be derived from the calculation by using the above result. Therefore we consider Theorem 1 of this section as the extension of the theorem in the paper [8]. In the following we observe the property of these game when N gets large.

Next we observe the asymptotic property of this game about the case stated in Remark 3. Consider the special version in which each man has an associated p 's equal to 1. In the following we use the notation $G_n(r, s)$ and $v_n(r, s)$ which are defined in Remark 3.

Suppose $0 < k_1 < k_2 < 1$, and let $s = N - [k_2 N]$, $r = [k_1 N]$, and $n = N - r - s$.

Then $v_n(r,s)$ of this N-game $G_n(r,s)$ is (see (8))

$$(9) \quad v_n(r,s) = \sum_{i=[k_1 N]+1}^{[k_2 N]} \alpha_{i,N+1}.$$

Here we state the following theorem of Albright and Derman [2].

Theorem 2. (Albright and Derman) If F is a continuous function, then for each $0 < \pi < 1$, we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=[n\pi]+1}^n \alpha_{i,n+1} = \int_{F^{-1}(\pi)}^{\infty} x dF(x),$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{[n\pi]} \alpha_{i,n+1} = \int_{-\infty}^{F^{-1}(\pi)} x dF(x),$$

$$\text{and } \lim_{n \rightarrow \infty} \alpha_{[n\pi],n+1} = F^{-1}(\pi),$$

where $F^{-1}(\pi)$ denotes the quantile of order π of the cdf $F(x)$.

By combining the equation (9) and Theorem 2, we can obtain the following corollary concerning the asymptotic property of this game.

Corollary 1. If F is a continuous distribution function, then for each $0 < k_1 < k_2 < 1$, we have for the N-game $G_n(r,s)$ where $N=r+n+s$, $s=N-[k_2 N]$ and $r=[k_1 N]$,

$$\lim_{N \rightarrow \infty} \frac{1}{N} v_n(r,s) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=[k_1 N]+1}^{[k_2 N]} \alpha_{i,N+1}$$

$$= \int_{F^{-1}(k_1)}^{F^{-1}(k_2)} x dF(x).$$

Acknowledgement.

The author wishes to thank Professor M. Sakaguchi of Osaka University for his guidance and encouragement and the referees for their helpful comments.

References.

- [1] Albright, S.C.: "A Markov chain version of the secretary problem", *Naval Research Logistics Quarterly*, vol.23 (1976), pp. 151-159.
- [2] Albright, S.C. and Derman, C.: "Asymptotic optimal policies for the stochastic sequential assignment problem", *Management Science*, vol.19 (1972), pp. 46-51.
- [3] Brams, S.J. and Davis, M.D.: "Optimal jury selection: A game theoretic model for the exercise of peremptory of challenges", *Operations Research*, vol.26 (1978), pp. 966-981.

- [4] Derman, C., Lieberman, G.J. and Ross, S.M.: " A sequential stochastic assignment problem", *Management Science*, vol.18 (1972), pp. 349-355.
- [5] Gilbert, J. and Mosteller, F.: " Recognizing the maximum of a sequence", *Journal of the American Statistical Association*, vol.61 (1966), pp. 35-73.
- [6] Hardy, G.H., Littlewood, J.E. and Polya, G.: *Inequalities*, Cambridge University Press, 1934.
- [7] Sakaguchi, M.: "A bilateral sequential game for sums of bivariate random variables", *Journal of the Operations Research Society of Japan*, vol.21 (1978), pp. 486-508.
- [8] Sakaguchi, M.: "Note: A time sequential game for sums of random variables", *Journal of the Operations Research Society of Japan*, vol.23 (1980), pp. 81-89.

Tōru NAKAI: Department of Mathematics,
College of Integrated Arts and Sciences,
University of Osaka Prefecture, Sakai,
Osaka, 591, Japan.