Math Meth Oper Res (1999) 49:111–123

Mathematical Methods
of Operations Research

The Elimination algorithm for the problem of optimal stopping

Isaac Sonin

Department of Mathematics, University of North Carolina at Charlotte, Charlotte, NC 28223, USA (e-mail: imsonin@email.uncc.edu)

Abstract. We present a new algorithm for solving the optimal stopping problem. The algorithm is based on the idea of elimination of states where stopping is nonoptimal and the corresponding correction of transition probabilities. The formal justification of this method is given by one of two presented theorems. The other theorem describes the situation when an aggregation of states is possible in the optimal stopping problem.

Key words: Markov chain, optimal stopping, the elimination algorithm, the state reduction approach, secretary problem

1 Introduction

The optimal stopping problem (OSP) is one of the most refined and extensively studied fields in the general theory of stochastic control. It was initiated in the context of optimal stopping times in Bayes decision procedures by classical works of Wald (1947), Wald and Wolfowitz (1948). They were followed by numerous publications of many authors. The general problem of optimal stopping of random processes in discrete time was formulated by Snell in 1953. One of the most interesting chapters of this theory is the so called "Secretary Problem" (see one of the pioneering works Chow, Moriguti, Robbins and Samuels (1964), surveys of Rose (1982), Freeman (1983), Ferguson (1989), Samuels (1991), and proceedings of conferences Amherst (1991) and Nagoya (1994)). Most books on stochastic control discuss OSP, see Çinlar (1975), Bertsekas (1987), Puterman (1994), Feldman and Valdez-Flores (1996) and many others. The resurgent interest in OSP and corresponding outpouring of publications is due mainly to demands from three areas: sta-

Manuscript received: April 1997/final version received: July 1998

tistical applications which require the termination of observations at some moment (see the important paper of Berger and Berry (1994); computational algorithms for which traditional stopping rules are inadequate, e.g. simulated annealing; and relatively recently, the appearance of new problems from options pricing theory.

There are two different, though essentially equivalent, approaches to OSP, usually called "the martingale theory of OSP of general stochastic sequences (processes)" (formulated by Snell) and "the OSP of Markov chains", represented by the two classical monographs Chow, Robbins and Sigmund (1971) and Shiryayev (1978). (See also Dynkin and Yushkevich (1969)). They lay the theoretical background which, in the case of finite or countable state spaces, has a relatively simple basic structure which will be briefly described in Section 3.

The main goal of this paper is to present a new method for solving the OSP of Markov chains with countable state spaces initially proposed in Sonin (1995). We call this new method the *elimination algorithm (EA)* because it is based on the idea of elimination of the states where stopping is nonoptimal in the one step problem and the corresponding recomputation of transition probabilities.

A better understanding of the EA algorithm is provided by the framework of what we shall call the State Reduction Approach. This approach is based on the simple probabilistic idea that appeared in the pioneering works of Kolmogorov and Döeblin more than sixty years ago, and described in Proposition 1 below. This idea has been used since in Probability Theory in several contexts and on numerous occasions. As a basis for computational algorithms it has been in use starting from 1985, when a new algorithm for the sequential calculation of the invariant distribution of a Markov chain was proposed independently in Sheskin (1985) and in Grassman, Taksar and Heyman (1985). Later it become known as the GTH algorithm. Taking into account the short but very precise paper of Sheskin we call this algorithm GTH/S. The computational properties of this algorithm and some generalizations have been studied in numerous papers. We refer the reader interested in this subject to the volume Meyer and Plemmons (1993) or to the forthcoming paper Sonin (1999) where the further development of the SR Approach and its relation to the tree decomposition is discussed. We briefly describe this approach in Section 2 as a necessary step for the EA.

As some examples show, at least in some situations the EA is better and simpler to apply than such well-known methods of solving the OSP as the value iteration method or linear programming (for the finite case). A more detailed comparison of our elimination method with traditional ones and computational aspects of its application will be presented in a subsequent paper.

In Section 3 we present all necessary definitions and statements about the optimal stopping problem and a brief review of the existing algorithms.

In Section 4 the EA is presented, Theorem 1 is proved and three illustrative examples are given. Section 5 contains the proof of basic Lemma 1 and some other statements. In Section 6 Theorem 2 is presented. Its proof is omitted as rather simple. In Section 7 we consider one more example, the Classical Secretary Problem.

The author would like to thank Joseph Quinn, Albert Shiryayev, Alexander Yushkevich and Ernest Presman who read the first version of this paper and made valuable comments.

2 The state reduction approach

The general idea of The State Reduction Approach is based on a sequential application of the following simple probabilistic statement, which is well-known though under different forms.

Let us assume that a Markov model M_1, i.e. a pair (X_1, P_1), where X_1 is a countable state space and $P_1 = \{p_1(x, y)\}$ is a transition matrix, is given and let (Z_n) be a Markov chain specified by the model M_1. Let $D \subset X_1$ and let us introduce the sequence of Markov times $\tau_1, \tau_2, \ldots, \tau_n, \ldots$, the moments of first, second and so on visits of (Z_n) to the set $X_1 \backslash D$, i.e. $\tau_1 = \min\{k > 0, Z_k \in (X_1 \backslash D)\}$, $\tau_{n+1} = \min\{k > \tau_n, Z_k \in (X_1 \backslash D)\}, 1 \le \tau_1 < \tau_2 < \cdots$. Let $X_2 \equiv (X_1 \backslash D)$ and consider the random sequence $Y_n = Z_{\tau_n}, n = 1, 2, \ldots$. (For the sake of brevity we assume that $\tau_n < \infty$ for all $n = 1, 2, \ldots$ with probability one. Otherwise we can complement X_2 by an additional absorbing point z_* and correspondingly modify transition probabilities participating in Proposition 1). Let us denote by $u_1(z, X_2, \cdot)$ the distribution of the Markov chain (Z_n) for the initial model M_1 at the moment τ_1 of first visit to set X_2 (first exit from D) starting at z, $z \in D$. The strong Markov property and standard probabilistic reasoning imply

Proposition 1. *(a) The random sequence (Y_n) is a Markov chain in a model $M_2 = (X_2, P_2)$, where*

(b) the transition matrix $P_2 = \{p_2(i, j)\}$ is given by the formula

$$p_2(x, y) = p_1(x, y) + \sum_{z \in D} p_1(x, z) u_1(z, X_2, y), \quad (x, y \in X_2). \tag{1}$$

We call model M_2 the D-reduced model of M_1.

An important case is when the set D consists of one nonabsorbing point z. In this case formula (1) obviously takes the form

$$p_2(x, \cdot) = p_1(x, \cdot) + p_1(x, z)(1 - p_1(z, z))^{-1} p_1(z, \cdot), \quad (x \in X_2). \tag{2}$$

According to this formula the row-vectors of the new stochastic matrix P_2 are linear combinations of row vectors of the truncated initial stochastic matrix P_1.

It is clear that although (Z_n) and (Y_n) are two different Markov chains, having two different state spaces and transition matrices, some of their characteristics will coincide or will have a simple relation. We formulate only one lemma which will be used as an auxiliary result in the sequel.

Lemma 1. *Let the models M_1, M_2 be defined as above, $G \subset X_2 = X_1 \backslash D$, and $\tau = \tau_G$, $(\tau' = \tau'_G)$ be the moment of first visit to G in the model $M_1, (M_2)$. Then for any $x \in X_2$*

$$u_1(x, G, y) = u_2(x, G, y), \quad (x \in X_2, y \in G). \tag{3}$$

We will prove Lemma 1 in Section 5, but now we will discuss its practical implications.

The reduced model M_2 has fewer states, and we can expect that the calculation of corresponding characteristic of the new Markov chain, for example $u_2(x, G, y)$, will be simpler in this model. If the calculation is still difficult, we can eliminate some subset of the state space again, and so on, until we reach a model where they can be calculated.

Generally we also should calculate at each step the term corresponding to $u_1(z, G, y)$ in (1). In a finite model, when we eliminate one state at a time, the calculation of this term and the whole recursive procedure become very simple since on each step we have a formula similar to (2), i.e. a linear transformation of a stochastic matrix. If, for example $|X_1 \backslash G| = m$, i.e. there are m states outside of G, Lemma 1 implies that after $(m-1)$ elimination steps we reduce the initial model M_1 to the model $M_m = (X_m, P_m), X_m = G \cup \{x_m\}$ and for any $y \in G$ we have $u_1(x_m, G, y) = u_2(x_m, G, y) = \cdots = u_m(x_m, G, y)$. But since the model M_m has only one state x_m outside of G, then obviously $u_m(x_m, G, y) = p_m(x_m, y)/(1 - p_m(x_m, x_m))$. More elaborate reasoning shows also that by using the same calculations and reversing the elimination steps we can get the distribution $u_1(x, G, y)$ for other initial points.

The relationship between the invariant distributions in an initial model and a model with one state eliminated can be described similarly. This relationship serves as a basis for the GTH/S algorithm which consists of two stages. On the first stage it constructs a sequence of stochastic matrices, each having dimension one less than the previous. After that using the relations between invariant distributions in two subsequent models they are calculated in reverse order.

3 The optimal stopping problem (OSP)

Let the triplet $M = (X, P, g)$ denote a problem of optimal stopping (of a Markov chain) with countable state space X, transition matrix $P = \{p(x, y)\}$ and *reward function* g. Let us denote by P_x, E_x the measure and expectation for a Markov chain (Z_n) specified by X and P and an initial point x. We will suppress the x, when the initial point is clear. Let v be a *value function* for this OSP, i.e. $v(x) = \sup E_x g(Z_\tau)$, where the sup is taken over all stopping times τ, $(\tau < \infty)$. Let $Tf(x) = \sum_y p(x, y) f(y)$ be the *averaging operator*. It is well-known that the value function v satisfies the Bellman (optimality) equation

$$v(x) = \max(g(x), Tv(x)), \tag{4}$$

and that v is the minimal excessive function which majorizes function g, i.e. the minimal function satisfying $v(x) \geq g(x), v(x) \geq Tv(x)$ for all $x \in X$, (in the terminology of the martingale approach $v(Z_n)$ is the Snell's envelope). If $G \subseteq X$ let us denote by τ_G the moment of first visit to G, i.e. $\tau_G = \min(n \geq 1 : Z_n \in G)$. We call a set S an *optimal stopping set* if $S = \{x : v(x) = g(x)\}$ and $P_x\{\tau_S < \infty\} = 1$ for all $x \in X$. It is known that if such a set S exists then $\tau \equiv \tau_S$ is an optimal stopping time and $v(x) = E_x g(Z_\tau)$. To simplify the presentation we will assume that in all OSP's under consideration the optimal stopping sets do exist, though this assumption can be relaxed. This assumption always holds in the case of finite state space X.

Basically there are three methods of solving OSP.

The first one can be conventionally called "the direct solution of the Bellman equation", and is applied in situations when the transition probabilities have a very simple structure or when they have a very specific structure. An example of the first kind is a simple random walk on the finite interval with absorbing end points. In this case the excessive majorant $v(x)$ is the smallest convex upward function "spanning function $g(x)$ from above". An example of the second kind is the Classical Secretary Problem which can be reduced to the OSP with $X = \{1, 2, \ldots, n\}$ and a Markov chain that satisfies the "monotonicity" property, $p(x, y) = 0$ if $y \leq x$. In this case the value of $v(x)$ can be calculated sequentially starting from $v(n)$ in inverse order. More generally, in any OSP in the finite time interval $[0, n]$ one can include time as a coordinate of the state space and after that using this space-time representation to apply backward induction to the sequential calculation of the value function.

The second method is the value iteration method when instead of solving the equation (4) one considers the sequence of functions $v_n(x)$ satisfying the relations

$$v_{n+1}(x) = \max(g(x), Tv_n(x)), v_0(x) = g(x). \tag{5}$$

In this case $v_0(x) \leq v_1(x) \leq \cdots$ and the sequence $v_n(x)$ converges to $v(x)$. Of course this is equivalent to considering the OSP on the finite time interval $[0, n]$ and after that letting $n \to \infty$.

The third well-known method is applicable when the state space X is a finite set. In this case, using the Bellman equation it is easy to prove that $v(x)$ is a solution of the linear programming problem

$$\min \sum_{y \in X} v(y), \textit{subject to } v(x) \geq \sum_{y} p(x, y)v(y), v(x) \geq g(x), \quad x \in X. \tag{6}$$

There is also an original approach to OSP described in Davis and Karatzas (1994) which provides an interesting interpretation of the Doob-Meyer decomposition of the Snell's envelope but it seems difficult to use for computational purposes.

The three methods described above have their advantages, drawbacks and limitations. Briefly they are the following. The first one is the best in all cases where it can be applied, but there are obvious limitations to its application. The second one is a universal, but the convergence of $v_n(x)$ to $v(x)$ can be very slow even in a state space with a few points if there is a cycle that is followed with probability nearly one. A further difficulty is that the presence of such cycles is often unknown in advance of calculation. The third method can be applied only for finite X and is unrelated to the essence of the problem. In the case where the state space contains only four points two of which are absorbing, the first method is not applicable and the second one works badly if two nonabsorbing points form a cycle of the type just described. Linear programming is applicable, but cumbersome to use without software.

The new algorithm, besides other useful properties, easily solves examples of such type using a few simple calculations.

4 The Elimination algorithm

This algorithm is based on the following simple consideration. Though in OSP it may be difficult to find the states where it is optimal to stop, it is relatively easy to find a state (states) where it is optimal *not to stop*. For example, it is optimal not to stop at all states where $Tg(\cdot) > g(\cdot)$, i.e. the expected reward of doing *one more step* is larger than the reward from stopping. Generally, it is optimal not to stop at any state where the expected reward of doing some, perhaps random, number of steps, is larger than the reward from stopping. Now we can exclude such states and recalculate the transition matrix using (2) if one state is eliminated or (1) if a larger subset of the state space is eliminated. According to Lemma 1 that will keep the distributions at the moments of visits to any subset of remaining states the same and the excluded states do not matter since it is not optimal to stop there. After that in the reduced model we can repeat the first step and so on. The formal justification of the transition from model M_1 to the model M_2 is given by Theorem 1 which we prove here assuming that Lemma 1 is true.

Theorem 1 (Elimination theorem): *(Sonin (1995). Let $M_1 = (X_1, P_1, g)$ be an optimal stopping problem, $D \subseteq \{z \in X_1 : T_1 g(z) > g(z)\}$ and $P_{1,x}(\tau_{X_1 \setminus D} < \infty) = 1$ for all $x \in D$. Consider an optimal stopping problem $M_2 = (X_2, P_2, g)$ with $X_2 = X_1 \setminus D$, $p_2(x, y)$ defined by* (1). *Let S be the optimal stopping set in problem M_2. Then S is the optimal stopping set in the problem M_1 also and $v_1(x) = v_2(x)$ for all $x \in X_2$.*

Proof: Let S_i be the optimal stopping sets in models $M_i, i = 1, 2$ (with reward function g), defined in Theorem 1. From the definition of D, G and S_1 it follows that $S_1 \cap D = \emptyset$ and hence $S_1 \subset X_2 = X_1 \setminus D$. Let us notice now that from Lemma 1 it follows that under its assumptions for any bounded function f we have

$$E_{1,x} f(x_\tau) = E_{2,x} f(x_{\tau'}). \tag{7}$$

Hence for $f \equiv g$, $\tau = \tau_{S_1}$, $\tau' = \tau'_{S_1}$ and for $x \in X_2$ we have $v_2(x) \geq E_{2,x} g(x_{\tau'}) = E_{1,x} g(x_\tau) = v_1(x)$. On the other hand, obviously $S_2 \subset X_2$ also and by (7) for $f \equiv g, \tau = \tau_{S_2}, \tau' = \tau'_{S_2}$ and for $x \in X_2$ we have $v_1(x) \geq E_{1,x} g(x_\tau) = E_{2,x} g(x_{\tau'}) = v_2(x)$. Thus $v_1(x) = v_2(x)$ for all $x \in X_2$. Hence $S_2 = \{x : g(x) = v_2(x) = v_1(x)\} = S_1$. Thus S_2 is the optimal stopping set for the model M_1 also. Theorem 1 is proved.

Let $M_1 = (X_1, P_1, g)$ be an OSP with finite $X_1 = \{x_1, \ldots, x_m\}$ and T_1 be a corresponding averaging operator. If each time only one state will be eliminated the EA has an especially simple structure. Its implementation consists of the sequential application of two basic steps. The first is to calculate the differences $g(x_i) - T_1 g(x_i), i = 1, 2, \ldots, m$ until the first state occurs where this difference is negative. If there is no such state, i.e. if all differences are nonnegative, it means that function $g(x)$ is an excessive function (for stochastic matrix P_1) and therefore $g(x)$ is a minimal excessive majorant of $g(x)$, i.e. $g(x) = v(x)$ for all x and X_1 is a stopping set. Otherwise there is a state, let us say z, where $g(z) < T_1 g(z)$. This implies (by (4)) that $g(z) < v(z)$ and hence z is not in the stopping set. Then we apply the second basic step of EA: we consider the new, "reduced" model of OSP $M_2 = (X_2, P_2, g)$ with state set $X_2 = (X_1 \backslash \{z\})$ and transition probabilities $p_2(x, y), x, y \in X_2$, recalculated by (2). By Theorem 1 this will guarantee that the stopping set in the reduced model M_2 coincides with optimal stopping set in the initial model M_1.

Now we repeat both steps in the model M_2, i.e. check the differences $g(x) - T_2g(x)$ for $x \in X_2$, where T_2 is an averaging operator for stochastic matrix P_2, and so on. Obviously, in no more than $(m-1)$ steps we shall come to the model $M_k = (X_k, P_k, g)$, where $g(x) - T_kg(x) \geq 0$ for all $x \in X_k$ and therefore X_k is a stopping set in this and in all previous models, including the initial model M_1.

Remark 1: It can be shown that by reversing the elimination algorithm we also can calculate recursively the values of $v(x)$ for all $x \in X_1$.

Remark 2: The idea of solving OSP by excluding some of the states can be traced to Dynkin and Yushkevich (1969) and Presman and Sonin (1972), where the Secretary problem and the Secretary problem with a random number of applicants were treated. In these papers all the states where the (nonnegative) reward function g was equal zero were excluded and the new transition probabilities were calculated.

Remark 3: We exclude points where $Tg(z) > g(z)$ because this condition can be easily checked. Generally, all points z, where $E_zg(x_\tau) > g(z)$ for some $\tau \equiv \tau(z)$, are candidates for elimination.

Remark 4: The idea of recursive elimination of some subsets of a state space can also be applied to the OSP in continuous time, for example for diffusion processes, where the inequality $Tg > g$ can be replaced by the inequality for the infinitesimal operator A. The determination of the form and the calculation of the appearing boundary conditions is an open problem.

As an illustration of this technique in this section we consider three examples. The first is taken from a comprehensive monograph on Markov Decision Processes Puterman (1994, example 7.2.6).

Example 1: In this example $X_1 = \{x_1, x_2, x_3, x_4, x_5\}$, $p_1(x_1, x_1) = p_1(x_1, x_2) = .1$, $p_1(x_1, x_4) = .6$, $p_1(x_1, x_5) = .2$, $p_1(x_2, x_1) = .2$, $p_1(x_2, x_3) = .3$, $p_1(x_2, x_5) = .5$, $p_1(x_3, x_3) = p_1(x_4, x_4) = 1$, $p_1(x_5, x_2) = .1$, $p_1(x_5, x_3) = .3$, $p_1(x_5, x_4) = .6$, and $g(x_1) = 6$, $g(x_2) = 1$, $g(x_3) = g(x_4) = 0$, $g(x_5) = 1$.

The traditional approach (applied in Puterman (1994)) to solving this problem is to find the value function $v = v_1$ solving the linear program, which follows immediately from the Bellman equation (1). Minimize $v(x_1) + v(x_2) + v(x_3) + v(x_4) + v(x_5)$ subject to $.9v(x_1) - .1v(x_2) - .6v(x_4) - .2v(x_5) \geq 0$, $-.2v(x_1) + v(x_2) - .3v(x_3) - .5v(x_5) \geq 0$, $-.1v(x_2) - .3v(x_3) - .6v(x_4) + v(x_5) \geq 0$, and $v(x_1) \geq 6, v(x_2) \geq 1, v(x_5) \geq 1$, and $v(x_i) \geq 0$ for $x_i \in X$. Using the simplex algorithm, the value function is found to be $v(x_1) = 6$, $v(x_2) = 1.7$, $v(x_3) = v(x_4) = 0$ and $v(x_5) = 1$.

Using the Elimination method we have on the first step $T_1g(x_i) \leq g(x_i)$ for $i = 1, 3, 4, 5$ and $T_1g(x_2) > g(x_2)$. Thus the state x_2 can be eliminated and the new transition probabilities according to (2) are $p_2(x_1, x_1) = .1 + (.1)(.2) =$

.12, $p_2(x_1, x_5) = .2 + (.1)(.5) = .25$, $p_2(x_5, x_1) = (.1)(.2) = .02$, $p_2(x_5, x_5) = (.1)(.5) = .05$. (We do not have to calculate other transition probabilities since at other points $g(x) = 0$). On the second stage we can easily check that now $T_2 g(x_i) \le g(x_i)$ for all remaining points $i = 1, 3, 4, 5$. Hence set $\{1, 3, 4, 5\}$ is the optimal stopping set in this and the initial problem and v at these points coincides with g. For x_2 we have $v(x_2) = T_1 v(x_2) = 1.7$.

In our examples 2A and 2B the state set X is *countable*. The first of these examples is a classical example solved in (Shiryayev (1978), Sect. 2.6). This is a simplified version of the OSP that appears in option pricing theory. The third is a generalization of this example when a discount factor is introduced.

Example 2A: Let us consider the problem of optimal stopping of a random walk on a line with $X = \{0, \pm 1, \pm 2, \ldots\}$ and with transition probabilities $p_1(x, x+1) = p$, $p_1(x, x-1) = q$, $p + q = 1$, and reward function $g(x) = \max(0, x)$. It is easy to see that in the case of $p \ge q$ there is no optimal stopping time and $v(x) = \infty$ for all x. So we consider the case when $c = (q/p) > 1$. We have $g(x) - T_1 g(x) \equiv x - (p(x+1) + q(x-1)) = q - p > 0$ for all points $x > 0$, $g(0) - T_1 g(0) = -p < 0$ and $g(x) - T_1 g(x) = 0$ for all points $x < 0$. Slightly modifying our algorithm on the first step we exclude all points $D = \{x : x < 0\}$, though instead of the strict inequality $g(x) < Tg(x)$ at these points we have only the equality $g(x) = Tg(x) = 0$. Let $\tau = \tau_{X \setminus D}$ be the moment of first visit to 0. It is known (see, for example Shiryayev (1984) that $P_x(Z_\tau = 0) \equiv \beta(x) = c^x < 1$ for $x < 0$. Hence we also must introduce an absorbing point z_*. Thus $X_2 = \{0, 1, 2, \ldots\} \cup z_*$ and by formula (2) $p_2(x, \cdot) = p_1(x, \cdot)$ for all $x > 0$ and $p_2(0, 1) = p_1(0, 1) = p$, $p_2(0, 0) = p_1(0, -1)\beta(-1) = qc^{-1} = p$, $p_2(0, z_*) = 1 - (p_2(0, 0) + p_2(0, 1)) = 1 - 2p$. Note that at the model $M_2 = (X_2, p_2)$ the only candidate for elimination is the extreme left point, $x_2 = 0$ because for all other points the new transition probabilities p_2 are the same and hence $g(x) - T_2 g(x) = g(x) - T_1 g(x) = q - p > 0$. At point $x_2 = 0$ we have $g(x_2) - T_2 g(x_2) = 0 - p \cdot 1 < 0$ and hence this point should be eliminated. Similarly, at each subsequent model M_n, $n = 3, \ldots$ we have $X_n = \{n-2, n-1, \ldots\} \cup z_*$ and $p_n(x, \cdot) = p_1(x, \cdot)$ for all $x > x_n = n - 2$. Thus again the only candidate for elimination is the left point x_n and for this point we have $p_n(x_n, x_n + 1) = p$, $p_n(x_n, x_n) = p_{n-1}(x_n, x_n - 1) \times p_{n-1}(x_n - 1, \ x_n)(1 - p_{n-1}(x_n - 1, x_n - 1))^{-1} = q(1-p)^{-1} p = p$. Therefore $g(x_n) - T_n g(x_n) = x_n - p x_n - p(x_n + 1)$ and the elimination process will continue until we reach point $x_n = x_* = \min(x : x - 2px - p \ge 0)$. At this stage $(n = x_* + 2)$ we have $g(x) - T_n g(x) \ge 0$ for all remaining points and hence the optimal stopping set is $X_n = \{x_*, x_* + 1, \ldots\}$. The same answer in slightly different form is given in Shiryayev (1978).

Example 2B: This is a previous example with a discount factor $\rho < 1$.

First of all let us notice that the discounted case with transition probabilities $p^*(x, y)$ can be treated as undiscounted if we introduce the absorbing point z^* and the new transitional probabilities $p(x, y) = \rho p^*(x, y)$ for $x, y \in X$, $p(x, z^*) = 1 - \rho, p(z^*, z^*) = 1$ and will assume that $g(z^*) = 0$ whereas at all other states $g(x) \ge 0$. Thus our problem is equivalent to the problem when there is no discount but $p + q = \rho < 1$ and there exists an absorbing point z^*.

We have

$$g(x) - T_1 g(x) \equiv x - (p(x+1) + q(x-1)). \tag{8}$$

Obviously this expression is positive for all sufficiently large points $x > 0$. We also have that $g(x) - T_1 g(x) = 0$ for all points $x < 0$. Again on the first step we exclude all points $D = \{x : x < 0\}$. Let $\tau = \tau_{X \setminus D}$ be the moment of first visit to 0. It is known in the case of $p + q = 1$ (see, for example Shiryayev (1984)) that for $x < 0$ we have $P_x(Z_\tau = 0) \equiv \beta(x) = 1$ if $c = (q/p) \leq 1$ and $\beta(x) = c^x < 1$ if $c = (q/p) > 1$. In the case of $p + q = \rho < 1$ similar reasoning shows that now in both cases $\beta(x) = d_2$, where d_2^x is the largest root of the quadratic equation

$$pd^2 - d + q = 0. \tag{9}$$

(It can be checked that $d_2 > 1$). Therefore $X_2 = \{0, 1, 2, \ldots\} \cup z^*$ and by formula (2) $p_2(x, \cdot) = p_1(x, \cdot)$ for all $x > 0$ and $p_2(0, 1) = p_1(0, 1) = p$, $p_2(0, 0) = p_1(0, -1)\beta(-1) = q/d_2$, $p_2(0, z^*) = 1 - (p_2(0, 0) + p_2(0, 1)) = 1 - p - q/d_2$. For all other points the new transition probabilities p_2 are the same and hence $g(x) - T_2 g(x) = g(x) - T_1 g(x)$. At point $x_2 = 0$, using also (9), we have $g(x_2) - T_2 g(x_2) = x_2 - p(x_2 + 1) - x_2 q/d_2 = p[x_2(d_2 - 1) - 1] = -p$ and since this expression is negative, this point should be eliminated. Similarly, at each subsequent model M_n, $n = 3, \ldots$ we have $X_n = \{n-2, n-1, \ldots\} \cup z^*$ and $p_n(x, \cdot) = p_1(x, \cdot)$ for all $x > x_n = n - 2$. Hence for all such points $g(x) - T_n g(x)$ coincide with (8). It is easy to see that the transition probabilities for the left point x_n are the same for all n. Therefore $g(x_n) - T_n g(x_n) = p[x_n(d_2 - 1) - 1]$ and the elimination process will continue until we reach the point $x_n = x_* = \min(x : x(d_2 - 1) - 1 \geq 0)$. It can be checked that (8) will be positive for all $x > x_*$. Therefore the optimal stopping set is $X_n = \{x_*, x_* + 1, \ldots\}$.

An important feature of EA is that it can also be applied easily to the generalized random walk with $p_1(x, x+1) = p(x)$, $p_1(x, x-1) = q(x)$, $p_1(x, x) = r(x)$, $p + q + r = 1$, $x = 0, \pm 1, \ldots$. At each step we can eliminate all intervals where $g(x) - T_n g(x) < 0$ and after that recalculate the new transition probabilities for the endpoints of the eliminated intervals, applying the well-known formulas for the probabilities of the first exit to the left or right end of the interval. Note also that the process of elimination can be carried out in parallel mode.

5 The proof of Lemma 1

To prove Lemma 1 we need to complement Proposition 1 by some other simple but useful statements.

Let us assume that a Markov model $M_1 = (X_1, P_1)$ is given and let $D \subset X_1$. Consider the (reduced) model $M_2 = (X_2, P_2)$, with $X_2 = (X_1 \setminus D) \cup z_*$, where z_* is an absorbing point and transition function $p_2(x, y)$ is defined by the formula (2). Let us denote by $P_{i,x}, E_{i,x}$ the measure and expectation for a Markov chain with an initial point x in model M_i, $i = 1, 2$. We again will suppress the x, when it is clear what the initial point is. Let us assume for simplicity that $P_{1,x}(\tau_{X_1 \setminus D} < \infty) = 1$ for all $x \in D$.

Denote

$H_i = \{(x_1x_2 \dots x_s \dots), x_s \in X_i\}$, the set of all (infinite) trajectories in model $M_i, i = 1, 2$ and let $\mathscr{B}_i, i = 1, 2$ are corresponding Borel σ-algebras. Let us introduce the projection mapping

$F : H_1 \to H_2$ by $F(x_1x_2 \dots x_s \dots) = (x_{\tau_1} \dots x_{\tau_k}, \dots)$, where $\tau_1, \tau_2, \dots, \tau_n, \dots$, the moments of first, second and so on visits to the set $X_1 \backslash D$, and $x_{\tau_k} = z_*$ if $\tau_k = \infty$. Thus all $x_i \in D$, $1 \leq i$, are deleted from the trajectory $(x_1x_2 \dots x_s \dots.)$ in H_1.

The proposition 1 and the definition of mapping F imply immediately

Proposition 2. *Let the models M_1, M_2 and mapping F be defined as above. Then for any set $B \in \mathscr{B}_2$*

$$P_2(B) = P_1(F^{-1}(B)). \tag{10}$$

Let $L_i, i = 1, 2$ are two functionals defined on H_i, $i = 1, 2$ and let them satisfy the condition

$$L_1(x_1x_2 \dots x_s \dots) = L_2(F(x_1x_2 \dots x_s \dots)) \quad \text{for all trajectories} \quad (x_1x_2 \dots x_s \dots) \in H_1. \tag{11}$$

Proposition 3. *Let the models M_1, M_2 be defined as above and $L_i, i = 1, 2$ be functionals satisfying condition (11). Then for any $x \in X_2$*

$$E_{1,x}L_1 = E_{2,x}L_2. \tag{12}$$

Proof: The statement is obviously true for all functionals having the form $L_2(\cdot) = I_B(x_1x_2 \dots x_s \dots), B \in \mathscr{B}_2, L_1(\cdot) = L_2(F(x_1x_2 \dots x_s \dots))$, where $I_B(\cdot)$ is a characteristic function of a set B. After that (12) follows easily by a limit transition. Now we can easily prove Lemma 1.

Proof of Lemma 1: Let an initial model M_1 and reduced model M_2 are given, $G \subset X_1 \backslash D$ and τ, (τ') is the moment of first visit to set G in model $M_1, (M_2)$. Let us define functional L_1 on $H_1 : L_1(x_1x_2 \dots x_s \dots) = f(Z_\tau)$. Let us define functional L_2 on H_2 similarly replacing τ by τ'. Then as easy to see condition (11) is fulfilled and proposition 3 immediately implies Lemma 1.

6 Reduction theorem

Our second theorem deals with the situation when in the initial OSP all states can be divided into disjoint classes that have the property that for any class K the probability of transition from each state in K to any other class are the same for all states in K and the reward function is a constant inside of each of these classes. In the terminology of Howard (1971) the Markov chain with the first property is a 'mergeable Markov chain'. The corresponding partition in our case is the inverse image partition for some mapping f. Statements similar to the Theorem 2 about classes of sufficient strategies can be found in different fields of stochastic control. We present this theorem in a form convenient for OSP.

Theorem 2. *Let $M_1 = (X_1, P_1, g)$ and $M_2 = (X_2, P_2, g)$ be two optimal stopping problems and let f be a mapping of X_1 onto X_2 such that*

(a) $P_1(x, f^{-1}(y)) = p_2(f(x), y)$ *for all* $x \in X_1, y \in X_2$,
(b) $g_1(x) = g_2(f(x))$ *for all* $x \in X_1$.

Then (i) $v_1(x) = v_2(f(x))$ *for all* $x \in X_1$,
(ii) if S_2 *is an optimal stopping set for the problem* M_2 *then* $S_1 = \{f^{-1}S_2\}$ *is an optimal stopping set for the problem* M_1.

We omit the proof of this theorem as rather simple. The only technical difficulty is the necessity to consider randomized stopping times.

7 Secretary problem (Example 3)

We now present a short and a rigorous proof of the solution of the Classical Secretary Problem. Its formulation is as follows.

We wish to select one applicant from a set of n rankable applicants, i.e. one of them is "best", one is second best, etc. The applicants are interviewed sequentially in random order, each order being equally likely. The decision to accept or reject an applicant must be based only on the *relative* ranks of those applicants interviewed so far. An applicant once rejected cannot later be recalled. The problem consists of maximizing the probability of choosing the best applicant. Let us call an applicant a *leader* if he/she is better than all those who precede him/her. The optimal strategy is to choose the next leader after the moment $(k_n - 1)$, where $k_n = (\min k : 1/k + 1/(k+1) + \cdots + 1/(n-1) \le 1)$, $(\lim_n k_n/n = e)$.

As we mentioned in the Introduction there is a vast literature about numerous generalizations of the Secretary Problem. The solution of the Secretary Problem itself is very simple but nevertheless two points in the proof of the optimality of the described strategy are usually presented vaguely or skipped as rather obvious (as they really are). The first point is that one can skip all past information about the relative ranks of all applicants other than leaders because the relative ranks are independent random variables. The second point is that one can transform the initial model where at moment k one observes the kth *applicant* (in this form the first published solution of the Secretary problem was given by Lindley (1961)) to the model where at moment k one observes the kth *leader* (in this form the solution was presented in Dynkin and Yushkevich (1969)). Now we can provide an absolutely rigorous solution of this problem.

In our initial setting, the information that we have at moment k is a tuple $(y_1, y_2, \ldots, y_k)$, where y_i is the relative rank of an applicant who appears at moment $i = 1, 2, \ldots, n$. Hence the states of the initial Markov chain are the tuples $(y_1), (y_1, y_2), \ldots, (y_1, y_2, \ldots, y_n)$. It is well known that the assumption that all $n!$ permutations of applicants are equally probable implies a simple combinatorial lemma which states that:

1) the relative ranks $y_1, y_2, \ldots, y_n$ are independent random variables,
2) $P(y_k = i) = 1/k, i = 1, \ldots, k$, and

3) $g(y_1, y_2, \ldots, y_k) \equiv P$ (a kth applicant is the best among all applicants $|y_1, y_2, \ldots, y_k) = P(y_{k+1} > 1, \ldots, y_n > 1 | y_1, y_2, \ldots, y_k) = k/n$ if $y_k = 1$ and 0 otherwise.

Now let us introduce the model $M_2 = (X_2, p_2(x, y), g)$ with state space $X_2 = \{(i, k), i = 0, 1; k = 1, \ldots, n\}$, the transition probabilities $p_2((j, k-1), (i, k)) = 1/k$ if $i = 1$, and $p_2((j, k-1), (i, k)) = (k-1)/k$ if $i = 0$, and with a reward function $g(i, k) = k/n$ if $i = 1$ and zero otherwise. (Thus all relative ranks larger than one are compressed to zero). Let us consider the mapping $f(y_1, y_2, \ldots, y_k) = (1, k)$ if $y_k = 1$ and $(0, k)$ otherwise. Now using 1)–3) it is easy to check that the assumptions of Theorem 2 are fulfilled and hence instead of M_1 we can consider the simpler model M_2. In this form the Secretary problem was solved in Lindley (1961). The next step is to eliminate all points $x = (0, k)$, where $g(x) = 0$ and where hence $T_2 g(x) > g(x)$. In the new model M_3 the remaining states have the form $(1, k) \equiv k, k = 1, 2, \ldots, n$; and using 1) it is easy to obtain that the new transition probabilities are $p_3(k, s) = p_2((1, k), (0, k+1)) \Pi_{i=k+1}^{s-2} p_2((0, i), (0, i+1)) \cdot p_2((0, s-1), (1, s)) = k/(s-1)s$, $s = k+1, \ldots, n$. The OSP for this model can be solved using backward induction (the solution presented in Dynkin and Yushkevich (1969)) or by applying Theorem 2 once more. In the latter case we see that if k_n is defined as in the Introduction then $T_3 g(k) = \sum_{s=k+1}^{n} p_3(k, s) g(s) = (1/k + \cdots + 1/(n-1)) k/n \leq g(k) = k/n$ for all $k \geq k_n$ and the inverse inequality otherwise. Hence we can eliminate all points $k < k_n$ and for the remaining points the transition probabilities will be unchanged. (This of course is a very specific property of the Secretary Problem where transitions are possible only in one direction). Therefore the set $\{k_n, k_n + 1, \ldots, n-1, n\}$ is the optimal stopping set in the initial problem also and $v(x) = g(x)$ for all x from this set. For other x the value function can be calculated in the usual way using the latter equality and the form of $p_3(k, s)$.

References

Bertsekas DP (1987) Dynamic programming: Deterministic and stochastic models. Prentice-Hall, Englewood Cliffs, NJ

Berger JO, Berry DA (1986) The relevance of stopping rules in statistical inference. With discussion by Robert Keener, Adam T. Martinsek and Carl N. Morris and a reply by the authors. In: Statistical decision theory and related topics, IV, Vol. 1 Springer, New York-Berlin, pp. 29–72

Çinlar E (1975) Introduction to stochastic processes. Prentice-Hall, Englewood Cliffs, NJ

Chow YS, Moriguti S, Robbins H, Samuels SM (1964) Optimal selection based on optimal rank (the "secretary" problem). Israel J. Math. 2:81–90

Chow YS, Robbins H, Sigmund D (1971) Great expectations: The theory of optimal stopping. Houghton Mifflin Co., NY

Davis MHA, Karatzas I (1994) A deterministic approach to optimal stopping. In: Kelly FP (ed.) Probability, statistics and optimization. Wiley & Sons, NY

Dynkin EB, Yushkevich AA (1969) Markov processes. Theorems and problems. Plenum Press, NY

Feldman RM, Valdez-Flores C (1996) Applied probability and stochastic processes. PWS Publishing Company, Boston

Ferguson TS (1989) Who solved the secretary problem? Statistical Science 4:282–296

Freeman PR (1983) The secretary problem and its extensions; A review. Int. Statistical Review 51: 189–206

Gilbert J, Mosteller F(1966) Recognizing the maximum of a sequence. J. Amer. Stat. Assoc. 61: 35–73

Grassmann WK, Taksar MI, Heyman DP (1985) Regenerative analysis and steady-state distributions for Markov chains. Operat. Res. 33:1107–1116

Howard RA (1971) Dynamic probabilistic systems: v.1 Markov models. Wiley&Sons, NY

Kemeny JG, Snell JL (1960) Finite Markov chains. Van Nostrand-Reinhold, NY

Lindley DV (1961) Dynamic programming and decision theory. Appl. Statist. 10:39–52

Meyer CD, Plemmons RJ (Eds.) (1993) Linear algebra, Markov chains, and queueing models. IMA Vol. Math. Appl., 48, Springer, New York

Presman EL, Sonin IM (1972) The best choice problem for a random number of objects. Theory Probab. Applic. 17:657–668

Puterman M (1994) Markov decision processes: Discrete stochastic dynamic programming. John Wiley & Sons, NY

Rose JS (1982) Twenty years of secretary problem: a survey of developments in the theory of optimal choice. Adv. in Management Studies 1:53–64

Samuels SM (1991) Secretary problems. In: Ghosh BK, Sen Pk (eds.) Handbook of sequential analysis, Marcel Dekker, Boston, pp. 381-4-5

Sheskin TJ (1985) A Markov partitioning algorithm for computing steady-state probabilities. Operat. Res. 33:228–235

Shiryayev AN (1978) Optimal stopping rules. Springer-Verlag, NY

Shiryayev AN (1984) Probability. Springer-Verlag, NY

Sonin IM (1995) Two simple theorems in the problems of optimal stopping. Proceed. INFORMS Appl. Probab. Conf., Atlanta, Georgia

Sonin IM (1998) The state reduction and related algorithms and their applications to the study of Markov chains, graph theory and the optimal stopping problem. Accepted for publication in Advances in Mathematics

Stochastic models and optimal stopping (1994) Proceed. Int. Conf., Nanzan Univ., Nagoya, Japan

Strategies for Sequential Search and Selection in Real Time (1992) Amer. Math. Soc., Providence, RI, In: Bruss F, Ferguson T, Samuels S. (eds.) Contemp. Math. 125, Amherst, MA