The Optimal Stopping of Markov Chain and Recursive Solution of Poisson and Bellman Equations

Isaac Sonin

Dept. of Mathematics, Univ. of North Carolina at Charlotte, Charlotte, NC, 28223, USA imsonin@email.uncc.edu

Summary. We discuss a modified version of the Elimination algorithm proposed earlier by the author to solve recursively a problem of optimal stopping of a Markov chain in discrete time and finite or countable state space. This algorithm and the idea behind it are applied to solve recursively the discrete versions of the Poisson and Bellman equations.

Key words: Markov chain, Optimal stopping, The Elimination Algorithm, The State Reduction approach, discrete Poisson and Bellman equations.

Mathematics Subject Classification (2000): 60J22, 62L15, 65C40

To Albert N. Shiryaeyev - one of the pioneers of Optimal Stopping, with appreciation and gratitude, on the occasion of his 70th birthday. 20 Aug

1 Introduction

The main goal of this paper is to present a unified recursive approach to the following two related but nevertheless different problems.

Problem 1. To find the solution f of the discrete *Poisson equation*

$$f = c + Pf, \tag{1}$$

where $Pf(x) = \sum_y p(x, y)f(y)$ is the averaging operator, defined by a transition matrix P, and c is a given function defined on a countable (finite) state space X.

Problem 2. To solve the problem of optimal stopping (OS) for a Markov chain (MC) with an immediate reward (one-step cost) function c and terminal reward function g. This means to describe an optimal strategy (or ε-optimal strategies if there is no optimal strategy), and to find the value function v, which is a minimal solution of a corresponding *Bellman (optimality) equation*

$$v = \max(g, c + Pv). \tag{2}$$

The main tool to study these problems in this paper is the recursive algorithm for Problem 2, which we call the *Elimination algorithm* (EA), described in the papers of the author ([13]) (see also ([12]) and ([14])). We present EA here in a modified form, and we prove also a new important Lemma 3. We limit our presentation to the case of a finite state space though one of the advantages of our approach is that in many cases it can be applied also to the countable state space. This algorithm is better understood in the context of a group of algorithms which are based on a similar idea and can be called the State Reduction algorithms. We will refer to this idea as the State Reduction (SR) approach. Problem 1 was analyzed earlier in Sheskin (1999) [8] on the basis of this approach, (see also the references there to the earlier works of Kohlas (1986) and Heyman and Reeves (1989)).

Note, that formally, Problem 1 can be considered as a special case of Problem 2 when $g(x) = -\infty$ but we will treat them separately. We start with Problem 2.

The author would like to thank Robert Anderson who read the first version of this paper and made valuable comments.

2 Optimal stopping of a MC.

The optimal stopping problem (OSP) is one of the most developed and extensively studied fields in the general theory of stochastic control. There are two different approaches to OSP, usually called "the martingale theory of OSP of general stochastic sequences (processes)" (formulated by Snell) and "the OSP of Markov chains". The first one is represented by the classical monographs Chow, Robbins and Sigmund (1971) [2] (see also the book of T. Fergusson on his website for a modern presentation). The second approach is represented by Shiryayev (1969,1978), [9],[10]. (See also Dynkin and Yushkevich (1969), [4]). There are also tens of books and monographs with chapters or sections on OSP (see e.g. [1], [7]), and more than a thousand papers on this topic. These two approaches basically represent nonstationary and stationary (nonhomogeneous versus homogeneous) situations and though formally they are equivalent, the second approach is more transparent for study and discussion.

Formally, OSP of MC is specified by a tuple $M = (X, P, c, g, \beta)$, where X is a finite (countable) state space, $P = \{p(x, y)\}$ is a transition matrix,$c(x)$ is a *one step cost function*, $g(x)$ is a *terminal reward function*, and β is a discount factor, $0 < \beta \le 1$. We call such a model *OS model*. A tuple without a terminal reward $M = (X, P, c, \beta)$, we call *reward model*, and a tuple $M =$

(X, P), we call a *Markov model*. A Markov chain (MC) from a family of MCs defined by Markov model is denoted by (Z_n). A probabilistic measure for the Markov chain with initial point x and corresponding expectation are denoted by P_x and E_x. The *value function* $v(x)$ for OS model is defined as $v(x) = \sup_{\tau \geq 0} E_x[\sum_{i=0}^{\tau-1} \beta^i c(Z_i) + \beta^\tau g(Z_\tau)]$, where the sup is taken over all stopping times $\tau, \tau \leq \infty$. If $\tau = \infty$ with positive probability we assume that $g(Z_\infty) = 0$. We also assume that the state space X is countable (finite) and the functions c and g are such that $v(x) < \infty$ for all x. This holds e.g. if $c(x) \leq 0$ and $g(x)$ is bounded from above.

It is well-known that the discounted case can be treated as undiscounted if an absorbing point x_* and new transition probabilities are introduced, $p^\beta(x, y) = \beta p(x, y)$ for $x, y \in X$, $p^\beta(x, x_*) = 1 - \beta, p^\beta(x_*, x_*) = 1$. In other words, with probability β the Markov chain "survives" and with complimentary probability it transits to an absorbing state x_*. Further we will assume that this transformation is made and we skip the superscript β. We will aslo assume that More, all subsequent results are valid if the constant β is replaced by a function $\beta(x), 0 \leq \beta(x) \leq 1$, the probability of "survival", $\beta(x) = P_x(Z_1 \neq x_*)$.

Let $Pf(x)$ be the *averaging operator*, and $Ff(x) = c(x) + Pf(x)$ be the *reward operator*. If $G \subseteq X$ let us denote by τ_G the moment of the first visit to G, i.e., $\tau_G = \min(n \geq 0 : x_n \in G)$. The following result is the main result for OSP with countable X.

Theorem 1. (Shiryaev, [9]) (a) *The value function $v(x)$ is the minimal solution of Bellman (optimality) equation* (2), *i.e. the minimal function satisfying $v(x) \geq g(x), v(x) \geq Fv(x)$ for all $x \in X$, and $v(x) = lim_n v_n(x)$, where $v_n(x)$ is the value function for the OSP on a finite time interval of length n;*

(b) *for any $\varepsilon > 0$ the random time $\tau_\varepsilon = \min\{n \geq 0 : g(Z_n) \geq v(Z_n) - \varepsilon)\}$, is an ε - optimal stopping time;*

(c) if $P_x(\tau_0 < \infty) = 1$ then the random time $\tau_0 = \min\{n \geq 0 : g(Z_n) = v(Z_n)\}$ is an optimal stopping time;

(d) if state space X is finite then set $S = \{x : g(x) = v(x)\}$ is not empty and τ_0 is an optimal stopping time.

The classical tools to solve the OSP of a MC are: the direct solution of the Bellman equation, which is possible only for very specific MCs; the value iteration method based on the equality $v(x) = lim_n v_n(x)$, mentioned in point (a) of Theorem 1; and for finite X, the value function $v(x)$ can be found as a solution of the linear programming problem.

The Elimination Algorithm (EA) solves the finite space OS problem in no more than $|X|$ steps, and allows us also to find the distribution of MC at the moment of stopping in an optimal stopping set S, and the expected number of visits to other states before stopping. Using the EA we also can prove in a new and shorter way Theorem 1. As a byproduct we also obtain a new recursive way to solve the Poisson equation. It works also for many OSP with countable X.

Before describing the EA in the next section we describe a more general framework of the State Reduction (SR) approach. This is a brief version of a section from ([14]).

3 Recursive Calculation of characteristics of a MC and the SR approach.

Let us assume that a *Markov model* $M_1 = (X_1, P_1)$ is given and let $D \subset X_1$, $X_2 = X_1 \setminus D$. Let us denote (Z_n) a Markov chain specified by the model M_1 with an initial point $\in X_2$. Let us introduce the sequence of Markov times $\tau_0, \tau_1, ..., \tau_n, ...$, the moments of zero, first, and so on, visits of (Z_n) to the set $X_2 = X_1 \setminus D$, i.e., $\tau_0 = 0$, $\tau_{n+1} = \min\{k > \tau_n, Z_k \in (X_1 \backslash D)\}, 0 < \tau_1 < ...$. Let us consider the random sequence $Y_n = Z_{\tau_n}, n = 0, 1, 2, ...$. For the sake of brevity we assume that $\tau_n < \infty$ for all $n = 0, 1, 2, ...$ with probability one. Otherwise we can complement X_2 by an additional absorbing point x_* and correspondingly modify transition probabilities participating in Lemma 1. Let us denote by $u_1(z, \cdot) \equiv u_1(z, \cdot|X_2)$ 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 the following basic lemma of the SR approach

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

(b) *the transition matrix* $P_2 = \{p_2(x, y)\}$ *is given by the formula*

$$p_2(x, y) = p_1(x, y) + \sum_{z \in D} p_1(x, z) u_1(z, y), \quad (x, y \in X_2). \tag{3}$$

Part (a) is immediately implied by the strong Markov property for (Z_n), while the proof of (b) is straightforward. Formula (3) can be represented in matrix form (see e.g. [6]). If matrix P_1 is decomposed as follows

$$P_1 = \begin{bmatrix} Q_1 & T_1 \\ R_1 & P_1' \end{bmatrix} \tag{4}$$

where substochastic matrix Q_1 describes the transitions inside of D, P_1' describes the transitions inside of X_2 and so on, then

$$P_2 = P_1' + R_1 U_1 = P_1' + R_1 N_1 T_1. \tag{5}$$

In this formula U_1 is a matrix of the distribution of MC at the moment of first exit from D (exit probabilities matrix), and N_1 is a *fundamental matrix* for the substochastic matrix Q_1, i.e.

$$N_1 = \sum_{n=0}^{\infty} Q_1^n = (I - Q_1)^{-1}, \tag{6}$$

where I is the $|D| \times |D|$ identity matrix. Formula (6) implies obviously

$$N_1 = I + Q_1 N_1 = I + N_1 Q_1. \tag{7}$$

Both equalities in (7) have relatively simple probabilistic interpretations. The first is almost trivial statement while the second recalls the words of Kai Lai Chung "Last exit is a deeper concept than first entrance".

Given set D, matrices N_1 and U_1 are related by

$$U_1 = N_1 T_1. \tag{8}$$

We call model M_2 the X_2-reduced model of M_1. For the sake of brevity we will call two such models *adjacent*. An important case is when the set D consists of one nonabsorbing point z. In this case formula (3) obviously takes the form

$$p_2(x, \cdot) = p_1(x, \cdot) + p_1(x, z) n_1(z) p_1(z, \cdot), \quad (x \in X_2), \tag{9}$$

where $n_1(z) = 1/(1 - p_1(z, z))$.

According to this formula, each row-vector of the new stochastic matrix P_2 is a linear combination of two rows of P_1 (with the z-column deleted). For a given row of P_2, these two rows are the corresponding row of P_1 and the z^{th} row of P_1. This transformation corresponds formally to one step of the Gaussian elimination method for the solution of linear system.

If an initial Markov model $M_1 = (X_1, P_1)$, is finite, $|X_1| = k$, and only one point is eliminated each time, then, a sequence of *stochastic* matrices $(P_n), n = 2, ..., k$, can be calculated recursively on the basis of formula (9), in which the subscripts "1" and "2" are replaced by "n" and "$n+1$" respectively.

This sequence provides an opportunity to calculate many characteristics of the initial Markov model M_1 recursively starting from some reduced model $M_s, 1 < s \le k$. For this goal one needs also a relationship between a characteristic in a reduced model and a model with one more point. Sometimes this relationship is obvious or simple, sometimes it has a complicated structure.

The EA algorithm for the problem of optimal stopping (OS) of a Markov chain was developed independently of other SR algorithms and shares with them the common idea of elimination. It also has very distinct specific features. The number of points to be eliminated and the order in which they are eliminated depend on some auxiliary procedure, and the value function of the problem to be recovered on the second stage.

For the problem of OS it is natural to try to find not only the optimal stopping set but as well the distribution of the stopping moment and the distribution of a MC at the moment of stopping. The next lemma provides tools for the sequential calculation of these characteristic.

Lemma 2. (Lemma 3 in ([14])). *Let the models M_1, M_2 be defined as in Lemma 1, set $G \subset X_2 = X_1 \setminus D$, and $u_i(x, \cdot)$ be the distribution of the Markov* *chain* (Z_n) *for the model* M_i *at the moment of first visit to* G *in the model* $M_i, i = 1, 2$, *and* $n_i(x, v)$ *the mean time spent at point* v *till such a visit with an initial point* $x \in X_2 \setminus G$. *Then for any* $x \in X_2$

$$u_1(x, y) = u_2(x, y), \qquad y \in G, \tag{10}$$

$$n_1(x, v) = n_2(x, v), \qquad v \in X_2 \setminus G. \tag{11}$$

4 The Elimination Algorithm.

The Elimination algorithm for the OSP of a MC is based on the three following facts.

1. Though in the OSP it may be *difficult* to find the states where it is optimal *to stop*, it is *easy* to find a state (states) where it is optimal *not to stop*. Really, it is optimal to stop at z if $g(z) \geq c(z) + Pv(z) \equiv Fv(z)$, but v is unknown until the problem is solved. On the other side, it is optimal not to stop at z if $g(z) < Fg(z)$, 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).

2. After we have found states (state) which are not in the optimal stopping set, we can eliminate them and recalculate the transition matrix using (9) if one state is eliminated or (3) if a larger subset of the state space is eliminated. According to Lemma 2 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.

3. Finally, though if $g(z) \geq Fg(z)$ at a particular point z, we can not make a conclusion about whether this point belongs to the stopping set or not, but if this inequality is true for *all* points in the state space then we have the following simple statement

Proposition 1. *Let* $M = (X, P, g)$ *be an optimal stopping problem, and* $g(x) \geq Fg(x)$ *for all* $x \in X$. *Then* X *is the optimal stopping set in the problem* M, *and* $v(x) = g(x)$ *for all* $x \in X$.

Proposition 1 follows immediately from the Theorem 1 because the function $g(x)$ in this case is its own excessive majorant.

The formal justification of the transition from the initial model M_1 to the reduced model M_2 is given by Theorem 2 below. This theorem was formulated in Sonin 1995 [12] and its proof is given in [13] for the case when $c(x) = 0$ for all x. Here we prove this theorem in a shorter way and for any $c(x)$ but for simplicity only for the case of finite X.

Let us introduce also a *transformation of the cost function* $c_1(x)$ (or any function $f(x)$) defined on X_1 into the cost function $c_2(x)$ defined on X_2, under the transition from model M_1 to model M_2.

Given the set $D, D \subset X_1$, let τ be the moment of the first *return* to X_2, i.e. $\tau = \min(n \geq 1, Z_n \in X_2)$. Then given the function $c_1(x)$ defined for $x \in X_1$ let us define function $c_2(x)$ defined on $x \in X_2$ as

$$c_2(x) = E_x \sum_{n=0}^{\tau-1} c_1(Z_n) = c_1(x) + \sum_{z \in D} p_1(x,z) \sum_{w \in D} n(z,w) c_1(w). \tag{12}$$

In other words, the new function $c_2(x)$ represents the expected cost (reward) gained by a MC starting from point $x \in X_2$ up to the moment of first return to X_2. For a function $f(x)$ defined on a set X_1 and a set $B \subset X_1$ denote by f_B the column-vector function reduced to a set B. Then formula (12) can be written in a matrix form as

$$c_2 = c_{1,X_2} + R_1 N_1 c_{1,D}. \tag{13}$$

If the set $D = \{z\}$ then the function $c_1(x)$ is transformed as follows

$$c_2(x) = c_1(x) + p_1(x,z) n_1(z) c_1(z), \quad x \in X_2. \tag{14}$$

Remark 1. This formula was obtained earlier in Sheskin (1999).

Theorem 2. *(Elimination theorem,* ([13]) *). Let* $M_1 = (X_1, P_1, c_1, g)$ *be an OS model,* $D \subseteq C_1 = \{z \in X_1 : g(z) < F_1 g(z)\}$. *Consider an OS model* $M_2 = (X_2,\ F_2, c_2, g)$ *with* $X_2 = X_1 \setminus D$, $p_2(x,y)$ *defined by* (5), and c_2 *is defined by* (13). *Let* S *be the optimal stopping set in* M_2. *Then*

a) S *is the optimal stopping set in* M_1 *also*,

b) $v_1(x) = v_2(x) \equiv v(x)$ *for all* $x \in X_2$, *and for all* $z \in D$

$$v_1(z) = E_{1,z}[\sum_{n=0}^{\tau-1} c_1(Z_n) + v(Z_\tau)] = \sum_{w \in D} n_1(z,w) c_1(w) + \sum_{y \in X_2} u_1(z,y) v(y), \tag{15}$$

where $u_1(z,\cdot)$ *is the distribution of a MC at the moment* τ *of first visit to* X_2, and $N_1 = \{n_1(z,w), z, w \in D\}$ *is the fundamental matrix for the substochastic matrix* Q_1.

Remark 2. Using (8) formula (15) can be written in a matrix form as

$$v_D = N_1[c_{1,D} + T_1 v_{X_2}]. \tag{16}$$

If set $D = \{z\}$ then formula (15) can be written as

$$v_1(z) = n_1(z)[c_1(z) + \sum_{y \in X_2} p_1(z,y) v(y)]. \tag{17}$$

The EA algorithm can be described as a sequence of steps when each time a subset of states such that none of them belong to a stopping set, is eliminated until a stopping set is achieved. The selection of these steps in the countable case is dictated by the structure of the problem and convenience of the calculation of matrices U. The algorithm has an especially simple structure if the state space is finite, and if each time only one state is eliminated.

Let $M_1 = (X_1,\ P_1, g)$ be an OSP with finite $X_1 = \{x_1, ..., x_k\}$ and P_1 be a corresponding averaging operator. The implementation of the EA consists of two stages: reduction and backward stage. The first one, consists of sequential application of two basic steps. The first is to calculate the differences $g(x_i)-F_1g(x_i), i = 1, 2, ..., k$ until the first state occurs where this difference is negative. If all differences are nonnegative then by Proposition 1, $g(x) = v(x)$ for all x and X_1 is a stopping set. Otherwise there is a state, let say z, where $g(z) < F_1g(z)$. This implies (by (2)) 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 new, "reduced" model of OSP $M_2 = (X_2,\ P_2, c_2, g)$ with state set $X_2 = (X_1\backslash\{z\})$ and transition probabilities $p_2(x, y), x, y \in X_2$, recalculated by (5). By Theorem 2 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 $g(x) - F_2g(x)$ for $x \in X_2$, where F_2 is an averaging operator for stochastic matrix P_2, and so on. Obviously, in no more than k steps we shall come to the model $M_k = (X_k, P_k, c_k, g)$, where $g\,(x) - F_kg(x) \geq 0$ 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.

Finally, by reversing the elimination algorithm we can calculate recursively the values of $v(x)$ for all $x \in X_1$, using sequentially formula (15) or (17), starting from the equalities $v(x) = g(x)$ for $x \in S = X_k$, where k is the number of the iteration where the reduction stage of the algorithm stops.

In the next section we obtain some useful formulas relating $g(\cdot) - F_ig(\cdot)$ in two adjacent models (Lemma 3). After that we prove Theorems 3 and 4 that serve as a basis for the recursive solution of Poisson and Bellman equations and give an opportunity to prove easily Theorems 1 and 2.

5 Recursive solution of a Poisson equation.

First we prove Lemma 3 which was not described in the original version of EA.

Lemma 3. *Let M_1 and M_2 be two adjacent models with state spaces X_1* and $X_2 = X_1 \setminus D$, where $D \subseteq X_1$, *P_i and $F_i, i = 1, 2$ be the corresponding averaging and reward operators, where functions c_1 and c_2 are related by* (13), *matrices R_1, T_1 are as in* (4) *and matrix N_1 is a fundamental matrix for Q_1. Let f be the function defined on X_1. Then*

$$(f - P_2f)_{X_2} = (f - P_1f)_{X_2} + R_1N_1(f - P_1f)_D, \tag{18}$$

$$f_D = N_1[T_1 f_{X_2} + (f - P_1 f)_D]. \tag{19}$$

The formula similar to (18) *holds if operators* P_i *are replaced by operators* F_i, *i.e.*

$$(f - F_2 f)_{X_2} = (f - F_1 f)_{X_2} + R_1 N_1 (f - F_1 f)_D. \tag{20}$$

Remark 3. If set $D = \{z\}$ these formulas take the form ($x \in X_2$)

$$(f - P_2 f)(x) = (f - P_1 f)(x) + p_1(x, z) n_1(z)(f - P_1 f)(z), \tag{21}$$

$$f(z) = n_1(z)(\sum_{y \in X_2} p_1(z, y) f(y) + f(z) - P_1 f(z)), \tag{22}$$

$$(f - F_2 f)(x) = (f - F_1 f)(x) + p_1(x, z) n_1(z)(f - F_1 f)(z). \tag{23}$$

Proof. Using (5) we have for $x \in X_2$

$$P_2 f_{X_2} = (P_1' + R_1 N_1 T_1) f_{X_2}.$$

Subtracting and adding from the right side $R_1 f_D$ and using (see (4)) the trivial equality $P_1 f_{X_2} = R_1 f_D + P_1' f_{X_2}$ we obtain

$$-P_2 f_{X_2} = -(P_1 f)_{X_2} + R_1[I f_D + N_1(-T_1 f_{X_2} + Q_1 f_D - Q_1 f_D]. \tag{24}$$

Formula (4) implies

$$(P_1 f)_D = Q_1 f_D' + T_1 f_{X_2}. \tag{25}$$

Using this equality, the equality $I + N_1 Q_1 = N_1$ (see (7)), and adding $f(x)$ to both sides of (24) we obtain (18).

To prove (19) note that the equality (25) implies that

$$(I - Q_1) f_D = T_1 f_{X_2} + (f - P_1 f)_D. \tag{26}$$

Multiplying both sides of this equality by $N = (I - Q)^{-1}$ we obtain (19).

Using equality $f - P_i f = f - F_i f + c_i$, formula (13) and the trivial equality $(f + g)_B = f_B + g_B$, valid for any B, we immediately obtain (20).

Remark 4. Formula (23) helps also to organize the recursive steps of the EA in a more efficient way.

Now we can prove Theorem 3.

Theorem 3. *Let* M_1 *and* M_2 *be two adjacent models with state spaces* X_1 and $X_2 = X_1 \setminus D$, where $D \subseteq X_1$, *with corresponding averaging operators* P_1 *and* P_2, *and matrices* R_1, T_1, N_1. *Let* c_1 *be a function defined on* $X_1 = X_2 \cup D$, and c_2 *be a function defined on* X_2 *by formula* (13). *Then*

(a) *if a function* f *satisfies an equation*

$$f = c_1 + P_1 f \equiv F_1 f \tag{27}$$

on X_1 *then its restriction to* X_2 *satisfies an equation*

$$f = c_2 + P_2 f \equiv F_2 f \tag{28}$$

and the restrictions f_{X_2} *and* f_D *are related by formula (see* (19))

$$f_D = N_1(T_1 f_{X_2} + c_{1,D}). \tag{29}$$

(b) *if a function* f *satisfies an equation* (28) *on* X_2 *and at points* $z \in D$ *satisfies the equality* (29) *then function* f *satisfies an equation* (27) *on* X_1.

Proof. Point (a) follows from (20) applied to a function f that satisfies (27).

To prove (b) note that (29) implies the equality

$$N_1^{-1} f_D = T_1 f_{X_2} + c_{1,D}.$$

Since $N_1^{-1} = I - Q_1$, using (25) we obtain $f_D = (P_1 f)_D + c_{1,D}$, i.e. f satisfies an equation (27) on D. This equality, (20), and (28) imply (27) on X_2. Thus (27) holds for all $x \in X_1$.

Remark 5. If set $D = \{z\}$ then formula (29) takes the form

$$f(z) = n_1(z)[\sum_{y \in X_2} p_1(z,y) f(y) + c_1(z)]. \tag{30}$$

Despite its simplicity Theorem 3 immediately provides a new recursive algorithm to solve the Poisson equation (1). Given an equation (27), let us consider a sequence of models $M_n = (X_n, P_n, c_n)$, $n = 1, ..., k$, where P_n and c_n are obtained from P_1 and c_1 sequentially using correspondingly (9) and (13). Then f can be calculated by formula (30), i.e.

$$f(x_n) = n_n(x_n)[\sum_{y \in X_{n+1}} p_n(x_n, y) f(y) + c(x_n)]. \tag{31}$$

Example 1. Let $X_1 = \{1,2,3\}$ and transition probabilities are given by a matrix P_1. Then the invariant distribution $\pi(1) = \frac{12}{35}$, $\pi(2) = \frac{14}{35}$ and $\pi(2) = \frac{9}{35}$. Function $c_1(x)$ must satisfy $(c, \pi) = 0$, and is defined up to a constant factor so let say $c_1(1) = 3, c_1(2) = 2$, and $c_1(3) = -\frac{64}{9}$. By eliminating state 1 we obtain transition matrix P_2, and function $c_2, c_2(2) = 2 + \frac{1}{4}\frac{3}{2}3 = \frac{25}{8}$, and $c_2(3) = -\frac{64}{9} + \frac{1}{2}\frac{3}{2}3 = -\frac{175}{36}$. Then $P_3 = \{1\}, c_3(3) = 0$ and we can select $f(3)$ equal to any constant, e.g. $f(3) = 0$. Applying formula (31) for $n = 2$, we obtain $f(2) = \frac{8}{3}\frac{25}{8} = \frac{25}{3}$. Applying formula (31) again for $n = 1$, we obtain $f(1) = \frac{3}{2}[\frac{1}{3}\frac{25}{3} + 3] = \frac{26}{3}$. Note that function f is defined up to an additive constant c. To normalize f, i.e. to make f satisfy $(f, \pi) = 0$, we can set $f(1) = \frac{26}{3} + c, f(2) = \frac{25}{3} + c, f(3) = c$ and to find $c = -\frac{662}{105}$. Then, finally $f(1) = \frac{248}{105}, f(2) = \frac{213}{105}$, and $f(3) = -\frac{662}{105}$.

$$P_1 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{bmatrix}, \; P_2 = \begin{bmatrix} \frac{5}{8} & \frac{3}{8} \\ \frac{7}{12} & \frac{5}{12} \end{bmatrix}.$$

6 Recursive solution of a Bellman equation.

Now we return to the OSP models $M_i = (X_i, P_i, c_i, g)$. First we obtain

Theorem 4. *Let M_1 and M_2 be two adjacent models with state spaces X_1 and $X_2 = X_1 \setminus D$, where $D \subset C_1 = \{z : g(z) \le F_1 g(z)\}$ with corresponding averaging operators P_1 and P_2, and matrices R_1, T_1, N_1.*

(a) *if function f is a (minimal) solution of the Bellman equation*

$$f = \max(g, c_1 + P_1 f) \equiv \max(g, F_1 f) \tag{32}$$

on X_1 then its restriction to X_2 is a (minimal) solution of the Bellman equation

$$f = \max(g, c_2 + P_2 f) \equiv \max(g, F_2 f), \tag{33}$$

on X_2, $f = c_1 + P_1 f$ on D, and the restrictions f_{X_2} and f_D are related by formula (29)

(b) *if function f is a (minimal) solution of Bellman equation* (33) *and it is defined on D by formula (29) then f is a (minimal) solution of* (32).

Proof. If a function f satisfies (32) then $f \ge g$ and therefore $F_1 f \ge F_1 g$. Combined with the assumption that $D \subset C_1$, this implies that $f = F_1 f \ge g$ on D, i.e. $(f - F_1 f)_D = 0$. Hence by (20) $(f - F_1 f)_{X_2} = (f - F_2 f)_{X_2}$ and $F_1 f = F_2 f$ on X_2. Therefore $\max(g, F_1 f) = \max(g, F_2 f)$, i.e. f satisfies (33) on X_2 also.

Now, suppose that function f satisfies (33) and is defined on D by formula (29). This function by Lemma 3 satisfies (19). Comparing (19) and (29) we obtain $(f - P_1 f)_D = c_{1,D}$. Therefore by (20) $(f - F_1 f)_{X_2} = (f - F_2 f)_{X_2}$ and thus $\max(g, F_1 f) = \max(g, F_2 f)$, and $f \ge g$ on X_2. Applying formula (19) to functions g and f, we obtain

$$g_D = N_1[T_1 g_{X_2} + (g - P_1 g)_D, \quad f_D = N_1[T_1 f_{X_2} + (f - P_1 f)_D. \tag{34}$$

Since $(f - P_1 f)_D = c_{1,D}$ and $g \le f$ on X_2, formula (34) implies that $g_D \le f_D$ and thus $f = \max(g, F_1 f) = F_1 f)$ on D. We proved earlier that f satisfies (32) on X_2. Point b) of Theorem is proved.

Suppose that f_1 is the minimal solution of (32) and f_2 is a solution of (33). As we proved in b), function f_2 can be extended to X_1 to be a solution for (32). Then $f_1 \le f_2$.

For $c \equiv 0$, Theorem 2 was proved in Sonin (1999) using Lemma 2. Here we prove the general case for finite X differently using the fact that value function satisfies the Bellman equation.

It is sufficient to note now that the value function v_2 for the model M_2 is the minimal solution of the Bellman equation (33). Therefore, by point (b) of Theorem 4, function v_1 equal to v_2 on X_2 and defined at z by formula (30) will be the minimal solution for (32) and hence the value function for model M_1. By the assumption of Theorem 2 at point z we have $g(z) < F_1 g(z) \le F_1 v_1(z)$, and hence $v_1(z) = \max(g(z), F_1(v_1(z)) = F_1(v_1(z) > g(z)$. Therefore the optimal stopping sets $S_i = \{x : v_i(z) = g\}$ coincide for both models.

The recursive algorithm to solve the Bellman equation is basically the same as the EA for the OS problem and coincides with the algorithm to solve the Poisson equation on its backward stage, i.e both use the same formula (30). The reason for that is the fact that outside of an optimal stopping set the Bellman equation takes a form $v = c + Pv$, i.e. coincides with the Poisson equation.

Example 2. Let $X_1 = \{1, 2, 3\}$, transition probabilities are given by a matrix P_0 from an example 1, cost function $c(x)$ is: $c_1(1) = 1, c_1(2) = -.5, c_1(3) = .5$, the terminal reward function $g(x)$ is: $g(1) = -1, g(2) = 2, g(3) = 3.5$, and discount factor $\beta = .9$. We intriduce an absorbing state x_* with $c_1(x_*) = g(x_*) = 0$ and then the transiton matrix becomes matrix P_1. On the first step we consider $g(x) - F_1 g(x) \equiv g(x) - (c_1(x) + P_1 g(x))$ and obtain that $g(1) - F_1 g(1) = -3.35 < 0$ and therefore state 1 can be eliminated. After this elimination we obtain (apprixImately) transition matrix P_2, and function $c_2, c_2(2) = -.18$, and $c_2(3) = 1.14$. After the second step we obtain that $g(2) - F_2 g(2) = -.04 < 0$ and therefore state 2 can be eliminated. After this elimination we obtain (apprixImately) transition matrix P_3, and function $c_3(3) = .95$. In this model $g(3) - F_3 g(3) = .13 > 0$ and therefore optimal stopping S set in this and two previous models $S = \{3, x_*\}$, and $v(3) = g(3) = 3.5$. Now applying formula (30) for $n = 2$, we obtain $v(2) = \frac{1}{.46}[.32(3.5) - .18] = 2.043$. Applying formula (30) again for $n = 1$, we obtain $v(1) = \frac{1}{.7}[.3(2.043) + .3(3.5) + 1] = 3.804$.

$$P_0 = \begin{array}{|c|c|c|}\hline \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \hline \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \hline \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \hline \end{array}, P_1 = \begin{array}{|l|l|l|l|}\hline .3 & .3 & .3 & .1 \\ \hline .225 & .45 & .225 & .1 \\ \hline .45 & .3 & .15 & .1 \\ \hline 0 & 0 & 0 & 1 \\ \hline \end{array}, P_2 = \begin{array}{|l|l|l|}\hline .54 & .32 & .13 \\ \hline .50 & .34 & .16 \\ \hline 0 & 0 & 1 \\ \hline \end{array}, P_3 = \begin{array}{|l|l|}\hline .69 & .31 \\ \hline 0 & 1 \\ \hline \end{array}$$

Note that for other values of parameters in this example, it may be optimal not to stop at all, i.e. wait until the MC will enter the absorbing state.

Remark 6. In addition to problems 1 and 2, the EA can also serve as a foundation for the recursive algorithm with a transparent probabilistic interpretation that allows the calculation of the *Gittins index* $\gamma(x)$. For a MC starting from x, this index can be defined as the maximum expected discounted reward per expected discounted unit of time

$$\gamma(x) = \sup_{\tau > 0} \frac{E_x \sum_{n=o}^{\tau-1} \beta^n c(Z_n)}{E_x \sum_{n=o}^{\tau-1} \beta^n}, \tag{35}$$

where β be a discount factor, $0 < \beta < 1$, and τ is a stopping time.

References

1. Çinlar, E. (1975), *Introduction to Stochastic Processes,* Prentice-Hall, Englewood Cliffs, NJ.
2. Chow, Y. S., Robbins H., Sigmund D. (1971), *Great Expectations: The Theory of Optimal Stopping,* Houghton Mifflin Co., NY.
3. Davis, M., Karatzas I. (1994), A deterministic Approach to Optimal Stopping. In *Probability, Statistics and Optimization.* Ed. F.P. Kelly, Wiley&Sons, NY.
4. Dynkin, E.B., Yushkevich A.A. (1969), *Markov Processes. Theorems and Problems,* NY: Plenum Press.
5. Ferguson, T. S. *Optimal Stopping and Applications.* (Electronic text on a website).
6. Kemeny, J. G., Snell J. L. (1960), *Finite Markov chains*, Van Nostrand Reinhold, Princeton, NJ.
7. Puterman, M. L. (1994), *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley, New York.
8. Sheskin, T.J. (1999), State reduction in a Markov decision process. *Internat. J. Math. Ed. Sci. Tech.* **30,** no. 2, 167–185.
9. Shiryaev, A. N. (1969), Statistical sequential analysis: Optimal stopping rules, (Russian) Izdat. "Nauka", Moscow, 231 pp.
10. Shiryayev, A.N. (1978), *Optimal Stopping Rules*, Springer-Verlag, NY.
11. Shiryayev, A. N. (1984), *Probability*, Springer-Verlag, NY.
12. Sonin, I. M. (1995), Two simple theorems in the problems of optimal stopping, in Proc. INFORMS Appl. Prob. Conf., Atlanta, Georgia.
13. Sonin, I. M. (1999), The Elimination Algorithm for the Problem of Optimal Stopping, *Math. Meth. of Oper. Res.*, pp.111-123.
14. Sonin, I. M. (1999), The State Reduction and related algorithms and their applications to the study of Markov chains, graph theory and the Optimal Stopping problem, *Advances in Mathematics* **145**, 159-188, Princeton, NJ.