BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 1(50), 2006, Pages 31–38
ISSN 1024–7696

Dynamic Programming Approach for Solving Discrete Optimal Control Problem and its Multicriterion Version *

D. Drucioc, D. Lozovanu, M. Popovici

Abstract. Time discrete systems determined by systems of difference equations are considered. The characterizations of their optimal trajectories with given starting and final states is studied. An algorithm based on dynamic programming technique for determining such trajectories is proposed. In additional multicriterion version for considered control model is formulated and a general algorithm for determining Pareto solution is proposed.

Mathematics subject classification: 90B10, 90C35.
Keywords and phrases: Dynamic networks, discrete optimal control, game control model, Pareto solution.

1 Introduction and Problem formulation

In [1] the following discrete optimal control problem is formulated and studied.

Let L be the dynamical system with the set of the states $X \subseteq \mathbb{R}^n$ where at every moment of time $t = 0, 1, 2, \dots$ the state of L is $x(t) \in X$, $x(t) = (x_1(t), x_2(t), \dots, x_n(t)) \in \mathbb{R}^n$. The dynamics of the system L is described as follows

$$x(t+1) = g_t(x(t), u(t)),\ t = 0, 1, 2 \dots, \tag{1}$$

where

$$x(0) = x_s \tag{2}$$

is the starting point of system L and $u(t) = (u_1(t), u_2(t), \dots, u_m(t)) \in \mathbb{R}^m$ represents the vector of control parameters [2–4]. For vectors of control parameters $u(t)$, $t = 0, 1, 2$, the admissible sets $U_t(x(t))$ are given, i.e.

$$u(t) \in U_t(x(t)), t = 0, 1, 2, \dots \ . \tag{3}$$

We assume that in (1) the vector function

$$g_t(x(t), u(t)) = (g_t^1(x(t), u(t)), g_t^2(x(t), u(t)), \dots, g_t^n(x(t), u(t)))$$

is determined uniquely by $x(t)$ and $u(t)$. So, $x(t+1)$ is determined uniquely by $x(t)$ and $u(t)$ at every moment of time $t = 0, 1, 2, \dots$.

*Supported by CRDF-MRDA 008 award MOM2-3049-CS-03

Let

$$x(0), x(1), \dots, x(t), \dots \tag{4}$$

be a process generated according to (1)–(3).

For each state $x(t)$ we define the numerical determination $F_t(x(t))$ by using the following recursive formula

$$F_{t+1}(x(t+1)) = f_t(x(t), u(t), F_t(x(t))),\ \ t = 0, 1, 2, \dots$$

and

$$F_0(x(0)) = F_0.$$

In this model $F_t(x(t))$ expresses the cost of system's passage from x_0 to $x(t)$.

Optimization Problem 1. For a given T determine the vectors of control parameters $u(0), u(1), \dots, u(T-1)$, which satisfy the conditions

$$\begin{cases} x(t+1) = g_t(x(t), u(t)), t = 0, 1, 2, \dots, T-1; \\ x(0) = x_0, x(T) = x_f, \\ u(t) \in U_t(x(t)), t = 0, 1, 2 \dots, T-1; \\ F_{t+1}(x(t+1)) = f_t(x(t), u(t), F_t(x(t))), t = 0, 1, 2 \dots, T-1; \\ F_0(x(0)) = F_0 \end{cases} \tag{5}$$

and minimize the object function

$$I_{x_0 x(T)}(u(t)) = F_T(x(T)). \tag{6}$$

Optimization Problem 2. For given T_1 and T_2 determine $T \in [T_1, T_2]$ and a control sequence $u(0), u(1), .., u(T-1)$ which satisfy condition (5) and minimize the object function (6).

Remark 1 . It is obvious that the optimal solution of problem 2 can be obtained by reducing to problem 1 fixing the parameter $T = T_1$, $T = T_1 + 1, \dots, T = T_2$. By choosing the optimal value of solutions of problems of type 1 with $T = T_1, T = T_1 + 1, \dots, T = T_2$ we obtain the solution of problem 2 with $T \in [T_1, T_2]$.

It is easy to observe that a large class of dynamic optimization problems can be represented as a problem mentioned above. As example if

$$f_t(x(t), u(t), F_t(x(t))) = F_t(x(t)) + c_t(x(t), u(t)),$$

where $F_0(x_0) = 0$ and $c_t(x(t), u(t))$ represents the cost of system's passages from state $x(t)$ to state $x(t+1)$, then we obtain the discrete control problems with integral-time which are introduced and treated in [2–7]. Some classes of control problems from [2, 3] may be obtained if

$$F_0(x_0) = 1, \quad f_t(x(t), u(t), F_t(x(t))) = F_t(x(t)) \cdot c_t(x(t), u(t)),\ \ t = 1, 2, \dots$$

and if

$$F_0(x_0) = 0 \quad f_t(x(t), u(t), F_t(x(t))) = \max\{F_t(x(t)), c_t(x(t), u(t))\}.$$

In this paper we formulate the multicriterion version of the discrete control problem and derive an algorithm for determining Pareto solution. The proposed algorithm represents an extension of single objective problem and its algorithm.

2 Algorithm for determining optimal solution

Let us assume that the starting and final states are fixed, $f_t(x, u, F)$, $t = 0, 1, 2, \dots,$ are non-decreasing function with respect to the third argument, i.e. with respect to F.

$$f_t(x, u, F^{'}) \leq f_t(x, u, F^{''}) \text{ if } F^{'} \leq F^{''}. \tag{7}$$

Algorithm 1

1. Set $F_0^*(x(0)) = F_0; F_t^*(x(t)) = \infty; x(t) \in X, t = 1, 2, \dots; X_0 = \{x_0\}$.
2. For $t = 1, 2, \dots, T$ determine:

$$X_{t+1} = \{x(t+1) \in X \mid x(t+1) = g_t(x(t), u(t)),$$

$$x(t) \in X_t, u(t) \in U_t(x(t))\}$$

and for every $x(t+1) \in X_{t+1}$ determine

$$F_{t+1}^*(x(t+1)) = \min\{f_t(x(t), u(t), F_t^*(x(t))) \mid x(t+1) = g_t(x(t), u(t)),$$

$$x(t) \in X_t, u(t) \in U_t(x(t))\};$$

3. Find the sequence

$$x_T = x^*(T), x^*(T-1), x^*(T-2), \dots, x^*(1), x^*(0) = x_0,$$

$$u^*(T-1), u^*(T-2), \dots, u^*(1), u^*(0),$$

which satisfy the conditions

$$F_{T-\tau}^*(x^*(T-1)) = f_{T-\tau-1}(x^*(T-\tau-1), u^*(T-\tau-1),$$

$$F_{T-\tau-1}^*(x(T-\tau-1))),\ \tau = 0, 1, 2, \dots, T.$$

Then $u^*(0), u^*(1), u^*(2), \dots, u^*(T-1)$ represent the optimal solution of problem 1.

Theorem 1. *If $f_t(x, u, F)$, $t = 0, 1, 2, \dots, T$, are non-decreasing functions with respect to the third argument F, i.e. the functions $f_t(x, u, F)$, $t = 0, 1, 2 \dots, T$, satisfy condition (7), then the algorithm determines the optimal solution of problem 1. Moreover, an arbitrary leading part $x^*(0), x^*(0), \dots, x^*(k)$ of the optimal trajectory $x^*(0), x^*(0), \dots, x^*(k), \dots, x^*(T)$ is again an optimal one.*

Proof. We prove the theorem by using the induction principle on number of stages T. In the case $T \leq 1$ the theorem is evident. We consider that the theorem holds for $T \leq k$ and let us prove it for $T = k + 1$.

Assume by contrary that $u^*(0), u^*(1), \dots, u^*(T-2), u^*(T-1)$ is not an optimal solution of problem 1 and $u^{'}(0), u^{'}(1), \dots, u^{'}(T-2), u^{'}(T-1)$ is an optimal solution of problem 1, which differs from $u^*(0), u^*(1), \dots, u^*(T-2), u^*(T-1)$. Then $u^{'}(0), u^{'}(1), \dots, u^{'}(T-2), u^{'}(T-1)$ generate a trajectory $x_0 = x^{'}(0), x^{'}(1), \dots, x^{'}(T) = x_T$ with corresponding numerical evaluations of states

$$F^{'}_{t+1}(x^{'}(t+1)) = f_t(x^{'}(t), u^{'}(t), F^{'}_t(x^{'}(t))), t = 0, 1, 2, \dots, T-1;$$

where

$$F^{'}_0(x^{'}(0)) = F_0 \quad \text{and} \quad F^{'}_T(x^{'}(T)) < F^*_T(x^{'}(T)), \tag{8}$$

because $x^{'}(T) = x^*(T)$. According to the induction principle for problem 1 with $T-1$ stages the algorithm finds the optimal solution. So, for arbitrary $x(T-1) \in X$ we obtain the optimal evaluations $F^*_{T-1}(x(T-1))$ for $x(T-1) \in X$. Therefore

$$F^*_{T-1}(x^{'}(T-1)) \leq F^{'}_{T-1}(x^{'}(T-1)).$$

According to the algorithm

$$f_{T-1}(x^*(T-1), u^*(T-1), F^*_{T-1}(x^*(T-1))) \leq$$
$$\leq f_{T-1}(x^{'}(T-1), u^{'}(T-1), F^*_{T-1}(x^{'}(T-1))). \tag{9}$$

Since $f_t(x, u, F), t = 0, 1, 2 \dots$ are non-decreasing functions with respect to F then

$$f_{T-1}(x^{'}(T-1), u^{'}(T-1), F^*_{T-1}(x^{'}(T-1))) \leq$$
$$\leq f_{T-1}(x^{'}(T-1), u^{'}(T-1), F^{'}_{T-1}(x^{'}(T-1))). \tag{10}$$

Using (9) and (10)we obtain

$$F^*_T(x(T)) = f_{T-1}(x^*(T-1), u^*(T-1), F^*_{T-1}(x^*(T-1))) \leq$$
$$\leq f_{T-1}(x^{'}(T-1), u^{'}(T-1), F^*_{T-1}(x^{'}(T-1))) \leq$$
$$\leq f_{T-1}(x^{'}(T-1), u^{'}(T-1), F^{'}_{T-1}(x^{'}(T-1))) = F^{'}_T(x(T)),$$

i.e

$$F^*_T(x(T)) \leq F^{'}_T(x(T)),$$

which contradicts (8). So the algorithm finds the optimal solution of problem 1 with $T = k+1$. □

Theorem 2. *Let X and $U_t(x)$, $x \in X$, $t = 0, 1, 2, \dots, T-1$, be the finite sets, and $M = \max_{x \in X, t=0,1,2,\dots,T-1} |U_t(x)|$. Then the algorithm uses at most $M \cdot |X| \cdot T$ elementary operations (without operations for calculating the values of functions $f_t(x, u, F)$ for given x, u, F).*

Proof. It is sufficient to prove that at step t the algorithm uses not more than $M \cdot |X|$ elementary operations. Indeed for finding the value $F_{t+1}(x(t+1))$ for $x(t+1) \in X$ it is necessary to use $\sum_{x \in X} |U_t(x)|$ operations. Since $\sum_{x \in X} |U_t(x)| \leq |X| \cdot M$ then at step t the algorithm uses not more than $|X| \cdot M$ elementary operations. So in general the algorithm uses not more than $|X| \cdot M \cdot T$ elementary operations. □

3 The discrete optimal control problem on network

Let L be a dynamical system with a finite set of states X, and at every moment of time $t = 0, 1, 2, \dots$ the system L is described by a directed graph $G = (X, E)$, where the vertices $x \in X$ correspond to the states of L and an arbitrary edge $e = (x, y) \in E$ identifies the possibility of the system passage from the state $x = x(t)$ to the state $y = x(t + 1)$. So, the set of edges $E(x) = \{e(x, y)|(x, y) \in E\}$ originated in $x(t)$ corresponds to an admissible set of control parameters $U_t(x(t))$ which determines the next possible state $y = x(t+1)$ of L at the moment of time t. Two states x_0 and x_f are chosen, where $x_0 = x(0)$ is the starting state and $x_f = x(T)$ is the final state of system L. In addition we assume that to each edge $e = (x, y) \in E$ a cost function $c_e(t)$ is associated which depends on time and which expresses the cost of system L to pass from the state $x = x(t)$ to the state $y = x(t + 1)$ at the stage $[t, t + 1]$ (like a *transition*). For given dynamic network we regard the problem of finding a sequence of system transitions $(x(0), x(1)), (x(1), x(2)), \dots, (x(T(x_f - 1)), x(T(x_f)))$ which transfers the system from the starting state $x_0 = x(0)$ to the final state $x_f = x(T(x_f))$ with minimal integral-time cost. Like in Section 1 we will discuss two variants of problem. First when time T is fixed and second when $T \in [T_1, T_2]$. It is easy to observe that for solving these problems we can use algorithm 1. We put $F_0(x(0)) = 0$ and $F_{t+1}(x(t + 1)) = F_t(x(t)) + c_{(x(t),x(t+1))}(t)$. A more general model is obtained if for each edge $e \in E$ a function $f_{e_t}(x(t), F_t(x(t)))$ is associated. Here we put $u(t) = e_t$ and we have the same function like in Section 1, i.e. $f_t(x(t), u(t), F_t(x(t))) = f_{e_t}(x(t), F_t(x(t)))$. For the trajectory $x(0), x(1), \dots, x(t), x(t + 1), \dots$ of system passages we have the following recursive formula $F_{t+1}(x(t + 1)) = f_{e_t}(x(t), F_t(x(t)))$, $t = 0, 1, 2, \dots$, and $F_0(x(0)) = F_0$.

4 Multicriterion Discrete Control Problem: Pareto Optimum

In this section we extend the control model from Section 1 using the concept of cooperative games.

4.1 General Statement of the Problem

We assume that the dynamics of the system L is controlled by p players, who coordinate their actions using the common vector of control parameters $u(t)$. So the dynamics of the system L is described according to (1)–(3).

Let $x(0), x(1), \dots, x(t), \dots$ be a process generated according to (1)–(3) with the given vector of control parameter $u(t)$, $t = 0, 1, 2, \dots$. For each state we define the quantities $F_t^i(x(t))$, $i = 1, 2, \dots, p$, in the following way:

$$F_{t+1}^i(x(t + 1)) = f_t^i(x(t), u(t), F_t^i(x(t))), \tag{10}$$

where

$$F_0^i(x(0)) = F_0^i,\ i = 1, 2, \dots, p, \tag{11}$$

are given representations of the starting state $x(0)$ of the system L; $f_t^i(x(t), u(t), F_t^i(x(t)))$, $t = 0, 1, 2, \dots$, are arbitrary functions. So, $F_t^i(x(t))$ expresses the cost of system's passage from the state $x(0)$ to the state $x(t)$ for player i.

In this model we assume that players choose vectors of control parameters in order to achieve the final state x_f from the starting state x_0 at the moment of time $T(x_f)$, where $T_1 \le T(x_f) \le T_2$.

For the given $u(t)$ the cost of system's passage from x_0 to x_f for player i is calculated on the basis of (1)–(3), (10), (11) and we put

$$I^i_{x_0 x_f}(u(t)) = F^i_{T(x_f)}(x_f),$$

if the trajectory passes through x_f at the time moment $T(x_f)$ such that $T_1 \le T(x_f) \le T_2$; otherwise we put

$$I^i_{x_0 x_f}(u(t)) = \infty.$$

We consider the problem of finding Pareto solution $u^*(t)$, i.e. there is no other vector $u(t)$ for which

$$\Big(I^1_{x_0 x_f}(u(t)), I^2_{x_0 x_f}(u(t)) \dots, I^p_{x_0 x_f}(u(t))\Big) \le$$

$$\le \Big(I^1_{x_0 x_f}(u^*(t)), I^2_{x_0 x_f}(u^*(t)) \dots, I^p_{x_0 x_f}(u^*(t))\Big)$$

and for any $i_0 \in \{1, 2, \dots, p\}$

$$I^{i_0}_{x_0 x_f}(u(t)) < I^{i_0}_{x_0 x_f}(u^*(t)).$$

4.2 Multicriterion Problem on Network and Algorithm for its Solving on T-Partite Networks

We formulate the multicriterion control model on network in general form on the basis of the control model from Section 3.

Let $G = (X, E)$ be a directed graph of transactions for the dynamical system L with the given starting state $x_0 \in X$ and the final state $x_f \in X$. In addition, for the state x_0 starting representations $F_0^1(x_0) = F_0^1$, $F_0^2(x_0) = F_0^2$, $\dots$, $F_0^p(x_0) = F_0^p$ are given, which express the payoff functions of players at the time moment $t = 0$. We define the control u^* on G as a map

$$u : (x, t) \to (y, t+1) \in X_G(x) \times \{t+1\} \quad \text{for} \quad x \in X \setminus \{x_f\},\ t = 1, 2, \dots.$$

For an arbitrary control u we define the quantities:

$$I^1_{x_0 x_f}(u),\ I^2_{x_0 x_f}(u),\ \dots, I^p_{x_0 x_f}(u)$$

in the following way.

Let

$$x_0 = x(0), x(1), x(2), \dots, x(T(x_f)) = x_f$$

be a trajectory from x_0 to x_f generated by control u, where $T(x_f)$ is the time moment when the state x_f is reached. Then we put

$$I^i_{x_0x_f}(u) = F^i_{T(x_f)}(x_f) \quad \text{if} \quad T_1 \le T(x_f) \le T_2,\ i = \overline{1,p},$$

where $F^i_t(x(t))$ are calculated recursively by using the following formula

$$F^i_{t+1}(x(t+1)) = f^i_{(x(t),x(t+1))}(x(t), F^i_t(x(t))),\ t = \overline{0, T(x_f)-1};$$

$$F^i_0(x(0)) = F^i_0,$$

where $f^1_e(\cdot,\cdot),\ f^2_e(\cdot,\cdot),\ \dots,\ f^p_e(\cdot,\cdot)$ are arbitrary functions. If $T(x_f) \notin [T_1, T_2]$ then we put

$$I^i(u) = \infty,\ i = \overline{1,p}.$$

We regard the problem of finding Pareto solution u^*.

In the following let us show that if the graph G has the structure of $(T+1)$-partite graph and $T_1 = T_2 = T$, then the algorithm from Sect. 2 can be extended for the multicriterion control problem on network.

So, assume that the vertex set X is represented as $X = Z_0 \cup Z_1 \cup \dots \cup Z_T$, $Z_i \cap Z_j = \emptyset$, $i \ne j$, and the edge set E is divided into T non-empty subsets $E = E_0 \cup E_1 \cup \dots \cup E_{T-1}$ such that an arbitrary edge $e = (y,z) \in E_t$ begins in $y \in Z_T$ and enters $z \in Z_{t+1}$, $t = \overline{0, T-1}$.

In this case for the nondecreasing function $f^i_e(\cdot,\cdot)$ with respect to the second argument the values $I^i(u) = F^i_t(x_t)$ can be calculated by using the following algorithm.

Algorithm 2

Preliminary step (Step 0): For the starting position $x(0) = x_0$ set $F^i_0(x(0)) = F^i_0$, $i = \overline{1,p}$; for any $x \in X \setminus \{x_0\}$ put $F^i_t(x(t)) = \infty$, $i = \overline{1,p}$, $t = \overline{1,T}$.

General step (Step t, $t \ge 0$): For an arbitrary state $x(t+1) \in X_{t+1}$ find a vertex $x'(t) \in X_t$ such that there is no other vertex $x(t) \in X_t \setminus \{x_f\}$ for which

$$\Big(f^1_{(x(t),x(t+1))}(x(t), F^1_t(x(t))), f^2_{(x(t),x(t+1))}(x(t), F^2_t(x(t))), \dots$$

$$\dots, f^p_{(x(t),x(t+1))}(x(t), F^p_t(x(t)))\Big) \le$$

$$\le \Big(f^1_{(x'(t),x(t+1))}(x'(t), F^1_t(x'(t))), f^2_{(x'(t),x(t+1))}(x'(t), F^2_t(x'(t))), \dots$$

$$\dots, f^p_{(x'(t),x(t+1))}(x'(t), F^p_t(x'(t)))\Big)$$

and

$$f^{i_0}_{(x(t),x(t+1))}(x(t), F^{i_0}_t(x(t))) < f^{i_0}_{(x'(t),x(t+1))}(x'(t), F^{i_0}_t(x'(t)))$$

for any $i_0 \in \{1, 2, \dots, p\}$.

Then calculate

$$F^i_{t+1}(x(t+1)) = f^i_{(x'(t),x(t+1))}(x'(t), F^i_t(x'(t))),\ i = \overline{1,p}.$$

If $t < T-1$ then go to the next step; otherwise STOP. □

If $F^i_t(x(t))$ are known for every vertex $x(t) \in X$ then Pareto optimum u^* can be found starting from the end position x_f by fixing each time $u^*(x(t)) = x(t+1)$ for which

$$F^i_{t+1}(x(t+1)) = f^i_{(x(t),x(t+1))}(x(t), F^i_t(x(t))),\ i = \overline{1,p}.$$

References

[1] Lozovanu D., Pickl S. *General approach for an algorithmic solution of discrete decision problems and its game-theoretical extension.* Central European Journal of Operation Research (submitted for publication).

[2] Bellman R., Kalaba R. *Dynamic programming and modern control theory.* Academic Press, New York-London, 1965.

[3] Boltyanskii V.G. *Optimal control of discrete systems.* Moscow, Nauka, 1973.

[4] Krabs W., Pickl S. *Analysis, Controllability and Optimization of Time-Discrete Systems and Dynamical Games.* Lecture Notes in Economics and Mathematical Systems, Springer Verlag, Berlin-Heidelberg, 2003.

[5] Lozovanu D.D. *The network models of discrete optimal control problem and dynamic games with p players.* Discrete mathematics and Applications, 2001, N 4(13), p. 126–143.

[6] Lozovanu D., Pickl S. *Polynomial time algorithm for determining optimal strategies.* Electronic Notes in Discrete Mathematics, 2003, **13**, p. 154–158.

[7] Lozovanu D., Pickl S. *A Special Dynamic Programming technique for Multiobjective Discrete Control and for Dynamic Games on Graph – Based Networks.* CTWO4 Workshop on Graphs and Combinatorial Optimization, Milano, 2004, p. 184–188.

D. Lozovanu
Institute of Mathematics and Computer Science
Academy of Sciences of Moldova
Academiei str. 5, MD-2028 Chisinau
Moldova
E-mail:*lozovanu@math.md*

Received February 3, 2006