15th Triennial World Congress, Barcelona, Spain

SET-VALUED CONTROL LAWS IN TEV-DC CONTROL PROBLEMS

Annelise Aufiero * **Rodolfo Soncini Sessa** *
Enrico Weber *

* *Dipartimento Elettronica e Informazione, Politecnico di Milano*

Abstract: In this paper we dealt with a new formulation of the classic total expected value - discounted cost (TEV-DC) control problem, over an infinite time horizon, where a set-valued control policy and a set of initial states must be found. The control problem is applied to a periodic discrete-time dynamic system, affected by stochastic disturbances. The solution is obtained by manipulating the solution of an analogous, but point-valued, TEV-DC control problem.

Keywords: Environmental systems, infinite time horizon, control law.

1. INTRODUCTION

In the automatic control theory the feed-back control law solving an optimal control problem has been always assumed to be point-valued, i.e. a function $m_t(\cdot)$ that, given a state value x_t at time t, produces a single value for the control u_t. The rationale of this assumption is that the controller is a device which must control the system without the human intervention. However automatic control theory has been applied also to the economic and environmental fields, where the controller rarely substitutes the human, more frequently, it suggests a "suitable" (optimal) control action, while the final decision is left to the decision maker (DM). Moreover, when solving an optimal control problem, it may happen that for a given state x_t the optimal control value u_t is not unique; i.e. there exists a set $M_t(x_t)$ of equivalent (optimal) controls. Then the DM would take advantage in knowing the set $M_t(x_t)$ of the "equivalent" controls, since (s)he may choose the more suitable one to the situation. In fact in the definition of the control problem it is generally impossible to completely specify the complexity of the real world and therefore only a limited set of all the possible goals is actually taken into account. Therefore the set $M_t(x_t)$ is the set of controls that at time t are equivalent from the point of view of these goals, but they might not be equivalent with regard to other goals not included in the original problem formulation. Thus the DM may usefully discriminate in $M_t(x_t)$.

On the basis of these considerations in a previous paper (Aufiero *et al.*, 2001*a*) we proposed to broaden the notion of control law by assuming that it is a set-valued function $M_t(x_t)$, of the state x_t, and that the DM may arbitrarily choose a control u_t out of it. In this paper we explore the optimal design problem of set-valued control policy. The system under control is a discrete-time, periodic system affected by an uncertain disturbance, and we assume that the disturbance has a cyclostationary stochastic description. More precisely the problem is to find the set-valued policy that minimizes the total, expected value, discounted cost (TEV-DC) occurring over an infinite-time horizon, i.e. attention is paid to the average performance of the system in the long run. Two reasons justify the adoption of the total discounted cost as performance index: first the introduction of the discount factor is often mandatory when the cost has a monetary interpretation, secondly the TEV-DC problem is by far the simplest and most well-behaved infinite horizon problem. This is due to the contraction property induced by the presence of the discount factor.

The aforementioned control problem is hereafter denoted as set-valued TEV-DC control problem. Its point-valued version has been analyzed by many authors: most notably Bellman (1957), Howard (1960), Blackwell (1965), Denardo (1967), and Bertsekas (1976 and 1977). All of them adopted Stochastic Dynamic Programming (SDP) to solve the problem and it is on the base of the solution algorithm proposed by Bertsekas (1977) that we will determine the solution to the set-valued TEV-DC control problem.

The paper is organized as follows. The system, and its point-valued and set-valued control policies are described in Section 2. The point-valued TEV-DC control problem is formulated and solved in Section 3. In Section 4 an alternative representation of the system is proposed, that turns out to be particularly useful to formulate and solve the set-valued TEV-DC control problem in Section 5. The conclusions complete the paper.

Due to space limitation all proofs will be omitted. They can be found in Aufiero *et al.* (2001b).

2. THE SYSTEM AND THE CONTROL POLICY

Consider the following periodic, discrete-time, discrete-space, dynamic system described by the state equation:

$$x_{t+1} = f_t(x_t, u_t, \varepsilon_t) \qquad t = 0, 1, \ldots \tag{1}$$

where the state x_t, the control u_t and the disturbance ε_t are elements of discrete spaces S_{x_t}, S_{u_t} and S_{ε_t}, respectively. Owing to the periodicity of the system we have that both the function $f_t : S_{x_t} \times S_{u_t} \times S_{\varepsilon_t} \to S_{x_t}$ and the spaces S_{x_t}, S_{u_t} and S_{ε_t} are periodic of period T. We pose some assumptions:

- the spaces S_{x_t}, S_{u_t} and S_{ε_t} are finite sets and therefore the system is an automaton;
- the control vector u_t is constrained to take value in a periodic not empty subset $U_t(x_t)$ of S_{u_t};
- the disturbance ε_t is assumed to be a random variable at any time t, described by a periodic probability distribution $\phi_t(\cdot \mid x_t, u_t)$;

A *point-valued control policy* (pv-policy)

$$p = [m_t(\cdot), \quad t = 0, 1, \ldots] \tag{2}$$

for the system (1) is an infinite sequence of *point-valued control laws* $m_t(\cdot) : S_{x_t} \to S_{u_t}$ (pv-laws). It is *periodic* with period T.

A *set-valued control policy* (sv-policy)

$$P = [M_t(\cdot), \quad t = 0, 1, \ldots] \tag{3}$$

for the system (1) is a infinite sequence of *set-valued control laws* $M_t(\cdot) : S_{x_t} \to$ (power set of S_{u_t}) (sv-laws). It is *periodic* with period T.

Remark 1. A set-valued control law may have two interpretations: it is defined as a set-valued function, but it may also be seen as a set of point-valued control laws: $M_t(\cdot) = \{m_t(\cdot)\}$. Analogously a sv-policy is defined as a sequence of set-valued functions, but it may also be thought as a set of pv-policies, i.e. $P = \{p\}$.

3. THE POINT-VALUED TEV-DC CONTROL PROBLEM

A step-cost function $g_t(x_t, u_t, \varepsilon_t)$, is associated with the transition from the state x_t to x_{t+1}. We assume this function is time variant, periodic of period T and bounded. Given an initial state x_0 and a pv-policy p, the cost functional that derives from the application of Laplace's criterion to the total discounted cost over an infinite time horizon is

$$J(x_0, p) = \lim_{N \to \infty} \sum_{t=0,\ldots,N} \underset{\varepsilon_0,\ldots,\varepsilon_t}{E} \left[\alpha^t \cdot g_t(x_t, u_t, \varepsilon_t) \right] \tag{4}$$

The discount factor α satisfies $0 < \alpha < 1$. Note that, in the expected operator E in (4), the sequence of disturbances is considered only up to t, since the step-cost $g_t(\cdot)$ does not depend on the realization of the disturbance from $t+1$ on. The cost functional (4) has a very direct interpretation: if the initial state is x_0 and the control policy p is applied, $J(x_0, p)$ represents the expected discounted total cost over an infinite time horizon. The introduction of a discount factor is often due, particularly when the step-cost has a monetary interpretation. From the mathematical point of view the presence of the discount factor guarantees the finiteness of the cost value.

3.1 *Formulation of the point-valued TEV-DC control problem*

Given an initial state $x_0 \in S_{x_0}$, determine an optimal control policy p^* such that

$$J^*(x_0) = J(x_0, p^*) = \min_p J(x_0, p) \tag{5a}$$

$$x_{t+1} = f_t(x_t, u_t, \varepsilon_t) \tag{5b}$$

$$u_t = m_t(x_t) \in U_t(x_t) \tag{5c}$$

$$\varepsilon_t \sim \phi_t(\varepsilon_t | x_t, u_t) \tag{5d}$$

$$p = [m_t(\cdot), \quad t = 0, 1, \ldots] \tag{5e}$$

Any control policy p^* is said to be optimal when $J(x_0, p^*) = J^*(x_0)$.

3.2 *Solution of the point-valued TEV-DC control problem*

Prob. (5) belongs to the class of problems considered by Bertsekas (1977), from which we may deduce the following proposition.

Proposition 1. The optimal value function $J^*(\cdot)$ satisfies the following optimality conditions:

$$J^*(x) = h_0^*(x) \qquad \forall x \in S_{x_0} \tag{6}$$

$$h_t^*(x) = \min_{u_t \in U_t(x)} \mathop{E}_{\varepsilon_t \sim \phi_t} \big[g_t(x, u_t, \varepsilon_t) + \\ + \alpha \cdot h^*_{(t+1) \mathrm{mod} T}(f_t(x_t, u_t, \varepsilon_t)) \big] \tag{7}$$

where $h_t^* : S_{x_t} \to R^+ \quad t = 0, \dots, T-1$. In the following, these functions will be called Bellman's functions. Moreover, if there exist pv-policies, then at least one of them is T-periodic.

Remark 2. Bertsekas (1977) proposed an algorithm to solve (7) in the stationary case. It is possible to derive from it a new algorithm to solve (7) in the period case, by applying the usual procedure to transform a periodic system into an equivalent stationary one. However in Aufiero *et al.* (2001a), in a minimax framework, we have shown that such algorithm is computationally redundant and that it can be usefully substituted by a slightly modified algorithm.

4. ALTERNATIVE REPRESENTATION OF THE SYSTEM

In the previous section we described the system by means of its state x_t, the dynamics of which is governed by the one step transition function (1). This latter, given the state, control and disturbance values at time t, uniquely specifies the state x_{t+1}. However when the disturbance value is unknown and its stochastic description is given, a probability distribution π_{t+1} is associated to the state x_{t+1}. Then it is convenient to describe the system by means of the state probability π_t, the dynamic of which is governed by a dynamic law that can be derived from the one-step transition function (1). In this section we present this law and some of its properties that will appear to be useful in the formulation of the TEV-DC sv-problem.

4.1 *State probability*

Since the state set S_{x_t} is finite, π_t is a vector, with as many elements as the ones of S_{x_t}. The value $[\pi_t]_x$ of the x-component of π_t is the probability of occurrence of the state x, at time t. To describe the time evolution of π_t, it is convenient to introduce a notation that is better suited to finite state system. For $t = 0, \dots, T-1$, let us ordinate the (finite) set S_{x_t}, so that a cardinal number i is associated to each state x_t. Thus a one to one correspondence is established between the element of S_{x_t} and the set $S_t = \{1, 2, \dots, C(S_{x_t})\}$ of natural numbers (where $C(S_{x_t})$ denotes the cardinality of the set S_{x_t}). We will denote this one to one correspondence by $x_t \leftrightarrow i$. To each state $i \in S_t$ and each control $u_t \in U_t(x_t)$, with $x_t \leftrightarrow i$, there corresponds a transition probability $\delta_t^{ij}(u_t)$, for any $j \in \{1, 2, \dots, C(S_{x_{t+1}})\}$, where $\delta_t^{ij}(u_t)$ denotes the probability that the next state will be j, given the present state is i and the control u_t is applied. These transition probabilities are univocally determined by the system equation and the probability distribution of the disturbances. Given the probability distribution π_t and a pv control law $m_t(\cdot) : S_t \to S_{u_t}$ (for simplicity we adopt the same symbol for $m_t(i)$ and $m_t(x_t)$), the probability distribution π_{t+1} at time $t+1$ is computable by

$$\pi_{t+1} = f_t^\pi(\pi_t, m_t(\cdot)) \tag{8a}$$

The function $f_t^\pi(\cdot, \cdot)$ has the form (Markov chain)

$$f_t^\pi(\pi_t, m_t(\cdot)) = \Delta_t(m_t(\cdot))' \pi_t \tag{8b}$$

where $'$ denotes transposition and Δ_t is a $C(S_{x_t}) \times C(S_{x_{t+1}})$ stochastic matrix (i.e. all its elements are non negative and the sum of the elements of each one of its rows equals unity, moreover its elements are the transition probabilities $\delta_t^{ij}(m_t(i))$). Eq.(8) specifies completely the system and the disturbances. Moreover Eq.(8) can be seen as the transition function of a deterministic system with state π_t and control $m_t(\cdot)$ (that is fixed open loop). In particular, when, at time t, the state $x_t \leftrightarrow i$ is known (and therefore all the components of π_t are null, but the i-th component that equals unity), Eq.(8) tell us that π_{t+1} is given by

$$\pi_{t+1} = \Delta_t^i(m_t(i))'$$

where $\Delta_t^i(m_t(i))$ is the i-row of Δ_t. Note that the state π_t takes values in an uncountable set, and therefore system (8) is not an automaton.

4.2 *Alternative representation of the cost functional*

We reformulate now the cost functional (4) by taking advantage of the notion of state probability. Consider the sequence of disturbances $[\varepsilon_0, \dots, \varepsilon_t]$ that is the argument of the *expected operator* E in (4). This sequence can be considered as the concatenation of two subsequences: $[\varepsilon_0, \dots, \varepsilon_{t-1}]$ and $[\varepsilon_t]$. Given the initial state x_0 and the feasible control policy $p_{[0,t)}$, at time t the first subsequence maps into a state x_t, to which is associated a probability $[\pi_t]_{x_t}$ (computable by recursive use of (8)). Therefore, in (4) the expected value with respect to $[\varepsilon_0, \dots, \varepsilon_{t-1}]$ can be substituted by the expected value with respect to x_t. Then (4) is equivalent to

$$J(x_0, p) = \lim_{N \to \infty} \sum_{t=0,\dots,N} \mathop{E}_{\substack{x_t \sim \pi_t \\ \varepsilon_t \sim \phi_t}} \left[\alpha^t \cdot g_t(x_t, u_t, \varepsilon_t) \right] \tag{9}$$

In (9) the expression in brackets is the discounted cost that is expected to occur in the t-th transition (given the probability π_t that results from the previous decisions and disturbances). To set in evidence this fact it is convenient to rewrite (9) in the following form:

$$J(x_0, p) = \lim_{N\to\infty} \sum_{t=0,\dots,N} j_t(\pi_t, m_t(\cdot)) \qquad (10a)$$

where

$$j_t(\pi_t, m_t(\cdot)) = \underset{\substack{x_t\sim\pi_t \\ \varepsilon_t\sim\phi_t}}{E} \left[\alpha^t \cdot g_t(x_t, u_t, \varepsilon_t)\right] \qquad (10b)$$

subject to

$$\pi_0 = \pi_{in}(x_0) \qquad (10c)$$

$$\pi_{t+1} = f_t^{\pi}(\pi_t, m_t(\cdot)) \qquad (10d)$$

$$u_t = m_t(x_t) \in U_t(x_t) \qquad (10e)$$

$$\varepsilon_t \sim \phi_t(\varepsilon_t \mid x_t, u_t) \qquad (10f)$$

$$p = [m_t(\cdot) \quad t = 0, 1, \dots] \qquad (10g)$$

where π_{in} is defined by:

$$\pi_{in}(x_0) : [\pi]_{x_0} = 1,\ [\pi]_x = 0\ \forall x \in S_{x_0} \text{ and } x \neq x_0$$

Remark 3. Notice that for every t the value of the cost functional (10a) is unaffected by the values that the control law $m_t(x_t)$ assumes in states x_t such that $[\pi_t]_{x_t} = 0$.

5. THE SET-VALUED TEV-DC CONTROL PROBLEM

The cost functional of the pv-problem has two arguments: the initial state and the pv-policy. In the sv-problem these arguments are enlarged to a set of initial states and a sv-policy.

5.1 *Set-valued control policies*

Consider a sv-policy

$$P = [M_t(\cdot), \quad t = 0, 1, \dots]$$

At any time t, once the state x_t is known, the DM chooses his(er) decision u_t by arbitrarily selecting an element u_t out of the set $M_t(x_t)$, i.e. $u_t \in M_t(x_t)$. We do not know how the DM will make his(er) choice, thus our system is affected by one more uncertainty: the DM's choice. This uncertainty is naturally described by the set-membership description ($u_t \in M_t(x_t)$).

The initial state set O We assume that the initial state x_0 is no more given, we only know that it belongs to a set O, i.e. x_0 is assumed to have a set-membership description:

$$x_0 \in O \subseteq S_{x_0}$$

5.2 *Again on state probability*

When the control law is set-valued, the probability distribution π_{t+1} is not anymore univocally determined given the probability distribution π_t, since it is unknown which control the DM will select out of $M_t(x_t)$. This implies that $m_t(x_t)$ in (8) (for $x_t \leftrightarrow i$) has to be treated as a disturbance of which a set-membership description is given. Then there exists an analogy between system (8), controlled by a sv-policy, and a system of type (1) affected by a set-membership disturbance (Aufiero *et al.*, 2001*a*). Due to the presence of the sv control law, the control $u_t = m_t(x_t)$ plays in the Markov chain (8) the same role played by a set-membership disturbance ε_t in system (1). Therefore we know that, given π_t, π_{t+1} will belong to a set Π_{t+1} *(reachable(probability) set)*

$$\Pi_{t+1} = \{\pi_{t+1} : \exists m_t(\cdot) \in M_t(\cdot) : \pi_{t+1} = f_t^{\pi}(\pi_t, m_t(\cdot)) \text{ with } \pi_t \text{ given}\} \qquad (11)$$

More in general, if, at time t, we do not known the probability distribution π_t, but only a set Π_t to which it belongs, the set Π_{t+1} is given by

$$\Pi_{t+1} = F_t^{\pi}(\Pi_t, M_t(\cdot)) = \{\pi_{t+1} : \exists \pi_t \in \Pi_t, m_t(\cdot) \in M_t(\cdot) : \pi_{t+1} = f_t^{\pi}(\pi_t, m_t(\cdot))\} \qquad (12)$$

Notice that (12) describes a system with state Π_t and control $M_t(\cdot)$. This system is deterministic (in $F_t^{\pi}(\cdot,\cdot)$ does not appear any disturbances), since the uncertainty disappear when we move the attention from the probability distribution π_t to the reachable (probability) set Π_t. Observe that system (12) is not an automaton. Since system (12) is deterministic we may associate it a multi-step function

$$\Pi_t = \Pi(t, 0, \Pi_0, P_{[0,t)}) \qquad t = 1, 2, \dots \qquad (13a)$$

where, the set-valued function $\Pi(\cdot,\cdot,\cdot,\cdot)$ is computed by the following recursive procedure

$$\Pi_0 \text{ given} \qquad (13b)$$

$$\Pi_{\tau+1} = F_\tau^{\pi}(\Pi_\tau, M_\tau(\cdot)) \quad \tau = 0, \dots, t-1 \qquad (13c)$$

The definition of the transition function $\Pi(\cdot,\cdot,\cdot,\cdot)$ is extended at time $t = 0$ by setting

$$\Pi(0, 0, \Pi_0, P_{[0,0)}) = \Pi_0 \qquad (13d)$$

As said we assume that the initial state x_0 has a set-membership description, i.e. $x_0 \in O$ where the set O is a design variable. Consistently the set Π_0 of initial probability π_0 is a function of O, given by

$$\Pi_0 = \Pi_{in}(O) = \{\pi_0 : \pi_0 = \pi_{in}(x_0)\ \forall x_0 \in O\}$$

When the control policy is set-valued, given the state $\bar{x}_{t-1}$, a state x_t will be reachable at time t if its probability $[\pi_t]_{x_t}$ is not null. Hence the reachable state set at time t (conditioned to $\bar{x}_{t-1}$) is given by

$$\{x_t \in S_{x_t} : \exists \pi_t \in \Pi(t, t-1, \pi_{in}(\bar{x}_{t-1}), P_{[t-1,t)}) : [\pi_t]_{x_t} > 0\} \quad (14)$$

More in general, if the system starts at time $t = 0$ from the set O and it is controlled by P, the state x_t will belong to the set

$$X_t = X(t, 0, O, P_{[0,t)}) = \{x_t \in S_{x_t} : \exists \pi_t \in \Pi(t, 0, \Pi_{in}(O), P_{[0,t)}) : [\pi_t]_{x_t} > 0\} \quad (15)$$

5.3 *The cost functional*

When the cost functional depends on a set of initial states and on a sv-policy, its structure has to be modified with respect to the form (10) to take into account the non uniqueness of the initial state and the set-valued nature of the control policy. These differences are fixed up as follows:

(1) Since the initial state is not given, but it is known that it belongs to the set O, in the definition of the cost functional it is sensible to consider the maximum with respect to x_0 in O.
(2) Since the proposed control is not unique and we have no knowledge on the criterion adopted by the DM to select a value u_t out of $M_t(x_t)$, it is sensible to consider the worst case, i.e. the maximum expected cost with respect to u_t in $M_t(x_t)$.
(3) As we already noticed, when the control law is set-valued, the state probability, at time t, is not uniquely determined. We only know that it will belong to a set Π_t. Therefore in (10b) in the expectation with respect to π_t, it is sensible to consider the worst case, i.e. the maximum expected cost with respect to π_t in Π_t.

Given these positions the cost functional of the set-valued TEV-DC control problem is defined by

$$L(O, P) = \lim_{N\to\infty} \sum_{t=0,\dots,N} l_t(\Pi_t, M_t(\cdot)) \quad (16a)$$

where

$$l_t(\Pi_t, M_t(\cdot)) = \max_{\pi_t\in\Pi_t} \mathop{E}_{x_t\sim\pi_t} \left[\max_{u_t\in M_t(x_t)} \mathop{E}_{\varepsilon_t\sim\phi_t} \left[\alpha^t \cdot g_t(x_t, u_t, \varepsilon_t)\right]\right] \quad (16b)$$

subject to

$$\Pi_0 = \{\pi_0 : \exists x_0 \in O : \pi_0 = \pi_{in}(x_0)\} \quad (16c)$$

$$\Pi_{t+1} = F_t^{\pi}(\Pi_t, M_t(\cdot)) \quad (16d)$$

$$u_t \in M_t(x_t) \subseteq U_t(x_t) \quad (16e)$$

$$\varepsilon_t \sim \phi_t(\varepsilon_t \mid x_t, u_t) \quad (16f)$$

$$P = [M_t(\cdot) \quad t = 0, 1, \dots] \quad (16g)$$

In plain words, $l_t(\Pi_t, M_t(\cdot))$ turns out to be the maximum, over the set Π_t, of the expected value with respect to x_t of the maximum cost over the sets $M_t(x_t)$ of the expected discounted cost with respect to ε_t. Observe that when $t = 0$, thanks to (16c), the first max operator of (16b) performs the maximum with respect to x_0 in O, as required by point 1; while, when $t \geq 1$, the same maximum produce the effects required by point 3. Finally, for all t, point 2 is operationally performed by the second max operator in (16b).

Remark 4. Notice that, for any t, the value of the cost function (16a) is unaffected by the values that the control law $M_t(x_t)$ assumes in states x_t such that $[\pi_t]_{x_t} = 0$ (observe the analogy with Rem. 3).

5.4 *Some definitions and theorems*

Definition 1. (*feasible couples*) A couple (O, P) is said to be *feasible* if it satisfies the following constraints

$$O \subseteq S_{x_0}$$
$$P : M_t(x_t) \subseteq U_t(x_t)$$

Theorem 1. (*property of feasible couples*) Given a feasible couple (O, P) the following relation holds

$$L(O, P) = \max_{x_0\in O} \max_{p\in P} J(x_0, p)$$

Definition 2. (*larger couples*) A feasible couple (O, P) is said to be *larger* than a feasible couple (O', P') and we write $(O, P) > (O', P')$ if

$$O \supseteq O' \quad (17a)$$
$$M_t(x) \supseteq M'_t(x)\ \forall x \in X(t, 0, O', P'_{[0,t)}) \quad (17b)$$

and either (17a) is satisfied with the strict inequality sign or there exist at least one t and one x such that the condition (17b) is satisfied with the strict inequality sign.

Notice that Def. 2 requires that the condition (17b) is verified *only* for those states that are reachable from O', with the policy P'. The set of those states is a subset of the states that are reachable from O with the policy P, as the following theorem proves.

Theorem 2. (*property of larger couples*) If the couple (O, P) is *larger* than (O', P') then

$$X(t, 0, O, P_{[0,t)}) \supseteq X(t, 0, O', P'_{[0,t)})$$

and

$$L(O, P) \geq L(O', P')$$

Definition 3. (*optimal couples*) A couple (O^*, P^*) is said to be optimal if

$$L(O^*, P^*) = \min_{(O,P)\,\text{feasible}} L(O, P)$$

where $L(O, P)$ is defined by (16).

Finally we can state the following

Theorem 3. (*property of the optimal couples*) Given an optimal couple (O^*, P^*), between the optimal value $L(O^*, P^*)$ and the value $J^*(\cdot)$ of the pv-problem the following relation holds:

$$L(O^*, P^*) = \min_{x_0 \in S_{x_0}} J^*(x_0)$$

Remark 5. From the Th. 3 it follows, that, if the optimal couple is not unique, all the optimal couples share the same value.

5.5 *Formulation of the set-valued TEV-DC control problem*

Determine the largest optimal couple $(\bar{O}, \bar{P})$ such that

$$L(\bar{O}, \bar{P}) = \min_{(O,P)\,\text{feasible}} L(O, P) \qquad (18)$$

subject to (16).

5.6 *Solution of the set-valued control problem*

Prob. (18) is solvable by the procedure stated in the following

Proposition 2. Let $h_t^*(\cdot)$, $t = 0, \dots, T-1$ be the Bellman's functions solution of (7). Pose

$$l^* = \min_{x_0 \in S_{x_0}} h_0^*(x_0) \qquad (19)$$

Consider the following sets

$$\tilde{O} = \{x_0 \in S_{x_0} : h_0^*(x_0) = l^*\} \qquad (20a)$$

$$\tilde{M}_t(x_t) = \{u_t \in U_t(x_t) : \mathop{E}_{\varepsilon_t \sim \phi_t} [g_t(x_t, u_t, \varepsilon_t) + $$
$$+ \alpha \cdot h^*_{(t+1) \bmod T}(f_t(x_t, u_t, \varepsilon_t))\Big] \le l^*\} \qquad (20b)$$

Then the couple $(\tilde{O}, \tilde{P})$ with

$$\tilde{P} = \Big[\tilde{M}_0(\cdot), \dots, \tilde{M}_{T-1}(\cdot), \tilde{M}_0(\cdot), \dots, \Big] \qquad (20c)$$

is the largest optimal couple, i.e. the solution of sv-problem.

6. CONCLUSION

This paper deals with the optimal design of sv-policies. This type of policies is of definitive importance in various fields, such as water and natural resources management, environmental quality control and storage system management; i.e. in all the fields where the notion of Decision Support System is adopted. The design problem of such a type of control policy has been formulated and solved as an optimal periodic control problem in the case of a discrete-time, periodic system affected by a stochastic disturbance. The problem is to find the "largest", periodic, sv-policy that minimizes the total expected value discounted cost over an infinite-time horizon. This problem very often emerges in the previously mentioned fields.

The problem is solved by an algorithm that in turn requires to solve an analogous problem, obtained from the previous one by substituting the sv-policy by a classical point-valued one. The analysis of the algorithm offers a deep insight into the relationship between sv- and pv-policies.

The system controlled by the optimal sv-policy turns out to be a Markov chain. Current research is working to prove that this chain is characterized by a unique ergodic class and to determine a procedure to compute it from the Bellman functions $h_t^*(\cdot)$.

7. REFERENCES

Aufiero, A., R. Soncini-Sessa and E. Weber (2001*a*). Set-valued control laws in minmax control problem. Technical Report 2001.51. Dip. Elettronica e Informazione, Politecnico di Milano. Milano, Italy.

Aufiero, A., R. Soncini-Sessa and E. Weber (2001*b*). Set-valued control laws in TEV-DC control problem. Technical Report 2001.60. Dip. Elettronica e Informazione, Politecnico di Milano. Milano, Italy.

Bellman, R. (1957). *Dynamic Programming*. Princeton University Press. Princeton, NJ.

Bertsekas, D.P. (1976). *Dynamic Programming and Stochastic Control*. Academic Press. New York, NY.

Bertsekas, D.P. (1977). Monotone mappings with application in dynamic programming. *SIAM Journal on Control and Optimization* **15**, 438–464.

Blackwell, D. (1965). Discounted Dynamic Programming. *The Annals of Mathematical Statistics* **36**(1), 226–23 5.

Denardo, E. V. (1967). Contraction mappings in the theory underlying dynamic programming. *SIAM Review* **9**(2), 165–177.

Howard, R. (1960). *Dynamic Programming and Markov Processes*. MIT Press. Cambrige.