Research Inventy: International Journal of Engineering And Science
Vol.8, Issue 3 (July 2018), PP -33-38
Issn (e): 2278-4721, Issn (p):2319-6483, www.researchinventy.com

Distruption of Semi-Markov Control Problems

Abdesselam Elbahja[1],SoulhiAziz[2]
[1]*Professor at National Higher School of Mines, Rabat, Morocco*
IT Department B.P: 753, Morocco
[2]*Professor at National Higher School of Mines, Rabat, Morocco*
[2]*ORCID: 0000-0003-1904-513X and the Scopus Author id: 42162322200*
IT Department B.P: 753, Morocco
Corresponding Author: Abdesselam Elbahja

ABSTRACT: *In this paper, we consider an extension of perturbations to semi-Markov control problems. We then study process perturbations in which the instants between transitions are random variables. The discounted semi-Markov control problem and the limiting average semi-Markov control problem are considered. We proceed to a perturbation on the law transition probabilities and the discounted factor. For this, we consider the particular case where the transition time of the original semi-Markov process is a random variable that follows an exponential law.*

KEYWORDS: *Semi-Markov decision problem, discounted expected criterion, limiting average criterion, perturbation, discounted factor.*

Date Of Submission: 03-02-2019 Date Of Acceptance: 19-02-2019

I. INTRODUCTION

In the case of Markov decision problems, the times between two consecutive decision time points are equidistant. In this paper, we consider perturbations of processes, such as the times between transitions are random variables. These processes, called semi-Markov control (or decision) processes, were introduced by De Cani [18], Howard [22], Jewell [26] and Schweitzer [27].

II. DEFINITIONS AND PRELIMINARIES

A Semi-Markov Control Process (SMCP, for short) is observed at decision time points t= 0, 1, …; starting at t=0. At each decision time point, the system is in once of a finite number of states and an action has to be chosen.

Let S :={1, 2, …, N} be the state space, and for each s ∈ S let A(s) be the finite set of possible actions in state s.

If the system is in state s ∈ S and an action a ∈A(s) is chosen, then the following occurs independently of the history of the process:

i) The next state s' of the process is chosen according to the transition probability p (s'/s, a).
ii) Conditional on the event that the next state is s', the time until the transition from s to s' occurs is a random variable with probability distribution F (. /s, a, s').
iii) If the next decision time point falls after τ units of times, then the reward in this epoch is denoted by r (τ, s, a).

The transition law p satisfies:

p (s'/s, a) ≥ 0, s, s' ∈ S, a ∈ A(s) and $\sum_{s'\in S} p(s'/s, a) = 1$, s ∈ S, a ∈ A(s).

A decision rule π^t at time t is a function which assigns a probability to the event that any particular action is taken at time t. In general π^t may depend on all realized states up to time t, and on all realized actions up to time t-1.

Let $h_t = (s_0, a_0, s_1, a_1, \ldots, a_{t-1}, s_t)$ be the history up to time t where $a_0 \in A(s_0), \ldots, a_{t-1} \in A(s_{t-1})$, then $\pi^t(h_t, .)$ is the probability distribution on $A(s_t)$, that is, $\pi^t(h_t, a_t)$ is the probability of selecting the action a_t time t, given the history h_t.

A strategy π is a sequence of decision rules $\pi = (\pi^0, \pi^1, \ldots, \pi^t, \ldots)$.

A semi-Markov strategy is one in which π^t depends only on the current state at time t.

A stationary strategy is semi-Markov strategy with identical decision rules.

A deterministic strategyπ is a stationary strategy whose single decision rule is nonrandomized: (For any s∈ S, there exist $a_s \in A(s)$ such that, $\pi(s, a_s) = 1$).

Let C, C_S and C_D denote the sets of all strategies, all stationary strategies and all deterministic strategies respectively.

If C_M is the let of all semi-Markov strategy, then:
$C_D \subset C_S \subset C_M \subset C$.

III. DISCOUNTED EXPECTED CRITERION

In order to insure that an infinite number of transitions does not occur in a finite interval, we shall assume throughout that the following condition holds:
For all s, s' ∈ S, and a ∈ A(s),
$\int_0^{+\infty} e^{-\alpha t} dF(t/s, a, s') < 1$; (where α is a fixed positive real number).(3.0)
Note that $\int_0^{+\infty} e^{-\alpha t} f(t/s, a, s')\, dt$ represents the expected discount rate.
For any strategy $\pi \in C$ and any initial state s ∈ S, we define the expected discounted reward V(s, π) by:
$V(s, \pi) := E_\pi[\sum_{n=0}^{\infty} (e^{-\alpha})^{\tau_0+\tau_1+\cdots+\tau_{n-1}} r(\tau_n, X_n, Y_n) / X_0 = s]$; where $\tau_0 + \tau_1 + \cdots + \tau_{n-1} := 0$ for n=0, and τ_n is the time between the n-th and the (n+1)-th transition.
X_n is the observed state at time point $\tau_0 + \tau_1 + \cdots + \tau_{n-1}$
Y_n is the chosen action at time point $\tau_0 + \tau_1 + \cdots + \tau_{n-1}$
For any s ∈ S and π ∈ C, we pose:
$V_\alpha(s, \ ) = (1 - e^{-\alpha})\, V(s, \pi)$
The discounted semi-Markov control problem is defined by the following optimization problem:
$V_\alpha(s) := \max_{\pi \in C} V_\alpha(s, \pi), s \in S$.
A strategy π^0 is called optimal (or α-optimal) if for all s ∈ S,
$V_\alpha(s) = V_\alpha(s, \pi^0)$.

Remark 3.1
It is well known that there exists an optimal deterministic strategy and there are a number of finite algorithms for its computation (e.g., Kallenberg [15], Ross [24]).
For every s, s'∈ S and a∈ A (s), we denote by f(t/ s, a, s') the probability density of the distribution F (t/ s, a, s').
We define: for all s, s' ∈ S and a∈ A (s)

$$\bar{r}(s, a) := \sum_{s' \in S} p(s'/ s, a) \int_0^{+\infty} r(t, s, a) f(t /s, a, s')\, dt \quad (3.1)$$

$$\bar{p}(s'/ s, a) := p(s'/ s, a) \int_0^{+\infty} e^{-\alpha t} f(t /s, a, s')\, dt. \quad (3.2)$$

The following two results can be derived (using analogous proofs) from Kallenberg [15] or Ross [24].

Lemma 3.1
For any deterministic strategy $\pi \in C_D$ and s ∈ S,
$V_\alpha(s, \pi) = (1 - e^{-\alpha}) \bar{r}(s, \pi(s)) + \sum_{s' \in S} \bar{p}(s'/ s, \pi(s))\, V_\alpha(s, \pi)$.

Lemma 3.2
For any $s \in S, V_\alpha(s) = \max_{a \in A(s)}\{(1 - e^{-\alpha})\, \bar{r}(s, a) + \sum_{s' \in S} \bar{p}(s'/ s, a))\, V_\alpha(s')\}$.
Let $\bar{\Gamma}$ be the Markov control process defined by:
$\bar{\Gamma} := < S, \{A(s), s \in S\}, \bar{p}, (1 - e^{-\alpha})\bar{r} >$.
We define:
$\lambda := \max\{\int_0^{+\infty} e^{-\alpha t} f(t /s, a, s')\, dt, s, s' \in S, a \in A(s)\}$.

Remark 3.2
Note that from the condition (3.0), it follows that $\lambda < 1$.
By using the definition (3.2) of $\bar{p}$, we have that for any s ∈ S and $a \in A(s)$,
$1 - \sum_{s' \in S} \bar{p}(s'/ s, a) = 1 - \sum_{s' \in S} p(s'/ s, a) \int_0^{+\infty} e^{-\alpha t}.f(t /s, a, s')\, dt$
$> 1 - \lambda$
$> 0.$
Then $\sum_{s' \in S} \bar{p}(s'/ s, a) < 1$, for all $s \in S$ and $a \in A(s)$.
From lemma 3.2, we derive that for any $s \in S$, $V_\alpha(s)$ can be interpreted as the optimal value in state s for the MCP $\bar{\Gamma}$.

IV. LIMITING AVERAGE CRITERION

For any strategy $\pi \in C$ and initial state s∈ S, the limiting average reward J (s, π) is defined by:

J (s, π): = $\lim_{T\to\infty} inf [\frac{1}{T} J_T(s,\pi)$;

where $J_T(s,\pi)$ denotes the expected reward in the interval [0, T] when the strategy π is used and the initial state is s. That is:

J (s, π): = $E_\pi [\sum_{n=0}^{n(T)} r(\tau_n, X_n, Y_n) / X_0 =$ s],

where, n(T): = max { n / $\tau_0 + \tau_1 + \cdots + \tau_{n-1} + \tau_n$ < T }.

For any s ∈ S and a ∈ A (s); the holding time and the immediate reward are defined respectively by:

$\tau(s,a) := \sum_{s' \in S} p(s'/$ s, a) $\int_0^{+\infty} t\, f(t\ /$s, a, s') dt, and c' (s, a): = r ($\tau(s,a)$, s, a).

Throughout this Section, it is assumed that: for all $s \in S$ and $a \in A(s)$

0<$\tau(s,a)$<∞.

The limiting average semi-Markov control problem is defined by the following optimization problem:

J (s) : = $\max_{\pi \in C} J(s,\pi), s \in S.$

A strategy π^0 is called optimal if:

J (s, π^0) = J (s) for alls ∈ S.

It is well known that there exists an optimal deterministic strategy and there are a number of finite algorithms for its computation (e.g., see Kallenberg [15], Ross [24]).

We note that any limiting average semi-Markov control problem can be described by:

S; {A (s), s ∈ S }, p, τ, c'.

5- Perturbations (DiscountedCase)

We consider a S.M.C.P Γ: = < S, {A(s), s∈ S }, p, τ, r >.

The case of the disturbance of the transition probabilities and the probability density associated with the transition time has been studied (see [3]).

We consider the situation where the law of transition probabilities p and the discounted factor are disturbed, and we pose:

$p_\varepsilon(s'/$ s, a):= p(s'/ s, a) + ε d(s'/ s, a); s, s' ∈ S, a∈ A(s);

$e^{-\alpha} = \frac{1}{(1+\mu\varepsilon)}$; where,$\mu > 0$,$\varepsilon \in]0,\varepsilon_0[$, $\varepsilon_0$is a fixed positive real number. (4.0)

It is assumed, as in the case of M.C.P, that the perturbed probabilities:

[$p_\varepsilon(s'/$ s, a) / s, s' ∈ S, a∈ A(s)] are transition probabilities.

IfΓ is the original semi-Markov decision process, then we denote byΓ_ε the disturbed S.M.D.P.

According to Lemma 3.2, the valueW_ε (optimum ofΓ_ε) must verify the optimality equation:

For all s∈ S,

W_ε(s)=$\max_{a\in A(s)}\{\frac{\mu\varepsilon}{1+\mu\varepsilon}\bar{r}_\varepsilon(s,a)+\sum_{s'\in S}(p+\varepsilon d)$(s'/s,a)$\int_0^{+\infty}\frac{1}{(1+\mu\varepsilon)^t} f(t/$s,a,s')dt W_ε(s')},

Where$\bar{r}_\varepsilon(s,a) = \sum_{s'\in S}(p+\varepsilon d)$(s'/s,a)$\int_0^{+\infty} r(t$,s,a) f (t/ s, a, s') dt; for anys ∈ S, a∈ A(s).

We can write: for any s ∈ S, a∈ A(s);

$\bar{r}_\varepsilon(s,a) = \sum_{s'\in S} p$(s'/s, a)$\int_0^{+\infty} r(t$, s, a) f (t / s, a, s') dt

+ $\varepsilon\sum_{s'\in S} d$ (s'/s, a)$\int_0^{+\infty} r(t$, s, a) f real (t / s, a, s') dt.

If $\bar{r}_d$ (s, a) = $\sum_{s'\in S} d$ (s'/s, a)$\int_0^{+\infty} r(t$, s, a) f (t / s, a, s') dt, (s ∈ S, a∈ A(s)),

The optimality equation becomes:

For all s ∈ S,

W_ε(s)=$\max_{a\in A(s)}\{\frac{\mu\varepsilon}{1+\mu\varepsilon}[\bar{r}_\varepsilon(s,a)+\varepsilon\bar{r}_d$ (s,a)]

$+ \quad \sum_{s' \in S}(p + \varepsilon d)$(s'/s, a)$\int_0^{+\infty} \frac{1}{(1+\mu\varepsilon)^t} f(t/$s, a, s') dt W_ε(s')}. (4.1)

Remark 4.0

If$\int_0^{+\infty} \frac{1}{(1+\mu\varepsilon)^t} f(t/$s, a, s') dt = $\frac{1}{1+\mu\varepsilon}$, for all s, s'∈ S, a∈ A(s), then 4.1will become:

For any s ∈ S, $(1 + \mu\varepsilon)\ W_\varepsilon$(s) = $\max_{a\in A(s)}\{\mu\varepsilon\ [\bar{r}(s,a)$+ $\varepsilon\, \bar{r}_d$ (s, a)]

$+ \quad \sum_{s' \in S}(p + \varepsilon d)$(s'/s,a) W_ε(s')} (4.2)

Starting from (4.2), we obtain an analogous equationto [2.3] (see [1],(case of disturbance of Markov control problems) :

$-\varepsilon\mu V_\varepsilon^*(s)$ +$\max_{\pi\in Fs}$ {$[G_\varepsilon(\pi).V_\varepsilon^*]$(s)+

$\mu\varepsilon$ r(s,π)} $= 0$ for all s ∈S. [2.3]

Particular case

We will now consider a special case that leads to equation [2.3], seen in [1].

For this, suppose that the transition time of the original semi-Markov process is a random variable that follows an exponential law of parameter:$\frac{\alpha}{e^\alpha - 1}$.

So, we can write:

For all s, s' ∈ S, a∈ A(s) and t≥ 0,

f (t / s, a, s') = $\frac{\alpha}{e^\alpha - 1} e^{-(\frac{\alpha}{e^\alpha - 1})t}$. (4.3)

Furthermore: $\int_0^{+\infty} e^{-\alpha t} f(t/s,a,s')dt = e^{-\alpha}$.

The expression (4.3) induces a perturbation of the probability density f which depends on the factor $e^{-\alpha} := \frac{1}{(1+\mu\varepsilon)}$(see (4.0).

Considering equation (4.1), write the following whole series development:

$$(1 + \mu\varepsilon)^{-t} = \sum_{n\geq 0} \frac{(-1)^n}{n!} t(t-1)\ldots(t-n+1)\mu^n\varepsilon^n.$$

We have that:

$\sum_{s' \in S}(p + \varepsilon d)$(s'/s,a)$\int_0^{+\infty} \frac{1}{(1+\mu\varepsilon)^t} f(t/$s,a,s') dt W_ε(s') =

$\sum_{s' \in S}(p + \varepsilon d)$(s'/s,a)$\int_0^{+\infty}(\sum_{n\geq 0} \frac{(-1)^n}{n!} t\,(t-1)\ldots(t-n+1)\mu^n\varepsilon^n) f(t/$ s, a, s') dt W_ε(s').

Under the convergence hypotheses and according to the theory of generalized integrals, the sum and integral signs can be inverted in the previous equality; he comes then:

$\sum_{s' \in S}(p + \varepsilon d)$(s'/s,a)$\int_0^{+\infty} \frac{1}{(1+\mu\varepsilon)^t} f(t/$s,a,s') dt W_ε(s') =

$\sum_{s' \in S}(p + \varepsilon d)$(s'/s,a) W_ε(s')+ $\sum_{s' \in S}(p + \varepsilon d)$(s'/s,a)$(-\mu\varepsilon)\int_0^{+\infty} t f(t/$s, a, s') dt W_ε(s')

+($\frac{\mu^2\varepsilon^2}{2!}$)$\sum_{s' \in S}(p + \varepsilon d)$(s'/s,a)$\int_0^{+\infty} t\,(t-1) f(t/$s, a, s') dt W_ε(s')

+($\frac{\mu^3\varepsilon^3}{3!}$)$\sum_{s' \in S}(p + \varepsilon d)$(s'/s,a)$\int_0^{+\infty} t\,(t-1)(t-2) f(t/$s, a, s') dt W_ε(s') + …

+($\frac{\mu^n\varepsilon^n}{n!}$)$\sum_{s' \in S}(p + \varepsilon d)$(s'/s,a)$\int_0^{+\infty} t\,(t-1)(t-2)\ldots(t-n+1) f(t/$s, a, s') dt W_ε(s')+ $R_n\,(\mu\varepsilon)$; where$R_n\,(\mu\varepsilon)$is the rest of order n.

In this last equality we used the relationship:

$\int_0^{+\infty} f(t/$ s, a, s') dt = 1, (For all s, s' ∈ S and a ∈ A(s)).

We then obtain:

For all s, s' ∈ S and a∈ A(s),

$\sum_{s' \in S}(p + \varepsilon d)$(s'/s, a) $\int_0^{+\infty} \frac{1}{(1+\mu\varepsilon)^t} f(t/$s, a, s') dt W_ε(s') =

$\sum_{s' \in S}[\, \boldsymbol{p} + \boldsymbol{\varepsilon}\ (\boldsymbol{d}(s'/s, a) - \boldsymbol{\mu}\ p\ (s' / s, a) \int_0^{+\infty} \boldsymbol{t f(t}/s, a, s')\ dt) + \boldsymbol{\varepsilon}^2(-\boldsymbol{\mu}\ d\ (s' / s, a) \int_0^{+\infty} \boldsymbol{t f(t}/s, a, s')\ dt$

$+\frac{\mu^2}{2}\ p\ (s' / s, a) \int_0^{+\infty} \boldsymbol{t(\ t-1) f(t}/s, a, s')\ dt)$

$+\ \boldsymbol{\varepsilon}^3((\frac{\mu^2}{2}\ d\ (s' / s, a) \int_0^{+\infty} \boldsymbol{t\ (t-1) f(t}/s, a, s')\ dt$

$-(\frac{\mu^3}{3!})\ d\ (s' / s, a) \int_0^{+\infty} \boldsymbol{t\ (t-1)(t-2) f(t}/s, a, s')\ dt)$

$+ \ldots + \boldsymbol{\varepsilon}^n (\ \frac{(-\mu)^{n-1}}{(n-1)!}\ d\ (s' / s, a) \int_0^{+\infty} \boldsymbol{t\ (t-1) \ldots\ (t-n+1) f(t}/s, a, s')\ dt$

$+\frac{(-\mu)^n}{n!} p(s'/s,a) \int_0^{+\infty} \boldsymbol{t\ (t-1) \ldots\ (t-n+1) f(t}/s, a, s')\ dt + \boldsymbol{R_n}\ (\boldsymbol{\mu\varepsilon})]\ \boldsymbol{W_\varepsilon}(s')$.

Let $n \in \boldsymbol{N}^*$. For $i \in [2, n]$, we pose: (for all $s, s' \in S$, $a \in A(s)$);

$\boldsymbol{d_i}(\ s' / s, a\ ) = \frac{(-\mu)^{i-1}}{(i-1)!}\ d\ (s' / s, a) \int_0^{+\infty} \boldsymbol{t\ (t-1) \ldots\ (t-i+2) f(t}/s, a, s')\ dt$

$+\frac{(-\mu)^i}{(i)!} p(s'/s,a) \int_0^{+\infty} \boldsymbol{t\ (t-1) \ldots\ (t-i+1) f(t}/s, a, s')\ dt$, and:

$\boldsymbol{d_1}(s' / s, a) = d\ (s' / s, a) - \boldsymbol{\mu}\ p\ (s' / s, a) \int_0^{+\infty} \boldsymbol{t f(t}/s, a, s')\ dt$.

We can then write:

For any $s \in S$;

$\boldsymbol{W_\varepsilon}(s) = \mathbf{max}_{a \in A(s)} \{ \frac{\mu\varepsilon}{1+\mu\varepsilon}\ (\ \bar{r}(s,a) + \varepsilon \bar{r}_d\ (s, a)) + \sum_{s' \in S}[\boldsymbol{p(s'/s, a)} + \boldsymbol{\varepsilon d_1}(s'/s, a)$

$+\boldsymbol{\varepsilon}^2 \boldsymbol{d_2}(s' / s, a) + \boldsymbol{\varepsilon}^3 \boldsymbol{d_3}(s' / s, a) + \ldots + \varepsilon^n d_n(s'/ s, a) + \boldsymbol{R_n}\ (\boldsymbol{\mu\varepsilon})]\ \boldsymbol{W_\varepsilon}(s')\}$. (4.4)

We note that in (4.4), the term in brackets represents a perturbation of order greater than 2 in ε.
Moreover, the ergodic structure of the semi-Markov process can be modified by a strategy of C.
Thus, if the condition of decomposability is not verified, we will not be able to apply a method analogous to the algorithm of the improved strategy that we developed in the case of Markov Decision Problems (see [1]).
We can conclude that the perturbations of Semi-Markov processes present more difficulties than the case of Markov Control Problems.

REFERENCES

[1]. A. Elbahja and A. Soulhi (2018), "Strategy Improvement Algorithm for Singularly Perturbed Discounted Markov Decision Processes", ARPN Journal of Engineering and Applied Sciences, vol.13,NO. 16, AUGUST 2018.

[2]. M. Abbad, Jerzy A. Filar (1992), "Perturbation and Stability theory for Markov Control Problems", IEEE Tranc. Automat. Control, Vol.37, No. 9.

[3]. M. Abbad and Jerzy A Filar (1991), " Perturbation Theory for Semi Markov Control Problems", Proceedings of the 30th Conference on Decision and Control, December 1991 (England).

[4]. M. Tkiouat (1991) "Paramétrisation dans le problème de contrôle optimal de processus markoviens et semi-markoviens. Algorithmes et quelques domaines d'application". Thèse d'état ; Université de Bruxelles, Faculté des sciences.

[5]. M. Abbad, Jerzy A .Filar, and Tomasz R. Bielecki (1992), ''Algorithms for Singularly Perturbed Limiting Average Markov Control Problems", IEEE Tranc. Automat. Control, Vol.37, No. 9.

[6]. R. Aldhaheri and H. Khalil (1989),"Aggregation and optimal control of nearly completely decomposable Markov chains'', in Proceedings of 28th CDC, IEEE, pp.1277-1282.

[7]. Arthur F.Veinott. JR. (Stanford University) (1969), " Discrete Dynamic Programming with sensitive Discount Optimality Criteria ".

[8]. T.R. Bielecki and J.A. Filar (1991), " Singularly Perturbed Markov Control Problem: Limiting Average Cost ", Annals of Operations Research, Vol.28, pp.153-168.

[9]. D. Blackwell (1962), "Discrete Dynamic Programming, Annals of Mathematical Statistics'', 33, pp.719-726

[10]. P.J Courtois (1977), "Decomposability: Queuing and Computing Systems", academic Press, N.Y.

[11]. F. Delebecque (1983), "A Reduction Process for Perturbed Markov Chains", SIAM J. App. Math., vol AC - 48, pp.325 -350.

[12]. F. Delebecque and J. Quadrat (1981), ''Optimal control of Markov chains admitting strong and weak interactions ", Automatica, 17, pp.281-296.

[13]. E.V Denardo (1982), " Dynamic Programming'', Prentice- Hall Egle -Wood Cliffs, New Jersey.

[14]. C.Derman (1970), " Finite State Markovian Decision Process", Academic Press, N.Y.

[15]. L.C.M. Kallenberg (1983), ''Linear Programming and Finite Markovian Control Problem ", Mathematical Center Tracts 148, Amsterdam.

[16]. T. Kato (1980), "Perturbation Theory for Linear Operators", Spring-Verlag, Berlin.

[17]. Kemeny and J.L.Snell (1960), "Finite Markov Chains", Van Nostrand, New York.

[18]. J.S De Cani (1964) Ronald A. Howard, "A Dynamic programming Algorithm for Embedded Markov Chains when the Planning Horizon is infinity'', Management science, 10, pp.716-733.

[19]. R.G Phillips and P. Kokotovic (1981), " A Singular Perturbation Approach to Modeling and Control of Markov Chains"; IEEE Transactions on Automatic Control, (AC-26) (pp.1087-1094).

[20]. Pravin. Varaiya (1978), ``Optimal and suboptimal stationary controls for Markov Chains``; IEEE Transactions on Automatic control vol AC 23 No 3, June 1978.

[21]. Ronald A. Howard (1960), ``Dynamic Programming and Markov Processes``. The M.I.T. Press Massachusetts institute of Technology Cambridge, Massachusetts N.Y.

[22]. Ronald A. Howard, Semi-Makovian Decision Processes, Proceedings International Statistical Institute, Ottawa, 1963.

[23]. Sheldom M. Ross (1971), ``Applied Probability Models with Optimization Applications`` (Holden-day.San Francisco).

[24]. Sheldom M. Ross (1983), ''Introduction to Stochastic Dynamic Programming''.N.Y: Academic 1983.

[25]. H.A. Simon and A. Ando (1961), ``Aggregation of variables in Dynamic Systems ``; Econometrica, vol .29, 111-138.

[26]. W. Jewel (1963), ``Markov Renewal Programming``, Op. Res, 11, pp. 938-971.

[27]. Paul J. Schweitzer (1986), "Iterative Solution of the Functional Equations of Undiscounted Markov Renewal Programming, Journal of Mathematical Analysis and Applications, 34, pp. 495-501, 1971".

Abdesselam Elbahja" Distruption of Semi-Markov Control Problems" Research Inventy: International Journal of Engineering And Science, vol. 08, no. 03, 2018, pp. 33-38