Polynomial RST Control for Blood Pressure Regulation

Ciprian Sandu, Dumitru Popescu, Cătălin Dimon
Automatic Control and Systems Engineering
Faculty of Automatic Control and Computers, University Politehnica of Bucharest
Bucharest, Romania
ciprian.sandu.cs@gmail.com, popescu_upb@yahoo.com, catalin.dimon@gmail.com

Abstract — ***This paper proposes polynomial robust RST control for the regulation of the blood pressure in post cardiac surgery patients. Results using Reinforcement Learning is also presented. Clinical applications require simple and effective controllers. This study's focus is to show that polynomial robust RST control is suitable for the blood pressure regulation in post cardiac surgery patients. A realistic and detailed model of a hypertensive patient is used; this model offers a drug response (here we use the model of Slate et al. (1980)), a model for an internal reflex of the body (the baroreceptor reflex) and disturbances. We report simulation results obtained by using polynomial robust RST control, because it ensures the tolerance of the system to the structure disturbances of the model, as well as to the inherent approximations. A result obtained with Reinforcement Learning is also presented.***

Keywords — blood pressure control, reinforcement learning, polynomial robust RST control, simulation results.

I. Introduction

Some of the natural control systems of the body are disrupted when a cardiac surgical intervention is performed. In this situation, the patients need to keep their mean arterial pressure within safe limits by using a drug, a vasodilator. There are advantages to including automatic control in the mean arterial pressure regulation of patients that suffered cardiac surgery. In such patients, the infusion of the fast-acting vasodilator - sodium nitroprusside (SNP) is adjusted as necessary, according to the levels of the mean arterial pressure.

The precision of the amount of the administrated vasodilator is improved by using automatic control. Hence, the goals of reducing costs, effort and safety are met faster.

The human body contains many natural control systems that are called homeostasis; these systems are complex, and may be disrupted when surgery is performed. Homeostasis is the maintenance of almost constant levels of various parameters in the body. Many parts of the body help maintain the homeostasis. By using automatic control we can provide better patient safety and reduce care effort. Automatic control uses the feedback concept. Automatic feedback control systems are used to control crucial variables by correcting the levels of manipulated variables, such as drug infusion rate based on measurements [11].

In the case of blood pressure regulation in post cardiac surgery patients, automatic control is very attractive.

Asla and others (Isaka and Sebald, 1993) had shown that automatic control could keep blood pressure within ±5 mm Hg in about 94% of the 110 hours of testing as compared to the 52% - in the case of the healthcare staff.

This paper focuses on showing an alternative to the previously used Reinforcement Learning and PID for blood pressure regulation, namely a polynomial robust RST control algorithm.

In section 2 we explain the reasons behind using automatic control for blood pressure regulation and present the problem formulation. Then we show the solutions that use Reinforcement Learning and PID. Section 3 presents the simulation results obtained with the polynomial robust RST control algorithm and a robustness analysis is also performed, in order to show that the proposed control solution can handle successfully model uncertainties, approximations used and system nonlinearities. In the end we mention the conclusions.

II. Problem formulation

A. *Reasons for using automatic control*

Automatic control of blood pressure in post cardiac surgery patients is suitable, because it increases safety, reduces costs effort: a controller maintains the arterial pressure within desired limits longer than the experienced healthcare staff.

The purpose is to regulate the arterial pressure of a patient who suffered cardiac surgery. For reaching this goal we use a regular, using feedback loop control [10].

The system identification for obtaining a mathematical model of a human is complex. We will use such a mathematical model in a closed feedback loop, with Reinforcement Learning. Then we will use polynomial robust RST control.

978-1-4799-1780-8/15 $31.00 © 2015 IEEE
DOI 10.1109/CSCS.2015.103

B. The mathematical model for the patient

A mathematical model for the patient, in order to be realistic – in the context of blood pressure regulation – has four components:

- a drug response model,
- models for internal reflexes,
- measurement dynamics (random noise due to respiration),
- patient movements.

We use the drug response model of Slate et al. (1980) [7]:

$$\frac{\Delta P_d(s)}{I(s)} = \frac{Ke^{-T_i s}(1+\alpha e^{-T_c s})}{\tau s+1}, \quad (1)$$

where:

$\Delta Pd(s)$ is the change in MAP (mmHg)

$I(s)$ is the infusion rate of SNP (ml h^{-1}),

K gives the patient's sensitivity (a high value of K implies a sensitive patient),

α is a recirculation index,

τ is a time constant,

T_i is the initial transport delay

T_c is the recirculation time.

The human body already contains internal reflexes that regulate the blood pressure. Some of them may be disrupted during surgery. Lee et al. (2005) identified that RAS (renin-angiotensin system) and BRS (baroreceptor reflex system) are necessary for a complete model of a patient for good blood pressure control. The former, as shown by Lee et al. (2005), "is an internal blood pressure buffering system that is activated when MAP drops below a threshold value. Through a series of chemical reactions from renin to angiotensin II, RAS can alter the total peripheral resistance of arterioles and hence increase blood pressure".

The typical range of threshold for activation of RAS is between 70 to 75 mmHg for the general population. Hahn et al. (2002) noted the existence of the threshold and range for RAS, but he did not use them [8].

Lee's model used the threshold value 72 mmHg and the range of 50-110: the "switch" simulates the RAS threshold of 72 mmHg whereas the saturation block represents the range of MAP (50 - 110 mmHg) where RAS remains effective [9].

The model for RAS proposed by Lee is shown in figure 1:

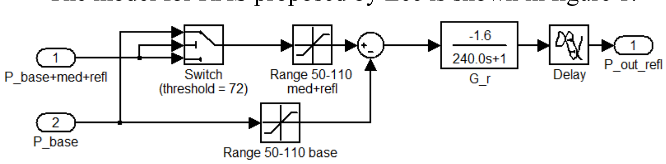

Fig.1: Lee's model for RAS

The model for RAS proposed in a previous paper that will be used here in the feedback control loop for blood pressure regulation is shown in figure 2.

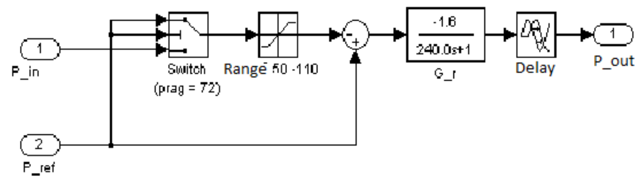

Fig.2: The improved model for RAS

Another important internal reflex of the human body is the baroreceptor reflex. The response of the baroreceptor reflex is shown in figure 3:

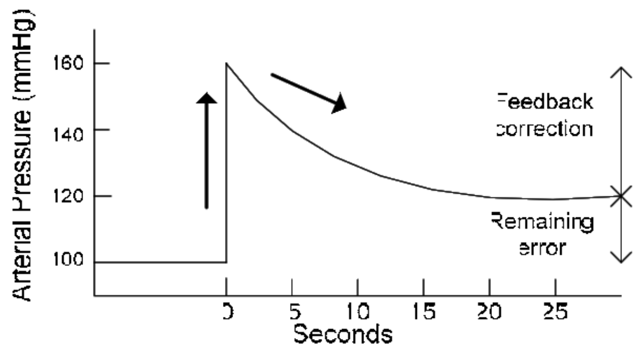

Fig.3: The theoretical response of the BRS reflex (Guyton, 1980)

C. Reinforcement Learning

The control problem is defined by a set of states in which the patient can be, a set of actions taken in order to influence the patient (the blood pressure), a transition function that makes the connections between actions and states and a reward function (that evaluates the performance). The goal is for the system to reach optimum states.

We are in the context of an optimal control problem, a Markov Decision Process (which is a mathematic framework for optimal decision making used for systems with uncertainty): a set of states to represent the environment, a set of actions to control it, the transition function for connecting the two and the reward function.

We will show results obtained with Reinforcement Learning (RL). The Markov property is critical for the RL algorithms (in terms of theoretical guarantees) [14], [15], [16].

Reinforcement Learning is a technique that belongs to the category of intelligent systems. It is a mathematical method of reaching optimal control in systems with uncertainty, disturbances, nonlinearity, delay, etc.

The controller is a Reinforcement Learning agent.

The Reinforcement Learning agent applies actions on a patient model. The agent receives the control error as an input, the difference between the measured blood pressure and the setpoint.

The states set x is defined by the error (the difference between the measured blood pressure and the setpoint) and it is a discrete set of values:

$$X = [1,2,\ldots,200] \qquad (2)$$

The action set u is the SNP infusion rates set (in ml per hour):

$$U = [1, 2, \ldots, 350] \qquad (3)$$

In reality, the values will be much smaller than 100.

As an effect of the action u_k, while being in the state x_k, the state will become the next state x_{k+1}, in conformity with the transition function:

$$f : X \times U \rightarrow X :$$

$$x_k + 1 = f(x_k, u_k) \qquad (4)$$

The utility of the states is given by the reward function; it decides the state value. The controller receives the reward r_{k+1}, in conformity with the reward function

r : X x U -> **R**

$$r_{k+1} = r(x_k, u_k) \qquad (5)$$

The reward shows the effect of action u_k, the transition from x_k to x_{k+1}, but does not estimate the long-term effects.

The controller takes the actions in conformity with its *policy*

$$h : X \rightarrow U :$$

$$u_k = h(x_k) \qquad (6)$$

Given f and r, the state x_k and the action u_k are enough to calculate the next state x_{k+1} and the reward r_{k+1}. This is the Markov property, which is critical for the RL algorithms (in terms of theoretical guarantees).

The policy is defined (after many calculations, to cover as many scenarios as possible) as follows:

$$u = \text{ref} - 145 + x^{*} \qquad (7)$$

where ref is the setpoint and x* is number that is defined using on the state x, that can take the following values:

[x/10], [x/5], x, 2x, 3x, 4x, 5x, 6x

where "[ ]" is the nearest integer.

The action takes a random value 20% of the time and the value that optimizes the reward – 80% of the time (after the learning session).

The reward function is defined as follows:

$$r(x, u) = 120000 - 1000(x + xx")/2$$
$$1000 + 10u - xx" - x \qquad (8)$$

where $x"$ is the previous state.

The first one is for small values for the states (lower than 10), the second one for big values for the states (10 or above). The reasons behind choosing the reward function are complex: detailed calculations, as well as empirical results obtained using a large number of scenarios. They will not be detailed here.

We applied a SARSA algorithm:

$$Q(x,u) = (1-\alpha)Q(x,u) + \alpha[r(x,u) + \gamma Q(x',u')] \quad (9)$$

where u', x' are the next action and respectively the next state. Each time the agent chooses an action and receives a reward, the Q function is updated. The algorithm gives total control over the duration of the actions.

We use the model of a patient (including the RAS reflex) with the only purpose of generating the input data (instead of gathering the data from a real patient). The closed loop used for Reinforcement Learning blood pressure regulation is shown in figure 4:

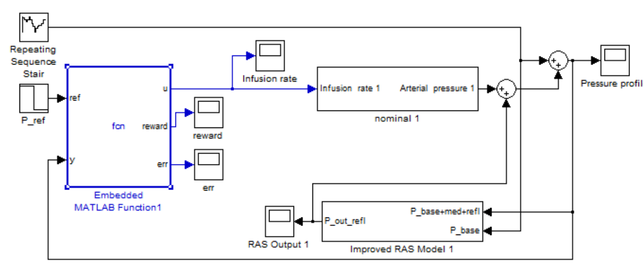

Fig 4.The feedback control loop with RL and RAS

In figure 5 we present the arterial pressure when controlled with the SARSA algorithm:

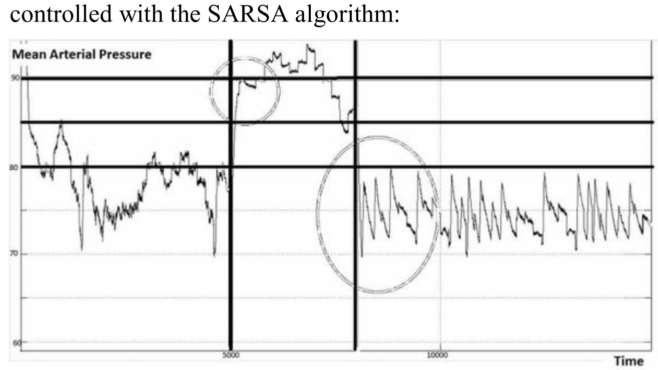

Fig.5: The arterial pressure when controlled with RL - the SARSA algorithm

On the vertical axis there is the arterial pressure.

On the horizontal axis there is the time.

Two important moments have been highlighted with circles and with vertical lines:

5000 seconds – when the agent switches from the random actions that make it learn to applying what it learned (20% random actions, 80% intelligent, informed actions)

8000 seconds – when the pressure setpoint changes from 90 to 70 mmHg.

We have also highlighted

The upper horizontal line the 90 mmHg level,

The middle horizontal line the 85 mmHg level,

The lower horizontal line the 80 mmHg level.

Between the 5000 and the 8000 seconds moments, the error is greater than 5 mmHg only once, and not with much (it is acceptable from a medical point of view). The behavior after 8000 seconds is due to the RAS reflex: when the pressure goes below 72 mmHg, the reflex increases the pressure. Again, the error is acceptable and the overshoot within desired limits.

D. *Discrete system transfer function*

In equation (1) the non-standard structure of the transfer function can be observed. We will use the following values for its parameters:

K = 0.714, α = 0.4, τ = 40, T_i = 30 [s], T_c = 45 [s]

After adding the BRS reflex to the model, approximating the lag times and using the “zero order hold” discretization method, we obtain the following discrete transfer function of the model (with a sampling time equal to 5 seconds):

$$H_p(z^{-1}) = \frac{0.01112\, z^{-1} + 0.00199\, z^{-2} - 0.00530\, z^{-3}}{1 - 2.14500\, z^{-1} + 1.44400\, z^{-1} - 0.29050\, z^{-3}}$$

The dynamic of the process can be observed in figure 6:

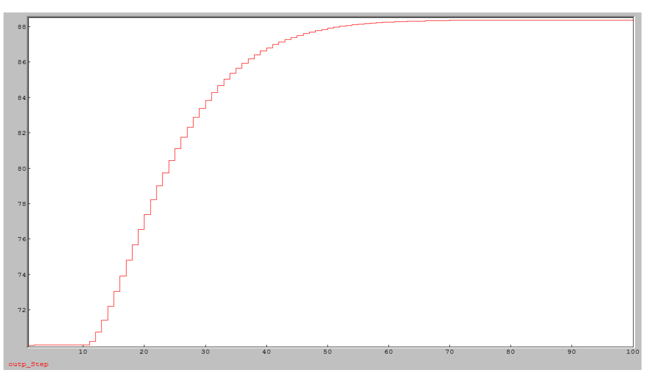

Fig.6: Step response of the discretized plant model
(Mean Arterial Pressure – mmHg, time – number of samples)

The step response behaves as expected in terms of the delay and of the time constant, as well as in terms of the shape of the curve.

III. DESIGNING THE POLYNOMIAL RST ALGORITHM

A. *The polynomial RST structure*

In order to achieve good tolerance to structure disturbances of the model, as well as to its approximations, we propose an RST polynomial control algorithm, detailed in figure 7:

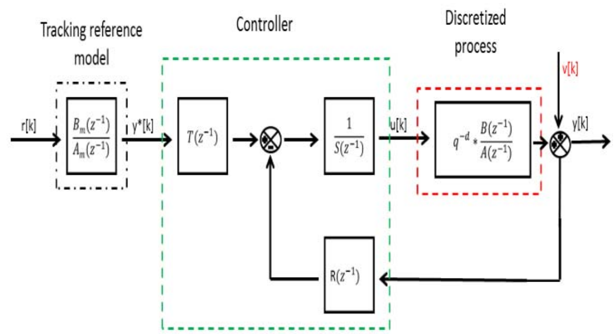

Fig.7: The polynomial RST control structure

where:

$$R(z^{-1}) = r_0 + r_1 z^{-1} + \dots + r_{nr} z^{-nr}$$

$$S(z^{-1}) = 1 + s_1 z^{-1} + \dots + s_{ns} z^{-ns}$$

$$T(z^{-1}) = t_0 + t_1 z^{-1} + \dots + t_{nt} z^{-nt}$$

$$B(z^{-1}) = b_1 z^{-1} + b_2 z^{-2} + \dots + b_{nb} z^{-nb}$$

$$A(z^{-1}) = 1 + a_1 z^{-1} + \dots + a_{na} z^{-na}$$

The closed loop transfer function has the following expression:

$$H_{cl}(z^{-1}) = \frac{z^{-d}\, T(z^{-1}) B(z^{-1})}{A(z^{-1})\, S(z^{-1}) + z^{-d} B(z^{-1})\, R(z^{-1})}$$

$$= \frac{z^{-d}\, T(z^{-1}) B(z^{-1})}{P(z^{-1})}$$

$$\text{with } P(z^{-1}) = p_1 * z^{-1} + p_2 * z^{-2} + \dots + p_{np} * z^{-np}$$

where P is the characteristic polynomial

The characteristic polynomial P contains the dominant and auxiliary poles of the closed loop system. A second degree transfer function is usually used for imposing the dominant poles, thus the desired behavior of the closed loop system [1].

The coefficients of the R and S polynomials are determined using the pole placement method from the equation:

$$P(z^{-1}) = A(z^{-1})S(z^{-1}) + z^{-d}B^{*}(z^{-1})R(z^{-1})$$

by solving the algebraic equation,

$$Mx = p$$

where:

$$x^T \stackrel{\text{def}}{=} [1\; s_1 s_2 \dots s_{nb+d} r_0 r_1 \dots r_{na-1}]$$
$$p^T \stackrel{\text{def}}{=} [1\; p_1 p_2 \dots p_{na+nb+d}]$$

M is the Sylvester matrix, associated to the coefficients of the polynomials A and B.

The coefficients of the T polynomial are determined from the following expression:

$$T(z^{-1}) = K * P(z^{-1})$$

where:

$$K(z^{-1}) = \begin{cases} \dfrac{1}{B(1)}, & B(1) \neq 0 \\ 1, & B(1) = 0 \end{cases}$$

Using the dedicated software WinReg for computing the RST control, the following results are obtained:

$$R(z^{-1}) = 125.187943 - 120.395544\, z^{-1}$$

$$-67.996674\, z^{-2} + 109.386953\, z^{-3} - 27.658607\, z^{-4}$$

$$S(z^{-1}) = 1 - 0.639167\, z^{-1}$$

$$-1.188589\, z^{-2} + 0.323141\, z^{-3} + 0.504615\, z^{-4}$$

$$T(z^{-1}) = 128.040973 - 178.242894\, z^{-1}$$

$$+ 68.725992\, z^{-2}$$

Figure 8 shows the step response of the closed loop system, for a setpoint change from 70 mm Hg to 90 mm Hg, and for a step disturbance response applied after 1900 seconds, using the polynomial robust RST controller.

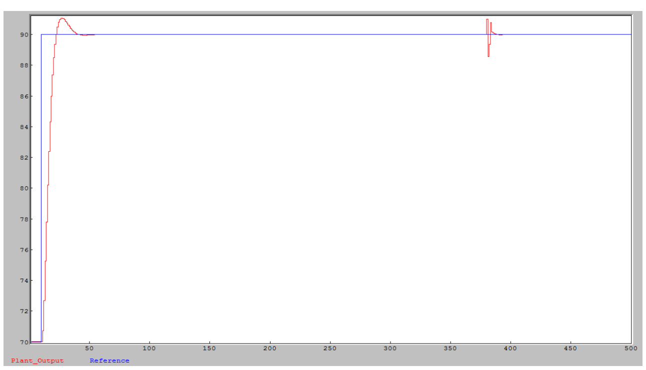

Fig.8: System tracking and rejecting disturbances
(Mean Arterial Pressure – mmHg, time – samples)

The associated control algorithm is shown in figure 9:

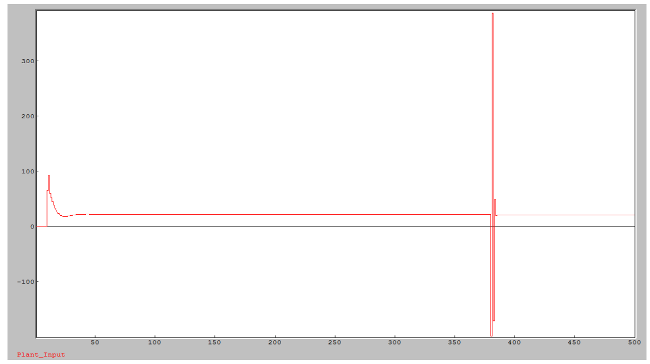

Fig.9: Nominal control algorithm
(quantity of administered drug – ml, time – samples)

Considering the step response, a good dynamic evolution can be observed, with an appropriate stationary behaviour.

It can be observed that the disturbance is rejected successfully. Since the behaviour of the control algorithm means nothing more than the rate and quantity of adminstrated drug, it does not affect the patient in other ways than what can be observed in the output characteristic.

B. Robustness analysis

For the purposes of a good robustness analysis, the following indicators are used [13]:

- the sensitivity function S_{yp} for the closed loop system, that has the expression:

$$S_{yp}(j\omega) = \frac{A(j\omega) * S(j\omega)}{A(j\omega) * S(j\omega) + z^{-d} * B(j\omega) * R(j\omega)}$$

and

- the associated robustness (modulus) margin that is defined as:

$$\Delta M \stackrel{\text{def}}{=} \min_{\omega \in R} |1 + H(j\omega)|$$

The relationship between the two indicators is given by the following relation:

$$\Delta M|_{dB} = -\max_{\omega \in R} \left| S_{yp}(j\omega) \right| \Big|_{dB}$$

A bigger value of ΔM implies a better robustness. A high value of ΔM implies a small maximum value of S_{yp}. Therefore, a small value of the maximum value of S_{yp} means a better robustness [12].

A representation of the sensitivity function is given in figure 10:

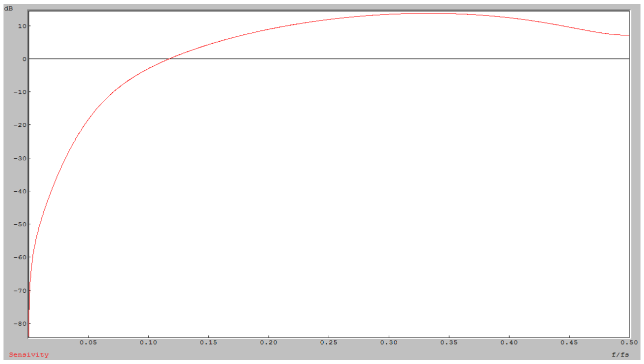

Fig.10: Sensitivity function Syp (RST)

After computing with the WinReg *Robustness Analysis* routine, ΔM= -13.74 *dB*.

In order to have a better robustness, a correction must be applied to the algorithm. Using the WinREG software, we can prespecify a pole for the algorithm. We will alter the S polynomial:

$$S(q^{-1}) = HS(q^{-1}).S'(q^{-1})$$

where S' is computed by the software.

After this correction, the S polynomial becomes:

$$S(z^{-1}) = 1 - 1.697906\, z^{-1}$$

$$-0.319318\, z^{-2} + 1.017224\, z^{-3}$$

The sensitivity function becomes:

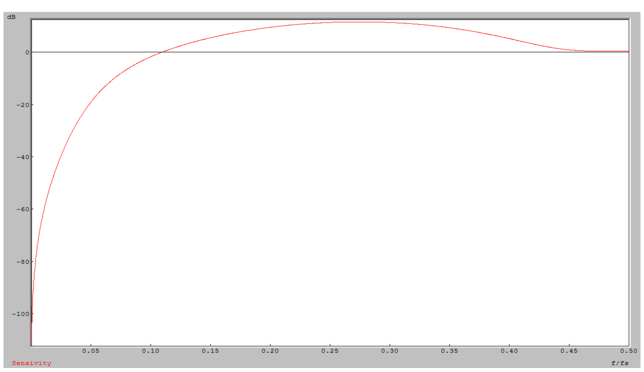

Fig.11: Sensitivity function Syp (RST) after correction

After applying the correction, ΔM= -11.63 *dB*. Therefore, an improvement has been achieved.

IV. CONCLUSIONS

The study presented a drug response model used firstly as a source a pseudo-clinical data for a reinforcement learning algorithm based closed loop control system. The goal of the study was to show that the polynomial robust RST control is also suitable for blood pressure regulation. Also the closed loop system performances are imposed in order to be in conformity with the medical practices. The method used for the control law computation is the pole allocation. A robustness analysis was made, taking into account the frequency characteristics of the sensitivity function, in order to show that the proposed control solution can handle successfully model uncertainties and system nonlinearities.

In the future, our results obtained in simulation can be compared with the other advanced control techniques such as adaptive control.

References

[1] C.Lupu,D.Popescu,C.Dimon, *Solu□ii Practice de Conducere a Proceselor Neliniare*, Editura Politehnica Press, Bucure□ti 2010

[2] I.D.Landau,Gianluca Zito, *Digital Control Systems-Design, Identification and Implementation*, Springer-Verlag London Limited, 2006.

[3] I.Dumitrache, T. Dragomir, *Automatica Volumul II*, Editura Academiei Române, Bucure□ti, 2013.

[4] D.Popescu, R.Dobrescu, F.Ionescu, D.Stefanoiu, *Modelare in Ingineria Proceselor Industriale*, Editura AGIR, Bucuresti 2011.

[5] M.Voicu, *Teoria Sistemelor*, Editura Academei Romane, Bucuresti 2008.

[6] K.Y.Astrom,R.M. Murraay, *Feedback Systemes : An introduction for Scientists and Engineers*, Princeton University Press,2008.

[7] Slate, J. B., Sheppard, L. C., Rideout, V. C. and Blackstone, E. H. (1980). *Closed-loop nitroprusside infusion: Modeling and control theory for clinical application*.

[8] Hahn, J., Edison, T. and Edgar, T. F. (2002). *Adaptive IMC control for drug infusion for biological systems*. Control Engineering Practice. Vol. 10. pp 45 – 56.

[9] Lee, H. W. (2004). *Analysis of Models and Techniques for Blood Pressure Control in Patients*. *B. Eng. Thesis*. National University of Singapore.

[10] H.W. Lee, S. Lakshminarayanan, G.P. Rangaiah (2005). *Models and Simple Controllers for Blood Pressure Regulation in Post Cardiac Surgery Patiens.* Journal of The Institution of Engineers, Singapore. Vol. 45 Issue 6

[11] Arthur C. Guyton, John E. Hall (2006). *Textbook of Medical Physiology* – Eleventh Edition.

[12] P. Borne, F. Filip, D. Popescu, M. Benrejeb. *Automatique Avancee et Informatique Appliquee,* Ed. Academiei Romane, Bucuresti, Romania, 2009, ISBN 978-973-27-1806-3, pg. 250.

[13] L.Foulloy, D. Popescu, G. Dauphin-Tanguy. *Modelisation,Identification et commande des Systemes, Ed. Academiei Romane, Bucuresti, Romania, 2004*

[14] Brett L. Moore, PhD, Anthony G. Doufas, MD, PhD and Todd M. Quasny, MS, *Reinforcement Learning: A novel Method for Optimal Control of Propofol-Induced Hypnosis*, International Anesthesia Research Society, Anesthesia – Analgesia, February 2011, Vol. 112, Number 2.

[15] Brett L. Moore, B.S., B.A., M.S., *Reinforcement Learning for Patient-Specific Propofol Anesthesia: A Human Volonteer Study*, 2010.

[16] Lucian Busoniu, Robert Babuska, Bart De Schutter and Damien Ernst. *Reinforcement Learning and Dynamic Programming Using Function Approximators.*