Multi-Critics Based Intelligent Control

MEHRDAD FATOURECHI
Department of Electrical and Computer Engineering
University of British Columbia
Vancouver, CANADA

CARO LUCAS
Department of Electrical and Computer Engineering
University of Tehran
Tehran, IRAN

ALI KHAKI SEDIGH
Department of Electrical Engineering, K.N.Toosi
University of Technology
Tehran, IRAN

Abstract: - A new approach for the multi-objective control of dynamic plants is presented based on the agent concept. The control system consists of a neurofuzzy controller whose weights are adapted according to emotional signals provided by blocks called emotional critics. Each critic is assigned to assess the situation of its corresponding control objective. Simulation results are provided for the control of different dynamical systems in order to clarify the matter further.

Key-words: -Intelligent control, multi-objective systems, emotional learning, neurofuzzy control, agents

1 Introduction

It is widely believed that decision making, even in the case of human agents, should be based on full rationality and emotional cues should be suppressed in order to not influence the logic of arriving at proper decisions. The assumption of full rationality, however, has sometimes been abandoned in favor of satisfying or bounded rationality models [1], and in recent years, the positive and important role of emotions have been emphasized not only in psychology, but also in AI and robotics ([2]-[4]). Very briefly, emotional cues can provide an approximate method for selecting good actions when uncertainties and limitations of computational resources render fully rational decision-making based on Bellman-Jacobi recursions impractical. In past researches ([5-8]), a simple cognitive/emotional state designated as *stress* has been successfully utilized in various control applications. This approach is actually a special case of the popular reinforcement learning technique. However, in this case it is believed that since the continual assessment of the present situation in terms of overall success or failure is no longer simple behaviorist type of conditioning but it is closer to the definition of cognitive state modification and adaptation learning, the designation of emotional learning seems more appropriate. We should emphasize that here emotion merely refers to *stress* cue, and use of other, and perhaps higher emotional cues are left for future researches.

On the other hand, multi-objective control can be thought as the primary target of every control scheme because not only we want to control the outputs of the control system, but also the control signal and internal states of the system. In this paper, we've applied emotional learning for the multi-objective of both the control signal and the internal states of the system.

In the proposed method, a *critic* is assigned to each *control objective*. Each critic's task is to assess the control situation of its corresponding control objective and to provide the appropriate *stress (emotional)* signal which describes the extent of the success or failure of that particular part of the control system. Then based on these stress signals the controller provides the appropriate control signal to be applied to the plant.

The organization of the paper is as follows: The focus of section 2 is on the emotional learning. The structure of the proposed controller and its adaptation law are developed in section 3 and section 4 is divided into two sections itself. In 4.1 we consider control of the amplitude of the control signal as the secondary control objective (besides input tracking) and in 4.2, stabilization of internal states is considered as the secondary objectives. In each sub-section, simulation results are provided to clarify the matter further. Finally, conclusions are addressed in section 5.

2 Emotional Learning

According to psychological theories, some of the main factors of human beings' learning are emotional elements such as satisfaction and stress. Emotions can be defined as states elicited by instrumental reinforcing stimuli, which

if their occurrence, termination or omission is made contingent upon the making of a response, alter the course of future emission of that response [9].

Emotions can be accounted for, as a result of the operation of a number of factors, including the following [9]:

1. The reinforcement contingency (e.g. whether reward or punishment is given, or withheld).

2. The intensity of reinforcement

3. Any environmental stimuli might have a number of different reinforcement associations.

4. Emotions elicited by stimuli associated with different reinforcers will be different.

It should also be mentioned that in this paper, emotion merely refers to *stress* cue and other (and perhaps higher) emotions are not considered here. In our proposed approach, which in a way is a cognitive restatement of reinforcement learning in a more complex continual case (where reinforcement is no longer a binary signal), there exists an element in the control system called *emotional critic* whose task is to assess the present situation which has resulted from the applied control action in terms of satisfactory achievement of the control goals and to provide the so called emotional signal (the stress). The controller should modify its characteristics so that the critic's stress is decreased. This is the primary goal of the proposed control scheme, which is similar to the learning process in the real world because in the real world, we also search for a way to lower our stress with respect to our environment ([10-11]).

As seen, emotional learning is very close to reinforcement learning, but the main difference between them is that in the former the reinforcement signal is an analog emotional cue that represents the cognitive assessment of future costs given the present state. So here the system does not wait for a total failure to occur before it starts learning. Instead, it continues its learning process at the same time as it applies its control action. The resulting analog reinforcement signal constitutes the stress cue, which has been interpreted as cognitive/emotional state.

Based on these concepts, we have proposed an emotion-based approach for the control of dynamic systems, which will be discussed in the next section.

3 An Emotion-based Approach to the Multi-Objective Control

The general structure of the proposed multi-objective control system is as follows: It consists of a neurofuzzy controller, which has four layers. The first layer's task is the assignment of inputs' scaling factors in order to map them to the range of [-1, +1] (the inputs are chosen as the error and of the output and its derivative, respectively). In the Second layer, the fuzzification is performed by assigning five labels for the input. For decision-making, max-product law is used in layer 3. Finally, in the last layer, the crisp output is calculated using Takagi- Sugeno formula [12],

$$y = \frac{\sum_{l=1}^{p} u_l (a_l x_1 + b_l x_2 + c_l)}{\sum_{l=1}^{p} u_l} \qquad (1)$$

Where x_1 and x_2 are inputs to the controller (the error of the output and its derivative, respectively), u_l, p, and y are the l'th input of the last layer, number of rules in the third layer and output of the controller, respectively and a_l's, b_l's and c_l's are parameters to be determined via learning.

For each control objective, a critic is assigned whose task is to assess the control situation of that objective and to provide the appropriate emotional signal. The emotional signals provided by these critics contribute collaboratively to update output layer's learning parameters of each controller. The aim of the control system is the *minimization of the sum of squared emotional signals*. Accordingly, first we describe the error function E as follows,

$$E = \sum_{j=1}^{m} K_j (\tfrac{1}{2} r_j^{\,2}) \qquad (2)$$

Where r_j is the output signal of j's critic, K_j is the corresponding weights and m is the total number of control objectives (for the special case of single-objective systems, K_j=1 and m=1)

For the adjustment of controllers' weights the steepest descent method is used,

$$\Delta \omega = -\eta \frac{\partial E}{\partial \omega} \qquad (3)$$

where η is the *learning rate* of the neurofuzzy

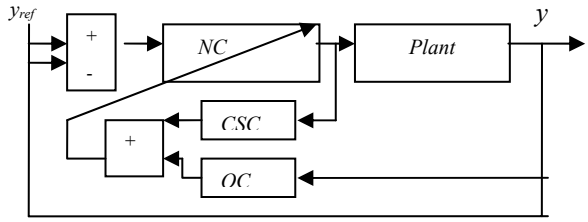


Fig.1. The control loop in the case of reduction of control signal

controller .

In order to calculate the RHS of (3), the chain rule is applied,

$$\frac{\partial E}{\partial \omega} = \frac{\partial E}{\partial r_j} \cdot \frac{\partial r_j}{\partial y} \cdot \frac{\partial y}{\partial u} \cdot \frac{\partial u}{\partial \omega} \qquad (j = 1,2,\ldots, m) \quad (4)$$

From (2), we have,

$$\frac{\partial E}{\partial r_j} = K_j \cdot r_j \qquad (j = 1,2,\ldots,m) \qquad (5)$$

Also,

$$\frac{\partial y}{\partial u} = J \qquad (6)$$

where J is the Jacobian of the system. This yields:

$$\Delta\omega = \sum_{j=1}^{m} K_j . r_j \cdot \frac{\partial r_j}{\partial y} . J . \frac{\partial u}{\partial \omega} \qquad (j=1, 2, m) \quad (7)$$

Equation (7) is the general form used for updating the learning parameters a_l's, b_l's and c_l's in (1), which their calculations from this equation is straightforward. Calculation of the partial derivative term involving r_j depends on the type of the objectives that are considered in the multi-objective control scheme.

4. Applications of Emotion-based Multi-objective control

In this section two applications are considered for the proposed algorithm in the previous section. In the first application, we consider reduction of control signal as our second objective and in the second example; we consider stabilization of the internal states of the control system.

4.1 Reduction of control signal by means of emotional control

The magnitude of the control signal is of high importance in control system design because although our primary goal in the design of a control system is to achieve stability but this may not happen if the control signal's magnitude reaches out of the boundaries of controller's actuator, resulting in a chopped control signal and failure in the control system performance.

The schematic of the control system is shown in Fig.1. The neurofuzzy controller (NC) has the same structure as what was discussed in the previous section. Two critics are assigned for the assessment of the output signal and control signal; namely Output Critic (OC) and Control System Critic (CSC). Fuzzy critics are applied here. In order to obtain the learning rule for this case, first we define the following cost function:

$$E = K_1 E_1 + K_2 E_2 \qquad (8)$$

Where E_1 and E_2 show the errors corresponding to the OC and CSC respectively and K_1 and K_2 are their corresponding weights which determine the importance of the error functions in the calculations. These coefficients are non-zero but selection of their value is in the hands of the designer based on the level of importance of the goals.

We define the error function E_1 associated with the output as follows:

$$E_1 = \tfrac{1}{2} r_1^2 \qquad (9)$$

Where r_1 is the emotional signal describing the status of the output. In order to update the weights of the controller, we use (7) to obtain the weights updating rule ($m=1$).

Before that, notice:

$$\frac{\partial r_1}{\partial y} = -\frac{\partial r_1}{\partial e} = -k \qquad (k > 0) \qquad (10)$$

Also because the on-line calculation of J is accompanied with the computational burden, and since this parameter usually has a positive value in dynamical systems, we use the sign (+1) as a replacement. Hence, (7) becomes

$$\Delta\omega_1 = r_1 . \frac{\partial u}{\partial \omega} \qquad (11)$$

In a similar manner to our approach for the output error, we define the error function E_1 associated with the control signal as follows

$$E_2 = \tfrac{1}{2} r_2^2 \qquad (12)$$

Where r_2 is the emotional signal describing the status of the control signal. In order to update the weights of the

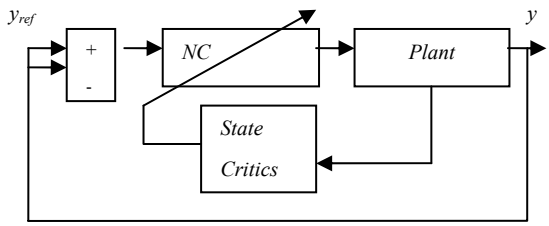


Fig.2.the control loop in the case of internal stabilization

controller, we use (7) to obtain the weights updating rule. Before that, notice if we define r_2 as the control signal itself, we have

$$\frac{\partial r_2}{\partial u} = 1 \qquad (13)$$

Hence, (7) becomes $\Delta\omega_2 = -r_2 \cdot \frac{\partial u}{\partial \omega}$ (14)

And the total weight update scheme will become as follows:

$$\Delta\omega = K_1 . \Delta\omega_1 + K_2 . \Delta\omega_2 \qquad (15)$$

4.2 Internal Stabilization of the Control System by means of emotional control

Internal stabilization, i.e. stabilization of internal states of a given control system, is one of the key factors which should be considered in the design of every control system. achieving output stability without internal stability leads to the destruction of some parts of the control system and ultimately failure of the control method.

Here we will apply our multi-critic methodology to address this problem. The structure of the control system is shown in Fig.2. Here State Critics (SCs) judge the states behavior and provide the so-called emotional (stress) signals. NC then applies the appropriate control action based on these stress signals. Considering $r_j = x_j$ (where x is the state of the system and j is the state index), We define the following cost function as the function to be minimized via learning:

$$\Delta\omega \propto \frac{\partial E}{\partial \omega} = \sum_{j=1}^{N} \frac{\partial E}{\partial x_j} \frac{\partial x_j}{\partial u} \frac{\partial u}{\partial \omega} \qquad (16)$$

Again because of on-line calculation problems, we'll assume

$$\frac{\partial x_j}{\partial u} = 1 \qquad (17)$$

Using (2) and (17), (16) becomes

$$\frac{\partial E}{\partial \omega} = -\sum_{j=1}^{N} (K_j x_j) \frac{\partial u}{\partial \omega} \qquad (18)$$

From (16) to (18) we'll have

$$\Delta\omega = -\eta \sum_{j=1}^{N} (K_j x_j) \frac{\partial u}{\partial \omega} \qquad (19)$$

This is the update rule for the weights of the control system.

5. Simulation Results

In this section, we'll apply the proposed method for the multi-objective control of two dynamical systems. In the first example, we'll control the amplitude of the control signal in the Van Der Pol system and in the second one we examine internal stabilization of the famous Inverted Pendulum system.

5.1. Example 1: Multi-Objective Control of Van Der Pol System

The nonlinear Van der Pol system is defined the following equation:

$$\ddot{x} + (1 - x^2)\dot{x} + x = u$$
$$y = x \qquad (20)$$

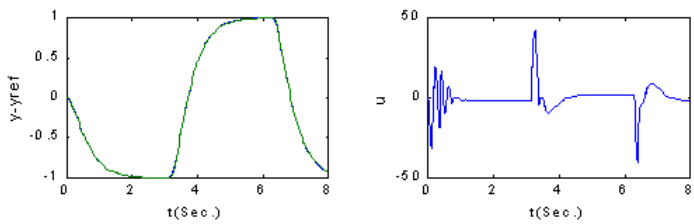
In (20), u is the control signal, x is the system state and y is the system's output. The objective is not only to control the output signal y but also to control the amplitude of the control signal.

Table.1 shows three different control situations in response to the reference input y_{ref} and Fig.3 shows the corresponding results. In the first case, only the OC is included in the simulations. As it can be seen, the maximum value of the control signal is 42.6774. Suppose this value is beyond the threshold levels of the actuators and causes their saturation. The second and the third cases show the results of the simulation after including CSC in the control system. As the results clearly show with the growth of the importance of CSC (increasing he corresponding K_2 coefficient), the maximum of the control signal is decreased to nearly the half of its original value while as Fig.3 shows at the same time, the tracking error of the output is slightly increased. This was already

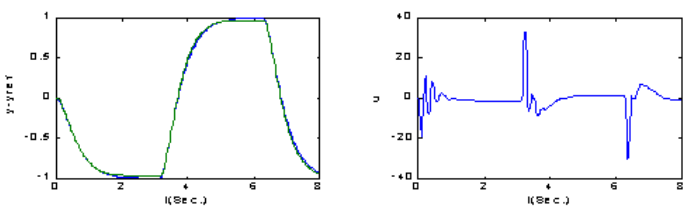
expected because increasing the value of the CSC weight means the decrease in the OC importance which results in poorer control situation.

Table.1. the different Control situations in Example 1

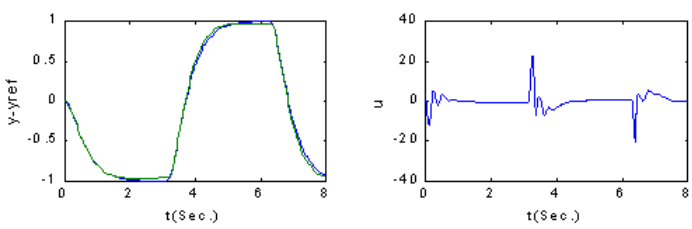
	$Max(\lvert u \rvert)$	$\int (\lvert u \rvert)$
$K_1=1,\ K_2=0$	42.6774	34.3234
$K_1=1.176,\ K_2=0.008$	33.1774	22.8981
$K_1=1.176,\ K_2=0.012$	22.3068	15.3917



(a)



(b)



(c)

Fig.3. Different Control Situations in Example.1

5.2 Example 2: Multi-Objective Control of an Inverted Pendulum:

The problem of balancing an inverted pendulum on a moving cart is a good example of a challenging control situation, due to its highly nonlinear equations, non-minimum phase characteristics and the problem of handling two outputs with only one control input. Here, the dynamics of the inverted pendulum are characterized by four variables: θ (angle of the pole with respect to the vertical axis), $\dot{\theta}$ (angular velocity of the pole), z (position of the cart on the track), and $\dot{z}$ (velocity of the cart). The behavior of these state variables is governed by the following two second-order differential equations [12]:

$$\ddot{\theta} = \frac{g * Sin\theta + Cos\theta * \left(\frac{-F - m*l*\dot{\theta}^2 * Sin\theta}{m_c + m}\right)}{l * \left(\frac{4}{3} - \frac{m * Cos^2\theta}{m_c + m}\right)} \tag{21}$$

$$\ddot{z} = \frac{F + m*l*(\dot{\theta}^2 * Sin\theta - \ddot{\theta} * Cos\theta)}{m_c + m} \tag{22}$$

Where g (acceleration due to gravity) is 9.8 m/s^2, m_c (mass of cart) is 1.0 kg, l (half-length of pole) is 0.5 m, and F is the applied force in Newton. Our control goal is to balance the cart, yet keep the z not further than 2.5 meters from its original position.

Considering the following notation $\theta=x_1$, $\dot{\theta}=x_2$, $z=x_3$ *and* $\dot{z}=x_4$, we'll use PID critics for the simplicity of the online calculation and examine three different control situation shown in Fig.4.As it's apparent from the situation, handling the internal states is easily done by the proposed control system.

6. Conclusions

In this paper, we introduced an agent-based structure for the multi-objective control of SISO plants based on the concepts of emotional learning and agents. Two major drawbacks of the control systems design (magnitude of control signal and stability of internal states) were addressed as the applications of the proposed methodology. The results of these simulations and other simulations as well [5-8] clearly shows effectiveness of the presented methodology.

Our future works include applying genetic algorithms for the optimal selection of the control system's coefficients (both the controller's and the

critics'). Also we can consider more complex structures for both the controller and the critics. Other aspects include changing the structure of the controller so that it could be applied to processes with unknown delays, applying multi-objective control to multivariable control systems, optimizing the structure of the controller, and finally considering more complex cues in the learning process.

References

[1] A. H. Simon and Associates, Decision making and problem solving, *Interfaces,* no.17, 1987.

[2] C. Balkenius and J. Moren, A Computational Model of Context Processing, *6th International Conference on Simulation of Adaptive Behavior*, Cambridge, 2000.

[3] M.El-Nasr, T. Loerger, and J.Yen, Peteei: A Pet with Evolving Emotional Intelligence, *Autonomous Agents99*, 1999, pp.9-15.

[4] J. Velasquez, a Computational Framework for Emotion-Based Control, *the Grounding Emotions in Adaptive Systems Workshop SAB'98*, and Zurich, Switzerland, 1998.

[5] M. Fatourechi, C. Lucas, and A. Khaki Sedigh, An Agent-based Approach to Multivariable Control, *IASTED International Conference on Artificial Intelligence and Applications*, Marbella, Spain, pp.376-381, 2001.

[6] M. Fatourechi, C. Lucas and A. Khaki Sedigh, Emotional Learning as a New Tool for Development of Agent-based Systems, Informatica, in press, 2001.

[7] M. Fatourechi, C. Lucas and A. Khaki Sedigh, Reduction of Maximum Overshoot by means of Emotional Learning, *6th Annual CSI Computer Conference*, Isfahan, Iran, 2001, pp.460-467.

[8] C. Lucas and S.A. Jazbi, Intelligent motion control of electric motors with evaluative feedback, *Cigre98*, Cigre, France, 11-104, 1998, pp.1-6.

[9] E. D. Rolls, *the Brain and Emotion*, Oxford University Press, 1998.

[10] K.M. Galloti, *Cognitive psychology in and out of laboratory (2nd ed.)*, Brooks/Cole, Pacific Grove, CA, 1999.

[11] R.W. Kentridge and J.P. Aggleton, Emotion: Sensory representations, reinforcement and the temporal lobe, *Cognition and Emotion 4*, 1990, pp.191-208.

[12] T. Takagi and M. Sugeno, Derivation of fuzzy control rules from human operator's control actions, *IFAC Symp.on Fuzzy Information, Knowledge Representation an dDecision analysis,* 1983, pp.55-60.

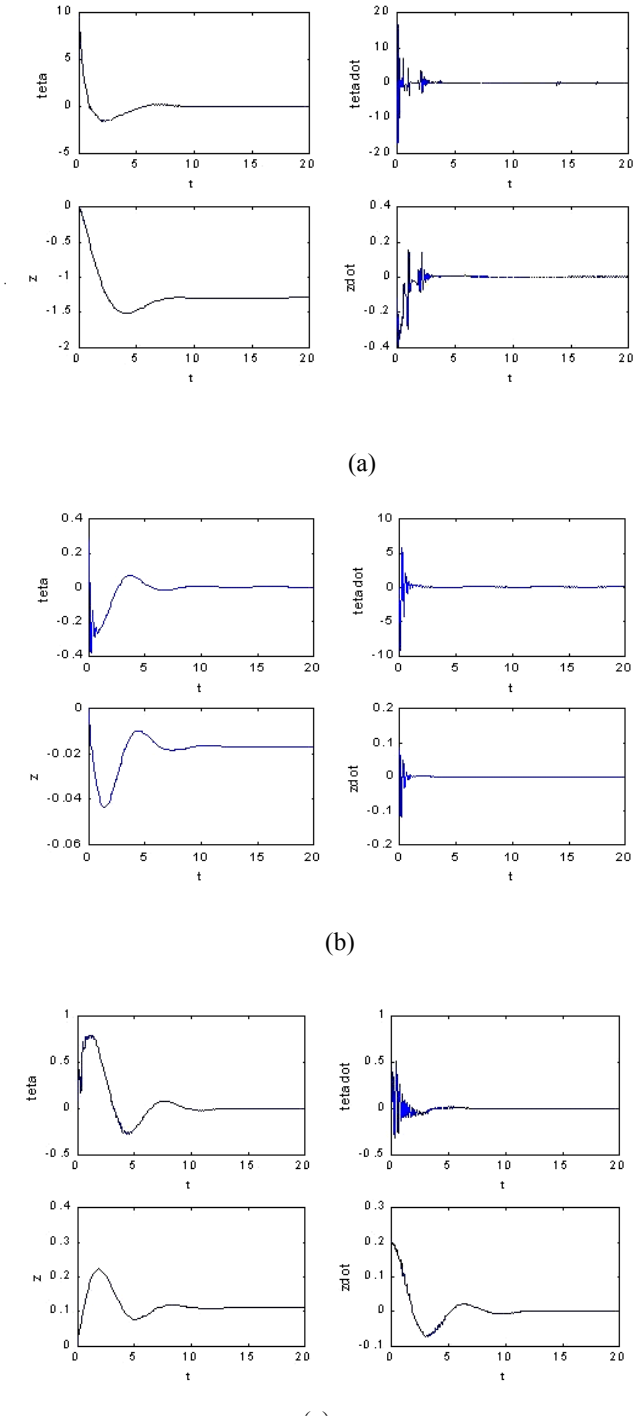


Fig.4. Results of simulation for three different state sets in Example 2 (a) $x_1(0) = 10$, (b) $x_2(0) = 5$, (c) $x_4(0) = 0.2$