Proceedings of the 4th WSEAS/IASME Int. Conf. on System Science and Simulation in Engineering, Tenerife, Spain, December 16-18, 2005 (pp53-59)

Design of a Multi Agent Adaptive Critic Based Neuro-Fuzzy Controller for Multi-objective Nonlinear Systems

MEHRAN RASHIDI FARZAN RASHIDI

Faculty of Engineering, Hormozgan University
Bandar-Abbas, IRAN
Email: mrashidi@mehr.sharif.edu

Abstract – In this paper, a multi agent controller for meeting different criteria, based on neuro-fuzzy controller is presented. The proposed controller is motivated by the affective and emotional faculties in human begins, which constantly evaluate the current states with respect to the achievement of the desired goals. For meeting different criteria, the controller consists of several critic agents that each agent tries to meet its goal. The combination of emotions of these agents applies on the controller in order to adapt the learning coefficients to achieve predefined criteria and goals. The proposed controller, also continuously evaluates the current states from critic agents and incremental achievement or disachievement of the set objectives, and self tune its control action accordingly. The controller is based on intelligent neuro-fuzzy architecture that suitable for online training algorithms. The effectiveness of the proposed method is demonstrated trough examples in which the proposed system is used for reducing control effort and tracking error simultaneously. The contribution of critic's emotions in multi criteria satisfaction is highlighted through these examples.

Key-Words: Multi Agent, Neural Network, Fuzzy Logic, Nonlinear Systems, Adaptive Critic

1 Introduction

In multi agent controllers, there are more than one controller or optimization/adaptation criteria that we called them agent and each agent is trying to reach to its goal independently [1, 2]. Each agent analyzes the control results with predefined criteria and produces an emotional feedback. These emotional feedbacks are used to make necessary changes in learning coefficients of controller in order to satisfy critic agent goals.

Emotional Learning is one of powerful learning methods that its dates back to the early research work in psychology, neuro science and computer science. Emotional learning can be categorized in reinforcement learning methods [3]. In the last twenty years, there have been rapidly increasing interest in reinforcement learning and mainly in emotional learning. Emotional learning is a kind of unsupervised learning methods for autonomous agents to acquire action rules to adapt clue of emotional reward and punishment. In emotional learning the teacher of conventional supervised learning is replaced by an intelligent critic that assesses the performance of controller and evaluates the current states of system and generates proper emotional reinforcement signal [4]. Fuzzy system theory provides a mathematical framework for modeling vagueness and imprecision data [5]. Neural networks have the ability to learn complex mapping, generalize known data and classify inputs [6]. Hybrid system utilizes the advantages of both, as well as other novel techniques, creating powerful tools for intelligent control [7].

In this paper we use a powerful architecture for the main controller [8] and use multi agent critics in order to train the main controller so that can satisfy all critics' criteria and emotions. By using of emotional learning, our proposed multi agent controller can be learned to satisfy all critics simultaneously. On line training, fast convergence of controller and robustness of proposed controller are other advantages of this method that be shown by some simulation results.

2 Multi Agent Adaptive Critic Based Neuro-Fuzzy Controller

This section, describes the proposed multi agent adaptive critic based neuro-fuzzy controller. The block diagram of the proposed system is shown in figure 1. As shown in this figure, it contains four main items as controller, plant, adaptive critic agents and learning mechanism. In the subsequent sections, a briefly discuss of the above elements is presented.

3. Neuro-fuzzy Controller

Two major approaches of trainable neuro-fuzzy models can be distinguished. The network based Takagi-Sugeno fuzzy inference system and the locally linear neuro-fuzzy model. It is easy to see that the locally linear model is equivalent to Takagi-Sugeno fuzzy model under certain conditions, and can be interpreted as an extension of normalized RBF network as well.

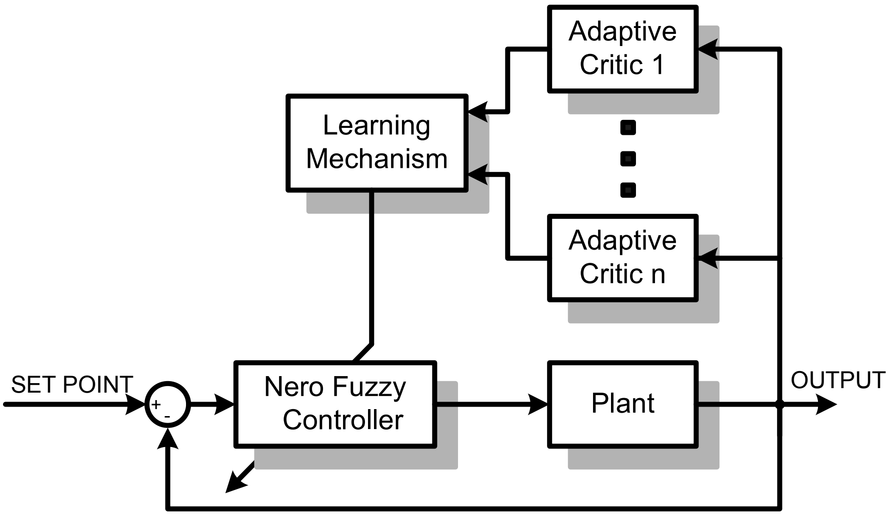


Fig 1. Block diagram of Multi agent adaptive critic based neuro-fuzzy controller

The Takagi-Sugeno fuzzy inference system is based on fuzzy rules of the following type:

$$\begin{aligned}&Rule_i: \quad If\ u_1 = A_{i1}\ And\ \ldots\ And\ u_p = A_{ip}\\ &then\ y = f_i(u_1, u_2, \ldots, u_p)\end{aligned} \quad \textbf{(1)}$$

Where $i = 1 \ldots M$ and M is the number of fuzzy rules. $u_1, \ldots, u_p$ are the inputs of network, each A_{ij} denotes the fuzzy set for input u_j in rule i and $f_i(.)$ is a crisp function which is defined as a linear combination of inputs in most applications

$$y = \omega_{i0} + \omega_{i1}u_1 + \omega_{i2}u_2 + \ldots + \omega_{ip}u_p \quad \textbf{(2)}$$

Thus, the output of this model can be calculated by:

$$y = \frac{\sum_{i=1}^{M} f_i(\underline{u})\mu_i(\underline{u})}{\sum_{i=1}^{M} \mu_i(\underline{u})}, \quad \mu_i(\underline{u}) = \prod_{j=1}^{p} \mu_{ij}(u_j) \quad \textbf{(3)}$$

A simple form of $f_i(u)$ can be as

$$f_i(u) = a_i u_1 + b_i u_2 + c_i \quad \textbf{(4)}$$

The output of controller is in the following form:

$$y = \frac{\sum_{i=1}^{n} \mu_i(a_i u_1 + b_i u_2 + c_i)}{\sum_{i=1}^{n} \mu_i} \quad \textbf{(5)}$$

Where n is number of controller fuzzy rules, μ_i is the firing strength of $i'th$ rules, u_1 is the first and u_2 is the second one for two input type controller (for example error and its derivative). In this paper we choose $u_1 = e$ and $u_2 = \dot{e}$. The neuro-fuzzy controller applied in this paper, is a standard Sugeno fuzzy controller composed of four layers. In the first layer, all inputs are mapped into the range of [-1, +1]. In the second layer, the fuzzification process is performed using Gaussian membership functions with five labels for each input. In layer 3, decision-making is done using Max-Product law and defuzzification is carried out in the fourth layer in order to calculate the proper control input using (5), a_i, b_i, c_i are parameters to be determined via learning mechanism.

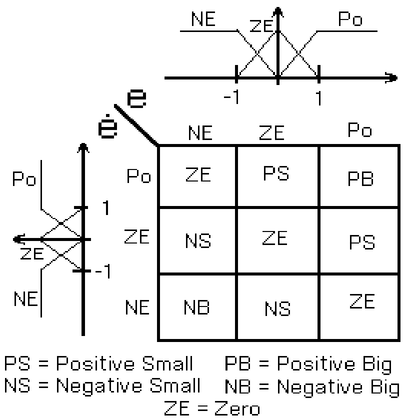
4. Adaptive Critic

The most important blocks in figure 1 are the emotional critics. Emotional Critic is the main part of any emotional learning system. The performance of the critic can be compared with the performance of emotional hue in humans. In absence of an exact evaluation of the present state in term of the objective value function, emotional cues like stress, satisfaction and etc. can be guide our control action into changing in the right direction so as to produce desired response. Similarly, the critic evaluates the state of system and generates a signal called emotion (r). In multi agent system, there are more than one critic that each of them evaluate the performance of system from their own point of view. For example in this paper, we proposed a multi agent controller with two critics.

One critic is satisfied by the low control error and another one is satisfied by low control cost and action. These reinforcemental signals are used to train and fine-tune the main controller. Basically, these critics act as intelligent guides for the controller. The learning mechanism will be adapted the controller parameters in order to satisfy all critics and reduce the stresses of them.

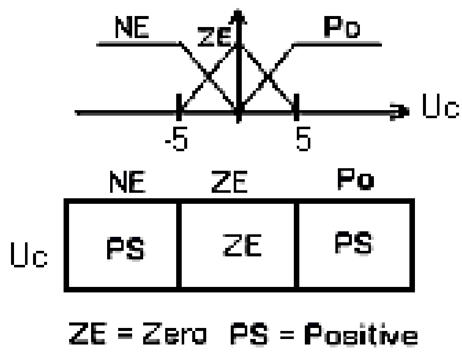
Both of these critics are define in fuzzy forms. Fuzzy systems are very useful for critic modeling because the critic just gives us an approximate evaluation of current states of system. The first critic satisfies when the control error reach to zero. For this plan, the inputs of critic are error of plant output from desired response and its derivative.

The output of critic is a signal between [-1, 1] and shows the performance of the system. If this signal becomes zero, it means that the critic is satisfied by the performance of controller. If the signal becomes larger, it shows the more stress and more dissatisfaction. The fuzzy sets and rules base of this critic is shown in figure 2. The second critic will be satisfied by the low control cost. If control action becomes larger, it causes more stress in the critic output. The output of this critic is a signal between [0, 1] that 0 indicate better controller performance. The fuzzy sets and rules base of this critic is shown in figure 3.



ė \ e	NE	ZE	Po
Po	ZE	PS	PB
ZE	NS	ZE	PS
NE	NB	NS	ZE

Fig. 2. Rule base and fuzzy sets of first critic



	NE	ZE	Po
Uc	PS	ZE	PS

Fig.3. Rule base and fuzzy sets of second critic

5. Learning Mechanism

The main objective of learning mechanism is to satisfy total emotion and reduces total stress. This aim can be extracted trough bellow energy function:

$$E = \frac{1}{2}\left(k_1 r_1^2 + k_2 r_2^2\right) = E_1 + E_2 \qquad \textbf{(6)}$$

By minimizing this energy function, we can reduce the total stress of the system and satisfy all critics. With applying Newton gradient decent method, the changes in weight must be followed by bellow general rule:

$$\Delta\omega_i = \Delta\omega_{i1} + \Delta\omega_{i2} = -\eta\frac{\partial E}{\Delta\omega_i} = -\eta\left(\frac{\partial E_1}{\Delta\omega_{i1}} + \frac{\partial E_2}{\Delta\omega_{i2}}\right) \qquad \textbf{(7)}$$

Where η is the learning rate of the corresponding neuro-fuzzy controller. In order to calculate the RHS of (7), the chain rule is used:

$$\Delta\omega_{i1} = -\eta\frac{\partial E_1}{\partial r_1}\cdot\frac{\partial r_1}{\partial y}\cdot\frac{\partial y}{\partial u}\cdot\frac{\partial u}{\partial \omega_i} \qquad \textbf{(8)}$$

From (6),

$$\frac{\partial E}{\partial r_1} = k_1 r_1 \qquad \textbf{(9)}$$

and also

$$e = y_{ref} - y \tag{10}$$

Then

$$\frac{\partial r_1}{\partial y} = -\frac{\partial r_1}{\partial e} \tag{11}$$

Since with the incrimination of error, r_1 will also be incremented and on the other hand, on-line calculation of $\frac{\partial r_1}{\partial e}$ is accompanied with measurement errors, thus producing unreliable results, we have:

$$\frac{\partial r_1}{\partial y} = -\frac{\partial r_1}{\partial e} = -\lambda \qquad (\lambda > 0) \tag{12}$$

That only the sign of it (-1) is used in our calculations. Also, $\frac{\partial y}{\partial u} = J$, where J is a Jacobian Matrix of the system. From (7) to (12), $\Delta\omega_{i1}$ will be calculated as follows:

$$\Delta\omega_{i1} = \eta k\ r_1 J \cdot \frac{\partial u}{\partial \omega_{i1}} \tag{13}$$

To follow up equation (6) to (13) $\Delta\omega_{i2}$ can be calculated as:

$$\Delta\omega_{i2} = \eta k_2 r_2 J \cdot \frac{\partial u}{\partial \omega_{i2}} \tag{14}$$

From equation (6), (13) and (14) $\Delta\omega_i$ is calculated as:

$$\Delta\omega_i = \Delta\omega_{i1} + \Delta\omega_{i2} = \eta(K_1 r_1 + K_2 r_2) \cdot \frac{\partial u}{\partial \omega_i} \tag{15}$$

Equation (15) is used for updating the learning parameters a_i's, b_i's and c_i's in (5), which is straightforward:

$$a_i(n+1) = a_i(n) + \eta\big(K_1 r_1(n) + K_2 r_2(n)\big)\frac{e(n)u_i}{\sum_{i=1}^{n} u_i} \tag{16}$$

$$b_i(n+1) = b_i(n) + \eta\big(K_1 r_1(n) + K_2 r_2(n)\big)\frac{\dot{e}(n)u_i}{\sum_{i=1}^{n} u_i} \tag{17}$$

$$c_i(n+1) = c_i(n) + \eta\big(K_1 r_1(n) + K_2 r_2(n)\big)\frac{u_i}{\sum_{i=1}^{n} u_i} \tag{18}$$

In above formula, k_1 and k_2 are importance coefficient of emotions and indicate that with critic is more important and η is the learning coefficient that must be select proper value in order to both increase the learning speed and prevent the learning instability. It is recommended that this value be assumed between [0, 10].

6. Simulation Results

The following simulation results illustrate the capabilities of proposed multi agent controller. In these simulations, we choose a nonlinear model of HVAC system which it has delayed behavior and is a multivariable, nonlinear non-minimum phase system, that its control is very difficult. The state space equations governing the model are as follows [9]:

$$\begin{aligned}
\dot{x}_1 &= 60u_1\alpha_1(x_3 - x_1) - 60u_1\alpha_1(W_s - x_2) + \alpha_3(Q_o - h_{fg}M_o) \\
\dot{x}_2 &= 60u_1\alpha_1(W_s - x_2) + \alpha_4 M_o \\
\dot{x}_3 &= 60u_1\beta_1(x_1 - x_3) - 15u_1\beta_1(T_o - x_1) - \\
&\quad 60u_1\beta_3(0.25W_o + 0.75x_2 - W_s) \\
y &= x_1
\end{aligned} \tag{19}$$

In which the parameters are:

$u_1 = f, u_2 = gpm, x_1 = T_3, x_2 = W_3, x_3 = T_2$
$\alpha_1 = 1/V_s, \alpha_2 = h_{fg}/C_pV_s, \alpha_3 = 1/\rho C_pV_s,$
$\alpha_4 = 1/\rho V_s, \beta_1 = 1/V_{he},$
$\beta_2 = 1/\rho C_pV_{he}, \beta_3 = h_w/C_pV_{he}$

And the numerical values are given in table 1.

Table1: Numerical Values for system parameters

$\rho = .074\ lb/ft^3$	$C_p = .24\ Btu/lb.°F$
$V_S = 58464\ ft^3$	$T_o = 85°F$
$M_O = 166.06\ lb/hr$	$V_{he} = 60.75\ ft^3$
$W_S = .007\ lb/lb$	$W_O = .0018\ lb/lb$
$Q_O = 289897$	$\tau = .008\ hr, k = 5$

In the first simulation, we compare the results that are taken by using importance factor K1=5 and K2=10 (Multi Agent) with Single Agent controller (K1=5, K2=0). The main controller started with random learning coefficient and learning rate (η) was selected 10 and learning mechanism updated the learning parameters 10 times in a second for each cases. The desired output and Plant output, Control effort and Error are shown in figure 4. As it shown, both of controllers can control and learn the control strategy without any instability and very fast. This on line and fast training are the most important key points of this method to other one. It is shown that multi agent controller can control the system with less control cost. The maximum of control effort for multi agent controller is 18 but in single agent controller is about 25. So the multi agent controller can be minimized both of error and control effort simultaneously.

In the next simulation, we choose k1=5 and k2=5 for multi agent controller. Indeed, we expected that the control effort became less than to the last simulation because the importance of second critic agent was increased. This simulation result is shown in fig 5. As we expected the control effort becomes less than (about 10) the last one but the error become greater.

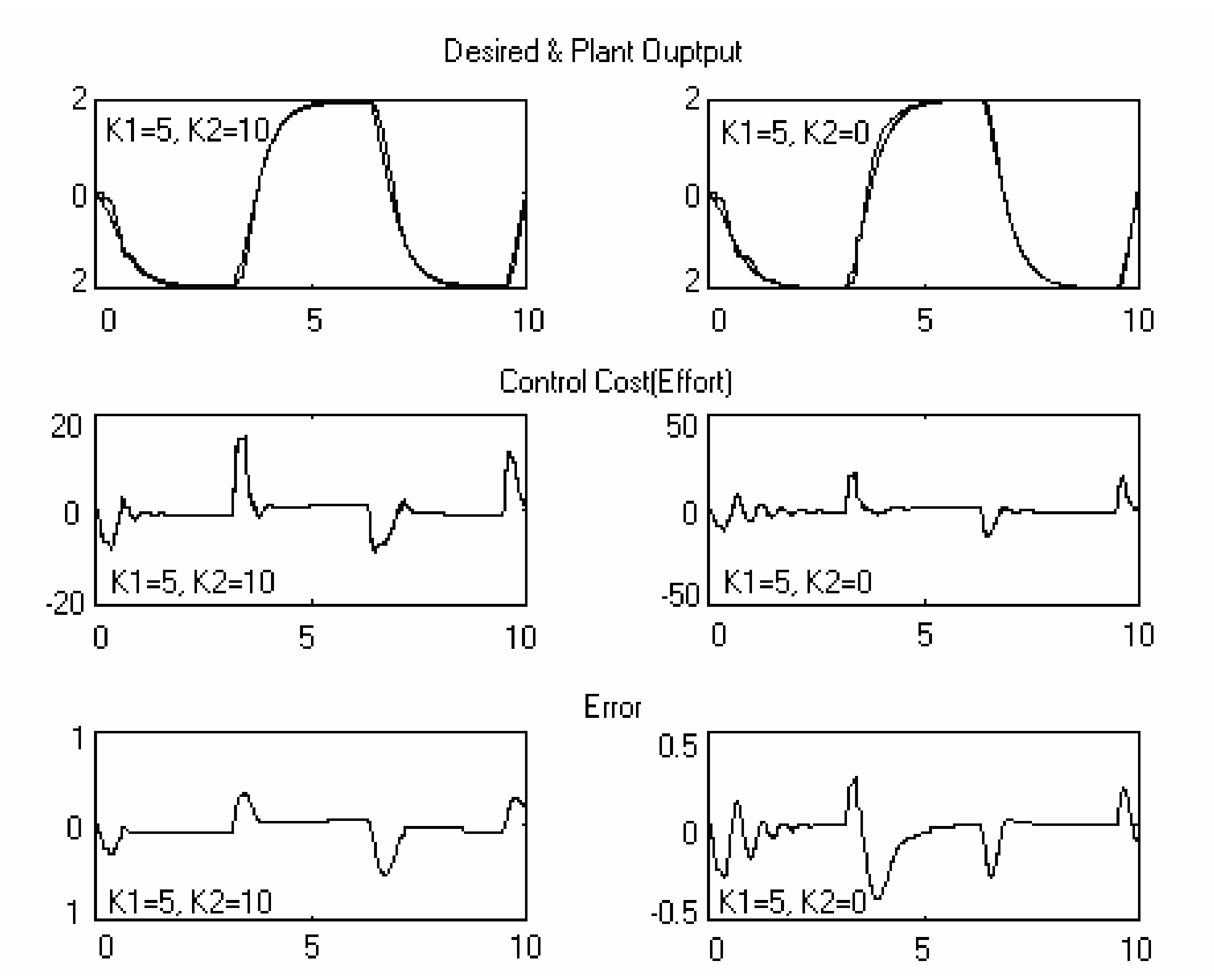


Fig 4 - Comparison of Multi agent controller (k1=5, k2=10) with single agent controller

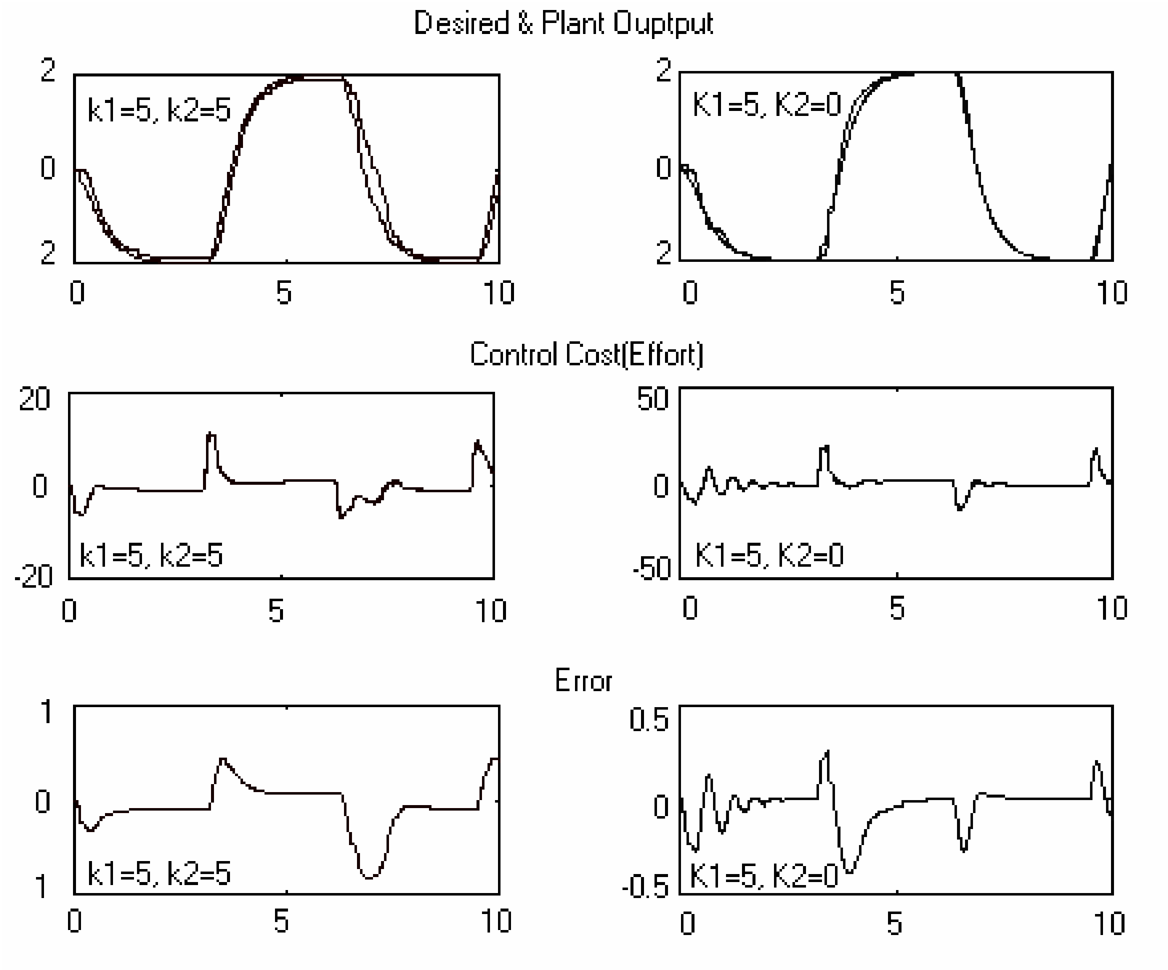


Fig 5 -Comparison of Multi agent controller (k1=5, k2=5) with single agent controller

Nomenclauture of HVAC System

h_w	Enthalpy of liquid water	h_{fg}	Enthalpy of water vapor
W_s	Humidity ratio of supply air	C_p	Specific heat of air
M_o	Moisture load	T_2	Temperature of supply air
V_s	Volume of thermal space	f	Volumetric flow rate of air
W_o	Humidity ratio of outdoor air	V_{he}	Volume of heat exchanger
W_3	Humidity ratio of thermal space	T_o	Temperature of outdoor air
Q_o	Sensible heat load	T_3	Temperature of thermal space
ρ	Air mass density	gpm	Flow rate of chilled water

7. Conclusion

In this paper, we developed a multi agent emotional controller that can achieve to multiple goals and aims. Online learning, fast convergence, and learning stability are the most important advantages of this controller that are shown in the simulation results. Other important advantage of this method is its robustness and relative independency to plant model that makes it more interesting for real application.

Reference

1. Kaiser, M., Dillman, R., Friedrich, H., Lin ,I., Wallner, F. and Weckesser, P., “Learning Coordination Skills in Multi-Agent Systems”, IEEE/RSJ Int. Conf. Intelligent robot and systems, Osaka, Japan, Nov. 1996.

2. Murciano, A., Millan, J. D. R. and Zamora J., “Specialization in multi-agent Systems through Learning”, Bio. Cyber., vol. 76, pp. 375-382, 1997.

3. Sutton, R. S. and Barto, A. G., Reinforcement Learning: An Introduction, MIT Press, 1998.

4. Lucas, C. and Jazbi, S. A., “Intelligent Motion Control of Electric Motors with Evaluative Feedback”, Cigre97, Cigre, France, 1997.

5. Wang, L. X., Adaptive Fuzzy Systems and Control, PTR Prentice-Hall, 1994.

6. Haykin, S., Neural Network, A compressive Foundation, Prentice-Hall Inc., 1994.

7. Jang, J. -S. R., Sun, C. -T. and Mizutani, E., Neuro-Fuzzy and Soft Computing, Prentice-Hall Inc., 1997.

8. Takagi, T. and Sugeno, M., “Derivation of fuzzy control rules from human operator’s control actions”, Proc. IFAC Symp. Fuzzy Information, Knowledge Representation and Decision Analysis, pp. 55-60, Jul. 1983.

9. Rashidi, F., Moshiri, B., “Design of a Robust Sliding Mode Fuzzy Controller for Nonlinear HVAC Systems”, LNCS, ISSN: 0302-9743, pp. 984-989