

Brain Emotional Learning Based Intelligent Controller via Temporal Difference Learning

Javad Abdi

*Department of Electrical and Computer Engineering,
NazarAbad Center, Islamic Azad University,
Karaj, Alborz, Iran
e-mail: javad.abdi@kiaiu.ac.ir*

Azam FamilKhalili

*Department of IT & Computer Engineering,
Amirkabir University of Technology
Tehran, Iran
e-mail: a.familkhalili@aut.ac.ir*

Modeling emotions has attracted much attention in recent years, both in cognitive psychology and design of artificial systems. Far from being a negative factor in decision-making, emotions have shown to be a strong faculty for making fast satisfying decisions. In this paper, we have adapted a computational model based on the limbic system in the mammalian brain for control engineering applications. Learning in this model based on Temporal Difference (TD) Learning. We applied the proposed controller (termed BELBIC) for a simple model of a submarine. The model was supposed to reach the desired depth underwater. Our results demonstrate excellent control action, disturbance handling and system parameter robustness for TDBELBIC. The proposal method, regarding the present conditions, the system action in the part and the controlling aims, can control the system in a way that these objectives are attained in the least amount of time and the best way.

Key Words: Artificial neural networks; Temporal difference; Brain Emotional Learning Based Intelligent Controller; heating, ventilating and air conditioning

I. INTRODUCTION

In these year, design of computational intelligent systems have received significant attentions. Control technique and application based on Artificial Neural Networks (ANNs) [1], Genetic Algorithms (GA) [2] and Fuzzy Inference System (FIS) [3] are among them. Emotional Learning is a psychologically motivated algorithm which is a family of intelligent algorithms [4].

Recently, biologically motivated intelligent computing has been successfully employed for solving different types of problems [5] [6]. Capability of learning is the greatest benefits of an intelligent system from a classical one. A common attribute of the learning process is the adaptation of the system parameters to better tackle the changing environment. One type of evaluation is based on emotional cues, which evaluate the impact of the external stimuli on the ability of the system both to function effectively in the short term and to maintain its long term prospects for survival [7]. Emotional learning is one of the learning strategies based on emotional evaluations. This learning process occurs in the brain Limbic system in mammalian brains [8].

Moren and Balkenius [9] presented a neurologically inspired computational model of the Amygdala and the Orbitofrontal Cortex in the Limbic System which is based on brains. A new model of control algorithm called Brain Emotional Learning Based Intelligent Controller

(BELBIC) has been suggested [10]. By direct and indirect approaches, two different approaches are developed.

Applying this controller for eliminating stator oscillations through fin placement was done in [11]. Application of BELBIC in Speed control of an interior permanent magnet synchronous motor was shown in [12] and in [13] a modified version of BELBIC was utilized in heating, ventilating and air conditioning (HVAC) control problem that is multivariable, nonlinear and non-minimum phase. In [14] other HVAC called micro-heat exchanger was the scope of BELBIC application. In [15], this controller used for controlling an identified model of a washing machine and in [16] this controller with multi objectives constraints was tuned for washing machine with evolutionary algorithms where it is possible to have a trade-off between energy consumption and other control objectives.

The BELBIC controller used for many application such as power system [17], active queue management [18], aerospace launch vehicle [19], interior permanent magnet synchronous motor system [20], flight simulation servo system [21], delayed systems [22] and other uncertain nonlinear systems [23].

Lucas has some reviews on successful control engineering and decision making applications, in which BEL has been used for satisfying action selection based on artificial emotions. He also demonstrated the capability of BELBIC including high levels of adaptability, disturbance rejection, and fault tolerance by implementing it with neural network [24].

As mentioned, in real time control and decision systems, Emotional Learning is a powerful methodology due to its simplicity, low computational complexity and fast training where the gradient based methods and evolutionary algorithms are hard to be applied because of their high computational complexity [25] [26].

Although this controller was applied in many simulation and real control tasks but the previous actions are not considered in these studies and the simulation and implementation were done in limited simulation or real time. In this article the TD learning algorithm used in BELBIC controller to improved version of this controller, and it is introduced and applied to control submarine plant.

The structure of the paper is as the following. In Section 2 the limbic model of mammalian brain and original model of BELBIC as an applied model is

presented. Section 3 present the control schema based on TD learning is demonstrated. In Section 4, the linear single input single output (SISO) system: submarine model and TD learning are presented, respectively. the implementation of the original BELBIC on control of Submarine Model is demonstrated in this section. The internal instability of controller in simulation plant model and real plant are shown, too. Finally in Section 6 the conclusion part is discussed.

II. AN APPLIED MODEL: STRUCTURE OF THE LIMBIC SYSTEM

Motivated by the success in functional modeling of emotions in control engineering applications [27] [28] [29], the main purpose of this research is to use a structural model based on the limbic system of mammalian brain, for decision making and control engineering applications.

The emotional processes is done by the Limbic System, as part of the mammalian creatures' brain. The Limbic System located in the cerebral cortex consists mainly of following components: Amygdala, Orbitofrontal Cortex, Thalamus, Sensory Cortex, Hypothalamus, Hippocampus and some other less important areas. Fig. 1 shows a simple graphical of limbic system in human brain [30].

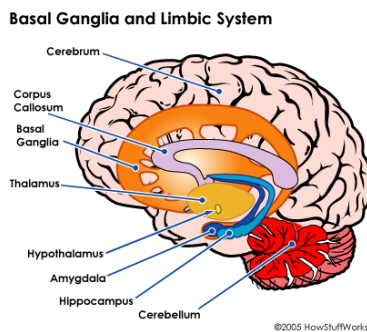


Fig. 1: The structure of the main components of Limbic System [30]

Amygdala which is a small almond-shaped in sub-cortical area is play role in the emotional system. This component is placed in a way to communicate with all other Sensory Cortices and areas within the Limbic System. For In this section the all parts of Limbic system are describe.

The studies show that a stimulus and its emotional consequences are associated in the Amygdala area [31]. A powerful area but small of the brain, the amygdala is located below the hypothalamus and helps the mammalian process the world at large. The amygdala is in charge of for regulating emotions, affecting our relationships and well-being. It also plays a part in causing arousal. Automatic reactions like fear are in large part caused by the amygdala. Amygdala are believed to mediate inherently emotionally charged

stimuli as well as coarsely resolved stimuli in general [32].

The Thalamus signal going to the Amygdala evades the processes involved in the Sensory Cortex and other components of the system. Therefore, Amygdala receives a non-optimal but fast stimulus from the Thalamus which among the input stimuli is often known as a characteristic signal [33].

Hippocampus is the next to the Amygdala and part of the forebrain, which is shaped somewhat like a seahorse. The hippocampus is key in creating new memories and helps mammalian with spatial orientation and sleep patterns.

The cerebral cortex is the largest, most apparent part of the brain. It is the outer layer of the brain that is the main source of human intelligence. The surface of the cortex is grey matter and has six different layers with many neural networks. Beneath these layers is white matter and when all of these are put together we are given a huge number of connections that facilitate our ability to think, feel, and reason.

The cerebral cortex has two hemispheres and each hemisphere helps to manage different things and perform various tasks. Both the hemispheres can communicate with one another, and can be divided into four different lobes.

The corpus callosum connects the brain's two hemispheres together. It is a huge bundle of nerve fibers that allows information to pass between the two parts of the brain. The corpus callosum allows for optimal performance from the brain.

As mentioned, there are two approaches to intelligent and cognitive control. In the indirect approach, the intelligent system is utilized for tuning the parameters of the controller. We have adopted the second, so called direct approach, where the intelligent system, in our case the computational model termed TDBELBIC, is used as the controller block. Fig. 2 present a structural engineering model of this System. TDBELBIC is essentially an action generation mechanism based on sensory inputs and emotional cues. In general, these can be vector valued, although in the benchmarks discussed in this paper for the sake of illustration, one sensory input and one emotional signal (stress) have been considered.

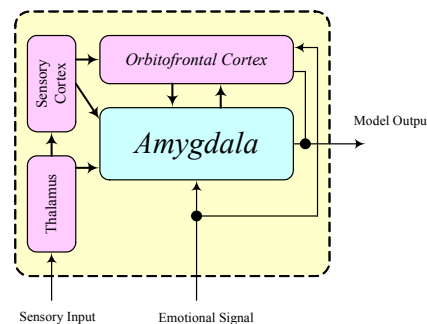


Fig. 2. The abstract structure of TDBELBIC

The emotional learning occurs mainly in amygdala. The learning rule of amygdala is given in formula (1).

$$\Delta G_a = k_1 \cdot \max(0, EC - A) \quad (1)$$

where G_a is the gain in amygdala connection, k_1 is the learning step in amygdala and EC and A are the values of emotional cue function and amygdala output at each time. The term \max in the formula (1) is for making the learning changes monotonic, implying that the amygdala gain can never be decreased. This rule is for modeling the incapability of unlearning the emotion signal (and consequently, emotional action), previously learned in the amygdala [34] [35]. Similarly, the learning rule in orbitofrontal cortex is shown in formula (2).

$$\Delta G_o = k_2 \cdot (MO - EC) \quad (2)$$

where G_o is the gain in orbitofrontal connection, k_2 is the learning step in orbitofrontal cortex and MO is the output of the whole model, where it can be calculated as formula (3):

$$MO = A - O \quad (3)$$

In fact, by receiving the sensory input S , the model calculates the internal signals of amygdala and orbitofrontal cortex by the relations in (4) and (5) and eventually yields the output.

$$A = G_a \cdot S \quad (4)$$

$$O = G_o \cdot S \quad (5)$$

Since amygdala does not have the capability to unlearn any emotional response that it ever learned, inhibition of any inappropriate response is the duty of orbitofrontal cortex.

III. IMPLEMENTATION

Controllers based on emotional learning have shown very good robustness and uncertainty handling properties [27] [29], while being simple and easily implementable. To utilize our version of the Moren-Balkenius model as a controller, we note that it essentially converts two sets of inputs into the decision signal as its output. We have implemented a closed loop configuration using this block (termed TDBELBIC) in the feed forward loop of the total system in an appropriate manner so that the input signals have the proper interpretations. The block implicitly implemented the critic, the learning algorithm and the action selection mechanism used in functional implementations of emotionally based (or generally reinforcement learning based) controllers, all at the same time [27] [28] [29]. The structure of the control circuit we implemented in our study is illustrated in Fig. 3. The functions we used in emotional cue and sensory input blocks are given in (6) and (7),

$$EC = W_1 \cdot e + W_2 \cdot CO \quad (6)$$

$$SI = W_3 \cdot PO + W_4 \cdot PO \quad (7)$$

where EC , CO , SI and PO are emotional cue, controller output, sensory input and plant output and the

W_1 through W_4 are the gains must tuned for designing a satisfactory controller.

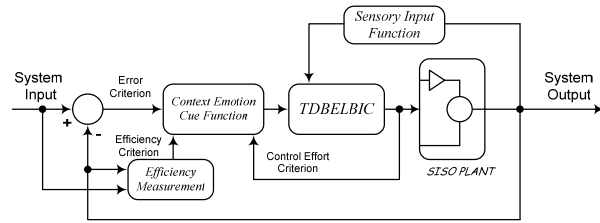


Fig. 3. Control system configuration using TDBELBIC

IV. SIMULATIONS

We confirmed the capability of TDBELBIC by performing some simulations. It must be mentioned that in the all simulations outlined below, we implemented the set-point control strategy with the desired value of 1. The descriptions of simulations are given below:

A LINEAR SISO SYSTEM: SUBMARINE MODEL

In this simulation, we considered a simple model of a submarine. The model was supposed to reach the desired depth underwater. The quantitative model is represented via (8).

$$G(s) = \frac{0.1(s+1)^2}{s(s^2+0.09)} = \frac{0.1s^2+0.2s+0.1}{s^3+0.09s} \quad (8)$$

We implemented the control circuits in MATLAB SIMULINK package. The output of the system with a simple feedback and the output of the system with a TDBELBIC controller are given in Fig. 4.

B TD(LAMBDA) LEARNING

Most of new learning algorithms like reinforcement learning, Q-learning and the method of temporal differences are characterized by their fast computation and in some cases lower error in comparison with the classical learning methods. Fast training is a notable consideration in some control applications. However, in prediction applications, two more desired characteristics of a good predictor are accuracy and low computational complexity.

In reinforcement learning, there is no teacher available to give the correct output for each training example, which is called unsupervised Learning. The output produced by the learning agent is fed to the environment and a scalar reinforcement value (reward) is returned. The learning agent tries to adjust itself to maximize the reward. [36] [37]

Often that the actions taken by the learning agent to produce an output will affect not only the immediate reward but also the subsequent ones. In this case, the immediate reward only reflects partial information about the action. It is called delayed-reward. [37] [38]

TD learning is a type of reinforcement learning for solving delayed-reward prediction problems. Unlike

supervised learning, which measures error between each prediction and target, TD uses the difference of two successive predictions to learn that is Multi Step Prediction. The advantage of TD learning is that it can update weights incrementally and converge to a solution faster [39].

In a delay-reward prediction problem, the observation-outcome sequence has the form x_1, x_2, \dots, x_m, z where each x_t is an observation vector available at time $t, 1 \leq t \leq m$ and z is the outcome of the sequence. For each observation, the learning agent makes a prediction of z , forming a sequence: P_1, P_2, \dots, P_m .

Assuming the learning agent is an artificial neural network, update for a weight w of the network with the classical gradient descent update rule for supervised learning is:

$$\Delta w = -\alpha \nabla_w E = -\alpha \sum_{t=1}^m (P_t - z) \nabla_w P_t \quad (9)$$

Where α is the learning rate and $\nabla_w E$ is the gradient vector, $\partial E / \partial w$ of the mean square error function:

$$E = \frac{1}{2} \sum_{t=1}^m (P_t - z)^2 \quad (10)$$

In [38], Sutton derived the incremental updating rule for equation (9):

$$\Delta w_t = \alpha (P_{t+1} - P_t) \sum_{k=1}^t \nabla_w P_k \quad (11)$$

For $t=1,2,\dots,m$ where $P_{m+1} \xrightarrow{def} z$

To emphasize more recent predictions, an exponential factor λ is multiplied to the gradient term:

$$\Delta w_t = \alpha (P_{t+1} - P_t) \sum_{k=1}^t \lambda^{t-k} \nabla_w P_k \quad (12)$$

where $0 \leq \lambda \leq 1$

This results in a family of learning rules, TD(Lambda), with constant values of λ .

But there are 2 special cases:

First, when $\lambda=1$, Eq. (12) falls back to Eq. (11), which produces the same training result as the supervised learning in Eq. (9). Second, when $\lambda=0$, since $0^0 = 1$, Eq. (12) becomes

$$\Delta w_t = \alpha (P_{t+1} - P_t) \nabla_w P_t \quad (13)$$

I can extended the Eq. (13) for BELBIC and made Eq. (14) for TDBELBIC.

$$\Delta G_{O_t} = \alpha (z - P_t) \nabla G_{O_t} P_t \quad (14)$$

Which has a similar form as Eq. (9). So the same training algorithm for supervised learning can be used for TD(0).

V. CONCLUSION

In Fig. 4, you can observe the results of simulating the diagram block Fig. 3. The results, based on TD learning, are compared to Orbitofrontal Cortex learning in a shared TDBELBIC structure. The outcomes suggest that TD based learning is faster than Orbitofrontal Cortex learning. But faster learning is increased for maximum overshoot. Both of the learning is incremental, however,

their memory output signals are presented in figures 5, 6. The increase rates represent their learning speed.

Paying attention to the achievements in the emotional controls founded a computational model, based on the Limbic system, for mammals' brain via time series learning. The paper tried to develop this method for answering more complicated issues and achieving difficult goals.

To do this, the ability of the learning module the emotional controller, was increased achieving based a brain computational model means of TD learning for credit assignment. TD learning, has easier computations because of using it's own experience. The methods resemble human behavioral learning.

APPENDIX: BODY OF BELBIC ALGORITHM

```
function E=belbic(x)
global OLD
A = OLD(1);
O = OLD(2);
V = OLD(3);
W = OLD(4);

S=x(1);

Rew=x(2);

%-----initial learning( alpha & beta )
alpha=0.1;

beta=0.4;
%-----
DV=alpha*max(0,Rew-A)*S;
y=A-O-Rew;
DW=beta*y*S;
V=V+DV;
W=max(0,W+DW);
A=S*V;
O=S*W;
OLD = [ A ; O ; V ; W ];
E=A-O;
```

REFERENCES

K.S. Narendra, and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Trans. on Neural Networks*, vol. 1, no. 1, pp. 4-27, 1990.

R.J. Wai, J.D. Lee, and K. H.Su, "Supervisory enhanced genetic algorithm control for indirect field-oriented induction motor drive," *Proceedings IEEE International Joint Conference on Neural Networks*, vol. 2, pp. 1239-1244, 2004.

T.M. Takagi, and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 15, no. 1, pp. 122-132, 1985.

H. Zhang, S. Liu, and S.Y. Yang, "An emotion-based behavior coordination approach to autonomous mobile robots," *Proceedings of the International Conference on Sensing, Computing and Automation*, vol. 1, pp. 249-254, 2006.

C. Lucas, D. Shahmirzadi, and M. Biglarbegian, "Co-evolutionary Approach to Graph-Coloring Problem," *Technical Journal of Amirkabir University of Technology*, vol. 14, nO.54, pp. 363-369, 2003.

- M. Sipper, and M. Tomassini, "Convergence to Uniformity in A Cellular Automaton via Local Co-evolution," *International Journal of Modern Physics*, vol. 8, no. 5, pp. 1013-1024, 1997.
- D. Shahmirzadi, *Computational Modeling Of The Brain Limbic System And Its Application In Control Engineering*, Master dissertation, Texas A&M University, U.S.A., 2005.
- S. Maren, "Long-Term Potentiation in the Amygdala: A Mechanism for Emotional Learning and Memory," *Trends in Neurosciences*, vol. 22, no. 12, pp. 561-567, 1999.
- J. Moren, and C. Balkenius, "A Computational Model of Emotional Learning in the Amygdala," *Cybernetics and Systems*, vol. 32, no. 6, pp. 611-636, 2000.
- C. Lucas, D. Shahmirzadi, and N. Sheikholeslami, "Introducing BELBIC: Brain Emotional Learning Based Intelligent Controller," *International Journal of Intelligent Automation and Soft Computing*, vol. 10, no.1, pp. 11-22, 2004.
- C. Lucas, D. Shahmirzadi, and H. Ghafoorifard, "Eliminating Stator Oscillations through Fin Placement," *Journal of Engineering Simulation*, vol. 3, no. 1, pp 3-7, March 2002.
- R.M. Milasi, C. Lucas, and B.N. Araabi, "Speed Control of an Interior Permanent Magnet Synchronous Motor Using BELBIC," *5th International Symposium on Intelligent Automation and Control, World Automation Congress, WAC 2004*, Seville, Spain, June 28-July 1, ISIAAC 116, 2004.
- N. Sheikholeslami, D. Shahmirzadi, E. Semsar, C. Lucas, and M. J. Yazdanpanah, "Applying Brain Emotional Learning Algorithm for Multivariable Control of HVAC Systems," *International Journal of Intelligent and Fuzzy Systems*, vol. 17, no. 1, pp. 35-46, 2006.
- H. Rouhani, M. Jalili, B.N. Araabi, W. Eppler, and C. Lucas, "Brain emotional learning based intelligent controller applied to neuro-fuzzy model of micro-heat exchanger," *Expert Systems with Applications*, vol. 32, pp. 911-924, 2007.
- R.M. Milasi, C. Lucas, and B.N. Araabi, "Intelligent Modeling and Control of Washing Machine Using Locally Linear Neuro- Fuzzy (LLNF) Modeling and Modified Brain Emotional Learning Based Intelligent Controller (BELBIC)," *Asian Journal of Control*, vol. 8, no. 4, 2005.
- R.M. Milasi, M.R. Jamali, and C. Lucas, "Intelligent Washing Machine: A Bioinspired and Multi-Objective Approach," *International Journal of Control, Automation, and Systems*, vol. 5, no. 4, pp. 436-443, August 2007.
- R. Mohammadi, C. Lucas, and B.N. Araabi, "A novel controller for a power system based BELBIC," *Proceedings of World Automation Congress*, vol. 18, 2004, pp. 409-420, 2001.
- M.J. Kharaajoo, "Application of brain emotional learning based intelligent controller (BELBIC) to active queue management," *International Conference on Computational Science, Lecture Notes in Computer Science*, vol. 3037, pp. 662-665, 2004.
- A.R. Mehrabian, C. Lucas, and J. Roshanian, "Aerospace launch vehicle control: An intelligent adaptive approach," *Aerospace Science and Technology*, vol. 10, no. 2, pp. 149-155, 2006.
- R. Mohammadi, C. Lucas, B.N. Arraabi, T.S. Radwan, and M.A. Rahman, "Implementation of emotional controller for interior permanent magnet synchronous motor drive," *Industry Applications Conference*, vol. 4, pp. 1767-1774, 2006.
- Z. Ziyang, W. Daobo, and W. Zhisheng, "Flight simulation servo system based on learning based intelligent controller," *Proceedings of the Fourth International Conference on Impulse and Hybrid Dynamic Systems*, pp. 1372-1375, 2007.
- A. Torabi, M. Boroushaki, and G. Vosoughi, "Design a brain emotional learning based intelligent controller (BELBIC) in delayed systems," *14th- International conference of Iranian Society of Mechanical Engineering (ISME2006)*, Isfahan, Iran, 2006.
- A.R. Mehrabian, and C. Lucas, "Emotional Learning Based Intelligent Robust Adaptive Controller for Stable Uncertain Nonlinear Systems," *International Journal of Intelligent Technology (IJIT)*, vol. 1, no. 1, pp. 34-40, 2005.
- C. Lucas, "BELBIC and Its Industrial Applications: Towards Embedded Neuro-emotional Control Codesign," *Chapter Book, Integrated Systems Design and Technology 2010*, M. Fathi, A. Holland (Ed.), F. Ansari and C. Weber (Co-Eds.), vol. 3, pp. 203-214, 2011.
- R.W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1175-1191, 2001.
- M. Fatourech, C. Lucas, and A.K. Sedigh, "Reducing Control Effort by means of Emotional Learning," *Proc. of 9th Iranian Conf. on Electrical Engineering, ICEE'01*, Tehran, Iran, vol. 41, pp. 1-8, 2001.
- C. Balkenius, and J. Moren, "A Computational Model Of Emotional Conditioning In The Brain," *Proceeding of the workshop on Grounding Emotions in Adaptive Systems*, Zurich, 1998.
- J.R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System," *IEEE Transaction on Systems, Man and Cybernetics*, vol. 23, no. 3, 1993.
- A.A. Savinov, "An Algorithm for Induction of Possibilities Set-Valued Rules by Finding Prime Disjunctions," *4th On-line World Conference on Soft Computing in Industrial Applications (WSCA)*, September 21-30, 1999.
- D. Purves, G.J. Augustine, D. Fitzpatrick, L.C. Katz, A. LaMantia, J.O. McNamara, and S.M. Williams, *Neuroscience*, Sinauer Associates, 2001.
- J.E. LeDoux, *In Search of an Emotional System in the Brain: Leaping from Fear to Emotion and Consciousness*. Cambridge, MA: MIT Press, 1995.
- A. Ohman, and S. Mineka, "Fears, Phobias, and Preparedness: Toward and Evolved Module of Fear and Fear Learning," *Journal of Psychological Review*, vol. 108, no. 3, pp. 483-522, 2001.
- J.P. Kelly, *The Neural Basis of Perception and Movement*, Principles of Neural Science, London. UK: Prentice Hall, 1991.
- M.H. Voigt, "Fuzzy evolutionary algorithms" *Technical Report 92-038, International Computer Science Institute (ICSI), 1947 Center Street, Suite 600, Berkeley, CA, 94704*, June 1992.
- J. Chen, E. Antipov, B. Lemieux, W. Cedeno, and D.H. Wood, "DNA Computing Implementing Genetic Algorithms," *Preliminary Proceedings DIMACS Workshop On Evolution As Computation*, pp: 39-49, DIMACS, Piscataway NJ, January 1999.
- J. Abdi, C. Lucas, A.F. Khalili, C. Lucas, and A.K. Sedigh, "The Role of Truncated Temporal Difference Learning (TTDL) with Functional Approximators in Reduction of Time and Computational Costs of Reinforcement Learning," *Journal of Iranian Association of Electrical and Electronics Engineers*, vol. 1, no. 1, P.P. 4-19, Spring 2004.
- D. Eby, R.C. Averill, William F. Punch III, and E.D. Goodman, "Evaluation of Injection Island GA Performance on Flywheel Design Optimization," *Proceedings, Third Conference on Adaptive Computing in Design and Manufacturing, Plymouth, England*, Springer Verlag, April 1998.
- R.S. Sutton, "Learning to Predict by the Methods of Temporal Differences," *Machine Learning*, vol. 3, pp. 9-44, 1988.
- R.S. Sutton, "Generalization in Reinforcement Learning: Successful Examples Using Sparse Coarse Coding," *To Appear in Advances in Neural Information Processing Systems 8*, 1995.

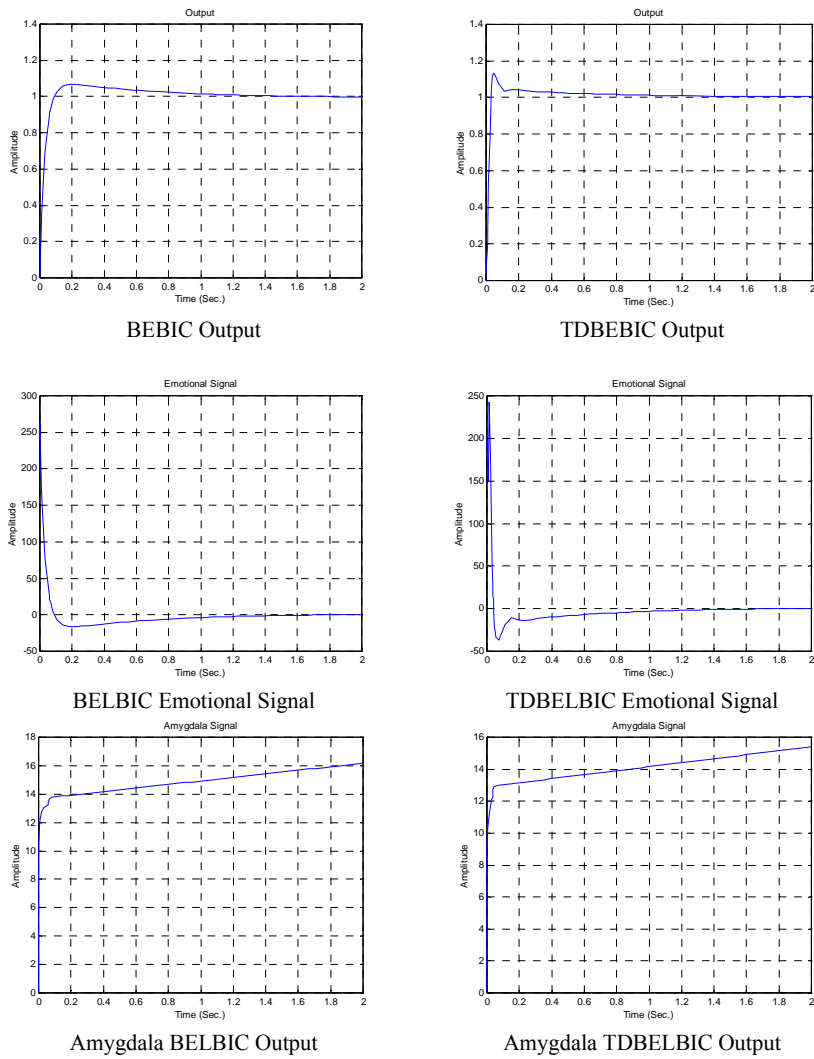


Fig. 5. Comparison of BELBIC and OFC Learning with BELBIC and TD Learning

(i)

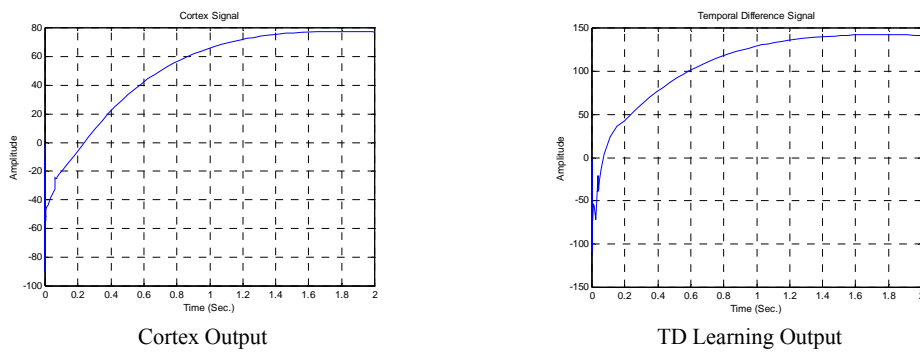


Fig. 6. Comparison of BELBIC and with TDBELBIC Memory