Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

A T-S Fuzzy based Adaptive Critic for Continuous-time Input Affine Nonlinear Systems

Prem Kumar P.*, Laxmidhar Behera*†, Nazmul H. Siddique† and Girijesh Prasad †
*Department of Electrical Engineering
Indian Institute of Technology Kanpur, Kanpur, India
†School of Computing and Intelligent systems
University of Ulster, UK

***Abstract*—This paper proposes a novel scheme of a Takagi-Sugeno (T-S) fuzzy based adaptive critic for the optimal control of the continuous-time input affine nonlinear system. A novel learning strategy is proposed to update the weights of critic network which resolves the issue of under-determined weight update equations discussed in [1]. The T-S Fuzzy based critic network approximates the global optimal cost as fuzzy average of local costs associated with local linear subsystems. *This work clearly demonstrates that the optimal cost of a nonlinear system can be represented as the fuzzy cluster of optimal costs of locally valid linear models in a T-S framework*. The proposed scheme has been simulated for four different dynamic systems. Simulation results clearly demonstrate that the T-S Fuzzy approximates the optimal cost, with subsystems in each fuzzy zone represents the optimal cost of locally valid linear model.**

***Index Terms*—Approximate dynamic programming, Continuous-time nonlinear system, T-S Fuzzy based Critic**

I. INTRODUCTION

Given a nonlinear dynamical system with performance cost index, the optimal control problem can be solved using Pontryagin's minimum principle and Bellman's Dynamic programming [2]–[4]. Dynamic programming gives a comprehensive computational technique based on the principle of optimality and gives an optimal controller in state feedback form. Though a feedback form is robust to noise and model uncertainties, the associated Hamilton-Jacobi-Bellman equation demands expensive computation and storage and is an off-line process where the problem is solved from the end point and is approached in backward direction.

Werbos [5] proposed Approximate Dynamic Programming (ADP) to overcome these issues, which solved the dynamic programming in forward direction and hence implemented online. ADP uses dual neural network architecture known as Adaptive Critic (AC) to solve the dynamic programming in forward direction. In general Adaptive Critic has two networks namely action network and critic network. The action network represents the relationship between state and input while the critic network represents the relationship between state and costate vector. The critic evaluates the performance of actor and the actor is improved, based on the feedback from the critic network. The readers may refer [6], [7] to know about various architectures of ADP. The stability proof of adaptive critic is derived in [8] and is shown that it converges to optimal values for linear systems.

Padhi et al. [9] proposed a single network adaptive critic (SNAC) for discrete-time system which works for optimal control problem where the optimal control input can be explicitly expressed in terms of state and costate variables. Typically, dynamic systems in input affine form with quadratic cost function can be effectively controlled using SNAC and is shown that the network converges to optimal values for linear systems. Swagat et al. [1] proposed single network adaptive critic for continuous-time systems and discussed the constraints in network architecture to achieve stability. In [1], the cost-to-go cost is approximated with a critic network in contrast to [9], where the costate dynamics is learnt directly with the critic network. The major limitation in [1], is that a scalar equation is derived for weight update, which results in an under-determined system and a suitable network architecture is not identified. In general, the existing adaptive critic methodologies employ a multi-layered perceptron as the critic network and the learned weights do not convey any meaningful insight into the optimality. Though the adaptive critic is successfully implemented in real-life problems [10]–[12], the network architectures which could give much information for control engineers is not analyzed much in the literature.

The following major issues are addressed in this work:

- Discussing about a network architecture which would give meaningful insight about optimality for control engineers.
- Developing learning methodology for continuous-time systems with a single network which would not result in under-determined system as discussed in [1].

This paper proposes a T-S fuzzy system based adaptive critic for continuous-time systems which addresses both of the above issues. The major motivation of this work comes from the fact that in a small operating region, any nonlinear system behaves like a linear system and the analytic expression for the optimal cost for linear system is well-known. Hence, *it is intuitive to express the optimal cost of a nonlinear system in terms of optimal costs associated with local linear models. Thus the critic network that predicts cost-to-go function is modelled as the T-S fuzzy network where each fuzzy zone is associated with a local cost function associated with the locally valid linear system. Thus the complex optimal cost in this paper is obtained as the fuzzy-average of local linear optimal cost functions.* With such an architecture, the local model of the T-S fuzzy system would represent the optimal weights in each zone of operation. In addition, a novel learning methodology is described which combines [9] and [1], considering the fact that the weight update is implemented as discrete steps in real-time. With such a modification, the weight update equation is not under-determined. The extensive simulations are performed which shows that the proposed T-S fuzzy based architecture learns the global nonlinear optimal cost as fuzzy-cluster of local linear optimal costs.

The rest of the paper is organized as follows. The proposed scheme is discussed in next section. The simulation results with four benchmark systems are presented in section III and finally concluded in section IV.

II. CONTINUOUS-TIME ADAPTIVE CRITIC

Consider a continuous-time nonlinear affine system given by

$$\dot{x} = f(x) + g(x)u \tag{1}$$

The task is to find a control input u^* which stabilizes the system, while minimizing the quadratic cost function

$$\begin{aligned} J &= \int_0^\infty (x^T Qx + u^T Ru)dt \\ &= \int_0^\infty \psi(x,u)dt \end{aligned} \tag{2}$$

where $\psi(x,u)$ is the utility function, and in general, Q is taken as a constant positive semi-definite matrix and R is a positive definite matrix. The reader should note that we are focussing on optimal control problem with boundary conditions $t_f = \infty$ and $x_f = 0$.

It is well known that the Hamiltonian of the above optimal control problem is given by,

$$H(x,\lambda^*,u) = \psi(x,u) + \lambda^{*T}[f(x) + g(x)u] \tag{3}$$

where, $\lambda^* = \frac{\partial J^*}{\partial x}$ is the costate vector along the optimal path and J^* is the optimal cost.

The optimal control law satisfies the necessary condition of extremum given by,

$$\frac{\partial H}{\partial u} = \frac{\partial \psi}{\partial u} + \lambda^{*T}\frac{\partial}{\partial u}[f(x) + g(x)u] = 0 \tag{4}$$

which results in the optimal controller,

$$\begin{aligned} u^* &= -R^{-1}g^T(x)\frac{\partial J^*}{\partial x} \\ &= -R^{-1}g^T(x)\lambda^* \end{aligned} \tag{5}$$

We know that the optimal cost satisfies the Hamilton-Jacobi-Bellman (HJB) equation:

$$\frac{\partial J^*}{\partial t} + \min_u H(x,\lambda^*,u) = 0 \tag{6}$$

The above HJB equation gives the solution to the optimal control problem for any class of system. However, the analytical solution to the HJB equation is difficult to obtain and it depends on optimal cost J^*. Hence we are motivated to approximate the optimal cost function through a suitable function approximator.

We have selected the function approximator based on the fact that in a small zone of operation, every nonlinear system behaves like a linear system. It is well known that for a linear time invariant (LTI) system with quadratic cost function (2), the optimal cost is given by,

$$J^* = \frac{1}{2}x^T Px \tag{7}$$

where, P is computed from algebraic Riccati equation (ARE). Since, nonlinear system behaves like a linear system in a small region, it is reasonably true that a quadratic cost would be a valid optimal function in a small zone of operation. We argue that the optimal cost in a global operating zone will vary from the local linear optimal cost function and can be represented by a cluster of optimal cost functions of linear subsystems. Hence, we are motivated to approximate the optimal cost function of a nonlinear system as a cluster of quadratic costs which are valid in local zone of operation. With this motivation, T-S Fuzzy system is chosen as our optimal cost approximator.

T-S fuzzy model is particularly chosen because, it approximates a complex nonlinear function as a fuzzy cluster of locally valid linear subsystems. The T-S fuzzy model for nonlinear quadratic optimal cost function is given by,

$$\text{IF } x_1(t) \text{ is } F_1^i \text{ AND } \cdots \text{ AND } x_n(t) \text{ is } F_n^i \text{ THEN}$$
$$\hat{V}_i(x) = x^T W_i x$$

where F_j^i, $j = 1,2,\cdots,n$, is the j^{th} fuzzy set of the i^{th} rule. Let

$$\mu_i = \prod_{j=1}^n \mu_i^j(x_j) \tag{8}$$

where $\mu_i^j(x_j)$ is the membership function of the fuzzy set F_j^i, $i = 1,2,\cdots,m$ and $\hat{V}_i(x)$ represents the cost in i^{th} zone.

Given the current state vector $x(t)$, the fuzzy model around this operating point is constructed as the weighted average of the local models and has the form

$$V(x) = \frac{\sum_{i=0}^m \mu_i x^T W_i x}{\sum_{i=0}^N \mu_i} \tag{9}$$

where μ_i is the membership in i^{th} zone and $V(x)$ is the optimal cost to stabilize from the state $x(t)$.

Since, the T-S fuzzy model approximates the cost-to-go function, it must satisfy the HJB equation, i.e.,

$$\frac{\partial V(x,t)}{\partial t} + \min_u H(x,\lambda^*,u) = 0 \tag{10}$$

Substituting, (3) in (10), we have,

$$\begin{aligned} \frac{\partial V(x,t)}{\partial t} + \psi(x,u) + \lambda^{*T}[f(x) + g(x)u] &= 0 \\ \frac{\partial V(x,t)}{\partial t} + \psi(x,u) + \left(\frac{\partial V}{\partial x}\right)^T [f(x) + g(x)u] &= 0 \end{aligned} \tag{11}$$

The above equation can be written as,

$$\dot{V}(x,t) = -\psi(x,u) \tag{12}$$

where,

$$\dot{V}(x,t) = \frac{\partial V(x,t)}{\partial t} + \left(\frac{\partial V}{\partial x}\right)^T \dot{x}$$
$$= \frac{\partial V(x,t)}{\partial t} + \left(\frac{\partial V}{\partial x}\right)^T [f(x) + g(x)u] \quad (13)$$

The above equation gives the expression for change in the optimal cost function as the system moves along the optimal path. The optimal weights of T-S Fuzzy model would satisfy (12), where $\frac{\partial V(x,t)}{\partial t}$ represents the explicit dependency of cost function on time, which occurs due to the evolution of weights during training. Hence, a suitable weight update law is derived to satisfy (12), i.e., the weights have to evolve such that along the trajectory (12) is satisfied.

The weight update will be performed in discrete steps in practice and such discrete time assumptions are always valid with high sampling rate. Hence, further discussions are continued in discrete steps of sampling time ΔT. The discrete time form of the above equation can be written as,

$$\Delta V(x(k)) = \psi(x(k), u(k))\Delta T \quad (14)$$

The above equation can be rewritten as,

$$V(x(k)) = V(x(k+1)) + \psi(x(k), u(k))\Delta T \quad (15)$$

where $\Delta V(x(k)) = V(x(k+1)) - V(x(k))$. The forward differentiation is considered since, we know from dynamic programming that the cost-to-go function at the k^{th} instant depends on $(k+1)^{th}$ instant, i.e.,

$$J^*(x(k), u^*(k)) = \psi(x(k), u^*(k)) + J^*(x(k+1), u^*(k+1)) \quad (16)$$

At any instant, critic network weights would satisfy (15). The weights are updated such that above equation is satisfied. At any instant k, we have weight vector $W(k)$. Then, along optimal path, we predict the optimal cost as,

$$V^d(W(k), x(k)) = V(W(k), x(k+1)) + \psi(x(k), u(k)) \quad (17)$$

The above equation is analogous to the costate vector dynamics obtained in [9], where costate vector at instant k depends on the costate vector on instant $k+1$. With such analogy, the weight update law is further derived as follows.

Then, weights are updated to minimize $\| V^d(W(k), x(k)) - V(W(k), x(k)) \|$. The reader should note that, with such discrete update, the weight update is not under-determined as discussed in [1].

After learning the optimal cost $J^* = V$, the costate vector $\frac{\partial J^*}{\partial x}$, can be obtained from the network to compute the optimal control input. The control scheme is shown in figure 1. The current state $x(k)$ is the input to the critic network and it learns the cost-to-go function. As the learning progresses, the cost-to-go function approaches the optimal cost J^*. Hence, V is considered as the approximate of J^* and the costate vector λ is computed from the network and is used to calculate the optimal control law.

The network is trained as follows:

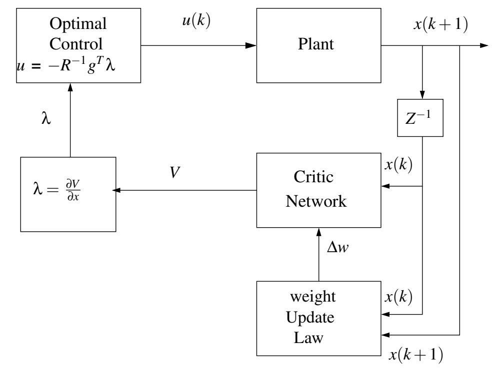


Fig. 1. Control Scheme with continuous-time Adaptive Critic

1) Initialize the states to a random initial state x_0 in the operating zone.
2) Compute the control input $u(t) = u(k)$ using (5), from the T-S Fuzzy model (9) .
3) Give the input $u(k)$ and obtain the next state $x(k+1)$.
4) Compute $\psi(x(k), u(k))$, $V(x(k+1))$ and $V(x(k))$, with current instant weight vector $W(k)$.
5) Compute $V^d(W(k), x(k))$ from $V(W(k), x(k+1))$ and $\psi(x(k), u(k))$ using (17).
6) Update the weights to minimize, $\| V^d(W(k), x(k)) - V(W(k), x(k)) \|$.
7) Repeat from step 2 for a fixed number of steps, say N for the initial state x_0.
8) Check for weight convergence. If it is not converged repeat from step 1.

Since, the T-S fuzzy model represents the optimal cost, it would be a positive definite function. Hence, the weight has to be properly initialized such that the T-S fuzzy model is positive definite from the beginning of training. In our approach, we initialized the weights in each zone to the weights corresponding to the optimal cost function of linear model around the origin, which can be easily obtained from algebraic Riccati equation.

III. Simulation Results

The proposed continuous-time adaptive critic is tested on four systems. At first, a second order LTI system is taken to show that the proposed continuous-time adaptive critic cost function is actually converging to the optimal cost of the LTI system. Then, a first order nonlinear system whose optimal control law is already known is chosen, and the optimal cost is approximated with T-S fuzzy model based critic network. It is shown through simulation that the weights in each zone finally settles to a value, which closely approximates the optimal cost. Then, the control scheme is tested on two benchmark systems and performance is compared with LQR obtained around the origin since the optimal cost is not known to us.

A. Linear Time Invariant System

Consider a LTI system [1] with dynamics,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0.4 & 0.1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \tag{18}$$

The task is to find the control u which minimizes the cost,

$$J = \frac{1}{2}\int_0^\infty [x^T Qx + u^T Ru]dt \tag{19}$$

where,

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } R = 1$$

The optimal quadratic cost of the system is obtained from ARE as,

$$P = \begin{bmatrix} 2.10456 & 1.4722 \\ 1.4722 & 2.09112 \end{bmatrix} \tag{20}$$

The critic network is considered as,

$$V = \frac{1}{2}(w_1 x_1^2 + w_2 x_2^2 + w_3 x_1 x_2) \tag{21}$$

The critic network is trained with 200 random initial states with sampling instant $0.01sec$ and $N = 200$. The weight evolution during the training phase is shown in figure 2.

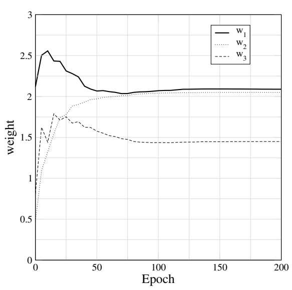


Fig. 2. Weights Evolution during training phase

The weights of the network finally converges to $w_1 = 2.0889$, $w_2 = 2.051$ and $w_3 = 1.449$. It is clear that the proposed weight update scheme ensures that the network finally converges to the optimal cost for linear systems.

B. First order Nonlinear System

Consider the first order nonlinear system dynamics given by,

$$\dot{x} = -x^3 + u \tag{22}$$

where x is the system state and u is the input to the system. The objective is to stabilize the system such that the input would minimize the cost function,

$$J = \frac{1}{2}\int_0^\infty (x^2 + u^2)dt \tag{23}$$

The analytic expression of the optimal controller is known for this system and is given by

$$u^* = x^3 - \sqrt{x^6 + x^2} \tag{24}$$

Given this optimal controller (24), the optimal cost can be computed from equation (23). Thus the cost associated with the proposed continuous-time adaptive critic based optimal control can be compared easily with the actual optimal cost. This is the main objective of this simulation example. The operating zone is considered as $(-1, 1)$ and the optimal cost function is approximated using 9 equally spaced fuzzy zones. The system is simulated with a sampling time of $0.1sec$ for a duration of $10sec$. The critic is trained with 20000 random initial states in the operating zone and the evolution of weights at different zone during training is shown in figure 3.

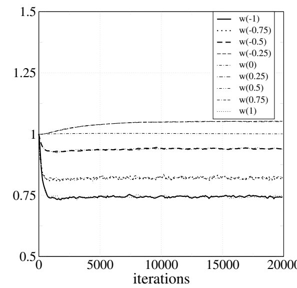


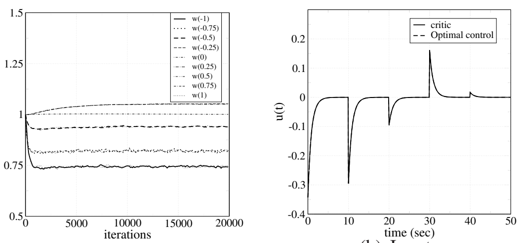
Fig. 3. Fuzzy Weights Evolution during training phase

The final values of weights in different operating zone is tabulated in Table I. It is clear from the table that the weight in each operating fuzzy zone smoothly varies from the optimal weight corresponding to origin. The optimal weight around the zone of origin is very closer to the quadratic cost function corresponding to the linearized system around origin, given by $J^* = 0.5x^T x$, which corresponds to $P = 1$. The learned weight around the fuzzy zone $x_1 = 0$ is 1.0009 which is closer to the optimal cost of local linear model, which confirms our claim. The global optimal cost varies smoothly from the cost of local linear model around the origin. After initial training the system is controlled from different initial conditions and the control result is compared with the optimal control law. The performance comparison of both the proposed controller and optimal controller is shown in figure 4.

TABLE I
FIRST ORDER NONLINEAR SYSTEM: OPTIMAL WEIGHTS

fuzzy zone	weight
-1	0.74513
-0.75	0.8185
-0.5	0.94148
-0.25	1.0525
0	1.0009
0.25	1.0511
0.5	0.93881
0.75	0.82113
1	0.74792

The corresponding optimal cost is tabulated in Table II. It is evident that, the proposed scheme approximates the optimal cost effectively and performs closer to the optimal controller.



(a) State (b) Input

Fig. 4. First Order System::Controller Performance

TABLE II
COMPUTED COST AT SIMULATED POINTS

x_0	Critic Cost	Optimal Cost
0.2915	0.0563	0.0540
0.2221	0.0415	0.0400
0.0762	0.0044	0.0046
-0.1237	0.01248	0.01244
-0.0149	0.00013	0.00014

To evaluate further, the optimal cost from various initial operating points x_0 in the operation zone is shown in figure 5. The figure clearly shows that the proposed T-S Fuzzy model based adaptive critic method approximates the optimal cost function effectively.

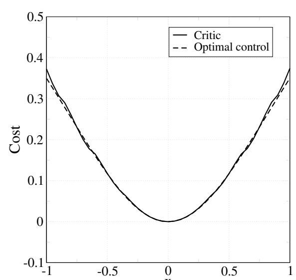


Fig. 5. First Order System:: Optimal Cost from different initial states

C. Vanderpol Oscillator

The objective of this simulation is to show that an improved cost can be obtained using T-S fuzzy clusters for nonlinear systems, compared to the conventional LQR controller obtained with the linearized model around the origin. With T-S Fuzzy based adaptive critic, the controller would guarantee the stability and optimality for a wide range of operation.

The Vanderpol oscillator system is considered for analysis because it is a benchmark problem. The system has an unstable equilibrium point at the origin and exhibits limit cycle too.

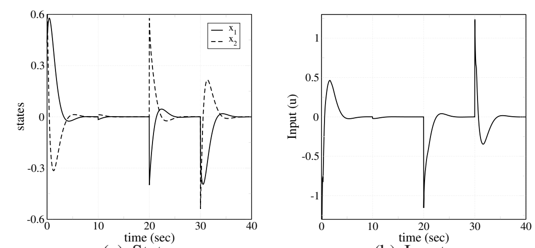
The dynamics of the system is given by,

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \alpha(1-x_1^2)x_2 - x_1 + (1+x_1^2+x_2^2)u \end{aligned} \tag{25}$$

where, α is taken as 0.5 in our simulation. The control task is to compute the input u such that it will derive the states $x = [0,0]$ while minimizing the quadratic cost function,

$$J = \frac{1}{2}\int_0^\infty [x^T Qx + u^T Ru]dt \tag{26}$$

where, $R = 1$ and Q is taken as Identity matrix. The continuous-time adaptive critic is employed and the optimum cost function is approximated by training a T-S fuzzy based critic network with 9 equally spaced fuzzy zones for both the states in the operating zone $[-1,1]$. Hence there will be 81 zones around the entire operating region. The critic network is trained with 50000 random points in the operating zone and then the controller performance is analyzed. The closed loop control performance from different initial operating point is shown in figure 6. It is easy to infer that the T-S Fuzzy based adaptive critic stabilizes the system in the considered operating zone.



(a) State (b) Input

Fig. 6. Vanderpol Oscillator:: Controller Performance

The cost incurred from different operating zone is computed and analyzed to evaluate further. The cost incurred due to LQR and T-S Fuzzy based critic is compared in figure 7. The figure shows the cost incurred from initial state $x = [x_1, 0]$ in the operating zone. These initial states are chosen in particular because, LQR is designed for origin, i.e., $x = [0,0]$. Hence, the figure would clearly show the improvement in optimality obtained by T-S fuzzy based critic network, as the system state x_1 varies around the origin. It is obvious from the figure that the T-S Fuzzy based critic network gives an optimal solution better than simple LQR based control as the zone of operation increases. The reader can also note that the optimal cost near the origin is closer to the optimal cost of linearized model around $x = [0,0]$.

D. Single Link Manipulator

The single link manipulator is considered for further analysis to check the performance of proposed scheme. The dynamics of the system is given by,

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -10sin(x_1) + u \end{aligned} \tag{27}$$

The task is to find the control law u, which minimizes the cost

$$J = \frac{1}{2}\int_0^\infty [x^T Qx + u^T Ru]dt \tag{28}$$

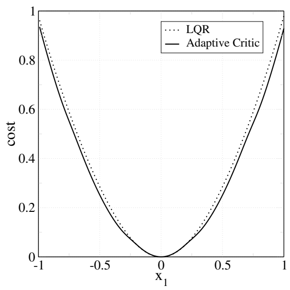


Fig. 7. Vanderpol Oscillator:: Optimal Cost from different initial states

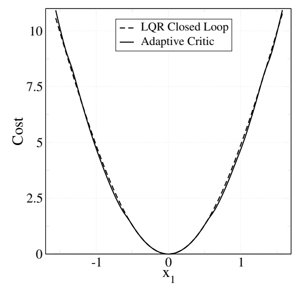
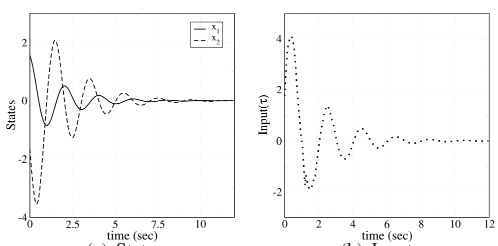


Fig. 9. Single Link Manipulator:: Optimal Cost from different initial states

where, $R = 1$ and Q is considered as Identity matrix. The region of operation is chosen as $[-\pi/2, \pi/2]$ for both x_1 and x_2. The optimal cost is learned with a T-S fuzzy system with 9 equally spaced fuzzy zones for both the states. The critic is learned initially with 50000 random points in the operating zone and then the performance is compared with LQR gains. The closed loop control performance with critic is shown in figure 8 which shows that both the system states and input are bounded and smooth.



(a) State (b) Input

Fig. 8. Single Link Manipulator:: Controller Performance

The cost incurred due to LQR and T-S Fuzzy based critic control is compared in figure 9. The figure shows the control cost incurred, from initial state $x = [x_1, 0.0]$. It can be easily seen from the figure that the adaptive critic optimizes cost better in most of the regions which confirms that the proposed T-S fuzzy based adaptive critic scheme gives a better optimal control performance over a wider range of operation.

IV. Conclusion

The major contribution of this work is that, it shows that for a given nonlinear system, there is closed relation exists between the optimal cost of local linear model and global cost and the optimal cost of nonlinear system can be expressed as a fuzzy cluster of optimal costs of locally valid linear models using T-S representation. Thus the proposed T-S fuzzy based adaptive critic for nonlinear input affine system predicts the complex nonlinear optimal cost as a fuzzy cluster of local linear model costs. With such an architecture, the learned network explains the meaningful correspondence between the local linear model cost and the global nonlinear system cost. A learning methodology is proposed to learn the optimal weights which resolves the issue of under-determined weight update equation as stated in [1]. Extensive simulations are performed and shown that the T-S fuzzy model approximates the optimal cost effectively, with local models representing the optimal cost of local linear models in each fuzzy zone.

Acknowledgment

This work has been supported by UK-India Education and Research Initiative (UKIERI) grant 'Innovations in Intelligent Assistive Robotics'.

References

[1] Swagat Kumar, Radhakant Padhi, and Laxmidhar Behera, "Direct adaptive control using single network adaptive critic," in *IEEE International Conference on Systems of Systems Engineering, 2007, SOSE'07*. IEEE, April 2007.
[2] A.E. Bryson and Y.C. Ho, *Applied Optimal Control*, Taylor and Francis, 1975.
[3] Desineni S. Naidu, *Optimal Control Systems*, CRC Press, 2003.
[4] Dimitry P. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific, Belmount, 1995.
[5] P. J. Werbos, *Approximate dynamic programming for real-time control and neural modeling*, In D. A. White, and D. A Sofge (Eds.), *Handbook of Intelligent control, Multiscience Press*, 1992.
[6] D. V. Prokhorov an D. C. Wunsch II, "Adaptive critic designs," *IEEE Tr. on Neural Networks*, vol. 8, no. 5, pp. 997–1007, September 1997.
[7] S. Ferrari and R. F. Stengel, *Model based adaptive critic designs.*, In Jennie Si, A. G. Barto, W. B. Powell, and D. Wunsch II (Eds.), *Handbook of learning and Approximate Dynamic Programming*, IEEE Press, 2004.
[8] X. Liu and S. N. Balakrishnan, "Convergence analysis of adaptive critic based optimal control," in *Proceedings of the American Control Conference*, Chicago, USA, 2000, pp. 1929–1933.
[9] Radhakant Padhi, Nishant Unnikrishnan, Xiaohua Wang, and S.N. Balakrishnan, "A single network adaptic critic (snac) architecture for optimal control," *Neural Networks*, vol. 19, pp. 1648–1660, 2006.
[10] P. Jung-Wook, R. G. Harle, and G. K. Venayagamoorthy, "Adaptive-critc based optimal neurocontrol for synchronous generators in a power system using mlp/rbf neural networks," *IEEE Transactions on Industry Applications*, vol. 39, no. 5, pp. 1529–1540, October 2003.
[11] R. Padhi, S. N. Balakrishnan, and T. Randolph, "Adaptive critic based optimal neuro control synthesis for distributed parameter systems," *Automatica*, vol. 37, pp. 1223–1234, 2001.
[12] S. N. Balakrishnan and V. Biega, "Adaptive critic based neural network for aircraft optimal control," *Journal of Guidance, Control and Dynamics*, vol. 19, no. 4, pp. 893–898, July-August 1996.