Proceedings of the 2007 IEEE Symposium on Approximate
Dynamic Programming and Reinforcement Learning (ADPRL 2007)

A Novel Fuzzy Reinforcement Learning Approach in Two-Level Intelligent Control of 3-DOF Robot Manipulators

Nasser Sadati, *Member, IEEE*
Intelligent Systems Laboratory
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
sadati@sina.sharif.edu

Mohammad Mollaie Emamzadeh
Intelligent Systems Laboratory
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
molaie@ee.sharif.edu

***Abstract* – In this paper, a fuzzy coordination method based on Interaction Prediction Principle (IPP) and Reinforcement Learning is presented for the optimal control of robot manipulators with three degrees-of-freedom. For this purpose, the robot manipulator is considered as a two-level large-scale system where in the first level, the robot manipulator is decomposed into several subsystems. In the second level, a fuzzy interaction prediction system is introduced for coordination of the overall system where a critic vector is also used for evaluating its performance. The simulation results on using the proposed novel approach ,for optimal control of robot manipulators show its effectiveness and superiority in comparison with the centralized optimization methods.**

I. INTRUDUCTION

The optimal control is one of the most important topics in control theory and optimization of large-scale systems. The problems such as complexity, high dimensionality of variables, geographical separation of subsystems, etc., usually are the burdens for solving the overall problem in a centralized fashion. That is way during past three decades, many approaches such as coordination strategies in multi-level systems and decentralized schemes have been proposed by researchers.

In decentralized methods, the system is decomposed into several subsystems where their optimization only depends on local variables while the effects and interactions of other subsystems are either ignored, or considered resulting in robust decentralized sub-optimal control schemes.

In the coordination methods, similar to the decentralized approaches, the system is first decomposed into several subsystems, while the effects among them are compensated through a coordinator. In this approach, the control of large-scale systems is done by using the hierarchical multi-level control scheme. So hierarchical multi-level control is a common approach that has been presented as an important and efficient method in control of large scale systems.

The basic principle of hierarchical control is decomposition of a given large-scale system into several smaller scale systems and then coordination of the resulted sub-systems to reach the optimum solution. In an attempt for improving this strategy, Mesarovic et al. presented one of the earliest formal quantitative treatments of hierarchical systems by postulating two coordination principles; Interaction Prediction Principle (IPP) and Interaction Balance Principle (IBP) [1],[2], where the coordination of large-scale systems are mainly based on these two principles.

In [3]-[6], using these two principles (IPP & IBP), two new gradient based coordination schemes are introduced that have much faster convergence rates than the classical methods. In [7],[8], two new neuro-fuzzy reinforcement strategies are introduced for intelligent coordination of large-scale systems based on IPP and IBP, where critic vectors are used for evaluation of their performances. In [9]-[11], using the new gradient based coordination schemes, the optimal control of robot manipulators have also been considered. In this paper, by using the novel strategy [7], the optimal control of robot manipulators is investigated. The simulation results are also presented.

In section II, the problem formulation and control problems are defined. In section III, the dynamic model of a robot arm with three degrees-of-freedom has been formulated. In section IV, decomposition of the overall problem, into m sub-problems and modeling the corresponding subsystems is done. In section V, the optimization of the first level subsystems are done using the gradient method. In section VI, the proposed fuzzy system is introduced to predict the change of interactions while a critic vector is used to evaluate its performance. In section VII, the proposed coordination approach based on the proposed fuzzy interaction prediction system has been applied to optimal control of the robot manipulator and the obtained results are compared with the centralized optimal control approach. Finally, in section VIII, some concluding remarks are briefly discussed.

II. STATEMENT OF THE PROBLEM

Let there be given an overall process $P : X \times U \rightarrow Y$ and a performance function $G : U \times Y \rightarrow V$ with U as a set of controls, X as a set of states, Y as a set of outputs and V as a set of performance values. Let also g be defined on U by the following equation

$$g(U) = G[U , P(U)] \tag{1}$$

The goal of the overall control problem which is denoted by D, is to find a control action $\hat{U}$ which minimizes g over U. Such a control action will be referred to as the overall optimum.

Let $U = U_1 \times \cdots \times U_m$, $Y = Y_1 \times \cdots \times Y_m$, and $X = X_1 \times \cdots \times X_m$. For each $i = 1, 2, \cdots, m$, the i-th subsystem is given by

$P_i : U_i \times Z_i \to Y_i$, where Z_i is the set of interactions of the other subsystems. Now the i-th infimal control problem can be formulated in terms of an objective function g_i, given on $U_i \times Z_i$, in terms of the i-th subsystem. Also a performance function $G_i : U_i \times Z_i \times Y_i \to V$ can be given by the following equation

$$g_i(U_i, Z_i) = \mathrm{G}[U_i, Z_i, P(U_i, Z_i)] \tag{2}$$

Let Z_p in $Z = Z_1 \times \cdots \times Z_m$ be the predicted input for interference subsystems that is formulated by $P_{iz_P} = P_i(U_i, Z_{pi})$; $1 \le i \le m$. Then for each Z_p in Z, the infimal control problem $D_i(Z_{pi})$ is to find a control $\hat{U}_i$ in U_i such that

$$g_i(\hat{U}_i, Z_{pi}) = \min_{U_i} g_i(U_i, Z_{pi}) \tag{3}$$

and the minimization is only over the set U_i of local controls.

Let $Z_p = (Z_{p1}, \cdots, Z_{pm})$ be the predicted interference inputs and let $Z = (Z_1, \cdots, Z_m)$ be the actual interference inputs occurring when the sub-optimal control $\hat{U}(Z_p) = [\hat{U}_1(Z_{p1}), \cdots, \hat{U}_m(Z_{pm})]$ is implemented. The overall optimum is then achieved if the predicted interface inputs are correct (i.e. $e_i = 0$ where $e_i := Z_i - Z_{pi}$ and $1 \le i \le m$). Alternatively, if e_i can not be made to be zero, the supremal control problem can be defined as minimization of an appropriate function of the errors $e_1, e_1, \cdots, e_m$ in the second level.

In Fig. 1, the block diagram of the coordination of two subsystems using the interaction prediction principle is shown.

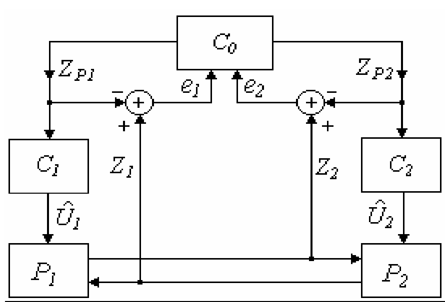

Fig. 1. Coordination of two subsystems using the interaction prediction principle.

III. Robot Modeling

The dynamic model of the robot arm with three degrees-of-freedom is shown in Fig. 2.

The dynamic model of the robot arm can be obtained by the Lagrangian method as follows

$$M(q)\ddot{q} + H(q, \dot{q}) + G(q) = \tau \tag{4a}$$

where q_i is the angle of joint i, $M(q)$ is the symmetric and positive definite inertia matrix, $H(q, \dot{q})$ is the coriolis and centrifugal vector, $G(q)$ is the gravity vector and τ is the torque vector.

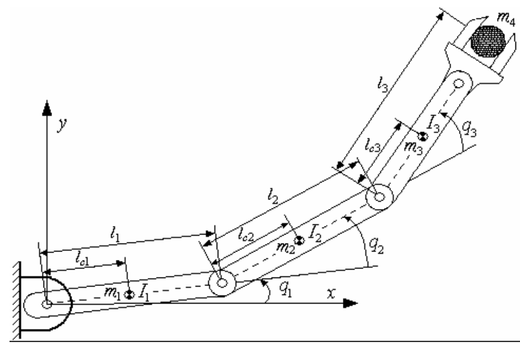

Fig. 2. the robot arm with three degrees-of-freedom.

Now using the following definitions

$$\begin{cases} p_1 := I_1 + m_1 l_{c1}^2 + m_2 l_1^2 + m_3 l_1^2 + m_4 l_1^2 & , \; c_i := \cos(q_i) \\ p_2 := I_2 + m_2 l_{c2}^2 + m_3 l_2^2 + m_4 l_2^2 & , \; s_i := \sin(q_i) \\ p_3 := I_3 + m_3 l_{c3}^2 + m_4 l_3^2 & , \; c_{ij} := \cos(q_i + q_j) \\ p_4 := [m_2 l_{c2} + m_3 l_2 + m_4 l_2] l_1 & , \; s_{ij} := \sin(q_i + q_j) \\ p_5 := [m_3 l_{c3} + m_4 l_3] l_2 & , \; c_{ijk} := \cos(q_i + q_j + q_k) \\ p_6 := [m_3 l_{c3} + m_4 l_3] l_1 & , \; s_{ijk} := \sin(q_i + q_j + q_k) \\ p_7 := [m_1 l_{c1} + m_2 l_1 + m_3 l_1 + m_4 l_1] g & , \; h_1 := p_5 s_3 \\ p_8 := [m_2 l_{c2} + m_3 l_2 + m_4 l_2] g & , \; h_2 := p_4 s_2 + p_6 s_{23} \\ p_9 := [m_3 l_{c3} + m_4 l_3] g & , \; h_3 := p_5 s_3 + p_6 s_{23} \end{cases} \tag{4b}$$

the parameters of equation (4a) can be written in the following forms

$$\begin{cases} q := \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}, \; \tau := \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix}, \; G = \begin{bmatrix} G_1 \\ G_2 \\ G_3 \end{bmatrix} = \begin{bmatrix} p_7 c_1 + p_8 c_{12} + p_9 c_{123} \\ p_8 c_{12} + p_9 c_{123} \\ p_9 c_{123} \end{bmatrix} \\ M(q) := \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}, \; H = \begin{bmatrix} -h_2(\dot{q}_2^2 + 2\dot{q}_1\dot{q}_2) - h_3(\dot{q}_3^2 + 2\dot{q}_1\dot{q}_3 + 2\dot{q}_2\dot{q}_3) \\ h_2\dot{q}_1^2 - h_1(\dot{q}_3^2 + 2\dot{q}_1\dot{q}_3 + 2\dot{q}_2\dot{q}_3) \\ h_3\dot{q}_1^2 + h_1(\dot{q}_2^2 + 2\dot{q}_1\dot{q}_2) \end{bmatrix} \\ \begin{vmatrix} M_{11} = p_1 + p_2 + p_3 + 2p_4 c_2 + 2p_5 c_3 + 2p_6 c_{23} & , \; M_{22} = p_2 + p_3 + 2p_5 c_3 \\ M_{12} = M_{21} = p_2 + p_3 + p_4 c_2 + 2p_5 c_3 + p_6 c_{23} & , \; M_{23} = M_{32} = p_3 + p_5 c_3 \\ M_{13} = M_{31} = p_3 + p_5 c_3 + p_6 c_{23} & , \; M_{33} = p_3 \end{vmatrix} \end{cases} \tag{4c}$$

IV. Decomosition of the Robot Model

For the purpose of decomposition, here the system of robot arm with three degrees-of-freedom is decomposed into three subsystems where each joint is assumed as one subsystem. This is given below

$$\begin{cases} X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}, \; U = \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix}, \; Z = \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix} \\ \begin{cases} X_i[k+1] = F_i(X_i[k], U_i[k], Z_i[k]) \\ X_i[0] = X_{i0} \qquad ; \qquad i = 1, 2, 3 \end{cases} \end{cases} \tag{5a}$$

and F_i is continuously double differential analytical function of the *i*-th subsystem, and X_i, U_i and Z_i ; $i = 1, 2, 3$ are defined as follows

$$\begin{cases} X_1[k] = \begin{bmatrix} x_{11}[k] \\ x_{12}[k] \end{bmatrix} = \begin{bmatrix} q_1[k] \\ \dot{q}_1[k] \end{bmatrix}, \; U_1[k] = u_1[k] = \tau_1[k] \\ Z_1[k] = \begin{bmatrix} z_{11}[k] \\ z_{12}[k] \\ z_{13}[k] \\ z_{14}[k] \\ z_{15}[k] \\ z_{16}[k] \end{bmatrix} = \begin{bmatrix} x_{21}[k] \\ x_{22}[k] \\ x_{31}[k] \\ x_{32}[k] \\ u_2[k] \\ u_3[k] \end{bmatrix} = \begin{bmatrix} q_2[k] \\ \dot{q}_2[k] \\ q_3[k] \\ \dot{q}_3[k] \\ \tau_2[k] \\ \tau_3[k] \end{bmatrix} \end{cases} \tag{5b}$$

$$\begin{cases} X_2[k]=\begin{bmatrix} x_{21}[k] \\ x_{22}[k]\end{bmatrix}=\begin{bmatrix} q_2[k] \\ \dot{q}_2[k]\end{bmatrix} \quad , \quad U_2[k]=u_2[k]=\tau_2[k] \\ Z_2[k]=\begin{bmatrix} z_{21}[k] \\ z_{22}[k] \\ z_{23}[k] \\ z_{24}[k] \\ z_{25}[k] \\ z_{26}[k]\end{bmatrix}=\begin{bmatrix} x_{11}[k] \\ x_{12}[k] \\ x_{31}[k] \\ x_{32}[k] \\ u_1[k] \\ u_3[k]\end{bmatrix}=\begin{bmatrix} q_1[k] \\ \dot{q}_1[k] \\ q_3[k] \\ \dot{q}_3[k] \\ \tau_1[k] \\ \tau_3[k]\end{bmatrix}\end{cases} \tag{5c}$$

$$\begin{cases} X_3[k]=\begin{bmatrix} x_{31}[k] \\ x_{32}[k]\end{bmatrix}=\begin{bmatrix} q_3[k] \\ \dot{q}_3[k]\end{bmatrix} \quad , \quad U_3[k]=u_3[k]=\tau_3[k] \\ Z_3[k]=\begin{bmatrix} z_{31}[k] \\ z_{32}[k] \\ z_{33}[k] \\ z_{34}[k] \\ z_{35}[k] \\ z_{36}[k]\end{bmatrix}=\begin{bmatrix} x_{11}[k] \\ x_{12}[k] \\ x_{21}[k] \\ x_{22}[k] \\ u_1[k] \\ u_2[k]\end{bmatrix}=\begin{bmatrix} q_1[k] \\ \dot{q}_1[k] \\ q_2[k] \\ \dot{q}_2[k] \\ \tau_1[k] \\ \tau_2[k]\end{bmatrix}\end{cases} \tag{5d}$$

V. Optimization with the Gradiant Method

Let us assume that we have a general nonlinear dynamic system described as follows

$$\begin{cases} X[k+1]=F(X[k],U[k]) \\ X[0]=X_0 \end{cases} \tag{6}$$

where X is the state vector, U is the control action and F is continuously double differential analytical function. Also the initial state X_0 is assumed to be known.

Now, the control problem is to find U which minimizes the cost function given by the following equation

$$J(X,U)=G_{n+1}(X[n+1])+\sum_{k=0}^{n} G_k(X[n],U[n]) \tag{7}$$

By decomposition of the overall system into m subsystems, each subsystem can be described by non-linear state space equation of the following form

$$\begin{cases} X_i[k+1]=F_i(X_i[k],U_i[k],Z_i[k]) \\ X_i[0]=X_{i0} \quad , \quad 1\le i\le m \end{cases} \tag{8}$$

The cost function can also be decomposed as follows

$$J(X,U)=\sum_{i=0}^{m} J_i(X_i,U_i,Z_i) \tag{9a}$$

where

$$J_i(X,U)=G_{i\ n+1}(X_i[n+1],Z_i[n+1])+\sum_{k=0}^{n} G_{i\ k}(X_i[n],U_i[n],Z_i[k]) \tag{9b}$$

Now, the necessary conditions for optimality of each subsystem can be written in terms of the Lagrangian L_i, which is in the following form

$$\begin{aligned} L_i(.) &= G_{i\ n+1}(X_i[n+1],Z_i[n+1])+\sum_{k=0}^{n} G_{i\ k}(X_i[k],U_i[k],Z_i[k]) \\ &+\sum_{k=0}^{n}[X_i[k+1]-F_i(X_i[k],U_i[k],Z_i[k])]^T \lambda_i[k] \end{aligned} \tag{10}$$

So the first order necessary conditions become

$$\begin{cases} \dfrac{\partial L_i(.)}{\partial \lambda[k]}=X_i[k+1]-F_i(X_i[k],U_i[k],Z_i[k])=0 \\ i=1,\ldots,m \quad , \quad k=0,\ldots,n \end{cases} \tag{11}$$

$$\begin{cases} \dfrac{\partial L_i(.)}{\partial X_i[n+1]}=\dfrac{\partial G_{i\ n+1}(X_i[n+1],Z_i[n+1])}{\partial X_i[n+1]}+\lambda_i[n]=0 \\ i=1,\ldots,m \end{cases} \tag{12}$$

$$\begin{cases} \dfrac{\partial L_i(.)}{\partial X_i[k]}=\dfrac{\partial G_{i\ k}(X_i[k],U_i[k],Z_i[k])}{\partial X_i[k]}+\lambda_i[k-1]-\left(\dfrac{\partial F_i(X_i[k],U_i[k],Z_i[k])}{\partial X_i[k]}\right)^T \lambda_i[k]=0 \\ i=1,\ldots,m \quad , \quad k=1,\ldots,n \end{cases} \tag{13}$$

$$\begin{cases} \dfrac{\partial L_i(.)}{\partial U_i[k]}=\dfrac{\partial G_{i\ k}(X_i[k],U_i[k],Z_i[k])}{\partial U_i[k]}-\left(\dfrac{\partial F_i(X_i[k],U_i[k],Z_i[k])}{\partial U_i[k]}\right)^T \lambda_i[k]=0 \\ i=1,\ldots,m \quad , \quad k=0,\ldots,n \end{cases} \tag{14}$$

Now for each subsystem, the control problem U_i which minimizes the Lagrangian L_i, can be obtained by the following updating rule

$$U_i^l[k]=U_i^{l-1}[k]-\eta\frac{\partial L_i()}{\partial U_i[k]} \tag{15}$$

VI. Fuzzy Prediction

The prediction of the change of interactions will be done using a fuzzy system. For evaluating its performance, a critic vector is used as shown in Fig. 3, to develop a method of training for this fuzzy system. The proposed training approach is based on minimizing the energy of the critic vector. In this reinforcement approach, the error of prediction and its rate of changes are used in such a way that they increase the speed of convergence of the algorithm.

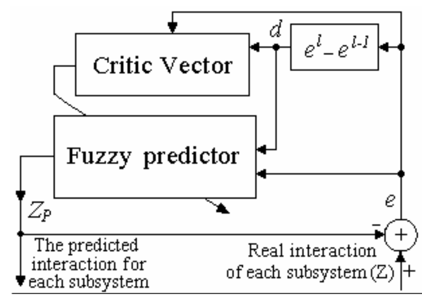

Fig. 3. The block diagram of fuzzy interaction prediction system.

A. Designing the Critic Vector

The critic vector is defined in the following forms

$$r_i[k]=R(e_i[k]\ ,\ d_i[k])\ ;\quad 1\le i\le m \quad , \quad 0\le k\le n \tag{16}$$

where

$$d_i[k]:=d_i^l[k]=e_i^l[k]-e_i^{l-1}[k] \tag{17}$$

$$e_i[k]=Z_i[k]-Z_{Pi}[k] \tag{18}$$

$$Z=[Z_1 \ \cdots \ Z_i \ \cdots \ Z_m]^T,\ Z_i=[Z_i[1] \ \cdots \ Z_i[k] \ \cdots \ Z_i[n]] \tag{19}$$

also Z_i is the real interactions and Z_{Pi} is the predicted value given by the coordinator. Now, the critic vector is assumed to be formulated as follows

$$r_i[k]:=R(e_i[k]\ ,\ d_i[k])=e_i[k]^2+\lambda\ d_i[k]^2 \tag{20}$$

where λ is a positive constant parameter $(i.e.\ 0\le\lambda\le 1)$.

B. Designing the Fuzzy Prediction System

The predicted values for the interactions are assumed to be updated in the following form

$$Z_{Pi}^l=Z_{Pi}^{l-1}+\Delta Z_{Pi} \tag{21}$$

To predict the change of interactions, a fuzzy system based on fuzzy inference is used that denotes the variation of the predicted values as follows

$$\Delta Z_i[k]:=s_i[k]:=S(e_i[k]\ ,\ d_i[k]) \tag{22}$$

where S is a fuzzy system based on the Takagi-Sugeno-Kang (TSK) model, that also has been used extensively in fuzzy modeling, and in this case, is defined by the following fuzzy sets and rules

$$\begin{cases} if\ \ e_i[k] is\ A_1\ \ and\ \ d_i[k] is\ B_1\ \ then\ \ \ s_1 = a_1 \cdot e_i[k] + b_1 \cdot d_i[k] + c_1 \\ \vdots \\ if\ e_i[k] is\ A_N\ and\ \ d_i[k] is\ B_N\ \ then\ \ s_N = a_N \cdot e_i[k] + b_N \cdot d_i[k] + c_N \end{cases} \tag{23}$$

In this paper, a_i and b_i are assumed to be zero. So the fuzzy sets and rules can be obtained easily, in the following forms

$$\begin{cases} if\ \ e_i[k] is\ A_1\ \ and\ \ d_i[k] is\ B_1\ \ then\ \ \ s_1 = c_1 \\ \vdots \\ if\ e_i[k] is\ A_N\ and\ \ d_i[k] is\ B_N\ \ then\ \ s_N = c_N \end{cases} \tag{24}$$

where μA_j and μB_j are the membership functions of A_j and B_j, respectively. Also c_j are the real constant parameters.

Now to summarize, the relation of $s_i[k]$ with $e_i[k]$ and $d_i[k]$ can be given by the following fuzzy inference

$$s_i[k] = S\left(e_i[k]\ ,\ d_i[k]\right) = \frac{\sum_{j=1}^{N} \mu A_j\left(e_i[k]\right) \cdot \mu B_j\left(d_i[k]\right) \cdot c_j}{\sum_{j=1}^{N} \mu A_j\left(e_i[k]\right) \cdot \mu B_j\left(d_i[k]\right)} \tag{25}$$

C. Training the Fuzzy Predictor

The goal of training is to minimize the energy of a critic vector related to the system parameters ; c_j, where

$$E = \frac{1}{2} \sum_{k=0}^{n+1} r[k]^T\, r[k] \tag{26}$$

also

$$r[k] = \begin{bmatrix} r_1[k] & r_2[k] & \cdots & r_m[k] \end{bmatrix}^T \tag{27}$$

Now to update the parameters of the proposed fuzzy system, the following updating rule can be used

$$\Delta c_j = -\eta \frac{\partial E}{\partial c_j} = -\eta \sum_{k=0}^{n+1} \frac{\partial r[k]^T}{\partial c_j} \cdot r[k] = -\eta \sum_{i=1}^{m} \sum_{k=0}^{n+1} \frac{\partial r_i[k]}{\partial c_j} \cdot r_i[k] \tag{28}$$

where η is a step size which has an effect on the rate of training and convergence of the fuzzy system parameters, during learning. Now, by using the chain rule, we have

$$\frac{\partial r_i[k]}{\partial c_j} = \frac{\partial r_i[k]}{\partial e_i[k]} \cdot \frac{\partial e_i[k]}{\partial c_j} + \frac{\partial r_i[k]}{\partial d_i[k]} \cdot \frac{\partial d_i[k]}{\partial c_j} \tag{29}$$

To calculate the right side of this equation, we need to calculate; $\frac{\partial r_i[k]}{\partial e_j}$, $\frac{\partial e_i[k]}{\partial c_j}$, $\frac{\partial r_i[k]}{\partial d_j}$ and $\frac{\partial d_i[k]}{\partial c_j}$. Now using (20), we can write

$$\frac{\partial r_i[k]}{\partial e_i[k]} = 2\, e_i[k] \qquad , \qquad \frac{\partial r_i[k]}{\partial d_i[k]} = 2\, \lambda\, d_i[k] \tag{30}$$

Moreover, the gradient of the prediction errors related to the system parameters can be given by

$$\frac{\partial e_i[k]}{\partial c_j} = \frac{\partial e_i[k]}{Z_{P\,i}[k]} \cdot \frac{Z_{P\,i}[k]}{\partial c_j} = \frac{\partial e_i[k]}{\partial \Delta Z_{P\,i}[k]} \cdot \frac{\partial \Delta Z_{P\,i}[k]}{\partial c_j} = -\frac{\partial \Delta Z_{P\,i}[k]}{\partial c_j} := T_{i,j}[k] \tag{31}$$

Now, in order to calculate $T_{i,j}[k]$, using (22) and (25), we have

$$T_{i,j}[k] := -\frac{\partial \Delta Z_{P\,i}[k]}{\partial c_j} = -\frac{\mu A_j\left(e_i[k]\right) \cdot \mu B_j\left(d_i[k]\right)}{\sum_{l=1}^{N} \mu A_l\left(e_i[k]\right) \cdot \mu B_l\left(d_i[k]\right)} \tag{32}$$

Also from (17) and (31), we can formulate $\frac{\partial d_i[k]}{\partial c_j}$ as follows

$$\frac{\partial d_i[k]}{\partial c_j} = \frac{\partial d_i[k]}{\partial e_i^l[k]} \cdot \frac{\partial e_i^l[k]}{\partial c_j} + \frac{\partial d_i[k]}{\partial e_i^{l-1}[k]} \cdot \frac{\partial e_i^{l-1}[k]}{\partial c_j} = T_{i,j}^l[k] - T_{i,j}^{l-1}[k] \tag{33}$$

So the training of the fuzzy predictor can be formulated in the following form

$$c_j^{l+1} = c_j^l[k] + \Delta c_j \tag{34a}$$

where

$$\Delta c_j = -\eta \sum_{i=1}^{m} \sum_{k=0}^{n+1} \frac{\partial r_i[k]}{\partial c_j} \cdot r_i[k] = -\eta \sum_{i=1}^{m} \sum_{k=0}^{n+1} \frac{\partial r_i[k]}{\partial c_j} \cdot r_i[k] \tag{34b}$$

$$\frac{\partial r_i[k]}{\partial c_j} = 2\, e_i[k] \cdot T_{i,j}^l[k] + 2\, \lambda\, d_i[k] \cdot \left(T_{i,j}^l[k] - T_{i,j}^{l-1}[k]\right) \tag{34c}$$

VII. Simulation Results

The proposed method is used to obtain the optimal control of a 3DOF robot manipulator. The parameters of the robot model are assumed to have the following values [12]

Table 1. The parameters of the robot manipulator.

	Mass	Length		Moment of Inertia
	(m_i , Kg)	(l_i , m)	(l_{ci} , m)	(I_i , $Kg.m^2$)
Joint 1	1.2	0.5	0.25	$43.33 * 10^{-3}$
Joint 2	1.5	0.4	0.2	$25.08 * 10^{-3}$
Joint 3	3.0	0.3	0.15	$32.68 * 10^{-3}$
$m_{load} = m_4 = 0\ Kg$, $g = 9.8\ N/Kg$				

Now using (9), the cost function can be written as

$$\begin{cases} G_{i\ n+1}(\cdot) = \frac{k_{n+1}^2}{2}\left(\left[X_{i1}[n+1] - X_{i1\ d}[n+1]\right]^2 + k_v^2\left[X_{i2}[n+1] - X_{i2\ d}[n+1]\right]^2\right) \\ G_{i\ k}(\cdot) = \frac{1}{2n}\left(\left[X_{i1}[k] - X_{i1\ d}[k]\right]^2 + k_v^2\left[X_{i2}[k] - X_{i2\ d}[k]\right]^2 + k_{ui}^2\ U_i[k]^2\right) \end{cases} \tag{35}$$

where k_{n+1} , k_u and k_v are the parameters of the cost function and defined by the conditions and limitations of the problem. In the simulation presented in this paper, they are defined to be

Table 2. The parameters of the cost function.

k_{n+1}	k_v	$k_u = [\ k_{u1}\ \ k_{u2}\ \ k_{u3}]$	T	t_0	t_f	n
1	10^{-2}	$[1.5\ \ 3\ \ 12] \times 10^{-3}$ or $[0\ \ 0\ \ 0]$	0.01 S	0 S	0.5 S	50

In critic vector, λ is assumed to be 0.2. The fuzzy membership functions of μA_j or μB_j are experimentally defined by three triangular membership functions as shown in Figure 4. By this number of membership functions, the fuzzy system has $N = 3^2 = 9$ rules.

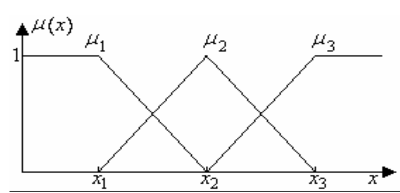

Fig. 4. The triangular fuzzy membership functions.

The values of x_1, x_2, *and* x_3 are shown in Fig. 4. These values are not equal for different inputs (the input of fuzzy membership functions; x, can be one of the three following forms; e_Z *or* d_Z for predicted angles, e_Z *or* d_Z for predicted velocities, and e_Z *or* d_Z for predicted control actions.). These are defined as follows

$$\boxed{\begin{array}{l} for\ x = e_{Z_{ij}}\ or\ d_{Z_{ij}}\ \left(Z_{ij} = q_1\ or\ q_2\ or\ q_3\right); x_1 = -0.1,\ x_2 = 0\ ,\ x_3 = 0.1 \\ for\ x = e_{Z_{ij}}\ or\ d_{Z_{ij}}\ \left(Z_{ij} = \dot{q}_1\ or\ \dot{q}_2\ or\ \dot{q}_3\right); x_1 = -1\ ,\ x_2 = 0\ ,\ x_3 = 1 \\ for\ x = e_{Z_{ij}}\ or\ d_{Z_{ij}}\ \left(Z_{ij} = \tau_1\ or\ \tau_2\ or\ \tau_3\right); x_1 = -50\ ,\ x_2 = 0\ ,\ x_3 = 50 \end{array}} \quad (36)$$

The proposed approach is applied to the robot manipulator to minimize the cost function defined by (9) and (35), where the desired target trajectories are assumed to be in three sinosoeid forms and the initial conditions are assumed to be zero [12].

The simulations are performed for two values; $k_u = 0$ and $k_u = [1.5 \quad 3 \quad 12]\times 10^{-3}$, after 70 iterations of coordination, where in each iteration the optimization of each subsystem is performed using the gradient method in 200 steps. For $k_u = 0$, the goal of minimization of cost function defined by (9) and (35) is only minimization of an energy function of tracking errors, so there are no limitations on the norm of the input actions. But for $k_u = [1.5 \quad 3 \quad 12]\times 10^{-3}$, the goal of minimization of cost function is minimization of both energy functions of tracking errors and input actions.

In Fig. 5, the calculated optimal control actions are presented for both values; $k_u = 0$ and $k_u = [1.5 \quad 3 \quad 12]\times 10^{-3}$. As shown in this figure, for $k_u = [1.5 \quad 3 \quad 12]\times 10^{-3}$, the obtained control actions have smaller values (norm) than the obtained control actions corresponding to $k_u = 0$. Also in Fig. 6, the obtained optimal trajectories of all joint angles are shown where for $k_u = 0$, the optimized trajectories of all joint angles have tracked the target trajectories with acceptable tracking errors. In contrast to obtained results using $k_u = 0$, the optimized trajectories of joint angles using $k_u = [1.5 \quad 3 \quad 12]\times 10^{-3}$ have not be able to track the target trajectories with acceptable tracking errors.

In Figs. 7, 8 and 9, respectively, the sum-squared interaction prediction errors of control actions $[\tau_1[k] \quad \tau_2[k] \quad \tau_3[k]]$, joint angles $[q_1[k] \quad q_2[k] \quad q_3[k]]$, and the joint angular velocities $[\dot{q}_1[k] \quad \dot{q}_2[k] \quad \dot{q}_3[k]]$; $1 \le k \le 50$ are shown. These figures present the convergence of the interactions predicted by fuzzy system in 50 iterations of coordination.

The simulation results of both methods; the centralized optimization approach (in this approach, the whole problem is solved in one shot, using a typical gradient optimization method), and the proposed strategy, are plotted together in Figs. 10 and 11. The centralized gradient based optimization is performed in 10000 steps for optimization of 150 variables; $U_1[1],\cdots,U_1[50]$, $U_2[1],\cdots,U_2[50]$ and $U_3[1],\cdots,U_3[50]$. In the proposed strategy, the coordination is done in 50 iterations and in each iteration, the optimization of each subsystem is performed in 200 steps, where each subsystem has 50 variables; $U_i[1] \cdots U_i[50]$; $i = 1,2,3$. As is shown in Figs. 10 and 11, the proposed approach has faster convergence rate rater than the centralized optimization approach.

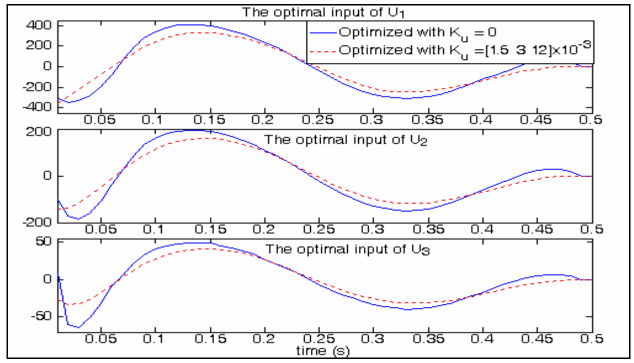

Fig. 5. The optimal trajectories of control actions obtained using fuzzy coordination scheme.

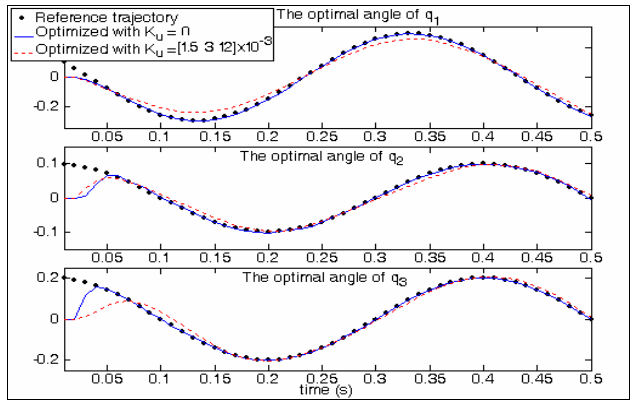

Fig. 6. The optimal trajectories of the joint angles obtained using fuzzy coordination scheme.

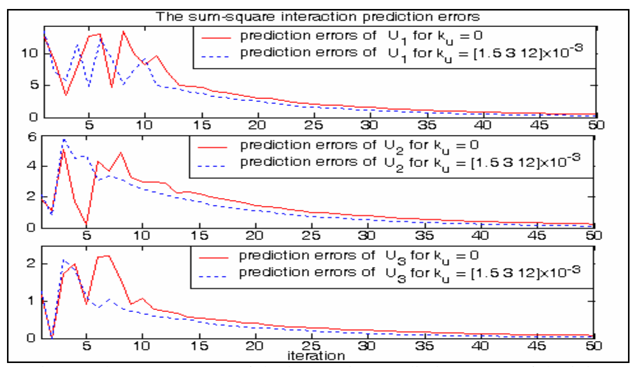

Fig. 7. The sum-squares of the interaction prediction errors of the joint torques.

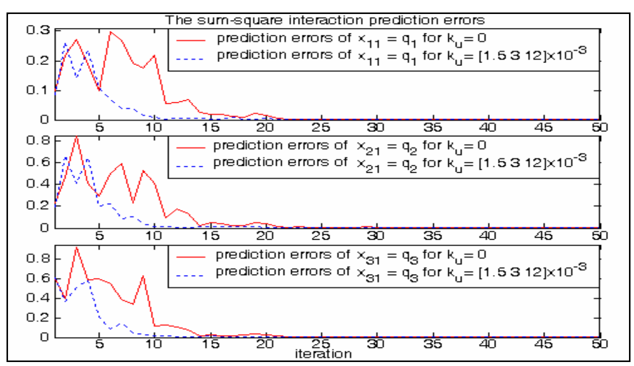

Fig. 8. The sum-squares of the interaction prediction errors of the joint angles.

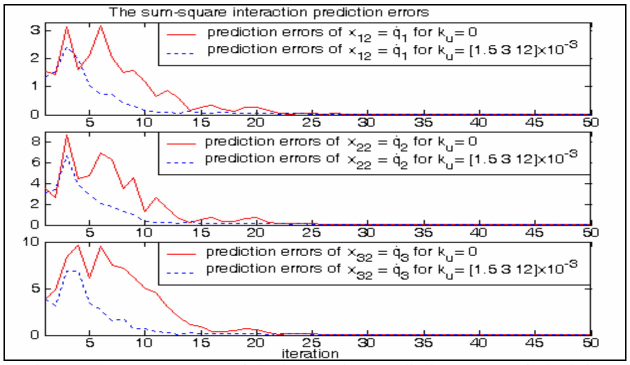

Fig. 9. The sum-squares of the interaction prediction errors of the joint angular velocities.

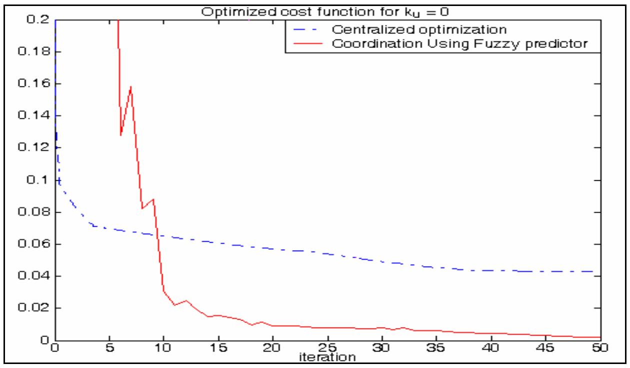

Fig. 10. The optimized cost function with $k_u = 0$, using two approaches ; the centralized optimization, and the proposed fuzzy coordination method.

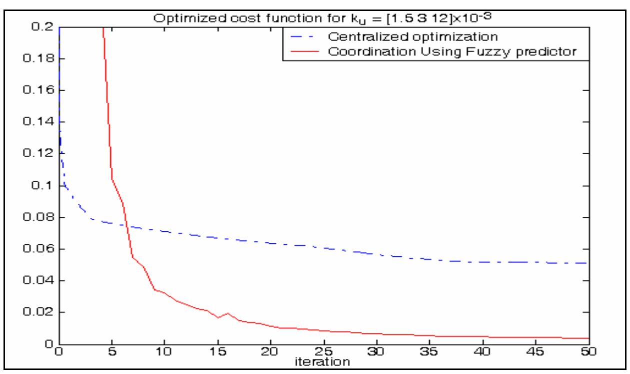

Fig. 11. The optimized cost function with $k_u = [1.5 \quad 3 \quad 12]\times 10^{-3}$, using two approaches; the centralized optimization, and the proposed fuzzy coordination method.

VIII. Conclusion

In this paper, a new approach for optimal control of robot manipulators, considered as two-level large-scale systems, is presented. It is shown how the specific feature of intelligent control including decision making, adaptation, planning and other characteristics of fuzzy logic and classical optimization methods, could help to design a unified hybrid approach for intelligent control of robot manipulators.

The fuzzy interaction prediction system is used to coordinate the overall system and also resolve the weakness of conventional coordination approaches in convergence. This fuzzy coordinator uses a critic vector to evaluate its own operation. The fuzzy coordinator learns its dynamics (parameters) through the minimization of an energy function relating to the critic vector. The minimization process, to train the fuzzy coordinator, is done by using the gradient of the energy function. The training of fuzzy predictor can be done in an off-line or on-line manner. Here, the parameters of fuzzy predictor are first calculated by an off-line manner and then, they are updated in each iteration of coordination in an on-line manner. It should be noted that in addition to the convergence of prediction, both critic vector and also the fuzzy interaction prediction system use the error and its variations, to predict the change of interactions for coordination of the overall system.

The obtained results present the performance and convergence of the proposed approach. By simulation of a robot manipulator, the superiority of the proposed method in comparison to the centralized method is shown.

The proposed approach could also be extended and used for coordination and cooperation of multi-agent systems.

References

[1] M. D. Mesarovic, D. Macko and Y. Takahara, *Theory of Hierarchical Multilevel System*. Academic press, New York, 1970.

[2] M. D. Mesarovic, D. Macko and Y. Takahara, “Two coordination criniciples and their application in large scale systems control,” *Pergamon Press*, New York, March 1970.

[3] N. sadati, “A new gradient based method for model coordination of large-scale systems,” Sharif University of Technology, Electrical Engineering Department, Report no. SUT-ICSL (18), 1997.

[4] N. Sadati, “A novel Approach to Coordination of Large-Scale Systems; part I - Interaction Prediction Principle,” *IEEE Int. Conf. on Industrial Technology,* 2005, pp. 641-647.

[5] N. Sadati “A Novel Approach to Coordination of Large-Scale Systems; part II - Interaction Balance Principle,” *IEEE Int. Conf. on Industrial Technology,* 2005, pp. 648-654.

[6] N. Sadati, “A two-level gradient based approach for intelligent coordination of large-scale systems. Part I,” *IEEE Int. Conf. on Tools with Artificial Intelligence*, Hong Kong, China, Nov. 2005.

[7] N. Sadati, "A neuro-fuzzy strategy to intelligent coordination of large-scale systems; Part I-Interaction prediction approach," Primary Accepted, under second review, IEE Proceedings, Control Theory and Applications, 2005.

[8] N. Sadati, "A neuro-fuzzy strategy to intelligent coordination of large-scale systems; Part II-Interaction balance approach," Primary Accepted, under second review, IEE Proceedings, Control Theory and Applications, 2005.

[9] N. Sadati, A. Babazadeh, “A new hierarchical approach for optimal control of robot manipulators,” *IEEE Conf. on Robotics, Automation and Mechatronics,* 2004, pp. 306-311.

[10] N. Sadati and A. Babazadeh, "Optimal control of robot manipulatros with a new two-level gradient-based approach," Journal of Electrical Engineering, Springer, July 2005.

[11] N. Sadati and A. Babazadeh, "A new hierarchical approach for optimal control of multiple-arm robotic systems," under publication, Journal of Intelligent and Robotic Systems, Springer, June 2006.

[12] E. Kim, “Output feedback tracking control of robot manipulators with model uncertainty via adaptive fuzzy logic,” *IEEE Trans. On fuzzy systems,* vol. 12, no. 3, pp.368-378, June 2004.