3

A Reinforcement Learning Approach to Intelligent Goal Coordination of Two-Level Large-Scale Control Systems

Nasser Sadati[1,2] and Guy A. Dumont[1]
[1]*Electrical and Computer Engineering Department*
The University of British Columbia, Vancouver
[2]*Electrical Engineering Department, Sharif University of Technology,*
[1]*BC, Canada*
[2]*Iran*

1. Introduction

Two principles for coordination of large-scale systems, namely Interaction Prediction Principle and Interaction Balance Principle were postulated by Mesarovic et al. [1], [2] to provide guidance in synthesizing structures for multi-level or hierarchical control of large-scale systems and obtain the optimal solution. Hierarchical structures are feasible structures which reduce the complexity of large-scale control systems and improve the solution through decomposition, coordination and parallel processing [3]-[6]. In two-level hierarchical approaches, the overall system is first decomposed into several interactive sub-systems, at the first level, where the optimization problem is redefined for each one of them. The interactions between these sub-systems, at the first level, and the coordinator, at the second level, called the coordination parameters, are used so that the overall solution is obtained. In compare to centralized approaches, where the whole problem is considered for the solution at once, the computational efforts in hierarchical approaches are based on sub-problems, having smaller order, requiring less computational time, in addition to the coordination strategy.

The Goal Coordination based on Interaction Balance Principle approach of Mesarovic et al. has already been applied to large-scale systems and the results are reported in [3]- [5]. In applying the Interaction Balance Principle, the supremal controller modifies the infimal (i.e. first-level) performance functions, compares the interface inputs (interactions) demanded by the infimal controllers and those which actually occur, then provides new performance modifications whenever the error is observed as being outside the acceptable bounds. A brief description of the Goal Coordination and Interaction Balance Principle is presented in the following section. Although a more detailed discussion of this principle can be found in [1] ,[2] and also, voluminous literature on large-scale systems theories and applications including survey articles, textbooks and monographs can be found in [6]-[12].

Based on Interaction Balance Principle, a new goal coordination scheme, as a foundation for intelligent coordination of large-scale systems is postulated in this chapter. The approach is formulated in an intelligent manner such that it provides the update of the coordination parameters so to reduce the coordination errors directly and improve the convergence rate of the solution. The proposed scheme is a neuro-fuzzy based reinforcement learning approach which can be used to synthesize a new supervisory coordination strategy for the overall two-level large-scale systems, in which the sub-systems, at the first level of hierarchy, and also the overall process control objectives are considered as optimization problems. So with the aim of optimization, the control problem is first decomposed into m sub-problems at the first level, where each sub-problem can be solved using a neuro-regulator. The neural networks which are capable of learning and reconstructing non-linear mappings could also be used for modeling each corresponding sub-system. By using the new methodology which is based on a Fuzzy Goal Coordination System and Reinforcement Learning; using TSK model, a critic vector and the gradient of the interaction errors (difference between the actual interactions and the optimum calculated interaction values) and also their rate of changes, appropriate change of coordination parameters are generated at the second level and the coordination of the overall large-scale system is done. The proposed scheme results in faster reduction of the interaction errors, which finally vanish to zero.

This chapter is organized into several sections. In Section 2, the problem formulation and control problems are defined. Also a brief review of the classical Goal Coordination and Interaction Balance Principle is presented. In Section 3, decomposition of the overall large-scale system into m sub-problems and modelling each corresponding subsystem is done. In Section 4, the first level sub-problems are solved with neuro-regulators, and in Section 5, the new Fuzzy Goal Coordination System based Reinforcement Learning is presented to generate the appropriate change of coordination parameters. In Section 6, the efficacy and advantages of the proposed approach is demonstrated in an open-loop power system consisting of a synchronous machine connected to an infinite bus bar through a transformer and a transmission line. It is shown how the convergence of the interaction errors exceeds substantially those obtained using the classical goal coordination method. Finally, Section 7 contains some concluding remarks.

2. Statement of the problem

As it was mentioned in the Introduction, two cases arise as how the coordination might be effected and the infimal control problems can be defined. In this chapter, a new approach for coordination of large-scale systems based on Interaction Balance Principle, which is more convergent than the previously suggested classical methods, has been presented.

2.1 Goal coordination and Interaction Balance Principle

Let $\boldsymbol{B}$ be a given set such that each β in $\mathbf{B}$ specifies, for each $i=1,\ldots, m$, a performance function $G_{i\beta} : \boldsymbol{U}_i \times \boldsymbol{Z}_i \times \boldsymbol{X}_i \rightarrow \mathbf{V}$ which is a modification of the original G_i. Let the mapping $g_{i\beta}$ be defined on $\boldsymbol{U}_i \times \boldsymbol{Z}_i$ in terms of P_i and $G_{i\beta}$. For each β in $\boldsymbol{B}$, the infimal control problems is to find a pair $(\hat{U}_i, \hat{Z}_{pi})$ in $\boldsymbol{U}_i \times \boldsymbol{Z}_i$ such that

$$g_{i\beta}(\hat{U}_i, \hat{Z}_{pi}) = \min_{\boldsymbol{U}_i \times \boldsymbol{Z}_i} g_{i\beta}(U_i, Z_i)\,, \tag{1}$$

where minimization is over both sets $\boldsymbol{U}_i$ and $\boldsymbol{Z}_i$; the interface inputs are treated as free variables. Let β in $\boldsymbol{B}$ be given; let $\hat{Z}_1(\beta), \ldots, \hat{Z}_m(\beta)$ be the interface inputs required by the infimal controllers to achieve local optimum; let $Z_1(\beta),....,Z_m(\beta)$ be the interface inputs that occur if the control $\hat{U}(\beta)=[\hat{U}_1(\beta),....,\hat{U}_m(\beta)]$ is implemented, then the overall optimum is achieved if the actual interface inputs are precisely those required by local optimization

$$\hat{Z}_i(\beta)=Z_i(\beta) \tag{2}$$

for each $i = 1 , \dots , m$ (Interaction Balance Principle). If the Interaction Balance principle applies, the supremal control problem is to find β in B such that $e_i = \hat{Z}_i(B) - Z_i(B) = 0$, for each $i = 1 , \dots , m$. The application of the Interaction Balance Principle is shown in Fig. 1.

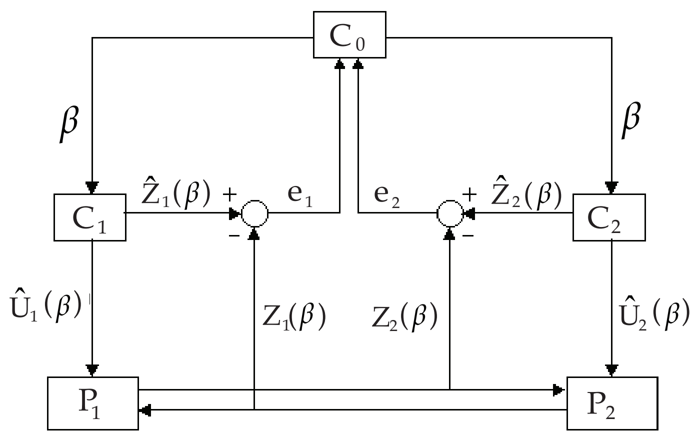


Fig. 1. Application of Interaction Balance Principle for coordination of two sub-problems.

Now, let us suppose that we have a general non-linear dynamic system described by the following state space equation

$$\underline{X}[k+1]=F\left(\underline{X}[k],\underline{U}[k]\right) \tag{3a}$$

$$\underline{X}[0]=\underline{X}_o \tag{3b}$$

where $\underline{X}$ is the state vector, $\underline{U}$ is the control vector and F is a continuously double differentiable analytical vector function which is going to be replaced by 2m neural models to describe the actual dynamics of m sub-systems and their interactions. The initial state $\underline{X}_o$ is also assumed to be known.
Now, the problem is to find $\underline{U}$ which minimizes the cost function given by

$$J=G_{n+1}\left(\underline{X}[n+1]\right)+\sum_{k=0}^{n}G_k\left(\underline{X}[k],\underline{U}[k]\right) \tag{4}$$

where G_k is in general, a scalar non-linear function of its arguments.

3. Decomposition of the overall problem into m sub-problems

Let us assume that the overall system comprises of m interconnected sub-systems. We assume that the sub-systems themselves can be described by non-linear state space equations of the following form

$$X_i[k+1] = F_i(X_i[k], U_i[k], Z_i[k]) \quad (5a)$$

$$X_i[0] = X_{i0} \quad (5b)$$

where X_i is the state, U_i is the control and Z_i is the interaction input of the ith sub-system that is assumed to be a non-linear function of the states of the m sub-systems

$$Z_i[k] = H_i(\underline{X}[k]) = H_i(X_1[k], \ldots, X_m[k]) \quad (6)$$

In Goal Coordination method, it is necessary for non-linear functions H_i to be separable. So the interaction variables Z_i must be defined in such a way that H_i functions to be separable, i.e.

$$Z_i[k] = H_i(X_1[k], \ldots, X_m[k]) = \sum_{j=1}^{m} H_{ij}(X_j[k]) \quad (7)$$

The interaction relations which can be expressed as $Z[k] = H(\underline{X}[k])$ are considered to be the optimization constraints. So the Lagrangian can be defined as

$$L = G_{n+1}\left(\underline{X}[n+1]\right) + \sum_{k=0}^{n} G_k\left(\underline{X}[k], \underline{U}[k]\right) + \sum_{k=0}^{n+1} \beta[k]^T \left(\underline{Z}[k] - H(\underline{X}[k])\right) \quad (8)$$

where β[k]'s are the Lagrange multipliers that we refer to them as the coordination parameters. Now, since the interaction function $H(\underline{X}[k])$ is separable, the Lagrangian can be decomposed as

$$L = \sum_{i=1}^{m} L_i \quad (9a)$$

where

$$L_i = G_{i_{n+1}}\left(X_i[n+1], Z_i[n+1]\right) + \sum_{k=0}^{n} G_{ik}(X_i[k], U_i[k], Z_i[k]) + \sum_{k=0}^{n+1}\left(\beta_i[k] Z_i[k] - \sum_{j=1}^{m} \beta_j[k] H_{ji}(X_i[k])\right) \quad (9b)$$

So the overall problem can be decomposed into m first level sub-problems of the following form

$$\min_{X_i, U_i, Z_i} L_i = G_{i_{n+1}}\left(X_i[n+1], Z_i[n+1]\right) + \sum_{k=0}^{n} G_{ik}(X_i[k], U_i[k], Z_i[k]) + \sum_{k=0}^{n+1}\left(\beta_i[k] Z_i[k] - \sum_{j=1}^{m} \beta_j[k] H_{ji}(X_i[k])\right) \quad (9b)$$

$$s.t. \quad X_i[k+1] = F_i(X_i[k], U_i[k], Z_i[k])$$

$$X_i[0] = X_{i0} \quad (10)$$

and also one second level problem expressed as:

Updating the coordination parameters $\beta_i[k]$ such that the interaction errors; $Z_i[k] - H_i(X_i[k], \ldots, X_m[k])$, become zero (Interaction Balance Principle).
Remark. In general, H(.) can be considered as a function of X[k] and U[k].

3.1 Modeling the corresponding sub-systems with neural networks

It should be noted that the dynamics of each sub-system and its interactions which are denoted by F_i and H_{ij}, respectively, could also be replaced by neural network models. So in this case, they can be denoted by NF_i and NH_{ij}, respectively.

$$X_i[k+1] = F_i(X_i[k], U_i[k], Z_i[k]) \triangleq NF_i(X_i[k], U_i[k], Z_i[k]) \tag{11}$$

$$Z_i[k] = H_i(\underline{X}[k]) \triangleq NH_i(X_1[k], \ldots, X_m[k]) = \sum_{j=1}^{m} NH_{ij}(X_j[k]) \tag{12}$$

The first step in identification of the sub-systems is to provide the training data using the actual system. To generate the training data, random inputs are applied to the actual system and the resulting state values, in addition to the input data are used for training the neural models.

4. Optimizing the first level sub-problems with neuro-regulators

In this approach, the first level sub-problems could be optimized with neuro-regulators [13]. The optimal control and interaction of each sub-system will be generated by non-linear feedback functions of the following forms

$$U_i[k] = NR_{Ui}(X_i[k], W_{Ui})\ ; \quad k = 0, 1, \ldots, n \tag{13}$$

$$Z_i[k] = NR_{Zi}(X[k], W_{Zi})\ ; \quad k = 0, 1, \ldots, n+1 \tag{14}$$

where NR_{Ui} and NR_{Zi} could be considered as multilayer perceptron (MLP) neural networks, and W_{Ui} and W_{Zi} are their parameters including weights and biases, respectively.
Now, the new Lagrangian L_i can be defined as follows

$$\begin{aligned} L_i = {} & G_{i_{n+1}}\left(X_i[n+1], Z_i[n+1]\right) + \sum_{k=0}^{n} G_{ik}(X_i[k], U_i[k], Z_i[k]) \\ & + \sum_{k=0}^{n+1}\left(\beta_i[k] Z_i[k] - \sum_{j=1}^{m} \beta_j[k] H_{ji}(X_i[k])\right) \\ & + \sum_{k=0}^{n} \lambda_i[k]\left(X_i[k+1] - F_i(X_i[k], U_i[k], Z_i[k])\right) \\ & + \sum_{k=0}^{n} \mu_{Ui}[k](U_i[k] - NR_{Ui}(X_i[k]; W_{Ui})) \\ & + \sum_{k=0}^{n+1} \mu_{Zi}[k](Z_i[k] - NR_{Zi}(X_i[k]; W_{Zi})) \end{aligned} \tag{15}$$

where $\lambda_i[k]$, $\mu_{Ui}[k]$ and $\mu_{Zi}[k]$ are the Lagrange multipliers.

Thus, the necessary conditions for optimality become

$$\frac{\partial L_i}{\partial U_i[k]} = \frac{\partial G_{ik}}{\partial U_i[k]} - \frac{\partial F_{ik}}{\partial U_i[k]}\lambda_i[k] + \mu_{Ui}[k] = 0\,; \quad k = 0,1,\ldots,n \tag{16}$$

$$\frac{\partial L_i}{\partial Z_i[n+1]} = \frac{\partial G_{i_{n+1}}}{\partial Z_i[n+1]} + \mu_{Zi}[n+1] = 0 \tag{17}$$

$$\frac{\partial L_i}{\partial Z_i[k]} = \frac{\partial G_{ik}}{\partial Z_i[k]} - \frac{\partial F_{ik}}{\partial Z_i[k]}\lambda_i[k] + \mu_{Zi}[k] = 0\,; \quad k = 0,1,\ldots,n \tag{18}$$

$$\frac{\partial L_i}{\partial X_i[n+1]} = \frac{\partial G_{i_{n+1}}}{\partial X_i[n+1]} + \lambda_i[n] - \frac{\partial NR_{Zn+1}}{\partial X_i[n+1]}\mu_{Zi}[n+1] = 0 \tag{19}$$

$$\frac{\partial L_i}{\partial X_i[k]} = \frac{\partial G_{ik}}{\partial X_i[k]} + \lambda_i[k-1] - \frac{\partial F_{ik}}{\partial X_i[k]}\lambda_i[k] - \frac{\partial NR_{Uk}}{\partial X_i[k]}\mu_{Ui}[k] - \frac{\partial NR_{Zk}}{\partial X_i[k]}\mu_{Zi}[k] = 0\,; \quad k = 1,2,\ldots,n \tag{20}$$

where

$$G_{ik} = G_i(X_i[k], U_i[k], Z_{pi}[k]) \tag{21}$$

$$F_{ik} = F_i(X_i[k], U_i[k], Z_{pi}[k]) \tag{22}$$

$$NR_{Uk} = NR_{Uk}(X_i[k]; W_{Ui}) \tag{23}$$

$$NR_{Zk} = NR_{Zk}(X_i[k]; W_{Zi}) \tag{24}$$

Now to train the neuro-regulators; NR_{Ui} and NR_{Zi} , based on preceding optimality conditions, the following algorithm can be suggested

1. Choose initial small values for neuro-regulator parameters, namely W_{Ui} and W_{Zi} .
2. Using initial state X_{i0} and equations (3), (13) and (14), find the values of $X_i[1], X_i[2], \ldots, X_i[n+1]$, $U_i[0], U_i[2], \ldots, U_i[n]$, and $Z_i[0], \ldots, Z_i[n+1]$.
3. Calculate $\lambda_i[k]$, $\mu_{Ui}[k]$, $\mu_{Zi}[k]$ for $k = n, n-1, \ldots, 0$, by using the following necessary conditions , backward in time;

$$\mu_{Zi}[n+1] = -\frac{\partial G_{i_{n+1}}}{\partial Z_i[n+1]} \tag{25}$$

$$\lambda_i[n] = -\frac{\partial G_{i_{n+1}}}{\partial X_i[n+1]} + \frac{\partial NR_{Zn+1}}{\partial X_i[n+1]}\mu_{Zi}[n+1] \tag{26}$$

$$\mu_{Ui}[k] = \frac{\partial F_{ik}}{\partial U_i[k]}\lambda_i[k] - \frac{\partial G_{ik}}{\partial U_i[k]}\,; \quad k = n, n-1, \ldots, 0 \tag{27}$$

$$\mu_{Zi}[k] = \frac{\partial F_{ik}}{\partial Z_i[k]}\lambda_i[k] - \frac{\partial G_{ik}}{\partial Z_i[k]}; \quad k = n, n-1, \ldots, 0 \tag{28}$$

$$\lambda_i[k-1] = \frac{\partial F_{ik}}{\partial X_i[k]}\lambda_i[k] + \frac{\partial NR_{Uk}}{\partial X_i[k]}\mu_{Ui}[k] + \frac{\partial NR_{Zk}}{\partial X_i[k]}\mu_{Zi}[k] - \frac{\partial G_{ik}}{\partial X_i[k]} \;\; ; \;\; k = n, n-1, \ldots, 1 \tag{29}$$

4. Calculate $\frac{\partial L_i}{\partial W_{Ui}}$ and $\frac{\partial L_i}{\partial W_{Zi}}$ for $k = n, n-1, \ldots, 0$, using $\mu_{Ui}[k]$ and $\mu_{Zi}[k]$

$$\frac{\partial L_i}{\partial W_{Ui}} = \sum_{k=0}^{n} \frac{\partial NR_{Uk}}{\partial W_{Ui}} \mu_{Ui}[k] \tag{30}$$

$$\frac{\partial L_i}{\partial W_{Zi}} = \sum_{k=0}^{n} \frac{\partial NR_{Zk}}{\partial W_{Zi}} \mu_{Zi}[k] \tag{31}$$

5. Update W_{Ui} and W_{Zi}, by adding $\Delta W_{Ui} = -\eta_U \frac{\partial L_i}{\partial W_{Ui}}$ and $\Delta W_{Zi} = -\eta_Z \frac{\partial L_i}{\partial W_{Zi}}$ to the prior values of W_{Ui} and W_{Zi}.
6. If $\left\| \frac{\partial L_i}{\partial W_{Ui}} + \frac{\partial L_i}{\partial W_{Zi}} \right\| < \varepsilon$ stop the algorithm, else go to step (2).

Remark. We should indicate that, in case the use of neural networks and neuro-regulators are not of interest, then the modelling and optimization process at the first level, can be easily done using the same approach as explained in Ref. [14]-[16].

5. Reinforcement Learning

To evaluate the operation of the fuzzy goal coordination system, with the use of reinforcement learning, we define a critic vector [17] and develop a method to train the new coordination strategy. The training is based on minimizing the energy of the critic vector. In this approach, we use both the errors and the rate of errors to increase the speed of convergence of the coordination algorithm.

5.1 Designing the critic vector

The critic vector includes m critic signals, where each of them evaluates the operation of the corresponding sub-system. The value of each critic signal is in the range of [-1 , 1] and is expressed by a fuzzy system of the following form

$$r_i[k] = \tilde{R}_i(e_i[k], d_i[k]) \; ; \quad i = 1, 2, \ldots, m \tag{32}$$

where $\tilde{R}_i$ is the fuzzy system, $e_i[k]$ is the interaction error and $d_i[k]$ is the rate of error, defined by

$$e_i[k] = Z_i[k] - Z_i^*[k] \tag{33a}$$

$$d_i[k] \triangleq d_i[k]^{(l)} = e_i[k]^{(l)} - e_i[k]^{(l-1)} \tag{33b}$$

also l is the iteration index.

The fuzzy system $\tilde{R}_i$ can now be defined by the fuzzy sets and rules as follows;

$$\underline{r = \tilde{R}(e,d)}$$

$$\begin{array}{ll} \text{if } e \text{ is } \tilde{E}_1 \text{ and } d \text{ is } \tilde{D}_1 \text{ then } r = R_1 \\ . & . \\ . & . \\ . & . \\ \text{if } e \text{ is } \tilde{E}_M \text{ and } d \text{ is } \tilde{D}_M \text{ then } r = R_M \end{array} \tag{34}$$

where R_j is a real value in the range of

$$-1 \le R_j \le 1 \;\; ; \;\; j = 1,2,...,M \tag{35}$$

The relation of r with e and d can also be given by the following fuzzy inference system

$$r = \tilde{R}(e,d) = \frac{\sum_{j=1}^{M} \mu E_j(e) \cdot \mu D_j(d) R_j}{\sum_{j=1}^{M} \mu E_j(e) \cdot \mu D_j(d)} \tag{36}$$

where μE_j and μD_j are the membership functions of E_j and D_j, respectively.

5.2 Updating the coordination parameters

To update the coordination parameters, we use a fuzzy system that calculates the variation of the coordination parameters as follows

$$\Delta \underline{\beta}[k] = S(\underline{e}[k], \underline{d}[k]) \tag{37}$$

where S is a fuzzy system based on Takagi-Sugeno-Kang (TSK) model [18], [19] and in this case, is defined by the fuzzy sets and rules as follows;

$$\underline{\underline{s} = S(\underline{e}, \underline{d})}$$

$$\begin{array}{ll} \text{if } \underline{e} \text{ is } A_1 \text{ and } \underline{d} \text{ is } B_1 \text{ then } \underline{s} = a_1 \underline{e} + b_1 \underline{d} + c_1 \underline{v} \\ . & . \\ . & . \\ . & . \\ \text{if } \underline{e} \text{ is } A_N \text{ and } \underline{d} \text{ is } B_N \text{ then } \underline{s} = a_N \underline{e} + b_N \underline{d} + c_N \underline{v} \end{array} \tag{38}$$

where

$$\underline{v} = [\ \overbrace{1\ 1\ ....\ 1}^{m}\ ]^T . \tag{39}$$

Also A_j and B_j are the m dimensional fuzzy sets, expressed as

$$A_j = A_{j1} \times A_{j2} \times ... \times A_{jm} \tag{40}$$

$$B_j = B_{j1} \times B_{j2} \times ... \times B_{jm} \tag{41}$$

where their membership functions are given by

$$\mu A_j(\underline{e}) = \mu A_{j1}(e_1) \cdot \mu A_{j2}(e_2) \cdot \cdots \cdot \mu A_{jm}(e_m) \tag{42}$$

$$\mu B_j(\underline{d}) = \mu B_{j1}(d_1) \cdot \mu B_{j2}(d_2) \cdot \cdots \cdot \mu B_{jm}(d_m) \tag{43}$$

also

$$\underline{e} = [e_1, e_2, ..., e_m]^T \tag{44}$$

$$\underline{d} = [d_1, d_2, ..., d_m]^T , \tag{45}$$

where μA_{jk} and μB_{jk} are the membership functions of A_{jk} and B_{jk}, respectively. Moreover, a_j , b_j and c_j are real constant parameters.
To summarize, the relation of $\underline{s}$ with $\underline{e}$ and $\underline{d}$ is given by the following fuzzy inference system;

$$\underline{s} = S(\underline{e}, \underline{d}) = \frac{\sum_{j=1}^{N} \mu A_j(\underline{e}) \cdot \mu B_j(\underline{d}).\ (a_j \underline{e} + b_j \underline{d} + c_j \underline{v})}{\sum_{j=1}^{N} \mu A_j(\underline{e}) \cdot \mu B_j(\underline{d})} \tag{46}$$

5.3 Training the fuzzy goal coordination system

The aim of training is to minimize the energy of the critic vector related to the system parameters; a_j , b_j and c_j , where

$$E = \frac{1}{2} \sum_{k=0}^{n+1} \underline{r}[k]^T \underline{r}[k] \tag{47}$$

also

$$\underline{r}[k] = [\ r_1[k], r_2[k], ..., r_m[k]\ ]^T \tag{48}$$

Now to update the fuzzy system parameters, we use the following updating rule

$$\Delta W = -\eta \frac{\partial E}{\partial W} = -\eta \sum_{k=0}^{n+1} \frac{\partial \underline{r}[k]}{\partial W}^T .\underline{r}[k] \tag{49}$$

where η is the training rate coefficient, and W can be considered as each of the fuzzy system parameters, given by

$$W = a_j, b_j, c_j \quad ; \quad j = 1,2,...,N \tag{50}$$

Now, using the chain rule, we can write

$$\frac{\partial \underline{r}[k]}{\partial W} = \frac{\partial \underline{r}[k]}{\partial \underline{e}[k]} \cdot \frac{\partial \underline{e}[k]}{\partial W} + \frac{\partial \underline{r}[k]}{\partial \underline{d}[k]} \cdot \frac{\partial \underline{d}[k]}{\partial W} \tag{51}$$

So to calculate the right side of this equation, we need to calculate $\frac{\partial \underline{r}[k]}{\partial \underline{e}[k]}, \frac{\partial \underline{e}[k]}{\partial W}, \frac{\partial \underline{r}[k]}{\partial \underline{d}[k]}$ and $\frac{\partial \underline{d}[k]}{\partial W}$, where

$$r_i[k] = \tilde{R}(e_i[k], d_i[k]) = \frac{\sum_{j=1}^{M} \mu E_j(e_i[k]).\mu D_j(d_i[k]) R_j}{\sum_{j=1}^{M} \mu E_j(e_i[k]).\mu D_j(d_i[k])} \triangleq \frac{NUM_i}{DEN_i} . \tag{52}$$

Hence, we have

$$\frac{\partial\, r_i[k]}{\partial\, e_i[k]} = \frac{\sum_{j=1}^{M} \mu E_j'(e_i[k]).\mu D_j(d_i[k]) R_j}{DEN_i} - \frac{NUM_i \sum_{j=1}^{M} \mu E_j'(e_i[k]) \mu D_j(d_i[k])}{DEN_i^{\,2}} \tag{53}$$

and

$$\frac{\partial\, r_i[k]}{\partial\, d_i[k]} = \frac{\sum_{j=1}^{M} \mu E_j(e_i[k]) \mu D_j'(d_i[k]) R_j}{DEN_i} - \frac{NUM_i \sum_{j=1}^{M} \mu E_j(e_i[k]).\mu D_j'(d_i[k])}{DEN_i^2} \tag{54}$$

where $\mu E_j'(.)$ and $\mu D_j'(.)$ denote the derivatives of the corresponding membership functions, respectively. Therefore

$$\frac{\partial \underline{r}[k]}{\partial \underline{e}[k]} = \begin{bmatrix} \frac{\partial r_1[k]}{\partial e_1[k]} & & & 0 \\ & \frac{\partial r_2[k]}{\partial e_2[k]} & & \\ & & \ddots & \\ 0 & & & \frac{\partial r_m[k]}{\partial e_m[k]} \end{bmatrix} \tag{55}$$

and

$$\frac{\partial \underline{r}[k]}{\partial \underline{d}[k]} = \begin{bmatrix} \frac{\partial r_1[k]}{\partial d_1[k]} & & & 0 \\ & \frac{\partial r_2[k]}{\partial d_2[k]} & & \\ & & \ddots & \\ 0 & & & \frac{\partial r_m[k]}{\partial d_m[k]} \end{bmatrix} \tag{56}$$

The gradient of the interaction errors related to the system parameters is also given by

$$\frac{\partial \underline{e}[k]}{\partial W} = \frac{\partial \underline{e}[k]}{\partial \underline{\beta}[k]} \cdot \frac{\partial \underline{\beta}[k]}{\partial W} = D[k] \frac{\partial \Delta \underline{\beta}[k]}{\partial W} \triangleq T[k] \tag{57}$$

where $D[k] \triangleq \frac{\partial \underline{e}[k]}{\partial \underline{\beta}[k]}$, as given in Appendix.

Now in order to calculate $\frac{\partial \Delta \underline{\beta}[k]}{\partial W}$, since we have

$$\Delta \underline{\beta}[k] = S(\underline{e}[k], \underline{d}[k]) \triangleq \frac{NUM}{DEN} = \frac{\sum_{j=1}^{N} \mu A_j(\underline{e}[k]) \mu B_j(\underline{d}[k])(a_j \underline{e}[k] + b_j \underline{d}[k] + c_j \underline{v})}{\sum_{j=1}^{N} \mu A_j(\underline{e}[k]) \mu B_j(\underline{d}[k])} \tag{58}$$

Thus we get

$$\frac{\partial \Delta \underline{\beta}[k]}{\partial a_j} = \frac{\mu A_j(\underline{e}[k]) . \mu B_j(\underline{d}[k]) \underline{e}[k]}{DEN} \tag{59}$$

$$\frac{\partial \Delta \underline{\beta}[k]}{\partial b_j} = \frac{\mu A_j(\underline{e}[k]) . \mu B_j(\underline{d}[k]) \underline{d}[k]}{DEN} \tag{60}$$

$$\frac{\partial \Delta \underline{\beta}[k]}{\partial c_j} = \frac{\mu A_j(\underline{e}[k]) . \mu B_j(\underline{d}[k]) \underline{v}}{DEN} , \tag{61}$$

where $\underline{e}[k]$ and $\underline{d}[k]$ are the values in the previous iteration i.e., $\underline{e}[k]^{(l-1)}$ and $\underline{d}[k]^{(l-1)}$.

Now to calculate $\frac{\partial \underline{d}[k]}{\partial W}$, since we have

$$\underline{d}[k]^{(l)} = \underline{e}[k]^{(l)} - \underline{e}[k]^{(l-1)} , \tag{62a}$$

according to the definition of $T[k]$, we get

$$\frac{\partial \underline{d}[k]}{\partial W} = T[k]^{(l)} - T[k]^{(l-1)} \tag{62b}$$

So we can calculate $\frac{\partial E}{\partial W}$.

The fuzzy system parameters can now be updated using the following updating rule

$$\Delta W = -\eta \frac{\partial E}{\partial W} = -\eta \sum_{k=0}^{n+1} \frac{\partial \underline{r}[k]}{\partial W}^T \underline{r}[k] \tag{63a}$$

where

$$W = a_j, b_j, c_j \quad ; \quad j = 1, 2, \ldots, N. \tag{63b}$$

Now, considering W as the fuzzy system parameters, we can update the coordination parameters with the following rule

$$\Delta \underline{\beta}[k] = S(\underline{e}[k], \underline{d}[k]; W) \tag{64}$$

where W is the updated value, given by

$$W^{(l+1)} = W^{(l)} + \Delta W^{(l+1)} \tag{65}$$

Thus, we can write

$$\begin{aligned} \underline{\beta}[k]^{(l+1)} &= \underline{\beta}[k]^{(l)} + \Delta \underline{\beta}[k]^{(l+1)} \\ &= \underline{\beta}[k]^{(l)} + S(\underline{e}[k]^{(l)}, \underline{d}[k]^{(l)}; W^{(l+1)}) \end{aligned} \tag{66}$$

The various steps of the new coordination algorithm based on Interaction Balance Principle using Fuzzy Goal Coordination System based Reinforcement Learning, can now be summarized as follows:

1. Choose initial values for $\underline{\beta}$ and W.
2. Solve the first level sub-problems using neuro-regulators (or the gradient technique, as described in [14], [15]).
3. Calculate the gradient matrices $\frac{\partial \underline{e}}{\partial \underline{\beta}}$ and $D[k]$. Then update W and consequently update the coordination parameters $\underline{\beta}$, using the Fuzzy Goal Coordination System.
4. Calculate the sum-squared error. If it is smaller than a small constant stop the algorithm, else go to step (2).

The new goal coordination strategy based on Fuzzy Goal Coordination System, Neural Modeling, Neuro-Regulators and Reinforcement Learning is shown in Fig. 2.

6. Simulation results

The application of this approach is demonstrated on an open-loop power system, consisting of a synchronous machine connected to an infinite bus bar through a transformer and a transmission line. For this system, Iyer and Cory [20], [4] have derived a sixth order non-linear dynamical model. The optimization problem is to minimize a cost function of the following form

$$J = \frac{T}{2} \sum_{k=0}^{K_f - 1} \left[\left\| (\underline{X}[k] - \underline{X}_f) \right\|_Q^2 + \left\| (\underline{U}[k] - \underline{U}_f) \right\|_R^2 \right] \tag{67}$$

where Q and R are the weighting matrices, with appropriate dimensions and definiteness. Now, the system can be decomposed into two sub-systems of orders 4 and 2, respectively, using the following sate vectors

$$\underline{X}_1[k] = [X_1[k] \quad X_2[k] \quad X_3[k] \quad X_4[k]]^T \tag{68}$$

$$\underline{X}_2[k] = [X_5[k] \quad X_6[k]]^T , \tag{69}$$

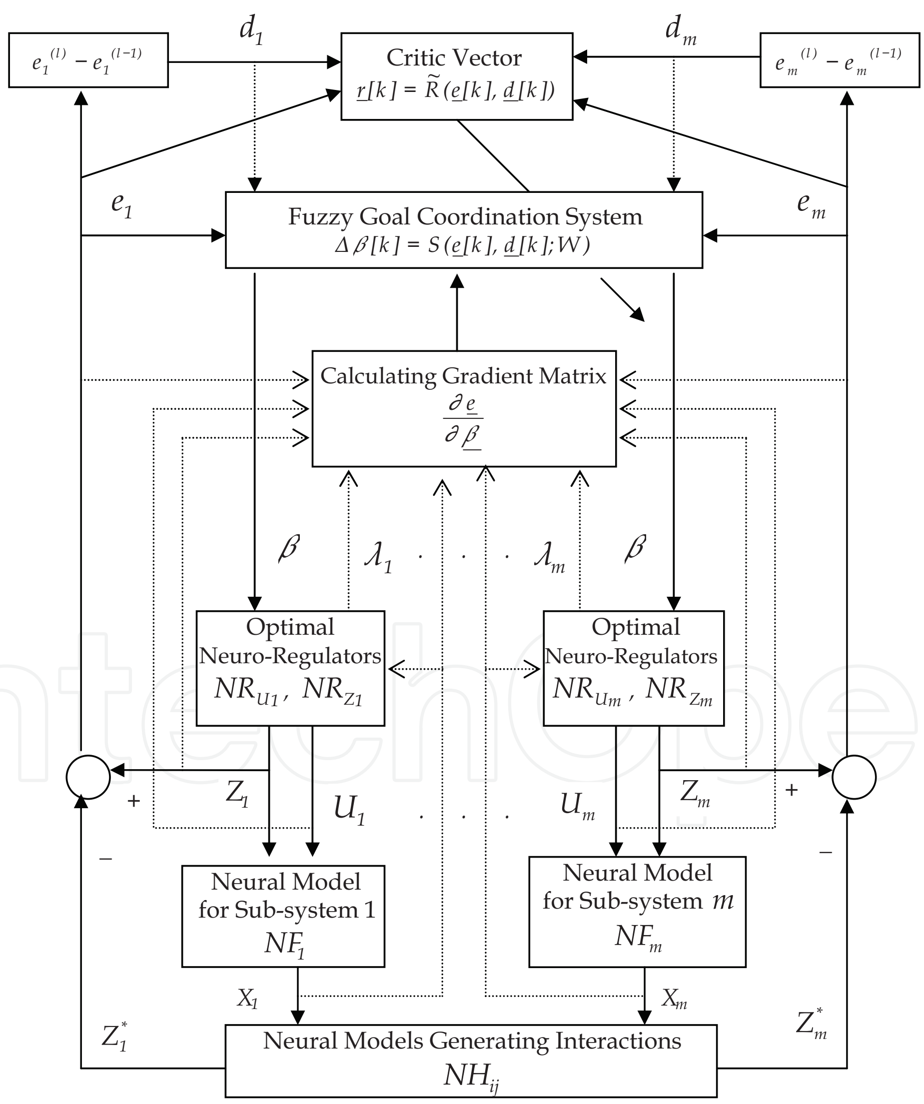


Fig. 2. Intelligent goal coordination strategy based on fuzzy goal coordination system and reinforcement learning.

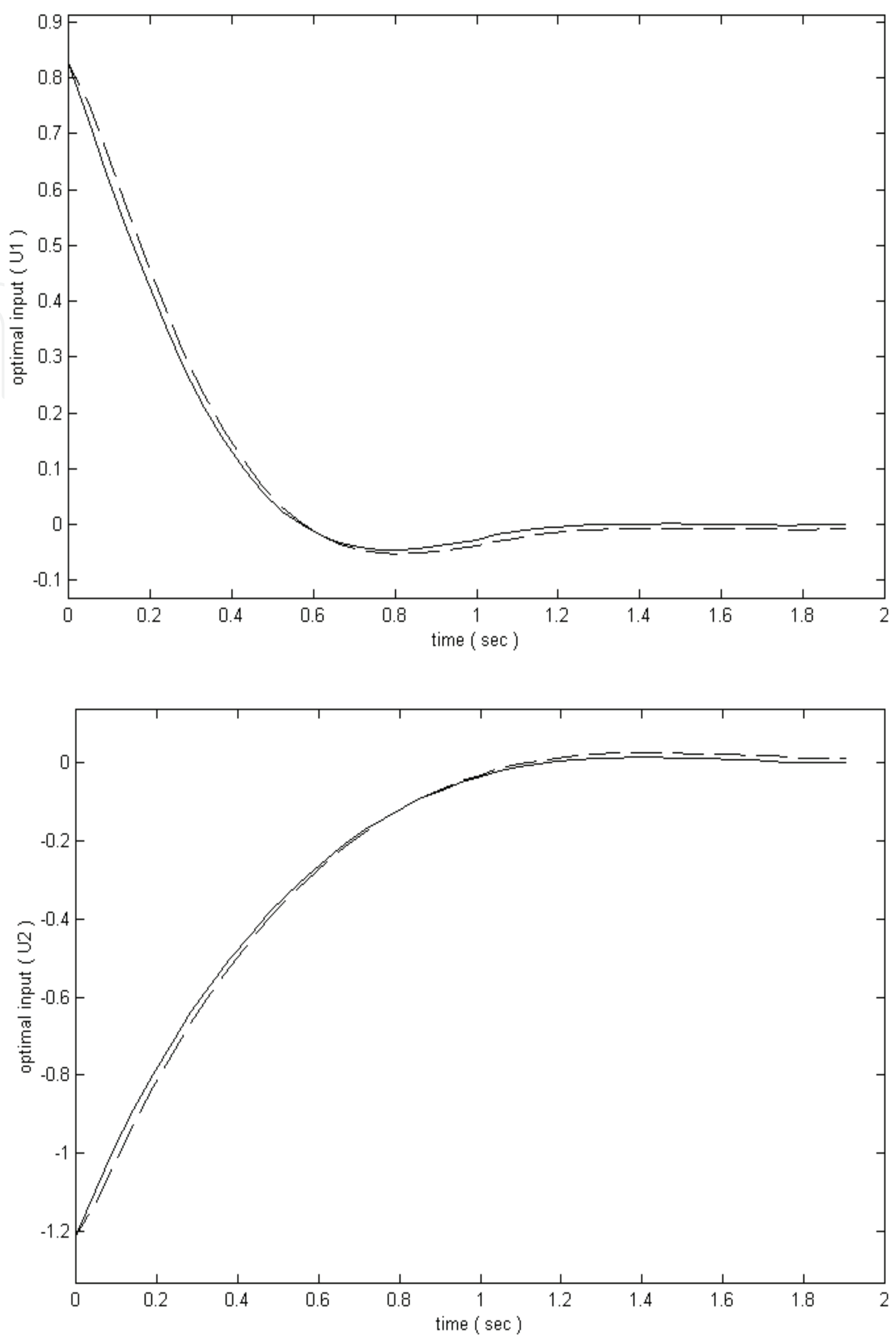


Fig. 3. Optimal control actions of sub-systems 1 and 2. Solid: The new goal coordination approach; Dot: classical method.

and four neural networks, as represented below, to model these two sub-systems and their interaction generators

$$\underline{X}_1[k+1] \underline{\underline{\Delta}} NF_1(\underline{X}_1[k], U_1[k], \underline{Z}_1[k], W_{F_1}) \tag{70}$$

$$\underline{X}_2[k+1] \underline{\underline{\Delta}} NF_2(\underline{X}_2[k], U_2[k], \underline{Z}_2[k], W_{F_2}) \tag{71}$$

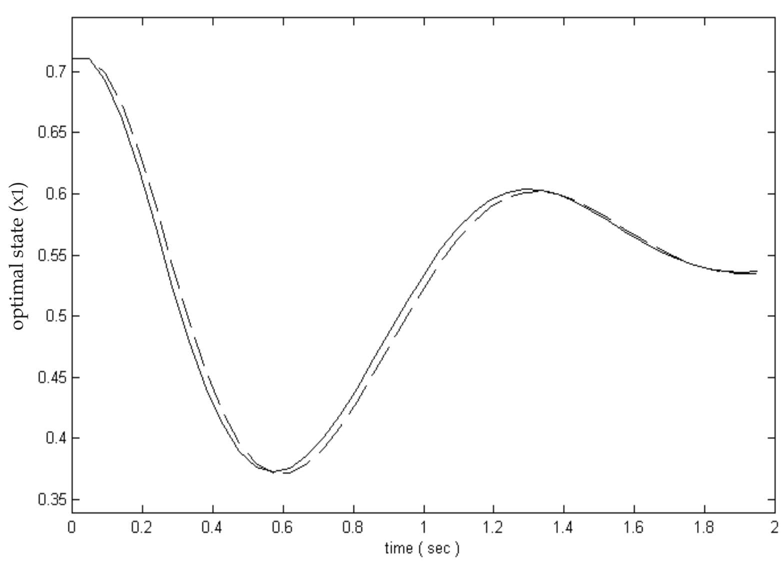
$$\underline{Z}_1[k] \underline{\underline{\Delta}} NH_1(\underline{X}_1[k], \underline{X}_2[k], W_{H_1}) \tag{72}$$

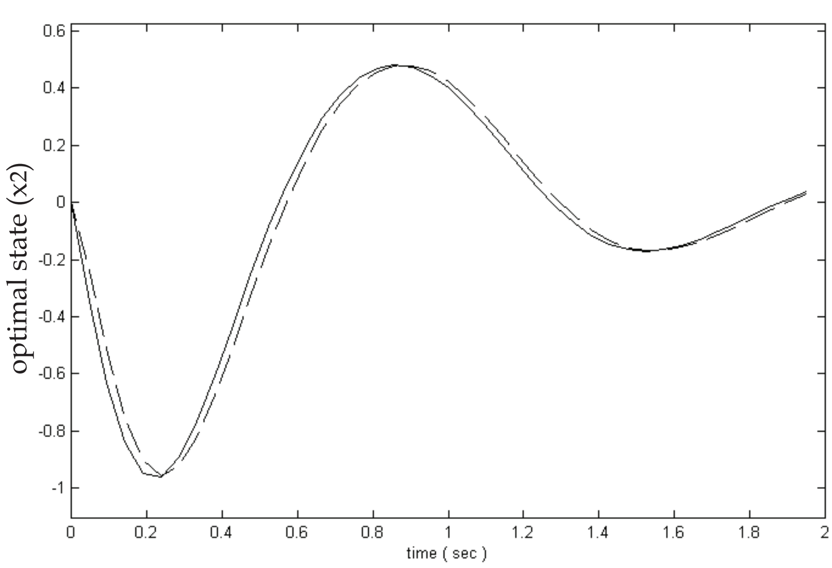
$$\underline{Z}_2[k] \underline{\underline{\Delta}} NH_2(\underline{X}_1[k], \underline{X}_2[k], W_{H_2}) \tag{73}$$

In this example, for describing the fuzzy system $\tilde{R}_i$ defined in (34) and also the Fuzzy Goal Coordination System defined in (38), triangular membership functions are used for fuzzy sets E_j and D_j, and also Gaussian membership functions are used for fuzzy sets A_j and B_j, respectively.

The resulting optimum control actions, state trajectories and the plot of the norm of the interaction errors, using the proposed approach and the classical goal coordination method are all shown in Figs. 3-5.

The results of using the goal coordination approach based on the proposed intelligent coordination strategy shows that the interaction errors vanish rapidly. The advantage of this method is its faster convergence rate in compare to the classical method. This is mainly because of using the new strategy which the update of the coordination parameters directly causes the reduction of the coordination error with the fuzzy goal coordination system based reinforcement learning.





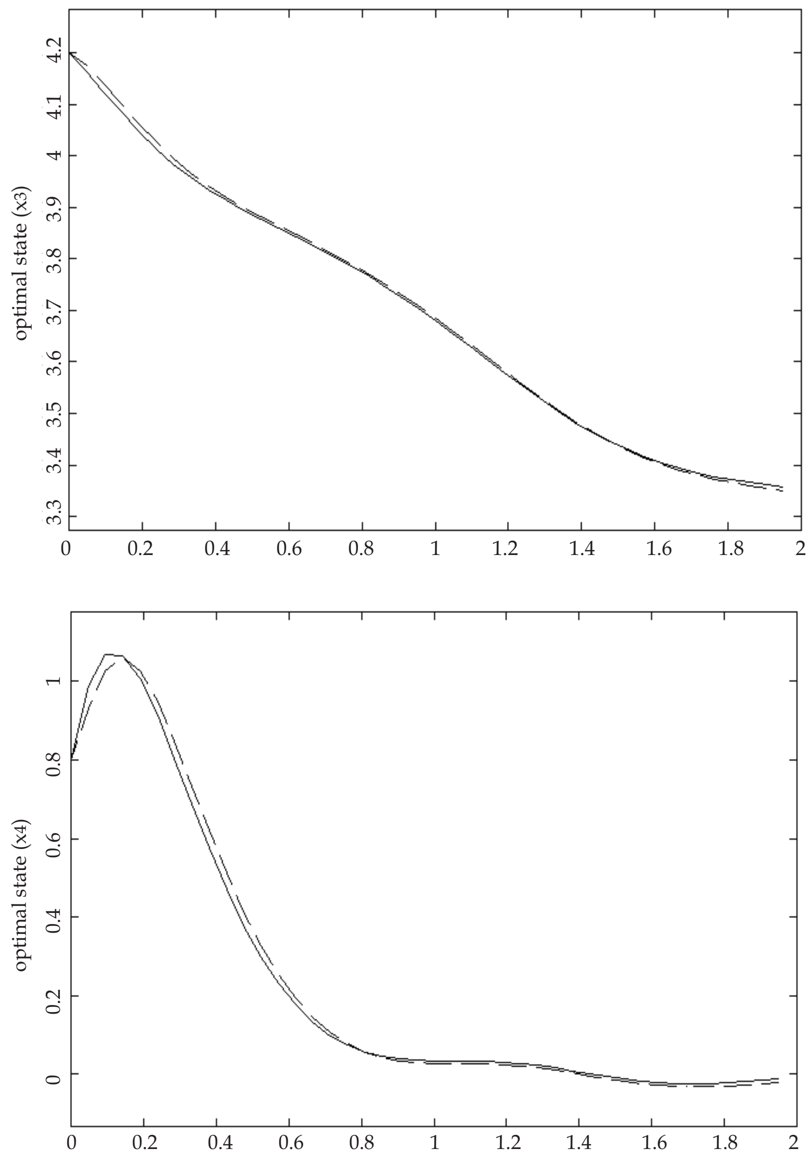


Fig. 4. Optimal state trajectories of sub-system 1. Solid: The new goal coordination approach; Dot: classical method.

7. Conclusion

In this chapter, a new intelligent approach for goal coordination of two-level large-scale control systems is presented. At the first level, sub-systems are modelled using neural networks, while the corresponding sub-problems are solved using neuro-regulators. Fuzzy Goal Coordination System based Reinforcement Learning is also used at the second level, to coordinate the overall large-scale control system. The fuzzy goal coordination system learns its dynamics through minimization of an energy function defined by a critic vector. The minimization process is done using the gradient of interaction errors, while in addition, both the critic vector and fuzzy goal coordination system use the variation of errors (rate of errors) to update their parameters.
As it can be seen, the proposed goal coordination approach, in compare to the classical one, results in much faster reduction of the interaction prediction errors.

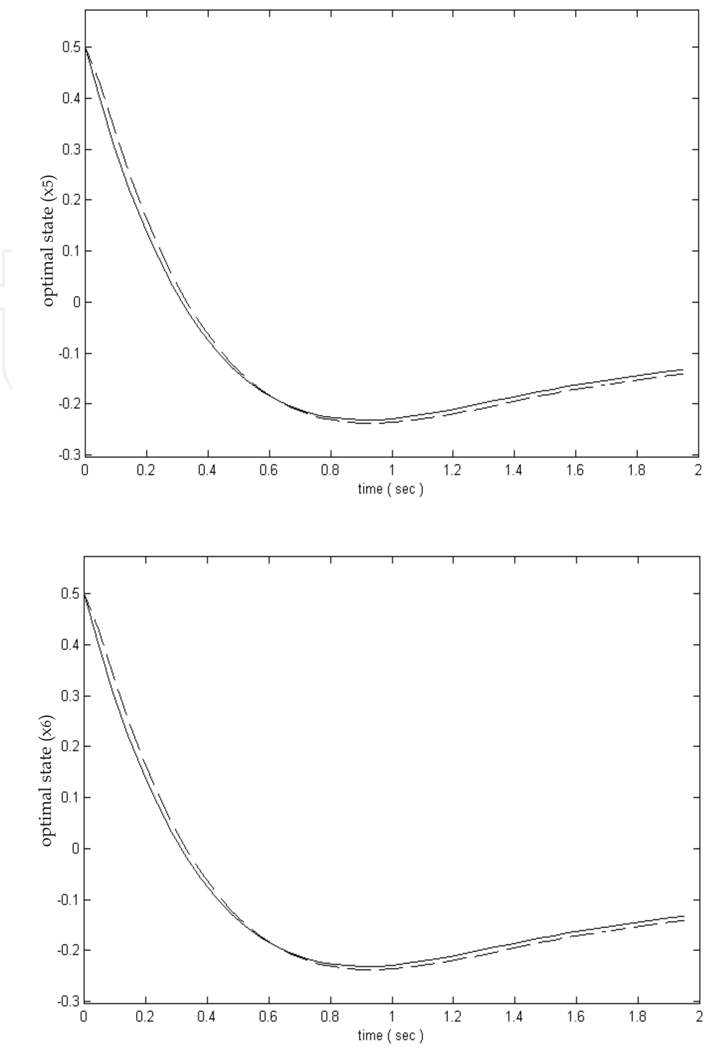


Fig. 5. Optimal state trajectories of sub-system 2. Solid: The new goal coordination approach; Dot: classical method.

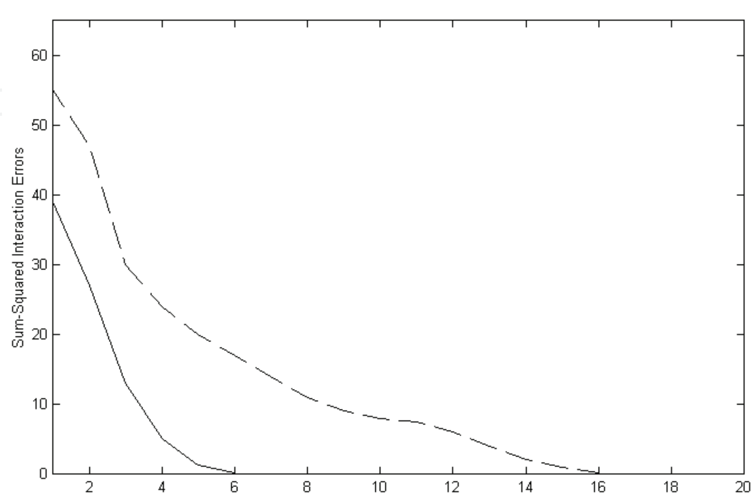


Fig. 6. Comparison between the norm of interaction errors using the proposed approach and the classical method. Solid: The new goal coordination approach; Dot: classical method.

With an extended version of the model coordination approach presented in [21], and the proposed goal coordination strategy of this chapter, the interaction prediction approach (mixed method) [22], [23], can also be extended to a new intelligent interaction prediction strategy.

8. Appendix

In the sequel, the elements of the matrix $D \triangleq \frac{\partial \underline{e}}{\partial \underline{\beta}}$ will be calculated

$$\frac{\partial \underline{e}}{\partial \underline{\beta}} = \frac{\partial \underline{Z}}{\partial \underline{\beta}} - \frac{\partial \underline{Z}^*}{\partial \underline{\beta}} \tag{74}$$

where,

$$\underline{e} = \begin{bmatrix} \underline{e}_1 \\ \underline{e}_2 \\ \vdots \\ \underline{e}_m \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \\ \vdots \\ \underline{\beta}_m \end{bmatrix}, \quad \underline{Z} = \begin{bmatrix} \underline{Z}_1 \\ \underline{Z}_2 \\ \vdots \\ \underline{Z}_m \end{bmatrix}, \quad \underline{Z}^* = \begin{bmatrix} \underline{Z}_1^* \\ \underline{Z}_2^* \\ \vdots \\ \underline{Z}_m^* \end{bmatrix} \tag{75}$$

and

$$\underline{e}_i = \begin{bmatrix} \underline{e}_i[0] \\ \underline{e}_i[1] \\ \vdots \\ \underline{e}_i[n+1] \end{bmatrix}, \quad \underline{Z}_i = \begin{bmatrix} \underline{Z}_i[0] \\ \underline{Z}_i[1] \\ \vdots \\ \underline{Z}_i[n+1] \end{bmatrix}, \quad \underline{Z}_i^* = \begin{bmatrix} \underline{Z}_i^*[0] \\ \underline{Z}_i^*[1] \\ \vdots \\ \underline{Z}_i^*[n+1] \end{bmatrix}, \quad \underline{\beta}_i = \begin{bmatrix} \underline{\beta}_i[0] \\ \underline{\beta}_i[1] \\ \vdots \\ \underline{\beta}_i[n+1] \end{bmatrix} \tag{76}$$

$$\underline{X}_i = \begin{bmatrix} \underline{X}_i[1] \\ \underline{X}_i[2] \\ \vdots \\ \underline{X}_i[n+1] \end{bmatrix}, \quad \underline{U}_i = \begin{bmatrix} \underline{U}_i[0] \\ \underline{U}_i[1] \\ \vdots \\ \underline{U}_i[n] \end{bmatrix}, \quad \underline{\lambda}_i = \begin{bmatrix} \underline{\lambda}_i[0] \\ \underline{\lambda}_i[1] \\ \vdots \\ \underline{\lambda}_i[n] \end{bmatrix} \tag{77}$$

Now using the optimization of the first level, we have

$$\underline{L}_i^x = \frac{\partial L_i}{\partial \underline{X}_i} = 0, \quad \underline{L}_i^u = \frac{\partial L_i}{\partial \underline{U}_i} = 0, \quad \underline{L}_i^\lambda = \frac{\partial L_i}{\partial \underline{\lambda}_i} = 0, \quad \underline{L}_i^z = \frac{\partial L_i}{\partial \underline{Z}_i} = 0 \tag{78}$$

where the corresponding variations can also be written as

$$\begin{cases} \frac{\partial \underline{L}_i^x}{\partial \underline{X}_i}\delta \underline{X}_i + \frac{\partial \underline{L}_i^x}{\partial \underline{U}_i}\delta \underline{U}_i + \frac{\partial \underline{L}_i^x}{\partial \underline{\lambda}_i}\delta \underline{\lambda}_i + \frac{\partial \underline{L}_i^x}{\partial \underline{Z}_i}\delta \underline{Z}_i + \frac{\partial \underline{L}_i^x}{\partial \underline{\beta}}\delta \underline{\beta} = 0 \\ \frac{\partial \underline{L}_i^u}{\partial \underline{X}_i}\delta \underline{X}_i + \frac{\partial \underline{L}_i^u}{\partial \underline{U}_i}\delta \underline{U}_i + \frac{\partial \underline{L}_i^u}{\partial \underline{\lambda}_i}\delta \underline{\lambda}_i + \frac{\partial \underline{L}_i^u}{\partial \underline{Z}_i}\delta \underline{Z}_i + \frac{\partial \underline{L}_i^u}{\partial \underline{\beta}}\delta \underline{\beta} = 0 \\ \frac{\partial \underline{L}_i^\lambda}{\partial \underline{X}_i}\delta \underline{X}_i + \frac{\partial \underline{L}_i^\lambda}{\partial \underline{U}_i}\delta \underline{U}_i + \frac{\partial \underline{L}_i^\lambda}{\partial \underline{\lambda}_i}\delta \underline{\lambda}_i + \frac{\partial \underline{L}_i^\lambda}{\partial \underline{Z}_i}\delta \underline{Z}_i + \frac{\partial \underline{L}_i^\lambda}{\partial \underline{\beta}}\delta \underline{\beta} = 0 \\ \frac{\partial \underline{L}_i^z}{\partial \underline{X}_i}\delta \underline{X}_i + \frac{\partial \underline{L}_i^z}{\partial \underline{U}_i}\delta \underline{U}_i + \frac{\partial \underline{L}_i^z}{\partial \underline{\lambda}_i}\delta \underline{\lambda}_i + \frac{\partial \underline{L}_i^z}{\partial \underline{Z}_i}\delta \underline{Z}_i + \frac{\partial \underline{L}_i^z}{\partial \underline{\beta}}\delta \underline{\beta} = 0 \end{cases} \tag{79}$$

or equivalently,

$$\begin{cases} \underline{L}_i^{xx}\delta\underline{X}_i + \underline{L}_i^{xu}\delta\underline{U}_i + \underline{L}_i^{x\lambda}\delta\underline{\lambda}_i + \underline{L}_i^{xz}\delta\underline{Z}_i + \underline{L}_i^{x\beta}\delta\underline{\beta} = 0 \\ \underline{L}_i^{ux}\delta\underline{X}_i + \underline{L}_i^{uu}\delta\underline{U}_i + \underline{L}_i^{u\lambda}\delta\underline{\lambda}_i + \underline{L}_i^{uz}\delta\underline{Z}_i + \underline{L}_i^{u\beta}\delta\underline{\beta} = 0 \\ \underline{L}_i^{\lambda x}\delta\underline{X}_i + \underline{L}_i^{\lambda u}\delta\underline{U}_i + \underline{L}_i^{\lambda\lambda}\delta\underline{\lambda}_i + \underline{L}_i^{\lambda z}\delta\underline{Z}_i + \underline{L}_i^{\lambda\beta}\delta\underline{\beta} = 0 \\ \underline{L}_i^{zx}\delta\underline{X}_i + \underline{L}_i^{zu}\delta\underline{U}_i + \underline{L}_i^{z\lambda}\delta\underline{\lambda}_i + \underline{L}_i^{zz}\delta\underline{Z}_i + \underline{L}_i^{z\beta}\delta\underline{\beta} = 0 \end{cases} \quad (80)$$

which can be summarized as

$$\begin{bmatrix} \underline{L}_i^{xx} & \underline{L}_i^{xu} & \underline{L}_i^{x\lambda} & \underline{L}_i^{xz} \\ \underline{L}_i^{ux} & \underline{L}_i^{uu} & \underline{L}_i^{u\lambda} & \underline{L}_i^{uz} \\ \underline{L}_i^{\lambda x} & \underline{L}_i^{\lambda u} & \underline{L}_i^{\lambda\lambda} & \underline{L}_i^{\lambda z} \\ \underline{L}_i^{zx} & \underline{L}_i^{zu} & \underline{L}_i^{z\lambda} & \underline{L}_i^{zz} \end{bmatrix} \begin{bmatrix} \delta\underline{X}_i \\ \delta\underline{U}_i \\ \delta\underline{\lambda}_i \\ \delta\underline{Z}_i \end{bmatrix} = -\begin{bmatrix} \underline{L}_i^{x\beta} \\ \underline{L}_i^{u\beta} \\ \underline{L}_i^{\lambda\beta} \\ \underline{L}_i^{z\beta} \end{bmatrix} \delta\underline{\beta} \quad (81)$$

where

$$\begin{cases} \underline{L}_i^{x\beta} = \begin{bmatrix} -\underline{H}_{1i}^{xT} & \cdots & -\underline{H}_{mi}^{xT} \end{bmatrix} \;,\; \underline{L}_i^{u\beta} = \begin{bmatrix} -\underline{H}_{1i}^{uT} & \cdots & -\underline{H}_{mi}^{uT} \end{bmatrix} \\ \underline{L}_i^{\lambda\beta} = 0 \;,\; \underline{L}_i^{z\beta} = \begin{bmatrix} \underline{L}_{i1}^{z\beta} & \cdots & \underline{L}_{im}^{z\beta} \end{bmatrix} \;,\; \underline{L}_{ij}^{z\beta} = \begin{cases} I & ; \; i = j \\ 0 & ; \; i \neq j \end{cases} \end{cases} \quad (82)$$

and

$$\begin{cases} \underline{L}_i^{xx} = M_{i31} \\ \underline{L}_i^{uu} = M_{i22} \\ \underline{L}_i^{\lambda\lambda} = 0 \\ \underline{L}_i^{zz} = -D_{Zi} \\ \underline{H}_{ji}^{x} = T_{\beta i2j}{}^{T} \\ \underline{H}_{ji}^{u} = T_{\beta i3j}{}^{T} \end{cases} , \quad \begin{cases} \underline{L}_i^{\lambda x} = \underline{L}_i^{x\lambda T} = M_{i11} \\ \underline{L}_i^{ux} = \underline{L}_i^{xuT} = M_{i21} \\ \underline{L}_i^{\lambda u} = \underline{L}_i^{u\lambda T} = M_{i12} \\ \underline{L}_i^{\lambda z} = \underline{L}_i^{z\lambda T} = -T_{Zi11} \\ \underline{L}_i^{uz} = \underline{L}_i^{zuT} = -T_{Zi21} \\ \underline{L}_i^{xz} = \underline{L}_i^{zxT} = -T_{Zi31} \end{cases} \quad (83)$$

In (83), the matrices M_{i11}, M_{i12}, M_{i21}, M_{i22}, M_{i31}, T_{Zi11}, T_{Zi21}, T_{Zi31}, $T_{\beta i2j}$, $T_{\beta i3j}$ and D_{Zi} can be find in Appendix of [22].
Now using (7), we can write

$$\delta\underline{Z}_i^* = \sum_{j=1}^{m} \begin{bmatrix} \underline{H}_{ij}^{x} & \underline{H}_{ij}^{u} & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta\underline{X}_j \\ \delta\underline{U}_j \\ \delta\underline{\lambda}_j \\ \delta\underline{Z}_j \end{bmatrix} \quad (84)$$

also $\delta\underline{Z}_i$ can be written as

$$\delta \underline{Z}_i = \begin{bmatrix} 0 & 0 & 0 & I \end{bmatrix} \begin{bmatrix} \delta \underline{X}_i \\ \delta \underline{U}_i \\ \delta \underline{\lambda}_i \\ \delta \underline{Z}_i \end{bmatrix} \tag{85}$$

Therefore, by using the following definitions

$$\underline{V} \triangleq \begin{bmatrix} \underline{V}_1 \\ \underline{V}_2 \\ \vdots \\ \underline{V}_m \end{bmatrix}, \quad \underline{V}_i \triangleq \begin{bmatrix} \underline{X}_i \\ \underline{U}_i \\ \underline{\lambda}_i \\ \underline{Z}_i \end{bmatrix} \tag{86}$$

$$\begin{cases} \underline{L}_i^{vv} \triangleq \begin{bmatrix} \underline{L}_i^{xx} & \underline{L}_i^{xu} & \underline{L}_i^{x\lambda} & \underline{L}_i^{xz} \\ \underline{L}_i^{ux} & \underline{L}_i^{uu} & \underline{L}_i^{u\lambda} & \underline{L}_i^{uz} \\ \underline{L}_i^{\lambda x} & \underline{L}_i^{\lambda u} & \underline{L}_i^{\lambda\lambda} & \underline{L}_i^{\lambda z} \\ \underline{L}_i^{zx} & \underline{L}_i^{zu} & \underline{L}_i^{z\lambda} & \underline{L}_i^{zz} \end{bmatrix}, \quad \underline{L}_i^{v\beta} \triangleq \begin{bmatrix} \underline{L}_i^{x\beta} \\ \underline{L}_i^{u\beta} \\ \underline{L}_i^{\lambda\beta} \\ \underline{L}_i^{z\beta} \end{bmatrix} \\ \underline{H}_{ij}^v \triangleq \begin{bmatrix} \underline{H}_{ij}^x & \underline{H}_{ij}^u & 0 & 0 \end{bmatrix}, \quad \underline{H}_i^v \triangleq \begin{bmatrix} \underline{H}_{i1}^v & \cdots & \underline{H}_{im}^v \end{bmatrix}, \quad \underline{I}_i^v \triangleq \begin{bmatrix} 0 & 0 & 0 & I \end{bmatrix} \end{cases} \tag{87}$$

and using (81), (84) and (85), for each subsystem, we obtain

$$\begin{cases} \underline{L}_i^{vv} \delta \underline{V}_i = -L_i^{v\beta} \delta \underline{\beta} \\ \delta \underline{Z}_i^* = \sum_{j=1}^{m} \underline{H}_{ij}^v \delta \underline{V}_j = \begin{bmatrix} \underline{H}_{i1}^v & \cdots & \underline{H}_{im}^v \end{bmatrix} \delta \underline{V} = \underline{H}_i^v \delta \underline{V} \\ \delta \underline{Z}_i = \underline{I}_i^v \delta \underline{V}_i \end{cases} \tag{88}$$

Now, for the overall system we have

$$\begin{cases} \underline{L}^{vv} \delta \underline{V} = -L^{v\beta} \delta \underline{\beta} \\ \delta \underline{Z}^* = \underline{H}^v \delta \underline{V} \\ \delta \underline{Z} = \underline{I}^v \delta \underline{V} \end{cases} \tag{89}$$

where

$$\begin{cases} \underline{L}^{vv} \triangleq \begin{bmatrix} \underline{L}_1^{vv} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \underline{L}_m^{vv} \end{bmatrix}, \quad L^{v\beta} \triangleq \begin{bmatrix} L_1^{v\beta} \\ \vdots \\ L_m^{v\beta} \end{bmatrix} \\ \underline{I}^v \triangleq \begin{bmatrix} \underline{I}_1^v & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \underline{I}_m^v \end{bmatrix}, \quad \underline{H}^v \triangleq \begin{bmatrix} \underline{H}_1^v \\ \vdots \\ \underline{H}_m^v \end{bmatrix} \end{cases} \tag{90}$$

and using (89), it can be concluded that

$$\delta \underline{V} = -\left(\underline{L}^{vv}\right)^{-1} L^{v\beta} \delta \underline{\beta} \tag{91}$$

Therefore, by substituting (91) in (89), we obtain

$$\begin{cases} \delta \underline{Z}^* = -\underline{H}^v \left(\underline{L}^{vv}\right)^{-1} L^{v\beta} \delta \underline{\beta} \\ \delta \underline{Z} = -\underline{I}^v \left(\underline{L}^{vv}\right)^{-1} L^{v\beta} \delta \underline{\beta} \end{cases} \tag{92}$$

and as a result

$$\delta \underline{e} = \delta \underline{Z} - \delta \underline{Z}^* = -\left(\underline{I}^v - \underline{H}^v\right)\left(\underline{L}^{vv}\right)^{-1} L^{v\beta} \delta \underline{\beta} \tag{93}$$

Finally,

$$\frac{\partial \underline{e}}{\partial \underline{\beta}} = -\left(\underline{I}^v - \underline{H}^v\right)\left(\underline{L}^{vv}\right)^{-1} L^{v\beta} \tag{94}$$

9. References

[1] M. D. Mesarovic, D. Macko and Y.Takahara, "*Two coordination principles and their application in large-scale systems control*," Pergamon Press, New York, March 1970.
[2] M. D. Mesarovic, D. Macko and Y.Takahara, T*heory of Hierarchical Multilevel Systems*. Academic Press, New York, 1970.
[3] J. D. Pearson, *Dynamic decomposition techniques*, in Wismer, D. A., ed., *Optimization Methods for Large- Scale Systems*. McGraw-Hill, New York, 1971.
[4] M. G. Singh and A.Titly, *Systems*: *Decomposition, Optimisation and Control*. Pergamon Press, Oxford, 1978.
[5] M. G. Singh, *Dynamical Hierarchical Control*. rev. ed., North Holland, Amsterdam, 1980.
[6] M. Jamshidi, *Large-Scale Systems–Modeling and Control*. North Holland, New York, 1983.
[7] M. F. Hassan and M. G. Singh, "*The optimisation of non-linear systems using a new two-level method*," Automatica, 12: 359-363, 1976.
[8] M. F. Hassan and M. G. Singh,"*A two-level costate prediction algorithm for non-linear systems*," Cambridge University, Engineering Dept., Report No. CUEDF-CAMS/TR (124), 1976.
[9] M. S. Mahmoud, "*Multilevel systems control and applications: A survey*," IEEE Trans. Sys. Man. and Cyb., SMC-7:125-143, 1977.
[10] S. M. Mahmoud, M. F. Hassan, M. G. Darwish , *Large-Scale Control Systems: Theories and Techniques*. Marcel Dekker, New York, 1985.
[11] D. C. Siljak, *Large-Scale Dynamic Systems : Stability and Structure*. North-Holland, Amesterdam, 1978.
[12] E. J. Bauman, "*Multi-level optimization techniques with application to trajectory decomposition*," in C. T. Leondes, ed., Advances in Control Systems. Pp. 160-222, 1986.

[13] Y. M. Park, M. S. Choi and K. Y. Lee, "*An optimal tracking neuro-controller for non-linear dynamic systems*," IEEE Trans. Neural Networks, vol.7, no.5, Sept. 1996.

[14] N. Sadati , "*A novel approach to coordination of large-scale systems; Part I* – interaction prediction principle," IEEE Int. Conf. on Industrial Technology, pp. 641-647, 2005.

[15] N. Sadati, "*A novel approach to coordination of large-scale systems; Part II* – interaction balance principle," IEEE Int. Conf. on Industrial Technology, pp. 648-654, 2005.

[16] N. Sadati, A. Babazadeh, "*Optimal control of robot manipulators with a new two-level gradient based approach*," Journal of Electrical Engineering, Springer, 88, pp. 383-393, 2006.

[17] C.T. Lin and C.S.G. Lee, "Reinforcement structure/ parameter learning for neural network-based fuzzy logic control systems," IEEE Trans. on Fuzzy Systems, vol. 2, no. 1, Feb. 1994.

[18] T. Takagi, M. Sugeno, "*Fuzzy identification of systems and its applications to modelling and control,*" IEEE Trans. on Syst. Man and Cyber., vol. 15, 116-132, 1985.

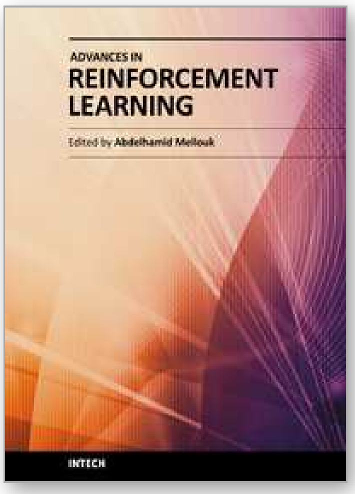
[19] M. Sugeno and G.T. Kang, "Structure identification of fuzzy *model*," Fuzzy Sets and Syst., vol. 28, pp. 15-33, 1988.

[20] S. N. Iyer and B. J. Cory, "*Optimization of turbo–generator transient performance by differential dynamic programming* ," IEEE Trans. Power Appar. Syst., PAS- 90: 2149-2157, 1971.

[21] N. Sadati, M. M. Emamzadeh, "*A novel fuzzy reinforcement learning approach in two-level control of 3-DOF robot manipulators,*" IEEE Symposium on Approximate Dynamic Programming and Reinforcement Learning, Honolulu, Hawaii, April 2007.

[22] N. Sadati, M. H. Ramezani, "*Optimization of large-scale systems using gradient type interaction prediction approach,*" Journal of Electrical Engineering, Springer, vol. 91, Dec. 2009.

[23] N. Sadati, M. H. Ramezani, "*Novel interaction prediction approach to hierarchical control of large-scale systems,*" IET Control Theory Application, vol. 4, issue 2, pp. 228-243, March 2010.



Advances in Reinforcement Learning
Edited by Prof. Abdelhamid Mellouk

ISBN 978-953-307-369-9
Hard cover, 470 pages
Publisher InTech
Published online 14, January, 2011
Published in print edition January, 2011

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821