‘State Déjà vu’ inter-agent learning adaptive control framework

Hongyi Qu* Ridong Zhang** Furong Gao***

*Department of Chemical and Biomolecular Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: hquaa@ust.hk).

**Department of Chemical and Biomolecular Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: kerzhang@ust.hk)

***Department of Chemical and Biomolecular Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: kefgao@ust.hk)

Abstract: An inter-agent learning adaptive control framework is proposed for mass production by using multi-agent system approach to enhance the convergence performance over the single agent control. The idea is to invoke agent-wise differences on estimated parameters of the online estimator in adaptive control. Each agent’s estimator selects estimated parameters from a corresponding ‘best’ agent among them for next iteration according to ‘State Déjà vu’ criterion mimicking psychological phenomenon of adopting experience from state similar to current one. The application of proposed framework on model free adaptive control shows to have robust convergence, good stability, and effective performances enhancement.

Keywords: Nonlinear system, Adaptive control, Parameter estimation, Learning control, Continuous systems, Regulation

1. INTRODUCTION

With the fast growing demand of industrial products over a short period of time, mass production, which refers to the production of a large amount of products by many machines working simultaneously, has become an inexorable trend in industrial application. Corresponding to this trend, process control should be developed accordingly. Efforts have been made by treating each individual machine/process as an agent, resulting in so-called multi-agent control to explore inter-agent information utilization. Multi-agent control has resulted in good applications due to the rapid development of and readily deplorable communication techniques that making intra and inter agent information sharing convenient. The main streams of current multi-agent control consist of distributed control (Camponogara et al., 2002; Wang et al., 2010), synchronization or coordination control (Zhang et al., 2015; Oh et al., 2015), multi-agent cooperative control (Choi and Horowitz, 2009; Chen et al., 2014; Su and Huang, 2012) etc. Those researches deal with the scenario that multiple agents cooperate for accomplishing a single task. For each agent, the information from others are used for adjusting this agent's control target such as set point or trajectory, while the realization of them, such as regulation or tracking, still depends on this agent's own information. This suggests that the state-of-the-art control law itself of each individual machine/process in the current mass production mode remains the same as in the conventional standalone (or single agent) manner in term of control tracking law performance.

The control performance of each single agent control depends on the quality and the way of unitizing information of the machine/process. Significant progresses have been achieved along with the development of many well-known single agent controllers such as PID controller, model predictive controller (MPC), and iterative learning controller (ILC). The single agent control approaches may have limited room for control performance improvement because it uses only local sources of information and uses little information diversity as all information comes from this single own agent. We define the information from own process (or single agent) as intra-agent information and information from other agent(s) as inter-agent information. In a multi agent context, it is obvious inter-agent information may be explored to enhance the diversity of utilizable information, so better control performance of each individual process/machine may be explored as each agent can have accesses to information of not only intra-agent but inter-agent. In this paper, such a kind of control law design is called as inter-agent learning (IAL) control. Inter-agent learning control is a kind of controller design framework in which each agent can utilize intra-agent and inter-agent information from all agents for control performance enchantment.

Different from the current streams of multi-agent control research, this paper focus on the control law design in a common mass production scenario, where multiple identical processes/machines are in operation under the same operation target (set-point) to be achieved. An example of such a mass production is in polymer processing industry,
where many same injection-moulding machines produce the identical plastic parts simultaneously. To develop IAL control scheme, the basic nature of IAL controller need to be figured out first. Motivated by ‘peers learning’ in human society, some differences among these controllers of those identical processes should be embraced. Moreover, each agent should adjust its own control strategy according to the acquired experiences of all other agents in real time. This suggests that the nature of controller should be adaptive, since the online inflexible type of controller such as robust control usually has one unique and unadjustable optimal design which is contradictory to the requirements for IAL stated above.

This paper proposes an inter-agent learning adaptive control of parallel running processes. The idea is to invoke the ‘differences’ among these agents (processes) by assigning different initial guesses of online recursively estimated parameters. During operation, each process chooses the ‘best’ estimated parameters and states generated from last iteration of all these processes, ‘peers’ and itself, according to a defined criterion function, and then iterates from above selected information. Furthermore, this paper designs a ‘naïve’ but effective ‘State Déjà vu’ based criterion function. This makes each process adopt the information from a ‘peer’ whose states of the last time step are closest to those of current time step of itself. The application to a model free adaptive control (MFAC) design of a nonlinear process (Hou et al., 2014) has been given to illustrate the effectiveness of the proposed approach. The corresponding robust convergence and stability are presented for the resulted control law. In addition, simulation is also given with results showing that the control performances after including the proposed inter-agent learning is superior to that of the conventional single agent control as the control performance sensitivity to initial guess has been much reduced.

The rest of this paper is organized as follows. Section 2 describes the proposed framework of inter-agent learning and the specified ‘State Déjà vu’ criterion. In section 3, the original single agent MFAC is briefly reviewed, then the inter-agent learning framework’s embedment is derived, robust convergence and stability are analysed. Simulation illustrations are given in section 4. Section 5 concludes.

2. INTER-AGENT LEARNING CONTROL FRAMEWORK AND ‘STATE DÉJÀ VU’ BASED CRITERION

2.1 General formulation

Without losing generality, this paper considers the following groups of SISO discrete time nonlinear processes operating in parallel:

\[ y(t+1) = f_y(t,i), \ldots, y(t-n_y,i), u(t,i), \ldots, u(t-n_u,i) \]  

where \( y(t,i) \in R \) and \( u(t,i) \in R \) are the process output and the control input of the \( i \)-th sub process at time \( t \) respectively, \( n_y \) and \( n_u \) are unknown integers, and \( f(\ldots) \) is an unknown nonlinear function. This type of processes is good representation of many typical nonlinear and linear processes. The common adaptive controller structure for nonlinear processes with recursive estimator in traditional single agent mode typically consists of three parts:

Model:

\[
y(t+1,i) = f_m[y(t,i),\ldots,y(t-n_y,i),u(t,i),\ldots,u(t-n_u,i),\tilde{\theta}(t,i)]
\]

Recursive Estimator:

\[
\hat{\theta}(t,i) = f_e[x(t,i),\hat{\theta}(t-1,i)]
\]

Feedback Control:

\[
u(t,i) = f_c[y(t,i),\ldots,y(t-n_y,i),u(t-1,i),\ldots,u(t-n_u,i)]
\]

where \( f_m(\ldots) \), \( f_e(\ldots) \) and \( f_c(\ldots) \) are respectively the representations of model, recursive estimator and feedback control law. \( \hat{\theta}(t,i) \) and \( \tilde{\theta}(t,i) \) are the model parameters vector and their corresponding estimates. As stated in introduction, the keys of IAL control scheme design are two-fold: first is to invoke differences among those identical agents, second is to design selection mechanism according to a certain specified criterion. The differences can be made by letting those processes adopt different initial guesses on estimated parameters vector \( \hat{\theta}(t,i) \), and involve inter-agent information selection by changing the recursive estimator into following:

\[
\tilde{\theta}(t,i) = f_{em}[\bar{x}_m(t,i),\bar{x}_m(t^*,i^*),\tilde{\theta}(t^*-1,i^*)]
\]

\[
\bar{x}_m(t,i) = [y(t,i),\ldots,y(t-n_{ym},i),u(t-1,i),\ldots,u(t-n_{um},i)]^T
\]

where \( f_{em}(\ldots) \) is a new estimator formulation designed for the inclusion of state \( \bar{x}_m(t^*-1,i^*) \), \( i^* \) and \( i^* \) are selected by solving following exhaustive search minimization problem:

\[
[t^*-1,i^*] = \arg \min_k c_{r_{ij}}(\bar{x}_{crl^{(i)}}(t,i),\ldots,\bar{x}_{erm^{(j)}}(t,i),\bar{x}_{crl^{(j)}}(k,j),\ldots,\bar{x}_{erm^{(j)}}(k,j))
\]

\[
t-d \leq k \leq t-1, 1 \leq j \leq n
\]

\[
t^* = t, i^* = i \text{ when } t = 2
\]
such as \( x_m(k, j) \) or \( \tilde{\theta}(k, j) \) etc. \( m_i \) and \( m_j \) are the states numbers of current sub process and its peers sub process relatively, \( d \) is the length of time range for selection.

### 2.2 Naïve 'State Déjà vu' based inter-agent learning criterion design

This section presents a simple IAL criterion which will be used throughout the paper. Define:

\[
m_1 = m_2 = 1
\]

\[
\dot{x}_{cr}^{(i)}(t_i) = \left[ y(t_i), \ldots, y(t-n_{c,r,i}), u(t-L_i), \ldots, u(t-n_{u,c,r,i}) \right]^T
\]

\[
\dot{x}_{cr}^{(j)}(k, j) = \left[ y(k, j), \ldots, y(k-n_{c,r,j}), u(k-1, j), \ldots, u(k-n_{u,c,r,j}) \right]^T
\]

Then the criterion function is designed as:

\[
cr_J \left( \dot{x}_{cr}^{(i)}(t_i), \dot{x}_{cr}^{(j)}(k, j) \right) = \left\| \dot{x}_{cr}^{(i)}(t_i) - \dot{x}_{cr}^{(j)}(k, j) \right\|_2
\]

\[ t - d \leq k \leq t - 1, \quad 1 \leq j \leq n \]

Above criterion function is called ‘State Déjà vu’ criterion, such name comes from its physical meaning: the agent considers \( x_c^{(i)}(t', i') \) and \( \dot{\theta}^{(i'-1, i')} \), whose corresponding criterion state is ‘closest’ to the current states, as the best peer’s information to be iterated from by online recursive estimator \( f_{em} \). It is analogous to the scenario that in human society, one is experiencing an event already similarly experienced by someone at sometime in the past history, which meets the definition of ‘Déjà vu’ (has happened) phenomenon.

The logic behind ‘State Déjà vu’ criterion is intuitive: in many recursive estimators such as projection algorithm based, least squares algorithm based etc. the convergence of \( \hat{\theta} \) to \( \theta \) can only be achieved when \( \hat{\theta} \) is time-invariant.

In time varying case, estimation error \( \hat{\theta} - \theta \) converges to certain range around zero and its varying amplitude is in concert with variation of \( \hat{\theta} \). The time varying property of \( \hat{\theta} \) is mainly contributed to its nonlinear relationship to the process variables caused by the differences between model structure \( f_m \) and true nonlinear process dynamic \( f \). ‘State Déjà vu’ criterion just takes the simple thought that if the distance between specified states \( \dot{x}_{cr}^{(i)}(t_i) \) and \( \dot{x}_{cr}^{(j)}(k, j) \) is closer, the corresponding \( \hat{\theta}(k, j) \) more likely equals to \( \hat{\theta}(t, i) \), thus, this criterion is expected to reduce the variation of \( \hat{\theta} \) and estimation error i.e. improve the accuracy of the estimation of time varying parameters.

The following sections of this paper present the application of above ‘State Déjà vu’ based inter-agent learning control framework to a compact form dynamic linearization (CFDL)-based model free adaptive control (MFAC) which has been a well recognized simple, effective adaptive control algorithm for regulation control problem of a nonlinear process.

### 3. CFDL BASED INTER-AGENT LEARNING MODEL FREE ADAPTIVE CONTROL ALGORITHM (IAL-MFAC) WITH ‘STATE DÉJÀ VU’ BASED CRITERION

Before presenting IAL-MFAC, the original MFAC is briefly reviewed first.

#### 3.1 Original MFAC

The MFAC handles the process \( (1) \) that can be transformed into following Compact Form Dynamic Linearization (CFDL) formulation:

\[
\Delta y(t + 1, i) = \theta(t, i) \Delta u(t, i)
\]

and \( \theta(t, i) \) is bounded. The concerning assumptions of process can be seen in (Hou et al., 2014).

Like most of the adaptive control algorithm, MFAC consists of two parts: a parameters recursive online estimator, and a closed-loop controller based on the system parameters estimated by the above estimator. In MFAC, pseudo partial derivative (PPD) is the system parameter need to be estimated recursively online, and the estimation can be carried out by an recursive projection estimator, which obtains the estimated \( \hat{\theta}(t, i) \) by minimizing following cost function:

\[
J(\hat{\theta}(t, i)) = \left\| y(t, i) - \hat{\theta}(t, i) \Delta u(t, i) \right\|^2 + \mu \left\| \hat{\theta}(t, i) - \hat{\theta}(t-1, i) \right\|^2
\]

with respect to \( \hat{\theta}(t, i) \) where \( \mu > 0 \). After taking Assumption 4 into consideration and adding parameter \( 0 < \eta \leq 2 \) to make estimation more flexible, the PPD estimation obtained is:

\[
\hat{\theta}(t, i) = \hat{\theta}(t-1, i) + \frac{\eta \Delta u(t-1, i)}{\mu + \Delta u(t-1, i)} (\Delta y(t, i) - \hat{\theta}(t-1, i) \Delta u(t-1, i))
\]

\[
\hat{\theta}(t, i) = \hat{\theta}(1, i)
\]

if \( \left| \hat{\theta}(t, i) \right| \leq \varepsilon \) or \( \left| \Delta u(t-1, i) \right| \leq \varepsilon \) or sign \( \hat{\theta}(t, i) \) \( \leq 0 \)

where \( \varepsilon \) is a very small positive constant which acts as a threshold indicating the value equals to zero approximately.

Substituting the estimated PPD into following cost function:

\[
J(\hat{u}(t, i)) = \left\| y(t, i) - \hat{\theta}(t, i) \Delta u(t, i) \right\|^2 + \lambda \left\| \Delta u(t, i) \right\|^2
\]

where \( \lambda > 0 \) is the penalty factor. Minimizing \( (11) \) with respect to \( \hat{u}(t, i) \), the controller input can be yielded as:
where $\rho \in (0, 1]$ is added to make the controller more general.

**Remark 1.** MFAC algorithm can be interpreted in form of (2) as:

\[
f_m(\ldots) = y(t, i) + \theta(t, i) \Delta u(t, i)
\]

\[
n_{ym} = 0, n_{ue} = 1, \tilde{\theta} = \theta
\]

\[
f_c(\ldots) = \theta(t-1, i) + \eta \Delta u(t-1, i)
\]

\[
x(t, i) = y(t, i), y(t-1, i), u(t-1, i), u(t-2, i)\]

\[
n_{yc} = 0, n_{ue} = 1, y_s(t+1) = y^*
\]

The equation (10) is omitted for convenience.

**3.2 The application of inter-agent learning control framework on CF DL based MFAC**

Specify the recursive estimator (3) as

\[
\hat{\theta}(t, i) = f_m \left( t, i \right) + \Delta u(t-1, i)
\]

\[
\hat{\theta}(t-1, i) = \theta(t-1, i) + \eta \Delta u(t-1, i)
\]

\[
\hat{x}(t, i) = y(t, i), y(t-1, i), u(t-1, i), u(t-2, i)
\]

\[
n_{yc} = 0, n_{ue} = 2
\]

Notice that above estimator only preserves $\tilde{x}_m(t, i)$, which can be considered as a special case. In this research, the open-loop inter-agent learning mode is adopted. Set the ‘State Déjà vu’ criterion as:

\[
\hat{\theta}(t, i) = \left[ y(t, i), y(t-1, i), u(t-1, i), u(t-2, i) \right]
\]

\[
\hat{x}(t, i) = y(t, i), y(t-1, i), u(t-1, i), u(t-2, i)
\]

\[
n_{yc} = 0, n_{ue} = 2
\]

Then, the formulation of IAL-MFAC can be concluded as follows:

\[
\sum_{IAL-MFAC} u(t, i) = u(t-1, i) + \frac{\rho \hat{\theta}(t, i)}{\lambda + \hat{\theta}(t, i)} (y^* - y(t, i))\]

and the vector $\text{rest} \left( \hat{\theta}(t, i) \right)$ is formed by deleting element $u(t, i)$ from $\hat{\theta}(t, i)$.

Then, the application of inter-agent learning control framework on CF DL based MFAC is:

**Theorem 1.** If a group of nonlinear processes (1) satisfies Assumptions 1-5, and are controlled by the IAL-MFAC algorithm for regulation problem, then there exists a large enough $\lambda$ such that:

a. Output tracking error converges monotonically i.e. $
\lim_{i \to \infty} \left| y^* - y(t, i) \right| = 0.$

b. $u(t, i)$ is bounded, which, along with (a), indicating that the sub-closed-loop system is bounded-input bounded output (BIBO) stable.

**Proof:** Define PPD estimation error as $\hat{\theta}(t, i) = \hat{\theta}(t, i) - \theta(t, i)$. According to (14), the dynamic of $\hat{\theta}(t, i)$ is:

\[
t^* = t, i^* = i \quad \text{when } \Delta t = 2
\]
\[ \hat{\theta}(t,i) = \left(1 - \frac{\eta \Delta u(t^*,t^*)}{\mu + \Delta u(t^*,t^*)} \right) \hat{\theta}(t^*,t^*) + \theta(t^*,t^*) - \theta(t,i) \] (17)

Then it is easy to yield that:
\[ \| \hat{\theta}(t,i) \| = \left[ - \frac{\eta \Delta u(t^*,t^*)^2}{\mu + \Delta u(t^*,t^*)} \right] \| \hat{\theta}(t^*,t^*) \| + \| \theta(t^*,t^*) - \theta(t,i) \| \] (18)

According to the properties of projection based recursive estimator stated in (HOU et al., 2014):
\[ 0 \leq 1 - \frac{\eta \Delta u(t^*,t^*)^2}{\mu + \Delta u(t^*,t^*)} \leq \frac{\bar{d}}{n} < 1 \] (19)

when \( \mu > 0 \) and \( 0 < \eta \leq 2 \). As stated in [ ], \( \theta(t,i) \) is bounded i.e. \( \theta(t,i) \leq \bar{b} \) which can lead to
\[ \| \theta(t^*,t^*) - \theta(t,i) \| \leq 2 \bar{b} \cdot \| \hat{\theta}(t^*,t^*) \| \] is generated by some optimally selected history \( \hat{\theta} \) from some sub process, which indicates that those \( \hat{\theta}(t^*,t^*) \)'s can be view as many special iteration sequences along time axis. Thus, it can be simply concluded as (Hou et.al, 2014) that:
\[ \| \hat{\theta}(t,i) \| \leq \bar{d}^{i-1} \| \hat{\theta}(1,i) \| + \frac{\bar{b} \left( 1 - \bar{d}^{i-1} \right)}{1 - \bar{d}} \] (20)

where \( \| \hat{\theta}(1,i) \| \) is the starting point from some agent. Thus, the absolute estimation error \( \| \hat{\theta}(t,i) \| \) is bounded. The rest of the proof is the same as Theorem 4.1 and Remark 4.6 in (Hou et al., 2014).

**Remark 3.** Above proof reveals the mechanism of ‘State Dèjà vu’ based inter-agent learning criterion on improving accuracy of online recursive estimation. \( \| \theta(t^*,t^*) - \theta(t,i) \| \) is expected to be smaller, and then the integrating PPD estimation error can be reduced.

4. ILLUSTRATIVE EXAMPLES

To verify the effectiveness of IAL-MFAC proposed in this paper, consider following discrete time nonlinear process (Narendra and Parthasarathy, 1990):
\[ y(t + 1, i) = \frac{y(t,i)}{1 + y(t,i)} + u^3(t,i) \] (21)

Set number of identical processes as \( n = 50 \), the final time spot \( N = 80 \). For fair comparison, the common parameters of original MFAC and PL-MFAC are set as the same:

\[ u(0,i) = 0, y(0,i) = 0, \hat{b}(0,i) = 1 + \frac{2(i - 1)}{(n - 1)}, i = 1, \ldots, n \] (22)

\[ \lambda = 22, \rho = 0.6, \eta = 1, \epsilon = 10^{-3} \]

The target for tracking problem is set as:
\[ y^* = 10 \] (23)

For IAL-MFAC, \( \pi_y = 0 \) and \( \pi_u = 0 \). To do the comparison quantitatively, define the average sum of absolute errors (ASAE) as:
\[ ASAE = \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{N} |e(t,i)| \] (24)

Fig.1 and Fig.2 show performances of original MFAC and IAL-MFAC in noise free case respectively. It can be obviously observed that under IAL-MFAC’s manipulation, all the processes finally converge to the set point almost as fast as the best controlled agent in the case of original MFAC. The ASSEs of original MFAC and IAL-MFAC are listed in Table 1, where the ASSE of IAL-MFAC is smaller than original MFAC in both case, thus the superiority of IAL-MFAC over original MFAC has gained further validation.

**Table 1 ASAE for both cases**
5. CONCLUSIONS

To enhance control performances of every individual in multiple parallel operating identical processes in mass production, this paper proposes a generalized inter-agent learning adaptive control scheme based on information selection according to certain design criterion under parallel operation of multiple identical agents. A simple but effective specified ‘State Déjà vu’ criterion is adopted to reduce the online estimation error accumulation. The application of this framework with ‘State Déjà vu’ criterion to a typical model free adaptive control (MFAC) resulted in a novel IAL-MFAC design. The resulted controller has been proven to have robust convergence and stability. The simulation application results on a discrete time nonlinear process illustrate the effectiveness of the resulted IAL-MFAC.

REFERENCES


