1Q−1

Comparison of Two Neural Network Approaches for Identification of Nonlinear Systems

Julio Tanomaru
Faculty of Engineering University of Tokushima

1. INTRODUCTION

Although the subject of system identification is well developed for linear systems, the same is not true for the nonlinear case. Since multilayer neural networks (MNNs) can be seen as very versatile feedforward blocks with great mapping capability and learning ability, their use for system identification has been the subject of several recent studies [1-2]. This paper presents two neuro-identifiers for general systems and compares their main characteristics. Numerical comparison based on simulation will be presented at the conference.

2. PROBLEM STATEMENT

Consider a P-input-Q-output discrete-time, unknown control plant with input vector **u**(t) and output vector **y**(t), where t is a discrete index. Given a sequence of vector pairs {**u**(t), **y**(t)} for t = 0, 1, ..., T , the identification problem is to devise a P-input-Q-output mathematical model which, when excited with the sequence {**u**(t)}, t = 1, 2, ..., T, will produce an output sequence $\{\hat{\mathbf{y}}(t)\}$, t = 1, 2, ..., T, in such a way that the total identification error, generally defined by the norm

$$E_I \triangleq \sum_{t=1}^{T} E_I(t) \triangleq 0.5 \sum_{t=1}^{T} \sum_{j=1}^{Q} \left[y_j(t) - \hat{y}_j(t)\right]^2 \tag{1}$$

is minimized. Furthermore, in most practical applications the identifier should be able to perform good generalization for inputs not included in the training set.

3. REGRESSIVE MODEL NEURO-IDENTIFIER

In this case, the output of plant at a given time is simply viewed as a function of the previous plant input and output vectors [1]. In other words, for any discrete-time control plant there would be suitable positive integers α and β and a multidimensional mapping **f**(.) in such a way that the plant output at a given instant could be approximated by

$$\hat{\mathbf{y}}(t+1) = \mathbf{f}\,[\mathbf{y}(t), \mathbf{y}(t-1), ..., \mathbf{y}(t-\alpha), \mathbf{u}(t), \mathbf{u}(t-1), ..., \mathbf{u}(t-\beta)]. \tag{2}$$

Such a model has been called series-parallel identification model [1] due to its block pictorial representation. An obvious counterpart is obtained by replacing the true plant output by the corresponding estimates produced by the model itself, resulting in the parallel model [1]:

$$\hat{\mathbf{y}}(t+1) = \mathbf{f}\,[\hat{\mathbf{y}}(t), \hat{\mathbf{y}}(t-1), ..., \hat{\mathbf{y}}(t-\alpha), \mathbf{u}(t), \mathbf{u}(t-1), ..., \mathbf{u}(t-\beta)]. \tag{3}$$

In both cases, the identification model can be thought of as a multidimensional function

$$\mathbf{f}(.): \Re^{(\alpha+1)Q+(\beta+1)P} \rightarrow \Re^{Q}. \tag{4}$$

A single MNN with [(α+1)Q+(β+1)P]-dimension input, Q-dimension output, and appropriate number of nonlinear nodes can be used to emulate **f**(.), constituting a simple *neuro-identifier*. Training of such network can be easily accomplished via pattern or batch learning in order to minimize the identification error. Since such an error is directly based on the MNN output, it is easy to compute the gradient of the error function with respect to the outputs of the MNN, enabling learning by backpropagation (BP). This regressive-model neuro-identifier is illustrated in Fig. 1, where z represents the time-lead operator. If all the traced rectangle illustrated in Fig. 1 is considered as a single network with built-in time delays, the identifier can be viewed as a simple time-delay neural network (TDNN).

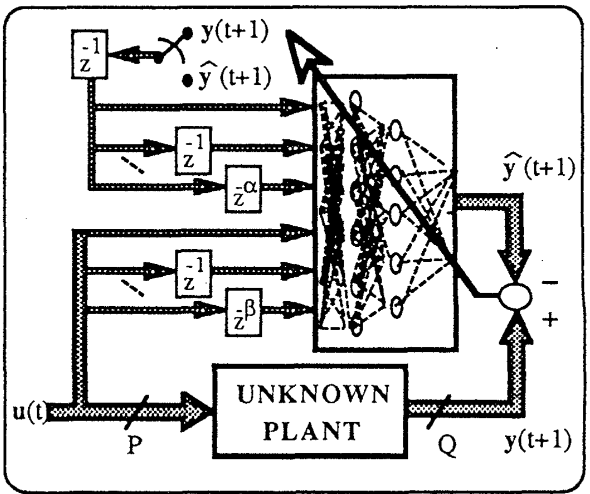

Fig. 1. Neuro-identifier based on the regressive approach to plant representation.

4. STATE VARIABLE NEURO-IDENTIFIER

The neurocontrol structure of the previous section was an immediate consequence of using a regressive model for the control plant. Other representation models may lead to different neurocontrol structures. A novel neurocontrol structure has been derived for the case in which the plant is modeled by a generalized form of the state variable representation [2]. By using this approach, any system can be represented by a set of first order differential (continuous case) or difference (discrete case) equations. For a MIMO discrete-time linear plant with P-dimensional input vector **u**(t) and Q-dimensional output **y**(t), the plant can is described by

$$\begin{cases} \mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \end{cases} \tag{5}$$

where **x**(t) is the N-dimensional plant state vector at instant t, and **A**, **B**, and **C** are real matrices of dimension NxN, NxP, and QxN, respectively, often dubbed state, input, and output matrices. In the nonlinear case, such a representation can be extended by replacing the matrices (linear transformations) by nonlinear mappings of properly defined dimensions. In the general case, the system equations in (5) can be rewritten as

$$\begin{cases} \mathbf{x}(t+1) = \psi\left[\mathbf{x}(t), \mathbf{u}(t)\right] \\ \mathbf{y}(t) = \varphi\left[\mathbf{x}(t)\right] \end{cases} \quad (6)$$

where two nonlinear mapping were defined as

$$\begin{cases} \psi(.): \Re^{N+P} \rightarrow \Re^{N} \\ \varphi(.): \Re^{N} \rightarrow \Re^{Q} \end{cases} . \quad (7)$$

Since the two nonlinear mappings in (7) are essentially feedforward ones, it is somehow straightforward to imagine that two MNNs properly structured could be trained to emulate $\psi(.)$ and $\varphi(.)$. This idea is depicted in Fig. 2, where the same symbol was deliberately used for each mapping and the MNN designed to emulate it, in such a way that the resulting neuro-identifier can be described by:

$$\mathbf{x}(t) = \psi\left[\mathbf{w}^{\psi}(t), \mathbf{x}(t-1), \mathbf{u}(t-1)\right] \quad (8)$$

and

$$\hat{\mathbf{y}}(t) = \varphi\left[\mathbf{w}^{\varphi}(t), \mathbf{x}(t)\right]. \quad (9)$$

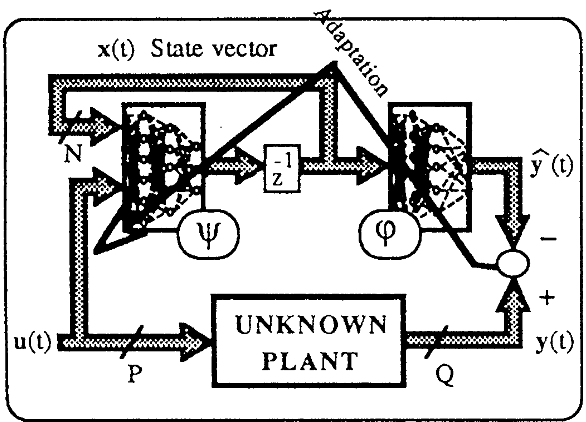

Fig. 2. Neuro-identifier conceived from the state-variable plant representation.

5. Comparative Discussion

Generality: Both approaches are very general in the sense that no constraints were imposed upon the system to be identified. In other words, with appropriate parameters, both approaches shall be able to identify virtually any system in the sense of (1).

Necessary Parameters: While in the regressive approach one must specify α and β in order to define the input-output of the MNN, in the state-variable case the parameter N (order of the state vector) suffices for defining both MNNs. In both cases, the internal structure of the networks (number of nodes, layers, connections, etc.) must be specified by the user.

Dimensions of Networks: A glance at the block diagrams of both neuro-identifiers suggests that a single, large MNN for the regressive case was replaced by two smaller MNNs in the state-variable case. Since high dimensions imply high number of weights and, consequently, difficult training and unreliable generalization ability, low order of the networks is a desirable property.

Robustness to errors in the system order: While small errors in α and β (usually chosen experimentally) can lead to large variation in the total number of weights of the neuro-identifier, the same does not happen in the state-variable case.

Customization: The regressive neuro-identifier offers many possibilities for embedding a priori knowledge about the system to be identified. For instance, if some of the components of the output vector are known to be independent of the others, the identifier could be divided into smaller MNNs, easing training and increasing reliability. The structure of the state-variable identifier is more rigid and virtually does not allow customization.

Training: The regressive identifier is a simple feedforward MNN that can be easily trained via BP. In the state variable case, however, training must be accomplished in two steps. This happens because the effects of the weights and outputs of the nodes of the MNN ψ at a given instant will only affect the identification error at the following instant. Moreover, since training of ψ is performed by backpropagating the identification error through φ, as if both were part of the same MNN.

Applicability: While both identifiers perform essentially the same identification task, the state-variable neuro-identifier has the advantage of producing a state vector as a by-product. Such a state vector has large applicability in the synthesis of controllers [2].

6. Concluding Remarks

This paper presented two configurations for system identifiers based on MNNs, and drafted a comparison between them. In general, while the regressive-model neuro-identifier is easier to train, the state-variable identifier results in smaller networks, is more robust to errors in the order specification, and has great potential applicability in the control field, due to the fact that the system state results as a by-product of the identification process. Currently computer simulation is being performed, and results shall be presented during the conference.

References

[1] K. Narendra and K. Parthasarathy, "Identification and Control of Dynamical Systems Using Neural Networks", *IEEE Trans. Neural Networks*, vol. 1, no. 1, pp. 4-27, 1990.

[2] J. Tanomaru and Y. Takahashi, "A General Structure for Neurocontrol of Dynamical Systems", to appear in *Proc. IECON'93*, Maui, 1993.