10-12-1997

On the stability of LAPART

Chaouki T. Abdallah

G.L. Heileman

M.J. Healy

On The Stability of LAPART [1]

C.T. Abdallah†, G.L. Heileman†, and M.J. Healy‡

†Department of Electrical and Computer Engineering
University of New Mexico, Albuquerque, NM 87131
{chaouki,heileman}@eece.unm.edu

‡The Boeing Company
P.O. Box 3707 MS 7L-66, Seattle, WA 98124
mjhealy@boeing.com

Abstract

In this paper we show how the stability of a LAPART neural network can be deduced as a result of a general theorem on the input/output stability of nonlinear systems. This result gives conditions on how to choose certain parameters in the LAPART network in order to guarantee stability, which has implications on LAPART's generalization properties and its noise robustness.

1. Introduction

The LAPART architecture [1] is a connection of two ART 1 modules connected via feedback and is used to learn input/output mappings or in learning class representations and class-to-class inferences. Such neural networks have recently been used to identify and control nonlinear systems, by calling on their ability to approximate signals or systems arbitrarily closely [2]. Since the interconnection of dynamical systems via feedback causes non-trivial stability issues [3], we try in this paper to study the bounded-input-bounded-output stability of the LAPART network using general theorems on input-output stability [4]. Such stability concepts have implications in terms of the generalizations capabilities and the noise tolerance of the network, since if a small change in the input causes the output changes to be drastic, the network learning becomes highly susceptible to noise and unforeseen input changes. Note that the results of this paper apply to more general classes of neural networks (any class which has feedback) [5, 6], or to the interconnection of a neural network with a dynamical system in a feedback structure. The issue of stability of neural networks when used in control applications has been brought to the forefront by surveys such as [7, 8, 9, 10, 2], and by researchers such as [11, 12, 13, 14, 15].

The paper starts in section 2 by reviewing the basic operation of the LAPART network. We then briefly review ideas from input/output stability in section 3. The stability of LAPART is then studied in section 4 and our conclusions are given in section 5.

2. The LAPART Network

To introduce notation and support our discussion of the LAPART architecture, we briefly review the function of an ART 1 system [16]. An ART 1 network autonomously classifies binary input patterns and represents each class by a prototype or *template* binary pattern. Each template is formed through the ART 1 unsupervised learning process and represented by a unique set of adaptive connections controlled by a classification node.

A binary pattern X can be regarded as a string of numerical 1's and 0's. For any two binary patterns X and Y having the same length (number of 0-1 components), let $X \wedge Y$ denote the binary pattern that constitutes their componentwise minimum, where the minimum operation on components has the properties $0 \wedge 0 = 0$, $1 \wedge 1 = 1$, $0 \wedge 1 = 0 = 1 \wedge 0$. For a set S of binary patterns all having the same length, with $S = \{X^1, X^2, \ldots, X^N\}$,

[1]The research of C.T. Abdallah and G.L. Heileman was supported by Grant W-300445 from Boeing Computer Services

0-7803-4053-1/97/$10.00 © 1997 IEEE

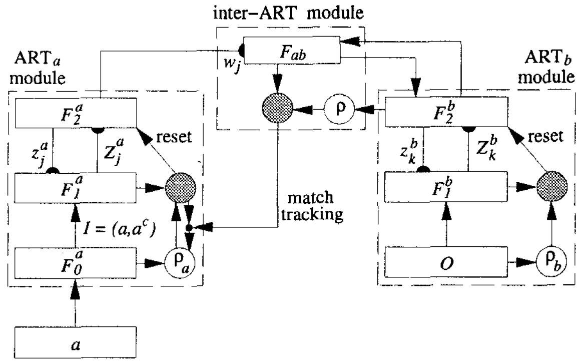

Figure 1: LAPART Architecture

let $\bigwedge S$ denote the minimum over the set, $\bigwedge S = X^1 \wedge X^2 \wedge \ldots \wedge X^N$ (this is unambiguous, since the $\wedge$ operation is associative). Finally, let $\| X \|$ denote the number of 1s in the binary pattern X.

For each ART 1 input pattern, unsupervised learning occurs in two phases: (1) recognition of the input pattern as a member of some class, and (2) updating of the class template through synaptic learning. The process can be summarized as solving the combinatorial optimization problem stated as follows:

$$\underset{j}{\text{maximize}} \quad \frac{\| I \wedge T^j \|}{\beta + \| T^j \|} \text{subject to} \quad \frac{\| I \wedge T^j \|}{\| I \|} \geq \rho. \tag{1}$$

The solution, i, is the index of the F_2 node $F_{2,i}$ that represents the class assigned to I, which has an associated ART 1 binary template pattern T^i.

The LAPART network is based upon the coupling of two ART 1 networks, referred to as A and B. The interconnects between these two subnetworks in a LAPART network force an interaction of their respective classifications of their inputs. This modifies their unsupervised learning properties to allow the learning of inferencing relationships between their learned input pattern classes. The learning takes place when the input nodes of the two ART 1 subnetworks are presented with a sequence of simultaneously-occurring input patterns I_A^k and I_B^k for subnetworks A and B, respectively. As A and B form class templates for their inputs, the LAPART network learns inference relations between their classes by forming strong $F_2^A \longrightarrow F_2^B$ interconnections.

To see how the LAPART network learns class-to-class inferences from example input pairs, we give a brief summary of its operation. We discuss the processing of a typical input pattern pair (I_A, I_B). Initially, networks A and B are untrained ART 1 networks, with their F_2 nodes linked by weak $F_2^A \longrightarrow F_2^B$ connections. There are two major cases to consider in describing the operation of a LAPART network when its two ART 1 subnetworks simultaneously receive their input patterns I_A and I_B, respectively.

Case 1: New A-class. Network A forms a new class for its current input I_A if it currently has no acceptable template for I_A. Then, a previously-uncommitted F_2^A node has its connections with the F_1^A layer recoded to have I_A as an incidence pattern. Denote the new class by A_i, where node $F_{2,i}^A$ is the corresponding classification node just selected. Following a delay controlled by network A, network B is allowed to read its input, I_B. This engages the unsupervised learning process in network B, and also the synaptic learning of a strong feedforward connection from node $F_{2,i}^A$ to the resonating F_2^B node, $F_{2,j}^B$, say. The class B_j could be either an existing class or a new class as in network A. We say that the LAPART network has formed the association $A_i \Longrightarrow B_j$, written as a logical implication between classes. We say this because the future presentation of an input pair for which A_i is the resonating class for the A input will result in the inference, through the strong $F_{2,i}^A \longrightarrow F_{2,j}^B$ connection, that class B_j is appropriate for the B input.

Case 2: Existing A-class. This case occurs when network A already contains a class representation A_i that resonates with I_A. Then it also has a previously-learned class-to-class relation $A_i \Longrightarrow B_j$. Thus, $F_{2,i}^A$ primes $F_{2,j}^B$ through the strong $F_{2,i}^A \longrightarrow F_{2,j}^B$ connection. Network A simultaneously releases its inhibition of network B's input field. Network B then reads out the class B_j template over the F_1^B layer, and simultaneously reads its input, I_B. The effect of this is to force network B to perform the vigilance pattern-matching test using the template pattern T_j^B instead of one that it would have selected through the ART 1 winner-take-all competition in layer F_2^B. Even though it does not control selection of a class for its input during this A-class-to-B-class inferencing, however, network B still has control of its vigilance node. As a consequence, network B can either confirm or disconfirm the inferred class for its input.

If the pattern match between the inferred class template T_B^j and the input pattern I_B is not acceptable,

$$\frac{\| I_B \wedge T^{B,j} \|}{\| I_B \|} < \rho_B ,$$

a lateral reset occurs in network B. Through the fixed, strong connection $VIG_B \longrightarrow VIG_A$ between the two vigilance nodes, network A is subsequently forced to also undergo a reset. This forces the choice of a new network A class to represent its input. This is followed by either a re-enactment of the Case 2 scenario or by Case 1, the learning of a new inferencing connection from a new A class. The Case 2 scenario can be re-enacted several times, ending with either Case 1 or, alternatively, the inference of an acceptable B class for input I_B.

Following completion of either a Case 1 scenario, or a Case 2 scenario ending in acceptance of the inferred template, both the I_A and I_B class templates are updated in the usual ART 1 fashion. The result of this process, operating upon a sequence of input pairs, is the formation of classes and inference relationships by which network A recognizes an input pattern and then infers a class, or template, for network B's input pattern. In summary, a LAPART network learns class representations and class-to-class inferences for for its input pairs by making and testing trial inferences, using a process similar to the ART 1 hypothesis-testing process for unsupervised classification..

The question then arises on how sensitive is the network to noise corrupted input/output patterns. In other words, having learned a class-to-class inference under normal circumstances, will it have a drastically different behavior if the input pattern changes by a small amount? In the next section, we discuss general concepts from input/output stability in order to answer these questions in section 4.

3. Input/Output Stability

Let us define the graph of a system with input u and output y as the ordered pair $G = (u, y)$, and the inverse graph as $G^I = (y, u)$. If we connect 2 systems in the standard feedback connection shown in Figure 2, and the result is a dynamical system (i.e. no algebraic loop), we say that we have a well-defined connection. We say that a system is stable if small input (in an appropriate sense) causes a small output. We also define the truncated signal x_τ as the signal equal to $x(t); \forall t \leq \tau$ and zero elsewhere. In order to state the theorem, we define $\delta x = x - x_{nominal}$ as the difference between the signal x and its nominal value. This will allow us to state our result in terms of the nominal behavior of the LAPART network and changes therein.

Theorem 1 *[4] A well-defined interconnection is stable if and only if there exists a gain function γ which gives a bound on the norm of truncated signals in the inverse graph of Σ_2^I as a function of the truncated distance from the signals to the graph of Σ_1, i.e.*

$$x \in G_2^I \Longrightarrow ||x_\tau|| \leq \gamma(d_\tau(x, G_1)); \forall \tau \qquad (1)$$

where $d_\tau(x, G_1) = \inf_{z \in G_1} ||(x - z)_\tau||$.

Note that this result applies in both continuous-time and discrete-time, i.e. $x(t)$ could denote a continuous-time signal or a sequence in discrete-time.

4. The Stability of LAPART

The stability of LAPART will be deduced from Theorem 1 above. Let us consider the following signals in the LAPART block diagram 2,

$$\begin{aligned} u_2 &= d_2 + y_1 \\ &= \begin{bmatrix} I_B \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ CF_2^A \end{bmatrix} \\ &= \begin{bmatrix} I_B \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & C & 0 \end{bmatrix} \begin{bmatrix} F_1^A \\ F_2^A \\ v_A \end{bmatrix} \end{aligned} \qquad (2)$$

and

$$\begin{aligned} u_1 &= d_1 + y_2 \\ &= \begin{bmatrix} I_A \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ v_B \end{bmatrix} \\ &= \begin{bmatrix} I_A \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} F_{1B} \\ F_2^B \\ v_B \end{bmatrix} \end{aligned} \qquad (3)$$

Then,

$$G_A = \begin{bmatrix} d_1 + y_2 \\ y_1 \end{bmatrix}; \quad G_B^I = \begin{bmatrix} y_2 \\ d_2 + y_1 \end{bmatrix} \tag{4}$$

Theorem 2 *If whenever*

$$\begin{bmatrix} -\delta I_A \\ \delta V_B - \delta V_A \\ \delta I_B \\ \delta F_2^B - C\delta F_2^A \end{bmatrix} \tag{5}$$

is small, leads to

$$\delta V_B \text{ and } \delta F_2^B \tag{6}$$

are small, the LAPART is BIBO stable.

The theorem basically states conditions under which small perturbations in the input patterns I_A do not cause large perturbations in internal LAPART signals. This has applications in trying to establish generalization properties and noise tolerances in the network structure. Moreover, if one uses the following interpretation of stability: *an input/output mapping is BIBO stable if it is continuous, i.e. if for any given 2 output signals which are close (in some suitable norm), one can find 2 corresponding input signals which cause the output signals to be close.*, then the theorem may be re-interpreted to mean that a stable LAPART learns continuous input/output mappings.

5. Conclusions

In this paper, we have introduced a general result from stability theory to study the BIBO stability of the LAPART network. Note that the same result may be used in many feedback neural networks structures, both in continuous and discrete-times. We are currently investigating the applications of such concept to the selection of different parameters in the ART1 modules.

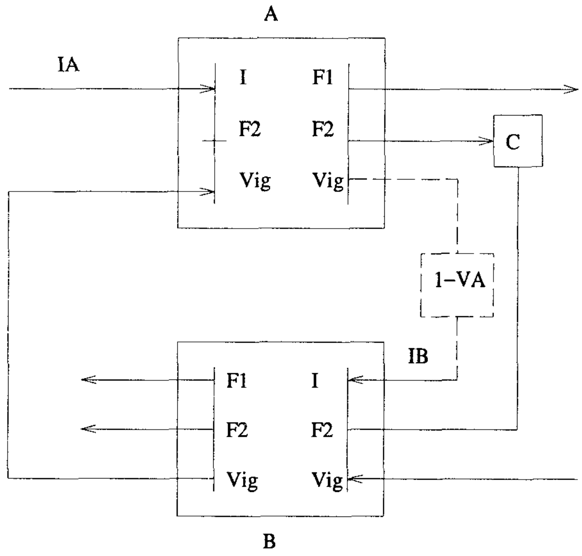

Figure 2: LAPART Feedback Structure

References

[1] M. J. Healy, T. P. Caudell, and S. D. G. Smith, "A neural architecture for pattern sequence verification through inferencing," *IEEE Transactions on Neural Networks*, vol. 4, no. 1, pp. 9–20, 1993.

[2] C. Abdallah, G. Heileman, M. Georgiopoulos, and D. Hush, "An overview of neural network results for systems and control," *Int. J. of Intelligent Control and Systems*, vol. 1, no. 2, pp. 177–194, 1996.

[3] C. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties.* New York, N.Y.: Academic Press, 1st ed., 1975.

[4] A. Teel, T. georgiou, L. Praly, and E. Sontag, "Input-output stability," in *The Control Handbook* (W. S. Levine, ed.), pp. 895–908, Boca Raton, Fl: CRC Press, 1995.

[5] M. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 13, pp. 815–825, September/October 1983.

[6] M. Zak, "Non–Lipschitzian dynamics for neural modeling," *Applied Mathematics Letters*, vol. 2, no. 1, pp. 69–74, 1989.

[7] K. Hunt, D. Sbarbaro, R. Żbikowski, and P. Gawthrop, "Neural networks for control systems–A survey," *Automatica*, pp. 1083–1112, 1995.

[8] W. Miller, R. Sutton, and P. Werbos, eds., *Neural Networks for Control.* Cambridge, MA: MIT Press, 1990.

[9] E. Sontag, "Neural networks for control," in *Essays on Control: Perspectives in the Theory and its Applications* (H. Trentelman and J.Willems, eds.), pp. 339–380, Boston, MA: Birkhauser, 1993.

[10] R. Żbikowski, K. Hunt, A. Dzielinński, R. Murray-Smith, and P. Gawthrop, "A review

of advances in neural adaptive control systems," Tech. Rep. ESPRIT III Project 8039: NACT, Dept. of Mathematics, Royal Institute of Technology, Daimler-Benz AG and University of Glasgow, 1994.

[11] F. Lewis, K. Liu, and A. Yeşildirek, "Neural net robot controller with guaranteed tracking performance," *IEEE Trans. Neural Nets*, vol. 6, no. 3, pp. 703–715, 1995.

[12] E. Tzirkel-Hancock and F. Fallside, "Stable control of nonlinear systems using neural networks," Tech. Rep. CUED/F–INFENG/TR.81, Cambridge University Engineering Department, 1991.

[13] A. Barto, "Connectionist learning for control: An overview," in *Neural Networks For Control* (W. Miller, R. Sutton, and P. Werbos, eds.), pp. 5–58, Cambridge, MA: MIT Press, 1990.

[14] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Trans. Neural Nets*, vol. 1, no. 1, pp. 4–27, 1990.

[15] F. Lewis, S. Jagannathan, and A. Yeşildirek, "Neural network control of robot arms and nonlinear systems," in *Neural Networks for Control* (D. Elliot, ed.), p. To appear, Ablex Publishing Corp., 1995.

[16] G. A. Carpenter and S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine," *Computer Vision, Graphics, and Image Processing*, vol. 37, pp. 54–115, 1987.