Ordered Fuzzy ARTMAP: A Fuzzy ARTMAP algorithm with a fixed order of pattern presentation

I. Dagher*, M. Georgiopoulos*, G. L. Heileman**, G. Bebis***

* Department of Electrical and Computer Engineering
University of Central Florida, Orlando, FL 32816
Phone: (407)823-5338 Fax: (407)823-5835
mng@ece.engr.ucf.edu

** Department of Electrical and Computer Engineering
University of New Mexico, Albuquerque, NM 87131
Phone: (505)277-4011 Fax: (505)277-1439
heileman@eece.unm.edu

*** Department of Computer Science
University of Nevada-Reno, Reno, NV 89557
Phone: (702)784-6463 Fax: (702)784-1877
bebis@cs.unr.edu

Abstract

In this paper we introduce a procedure that identifies a fixed order of training pattern presentation for Fuzzy ARTMAP. The resulting algorithm is named *Ordered Fuzzy ARTMAP*. Experimental results have demonstrated that Ordered Fuzzy ARTMAP achieves a network performance that is better than the average Fuzzy ARTMAP network performance (averaged over a fixed number of random orders of pattern presentations), and occasionally better than the maximum Fuzzy ARTMAP network performance (maximum over a fixed number of random orders of pattern presentations). What is also worth noting is that the computational complexity of the aforementioned procedure is only a small fraction of the computational complexity required to complete the training phase of Fuzzy ARTMAP for a single order of pattern presentation.

1 Introduction

Pattern classification is a key element to many engineering solutions. Sonar, radar, seismic, and diagnostic applications all require the ability to accurately classify data. Control, tracking and prediction systems will often use classifiers to determine input-output relationships. Because of this wide range of applicability, pattern classification has been studied a great deal. Simpson in his Fuzzy Min-Max paper (see [1]) identified a number of desirable properties that a pattern classifier should possess. These properties are listed below: **Property 1:** *On-Line Adaptation*, **Property 2:** *Non-Linear Separability*, **Property 3:** *Short Training Time*, **Property 4:** *Soft and Hard Decisions*, **Property 5:** *Verification and Validation*, **Property 6:** *Independence from Tuning Parameters*, **Property 7:** *Nonparametric Classification*, and **Property 8:** *Overlapping Classes*.

A neural network classifier that satisfies most of these properties is Fuzzy ARTMAP ([2]). Fuzzy ARTMAP is a member of the class of neural network architectures referred to as *ART-architectures* developed by Carpenter, Grossberg, and their colleaugues at Boston University. The ART-architectures are based on the *ART theory* introduced by Grossberg in ([3]). However, the performance of Fuzzy ARTMAP is not independent from tuning parameters (Property 6). Actually, Fuzzy ARTMAP performance depends on the values of two parameters (the choice parameter, and the vigilance parameter), and also on the order of pattern presentation in the training phase. To circumvent the first problem (i.e., dependence of Fuzzy ARTMAP on the dependence of the choice parameter and vigilance parameter), most Fuzzy ARTMAP simulations that have appeared in the literature assume zero values for these two parameters. One of the main reasons for the popularity of this choice is that we end up, after training is over, with a neural network architecture of the minimal size; this is quite desirable especially when performance comparisons are made between Fuzzy ARTMAP and other neural network architectures, which offer more compact representations of the data (such as the Back-Prop neural network). The second problem (dependence on the order of pattern presentation in the training phase of Fuzzy ARTMAP) is not as easy to solve. One way around it is to consider differ-

0-7803-4859-1/98 $10.00©1998 IEEE

ent orders of presentations of the training data and pick the one that maximizes the performance of the network. The drawbacks of this approach are that (i) considerable experimentation is required to find a random order of pattern presentation that achieves a good network performance, and (ii) finding a random order of pattern presentation that obtains good performance seems like a guessing exercise. In this paper we introduce a systematic procedure that identifies a fixed order of pattern presentation, which leads to a Fuzzy ARTMAP network performance that is better than the average network performance (averaged over a fixed number of random orders of pattern presentation), and occasionally better than the maximum network performance (maximum over a fixed number of random orders of pattern presentation). The resulting Fuzzy ARTMAP algorithm is named *Ordered Fuzzy ARTMAP*. What is worth noting is that the computational complexity of the procedure that generates the fixed order of pattern presentation for the Ordered Fuzzy ARTMAP is a small percentage of the computational complexity of the training phase of Fuzzy ARTMAP.

The organization of the paper is as follows: In Section 2, we discuss briefy the Fuzzy ARTMAP architecture its training phase, and its performance phase. In Section 3, we introduce the procedure that produces the fixed order of pattern presentation for Ordered Fuzzy ARTMAP. In Section 4, we experimentally demonstrate the superiority of Ordered Fuzzy ARTMAP compared to the average Fuzzy ARTMAP performance (averaged over a fixed number of random orders of pattern presentation), and occasionally compared to the maximum Fuzzy ARTMAP performance (maximum over a fixed number of random orders of pattern presentations). In Section 5, we provide a review of the paper and some conclusive remarks.

2 The Fuzzy ARTMAP Neural Network

A detailed description of the Fuzzy ARTMAP neural network can be found in [2]. For completeness, in the following, we present only the necessary details.

The Fuzzy ARTMAP neural network consists of two Fuzzy ART modules, designated as ART_a and ART_b, as well as an inter-ART module as shown in Figure 1. Inputs (I's) are presented at the ART_a module, while their corresponding outputs (O's) are presented at the ART_b module. The inter-ART module includes a MAP field whose purpose is to determine whether the correct mapping has been established from inputs to outputs.

Fuzzy ARTMAP can operate in two distinct phases: the *training phase* and the *performance phase*. In this paper we focus on classification tasks, where many inputs are mapped to a single, distinct output. It turns out that for classification tasks , the operations performed at the ART_b and inter-ART modules can be ignored, and the algorithm can be described by referring only to the top-down weights and the parameters of the ART_a module.

The *training phase* of Fuzzy ARTMAP works as follows: Given a list of training input/output pairs, such as $\{\mathbf{I}^1, \mathbf{O}^1\}, \ldots \{\mathbf{I}^r, \mathbf{O}^r\}, \ldots \{\mathbf{I}^{PT}, \mathbf{O}^{PT}\}$, we want to train Fuzzy ARTMAP to map every input pattern of the training list to its corresponding output pattern. In order to achieve the aforementioned goal, we present the training list repeatedley to the Fuzzy ARTMAP architecture. That is present $\mathbf{I}^1$ to ART_a and $\mathbf{O}^1$ to ART_b, then $\mathbf{I}^2$ to ART_a and $\mathbf{O}^2$ to ART_b, and finally $\mathbf{I}^{PT}$ to ART_a and $\mathbf{O}^{PT}$ to ART_b; this corresponds to *one list presentation*. We present the training list as many times as it is necessary for Fuzzy ARTMAP to correctly classify all the input patterns. The classification task is considered accomplished (i.e., the learning is complete) when the weights do not change during a list presentation. The aforementioned training scenario is called *off-line learning*. Note that an input pattern $\mathbf{I}$ presented to module ART_a of Fuzzy ARTMAP is complementary encoded. That is $\mathbf{I} = (\mathbf{a}, \mathbf{a}^c)$, where $\mathbf{a}$ is a pattern of arbitrary dimensionality with components in the interval $[0, 1]$, and $\mathbf{a}^c = \mathbf{1} - \mathbf{a}$, where $\mathbf{1}$ is the all ones vector of the same dimensionality as $\mathbf{a}$ and $\mathbf{a}^c$.

The *performance phase* of Fuzzy ARTMAP works as follows: Given a list of test input patterns, such as $\tilde{\mathbf{I}}^1, \ldots, \tilde{\mathbf{I}}^2, \ldots, \mathbf{I}^{\tilde{PS}}$, we want to find the Fuzzy ARTMAP output produced when each one of the aforementioned test patterns is presented at its F_1^a field. In order to achieve the aforementioned goal, we present the test list once to the trained Fuzzy ARTMAP architecture.

Note that an input pattern $\mathbf{I}$ presented to module ART_a of Fuzzy ARTMAP is complementary encoded. That is $\mathbf{I} = (\mathbf{a}, \mathbf{a}^c)$, where $\mathbf{a}$ is a pattern of arbitrary dimensionality with components in the interval $[0, 1]$, and $\mathbf{a}^c = \mathbf{1} - \mathbf{a}$, where $\mathbf{1}$ is the all ones vector of the same dimensionality as $\mathbf{a}$ and $\mathbf{a}^c$.

3 The Ordered Fuzzy ARTMAP Algorithm

The idea of Ordered Fuzzy ARTMAP is to identify the order according to which patterns are presented during the training phase of Fuzzy ARTMAP. This task is accomplished by following a systematic procedure that consists of three stages: In Stage 1, we choose the first pattern to be presented; this pattern corresponds to the first cluster center of the training data. In Stage 2, we choose the next $n_{clust} - 1$ patterns to be presented to Fuzzy ARTMAP; these patterns correspond to the next $n_{clust} - 1$ cluster centers of the training data, and they are identified through a clustering method known as the *Max-Min* algorithm ([4]). In Stage 3, we choose the remaining $PT - n_{clust}$ patterns to be presented to Fuzzy ARTMAP; these patterns are chosen according to the minimum Euclidean distance criterion from the n_{clust} centers defined in Stages 1 and 2. In the following, we describe, in more detail, each one of these stages. Note that the only weak link in the aforementioned procedure is the value of n_{clust}. Experimental results have shown that good values of n_{clust} are the *number of classes* or the *number of classes* +1 of the data in the training set.

Stage 1: *The first pattern*
For each pattern $\mathbf{I} = (a_1, \dots a_{M_a}, a_{M_a+1}, \dots, a_{2M_a})$ in the training set we compute

$$|a_{M_a+1} - a_1| + |a_{M_a+2} - a_2| + \cdots |a_{2M_a} - a_{M_a}| \quad (1)$$

The pattern from the training set that maximizes the above sum is chosen to be the first pattern presented to Fuzzy ARTMAP, and the first cluster center to be used in Stage 2.

Stage 2: *The next $n_{clust} - 1$ patterns*
As we have mentioned before, this stage involves the Max-Min clustering algorithm which defines appropriate cluster centers that constitute the next in line input patterns to be presented in the training phase of Fuzzy ARTMAP. The steps followed to define these cluster centers are described below.

1. Denote the first cluster center (input pattern) identified in Stage 1 by $\mathbf{I}_O^1$. Initialize the index r to the value of one.

2. Compute the Euclidean distance of each one of the input patterns in the training set S_T from the k-th cluster center, and find the minimum one, denoted by d_{min}^k. That is,

$$d_{min}^k = \min_{\mathbf{I} \in S_T} \{dist(\mathbf{I}, \mathbf{I}_O^k)\} \quad (2)$$

   Repeat the above step for all the cluster centers k, such that $1 \leq k \leq r$.

3. Find the input pattern from the training set S_T that maximizes d_{min}^k, where $1 \leq k \leq r$. Designate this input pattern by the generic name $\mathbf{I}$. Our next cluster center, designated by $\mathbf{I}_O^{r+1}$ is equal to $\mathbf{I}$, that is is

$$\mathbf{I}_O^{r+1} = \mathbf{I} \quad (3)$$

   This cluster center constitutes the next input pattern to be presented in the training phase of Fuzzy ARTMAP. Eliminate input pattern $\mathbf{I}$ from the training set S_T.

4. If $r = n_{clust} - 1$ this stage is completed. Otherwise, increase your index R by one, and go to Step 2.

At the end of stage 2, we have identified n_{clust} cluster centers that correspond to the input patterns $\mathbf{I}_O^r$ $1 \leq r \leq n_{clust}$ of the training set. The next stage identifies the order according to which the remaining input patterns are presented in Fuzzy ARTMAP.

Stage 3: *The remaining $PT - n_{clust}$ input patterns*
The steps followed in this stage are described below.

1. Set the index r to the value n_{clust}. The patterns in the training set S_T are all the training input patterns except the ones identified as cluster centers in Stages 1 and 2.

2. Calculate the Euclidean distance of every pattern $\mathbf{I}$ in the set S_T from the n_{clust} cluster centers.

3. Find the minimum of these distances. Assume that it corresponds to input pattern $\mathbf{I}$. This pattern is the next in line input pattern to be presented in the training phase of Fuzzy ARTMAP. Eliminate $\mathbf{I}$ from the set S_T. We define

$$\mathbf{I}_O^{r+1} = \mathbf{I} \quad (4)$$

4. If $r = PT - 1$ this stage is complete. Otherwise, increase the index r by one and go to Step 2.

After the end of Stage 3, we have identified the ordered set of patterns $\mathbf{I}_O^1, \mathbf{I}_O^2, \dots, \mathbf{I}_O^{PT}$. This is the order according to which the patterns in the in the training set will be presented to the Ordered Fuzzy ARTMAP.

4 Experimental Results – Comparisons

4.1 Measures of Performance

In order to demonstrate the superior performance of the Ordered Fuzzy ARTMAP compared to Fuzzy ARTMAP we chose to conduct experiments on a number of databases extracted from the UCI repository database ([5]). The databases chosen from the repository were: Sonar, Diabetes, Breast, Bupa, Iris, Wine, Balance, Cars, and Glass. The Sonar, Diabetes, Breast, and Bupa are two class classification problems, the Iris, Wine, Balance are three class classification problems, the Cars is a four class classification problem, and finally the Glass is a six class classification problem. It is worth noting that the data in each of the above datasets were split into a training set (2/3 of the data) and a test set (1/3 of the data). A description of each one of these databases is provided in [5].

One of the measures of performance that we intend to use in comparing Ordered Fuzzy ARTMAP and Fuzzy ARTMAP is the *generalization performance* of these networks. The generalization performance of a network is defined to be the percentage of patterns in the test set that are correctly classified by a trained network. Since the performance of Fuzzy ARTMAP depends on the order of pattern presentation in the training set, ten different random orders of pattern presentation will be investigated, and performance measures such as the average generalization performance, the minimum generalization performance, the maximum generalization performance, and the standard deviation of the generalization performance will be produced for Fuzzy ARTMAP. Other measures of comparison of Ordered Fuzzy ARTMAP and Fuzzy ARTMAP are the *sizes* of the networks that these two architectures create, and the *computational complexity* of Ordered Fuzzy ARTMAP and Fuzzy ARTMAP.

4.2 Comparisons of the Ordered Fuzzy ARTMAP and the Fuzzy ARTMAP network

As we have mentioned before the only weak link in the procedure, described in Section 3, that finds an ordered sequence of training patterns for Ordered Fuzzy ARTMAP, is the number of clusters identified in Stage 2. Experimental results have shown that a good rule of thumb for n_{clust} in Stage 2, is the *number of classes* or the *number of classes* +1 of the database. In Table 1, we show generalization performance comparisons between Fuzzy ARTMAP and ordered Fuzzy ARTMAP when the number of clusters (n_{clust}) in the Ordered Fuzzy ARTMAP is chosen to be equal to the *number of classes* +1. Looking at the results depicted in Table 1, we observe the following:

For the **Sonar** database the generalization performance of the Ordered Fuzzy ARTMAP is better than the minimum generalization performance of Fuzzy ARTMAP by 15.49%, better than the average generalization performance of Fuzzy ARTMAP by 9.38%, and better than the maximum generalization performance by 1.7%. For the **Diabetes** database the generalization performance of the Ordered Fuzzy ARTMAP is better than the minimum generalization performance of Fuzzy ARTMAP by 8.23%, better than the average generalization performance of Fuzzy ARTMAP by 3.17%, and better than the maximum generalization performance by 1.18%. For the **Breast** database the generalization performance of the Ordered Fuzzy ARTMAP is better than the minimum generalization performance o f Fuzzy ARTMAP by 1.29%, better than the average generalization performance of Fuzzy ARTMAP by 0.04%, and worse than the maximum generalization performance by 1.73%. For the **Bupa** database the generalization performance of the Ordered Fuzzy ARTMAP is better than the minimum generalization performance of Fuzzy ARTMAP by 9.64%, better than the average generalization performance of Fuzzy ARTMAP by 0.17%, and worse than the maximum generalization performance by 6.15%. For the **Iris** database the generalization performance of the Ordered Fuzzy ARTMAP is better than the minimum generalization performance of Fuzzy ARTMAP by 8.34%, better than the average generalization performance of Fuzzy ARTMAP by 2.92%, and better than the maximum generalization performance by 2.09%. For the **Wine** database the generalization performance of the Ordered Fuzzy ARTMAP is better than the minimum generalization performance of Fuzzy ARTMAP by 6.89%, better than the average generalization performance of Fuzzy ARTMAP by 2.58%, and worse than the maximum generalization performance by 0.01%. For the **Balance** database the generalization performance of the Ordered Fuzzy ARTMAP is better than the minimum generalization performance of Fuzzy ARTMAP by 3.85%, worse than the average generalization performance of Fuzzy ARTMAP by 0.43%, and worse than the maximum generalization performance by 3.37%. For the **Cars** database the generalization performance of the Ordered Fuzzy ARTMAP is better than the minimum generalization performance of Fuzzy ARTMAP by 4.29%, better than the average generalization performance of Fuzzy ARTMAP by 1.07%, and worse than the maximum generalization performance by 1.79%. For the **Glass** database the generalization performance of the Ordered Fuzzy ARTMAP is better than the minimum generalization performance of Fuzzy ARTMAP by 11.59%, better than the average generalization performance of Fuzzy ARTMAP by 5.79%, and worse than the maximum generalization performance by 7.25%.

It is also worth pointing out that the size of the Ordered Fuzzy ARTMAP architecture ranges between 0.84 and 1.52 of the average size of the Fuzzy ARTMAP architecture. We have also evaluated the computational overhead of identifying the order according to which patterns ought to be presented in Fuzzy ARTMAP. It turns out that for all of the above experiments the computational overhead of the ordering procedure in the ordered Fuzzy ARTMAP ranges between 0.48% and 22.8% of the average computational complexity of the training phase of Fuzzy ARTMAP. Note though, that for all the databases (except one) the computational overhead of the ordering procedure in the ordered Fuzzy ARTMAP is below 10%, and in most databases well below 10%, of the average computational complexity of the training phase of Fuzzy ARTMAP.

If we allow ourselves to pick the optimum size of the number of clusters n_{clust} in the ordering procedure of the Ordered Fuzzy ARTMAP, our experiments demonstrate that we do not have to search over an extensive range of values for n_{clust}. In all of the experiments conducted so far we found that the optimum Ordered Fuzzy ARTMAP happens for n_{clust} values between *number of classes* and *number of classes* +1 of the databases. The generalization performance of the optimum Ordered Fuzzy ARTMAP is shown in Table 2. Note that to identify the optimum number of clusters for the Ordered Fuzzy ARTMAP we need to order the training data twice, and we need to train Fuzzy ARTMAP twice for the two different orders of pattern presentations. Hence the computational complexity of the optimum Ordered Fuzzy ARTMAP is approximately two times the computational complexity of the Ordered Fuzzy ARTMAP. The advantage of the optimum Fuzzy ARTMAP is improved performance compared to the Ordered Fuzzy ARTMAP.

5 Review – Conclusions

In this paper we introduced a procedure that identifies a fixed order of pattern presentation in Fuzzy ARTMAP. We designated the resulting algorithm *Ordered Fuzzy ARTMAP*. Experiments with ten different classification problems have shown that Ordered Fuzzy ARTMAP attains a superior network performance compared to the average Fuzzy ARTMAP Performance (averaged over a fixed number of random orders of pattern presentations), and occasionally compared to the maximum Fuzzy ARTMAP performance (maximum over the performances of a fixed number of random orders of pattern presentations). It is also worth pointing out that the computational complexity of the proposed procedure is a small fraction of the computational complexity of the training phase of Fuzzy ARTMAP for a single order of pattern presentation. Furthermore, the size of network architectures that Ordered Fuzzy ARTMAP creates is comparable with the average size of the Fuzzy ARTMAP architectures created (averaged over the sizes of the architectures created by a fixed number of random orders of pattern presentations).

References

[1] P. K. Simpson, "Fuzzy min-max neural networks – Part 1: Classification," *IEEE Transactions on Neural Networks*, Vol. 3, No. 5, pp. 776-786, September 1992.

[2] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, and D. B. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Transactions on Neural Networks*, Vol. 3, No. 5, pp. 698-713, September 1992.

[3] S. Grossberg, "Adaptive pattern recognition and universal recoding II: Feedback, expectation, olfaction, and illusions, *Biological Cybernetics*, Vol. 23, pp. 187-202, 1976.

[4] J. T. Tou and R. C. Gonzalez. Pattern Recognition Principles. Addison Wesley, Reading, MA, 1976.

[5] P. Murphy, and D. Ada, UCI Repository of Machine Learning databases, *Technical Report*, Department of Computer Science, University of California, Irvine, CA, 1994.

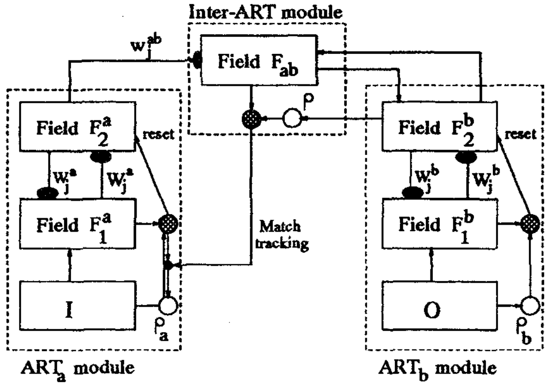

Figure 1: The Fuzzy ARTMAP Neural Network.

Table 1: Generalization performances of the Fuzzy ARTMAP and the Ordered Fuzzy ARTMAP with $n_{clust} = n_{class} + 1$

Database	Fuzzy ARTMAP				Ordered Fuzzy ARTMAP	
	min.Gen.	max.Gen.	avg.Gen.	std.dev.	n_{clust}	Gen.
Sonar	63.77	78.26	70.58	4.15	3	79.96
Diabetes	61.57	70.98	66.63	2.57	3	69.80
Breast	93.10	96.12	94.35	0.95	3	94.39
Bupa	47.37	63.16	56.84	4.22	3	57.01
Iris	89.58	95.83	95.00	1.91	4	97.92
Wine	91.38	98.28	95.69	2.70	4	98.27
Balance	71.63	78.85	75.91	2.42	4	75.48
Cars	63.21	69.29	66.43	2.13	5	67.50
Glass	57.97	76.81	63.77	6.18	7	69.56

Table 2: Generalization performances of the Ordered Fuzzy ARTMAP with the optimum number of clusters

Database	n_{clust}	Generalization
Sonar	3	79.96
Diabetes	3	69.80
Breast	3	94.39
Bupa	2	65.67
Iris	4	97.92
Wine	4	98.27
Balance	3	77.88
Cars	4	71.42
Glass	6	73.91