Properties of Learning in Fuzzy ART

Juxin Huang **Michael Georgiopoulos**
Department of Electrical and Computer Engineering
University of Central Florida, Orlando, FL 32816, USA

Gregory L. Heileman
Department of Electrical and Computer Engineering
University of New Mexico, Albuquerque, NM 87131, USA

Abstract-This paper presents some important properties of the Fuzzy ART neural network algorithm. The properties described in the paper are distinguished into a number of categories. These include template, access, and reset properties, as well as properties related to the number of list presentations needed for weight stabilization. These properties provide numerous insights as to how Fuzzy ART operates. Furthermore, the effect of the Fuzzy ART parameters α and ρ on the functionality of the algorithm is clearly illustrated.

1 Introduction

A neural network model that can be used to cluster arbitrary binary or analog data was derived by Carpenter et al. [1]. This model is termed Fuzzy ART in reference to the adaptive resonance theory. One of the major reasons for the development of Fuzzy ART was to remedy the inability of ART1, as well as Predictive ART architectures based on ART1 modules, to classify analog data. Although the learning properties of ART1 are well understood, the same cannot be said for the Fuzzy ART algorithm.

In this paper we present useful properties of the Fuzzy ART algorithm which facilitate the understanding of its operation. For clarity purposes we split the properties into four different categories: template properties (Section 3), access properties (Section 4), reset properties (Section 5), and properties related to the number of list presentations needed for the weight stabilization (Section 6). These properties are presented in the form of theorems and corollaries (their proofs are omitted in order to save space; the interested reader can find the proofs in [2]). Some of the properties discussed in this paper involve the size/similarities of templates created in Fuzzy ART, as well as the number of list presentations required to learn an arbitrary list of binary input patterns repeatedly presented to Fuzzy ART. For most of the Fuzzy ART properties mentioned in this manuscript, the effects of parameters α and ρ are clearly illustrated.

2 Preliminaries–Notations

The Fuzzy ART algorithm is described in detail by Carpenter et al. [1]. In this section we only provide information that is necessary to understand the results developed here. The Fuzzy ART architecture consists of two layers of nodes, designated F_1 and F_2. Inputs are presented at the F_1 layer. If $\mathbf{a} = (a_1, ..., a_M)$ denotes a vector, with each of its components in the interval $[0, 1]$, then the input to the F_1 layer is a vector $\mathbf{I}$ such that

$$\mathbf{I} = (\mathbf{a}, \mathbf{a}^c) = (a_1, ..., a_M, a_1^c, ..., a_M^c) \quad (1)$$

where

$$a_i^c = 1 - a_i \; ; \quad 1 \le i \le M \quad (2)$$

This type of transformation, called *complement coding*, is necessary for the successful operation of Fuzzy ART, especially when the input vector $\mathbf{I}$ is of analog nature (for more details see [1]). The F_2 layer in Fuzzy ART is usually referred to as the *category representation layer* because its nodes denote the categories to which the input patterns belong.

The F_1 layer has $2M$ nodes, while the F_2 layer has N nodes. We use the index i to designate nodes in the F_1 layer and the index j to designate nodes in the F_2 layer. There are *bottom-up weight* connections emanating from the nodes in the F_1 layer and converging to the nodes in the F_2 layer. Similarly, there are *top-down weight* connections emanating from the nodes in the F_2 layer and converging to the nodes in the F_1 layer. The bottom-up weights converging to a node in the F_2 layer can be completely defined by the top-down weights emanating from this F_2 node. Hence, only the set of top-down weights need to be defined. In particular, we let $\mathbf{W}_j = (W_{j1}, ..., W_{j(2M)})$ designate the vector of top-down weights emanating from node j in the F_2 layer.

When an input pattern $\mathbf{I}$ is applied at the F_1 layer, it produces an input $T_j(\mathbf{I})$ at node j in the F_2 layer. This input $T_j(\mathbf{I})$ is given by the following equation:

$$T_j(\mathbf{I}) = \frac{|\mathbf{I} \wedge \mathbf{W}_j|}{\alpha + |\mathbf{W}_j|} \quad (3)$$

The function $T_j(\mathbf{I})$ is referred to as the *choice function*. In the above equation, α is a Fuzzy ART network parameter, called the *choice parameter*, $\mathbf{I}\wedge\mathbf{W}_j$ is a vector whose i-th component ($1 \leq i \leq 2M$) is equal to the minimum of $\mathbf{I}_i$ and $\mathbf{W}_{ji}$ (the operator $\wedge$ is referred to as the MIN operator and $\mathbf{I}\wedge\mathbf{W}_j$ is called the MIN of $\mathbf{I}$ and $\mathbf{W}_j$), and $|\cdot|$ designates the size of a vector, where the size of a vector is defined to be the sum of its components.

The node in the F_2 layer which receives the largest input $T_j(\mathbf{I})$ will be chosen to represent the input pattern $\mathbf{I}$. Assume that node J in the F_2 layer receives the largest such input. The appropriateness of J to represent the input pattern is based on the *vigilance criterion*. This criterion is satisfied if

$$\frac{|\mathbf{I}\wedge\mathbf{W}_J|}{|\mathbf{I}|} \geq \rho \tag{4}$$

where ρ, the *vigilance parameter*, may take values in the interval $[0,1]$. If node J fails the vigilance criterion, it is reset and a search for another node in the F_2 layer to represent the input pattern starts. The reset of node J is accomplished by the *orienting subsystem* in Fuzzy ART. If node J passes the vigilance criterion, learning starts and the top-down weight vector $\mathbf{W}_J$ is updated as follows:

$$\mathbf{W}_J = (1-\beta)\mathbf{W}_J + \beta(\mathbf{I}\wedge\mathbf{W}_J) \tag{5}$$

where β is a third Fuzzy ART parameter, called *learning rate*, which may assume values in the interval $(0,1]$. If $\beta = 1$ the learning is called *fast learning*, and if $0 < \beta < 1$ it is called *slow learning*. If a node has previously coded an input pattern, then it is said to be *committed*; otherwise, it is said to be *uncommitted*. A special type of slow learning, called *fast-commit slow-recode*, is one in which fast learning occurs (i.e., $\beta = 1$) when the chosen node in the F_2 layer is uncommitted, and slow learning occurs (i.e., $0 < \beta < 1$) when the chosen node is committed.

The vector of top-down weights from a committed node in the F_2 layer is called a *template*. Consider now an input pattern $\mathbf{I}$ presented to the Fuzzy ART architecture, and an arbitrary template denoted by $\mathbf{W}_j$. A component of an input pattern $\mathbf{I}$ is indexed by i if it affects node i in the F_1 layer, and the corresponding component of template $\mathbf{W}_j$ is $\mathbf{W}_{ji}$. We can identify three types of templates with respect to an input pattern $\mathbf{I}$: subset templates, superset templates and mixed templates. A template $\mathbf{W}_j$ is a *subset template* of an input pattern $\mathbf{I}$ if each one of the $\mathbf{W}_j$ components is smaller than or equal to its corresponding components in $\mathbf{I}$. A template $\mathbf{W}_j$ is a *superset template* of an input pattern $\mathbf{I}$ if each one of the $\mathbf{W}_j$ components is larger than its corresponding components in $\mathbf{I}$. A template is a *mixed template* if some of the $\mathbf{W}_j$ components are smaller than or equal to its corresponding components in $\mathbf{I}$, and the rest of the $\mathbf{W}_j$ components are larger than its corresponding components in $\mathbf{I}$. With reference to an input $\mathbf{I}$, we designate a committed node in the F_2 layer as subset, superset or mixed depending on whether its corresponding template is a subset, superset or mixed with respect to $\mathbf{I}$. It is worth noting that in the case of fast-learning or fast-commit slow-recode learning we can only define subset, mixed, and uncommitted templates. Due to the complement coding nature of the input patterns, superset templates cannot be created in a Fuzzy ART architecture with fast-learning or fast-commit slow-recode learning.

The following assumptions will be used at various points in the remainder of the paper to guarantee the validity of specific results: (i) fast learning, (ii) fast-commit slow-recode learning, (iii) binary input patterns, (iv) repeated or cyclic presentations of an input list of patterns, and (v) a sufficient number of nodes in the F_2 layer. Assumptions (i) and (ii) have been discussed previously. Assumption (iii) implies that the input patterns presented to the Fuzzy ART architecture have binary (0 or 1) components. Most of the results in this paper are valid for analog or binary data. Only the properties presented in Section 6 require assumption (iii). Assumption (iv) corresponds to the case where we have a list of input patterns, designated $\mathbf{I}^1, \mathbf{I}^2, ..., \mathbf{I}^P$, which is presented either repeatedly or cyclically to Fuzzy ART. In *repeated presentations* of the list, the order of the pattern presentation within the list is of no consequence, but in *cyclic presentations* of the list the patterns are always presented in the same order within each list (e.g. $\mathbf{I}^1, \mathbf{I}^2, ..., \mathbf{I}^P, \mathbf{I}^1, \mathbf{I}^2, ..., \mathbf{I}^P$, and so on). Assumption (v) means that every time an input pattern $\mathbf{I}$ is presented to Fuzzy ART there is at least one uncommitted node available at the F_2 layer; this assumption is sufficient to guarantee that an appropriate node in the F_2 layer will always be found to represent the input pattern.

In the case where a list of input patterns is repeatedly presented to the Fuzzy ART architecture, it is reasonable to ask how many list presentations does Fuzzy ART need to learn the input list; or equivalently, how many list presentations are needed for the weights to stabilize. We say that the weights in the Fuzzy ART architecture are stabilized by the end of the n-th list presentation, if in subsequent presentations of the list (i.e., list presentations $\geq n+1$) weights cannot be modified. Under the aforementioned scenario (i.e., repeated presentations of a list of input patterns) when weights are stabilized we also say that *learning* (of the list) *is complete*. After stabilization of the weights occurs, each pattern from the input list will have *direct access* to a node in the F_2 layer (Assume there are enough nodes in the F_2 layer). We say that a pattern $\mathbf{I}$ has direct access to a node j in the F_2 layer if immediately after the presentation of $\mathbf{I}$ at the F_1 layer, j is chosen first and no reset of j occurs.

3 Template Properties

In this section we discuss properties related with the templates created in a Fuzzy ART architecture. In particular, Theorem 3.1 states that the templates in Fuzzy ART are distinct, and Theorem 3.2 deals with the smallest possible size of the templates. Corollary 3.2 shows how the range of α is related to the smallest possible template size. Theorem 3.3 focuses on the similarity among the templates in Fuzzy ART. Finally, Corollary 3.3 shows how this similarity is affected by the ranges of the α and ρ parameters.

THEOREM 3.1

In a Fuzzy ART architecture, all the templates are distinct.

REMARKS: This theorem shows one of the good properties of Fuzzy ART: the templates can never be the same. It applies to binary or analog patterns, fast or slow learning and for any values of the α and ρ parameters.

THEOREM 3.2

In a Fuzzy ART architecture with a sufficient number of nodes in the F_2 layer, the size of a template is larger than $\alpha M/(\alpha + M)$. For the binary patterns and fast learning case, the size of a template is larger than or equal to $(\alpha + 1)M/(\alpha + M)$.

Corollary 3.2

In a Fuzzy ART architecture with binary patterns, fast learning, and a sufficient number of nodes in the F_2 layer, if $\alpha > M(M - L - 1)/L$, then the smallest possible template size is equal to $M - L + 1$ and there are at most L different template sizes, where L is an integer in the interval $[1, M - 1]$.

REMARKS: Theorem 3.2 and Corollary 3.2 are valid independently of the value of the vigilance parameter ρ. The smallest possible template size increases as α increases. Furthermore, it is worth observing that with a sufficient number of nodes in F_2 layer size–1 templates cannot be created.

THEOREM 3.3

In a Fuzzy ART architecture with either fast-commit slow-recode or fast learning, and a sufficient number of nodes in the F_2 layer, the size of the MIN of any two templates (the number of common 1's between any two templates in the binary input patterns and fast learning case)

$$|\mathbf{W}_1 \wedge \mathbf{W}_2| < \max\left\{\rho M, M\frac{\alpha + M}{\alpha + 2M}\right\} \quad (6)$$

Corollary 3.3

Under the conditions stated in Theorem 3.3, if $\alpha \leq (M - 2L)M/L$ and $\rho \leq 1 - L/M$, then $|\mathbf{W}_1 \wedge \mathbf{W}_2| < M - L$, where $0 \leq L < M/2$.

REMARKS: Corollary 3.3 is a direct result of Theorem 3.3. They both show that the closeness of any two templates increases as ρ or α increases. The consequences of Corollary 3.2 and Corollary 3.3 are depicted in Tables 1 and 2.

Table 1: Consequences of Corollary 3.2 for $M = 10$

Range of α	Max. Number of Template Sizes	Smallest Template Size
(0, 1.25]	9	2
(1.25, 20/7]	8	3
(20/7, 5]	7	4
(5, 8]	6	5
(8, 12.5]	5	6
(12.5, 20]	4	7
(20, 35]	3	8
(35, 80]	2	9
(80, ∞)	1	10

Table 2: Consequences of Corollary 3.3 for $M = 10$

Range of α	Range of ρ	$\|\mathbf{W}_1 \wedge \mathbf{W}_2\|$
(30, 80]	(0.8, 0.9]	<9
(40/3, 30]	(0.7, 0.8]	<8
(5, 40/3]	(0.6, 0.7]	<7
(0, 5]	(0, 0.6]	<6

4 Access Properties

This section is entitled "Access properties" because here we present properties related with what type of nodes in the F_2 layer will be chosen during a pattern's presentation. In particular, Theorem 4.1 discusses the order of search among the nodes in the F_2 layer during a pattern's presentation. Theorem 4.2 states that with fast learning, uncommitted nodes in the F_2 layer will not be chosen after the first presentation of a list of input patterns. Theorem 4.3 states that a pattern will always directly access a node with a template equal to the pattern. Finally Theorem 4.4 verifies that under certain conditions, after learning of an input list of patterns is complete, there may exist committed nodes in the F_2 layer that are not directly accessed by any pattern from the input list.

In Fuzzy ART, the search order among the nodes in the F_2 layer depends on the choice parameter α. If α is small, a pattern tends to choose a node with the largest ratio $|\mathbf{I} \wedge \mathbf{W}_j|/|\mathbf{W}_j|$, regardless of the size of $|\mathbf{I} \wedge \mathbf{W}_j|$. Therefore, subset nodes always have priority over other nodes. If α is large, the size $|\mathbf{I} \wedge \mathbf{W}_j|$ plays a more important role in the choice of a node in the F_2 layer. For any α Fuzzy ART follows the rules stated in Theorem 4.1:

THEOREM 4.1

In a Fuzzy ART architecture, when an input pattern $\mathbf{I}$ is presented at the F_1 layer, a node in the F_2 layer is chosen according to the following rules:

(a) *A subset node (if there is one) will be chosen over an uncommitted node.*

(b) *Among all the subset nodes, the node with the largest template will be chosen first.*

(c) *If a mixed node j with template $\mathbf{W}_j$ is accessed prior to a subset node J with template $\mathbf{W}_J$ then $|\mathbf{I}\wedge\mathbf{W}_j| > |\mathbf{W}_J|$ must hold.*

(d) *If there are no subset nodes, and for every mixed node j: $\frac{|\mathbf{I}\wedge\mathbf{W}_j|}{|\mathbf{W}_j|} \leq 0.5$, then an uncommitted node will be chosen over any mixed node.*

THEOREM 4.2

In a Fuzzy ART architecture with fast learning and repeated presentations of a list of input patterns, no uncommitted node will be chosen after the first list presentation. As a result, the total number of committed nodes (or templates) cannot exceed the total number of patterns in the input list.

REMARKS: This theorem provides an upper bound for the number of nodes needed in the F_2 layer so that Fuzzy ART will learn all the patterns in an input list, provided that fast learning is employed. In practice, the number of categories (i.e., the number of nodes needed in the F_2 layer) is usually much less than the number of patterns in the input list, and is an increasing function of the choice parameter α and the vigilance parameter ρ.

THEOREM 4.3

In a Fuzzy ART architecture, if a node J in the F_2 layer has perfectly learned an input pattern $\mathbf{I}$ (i.e., $\mathbf{W}_J=\mathbf{I}$), then when $\mathbf{I}$ is presented it will directly access node J.

THEOREM 4.4

In a Fuzzy ART architecture with repeated presentations of a list of input patterns, after learning is complete, there may exist committed nodes in the F_2 layer that are not directly accessed by any pattern in the input list.

5 Reset Properties

The properties discussed in this section are byproducts of the theorems mentioned earlier. They are important to report though because they provide a different perspective of viewing these theorems; this perspective involves the orienting subsystem in Fuzzy ART. For example, Theorem 5.1 states that under certain assumptions, no reset events are possible after the first presentation of a list of input patterns, while Theorem 5.2 and Corollary 5.2 determine the effective range of the vigilance parameter, that is the range of ρ values that will allow reset events to occur. Theorem 5.2 and Corollary 5.2 are also useful in directing us to choose appropriate α and ρ values for Fuzzy ART simulations.

THEOREM 5.1

In a Fuzzy ART architecture with fast learning, and repeated presentations of a list of input patterns, no reset will occur after the first list presentation.

REMARKS: Theorem 5.1 tells us that with fast learning and repeated presentations of a list of input patterns, for list presentations ≥ 2, there is no need to check on the vigilance criterion. In terms of hardware, the orienting subsystem becomes inactive (automatically disengaged) after the first list presentation. In terms of a software simulation of Fuzzy ART, we can disregard the orienting subsystem after the first list presentation in order to speed up the learning.

THEOREM 5.2

In a Fuzzy ART architecture with a sufficient number of nodes in the F_2 layer, the vigilance parameter ρ should be larger than $\frac{\alpha}{\alpha+M}$. In other words, if $\rho \leq \frac{\alpha}{\alpha+M}$, no reset will ever occur. In the case of binary patterns and fast learning, the vigilance parameter ρ should be larger than $\frac{\alpha+1}{\alpha+M}$.

REMARKS: This theorem demonstrates that if α is large, small vigilance cannot be effective. For example, if $\alpha = M$, the vigilance should be larger than 0.5, because if $\rho \leq 0.5$, no reset will ever occur. On the other hand, given a vigilance parameter ρ, if $\alpha \geq \rho M/(1-\rho)$, no reset will ever occur. Theorem 5.2 illustrates that the Fuzzy ART parameters α and ρ should be carefully chosen if we want the vigilance parameter ρ to have an effect on the operation of the network.

Corollary 5.2

In a Fuzzy ART architecture with binary patterns, fast learning, and a sufficient number of nodes in the F_2 layer, if $\alpha > M(M-L-1)/L$, then the vigilance ρ should be larger than $1-(L-1)/M$, where L is an integer taking values in the interval $[2, M-1]$.

6 Number of the List Presentations Needed

In this section, we assume that a list of input patterns is repeatedly presented to the Fuzzy ART architecture, and we derive results related to the number of list presentations required by Fuzzy ART to learn this list. In particular, Theorem 6.1 states that if the choice parameter α is relatively small then learning in Fuzzy ART will be completed in one list presentation. Theorem 6.2 provides an upper bound on the number of list presentations needed by Fuzzy ART to learn the input list when α is relatively large (i.e., when α is not necessarily as small as it is required to validate Theorem 6.1).

THEOREM 6.1

In a Fuzzy ART architecture with binary patterns, fast learning, a sufficient number of nodes in the F_2 layer, and repeated presentations of a list of input patterns, if $\alpha \leq \rho/(1-\rho)$, then the weights will be stabilized in 1 list presentation.

REMARKS: (1) In the extreme case where $\rho = 1$, each pattern from the input list will choose a different node in the F_2 layer. In this case, for any value of α the weights will be stabilized in one list presentation. (2) By Corollary 3.2, for binary input patterns and fast learning, the smallest possible template size is greater than or equal to 2, and as a result the vigilance parameter ρ should be larger than $2/M$. Therefore, if $\rho \leq 2/M$ (including 0), Theorem 6.1 is valid for $\alpha \leq 2/(M-2)$. (3) A sufficient condition on the α values that will guarantee the validity of Theorem 6.1, even when the input patterns are analog, is

$$\alpha \leq \frac{\rho(|\mathbf{W}_j| - |\mathbf{I} \wedge \mathbf{W}_j|)}{1-\rho} \qquad (7)$$

where $\mathbf{W}_j$ is any mixed template of pattern **I**. Unfortunately, even if we know the exact ρ value, we cannot find a lower bound for the right hand side of (7) because $|\mathbf{W}_j| - |\mathbf{I} \wedge \mathbf{W}_j|$ can be arbitrarily small in the analog case. In conclusion, we can only state that Theorem 6.1 is also valid for analog input patterns except that α has to be chosen very, very small.

Before we present the next theorem on the number of list presentations, let us state a lemma first.

Lemma 1

In a Fuzzy ART architecture with binary patterns, fast learning, and repeated presentations of a list of input patterns, the following rules are valid in list presentations ≥ 2 ($|\mathbf{W}|$ is the minimum template size at the end of the first list presentation).

(1) *No template of size $|\mathbf{W}|$ or smaller can be created.*

(2) *A template of size $|\mathbf{W}|$ or $|\mathbf{W}|+1$ cannot be modified.*

(3) *A template of size $|\mathbf{W}|+L$ ($L \geq 2$) can be modified only by the patterns which have largest subset template of size $H \leq |\mathbf{W}|+L-2$. And the new template size should be greater than or equal to $H+1$.*

THEOREM 6.2

In a Fuzzy ART architecture with binary patterns, fast learning, a sufficient number of nodes in the F_2 layer, and repeated presentations of a list of input patterns, if $\alpha > \frac{1}{L}M(M-L-1)$ or $\rho > 1 - L/M$, then the weights will be stabilized in $L-1$ list presentations, where L is an integer taking the value in the interval $[2,4]$.

REMARKS: It is worth mentioning that Theorem 6.2 is valid for $L = 5$ under the additional assumptions that the input patterns are presented cyclically and $M \geq 9$. The consequences of Theorem 6.1 and Theorem 6.2 for $L \leq 5$, where M is arbitrary or $M = 10$, are depicted in Table 3. We could not prove Theorem 6.2 for $L > 5$ because its proof becomes very complicated. However, by looking at the results of Table 3, we are encouraged to believe that there seems to be a pattern relating the range of α or ρ values and the number of list presentations needed to learn an arbitrary binary list repeatedly presented to Fuzzy ART. Hence, we were tempted to formulate a conjecture that extends the results of Table 3 over the entire range of α and ρ value. We decided not to do so because of the additional assumptions needed to verify the validity of Theorem 6.2 for $L = 5$ (i.e., cyclic presentations of the input list and $M \geq 9$). Nevertheless it is worth mentioning that out of hundreds of simulations performed with random input patterns we found that the maximum number of list presentations needed for weight stabilization in Fuzzy ART was 3 for 2 simulations and 2 for the rest of the simulations.

7 Summary

We have examined the Fuzzy ART algorithm carefully from a number of different perspectives. For example, in Section 3 we demonstrated that Fuzzy ART templates are distinct , we calculated a lower bound on the template size and we found an upper bound on the similarity of the templates created. Furthermore, in Section 4, we focused on access properties, investigating the order of search of the F_2 layer nodes, finding an upper bound on the number of nodes needed in the F_2 layer of Fuzzy ART to learn an arbitrary list of input patterns, and proving the direct access property of patterns to perfectly learned templates. Also, in Section 5, we concentrated on the orienting subsystem and reset events and elaborated on the interrelationship between the α and ρ parameter values needed in Fuzzy ART simulations. Finally, in Section 6, we shifted our attention to the number of list presentations required by Fuzzy ART to learn an arbitrary list of patterns repeatedly presented to it. Most of the results presented in Section 6 were valid for binary input patterns and fast learning. The strongest result (Theorem 6.1) stated that for small α values (i.e., $\alpha \leq \rho/(1-\rho)$), learning will be complete in one list presentation. Weaker results were also presented in Section 6, where in order to come up with a definite upper bound on the number of list presentations needed we restricted either the α range or the ρ range.

Table 3: Consequences of Theorem 6.1 and Theorem 6.2 for $L \leq 5$

(a) General Case

Range of α		Range of ρ	Number of template sizes	Number of lists needed
$\alpha \in (0, max(\frac{\rho}{1-\rho}, \frac{2}{M-2})]$	and	$\rho \in (0, 1]$	$\leq M-1$	1
$\alpha \in (\frac{1}{2}M(M-3), \infty)$	or	$\rho \in (1-\frac{2}{M}, 1]$	≤ 2	1
$\alpha \in (\frac{1}{3}M(M-4), \frac{1}{2}M(M-3)]$	or	$\rho \in (1-\frac{3}{M}, 1-\frac{2}{M}]$	≤ 3	≤ 2
$\alpha \in (\frac{1}{4}M(M-5), \frac{1}{3}M(M-4)]$	or	$\rho \in (1-\frac{4}{M}, 1-\frac{3}{M}]$	≤ 4	≤ 3
$\alpha \in (\frac{1}{5}M(M-6), \frac{1}{4}M(M-5)]$	or	$\rho \in (1-\frac{5}{M}, 1-\frac{4}{M}]$	≤ 5	$\leq 4^*$

(b)Special Case for $M = 10$

Range of α		Range of ρ	# of template sizes	# of lists needed
$\alpha \in (0, max(\frac{\rho}{1-\rho}, 0.25)]$	and	$\rho \in (0, 1]$	≤ 9	1
$\alpha \in (35, \infty)$	or	$\rho \in (0.8, 1]$	≤ 2	1
$\alpha \in (20, 35]$	or	$\rho \in (0.7, 0.8]$	≤ 3	≤ 2
$\alpha \in (12.5, 20]$	or	$\rho \in (0.6, 0.7]$	≤ 4	≤ 3
$\alpha \in (8, 12.5]$	or	$\rho \in (0.5, 0.6]$	≤ 5	$\leq 4^*$

*Note: The last line (corresponding to the extention of Theorem 6.2 for $L = 5$) is valid under the assumptions that the input patterns are presented cyclicly and $M \geq 9$

References

[1] Carpenter, G.A., Grossberg, S., and Rosen, D.B. , "Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System," *Neural Networks* 4(6), pp.759-771, 1991.

[2] Huang, J., Georgiopoulos, M., and Heileman, G. L. , "Fuzzy ART Properties," submitted to *Neural Networks*, August 1993.