PERGAMON

Neural Networks 15 (2002) 1205–1221

Neural Networks

www.elsevier.com/locate/neunet

Category regions as new geometrical concepts in Fuzzy-ART and Fuzzy-ARTMAP

Georgios C. Anagnostopoulos[a,b,*], Michael Georgiopoulos[b]

[a]4132 Lake Underhill Road #303, Orlando, FL 32803, USA

[b]School of Electrical Engineering and Computer Science, University of Central Florida, 4000 Central Florida Boulevard, Engineering Building, Room 407, Orlando, FL 32816, USA

Received 7 December 2001; revised 8 May 2002; accepted 8 May 2002

Abstract

In this paper we introduce novel geometric concepts, namely category regions, in the original framework of Fuzzy-ART (FA) and Fuzzy-ARTMAP (FAM). The definitions of these regions are based on geometric interpretations of the *vigilance test* and the F_2 layer competition of committed nodes with uncommitted ones, that we call *commitment test*. It turns out that not only these regions have the same geometrical shape (polytope structure), but they also share a lot of common and interesting properties that are demonstrated in this paper. One of these properties is the shrinking of the volume that each one of these polytope structures occupies, as training progresses, which alludes to the stability of learning in FA and FAM, a well-known result. Furthermore, properties of learning of FA and FAM are also proven utilizing the geometrical structure and properties that these regions possess; some of these properties were proven before using counterintuitive, algebraic manipulations and are now demonstrated again via intuitive geometrical arguments. One of the results that is worth mentioning as having practical ramifications is the one which states that for certain areas of the vigilance-choice parameter space (ρ,a), the training and performance (testing) phases of FA and FAM do not depend on the particular choices of the vigilance parameter. Finally, it is worth noting that, although the idea of the category regions has been developed under the premises of FA and FAM, category regions are also meaningful for later developed ART neural network structures, such as ARTEMAP, ARTMAP-IC, Boosted ARTMAP, Micro-ARTMAP, Ellipsoid-ART/ARTMAP, among others.

Keywords: Adaptive resonance theory; Fuzzy-ART; Fuzzy-ARTMAP; Category regions; Match region; Choice region; Claim region; Commitment test

1. Introduction

Fuzzy-ART (FA) (Carpenter, Grossberg, & Rosen, 1991) and *Fuzzy-ARTMAP* (FAM) (Carpenter, Grossberg, Markuzon, Reynolds, & Rosen, 1992) are two self-organizing, neural network architectures based on the *Adaptive Resonance Theory* (ART) introduced by Grossberg (1976), which addresses the *stability–plasticity dilemma* occurring in learning systems. FA performs unsupervised clustering of its input data. On the other hand, FAM consists of two FA networks bridged via an *inter-ART module* and is capable of forming associative maps between clusters of its input and output domains in a supervised manner. As a special case, when the output domain is a finite set of class labels, FAM can be utilized as a classifier.

There are many desirable properties of learning and characteristics associated to FA/FAM. First, both networks are capable of learning in both off-line (batch) and on-line (incremental) training modes. Under batch mode and *fast learning rule* (Carpenter et al., 1991, 1992) assumptions, both exhibit fast, stable and finite learning: the networks' knowledge representation stabilizes (*self-stabilization property*) relatively fast after a finite number of *list presentations* (epochs). In the case of an FAM classifier, this last property makes it a consistent classifier (Bezdek, Reichherzer, Lim, & Attikiouzel, 1998), since using fast learning its resubstitution error (Bezdek et al., 1998) becomes zero. Furthermore, they both feature novelty detection mechanisms that identify input patterns not typical of previously experienced inputs. Also, due to the specific nature of their neural architecture, responses of FA and FAM to specified inputs are easily explained (Carpenter &

* Corresponding author. Address: 4132 Lake Underhill Road #303, Orlando, FL 32803, USA.

E-mail addresses: anagnostop@email.com (G.C. Anagnostopoulos), michaelg@mail.ucf.edu (M. Georgiopoulos).

0893-6080/02/$ - see front matter
PII: S0893-6080(02)00063-1

Nomenclature

$\varnothing$	the empty set
R	the set of real numbers
U	the set [0,1]
M	the feature space dimensionality
U^M	the set $[0,1]^M$
R^M	the M-dimensional Euclidian space
$\subseteq$	subset of
$\subset$	proper subset of
$\forall$	for all
$\lvert\ \rvert$	L_1 vector norm
$\mathbf{x}^c$	complement coded form of pattern $\mathbf{x}$
$\mathbf{1}$	all-ones vector
$\mathbf{w}_j$	template of category j
$\wedge$	Fuzzy-min operator
$s(\mathbf{w}_j)$	size of category j
$\mathrm{dis}(\mathbf{x},\mathbf{w}_j)$	L_1-norm distance of pattern $\mathbf{x}$ from category j
$R(\mathbf{w}_j)$	representation region of category j
ρ	baseline vigilance parameter value
a	choice parameter value
w_u	initial weight value for templates of uncommitted nodes
γ	learning rate parameter value
$\rho(\mathbf{w}_j\lvert\mathbf{x})$	match function value for category j with respect to pattern $\mathbf{x}$
$T(\mathbf{w}_j\lvert\mathbf{x})$	choice function value for category j with respect to pattern $\mathbf{x}$
T_u	choice function value of uncommitted nodes
$V(\mathbf{w}_j\lvert\rho)$	match (vigilance) region of category j
$d_V(\mathbf{w}_j\lvert\rho)$	match region radius of category j
$C(\mathbf{w}_j\lvert a,w_u)$	choice (commitment) region of category j
$d_C(\mathbf{w}_j\lvert a,w_u)$	choice region radius of category j
$L(\mathbf{w}_j\lvert\rho,a,w_u)$	claim region of category j
$d_L(\mathbf{w}\lvert\rho,a,w_u)$	claim region radius of category j

Tan, 1995), in contrast to other neural network models, where, in general, it is more difficult to explain why an input pattern **x** produced an output **y**. Properties of learning for FA and FAM can be found in their original references (Carpenter et al., 1991, 1992), as well as in the work of others (Georgiopoulos, Fernlund, Bebis, & Heileman, 1996; Huang, Georgiopoulos, & Heileman, 1995).

There have been many contributions in the ART literature over the last decade. We only refer to a limited number of them: ARTEMAP (Carpenter & Ross, 1995), Gaussian ARTMAP (Williamson, 1996), dART (Carpenter, 1997), dARTMAP (Carpenter, Milenova, & Noeske, 1998), ARTMAP-IC (Carpenter & Markuzon, 1998), Boosted ARTMAP (Verzi, Heileman, Georgiopoulos, & Healy, 1998), Micro-ARTMAP (Gomez Sanchez, Dimitriadis, Cano Izquierdo, & Lopez Coronado, 2000), Topographic Attentive Mapping network (Williamson, 2001) and finally Ellipsoid-ART/ARTMAP (Anagnostopoulos & Georgiopoulos, 2001). The above contributions revolve around modifications and enhancements as well as around new approaches based on the concepts of the original FA and FAM architectures. However, there are other, independent developments of similar ART-like structures like Fuzzy-Min–Max (Simpson, 1992), LAPART2 (Healy & Caudell, 1998), and σ-FLNMAP (Petridis, Kaburlasos, Fragkou, & Kehagias, 2001).

Both FA and FAM operate by summarizing similar training data into groups, which we define as *FA categories*. These categories constitute the building block of knowledge/memory representation for both architectures and are formed in a self-organizing manner. FA/FAM exhibit stability in learning by remembering the characteristics of the already-formed categories and exhibit plasticity in learning by allowing updates of existing categories in a non-destructive fashion or via the creation of new categories, if no update is possible. Significant insight has been gained in the past by attributing a geometrical interpretation to these categories (Carpenter et al., 1991, 1992). On the same tangent, we introduce in this paper a new geometrical perspective related to FA categories, which sheds more light into the process of determining eligible categories to compete for a newly presented input pattern. By eligible categories we mean categories that potentially may be chosen after the node competition.

The definitions of these regions are based on geometric interpretations of the *vigilance test* (VT, defines match regions), and the F_2 layer competition of committed nodes with uncommitted ones that we call *commitment test* (CT, defines choice regions). It turns out not only that these regions (match and choice) have the same geometrical shape (polytope structure), but also share a lot of common and interesting properties that are demonstrated in this paper. For example, the regions' definitions are of the same form, the regions' geometric representations are of the same shape (but may be of different size), ART network parameters enforce a maximum hyper-volume for these regions, both regions contain the representation region (corresponds to the hyper-box that every category in ART defines) and other commonalities, which reaffirms the clustering-by-similarity nature of the VT and CT. One of these common properties of category regions is the shrinking of the volume that each one of these polytope structures occupies, as training progresses, which alludes to the stability of learning of FA and FAM (a well-known result). These regions also exhibit different characteristics. For instance, a category update defines a new match region that is completely included in the old match region (the one corresponding to the category before its update). On the other hand, a category update defines a new choice region that is not completely included in the old choice region (the one corresponding to the category prior to its update); only the hyper-volume of the choice region decreases as we have emphasized above. A complete enumeration of the

Table 1
Comparison of match and choice region properties

Category match region properties	Category choice region properties
Match region size depends on category size and is tunable via ρ. Ranges from size 0 to the entire input domain.	Choice region size depends on category size and is tunable via a and w_u. Ranges from size 0 to the entire input domain.
Upon category update (expansion), the match region contracts.	Upon category update (expansion), the new choice region is not completely contained in the original one.
The match region enforces a maximum category size of $M(1-\rho)$.	The choice region enforces a (not-attainable) least upper bound for category sizes of $((2w_u-1)M^2)/((2w_u-1)M+a)$.
After a category update (expansion), patterns that used to be located outside the original match region, will also be outside the new match region.	After a category update (expansion), there will be patterns that used to be located outside the original choice region, but will be inside the new choice region.
The regions' definitions are of the same form, thus, they are of same shape, but have different radii.	
Both regions always contain the representation region.	
Upon category update (expansion), the regions' hyper-volume decreases.	

properties of the match and choice regions and their comparisons is depicted in Table 1.

Finally, it is worth mentioning that the intersection of the match region and the choice region defines a new region, named *claim region*, whose geometrical structure is the same as the geometrical structure of the match and choice regions (polytope shape). The claim region has an interesting interpretation. The claim region of a category in ART contains all the points in the pattern space that could be potentially encoded by this category. Due to its definition, no points outside a category's claim region can be encoded by this category. Hence, it can be thought of as the region of attraction of an ART category. A point in the pattern space can belong to more than one claim regions. The category that will eventually encode this point will be the category that wins the competition amongst all the categories whose claim region includes this point. The competition in ART is won by the category that produces the maximum choice function value. Defining geometrically, or through an equation, the region of attraction of cluster points (or prototypes) of any pattern classifier (not necessarily an ART neural network) is a worthwhile endeavor in the pattern recognition literature. Part of our work in this paper has accomplished this task for FA and FAM for both fast and slow learning. It is finally, worth mentioning that although the idea of category regions (choice, match, and claim regions) has been developed under the premises of FA and FAM, these category regions and their associated properties are valid for later-developed (than FA and FAM) ART neural network structures, such as ARTEMAP, ARTMAP-IC, Boosted ARTMAP, Micro-ARTMAP, Ellipsoid-ART/ARTMAP, among others. In the sequel, we will assume that the reader is already familiar with the basic elements and functionality of FA/FAM.

The rest of the paper is organized as follows. In Section 2, we provide limited, but necessary background regarding the concept of FA categories. Section 3 formally defines the CT as the competition with an uncommitted node in the F_2 layer of an FA module. In Section 4, we introduce the concept of category regions and mention some of their most interesting properties. Section 5 presents a group of results that are based on the idea of category regions and apply to both FA and FAM. A comprehensive summary of the presented material is being provided in Section 6. Finally, proofs of selected category region, properties and results can be found in Appendix A.

2. Fundamentals of FA categories

We begin this section by introducing some useful notation. If A and B are two sets, then by $A \subseteq B$ ($A \subset B$) we mean that A is a (proper) subset of B and $A - B$ denotes the set $\{x \in A \text{ but } x \notin B\}$. Let R be the set of real numbers and $U^M \in [0,1]^M$ denote the closure of the M-dimensional unit hyper-cube that serves as an input space for any FA module (network). A block diagram of such a module is displayed in Fig. 1. We define as $|\cdot| : U^{2M} \to R$ to be the L_1-norm for the U^{2M} domain; the same notation is also going to be used for the U^M domain L_1-norm. Additionally, we define the Fuzzy-min-operator $\wedge : U^{2M} \times U^{2M} \to U^{2M}$, such that, if $\mathbf{w}_1, \mathbf{w}_2 \in U^{2M}$ and $\mathbf{w}_3 = \mathbf{w}_1 \wedge \mathbf{w}_2$, then the mth component w_{3m} of vector $\mathbf{w}_3$ is $w_{3m} = \min\{w_{1m}, w_{2m}\}$. A pattern serving as an input to an FA module will be denoted as $\mathbf{x}$ and $\mathbf{x}^c = [\mathbf{x} \ \ \mathbf{1}-\mathbf{x}] \in U^{2M}$ will denote its complement coded form, where $\mathbf{1}$ is the all-ones vector. Note that $|\mathbf{x}^c| = M$ for all $\mathbf{x}$. Complement coding occurs in an FA module's F_0 layer, as shown in Fig. 1, and vector $\mathbf{x}^c$ serves as the input vector to the F_1 layer. All aforementioned quantities are row vectors. The information describing each FA category j of an FA module is stored in a *template*, which is a vector of the form $\mathbf{w}_j = [\mathbf{u}_j \ \ \mathbf{1}-\mathbf{v}_j] \in U^{2M}$ and $\mathbf{u}_j, \ \mathbf{v}_j \in U^M$. A template $\mathbf{w}_j$ is the top-down weight vector related to the connections from the jth node in the F_2 layer to all nodes in

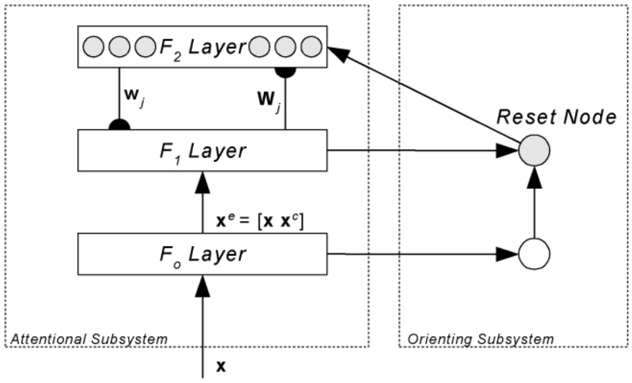

Fig. 1. Block diagram of a Fuzzy-ART module.

the F_1 layer. Due to FA/FAM's learning schemes, for every template it always holds that $u_{jm} \leq v_{jm}$ with $m = 1, \ldots, M$.

The F_2 layer consists of two kinds of nodes: committed and uncommitted. Committed nodes store information in their templates about previously experienced evidence gathered from the module's input domain. In particular, templates of committed nodes store the description of *FA categories*. These categories are the FA module's exemplars that summarize subsets of input patterns presented to the module during training. There is a one-to-one correspondence between FA categories and committed nodes in the F_2 layer of each FA module. On the other hand, uncommitted nodes in the F_2 layer of FA modules do not correspond to real categories and represent the 'blank' memory of the system. Moreover, uncommitted nodes feature a template of $\mathbf{w}_u = w_u \mathbf{1} \in U^{2M}$, where $w_u \geq 1$ is one of the module's parameters. A popular choice of this parameter is $w_u = 1$. Note that before the commencement of a module's training phase all nodes in the F_2 layer are uncommitted. As new, 'unseen' information is revealed to the networks about their environment, learning progresses in FA/FAM by gradually updating already existing categories and by committing uncommitted nodes, so that new categories are founded.

We continue by defining the *size* $s(\mathbf{w}_j)$ of a category j with template $\mathbf{w}_j$ as

$$s(\mathbf{w}_j) = M - |\mathbf{w}_j| = |\mathbf{v}_j - \mathbf{u}_j| = \sum_{m=1}^{M} (v_{jm} - u_{jm}) \tag{1}$$

where $|\mathbf{w}_j|$ is the template size of category j. Note that for every input pattern $\mathbf{x} \in U^M$ and template $\mathbf{w}_j$ it holds

$$|\mathbf{w}_j| = M - |\mathbf{v}_j - \mathbf{u}_j| = M - \sum_{m=1}^{M} (v_{jm} - u_{jm}),$$
$$|\mathbf{x}^c \wedge \mathbf{w}_j| = M - \sum_{m=1}^{M} [\max\{x_m, v_{mj}\} - \min\{x_m, u_{mj}\}] \tag{2}$$

Based on Eqs. (1) and (2), we define as the *distance* of a pattern $\mathbf{x} \in U^M$ from a category with template $\mathbf{w}_j$ the quantity

$$\mathrm{dis}(\mathbf{x}, \mathbf{w}_j) = |\mathbf{w}_j| - |\mathbf{x}^c \wedge \mathbf{w}_j| = \sum_{m=1}^{M} [(\max\{x_m, v_{jm}\} - v_{jm}) + (u_{jm} - \min\{x_m, u_{jm}\})] \tag{3}$$

Notice that for any $\mathbf{x} \in U^M$ and any category with template $\mathbf{w}_j$ it holds $0 \leq \mathrm{dis}(\mathbf{x}, \mathbf{w}_j) \leq M$. Utilizing Eqs. (1) and (3) we can reformulate Eq. (2) as

$$|\mathbf{w}_j| = M - s(\mathbf{w}_j), \qquad |\mathbf{x}^c \wedge \mathbf{w}_j| = M - s(\mathbf{w}_j) - \mathrm{dis}(\mathbf{x}, \mathbf{w}_j) \tag{4}$$

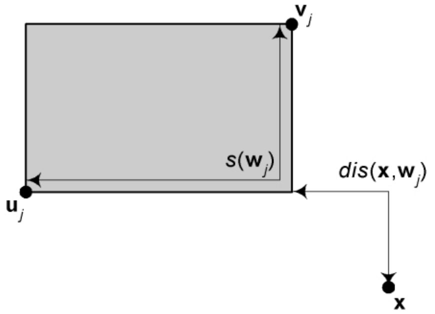

Fig. 2. Geometric representation of FA category j assuming a two-dimensional input domain. The shaded area constitutes the representation region of j and the size of j equals the city-block distance between the category's template elements $\mathbf{u}_j$ and $\mathbf{v}_j$. Also shown in this figure is the distance $dis(\mathbf{w}_j,\mathbf{x})$ of a pattern $\mathbf{x}$ from j, which is defined as the minimum city-block distance of $\mathbf{x}$ from the representation region of j. If $\mathbf{x}$ were inside the representation region, the distance from j would have been zero.

It has been shown by Carpenter et al. (1991, 1992) that FA categories can be geometrically represented as hyper-rectangles embedded in the FA module's input space U^M. An example, when $M = 2$, is shown in Fig. 2.

The union of the shaded area in Fig. 2 and the boundaries of the rectangle defined by $\mathbf{u}_j$ and $\mathbf{v}_j$, is called *representation region* of category j. Also depicted in the same figure, $\mathrm{dis}(\mathbf{x},\mathbf{w}_j)$ reflects the minimum L_1 distance (also known as *city-block* or *Manhattan* distance) between pattern $\mathbf{x}$ and category's j representation region. Note that, if $\mathbf{x}$ were inside or on the borders of the rectangle, its distance from category j would have been 0, therefore a more formal definition for the representation region can be stated as follows.

Definition 1. We define as *category representation region* $R_j = R(\mathbf{w}_j)$ of a category j with template $\mathbf{w}_j$ the following subset of U^M:

$$R(\mathbf{w}_j) = \{\mathbf{x} \in U^M | \mathbf{x}^c \wedge \mathbf{w}_j = \mathbf{w}_j\} \Leftrightarrow R(\mathbf{w}_j) = \{\mathbf{x} \in U^M | \mathrm{dis}(\mathbf{x}, \mathbf{w}_j) = 0\} \tag{5}$$

To incorporate new evidence in the form of a training pattern, an FA category increases its representation region size and simultaneously reduces the distance between the pattern and the region. The learning law that implements

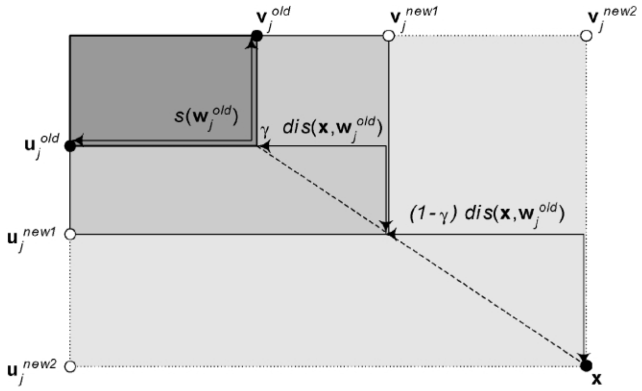

Fig. 3. FA category modifications—slow versus fast learning. Here we illustrate the two types of learning, when a pattern $\mathbf{x}$ updates category j with original template $\mathbf{w}_j^{\rm old} = [\mathbf{u}_j^{\rm old}\ \mathbf{1} - \mathbf{v}_j^{\rm old}]$; its representation region is displayed dark-shaded. Under fast learning assumption ($\gamma = 1$) category j's representation region will expand to include $\mathbf{x}$ and j's template will be updated to $\mathbf{w}_j^{\rm new2} = [\mathbf{u}_j^{\rm new2}\ \mathbf{1} - \mathbf{v}_j^{\rm new2}]$. Using slow learning ($0 < \gamma < 1$) j's will have its template updated to $\mathbf{w}_j^{\rm new1} = [\mathbf{u}_j^{\rm new1}\ \mathbf{1} - \mathbf{v}_j^{\rm new1}]$, which will cause its representation region to expand towards $\mathbf{x}$, so that its new size will become $s(\mathbf{w}_j^{\rm old}) + \gamma\,{\rm dis}(\mathbf{x}, \mathbf{w}_j^{\rm old})$ (according to Eq. (7)) and $\mathbf{x}$'s distance from j will reduce (according to Eq. (8)) to $(1-\gamma){\rm dis}(\mathbf{x}, \mathbf{w}_j^{\rm old})$. The template update in both cases of learning follows the learning rule depicted in Eq. (6).

this idea for a category j being updated due to a pattern $\mathbf{x}$ is

$$\mathbf{w}_j^{\rm new} = \gamma\left(\mathbf{x}^{\rm c} \wedge \mathbf{w}_j^{\rm old}\right) + (1-\gamma)\mathbf{w}_j^{\rm old} \tag{6}$$

where $\gamma \in (0,1]$ is a learning rate parameter. From Eqs. (1)–(3) and (6), we deduce that after a category has been updated due to a pattern $\mathbf{x}$ it holds

$$s\left(\mathbf{w}_j^{\rm new}\right) = s\left(\mathbf{w}_j^{\rm old}\right) + \gamma\,{\rm dis}\left(\mathbf{x}, \mathbf{w}_j^{\rm old}\right) \tag{7}$$

$${\rm dis}\left(\mathbf{x}, \mathbf{w}_j^{\rm new}\right) = (1-\gamma){\rm dis}\left(\mathbf{x}, \mathbf{w}_j^{\rm old}\right) \tag{8}$$

As a special case, when $\gamma = 1$ (*fast learning* assumption), a category j that will be modified upon presentation of pattern $\mathbf{x}$ will increase its size so that its new representation region contains $\mathbf{x}$. In such a case, we say that the updated category j *encodes* pattern $\mathbf{x}$. In all other cases, where $\gamma < 1$ (*slow learning* assumption), we will say that pattern $\mathbf{x}$ *updated* or *modified* category j. Fig. 3 shows these two cases in a two-dimensional setting.

In Fig. 3, under slow learning, pattern $\mathbf{x}$ modifies $\mathbf{w}_j^{\rm old}$ to $\mathbf{w}_j^{\rm new1}$, while under fast learning category j expands enough to encode $\mathbf{x}$ and updates $\mathbf{w}_j^{\rm old}$ to $\mathbf{w}_j^{\rm new2}$. Due to how learning occurs in FA and FAM, categories are never destroyed during the training phase and can only increase in size; destruction (elimination) of a category would be equivalent to partial loss of the FA module's memory. As can be noticed in Fig. 3, the new representation region of a category that has been updated includes the previous one for any value of $\gamma \in (0,1]$. Expressed in terms of sets, if $\mathbf{x} \notin R(\mathbf{w}_j^{\rm old})$ and category j is being updated due to $\mathbf{x}$ as in Eq. (6), then $R(\mathbf{w}_j^{\rm old}) \subset R(\mathbf{w}_j^{\rm new})$. Otherwise, if $\mathbf{x} \in R(\mathbf{w}_j^{\rm old})$, then $R(\mathbf{w}_j^{\rm old}) = R(\mathbf{w}_j^{\rm new})$ for any $\gamma \in (0,1]$. This implies that $R(\mathbf{w}_j^{\rm old}) \subseteq R(\mathbf{w}_j^{\rm new})$ for any $\mathbf{x} \in U^M$ and under any learning assumption.

Two important quantities related to FA categories are the *category match function* $\rho(\mathbf{w}|\mathbf{x})$ (CMF) and the *category choice function* $T(\mathbf{w}|\mathbf{x})$ (CCF—also known as *bottom-up input* or *activation function*) of a category with template $\mathbf{w}$ with respect to an input pattern $\mathbf{x}$, which are defined below:

$$\rho(\mathbf{w}|\mathbf{x}) = \frac{|\mathbf{x}^{\rm c} \wedge \mathbf{w}|}{M} \tag{9}$$

$$T(\mathbf{w}|\mathbf{x}) = \frac{|\mathbf{x}^{\rm c} \wedge \mathbf{w}|}{|\mathbf{w}| + a} \tag{10}$$

In Eq. (10), $a > 0$ is defined as the *choice parameter* of the module. Based on Eq. (4), the CMF and CCF can be alternatively expressed via geometry-based quantities as

$$\rho(\mathbf{w}|\mathbf{x}) = \frac{M - s(\mathbf{w}) - {\rm dis}(\mathbf{x}, \mathbf{w})}{M} \tag{11}$$

$$T(\mathbf{w}|\mathbf{x}) = \frac{M - s(\mathbf{w}) - {\rm dis}(\mathbf{x}, \mathbf{w})}{M - s(\mathbf{w}) + a} \tag{12}$$

The above functions play a central role in the two phases of operation of an FA module (training and performance). The CMF value of a category with respect to an input pattern is the quantity used in comparison to the FA module's *(baseline) vigilance parameter* $\rho \in [0,1]$. On the other hand, a category's CCF value of a category with respect to a pattern $\mathbf{x}$ is used to determine the winning node in the F_2 layer of FA modules during node competition for $\mathbf{x}$. Both committed and uncommitted nodes participate in the competition. The node featuring the highest CCF value is eventually chosen, unless it is committed and gets reset, as we will explain later. In the latter case, the competition will be repeated, until a non-reset node is found featuring the highest CCF value (*repetitive category search process*). During the training phase in FA modules only the non-reset winning node will become updated by training pattern $\mathbf{x}$ (*winner-take-all* scheme). In case of a tie among nodes, the one with smallest index j is finally chosen.

3. The commitment test

The comparison of CMF values to the vigilance parameter ρ, which we mentioned earlier, constitutes the VT. The VT acts as a screening device for categories after the node competition has taken place. A winning committed node j failing the VT with respect to a pattern $\mathbf{x}$ during either the training or performance phase can be interpreted as follows: $\mathbf{x}$ does not fit the characteristics of category j and, therefore, the node is being reset via the reset node in Fig. 1 and disqualified from the node competition for $\mathbf{x}$. If all committed nodes become reset after the competition for a pattern $\mathbf{x}$, an uncommitted node will be chosen, which signifies that none of the existing categories was able to explain the presence of **x**, therefore a new category must be created. Thus, the VT can be regarded as a tunable, novelty detection mechanism that points out atypical patterns with respect to existing categories in the FA module. The VT is expressed as

$$\rho(\mathbf{w}_j|\mathbf{x}) \geq \rho \tag{13}$$

Categories fail the test, when their CMF value is less than ρ. Since it can be shown that $0 \leq \rho(\mathbf{w}|\mathbf{x}) \leq 1$, for a value of $\rho = 0$ all categories will pass the VT for any pattern presented, which is practically equivalent to omitting the VT node/category filtering operation altogether. The higher the value of ρ, the more node filtering occurs via the VT. Due to Eq. (9), uncommitted nodes feature a constant CMF value of

$$\rho_u = \rho(\mathbf{w}_u|\mathbf{x}) = 1 \tag{14}$$

which means that they pass the VT for all values of $\rho \in [0,1]$ and will never get reset. This fact implies that if none of the committed nodes pass the VT for **x**, then an uncommitted node will be chosen. During training phase, the winning uncommitted node j will get committed by having its template updated to $\mathbf{w}_j = \mathbf{x}^c = [\mathbf{x}\ \mathbf{1} - \mathbf{x}]$ (*fast commit*) and will form a new category encoding a single pattern.

However, the VT is not the only novelty detection device in FA modules, as we will demonstrate shortly. Note that, due to Eq. (10), uncommitted nodes feature a constant CCF value of

$$T_u = T(\mathbf{w}_u|\mathbf{x}) = \frac{M}{2Mw_u + a} \tag{15}$$

for all patterns **x** of the input space. Occasionally the winning node will be an uncommitted one, which also means that the existing categories could not satisfactorily explain the just presented pattern (Georgiopoulos et al., 1996). Therefore, the competition of committed nodes against uncommitted can be thought as an implicit novelty detection mechanism, in contrast to the VT, which is an explicit one. In order for a category j to have a chance of being selected by a pattern and be the one that best explains the presence of this pattern, apart from having a CMF value larger or equal to ρ, it must also have a CCF value higher than or equal to the one of uncommitted nodes, that is,

$$T(\mathbf{w}_j|\mathbf{x}) \geq T_u \tag{16}$$

For this reason, we can view the above comparison of CCF values as a test similar to the VT, which determines if a category j has the potential to be chosen. Motivated from our discussion so far, we can formally define the competition between a committed and an uncommitted node as a test similar to the VT.

Definition 2. We define as CT of a category j featuring a template $\mathbf{w}_j$ with respect to an input pattern **x** the comparison of its CCF value, $T(\mathbf{w}_j|\mathbf{x})$, to the CCF value of an uncommitted node, T_u. We say that category j passes the CT, when $T(\mathbf{w}_j|\mathbf{x}) \geq T_u$.

It can be shown that $0 \leq T(\mathbf{w}_j|\mathbf{x}) \leq M/(M + a) < 1$ and $0 < T_u < 1/2$. Additionally, from Eq. (15), when $w_u \to \infty$, then $T_u \to 0$ and the CT is satisfied by any committed node. The smaller the value of w_u, the stricter the node filtering that is being performed via the CT. Comparing to the VT, w_u plays the same role as ρ in the VT, thus w_u can be regarded as the tuning parameter of the CT.

According to what we have presented so far, if a category does not pass the VT and/or the CT with respect to a pattern **x**, it is guaranteed that this particular category will not be chosen upon presentation of **x**. Furthermore, when both $\rho = 0$ and $w_u \to \infty$, we conclude that (i) during training of an FA module an uncommitted node will only be selected by the first training pattern presented, which will find a category (all remaining patterns will select this one and only category) and (ii) during performance no uncommitted node will ever be chosen. This particular setting of ρ and w_u must be used for the performance phase of FAM classifiers, when an 'unknown' class label response is unacceptable and all test patterns have to be classified as belonging to one of the existing classes. As a reminder, when a test pattern selects an uncommitted node during performance phase, this pattern is proclaimed as atypical in comparison to the training patterns seen by the network during training; in the case of an FAM classifier, the pattern cannot be classified.

4. Fuzzy-ART category regions

In Section 3, we have talked about the geometric representation of FA templates and we have described the update of a category due to a training pattern by geometric means. We also have formally defined a category's representation region, which corresponds to all the points in the input domain the category already encodes. A major contribution of our work is the introduction of additional regions as new geometric concepts related to FA categories. By expressing the VT and CT using geometric quantities via Eqs. (11) and (12), we will define FA category regions, which attribute a geometric facet to both tests. Their purpose is to explain the circumstances in a geometrical framework, under which a category has a potential to be chosen and increase our understanding of FA and FAM operations.

4.1. The category match region

Next, we will proceed with a definition that adorns the VT with a new geometrical interpretation. Proofs of selected properties regarding the match region, as well as regarding other regions presented in the sequel, are supplied in Appendix A.

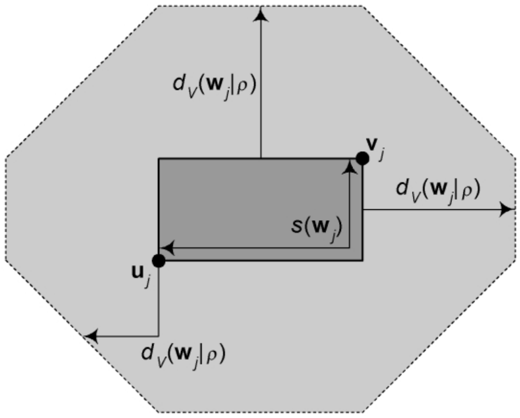

Fig. 4. Match region of a template j in two dimensions. The match region of j consists of all points of the input pattern space, for which j would pass the VT. As seen in the figure, the match region of j (the union of the light- and dark-shaded regions) contains its (dark-shaded) representation region in concordance with Property 3. Here, $d_V(\mathbf{w}_j|\rho)$ stands for the match region radius of j, which depends on the category's size $s(\mathbf{w}_j)$ as well as on ρ (as shown in Eq. (17)) and equals the maximum city-block distance a pattern can have from j, so that j will marginally pass the VT. The larger j's size and/or larger the value of ρ, the smaller the radius. When the radius equals zero, the match region coincides with the representation region and j is prohibited from further expanding due to other patterns, since it will fail the VT for any presented input pattern outside its representation region. Finally, note that, in the general case, j's choice region would have the same shape as the match region depicted in this figure featuring, however, a different radius.

Definition 3. We define as *category match (vigilance) region* $V_j = V(\mathbf{w}_j|\rho)$ of a category j with template $\mathbf{w}_j$ for a particular value ρ of the vigilance parameter the following subset of U^M:

$$V(\mathbf{w}_j|\rho) = \{\mathbf{x} \in U^M | \rho(\mathbf{w}_j|\mathbf{x}) \geq \rho\}$$

$$\Leftrightarrow \left\{ \begin{array}{l} V(\mathbf{w}_j|\rho) = \{\mathbf{x} \in U^M | dis(\mathbf{x}, \mathbf{w}_j) \leq d_V(\mathbf{w}_j|\rho)\} \\ d_V(\mathbf{w}_j|\rho) = M(1-\rho) - s(\mathbf{w}_j) \end{array} \right\} \quad (17)$$

We call the quantity $d_V(\mathbf{w}_j|\rho)$ the *radius of the match (vigilance) region*. It stands for the maximum L_1 distance that a pattern $\mathbf{x}$ can have from the category's representation region, so that category j passes the VT for a vigilance parameter value of ρ. Based on Definition 3 we can replace the algebraic definition of the VT, as shown in Eq. (13), with a geometric one.

Geometric definition of the vigilance test. An FA category j with template $\mathbf{w}_j$ passes the VT with respect to an input pattern $\mathbf{x} \in U^M$ for a particular value ρ of the vigilance parameter, if and only if $\mathbf{x} \in V(\mathbf{w}_j|\rho)$.

From the previous definition, a match region's size is parametrically affected only by ρ. More specifically, the match region radius is a monotonically decreasing function of ρ. The larger the value of the vigilance parameter is, the smaller the radius of the match region for constant category size. The following property reflects in some aspect this fact by examining the two extreme values for ρ.

Property 1. *For $\rho = 0$ the match region of any category j coincides with the entire input domain. For $\rho = 1$ a category's match region includes only the pattern that created the category. Stated in terms of sets:*

$$\text{If } \rho = 0 \Rightarrow V(\mathbf{w}_j|0) = U^M \qquad \forall \mathbf{w}_j,$$

$$\text{If } \rho = 1 \Rightarrow s(\mathbf{w}_j) = 0 \text{ and } V(\mathbf{w}_j|1) = \{\mathbf{u}_j\} = \{\mathbf{v}_j\} \qquad \forall \mathbf{w}_j \quad (18)$$

The first part of the property restates the fact that for $\rho = 0$ any category will pass the VT for any pattern of the input domain. The second part refers to the case where FA/FAM creates categories as many as training patterns during its training phase regardless of the value of a (Carpenter et al., 1992). Under this condition, no category will pass the VT for patterns that are not already encoded.

Assuming a constant value of ρ, we observe that the match region radius decreases with increasing category size. Also, if $d_V(\mathbf{w}_j|\rho) < 0$ the match region is the empty set $\emptyset$, which implies that category j will never pass the VT for any pattern, even for the ones it already encodes in its representation region. However, this can never be the case in an FA network. The above observations hint that the match region enforces a maximum category size, which is controlled by the value of ρ.

Property 2. *The category match region imposes a maximum on category sizes equal to $M(1-\rho)$ for all FA categories.*

We continue with a property that relates the match region to the representation region of a category.

Property 3. *For an FA category j with template $\mathbf{w}_j$ it holds that $R(\mathbf{w}_j) \subseteq V(\mathbf{w}_j|\rho)$ $\forall \rho \in [0,1]$. It holds that $R(\mathbf{w}_j) = V(\mathbf{w}_j|\rho)$, if and only if $s(\mathbf{w}_j) = M(1-\rho)$.*

We know at this point that the match region always contains the representation region, which explains a previous result that a category will pass the VT for all patterns it already encodes. In the case, where $R(\mathbf{w}_j) = V(\mathbf{w}_j|\rho)$, category j will never pass the VT for any pattern outside its representation region. Also, upon presentation of patterns inside its representation region, due to the learning law in Eq. (6), the category will not get modified. Therefore, if $R(\mathbf{w}_j) = V(\mathbf{w}_j|\rho)$, category j cannot be updated due to the presentation of any training pattern, since it has reached its maximum size with respect to the VT, and is called *stagnant*.

Category region illustrations for a general case of $\mathbf{w}_j$ when $M = 2$ and $M = 3$ are given in Figs. 4 and 5, respectively. In Fig. 4 the union of both shaded areas constitutes the match region of the category depicted, while

$d_V(\mathbf{w}_j|\rho)$ denotes the match region radius. The match region's boundary represents all points, for which the category will barely pass the VT. The most general case of a match region embedded in a three-dimensional space has 26 facets and is depicted in Fig. 5. We state here without proof that for higher dimensionalities of the input space (higher values of M) the match region's boundary is a convex polytope with its axes of symmetry parallel to the corresponding ones of the coordinate system. We continue with Property 4, which describes the relationship of a category's match region before and after an update due to a pattern belonging to the category's original match region. Note that, when we refer to *slow learning*, we actually mean *fast-commit/slow-recode learning* (Carpenter et al., 1991, 1992), i.e. only committed nodes are being updated via Eq. (6), while uncommitted ones become committed in the way we described earlier.

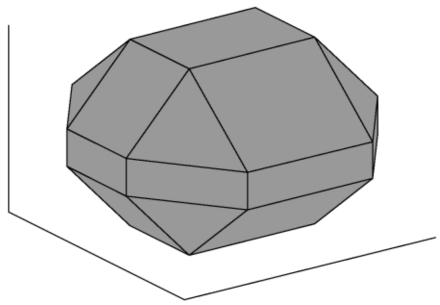

Fig. 5. General form of a category's match region in three dimensions. The geometric shape consists of 26 faces, its axes of symmetry coincide with the ones of the coordinate system and it encloses the category's representation region (as expected by Property 3). In the general case, the category's choice region would feature a similar shape in a three-dimensional pattern space.

Property 4. *During the training phase and for fast learning or slow learning using fast-commit, the match region of any FA category contracts, whenever the category is being updated due to a training pattern located inside its match region, but outside the category's representation region. Stated in terms of sets, for any FA category j with template* $\mathbf{w}_j^{\text{old}}$ *and any pattern* $\mathbf{x} \in V(\mathbf{w}_j^{\text{old}}|\rho)$ *it holds that* $V(\mathbf{w}_j^{\text{new}}|\rho) \subseteq V(\mathbf{w}_j^{\text{new}}|\rho) \subseteq V(\mathbf{w}_j^{\text{old}}|\rho)\ \ \forall \rho \in [0,1]$ *and any* $\gamma \in (0,1]$, *where* $\mathbf{w}_j^{\text{new}} = (1-\gamma)\mathbf{w}_j^{\text{old}} + \gamma(\mathbf{x}^{\text{c}} \wedge \mathbf{w}_j^{\text{old}})$. *Specifically, it holds that*

$$\begin{aligned} &\text{(a) } \mathbf{x} \in R(\mathbf{w}_j^{\text{old}}) \Leftrightarrow V(\mathbf{w}_j^{\text{new}}|\rho) = V(\mathbf{w}_j^{\text{old}}|\rho), \\ &\text{(b) } \mathbf{x} \in V(\mathbf{w}_j^{\text{old}}|\rho) - R(\mathbf{w}_j^{\text{old}}) \Leftrightarrow V(\mathbf{w}_j^{\text{new}}|\rho) \subset V(\mathbf{w}_j^{\text{old}}|\rho) \end{aligned} \tag{19}$$

Since match regions are contracting, whenever their related representation regions expand, an immediate result of Property 4 is the following.

$$C(\mathbf{w}_j|a,w_{\text{u}}) = \{\mathbf{x} \in U^M | T(\mathbf{w}_j|\mathbf{x}) \ge T_{\text{u}}\} \Leftrightarrow \left\{ \begin{array}{l} C(\mathbf{w}_j|a,w_{\text{u}}) = \{\mathbf{x} \in U^M | dis(\mathbf{x},\mathbf{w}_j) \le d_C(\mathbf{w}_j|a,w_{\text{u}})\} \\ d_C(\mathbf{w}_j|a,w_{\text{u}}) = \dfrac{(2w_{\text{u}}-1)M + a}{2Mw_{\text{u}} + a}\left[\dfrac{(2w_{\text{u}}-1)M^2}{(2w_{\text{u}}-1)M + a} - s(\mathbf{w}_j)\right] \end{array} \right\} \tag{20}$$

Property 5. *During the training phase and under fast or slow learning with fast-commit, the match region's hyper-volume of any FA category decreases, whenever the category experiences an update due to a pattern located inside its match region, but outside the category's representation region, i.e. if* $\mathbf{x} \in V(\mathbf{w}_j^{\text{old}}|\rho) - R(\mathbf{w}_j^{\text{old}})$, *then* $\text{Vol}(V(\mathbf{w}_j^{\text{new}}|\rho)) < \text{Vol}(V(\mathbf{w}_j^{\text{old}}|\rho))\ \ \forall \rho \in [0,1]$ *and* $\gamma \in (0,1]$, *where* $\mathbf{w}_j^{\text{new}} = (1-\gamma)\mathbf{w}_j^{\text{old}} + \gamma(\mathbf{x}^{\text{c}} \wedge \mathbf{w}_j^{\text{old}})$.

An example of a two-dimensional match region contracting is shown in Fig. 6, where a representation region expands due to category's j update and its match region decreases in volume (for two dimensions, in surface), while it remains contained in the original match region. Observe that, although $\mathbf{x}_1$ was inside the original match region before the expansion, the category will fail the VT with respect to $\mathbf{x}_1$ after the update, since $\mathbf{x}_1$ does not belong to the new match region. Hence, $\mathbf{x}_1$ will never cause an update to this category, if it is presented again in the future. Finally, notice that the category will never pass the VT for $\mathbf{x}_2$ regardless of future updates, since its match region always contracts.

4.2. The category choice region

So far we have highlighted some aspects of the match region, which relate directly to the notion of the VT. A similar development as in Definition 3 can be performed for the CT by using Eq. (12), which expresses the CCF in geometric quantities.

Definition 4. We define as *category choice (commitment) region* $C(\mathbf{w}_j|a,w_{\text{u}})$ of a category j with template $\mathbf{w}_j$ for particular values a of the choice parameter and w_{u} the subset of U^M depicted in Eq. (20)

In other words, $C(\mathbf{w}_j|a,w_{\text{u}})$ stands for all points of the input space, for which the category j with template $\mathbf{w}_j$ would satisfy the CT, for parameter values a and w_{u}. Points, for which $T(\mathbf{w}_j|\mathbf{x}) = T_{\text{u}}$, lie exactly on the boundary of the choice region. Due to the region's definition, a category would loose the competition against an uncommitted node with respect to any pattern outside its choice region. The quantity $d_C(\mathbf{w}_j|a,w_{\text{u}})$ in Eq. (20) is called the *radius of the choice (commitment) region*. In light of Definition 4, the CT can be geometrically redefined as shown below.

Geometric definition of the commitment test. An FA

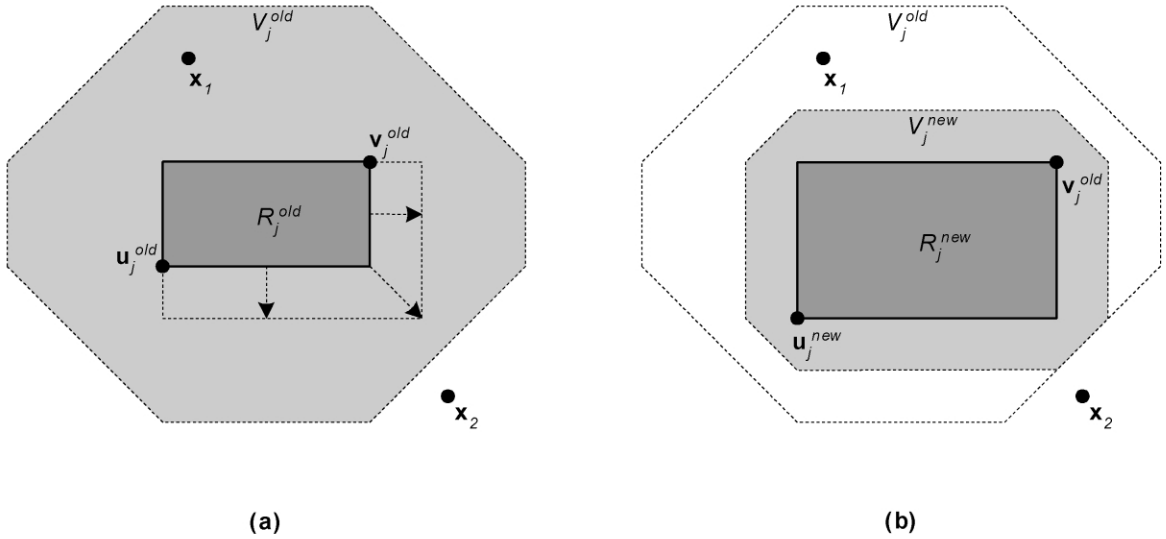

Fig. 6. Contraction of match region in two dimensions. The figure depicts a category j having a template of $\mathbf{w}_j^{\text{old}} = [\mathbf{u}_j^{\text{old}}\ \mathbf{1} - \mathbf{v}_j^{\text{old}}]$ and dark-shaded representation region. The category is being updated using either slow or fast learning due to a pattern, which is not shown here, and its template becomes $\mathbf{w}_j^{\text{new}} = [\mathbf{u}_j^{\text{new}}\ \mathbf{1} - \mathbf{v}_j^{\text{new}}]$. Both initial ($V_j^{\text{old}}$) and post-update ($V_j^{\text{new}}$) match regions are represented as the union of light- and dark-shaded areas in (a) and (b), respectively. (b) demonstrates that V_j^{new} is fully contained within V_j^{old}, which is in agreement with Property 4. In other words, a category's match region contracts after being updated, which also implies that the match region's hyper-volume (for two-dimensional pattern space, surface area) decreases after an update (see Property 5). The contraction effect may cause the updated category j to fail the VT with respect to some specific patterns (like $\mathbf{x}_1$ in the figure), even though it would have passed it with respect to the same patterns prior to the update. Also due to the match region's contraction, j will obviously fail the VT for a pattern $\mathbf{x}_2$ after the update, if it would have failed it with respect to the same pattern prior to the update (relates to Result 1).

category j with template $\mathbf{w}_j$ passes the CT with respect to a pattern $\mathbf{x} \in U^M$ for a particular value a of the choice parameter and a particular value of w_u if and only if $\mathbf{x} \in C(\mathbf{w}_j|a,w_u)$.

Due to their definitions, match and choice regions are very similar to each other (apart from some minor differences), as it will become obvious in the sequel. For example, observations similar to the ones that we have stated for the match region radius can be stated for $d_C(\mathbf{w}_j|a,w_u)$ as well. A choice region's radius is also a monotonically decreasing function of a; the larger the value of the choice parameter, the smaller the radius for constant category size. This verifies a well-known result in the FA and FAM literature that as a increases it is more likely to access an uncommitted node than an already committed node. The opposite holds for the value of w_u. Property 6 is the counterpart of Property 1 for choice regions.

Property 6. *For $w_u \to \infty$ the choice region of any category j coincides with entire input domain. For $a \to \infty$ a category's choice region includes only the pattern that created the category. Stated in terms of sets:*

$$\text{If } w_u \to \infty \Rightarrow C(\mathbf{w}_j|a,\infty) = U^M \qquad \forall \mathbf{w}_j,$$

$$\text{If } a \to \infty \Rightarrow s(\mathbf{w}_j) = 0 \text{ and } C(\mathbf{w}_j|\infty, w_u) = \{\mathbf{u}_j\} = \{\mathbf{v}_j\} \quad \forall \mathbf{w}_j \tag{21}$$

According to the previous property, when $w_u \to \infty$, any category will pass the CT for any input pattern and when $a \to \infty$, FA/FAM creates categories as many as training patterns during its training phase regardless of the vigilance's value, since no category will pass the CT for patterns that are not already encoded. An interesting relation between a category's match and choice region is given in the next statement.

Property 7. *For any FA category j and any $a > 0$, when $\rho = 0$ and $w_u \to \infty$, its match region coincides with its choice region, that is, $C(\mathbf{w}_j|a,\infty) = V(\mathbf{w}_j|0) = U^M$.*

For a two- and three-dimensional category, the choice region of a category would resemble in shape to the match region depicted in Figs. 4 and 5, respectively. This is because both sets are of the same form $\{\mathbf{x} \in U^M | \text{dis}(\mathbf{x}, \mathbf{w}) \le d(\rho, a, w_u)\}$; they only differ in radii. The choice region's counterpart of Property 2 is stated as follows.

Property 8. *During training using fast or slow learning with fast-commit, the category choice region imposes a least upper bound on category size equal to $((2w_u - 1)M^2)/((2w_u - 1)M + a)$ for all FA categories.*

Property 8 tells us that, if the VT had been disabled during the training phase of an FA module, all category sizes would satisfy $s(\mathbf{w}) < M^2/(M + a)$ due to the CT. Notice that this bound is not attained by any category, which contrasts the attainable bound of $M(1 - \rho)$ enforced by the VT in Property 2. In analogy to Property 3, the relationship

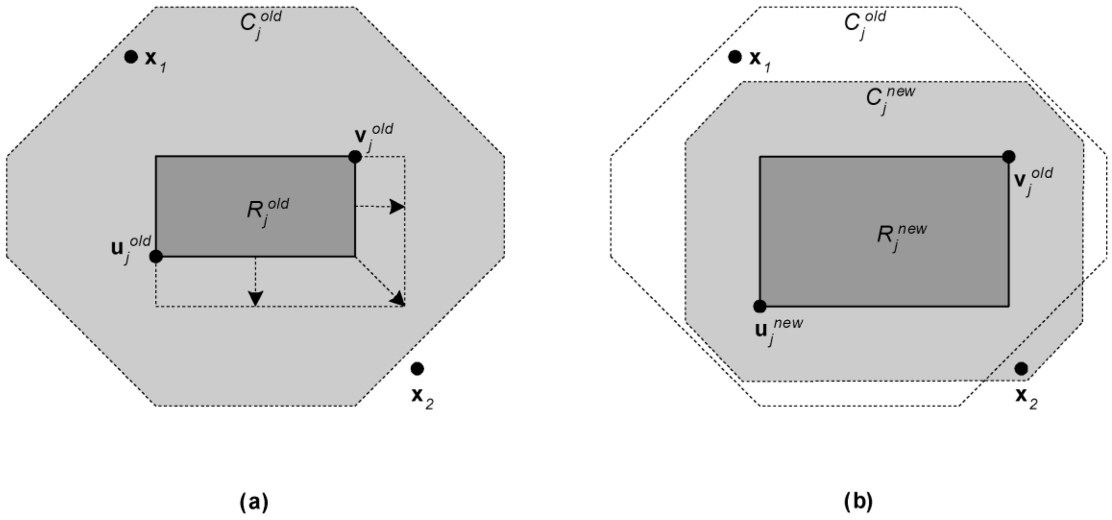

Fig. 7. Hyper-volume decrease of choice region in two dimensions. This figure is analogous to Fig. 6, but instead features choice regions. It displays a category j having a template of $\mathbf{w}_j^{\text{old}} = [\mathbf{u}_j^{\text{old}}\ \mathbf{1} - \mathbf{v}_j^{\text{old}}]$ and dark-shaded representation region. A pattern (not actually shown) causes the update of j either via slow or fast learning so that its template becomes $\mathbf{w}_j^{\text{new}} = [\mathbf{u}_j^{\text{new}}\ \mathbf{1} - \mathbf{v}_j^{\text{new}}]$. Both initial ($C_j^{\text{old}}$) and post-update ($C_j^{\text{new}}$) choice regions of j are represented as the union of light- and dark-shaded areas in (a) and (b), respectively. (b) demonstrates that, in agreement with Property 10, the hyper-volume (for two-dimensional pattern space, surface area) of j's choice region decreases after the category's update. In contrast to j's match region, its choice region does not contract after a template update, since C_j^{new} may include patterns like $\mathbf{x}_2$, which do not belong to C_j^{old}.

of the choice region with the representation region is expressed in the following.

Property 9. *For any FA category j with template* $\mathbf{w}_j$ *it holds* $R(\mathbf{w}_j) \subset C(\mathbf{w}_j|a, w_\text{u})\ \forall a > 0,\ w_\text{u} \geq 1$.

As stated above, a category's representation region will always be included in its choice region, which means that, for a pattern already encoded in a category, an uncommitted node will never be chosen during training. Notice that there is no counterpart to Property 4 for the choice region, because, as it can be illustrated graphically, after a category has been updated, the new choice region does not completely lie within the previous one, that is, choice regions do not contract after template updates take place. However, the following property is analogous to Property 5.

Property 10. *During the training phase and under fast or slow learning with fast-commit, the choice region's hyper-volume of any FA category decreases, whenever the category experiences an update due to a pattern located inside its choice region, but outside the category's representation region, i.e. if* $\mathbf{x} \in C(\mathbf{w}_j^{\text{old}}|a, w_\text{u}) - R(\mathbf{w}_j^{\text{old}})$, *then* $\text{Vol}(C(\mathbf{w}_j^{\text{new}}|a, w_\text{u})) < \text{Vol}(C(\mathbf{w}_j^{\text{old}}|a, w_\text{u}))$ $\forall a > 0,\ w_\text{u} \geq 1$ *and* $\gamma \in (0, 1]$, *where* $\mathbf{w}_j^{\text{new}} = (1 - \gamma)\mathbf{w}_j^{\text{old}} + \gamma(\mathbf{x}^\text{c} \wedge \mathbf{w}_j^{\text{old}})$.

Again, an example of the above statement in two dimensions is given in Fig. 7, which illustrates the fact that, although the choice region decreases in hyper-volume (surface, in two dimensions), it is not completely contained in the category's original choice region. As was the case with the match region in Fig. 6, $\mathbf{x}_1$ belongs to the original choice region, but after the category's update finds itself outside. The exact opposite happens to pattern $\mathbf{x}_2$, which can never happen for match regions.

A summary of the regions' properties is presented in Table 1, which highlights similarities, as well as a few differences between the two regions. A major difference is that upon a category update, although the hyper-volume of both its match and choice regions decreases, only its match region contracts.

So far we have examined both the VT and the CT separately in terms of their associated regions and the results imply that both perform analogous functionality: they regulate for which and for how many points a particular category will pass the VT or the CT. From that aspect the coexistence of the VT and CT seems to bear a redundancy in functional role. However, the absence of one or the other would produce different results during an FA module's training phase. We saw earlier that when $\rho = 0$ any category in an FA module will pass the VT, in essence inhibiting VT's novelty detection role. A similar statement holds for the CT, if $w_\text{u} \to \infty$. Training an FA module using $\rho = 0$ (no VT) in one case and training it using $w_\text{u} \to \infty$ (no CT) in another case would lead to two architectures most likely differing not only in the structure of the categories but also in the number of categories that would have been created. This is mainly because only match regions contract, while choice regions do not (see Figs. 6 and 7) resulting in two different ways of category expansion.

4.3. The category claim region

A category j may be chosen upon presentation of $\mathbf{x}$, if at least it passes both the VT and CT, as we have seen in

Section 3. According to the geometric definitions of the VT and the CT, we have given so far, this is equivalent to the pattern **x** belonging to both match and choice region of *j*. If **x** is located outside at least one of those two regions, then *j* will either fail the VT ($\mathbf{x} \notin V(\mathbf{w}_j|\rho)$), or *j* will lose the competition against the uncommitted node, since it fails the CT ($\mathbf{x} \notin C(\mathbf{w}_j|a)$), or both. Therefore, it is natural to define a new category region, which will include all points common to both the match and choice regions.

Definition 5. We define as *category claim region* $L(\mathbf{w}_j|\rho,a,w_u)$ of an FA category *j* with template $\mathbf{w}_j$ for a particular vigilance value ρ, choice parameter *a* and w_u the following subset of U^M:

$$L(\mathbf{w}_j|\rho,a,w_u) = V(\mathbf{w}_j|\rho) \cap C(\mathbf{w}_j|a)$$

$$\Leftrightarrow \begin{Bmatrix} L(\mathbf{w}_j|\rho,a,w_u) = \{\mathbf{x} \in U^M | \mathrm{dis}(\mathbf{x},\mathbf{w}_j) \le d_L(\mathbf{w}_j|\rho,a,w_u)\} \\ d_L(\mathbf{w}_j|\rho,a,w_u) = \min\{d_V(\mathbf{w}_j|\rho), d_C(\mathbf{w}_j|a,w_u)\} \end{Bmatrix} \quad (22)$$

As expected, the quantity $d_L(\mathbf{w}|\rho,a,w_u)$ is called the *radius of the claim region*, which also decreases, when a category's size increases. By virtue of Properties 1 and 6, an immediate observation regarding the category claim region is that $L(\mathbf{w}_j|0,a,\infty) = U^M$. The next property links the newly defined region with the match and choice regions.

Property 11. *The claim region of an FA category j with template* $\mathbf{w}_j$ *coincides either with the category's match region or its choice region depending on the value of the vigilance parameter* ρ*, the value of the choice parameter a, the value of* w_u *and, under certain circumstances, on the category's size* $s(\mathbf{w}_j)$*. In more detail, if we define as* $s_{\mathrm{thres}} = (2Mw_u + a)(1-\rho) - (2w_u - 1)M$, *then for* $a > 0$ *and* $w_u \ge 1$ *we discriminate three major cases:*

1. *If* $0 \le \rho \le a/[(2w_u - 1)M + a]$, *then* $L(\mathbf{w}_j|\rho,a,w_u) = C(\mathbf{w}_j|a,w_u)$.
2. *If* $a/[(2w_u - 1)M + a] < \rho < (M + a)/(2Mw_u + a)$
   (2a) *If* $s(\mathbf{w}_j) < s_{\mathrm{thres}}$, *then* $L(\mathbf{w}_j|\rho,a,w_u) = C(\mathbf{w}_j|a,w_u)$.
   (2b) *If* $s_{\mathrm{thres}} < s(\mathbf{w}_j)$, *then* $L(\mathbf{w}_j|\rho,a,w_u) = V(\mathbf{w}_j|\rho)$.
   (2c) *If* $s(\mathbf{w}_j) = s_{\mathrm{thres}}$, *then* $L(\mathbf{w}_j|\rho,a,w_u) = C(\mathbf{w}_j|a,w_u) = V(\mathbf{w}_j|\rho)$.
3. *If* $(M + a)/(2Mw_u + a) \le \rho \le 1$, *then* $L(\mathbf{w}_j|\rho,a,w_u) = V(\mathbf{w}_j|\rho)$.

Property 11 proclaims that the claim region, depending on which region it coincides with (match or choice region), exhibits analogous properties, as depicted in Table 1. It is also fundamental to some aspects related to the operation of FA/FAM networks, as it is shown in Section 5.

5. Property-based results for FA/FAM

The category regions along with their properties that we have presented so far are sufficient to describe under what conditions a category will be eligible to be chosen upon presentation of a particular pattern during the training and performance phase of FA/FAM. All the results of this section apply for FA modules with parameters $\rho \in [0,1]$, $a > 0$, $w_u \ge 1$ and $\gamma \in (0,1]$, unless otherwise specified. An immediate result stemming from Property 4 is provided below.

Result 1. During FA/FAM off-line training or performance phase, if in a particular list presentation a category *j* does not pass the VT for a pattern **x** and a specific value of the vigilance parameter ρ, then *j* will never pass the VT in future list presentations for the same pattern **x** and value of ρ.

The following statement comes as an immediate result of Definition 5 and Property 11.

Result 2. Upon presentation of pattern **x** during training or during the performance phase of an FA/FAM network, in order to determine if a particular FA category *j* with template $\mathbf{w}_j$ may have the potential to be chosen, it suffices to perform only one of the two tests (VT, CT), since satisfaction of one will imply concurrent satisfaction of the other one. If we define as $s_{\mathrm{thres}} = (2Mw_u + a)(1-\rho) - (2w_u - 1)M$, then the sufficient test to be performed depends on the values of the network parameters as follows:

1. If $0 \le \rho \le a/[(2w_u - 1)M + a]$, then it suffices to perform only the CT.
2. If $a/[(2w_u - 1)M + a] < \rho < (M + a)/(2Mw_u + a)$ and
   (2a) if $s(\mathbf{w}_j) < s_{\mathrm{thres}}$, then it suffices to perform only the CT.
   (2b) If $s_{\mathrm{thres}} < s(\mathbf{w}_j)$, then it suffices to perform only the VT.
   (2c) If $s(\mathbf{w}_j) = s_{\mathrm{thres}}$, then perform either the CT or the VT.
3. If $(M + a)/(2Mw_u + a) \le \rho \le 1$, then it suffices to perform only the VT.

The above result states that, when a category satisfies the VT, it will also automatically satisfy the CT (with respect to the same pattern) and vice versa depending on the circumstances outlined. From Result 2 we conclude that, when $w_u \to \infty$, only the VT is necessary for all categories and all (ρ,a). Some further implications of the previous result are presented in Results 3 and 4.

Result 3. Let us define as w_u-*insensitive parameter region* the set $\{(\rho,a) \in [0,1] \times (0,\infty) | a \in (0,\infty), (M + a)/(2Mw_u + a) \le \rho \le 1\}$. If some FA module, which is part of an FA/FAM network, operates in the w_u-insensitive

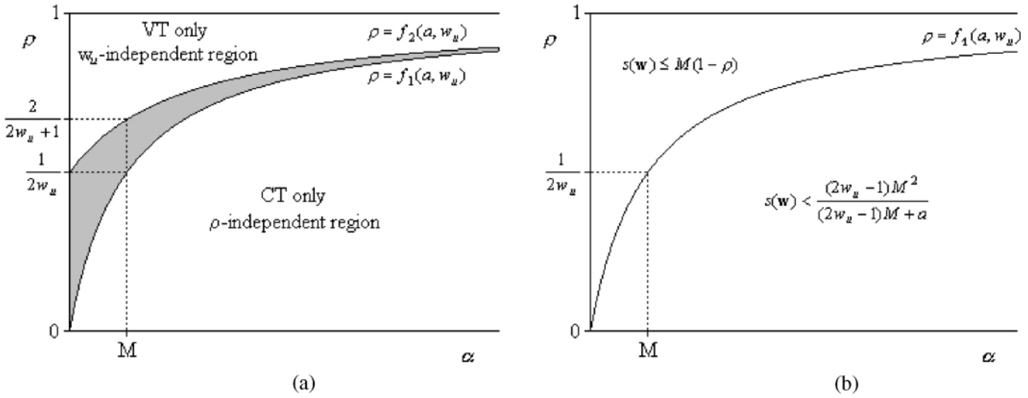

Fig. 8. Regions of interest in the (ρ,a) parameter plane. (a) shows the ρ- and w_u-insensitive (ρ,a) parameter regions (mentioned in Results 3 and 4) separated by a dark-shaded region. If a Fuzzy-ART module is being trained or performs on a testing set with choices of ρ and a values belonging to the ρ-insensitive region, the module's operation does not depend on the particular value of ρ. A similar statement holds for the w_u-insensitive parameter region regarding the lack of dependence on the particular value of w_u. The ρ-independence result is an immediate implication of the fact that for parameter values (ρ,a) belonging to the ρ-insensitive region, any category will pass the VT for any pattern, if it passes the CT for the same pattern. Therefore, in this case only the VT is necessary. The opposite holds, if the parameter values (ρ,a) belong to the w_u-insensitive region (Result 4). On the other hand (b) shows areas of the (ρ,a) parameter space and their corresponding restrictions they impose on the size of any category during training (Result 5).

region of the (ρ,a) parameter space, then the outcome of the training (under any learning assumption) and performance phases does not depend on the particular value of w_u.

When using values for the vigilance and choice parameters from within the w_u-insensitive region, an uncommitted node will be accessed and will form a new category only because all other existing committed nodes fail the VT for a presented pattern. Another immediate result of Result 2 follows in Result 4. Observe that the vigilance ρ is used in FA/FAM only for the performance of the VT. In case (1) of Result 2 we notice that, if a category passes the CT with respect to a certain pattern, then the corresponding VT will automatically be satisfied as well. This fact leads us to the following result.

Result 4. Let us define as *ρ-insensitive parameter region* the set $\{(\rho, a) \in [0,1] \times (0,\infty) | a \in (0,\infty), 0 \le \rho \le a/[(2w_u - 1)M + a]\}$. If some FA module, which is part of an FA/FAM network, operates in the ρ-insensitive region of the (ρ,a) parameter space, then for any $w_u \ge 1$ the outcome of the training (under any learning assumption) and performance phases does not depend on the particular value of ρ.

Result 4 tells us, for example, that, if an FA/FAM network operates in the ρ-insensitive region, the number and the structure of the categories it is going to create during training, as well as the number of list presentations required for fast learning convergence, does not depend on the specific value of ρ. Another example would be that, under the same parameter settings, the classification results of an FAM classifier are independent of the particular value of ρ used in its ART$_a$ module (Carpenter et al., 1992) during performance phase. An important observation is that the ρ-insensitive region vanishes, when $w_u \to \infty$. Finally, one more result directly derived from Property 11 is the following result pertaining to the maximum size of categories that can be constructed during training.

Result 5. For an FA/FAM network, where uncommitted nodes participate in the competition for pattern selection, the size of FA categories attained via training is limited by the following rules:

1. If $0 \le \rho \le a/[(2w_u - 1)M + a]$, then the size of any category has a least upper bound of $((2w_u - 1)M^2)/((2w_u - 1)M + a)$, that is, $s(\mathbf{w}) < ((2w_u - 1)M^2)/((2w_u - 1)M + a)$.
2. If $a/[(2w_u - 1)M + a] < \rho \le 1$, then $s(\mathbf{w}) \le M(1 - \rho)$.

Both statements can be combined in a single inequality

$$s(\mathbf{w}) \le M\left[1 - \max\left\{\rho, \frac{a}{(2w_u - 1)M + a}\right\}\right] \le M(1 - \rho) \tag{23}$$

The previous statement blends two independent results reported earlier in the literature. Case (1) is mentioned in Georgiopoulos et al. (1996), where it is not combined with case (2). Also, case (1) is a refinement of what has been reported in Carpenter et al. (1992), where it is implied that $s(\mathbf{w}) \le M(1 - \rho)$ for all values of $\rho \in [0,1]$ and $a \in (0,\infty)$. Fig. 8 displays different regions of the (ρ,a) parameter plane, which relate to the results of this section. The curves graphed are $\rho = f_1(a, w_u) = a/[(2w_u - 1)M + a]$ and $\rho = f_2(a, w_u) = (M + a)/(2Mw_u + a)$. In Fig. 8(a), the area above the curve $\rho = f_2(a, w_u)$ represents the w_u-insensitive region, for which the satisfaction of the VT implies the satisfaction of the CT for all categories with respect to any input pattern. The contrary holds for choices of (ρ,a) below the curve $\rho = f_1(a, w_u)$, namely inside the ρ-insensitive region. The shaded area signifies the choices of parameter values, where the necessity of the tests depends on the size of each category. For example, as long as $\rho > 2/(2w_u + 1)$ and $a = M$, then the particular value w_u does not affect the training or performance phase of FA/FAM. Also, if $a = M$, FA/FAM's training and performance phases are independent of ρ, as long as $\rho < 1/(2w_u)$. In Fig. 8(b), the curve $\rho = f_1(a, w_u)$ divides the plane into two different areas: for pairs (ρ,a) below $\rho = f_2(a, w_u)$ (ρ-insensitive region) only a least upper bound exists for the size of categories and for pairs above $\rho = f_2(a, w_u)$ the categories can reach a maximum size of $M(1 - \rho)$.

6. Summary and conclusions

In this paper, we have shown that the participation of uncommitted nodes in the competition for patterns during either FA/FAM training or performance phase, implements a category-filtering mechanism similar to the one exercised by the orienting subsystem of an FA module. This is in line with a well-known fact in the ART literature that the presence of uncommitted nodes in the competition at the F_2 layer enforces, by itself, clustering of the input patterns, even when the vigilance parameter in the network is set to zero. We called this filtering mechanism the commitment test (CT) since it boils down to a comparison of choice function values between a committed and an uncommitted node in the F_2-layer of an FA module. Based on this definition, we also have shown that both the CT and the vigilance test (VT) perform similar, but not identical, functions. More specifically, the two tests conjointly determine the eligibility of committed nodes to compete for presented patterns during the FA/FAM training or performance phase.

We have also given an interesting, visual, geometrical interpretation to the CT and VT through the concept of category regions. The VT defined the *match region*, while the CT defined the *choice region*. It was observed that not only these regions (match and choice) have the same geometrical shape (polytope structure), but they also share a lot of common and interesting properties that were demonstrated in this paper. A complete list of these properties was exhibited in Table 1. One of these common properties was the shrinking of the volume that each one of these polytope structures occupies, as training progresses, which alludes to the stability of learning in FA and FAM. These regions also exhibit different characteristics. For instance, a category update defines a new match region that is completely included in the old match region (the one corresponding to the category before its update). On the other hand, a category update defines a new choice region that is not completely included in the old choice region (the one corresponding to the category prior to its update); only the volume of the choice region decreases as we have emphasized above.

The intersection of the match region and the choice region led us to the definition of a new region, named *claim region*, whose geometrical structure is the same as the geometrical structure of the match and choice regions (polytope shape). The claim region has an interesting interpretation. The claim region of a category in ART contains all the points in the pattern space that can be encoded by this category; no points outside a category's claim region can be encoded by this category. Hence, it can be thought of as the region of attraction of an ART category. Defining geometrically, or through an equation, the region of attraction of cluster points (or prototypes) of any useful pattern classifier (not necessarily an ART neural network) is a worthwhile endeavor in the pattern recognition literature. Part of our work in this paper has accomplished this task for FA and FAM.

Based on the existence of these regions and after examining some of their major properties, we were led to a few results concerning FA/FAM, which has practical, as well as theoretical implications. For example, we have illustrated that ρ, a and w_u conjointly determine whether satisfaction of the VT by a category automatically implies satisfaction of the CT, and vice versa. This last observation unveiled the existence of ρ- and w_u-insensitive regions in the (ρ,a) parameter space, which are of practical interest. We have established that FA/FAM's behavior does not depend on the specific value of the vigilance ρ, when the network parameter selection is ρ-insensitive. Thus, in experimenting with different parameter values of an FA/FAM network one need only consider distinct values of the choice parameter a, and a constant value of ρ, when (ρ,a) belongs to the ρ-insensitive region. Similar observations can be stated for a w_u-insensitive selection of parameters, where w_u is the initial weight value for the uncommitted nodes. Furthermore, the introduction of category regions and their associated geometrical structure allowed us to demonstrate in a simple, and intuitive manner existing results in FA/FAM that were previously proven with complex, unintuitive, algebraic manipulations (e.g. see Result 5).

Apart from the fact that category regions give us an additional view into the inner processes of category selection in FA/FAM networks, they can also be used in deriving efficient software and hardware implementations for these architectures. For instance, efficient implementations may take advantage of the regions' properties (especially the shrinking of the claim region's hypervolume) in order to expedite the search for a category appropriate to encode a presented input pattern. Moreover, the same geometrical concepts can be utilized in the framework of virtually any other ART-based neural network architecture as an aid to understand these architectures and to derive theoretical results describing their behavior. Examples of such architectures include dART (Carpenter, 1997) and dARTMAP (Carpenter et al., 1998), Boosted-ARTMAP (Verzi et al., 1998), Micro-ARTMAP (Gomez

Sanchez et al., 2000), Gaussian ARTMAP (Williamson, 1996), Topographic Attentive Mapping network (Williamson, 2001), Ellipsoid-ART/ARTMAP (Anagnostopoulos & Georgiopoulos, 2001) and others.

Acknowledgments

The authors would like to thank the two anonymous referees, which reviewed this paper, for their thoughtful comments and suggestions. Their contribution helped in making the manuscript clearer and more complete.

Appendix A

Proof of Property 1. For $\rho = 0$ the match region radius equals $M - s(\mathbf{w})$ and $0 \le \mathrm{dis}(\mathbf{x},\mathbf{w}) \le M$ for all $\mathbf{x}$ and $\mathbf{w}$. Therefore, all points in U^M are within the radius of any category. Assume now an FA module prior to any training (all F_2 layer nodes are uncommitted) with parameters $\rho = 1$ and any values for $a > 0$. The first training pattern $\mathbf{x}_1$ will select an uncommitted node and form a category 1 with template $\mathbf{w}_1 = \mathbf{x}_1^c$ that has a match region radius equal to $d_V(\mathbf{w}_j|1) = -s(\mathbf{w}_1) = 0$ (from Eq. (17)). Consequently, when the second training pattern $\mathbf{x}_2 \neq \mathbf{x}_1$ is presented, category 1 will fail the VT, since $\mathbf{x}_2$ is outside the category's match region and $\mathbf{x}_2$ will create a new category. The scenario just described will repeat itself for all P training patterns. Within one list presentation the FA module will feature P categories of zero size with their associated match region containing only the pattern that initiated the creation of the category. □

Proof of Property 2. From Definition 3, $V(\mathbf{w}_j|\rho) \neq \varnothing$ when $d_V(\mathbf{w}_j|\rho) \ge 0$, which immediately implies that $s(\mathbf{w}_j) \le M(1-\rho)$. The maximum size of $M(1-\rho)$ is indeed attainable, as we will demonstrate. Assume a category j with $\mathbf{w}_j = \mathbf{x}_1^c$ and a pattern $\mathbf{x}_2$ such that $\mathrm{dis}(\mathbf{x}_2, \mathbf{w}_j) = \|\mathbf{x}_2 - \mathbf{x}_1\|_1 = M(1-\rho)$. Assuming that during the training phase the category wins the competition to encode $\mathbf{x}_2$, according to Eq. (6) the updated (via fast learning) category will feature the maximum size of $M(1-\rho)$. □

Proof of Property 3. The facts that $R(\mathbf{w}_j) \subseteq V(\mathbf{w}_j|\rho)\ \forall \rho \in [0,1]$ and that $R(\mathbf{w}_j) = V(\mathbf{w}_j|\rho)$, if and only if $s(\mathbf{w}) = M(1-\rho)$ are immediately concluded from Definitions 1 and 3. □

Proof of Property 4. We need to prove that $V(\mathbf{w}_j^{\mathrm{new}}|\rho) \subseteq V(\mathbf{w}_j^{\mathrm{old}}|\rho)\ \forall \rho \in [0,1]$ and any $\gamma \in (0,1]$, where $\mathbf{w}_j^{\mathrm{new}} = (1-\gamma)\mathbf{w}_j^{\mathrm{old}} + \gamma(\mathbf{x}^c \wedge \mathbf{w}_j^{\mathrm{old}})$ and $\mathbf{x} \in V(\mathbf{w}_j^{\mathrm{old}}|\rho)$. Pattern $\mathbf{x}$ needs to lie inside the old match region of j, otherwise j would not pass the VT with respect to $\mathbf{x}$ and no update would be possible. In order to prove the property it suffices to show that for any $\mathbf{y} \in V(\mathbf{w}_j^{\mathrm{new}}|\rho)$ it holds $\mathbf{y} \in V(\mathbf{w}_j^{\mathrm{old}}|\rho)$. Indeed, let $\mathbf{y} \in V(\mathbf{w}_j^{\mathrm{new}}|\rho)$, then

$$\mathbf{y} \in V(\mathbf{w}_j^{\mathrm{new}}|\rho) \overset{\text{Def 3}}{\Leftrightarrow} \mathrm{dis}(\mathbf{y}, \mathbf{w}_j^{\mathrm{new}}) \le d_V(\mathbf{w}_j^{\mathrm{new}}|\rho) \tag{A1}$$

Also, from the definition of the match region radius in Definition 3 we get

$$\left.\begin{aligned} d_V(\mathbf{w}_j^{\mathrm{old}}|\rho) &= M(1-\rho) - s(\mathbf{w}_j^{\mathrm{old}}) \\ d_V(\mathbf{w}_j^{\mathrm{new}}|\rho) &= M(1-\rho) - s(\mathbf{w}_j^{\mathrm{new}}) \end{aligned}\right\} \Rightarrow d_V(\mathbf{w}_j^{\mathrm{new}}|\rho)$$
$$= d_V(\mathbf{w}_j^{\mathrm{old}}|\rho) + s(\mathbf{w}_j^{\mathrm{old}}) - s(\mathbf{w}_j^{\mathrm{new}}) \tag{A2}$$

Additionally, it is easy to show that for any $\gamma \in (0,1]$ and $\mathbf{x},\mathbf{y} \in U^M$ it holds

$$|\mathbf{y} \wedge \mathbf{w}_j^{\mathrm{old}}| \ge |\mathbf{y} \wedge \mathbf{w}_j^{\mathrm{new}}| \overset{\text{Eq 4}}{\Leftrightarrow} \mathrm{dis}(\mathbf{y}, \mathbf{w}_j^{\mathrm{old}}) + s(\mathbf{w}_j^{\mathrm{old}}) - s(\mathbf{w}_j^{\mathrm{new}})$$
$$\le \mathrm{dis}(\mathbf{y}, \mathbf{w}_j^{\mathrm{new}}) \tag{A3}$$

Combining Eqs. (A1)–(A3) we arrive at

$$\mathrm{dis}(\mathbf{y}, \mathbf{w}_j^{\mathrm{old}}) \le d_V(\mathbf{w}_j^{\mathrm{old}}|\rho) \overset{\text{Def 3}}{\Leftrightarrow} \mathbf{y} \in V(\mathbf{w}_j^{\mathrm{old}}|\rho) \tag{A4}$$

Hence, we proved that $V(\mathbf{w}_j^{\mathrm{new}}|\rho) \subseteq V(\mathbf{w}_j^{\mathrm{old}}|\rho)\ \forall \rho \in [0,1]$ and any $\gamma \in (0,1]$.

(a) The proof goes as follows:

$$\mathbf{x} \in R(\mathbf{w}_j^{\mathrm{old}}) \overset{\text{Def 1}}{\Leftrightarrow} \mathbf{x}^c \wedge \mathbf{w}_j^{\mathrm{old}} = \mathbf{w}_j^{\mathrm{old}} \overset{\text{Eq 6}}{\Leftrightarrow} \mathbf{w}_j^{\mathrm{new}}$$
$$= \mathbf{w}_j^{\mathrm{old}} \overset{\text{Def 3}}{\Leftrightarrow} V(\mathbf{w}_j^{\mathrm{old}}|\rho) = V(\mathbf{w}_j^{\mathrm{new}}|\rho) \tag{A5}$$

(b) This statement is an immediate product of the previous two results. □

Proof of Property 6. Similar to Proof of Property 1. □

Proof of Property 7. For $\rho = 0$ we get $d_V(\mathbf{w}_j|0) = M - s(\mathbf{w}_j)$ and for $w_u \to \infty$ we have $d_C(\mathbf{w}_j|a,\infty) = M - s(\mathbf{w}_j)$ as well. From Definitions 3 and 4, we conclude that $V(\mathbf{w}_j|0) = C(\mathbf{w}_j|a,\infty)$. □

Proof of Property 8. From Definition 4 we establish an upper bound for a category's size equal to $((2w_u - 1)M^2)/((2w_u - 1)M + a)$, since for sizes exceeding this upper bound a category's choice region radius becomes negative and its choice region equals the empty set. We are now going to show that under any $\gamma \in (0,1]$, $a > 0$, $w_u \ge 1$, and for a finite number of training patterns, no category can reach a size of $((2w_u - 1)M^2)/((2w_u - 1)M + a)$. We

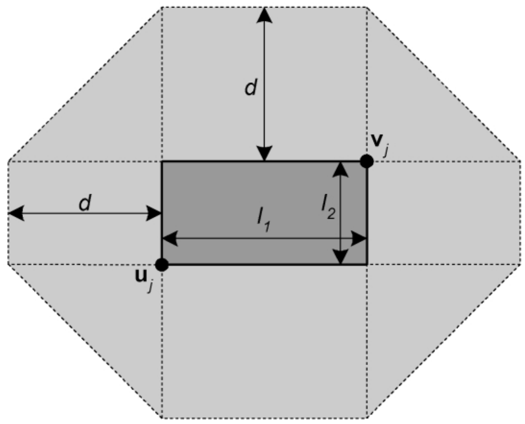

Fig. 9. Calculation of a surface area of a two-dimensional choice region.

first define the quantities

$$f_1 = \frac{(2w_u - 1)M^2}{2Mw_u + a}, \qquad f_2 = \frac{(2w_u - 1)M + a}{2Mw_u + a} \tag{A6}$$

Let there be a category j with initial size $s(\mathbf{w}_j) = s_0 < (((2w_u - 1)M^2)/((2w_u - 1)M + a))$. Furthermore, assume that category j is being updated and expands due to the sequence of patterns $\mathbf{x}_{(k)}$ with $k \geq 1$, for which j always satisfies the VT. Let $\mathbf{w}_{j(k)}$ be the template and s_k the size of j after the kth update. Additionally, assume that the patterns $\mathbf{x}_{(k)}$ are such that $\mathbf{x}_{(k)} \in C(\mathbf{w}_{j(k)}|a, w_u)$, hence from Definition 4 and Eq. (A6)

$$\mathrm{dis}(\mathbf{x}_{(k)}, \mathbf{w}_{j(k)}) \leq d_C(\mathbf{w}_{j(k)}|a, w_u) = f_1 - f_2 s_k \tag{A7}$$

Assume that category j is always being updated after each pattern $\mathbf{x}_{(k)}$ is presented. Then from Eq. (7) we have

$$s_{k+1} = s_k + \gamma\,\mathrm{dis}(\mathbf{x}_{(k)}, \mathbf{w}_{j(k)}) \tag{A8}$$

Under the described scenario, the generated sequence of sizes s_k is strictly increasing. The increase in size is maximum, when the distance of each pattern $\mathbf{x}_{(k)}$ from the representation region is the maximum possible, so that category j still passes the CT. This maximum distance equals the choice region radius of j. Any other increment will only slow down the convergence of s_k. Thus, from Eq. (A8) we have

$$\left.\begin{array}{l} s_{k+1} = s_k + \gamma\,\mathrm{dis}(\mathbf{x}_{(k)}, \mathbf{w}_{j(k)}) \\ \mathrm{dis}(\mathbf{x}_{(k)}, \mathbf{w}_{j(k)}) = d_C(\mathbf{w}_{j(k)}|a) = f_1 - f_2 s_k \end{array}\right\} \Rightarrow s_{k+1}$$

$$= \gamma f_1 - (1 - \gamma f_2) s_k \tag{A9}$$

If $\gamma \neq 0$, then from Eq. (A9) we conclude

$$\lim_{k\to\infty} s_k = \frac{f_1}{f_2} = \frac{(2w_u - 1)M^2}{(2w_u - 1)M + a} \tag{A10}$$

Eq. (A10) shows us that only with an infinite number of appropriate training patterns a category could reach the CT-implied upper bound. Thus, for a finite number of training patterns the choice region $((2w_u - 1)M^2)/((2w_u - 1)M + a)$ serves as a least upper bound for category sizes. □

Proof of Property 9. It follows directly from Definitions 1 and 4 that $R(\mathbf{w}_j) \subseteq C(\mathbf{w}_j|a, w_u)$ for all $a > 0$. Similar to Property 3, we would have that $C(\mathbf{w}_j|a, w_u) = R(\mathbf{w}_j)$, if and only if $s(\mathbf{w}_j) = ((2w_u - 1)M^2)/((2w_u - 1)M + a)$. However, we showed in Property 8 that the last equality is not possible. Therefore $R(\mathbf{w}_j) \subset C(\mathbf{w}_j|a, w_u)$ for all $a > 0$. □

Proof of Property 10. Consider a category with template $\mathbf{w} = [\mathbf{u}\ \ \mathbf{1} - \mathbf{v}]$ and $\mathbf{v} - \mathbf{u} = \mathbf{l}$, where $\mathbf{l}$ is the vector with components l_m $m = 1, \ldots, M$ equal to the length of each side of the category's associated hyper-rectangle. It can be shown that the hyper-volume of the category's choice region is given by

$$\mathrm{Vol}(C(\mathbf{w}|a, w_u)) = \prod_{m=1}^{M} (2d + l_m) - \frac{(2d)^M}{2} \tag{A11}$$

where in the case of the choice region we set $d = d_C(\mathbf{w}|a, w_u) = f_1 - f_2 s(\mathbf{w})$ and f_1, f_2 are again defined as in Eq. (A6). An example in two dimensions can be seen in Fig. 9, where the surface area of the region is given by

$$\mathrm{Vol}(C(\mathbf{w}|a, w_u)) = (2d + l_1)(2d + l_2) - 2d^2 \tag{A12}$$

In order to prove the property it suffices to show that the partial derivative of the hyper-volume with respect to any length l_p $p = 1, \ldots, M$ is negative for any $a \in (0,\infty)$. Taking the derivative of Eq. (A11) we get

$$\frac{\partial \mathrm{Vol}(C(\mathbf{w}|a, w_u))}{\partial l_p} = \sum_{k=1}^{M} (\delta_{k,p} - 2f_2) \prod_{m=1, m\neq k}^{M} (2d + l_m) + M f_2 (2d)^{M-1} \tag{A13}$$

In Eq. (A13) $\delta_{k,p}$ equals to 1 if $k = p$ and 0 otherwise. Notice that, if $l_1 = l_2 = \cdots = l_M = 0$, then

$$\sum_{k=1}^{M} \prod_{m=1, m\neq k}^{M} (2d + l_m) = M(2d)^{M-1} \tag{A14}$$

Since $l_1, l_2, \ldots, l_M \geq 0$, and taking into account Eq. (A14) it holds

$$\left.\begin{array}{l} M(2d)^{M-1} \leq \sum_{k=1}^{M} \prod_{m=1, m\neq k}^{M} (2d + l_m) \\ a \in (0,\infty) \Rightarrow f_2 \in (0.5, 1) \end{array}\right\}$$

$$\Rightarrow (1 - 2f_2) \sum_{k=1}^{M} \prod_{m=1, m\neq k}^{M} (2d + l_m) \leq (1 - 2f_2) M (2d)^{M-1} \tag{A15}$$

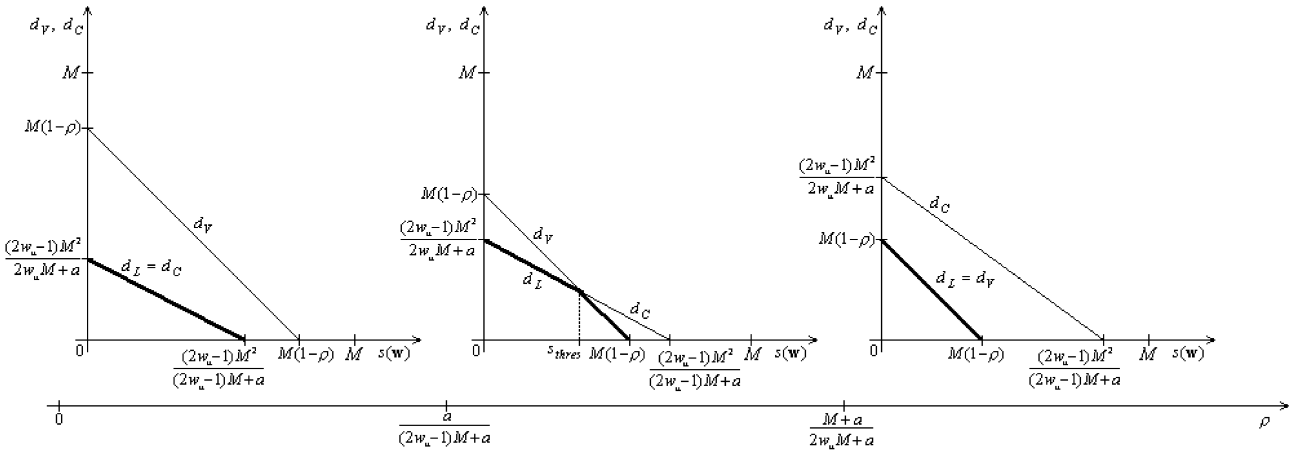

Fig. 10. Plots of match, choice and claim region radii versus category size $s(\mathbf{w})$ for different values of ρ. The figure displays the plots of the match (d_V), choice (d_C) and claim region radius (d_L) of a category versus its size $s(\mathbf{w})$ for three different ranges of the value of ρ. The graph for d_L is depicted as a thicker line in all three plots. In the leftmost plot d_L coincides with d_C, therefore a category's claim region will coincide with its choice region for the corresponding interval of ρ values. Using the same rationale, the rightmost plot indicates that a category's claim region coincides with its match region for the corresponding interval of ρ values.

It also holds that

$$\sum_{k=1}^{M}(\delta_{k,p}-2f_2)\prod_{m=1,m\neq k}^{M}(2d+l_m) \le (1-2f_2)\sum_{k=1}^{M}\prod_{m=1,m\neq k}^{M}(2d+l_m) \quad \text{(A16)}$$

Due to both Eqs. (A16) and (A17) and the fact that for $a>0$ and $w_u \ge 1$ it holds $f_2 > 1/M$ we get

$$\sum_{k=1}^{M}(\delta_{k,p}-2f_2)\prod_{m=1,m\neq k}^{M}(2d+l_m) \le (1-2f_2)M(2d)^{M-1} \overset{\text{Eq. (A13)}}{\Rightarrow} \frac{\partial \mathrm{Vol}(C(\mathbf{w}|a,w_u))}{\partial l_p} \le (1-Mf_2)(2d)^{M-1} < 0$$

□

Proof of Property 11. From Definition 5 it is obvious, since $d_L(\mathbf{w}_j|\rho,a,w_u)=\min\{d_V(\mathbf{w}_j|\rho),d_C(\mathbf{w}_j|a,w_u)\}$, that a category's claim region will either coincide with its match or its choice region depending on the values of ρ and a. Fig. 10 displays for different ranges of ρ the plots of the regions' radii versus category size. We distinguish three major cases:

(i) If $0\le\rho\le a/[(2w_u-1)M+a]$ and any $a>0$ it holds $((2w_u-1)M^2)/((2w_u-1)M+a)\le M(1-\rho)$. Therefore, $d_C(\mathbf{w}_j|a,w_u)\le d_V(\mathbf{w}_j|\rho)$ and from Definitions 3 and 4 we conclude that for this particular range $C(\mathbf{w}_j|a,w_u)\subseteq V(\mathbf{w}_j|\rho)$. Finally, it follows from Definition 5 that $L(\mathbf{w}_j|\rho,a,w_u)=C(\mathbf{w}_j|a,w_u)$.

(ii) Now assume that $a/[(2w_u-1)M+a]<\rho<(M+a)/(2Mw_u+a)$, where $a>0$. Under these circumstances, we have $M(1-\rho)<((2w_u-1)M^2)/((2w_u-1)M+a)$. We observe in Fig. 10 that for this range of ρ if $s(\mathbf{w}_j)<s_{\text{thres}}$, then $d_C(\mathbf{w}_j|a,w_u)\le d_V(\mathbf{w}_j|\rho)$, on which were based Definitions 3 and 4 means that $C(\mathbf{w}_j|a,w_u)\subset V(\mathbf{w}_j|\rho)$. Again, from Definition 5 we conclude that $L(\mathbf{w}_j|\rho,a,w_u)=C(\mathbf{w}_j|a,w_u)$. Similarly, when $s_{\text{thres}}<s(\mathbf{w}_j)$, $V(\mathbf{w}_j|\rho)\subset C(\mathbf{w}_j|a,w_u)$, thus $C(\mathbf{w}_j|\rho,a,w_u)=V(\mathbf{w}_j|\rho)$. Obviously, when $s(\mathbf{w}_j)=s_{\text{thres}}$, all three regions coincide.

(iii) If $(M+a)/(2Mw_u+a)\le\rho\le 1$, where $a>0$, then $M(1-\rho)<((2w_u-1)M^2)/((2w_u-1)M+a)$. Also, after examining once more Fig. 10 we observe that $d_V(\mathbf{w}_j|\rho)\le d_C(\mathbf{w}_j|a,w_u)$, hence $V(\mathbf{w}_j|\rho)\subset C(\mathbf{w}_j|a,w_u)$. Once more time, it follows from Definition 5 that $L(\mathbf{w}_j|\rho,a,w_u)=V(\mathbf{w}_j|\rho)$.

□

Proof of Result 4.

(i) According to Property 11(i), for $0\le\rho\le a/[(2w_u-1)M+a]$ and any $a>0$ the claim region coincides with the choice region for all categories. As a consequence, categories that pass the CT will also pass the VT for the same patterns. Thus, the upper bound for category sizes is solely determined by the upper bound, which the choice region allows. According to Property 8, only a least upper bound of $((2w_u-1)M^2)/((2w_u-1)M+a)$ exists.

(ii) According to Property 11(ii, iii) for $a/[(2w_u-1)M+a]<\rho\le 1$ and $a>0$ and for category sizes sufficiently large, the claim region radius coincides with the match region radius (see Fig. 10). In such a case the upper bound for category sizes is determined by the one imposed due to the match region, as stated in Property 2. Therefore, for this particular range of ρ the maximum size categories can achieve is $M(1-\rho)$.

To combine the above results in a single expression we observe that, if $0 \leq \rho \leq a/[(2w_u - 1)M + a]$, then $((2w_u - 1)M^2)/((2w_u - 1)M + a) \leq M(1 - \rho)$ and if $a/[(2w_u - 1)M + a] < \rho \leq 1$, then $M(1 - \rho) < ((2w_u - 1)M^2)/((2w_u - 1)M + a)$. Therefore, we can state that

$$s(\mathbf{w}) \leq \min\left\{M(1-\rho), \frac{(2w_u - 1)M^2}{(2w_u - 1)M + a}\right\} \Rightarrow s(\mathbf{w}) \leq M\left[1 - \max\left\{\rho, \frac{a}{(2w_u - 1)M + a}\right\}\right] \quad \text{(A17)}$$

However, let us note here that Eq. (A17) does not reflect the fact that $((2w_u - 1)M^2)/((2w_u - 1)M + a)$ is only a least upper bound for $s(\mathbf{w})$, when $0 \leq \rho \leq a/[(2w_u - 1)M + a]$ and $a > 0$. □

References

Anagnostopoulos, G. C., & Georgiopoulos, M. (2001). Ellipsoid ART and ARTMAP for incremental unsupervised and supervised learning. *Proceedings of the IEEE–INNS–ENNS International Joint Conference on Neural Networks* (Vol. 2, pp. 1221–1226). Washington, DC: IEEE, INNS, ENNS.

Bezdek, J. C., Reichherzer, T., Lim, G. S., & Attikiouzel, Y. (1998). Multiple-prototype classifier design. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *28*(1), 67–79.

Carpenter, G. A. (1997). Distributed learning, recognition, and prediction by ART and ARTMAP neural networks. *Neural Networks*, *10*(8), 1473–1494.

Carpenter, G. A., Grossberg, S., Markuzon, N., Reynolds, J. H., & Rosen, D. B. (1992). Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transaction on Neural Networks*, *3*(5), 698–713.

Carpenter, G. A., Grossberg, S., & Rosen, D. B. (1991). Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks*, *4*(6), 759–771.

Carpenter, G. A., Milenova, B. L., & Noeske, B. W. (1998). Distributed ARTMAP: A neural network for fast distributed supervised learning. *Neural Networks*, *11*, 793–813.

Carpenter, G. A., & Markuzon, N. (1998). ARTMAP-IC and medical diagnosis: Instance counting and inconsistent cases. *Neural Networks*, *11*(2), 323–336.

Carpenter, G. A., & Ross, W. D. (1995). ART-EMAP: A neural network architecture for object recognition by evidence accumulation. *IEEE Transactions on Neural Networks*, *6*, 805–818.

Carpenter, G. A., & Tan, A.-H. (1995). Rule extraction: From neural network architecture to symbolic representation. *Connection Science*, *7*(1), 3–27.

Georgiopoulos, M., Fernlund, H., Bebis, G., & Heileman, G. L. (1996). Order of search in Fuzzy ART and Fuzzy ARTMAP: Effect of the choice parameter. *Neural Networks*, *9*(9), 1541–1559.

Gomez Sanchez, E., Dimitriadis, Y.A., Cano Izquierdo, J.M., & Lopez Coronado, J. (2000). MicroARTMAP: Use of mutual information for category reduction in Fuzzy ARTMAP. *Proceedings of the IEEE–INNS–ENNS International Joint Conference on Neural Networks* (Vol. 6, pp. 47–52). Como, Italy: IEEE, INNS, ENNS.

Grossberg, S. (1976). Adaptive pattern recognition and universal encoding II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, *23*, 187–202.

Healy, M. J., & Caudell, T. P. (1998). Guaranteed two-pass convergence for supervised and inferential learning. *IEEE Transactions on Neural Networks*, *9*(1), 195–204.

Huang, J. X., Georgiopoulos, M., & Heileman, G. L. (1995). Fuzzy ART properties. *Neural Networks*, *8*(2), 203–213.

Petridis, V., Kaburlasos, V.G., Fragkou, P., & Kehagias, A. (2001). Text classification using the σ-FLNMAP neural network. *Proceedings of the IEEE–INNS–ENNS International Joint Conference on Neural Networks* (Vol. 2, pp. 1362–1367). Washington, DC: IEEE, INNS, ENNS.

Simpson, P. K. (1992). Fuzzy min max neural networks. 1. Classification. *IEEE Transactions on Neural Networks*, *3*(5), 776–786.

Verzi, S.J., Heileman, G.L., Georgiopoulos, M., & Healy, M.J. (1998). Boosted ARTMAP. *Proceedings of the IEEE–INNS–ENNS International Joint Conference on Neural Networks* (Vol. 1, pp. 396–401). Anchorage, AK: IEEE, INNS, ENNS.

Williamson, J. R. (1996). Gaussian ARTMAP: A neural network for fast incremental learning of noisy multidimensional maps. *Neural Networks*, *9*(5), 881–897.

Williamson, J. R. (2001). Self-organization of topographic mixture networks using attentional feedback. *Neural Computation*, *13*, 563–593.