2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

An Introduction to Tunable Equivalence Fuzzy Associative Memories

Estevo Esmi
University of Campinas
Campinas, Brazil
Email: eelaureano@gmail.com.br

Peter Sussner
University of Campinas
Campinas, Brazil
Email: sussner@ime.unicamp.br

Sandra Sandri
Brazilian National
Institute for Space Research (INPE)
S.J.Campos, Brazil
Email: sandra.sandri@inpe.br

***Abstract*—In this paper, we present a new class of fuzzy associative memories (FAMs) called *tunable equivalence fuzzy associative memories*, for short tunable E-FAMs or TE-FAMs, that belong to the class Θ-fuzzy associative memories (Θ-FAMs). Recall that Θ-FAMs represent fuzzy neural networks having a competitive hidden layer and weights that can be adjusted via a training algorithm. Like any associative memory model, Θ-FAMs depend on the specification of a fundamental memory set. In contrast to other Θ-FAM models, TE-FAMs make use of parametrized fuzzy equivalence measures that are associated with the hidden nodes and allow for the extraction of a fundamental memory set from the training data. The use of a smaller fundamental memory set than in previous articles on Θ-FAMs reduces the computational effort involved in deriving the weights without decreasing the quality of the results.**

I. Introduction

An *associative memory* (AM) is geared to storing a finite set of associations $(\mathbf{x}^\xi, \mathbf{y}^\xi)$, called fundamental memories, that are defined on arbitrary domains $\mathbf{X}$ and $\mathbf{Y}$. Thus, an associative memory yields a function $\Phi : \mathbf{X} \to \mathbf{Y}$. Ideally, $\Phi(\mathbf{x}^\xi) = \mathbf{y}^\xi$ and $\Phi(\tilde{\mathbf{x}}^\xi) = \mathbf{y}^\xi$ for corrupted or noisy versions $\tilde{\mathbf{x}}^\xi$ of $\mathbf{x}^\xi$. In the special case, where $\mathbf{X}$ and $\mathbf{Y}$ are respectively given by classes of fuzzy sets on some arbitrary domains X and Y, Φ is called a *fuzzy associative memory* (FAM).

The class of Θ-FAMs was recently introduced in the literature [1]. The activation functions of the ξth hidden nodes in a Θ-FAM are functions Θ^ξ that fulfil the condition $\Theta^\xi(x^\xi) = 1$ for all ξ. In contrast to many previous FAM models [2], [3], [4], [5], [6], [7], Θ-FAMs have a competitive hidden layer and, in this respect, they are similar to the Hamming net [8], that is also known as Hamming associative memories, and its extensions [9], [10].

The class of Θ-FAMs includes (weighted) subsethood, dual subsethood, and similarity measure FAMs whose hidden layer activation functions can be expressed in terms of subsethood or similarity measures [11], [12]. These particular Θ-FAM models as well as the aforementioned FAMs [2], [3], [4], [5], [6], [7] can be referred to as fuzzy morphological associative memories [13] since they perform elementary morphological operations in the complete lattice setting [14], [15], [16] of fuzzy mathematical morphology [17].

Weighted S-FAMs, dual S-FAMs, and SM-FAMs models were successfully applied to a number of classifications problems and to a problem of vision-based self-localization in robotics [1]. The tuning of the weights was performed using a specifically designed Θ-FAM training algorithm that is guaranteed to converge in finite number of steps and, under some weak conditions, to reach a local minimum of the proposed objective function. The high computational cost of tuning Θ-FAMs using this training algorithm, prohibits its use in applications with a large training set. Another option for tuning the weights of these Θ-FAM models would be the use of derivative-free non-linear optimization methods, some of which also require a high computational effort when dealing with problems involving a large number of variables [18].

In this work, we propose the use of a new type of Θ-FAMs, called *tunable equivalence fuzzy associative memories*, for short tunable E-FAMs or TE-FAMs, in which the Θ^ξs are modelled in terms of parametrized fuzzy equivalence measures such as the ones employed by Bedê et al. in the context of fuzzy case-based reasoning systems [19] and by Drummond et al. to restore consistency in fuzzy rule-based systems using residuated implication operators [20].

In contrast to previous Θ-FAM models, TE-FAMs allow for optimizing not only the weights corresponding to the contributions of the hidden nodes but also the contributions of the attributes of the data by adjusting the parameters of the equivalence measures that determine the hidden node activations. Before performing the optimization of the parameters, we extract a set of fundamental memories from the given training set. We illustrate the selection of the fundamental memory set and the optimization of the parameters using a specific example. A detailed description of the training algorithm can be found in [21], [22].

II. Mathematical Background

This section is organized as follows. First of all, we present well-established basic concepts that will be used throughout the text. We then introduce the idea of equivalence measure based on the works of Fodor's and Roubens' [23], [24] as well as of Bustince et al. [25].

A. Some Basic Concepts and Notations

In 1965, Lotfi A. Zadeh extended the classical notion of a set, by introducing the concept of fuzzy sets [26]. A fuzzy set A consists of a set X, the universe of discourse, and a membership function $\mu_A : X \to [0, 1]$, that assigns a membership degree $\mu_A(x) \in [0, 1]$ in A to each $x \in X$. In the following, the class of fuzzy sets on the universe X is denoted

978-1-4799-2072-3/14/$31.00 ©2014 IEEE

by $\mathcal{F}(X)$. A fuzzy set A is said to be normal if $\mu_A(x) = 1$ for some $x \in X$.

Mathematically speaking, a fuzzy set A on a universe X can be identified with its membership function μ_A and can simply be viewed as a function from the universe X to the unit interval $[0, 1]$. If X is finite, say $X = \{x_1, \ldots, x_n\}$, then $\mathcal{F}(X)$ can be identified with $[0, 1]^n$ via the bijection that maps $A \in \mathcal{F}(X)$ to the vector $(A(x_1), \ldots, A(x_n))^t \in [0, 1]^n$. In particular, if A (or, more precisely, μ_A) only produces values in $\{0, 1\}$, the A represents a classical or crisp set. Thus, a relation, which is simply a subset of $X \times Y$ for arbitrary universes X and Y, can viewed as a function $X \times Y \to \{0, 1\}$. Extending this classical concept of a relation to the fuzzy domain, a fuzzy relation is given by a function $R : X \times Y \to [0, 1]$, where $R(x, y)$ can be interpreted as the degree of relationship between x and y [27]. A fuzzy relation on $X \times Y$ is therefore nothing else than a fuzzy set on $X \times Y$.

Two fuzzy relations $R \in \mathcal{F}(X \times Y)$ and $S \in \mathcal{F}(Y \times Z)$, where X, Y, and Z are arbitrary universes, can be combined using a fuzzy relational composition that yields a fuzzy relation on the universe $X \times Z$. In this paper, we only address fuzzy relational compositions of the type sup-t defined in Equations 1 and 2 below, where $t : [0, 1]^2 \to [0, 1]$ stands for a t-norm operator (an associative, commutative, and increasing mapping with identity element 1).

The class of fuzzy sets over an arbitrary universe X, and in particular $[0, 1]$ and $[0, 1]^n$, represent partially ordered sets. Recall that a pair $(P, \leq)$ consisting of a non-empty set P together with a reflexive, antisymmetric, and transitive binary relation "$\leq$" is called a *partially ordered set* or *poset* [28]. If the partial order relation $\leq$ clearly arises from the context, then we simply refer to the poset under consideration using the symbol P instead of $(P, \leq)$. If $X \subseteq P$, then an element $l \in P$ is said to be a *lower bound* of X if $l \leq x$ for all $x \in X$. Similarly, $u \in P$ is said to be an *upper bound* of X if $x \leq u$ for all $x \in X$. The *infimum* of $X \subseteq P$, denoted using the symbol $\bigwedge X$, is defined as the greatest lower bound of X. Similarly, the *supremum* of $X \subseteq P$, denoted using the symbol $\bigvee X$, is defined as the least upper bound of X. If $X = \{x, y\}$, then we may alternatively write $x \vee y$ and $x \wedge y$ instead of $\bigvee X$ and $\bigwedge X$, respectively.

A partially ordered set $\mathbb{L}$ is called a *lattice* if every finite, non-empty subset of $\mathbb{L}$ has an infimum and a supremum in $\mathbb{L}$ [29]. A lattice $\mathbb{L}$ is *complete* if every subset of $\mathbb{L}$ has an infimum and a supremum in $\mathbb{L}$. The supremum of $r, s \in [0, 1]$ is s if and only if $r \leq s$. Similarly, $r \vee s = r$, if and only if $s \leq r$. If t is a t-norm, then the sup-t composition of $R \in \mathcal{F}(X \times Y)$ and $S \in \mathcal{F}(Y \times Z)$ yields fuzzy relation $R \circ_t S \in \mathcal{F}(X \times Z)$:

$$\forall x \in X, \forall z \in Z, R \circ_t S(x, z) = \bigvee_{y \in Y} R(x, y) \, t \, S(y, z), \quad (1)$$

and the sup-t composition of $R \in \mathcal{F}(X \times Y)$ and $S \in \mathcal{F}(Y)$ yields fuzzy set $R \circ_t S \in \mathcal{F}(X)$:

$$\forall x \in X, R \circ_t S(x) = \bigvee_{y \in Y} R(x, y) \, t \, S(y). \quad (2)$$

Let $\mathbb{L}^X$ denote the class of all functions $X \to \mathbb{L}$. If $\mathbb{L}$ represents a poset, then the partial ordering $\leq$ on $\mathbb{L}$ induces a partial ordering on $\mathbb{L}^X$, also denoted by $\leq$, is defined as follows, for all $f, g \in \mathbb{L}^X$:

$$\forall \, x \in X, f \leq g \Leftrightarrow f(x) \leq g(x). \quad (3)$$

The lattice structure on $\mathbb{L}$ induces the same lattice structure on $\mathbb{L}^X$. Note that $\mathcal{F}(X)$, is given by the complete lattice $[0, 1]^X$.

B. Fuzzy Equivalence

In this paper, we employ the notion of an equivalence measure according to Bustince et al. [25] who defined an equivalence measure as a binary operation $\mathcal{F}(X) \times \mathcal{F}(X) \to [0, 1]$ by extending Fodor's and Roubens' original definition on $[0, 1]$ [23], [24].

Definition 1: Let X be an arbitrary universe. An equivalence measure on $\mathcal{F}(X)$ is a function $E : \mathcal{F}(X) \times \mathcal{F}(X) \to [0, 1]$ that satisfies the following conditions:

- E1) $E(x, y) = E(y, x)$ for all $x, y \in \mathbb{L}$;
- E2) $E(0_{\mathbb{L}}, 1_{\mathbb{L}}) = 0$;
- E3) $E(x, x) = 1$ for all $x \in \mathbb{L}$;
- E4) if $x \leq y \leq z$, then $E(x, z) \leq E(x, y)$ and $E(x, z) \leq E(y, z)$.

In the literature of fuzzy set theory, one can find other concepts that are closely related to the one of (fuzzy) equivalence, such as fuzzy similarity measure and equality index [25], [30], [31], [32], [33], [34], [35]. In fact, these technical terms are either used interchangeably or their definitions are given by different, but slight modifications of Definition 1 (see [25] for a comparison). The main difference between the concepts of similarity and equivalence seems to be the requirement that the similarity between a crisp set and its complement should be zero [33], [34]. This property is crucial for the definition of entropy or fuzziness measures by means of similarity measures since it is reasonable to assume that the entropy of a crisp set equals 0 [25], [33], [34], [35]. Fuzzy equivalence and similarity measures have been extensively studied in the last few years and many authors have proposed formulas to produce fuzzy equivalence measures on the class of fuzzy sets over a universe U [25], [32], [33], [34]. In [ref sandra], a mapping $E : \mathbb{R}^2 \to [0, 1]$, named an *order compatible fuzzy relation* (OCFR), has been proposed, requiring conditions E1, E2 and E3. In that work, E4 is called "compatibility with total order", a property that distinguishes OCFRs from fuzzy relations that generalize the notion of equivalence relation from classical set theory by requiring t-transitivity [36].

An infinite number of equivalence measures on $\mathcal{F}(X)$ can be generated from fuzzy subsethood or inclusion measures [1], [11], [25], [34], [35]. The concept of equivalence measure on $\mathcal{F}(X)$ is also related to the concepts of fuzzy entropy and distance measures [32], [33], [34].

Below we define *strong equivalence measures* by adding another condition to the ones of Definition 1:

Definition 2: An equivalence measure E is said to be *strong* if $\forall x, y \in X, E(x,y) = 1 \Leftrightarrow x = y$.

Definition 2 yields Zeng's & Li's definition of a similarity measure [35]. We prefer to speak of a *strong equivalence* measure to avoid creating a confusion with several different definitions of similarity measure in the literature [25], [30], [31], [33], [35]. Below, we give an example of a strong equivalence measure, that is employed for restoring consistency in implication-based fuzzy rule-based systems [refs sandra] and in fuzzy CBRs [19].

Example 1: For each $\lambda \in (0,1]$, the function $E_\lambda : [0,1] \times [0,1] \to [0,1]$ given by

$$E_\lambda(x,y) = \max\left(0\,,\, 1 - \frac{|x-y|}{\lambda}\right) \tag{4}$$

yields a strong equivalence measure on $[0,1]$.

Example 2 below yields a class of strong equivalence measures that can be used to assign weights to attributes in applications involving multi-dimensional domains.

Example 2: Let E_i be equivalence measures on $[0,1]$ for $i = 1,\ldots,n$ and let $\mathbf{w} \in [0,1]^n$ such that $\sum_{i=1}^n w_i = 1$. A strong equivalence measure $E_\mathbf{w} : [0,1]^n \times [0,1]^n \to [0,1]$ can be defined as follows:

$$E_\mathbf{w}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^n w_i E_i(x_i, y_i) \;\forall \mathbf{x},\mathbf{y} \in [0,1]^n. \tag{5}$$

is an equivalence measure on $\mathbb{L}$.

In particular, we obtain the following (strong) tunable equivalence measure on $[0,1]^n$:

$$E_{\lambda,\mathbf{w}}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^n w_i E_{\lambda_i}(x_i, y_i) \;\forall \mathbf{x},\mathbf{y} \in [0,1]^n. \tag{6}$$

where $\lambda = (\lambda_1,\ldots,\lambda_n)^t \in (0,1]^n$.

Equivalence measures of the form given in Equation 6 have been employed in fuzzy case-based reasoning [19]. A particular case of an OCFR, proposed in [36], yields another example of a parametrized equivalence measure.

Example 3: Let $\mathbf{A} = \{A_1, ..., A_n\}$ be a collection of fuzzy sets on a universe X such that each A_i is normal, $\sum_i \mu_{A_i}(x) = 1$ for all $x \in X$, and each element in X has a membership degree $\neq 0$ in at most 2 fuzzy sets in $\mathbf{A}$. In other words, $\mathbf{A}$ is a fuzzy partition. Note that the elements of $\mathbf{A}$ may be either crisp sets or trapezoidal fuzzy sets. If $S^*(x,y) = \bigvee_{i=1}^n \min(A_i(x), A_i(y))$ and $S_L(x,y) = \bigwedge_{i=1}^n 1 - \mid \mu_{A_i}(x) - \mu_{A_i}(y) \mid$ for all $x, y \in X$, then

$$\forall x,y \in X, S^+(x,y) = \begin{cases} 0, \text{ if } S^*(x,y) = 0 \\ S_L(x,y), \text{ otherwise} \end{cases} \tag{7}$$

yields an equivalence measure on $[a,b]$. In particular, when $\mathbf{A}$ is formed solely by triangular fuzzy sets, E is a strong equivalence measure on $[a,b]$.

III. The Class of Θ-FAMs

The class of Θ-FAMs appeared only recently in the literature [1]. Let us briefly review the definition of Θ-FAMs that contain TE-FAMs as special cases.

Definition 3: Let $\mathcal{M} = \{(A^\xi, B^\xi) \in \mathcal{F}(X) \times \mathcal{F}(Y) : \xi = 1,\ldots,p\}$ be a fundamental memory set, where X and Y are arbitrary universes. Furthermore, let $\Theta^\xi : \mathcal{F}(X) \to [0,1]$ be operators such that $\Theta^\xi(A^\xi) = 1$ for $\xi = 1,\ldots,p$ and let $\mathbf{v}$ be an arbitrary vector in $\mathbb{R}^p$. The *Θ-FAM based on* Θ^ξ *and* $\mathbf{v}$ produces the following output $\mathcal{O}(A) \in \mathcal{F}(Y)$ upon presentation of an input fuzzy set $A \in \mathcal{F}(X)$:

$$\mathcal{O}(A) = \bigvee_{j \in I_\mathbf{v}(A)} B^j, \tag{8}$$

where

$$I_\mathbf{v}(A) = \left\{ j \in \{1,\ldots,p\} : v_j \Theta^j(A) = \max_{\xi=1,\ldots,p} v_\xi \Theta^\xi(A) \right\}. \tag{9}$$

Note that this model is well-defined because an index j than maximizes $\{v_\xi \Theta^\xi(A)\}_{\xi=1}^p$ is guaranteed to exist and therefore $I_\mathbf{v}(A) \neq \emptyset$. If $(A^\xi, B^\xi) \in [0,1]^n \times [0,1]^m$, then the resulting Θ-FAM corresponds to a two-layer neural network with p and m neurons in the hidden and output layers, respectively. Upon presentation of an input pattern $A = (A_1, A_2, \ldots, A_n)^t$, the jth hidden neuron computes $v_j \Theta^j(A)$ followed by a competitive binary activation function that yields 1 if and only if $v_j \Theta^j(A) \geq v_i \Theta^i(A)$ for all $i \neq j$. Hence, the hidden layer can be implemented in terms of a MAXNET network [37]. Figure 1 depicts the topology of the entire network for finite universes $X = \{x_1,\ldots,x_n\}$ and $Y = \{y_1,\ldots,y_m\}$.

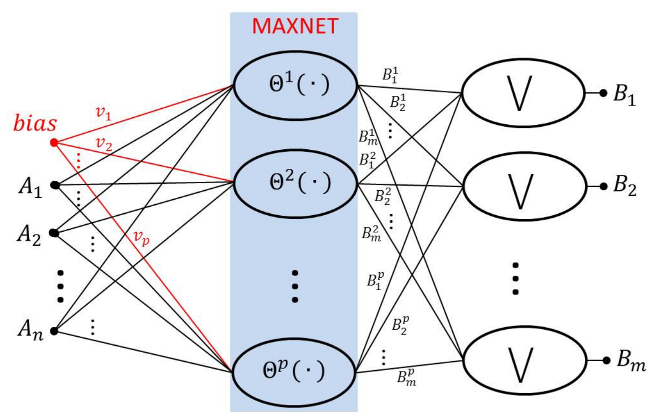

Fig. 1. Topology of a Θ-FAM for finite universes X and Y.

Esmi et al. [1] provided theoretical results concerning the storage capacities as well as the error correction capabilities of Θ-FAM models. These results include sufficient conditions for the perfect recall of all fundamental memories and a characterization of the basins of attraction around each A^ξ. The weight vector $\mathbf{v}$ can be optimized using a supervised training algorithm that was specifically designed for Θ-FAMs. Under some weak conditions, this supervised algorithm is guaranteed to converge to a local minimum of the objective function in a finite number of steps.

IV. TE-FAMs in the Context of Θ-FAMs

As mentioned before, the computational cost involved in the previous Θ-FAM training algorithms can be extremely high for large sets of fundamental memories. This drawback motivated us to develop a new supervised algorithm for a certain subclass of Θ-FAMs whose functions Θ^ξ are given in terms of equivalence measures. For example, training a weighted KS-FAM, i.e., a Θ-FAM based on the Kosko subsethood measure [2], using the algorithm proposed in [1], required a CPU time of about one hour for the Vowel dataset that comprises 990 instances [38].

Note that if E denotes an equivalence measure then the operator $E(\,\cdot\,, A^\xi)$ can play the role of Θ^ξ in Definition 3 since $\Theta^\xi(A^\xi) = E(A^\xi, A^\xi) = 1$ by property E3) of Definition 1. Thus, the notion of weighted SM-FAM can be generalized as follows.

Definition 4: Consider a Θ-FAM as defined in Equations 8 and 9. If each function $\Theta^\xi : \mathcal{F}(X) \to [0,1]$ of Definition 3 is given by $\Theta^\xi(\,\cdot\,) = E^\xi(\mathbf{x}^\xi, \,\cdot\,)$ for some equivalence measure E^ξon $\mathcal{F}(X)$, then the corresponding Θ-FAM is called a *equivalence measure fuzzy associative memory*, for short E-FAM. In the special case, where each E^ξ is parametrized, we speak of a *tunable equivalence measure fuzzy associative memory* (TE-FAM).

Throughout the rest of this manuscript we will concentrate on TE-FAMs based on tunable equivalence measures $E_{\mathbf{w}^\xi}$ on the product lattice $[0,1]^n$. Let us gather $\mathbf{w}^1, \ldots, \mathbf{w}^p \in [0,1]^n$ as column vectors of a matrix $W = [\mathbf{w}^1, \ldots, \mathbf{w}^p] \in [0,1]^{n \times p}$ and denote the corresponding TE-FAM using the symbol $\mathcal{E}_{W,\mathbf{v}}$. Thus, given an input pattern $\mathbf{x} \in \mathbb{L}$, the TE-FAM $\mathcal{E}_{W,\mathbf{v}}$ generates the following output $\mathcal{E}_{W,\mathbf{v}}(\mathbf{x}) \in \mathcal{F}(Y)$:

$$\mathcal{E}_{W,\mathbf{v}}(\mathbf{x}) = B \circ_t F(v_1 E_{\mathbf{w}^1}(\mathbf{x}, \mathbf{x}^1), \ldots, v_p E_{\mathbf{w}^p}(\mathbf{x}, \mathbf{x}^p)) \quad \forall \mathbf{x} \in [0,1]^n, \tag{10}$$

where

$$[F(x_1, \ldots, x_k)]_i = \begin{cases} 1 & \text{, if } x_i = \vee_{j=1}^k x_j \\ 0 & \text{, otherwise} \end{cases} \text{, for } i = 1, \ldots, p. \tag{11}$$

Note that the output $\mathcal{E}_{W,\mathbf{v}}(\mathbf{x})$ does not depend on the choice of the t-norm t in Equation 10. Figure 2 visualizes the topology of a TE-FAM of the form $\mathcal{E}_{W,\mathbf{v}}$ for a finite universe Y.

Esmi et al. proposed a training algorithm for TE-FAMs of the form $\mathcal{E}_{W,\mathbf{v}}$ that comprises two stages:

1) Extraction of the set of fundamental memories from the training set $\mathcal{T}$;
2) Optimization of the parameters $\mathbf{w}^\xi$ and $\mathbf{v}$.

Stage 1 was devised in order to generate a small fundamental memory set $\mathcal{M} \subseteq \mathcal{T}$ and determines the exact topology of $\mathcal{E}_{W,\mathbf{v}}$ since initially p and $\mathbf{x}^1, \ldots, \mathbf{x}^p$ are unknown. Under some weak conditions, the resulting TE-FAM with fundamental memory set $\mathcal{M}$ produces no training error. Stage 2 consists in optimizing a quadratic programming problem subject to linear inequality constraints in order to find suitable parameters for $\mathcal{E}_{W,\mathbf{v}}$. The interested reader can find more details in [1].

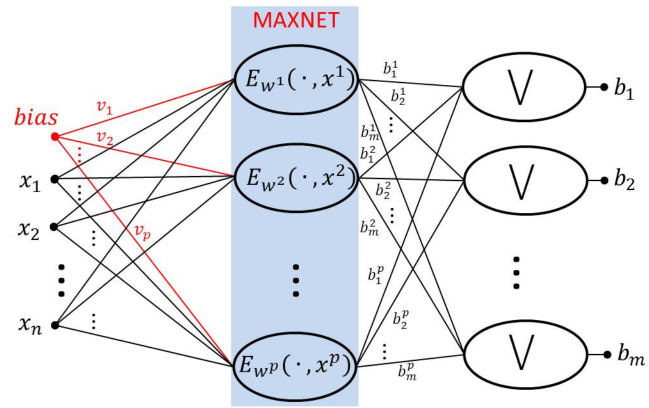

Fig. 2. Topology of a TE-FAM of the form $\mathcal{E}_{W,\mathbf{v}}$ for a finite universe Y.

V. An Illustration of the Training Algorithm for TE-FAMs of Type $\mathcal{E}_{W,\mathbf{v}}$

We have created a small synthetic binary classification problem in order to illustrate the training algorithm for TE-FAMs $\mathcal{E}_{W,\mathbf{v}}$. Let $\mathcal{T} \subset [0,1]^2 \times \{0,1\}$ be the following training set:

$$\mathcal{T} = \left\{ \begin{array}{c} (\mathbf{x}^1 = (0.66, 0.13), 0),\ (\mathbf{x}^2 = (0.93, 0.33), 0), \\ (\mathbf{x}^3 = (0.77, 0.37), 0),\ (\mathbf{x}^4 = (0.47, 0.33), 0), \\ (\mathbf{x}^5 = (0.27, 0.87), 1),\ (\mathbf{x}^6 = (0.17, 0.8), 1) \end{array} \right\}.$$

For a given fundamental memory set $\mathcal{M}$ and equivalence measures E^ξ, a TE-FAM can be defined as a map from a class of fuzzy sets on a universe X with $|X| = 2$ to a class of fuzzy sets on a universe Y with $|Y| = 1$. If $E_{\lambda,\mathbf{w}}$ denotes the equivalence measure on $\mathcal{F}(X) = [0,1]^2$ defined as in Equation 6 with $\lambda_1 = \lambda_2 = 1$ and $\mathbf{w} = (0.5, 0.5)$, then we have the following matrix $\mathbf{E}$ with components $\mathbf{E}_{j,k} = E_{\lambda,\mathbf{w}}(\mathbf{x}^j, \mathbf{x}^k)$, with $\mathbf{x}^j, \mathbf{x}^k \in \mathcal{T}$:

$$\mathbf{E} = \begin{pmatrix} 1.00 & 0.21 & 0.43 & 0.32 & 0.00 & 0.00 \\ 0.21 & 1.00 & 0.66 & 0.50 & 0.00 & 0.00 \\ 0.43 & 0.66 & 1.00 & 0.44 & 0.00 & 0.00 \\ 0.32 & 0.50 & 0.44 & 1.00 & 0.16 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.16 & 1.00 & 0.72 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.72 & 1.00 \end{pmatrix}$$

The fundamental memories, generated in Stage 1 of the training algorithm for TE-FAMs of type $\mathcal{E}_{W,\mathbf{v}}$, depend on the values $\mathbf{E}_{j,k}$ and are marked using red boxes in Figure 3. To be more precise, we obtained the following fundamental memory set: $\mathcal{M} = \{(\mathbf{x}^1, 0), (\mathbf{x}^2, 0), (\mathbf{x}^6, 1)\}$. Stage 2 of this algorithm produces the following parameters $W = [\mathbf{w}^1, \mathbf{w}^2.\mathbf{w}^3]$ and $\mathbf{v}$ for Equation 10:

$$W = \begin{pmatrix} 0 & 1 & 0.88 \\ 1 & 0 & 0.11 \end{pmatrix} \text{ and } \mathbf{v} = \begin{pmatrix} 0.97 \\ 1.05 \\ 0.87 \end{pmatrix}. \quad (12)$$

Hence, we obtain $\Theta^1(\cdot) = E_{\lambda,(0,1)}(\mathbf{x}^1, \cdot), \Theta^2(\cdot) = E_{\lambda,(1,0)}(\mathbf{x}^2, \cdot)$, and $\Theta^3(\cdot) = E_{\lambda,(0.88,0.11)}(\mathbf{x}^6, \cdot)$. For $\mathbf{x} \in [0,1]^2$, the resulting TE-FAMs $\mathcal{E}_{W,\mathbf{v}}$ produces the following output:

$$\mathcal{E}_{W,\mathbf{v}}(\mathbf{x}) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \circ_t F \begin{pmatrix} 0.97E_{\lambda,(0,1)}(\mathbf{x}^1, \mathbf{x}) \\ 1.05E_{\lambda,(1,0)}(\mathbf{x}^2, \mathbf{x}) \\ 0.87E_{\lambda,(0.88,0.11)}(\mathbf{x}^6, \mathbf{x}) \end{pmatrix}. \quad (13)$$

Figure 3 exhibits the decision surface produced by the resulting TE-FAM.

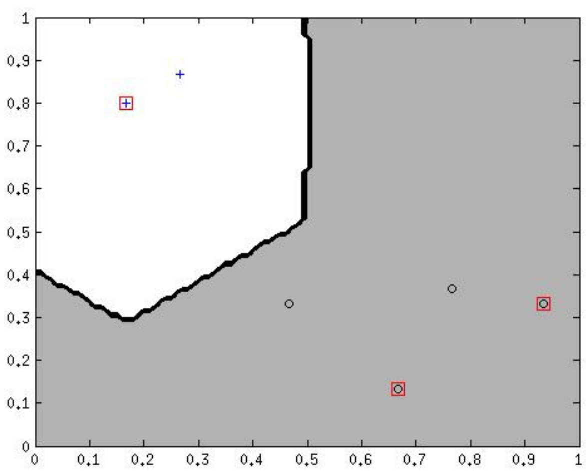

Fig. 3. Decision surface after the second stage of the training algorithm for $\mathcal{E}_{W,\mathbf{v}}$. Symbols $\circ$ and $+$ correspond to patterns belonging to the classes "0" and "1", respectively. Patterns marked by red boxes belong to the extracted fundamental memory set $\mathcal{M}$.

VI. Applications in Classification Problems

In this section, we describe the application of $\mathcal{E}_{W,\mathbf{v}}$ to two supervised classification problems.

Suppose that we are provided with a training set $\{(\mathbf{x}^\xi, y^\xi) \in \mathbf{X} \times L : \xi = 1, \ldots, k\}$, where $L = \{1, \ldots, c\}$ is a set of class labels. First, we need to convert the pairs $(\mathbf{x}^\xi, l^\xi)$ into pairs of fuzzy sets in order to be able to apply the TE-FAM approach. If $\mathbf{X} = \mathbb{R}^n$, then the vectors $\mathbf{x}^\xi$ can be converted into fuzzy sets $A^\xi \in [0,1]^n$ by normalizing each coordinate. The class labels y^ξ can be identified with the fuzzy sets $B^\xi \in \{0,1\}^c$ that satisfy $B^\xi_i = 1$ if and only if $\xi = i$.

A. Classification of Vertebral Column

In the following, we describe the use of T-E-FAMs in a classification corresponding to the Vertebral Column Data Set from the UCI Machine Learning Repository [39]. The patterns consist of biomedical data from 310 patients, 100 and 210 of which belong respectively to the classes "normal" (NO) and "abnormal" (AB), respectively. The patients are represented in the data set by six biomechanical attributes derived from the shape and orientation of the pelvis and lumbar spine: pelvic incidence, pelvic tilt, lumbar lordosis angle, sacral slope, pelvic radius and grade of spondylolisthesis.

In order to perform our experiments, we partitioned the entire set of data at random into three distinct subsets, Tr (training), Te (test) and Va (validation), containing respectively 50%, 25% and 25% of the patterns of each class.

The patterns $\mathbf{x}^\xi$ were mapped to the following fuzzy sets $A^\xi \in [0,1]^6$. Each attribute was normalized in order to obtain values in $[0,1]$. Then we applied the following function $\Psi : [0,1]^6 \to [0,1]^6$ whose ith coordinate is given by [12]:

$$\Psi(\mathbf{x})_i = \begin{cases} x_i & \text{if } x_i = 0 \text{ or } x_i = 1 \\ \dfrac{x_i - m_\mathbf{x} + 1.5}{3} & \text{otherwise.} \end{cases} \quad (14)$$

Here, $m_\mathbf{x}$ denotes the mean of $\mathbf{x} \in [0,1]^n$ over the restricted domain $I = \{i : 0 < x_i < 1\}$, that is

$$m_\mathbf{x} = \frac{\sum_{i \in I} x_i}{|\ I\ |}.$$

We evaluated the classification performances of TE-FAMs $\mathcal{E}_{W,\mathbf{v}}$ based on functions Θ^ξ that are determined by equivalence measures $E_{\lambda_\alpha,\mathbf{w}^\xi}$ as in Equation 6 with $\lambda_\alpha = (\alpha\sigma_1, \ldots, \alpha\sigma_6)$. Here, σ_i denotes the standard deviation of the set $\{A^\xi_i \mid \xi = 1, \ldots, k\}$ for $\alpha = 0.5, 0.51, \ldots, 1$. Each of these TE-FAMs was trained using the algorithm described in [1]. In Stage 2, we solve the proposed quadratic proggraming problem using the well-known *Trust-Region-Reflective Optimization* method [40], [41] with initial vectors $(\frac{1}{6}, \ldots, \frac{1}{6})^t \in [0,1]^6$.

Figures 4 and 5 show the percentage of fundamental memories in the training set and the error rate for the validation set, respectively, obtained by varying α in [0.5, 1]. Experiments performed for $\alpha > 1$ did not lead to an increase in the quality of the results. Figure 6 plots the validation error versus the percentage of fundamental memories in the training set. The best result is given by $\alpha = .88$, depicted in red in the figure.

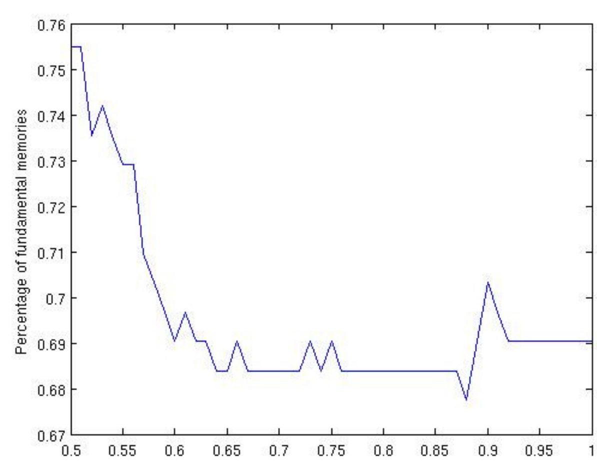

Fig. 4. Reduction rate of the fundamental memories set for $\alpha \in [.5, 1]$.

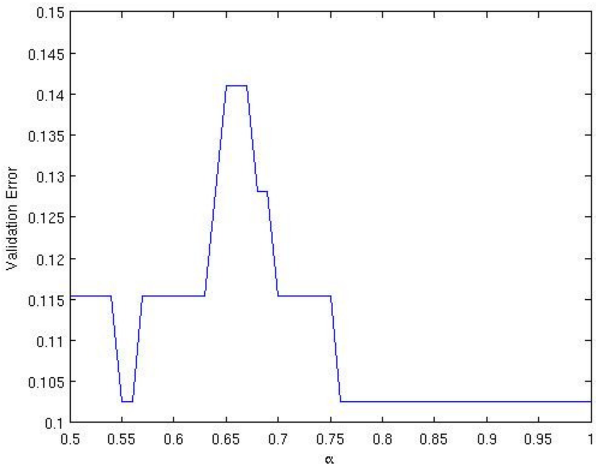

Fig. 5. Validation error for $\alpha \in [0.5, 1]$.

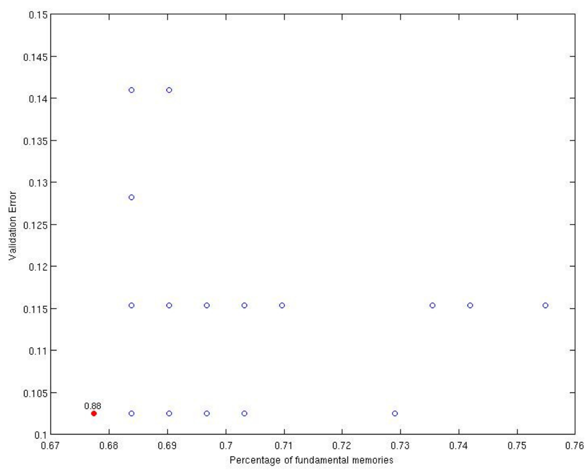

Fig. 6. Percentage of fundamental memories in the training set versus validation error for $\alpha \in [0.5, 1]$.

Table I exhibits the classification performance achieved by the TE-FAM with $\lambda_{0.88}$ and by other well-known models from the literature: kNN (*k-nearest neighbor*), MLP(*multi-layer perceptron*), SVM (*support vector machine*) [42], [43], and decision tree. To ensure a fair comparison, the parameters of these models were chosen by minimizing the validation errors within a range of values. For example, the classifier kNN was tested using $k = 1, 2, \ldots, 10$ and yielded the lowest validation error for $k = 1$. We also performed simulations using different MLPs with p hidden nodes for $p = 10, 20, 30$ and with sigmoid activation functions. Since the performance of an MLP depends on the initialization of weights in the training phase, we applied 10 MLPs with random initializations of weights for each p and selected the one that produced the lowest validation error. The MLPs were trained using backpropagation with early stopping, adaptive learning rate 0.01, increasing factor 1.2, reduction $0,7$, and gradient descent with *momentum* 0.9. The selected MLP has 30 hidden neurons and is denoted using the simbol MLP_{30}. We also performed experiments using

Model	$E_{tr}(\%)$	$E_{val}(\%)$	$E_{ts}(\%)$
TE-FAM	7.74	10,26	**14,29**
1NN	0.00	12.82	19.48
MLP_{30}	10.97	10.26	19.48
SVM	3.23	17.95	23.38
Tree	7.10	11.54	18.18

TABLE I. CLASSIFICATION ERRORS ON THE TRAINING SET (E_{tr}), VALIDATION SET (E_{val}), AND TEST SET (E_{ts}) PRODUCED BY DIFFERENT CLASSIFIERS.

10 different SVM models using gaussian kernels with spread factors $\sigma = 0, 5j$ for $j = 1, \ldots, 10$. The SVM classifier with $\sigma = 0.5$ yielding the lowest validation error produced a misclassification rate of 23.38% during the testing phase The last line of Table I exhibits the results obtained by a decision tree [44]. The TE-FAM approach achieved the best classification in this problem among all the classifiers that we tested.

B. *Classification of Time Series of Vegetation Indexes*

We also employed a TE-FAM model of type $\mathcal{E}_{W,\mathbf{v}}$ to associate a time series of vegetation indexes with one of the following classes: forest, deforestation, pasture and agriculture. Each time series datum represents a vector in $[0, 1]^{23}$ that corresponds to the value of a certain pixel of a multispectral image, given by a sequence of 23 images having values in $[0, 1]$. These multispectral images comprise 168 pixels, corresponding to an area of approximately $10.5km^2$ in the eastern part of the state of Mato Grosso, Brazil, and were captured from NASA's Aqua and Terra Satellites using a MODIS (Moderate Resolution Imaging Spectroradiometer) sensor. Pimentel et al. extracted two distinct datasets from 1848 patterns in $[0, 1]^{23}$, representing EVI-2 vegetation index data from August 2000 to June 2011 [45]. The first dataset has 205 data taken from areas having the following types of vegetation: 139 forest, 18 deforested areas, 35 pasture, and 13 of agricultural land. The second dataset contains 641 time series patterns of forest data, 100 of deforested areas, 279 of pasture data, and 343 of agricultural land.

We compared the results obtained by a TE-FAM model with the ones produced by MLP models in [46] and Θ-FAM models using the training algorithm presented in [1]. In order to ensure a fair comparison, we conducted our simulations following the instructions described in [46]. In particular, we applied the 4-fold cross-validation technique to each dataset in order to assess the classification performance of each model. We employed the same partitioning into 4 folds as in [46].

We considered a TE-FAM $\mathcal{E}_{W,\mathbf{v}}$ based on equivalence measures $E_{\lambda,\mathbf{w}^\xi}$ as in Equation 6 with $\lambda = (\sigma_1, \ldots, \sigma_{23})$, where σ_i denotes the standard deviation of the set $\{A_i^\xi \,|\, \xi = 1, \ldots, k\}$. As before, the TE-FAM $\mathcal{E}_{W,\mathbf{v}}$ was trained using the training algorithm presented in [21]. In addition, other types of Θ-FAM models were subjected to the training algorithm introduced in [1]. Specifically, we tested the Kosko subsethood measure [2] and the dual Kosko subsethood measure FAM as well as equivalence measure FAMs based on equivalence measures of the form $E_k(A, B) = \min\{S_k(A, B), S_k(B, A)\}$ [1] (E_k-FAMs). Recall that the Kosko subsethood measure $S_k : [0, 1]^n \times [0, 1]^n \to [0, 1]$ is defined as follows for all

Model	Dataset I		Dataset II	
	Train.	Test	Train.	Test
KS-FAM	**100.00 (± 0.00)**	98.98 (± 2.04)	99.19 (± 0.12)	91.30 (± 3.84)
Dual KS-FAM	**100.00 (± 0.00)**	98.98 (± 1.18)	99.36 (± 0.51)	94.40 (± 2.38)
E_k-FAM	**100.00 (± 0.00)**	97.96 (± 1.67)	**99.56 (± 0.13)**	95.06 (± 2.60)

TABLE II. CLASSIFICATION ERRORS ON THE TRAINING AND TEST SETS PRODUCED BY KS-FAM, DUAL KS-FAM, AND E-FAMS BASED ON E_k.

Model	Mean CPU time in sec.	
	Dataset I	Dataset II
TE-FAM	0.34	3.7
Θ-FAMs	3.67	5593.08

TABLE III. MEAN CPU TIME FOR TRAINING PHASE FOR TE-FAM $\mathcal{E}_{W,\mathbf{v}}$ AND $\Theta - FAM$ MODELS.

$A, B \in [0,1]^n$ [2]:

$$S_k(A,B) = \begin{cases} 1 & \text{if } A \subseteq B \\ \frac{\sum_{i=1}^n \min\{A_i, B_i\}}{\sum_{i=1}^n A_i} & \text{otherwise} \end{cases} .$$

Table II displays the classification rate produced by each Θ-FAM model in the training and testing phases.

After training the KS-FAM, dual KS-FAM, and E_k-FAM models, we chose the ones that produced the lowest missclassification rate with the lowest variance in the training phase for each dataset. Note that the selected models correspond to the ones whose classification rates for the training data are highlighted in bold face in Table II. Table III shows the mean CPU times required for training TE-FAMs and Θ-FAMs models using a Intel Core i5-3330 CPU with a processing speed of 3.00 GHz.

Table IV exhibits the mean classification rates for testing that were produced by the TE-FAM of type $\mathcal{E}_{W,\mathbf{v}}$, the selected Θ-FAM models [1], and several MLP models [46]. Note that the TE-FAM model yielded a competitive classification accuracy compared to the other classifiers we tested. The Θ-FAM approach together with training algorithm proposed in [1] produced the highest classification rates for both datasets. However, Table III reveals that training the Θ-FAM models using the algorithm of [1] requires a huge amount of CPU time in comparison to the training phase of the TE-FAM model. This observation applies especially to the simulations using the second dataset that comprises 1363 instances. In this case, the algorithm described in [1] has to cope with more than one thousand fundamental memories consisting of 75% of the entire dataset that is divided into 4 folds.

Model	Dataset I	Dataset II
TE-FAM	94.39	92.55
Θ-FAMs	**98.64**	**95.06**
MLP	89.22	93.16
MLP-G	93.45	93.50
MLP-D	92.50	92.05
MLP-GD	96.05	93.55

TABLE IV. MEAN CLASSIFICATION RATE ON THE TEST SET PRODUCED BY TE-FAM $\mathcal{E}_{W,\mathbf{v}}$, SELECTED $\Theta - FAM$ MODELS, AND MLPS MODELS.

VII. CONCLUDING REMARKS

In this article, we presented TE-FAMs, a novel approach to Θ-fuzzy associative memories.The first stage of the TE-FAM traininng algorithm automatically determines the topology of the resulting model by selecting a set of fundamental memories. Note that each hidden node is uniquely defined in terms of a fundamental memory (cf. Definition 3).

In contrast to the Θ-FAM approach presented in [1], the TE-FAM approach allows not only for tuning the weights corresponding to the contributions of the hidden nodes but also for adjusting the contributions of the attributes of the data. The latter is achieved by adjusting parametrized equivalence measures of the form $E_{\mathbf{w}}$ (cf. Equation 5).

Some preliminary experiments in classification also indicate a competitive accuracy. In addition, training the TE-FAMs using the algorithm outlined in Section V could be performed much faster than training Θ-FAMs using the previous algorithm [1].

VIII. ACKNOWLEDGMENT

The authors would like to thank Toni Pimentel and Dr. Fernando Manuel Ramos for providing the time series data used in Section VI.B. This work was partially supported by FAPESP under grants nos. 2009/16284-2 and 2014/04163-4.

REFERENCES

[1] E. Esmi, P. Sussner, H. Bustince, and J. Fernández, "θ-fuzzy associative memories," *IEEE Transactions on Fuzzy Systems*, 2014, accepted for publication, available online.

[2] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Englewood Cliffs, N. J.: Prentice Hall, 1992.

[3] F. Chung and T. Lee, "On fuzzy associative memory with multiple-rule storage capacity," *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 3, pp. 375–384, August 1996.

[4] F. Junbo, J. Fan, and S. Yan, "A learning rule for fuzzy associative memories," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 7, June 1994, pp. 4273 – 4277.

[5] P. Liu, "The fuzzy associative memory of max-min fuzzy neural networks with threshold," *Fuzzy Sets and Systems*, vol. 107, pp. 147–157, 1999.

[6] R. Belohlávek, "Fuzzy logical bidirectional associative memory," *Information Sciences*, vol. 128, no. 1-2, pp. 91–103, 2000.

[7] P. Sussner and M. E. Valle, "Implicative fuzzy associative memories," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 6, pp. 793–807, 2006.

[8] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-4, pp. 4–22, 1987.

[9] N. Ikeda, P. Watta, M. Artiklar, and M. H. Hassoun, "A two-level Hamming network for high performance associative memory," *Neural Networks*, vol. 14, no. 9, pp. 1189 – 1200, 2001.

[10] X. Mu, P. Watta, and M. Hassoun, "A weighted voting model of associative memory," *Neural Networks, IEEE Transactions on*, vol. 18, no. 3, pp. 756–777, 2007.

[11] E. Esmi, P. Sussner, M. Valle, F. Sakuray, and L. Barros, "Fuzzy associative memories based on subsethood and similarity measures with applications to speaker identification," in *Hybrid Artificial Intelligent Systems*, ser. Lecture Notes in Computer Science, E. Corchado et al., Ed. Springer Verlag, Berlin / Heidelberg, 2012, vol. 7209, pp. 479–490.

[12] P. Sussner, E. L. Esmi, I. Villaverde, and M. Graña, "The Kosko subsethood fuzzy associative memory (KS-FAM): Mathematical background and applications in computer vision," *Journal of Mathematical Imaging and Vision*, vol. 42, pp. 134–149, 2012.

[13] M. E. Valle and P. Sussner, "A general framework for fuzzy morphological associative memories," *Fuzzy Sets and Systems*, vol. 159, no. 7, pp. 747–768, 2008.

[14] G. J. F. Banon and J. Barrera, "Decomposition of mappings between complete lattices by mathematical morphology, part 1. general lattices," *Signal Processing*, vol. 30, no. 3, pp. 299–327, Feb. 1993.

[15] H. Heijmans and C. Ronse, "The algebraic basis of mathematical morphology I. dilations and erosions," *Computer Vision, Graphics, and Image Processing*, vol. 50, no. 3, pp. 245 – 295, 1990.

[16] J. Serra, *Image Analysis and Mathematical Morphology, Volume 2: Theoretical Advances*. New York: Academic Press, 1988.

[17] P. Sussner and M. E. Valle, "Classification of fuzzy mathematical morphologies based on concepts of inclusion measure and duality," *Journal of Mathematical Imaging and Vision*, vol. 32, no. 2, pp. 139–159, 2008.

[18] B. Schölkopf and A. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, ser. Adaptive Computation and Machine Learning. Mit Press, 2002.

[19] F. T. Martins-Bedé, L. Godo, S. Sandri, L. V. Dutra, C. C. Freitas, O. S. Carvalho, R. J. Guimarães, and R. S. Amaral, "Classification of Schistosomiasis prevalence using fuzzy case-based reasoning," in *Proceedings of the 10th International Work-Conference on Artificial Neural Networks: Part I: Bio-Inspired Systems: Computational and Ambient Intelligence*, ser. IWANN '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 1053–1060.

[20] I. Drummond, L. Godo, and S. Sandri, "Restoring consistency in systems of fuzzy gradual rules using similarity relations," *Lecture Notes in Computer Science*, vol. 2507, pp. 386–396, 2002.

[21] E. Esmi, P. Sussner, and S. Sandri, "Tunable equivalence fuzzy associative memories," 2014, submitted for publication.

[22] E. Esmi, "Estimation of mappings using θ-fuzzy associative memory for pattern recogniton," Ph.D. dissertation, Unversity of Campinas, 2014, in Portuguese.

[23] J. Fodor and M. Roubens, "Fuzzy preference modeling and multicriteria decision support," in *Theory and Decision Library*. Kluwer Academic Publishers, 1994.

[24] J. Fodor and R. R. Yager, "Fuzzy set-theoretic operators and quantifiers," in *Fundamentals of Fuzzy Sets*, ser. The Handbooks of Fuzzy Sets Series, D. Dubois and H. Prade, Eds. Springer US, 2000, vol. 7, pp. 125–193.

[25] H. Bustince, M. Pagola, and E. Barrenechea, "Construction of fuzzy indices from fuzzy DI-subsethood measures: Application to the global comparison of images," *Information Sciences*, vol. 177, no. 3, pp. 906 – 929, 2007.

[26] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[27] A. Di Nola, S. Sessa, W. Pedrycz, and E. Sanchez, *Fuzzy Relation Equations and Their Applications to Knowledge Engineering*. Norwell, MA: Kluwer Academic Publishers, 1989.

[28] G. Birkhoff, *Lattice Theory*, 3rd ed. Providence: American Mathematical Society, 1993.

[29] G. A. Grätzer, *Lattice Theory: First Concepts and Distributive Lattices*. San Francisco, CA: W. H. Freeman, 1971.

[30] D. Dubois and H. Prade, "A unifying view of comparison indices in a fuzzy set-theoretic framework," in *Recent Developments in Fuzzy Set and Possibility Theory*, R. R. Yager, Ed. Pergamon Press, 1982, pp. 3–13.

[31] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic; Theory and Applications*. Upper Saddle River, N. Y.: Prentice Hall, 1995.

[32] J. Fan and W. Xie, "Some notes on similarity measure and proximity measure," *Fuzzy Sets and Systems*, vol. 101, no. 3, pp. 403 – 412, 1999.

[33] X. Liu, "Entropy, distance measure and similarity measure of fuzzy sets and their relations," *Fuzzy Sets and Systems*, vol. 52, pp. 305–318, December 1992.

[34] H.-Y. Zhang and W.-X. Zhang, "Hybrid monotonic inclusion measure and its use in measuring similarity and distance between fuzzy sets," *Fuzzy Sets and Systems*, vol. 160, no. 1, pp. 107–118, Jan. 2009.

[35] W. Zeng and H. Li, "Inclusion measures, similarity measures, and the fuzziness of fuzzy sets and their relations," *International Journal of Intelligent Systems*, vol. 21, no. 6, pp. 639–653, 2006.

[36] S. Sandri and F. Martins-Bedê, "A method for deriving order compatible fuzzy relations from convex fuzzy partitions," *Fuzzy Sets and Systems*, to appear.

[37] R. P. Lippmann and E. A. Martin, "Neural classifiers useful for speech recognition," *1st International Conference on Neural Networks*, vol. IEEE, pp. 400–417, 1987.

[38] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, and S. García, "KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework." *Multiple-Valued Logic and Soft Computing*, vol. 17, no. 2-3, pp. 255–287, 2011.

[39] A. Asuncion and D. J. Newman, "UCI machine learning repository, University of California, Irvine, School of Information and Computer Sciences," 2007, *http://www. ics. uci. edu/∼mlearn/MLRepository. html*.

[40] T. Coleman and Y. Li, "On the convergence of interior-reflective newton methods for nonlinear minimization subject to bounds," *Mathematical Programming*, vol. 67, no. 1-3, pp. 189–224, 1994.

[41] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," Ithaca, NY, USA, Tech. Rep., 1993.

[42] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York, NY: John Wiley & Sons, 1973.

[43] V. N. Vapnik, *Statistical learning theory*, 1st ed. Wiley, Sep. 1998.

[44] J. F. R. O. Breiman, L. and C. Stone, "Classification and regression trees," 1984.

[45] "Laboratory of remote sensing in agriculture and forestry, Time-series derived from remotely sensed images," *www.dsr.inpe.br/laf/series/en*.

[46] T. Pimentel, F. M. Ramos, and S. Sandri, "Using fuzzy multilayer perceptrons for the classification of time series," in *Fuzzy Logic and Applications*, ser. Lecture Notes in Computer Science, F. Masulli, G. Pasi, and R. Yager, Eds. Springer International Publishing, 2013, vol. 8256, pp. 60–67.