8

Morphological and Certain Fuzzy Morphological Associative Memories for Classification and Prediction

Peter Sussner and Marcos Eduardo Valle

Institute of Mathematics, Statistics, and Scientific Computing, State University of Campinas sussner@ime.unicamp.br, mevalle@ime.unicamp.br

Summary. Morphological associative memories (MAMs) are based on a lattice algebra known as minimax algebra. In previous papers, we gained valuable insight into the storage and recall phases of gray-scale autoassociative memories. This article extends these results to the heteroassociative and to the fuzzy case in view of the fact that a gray-scale MAM model can be converted into a fuzzy MAM model that coincides with the Lukasiewicz IFAM by applying an appropriate threshold. The article includes experimental results concerning applications of MAM and fuzzy MAM models in classification and prediction.

8.1 Introduction

A number of recent approaches to neurocomputing are either explicitly or implicitly rooted in lattice theory [6]. These approaches include fuzzy lattice neural networks [31], morphological neural networks [34, 35], and fuzzy min-max neural networks [38, 39]. The morphological associative memory (MAM) model that we discuss in this article belongs to the class of morphological neural networks (MNNs).

The mathematical foundations of MNNs can be found in mathematical morphology [18] which represents a set theoretic approach to image processing. Mathematical morphology (MM) can be conducted very generally in a complete lattice setting [36]. In mathematical morphology, an operator that commutes with the lattice operation "meet" is called erosion and an operator that commutes with the lattice operation "join" is called dilation. An erosion can be formulated in terms of an inclusion measure and a dilation can be constructed from a given erosion via a relationship of duality. A fuzzification of an inclusion measure can be used to formulate a fuzzy erosion in the complete lattice $[0, 1]^n$ and a fuzzy dilation arises as the dual of fuzzy erosion. Other operators of mathematical morphology include anti-dilation, anti-erosion, opening and closing. The four morphological operators dilation,

P. Sussner and M.E. Valle: *Morphological and Certain Fuzzy Morphological Associative Memories for Classification and Prediction*, Studies in Computational Intelligence (SCI) **67**, 149–171 (2007)
www.springerlink.com

erosion, anti-dilation, and anti-erosion can be considered to be the elementary operators of mathematical morphology [4, 5].

In the context of artificial neural networks, we speak of a MNN if the first step in computing the next state of a neuron is given by one of the four elementary morphological operators. A MNN whose neurons perform operations in the fuzzy domain is called a fuzzy MNN. Applications of morphological and hybrid morphological/linear neural nets include automatic target recognition, land mine detection, handwritten character recognition, and prediction of financial markets [2, 15, 21, 22, 30].

The morphological associative memory model represents one of the first MNN models that appeared in the literature [35]. B. Raducanu, M. Graña *et al.* have applied this type of MNN to the problems of face-localization, self-localization, and hyperspectral image analysis [16, 17, 32]. Although initial research efforts have focused on the binary autoassociative case, the MAM model was proposed from the outset as a heteroassociative memory model for the storage and the recall of real-valued patterns and a number of notable features of autoassociative morphological memories (AMMs) such as optimal absolute storage capacity and one-step convergence have been shown to hold in the general case for real-valued patterns [35]. More importantly, these results remain valid for integer-valued patterns since MAMs can be applied in this setting without any roundoff errors.

In a recent paper, we presented a thorough analysis of gray-scale AMMs [43]. Specifically, we described the fixed points and the basins of attraction of real- and integer-valued AMMs. We also introduced a modified gray-scale AMM model that produces as an output a fixed point which is closest to the input pattern with respect to the Chebyshev distance. This article generalizes some of these results to include the heteroassociative case. In particular, we obtain a theorem that characterizes the output of a MAM for every input pattern. Furthermore, we show that a slightly modified version of this theorem describes the output patterns of a certain fuzzy morphological associative memory (FMAM). We construct this fuzzy model from the MAM model by applying appropriate thresholds and we point out that the new FMAM represents a special case of an implicative fuzzy associative memory (IFAM) [44, 45]. Finally, we undermine our theoretical results by applying the MAM and FMAM models to problems in classification and prediction.

8.2 Some Background Information on Lattice Theory, Mathematical Morphology, and Minimax Algebra

In contrast to traditional semi-linear neural network models, morphological neural networks perform the morphological operations of erosion or dilation at every node. Alternatively, we can describe the operations that are performed in each layer of a morphological neural network in terms of matrix products in

minimax algebra. Minimax algebra is a lattice algebra which originated from problems in operations research and machine scheduling [13, 14].

Lattice theory is concerned with algebraic structures that arise by imposing some type of ordering on a set [6, 18, 36]. A partially ordered X is called a *lattice* if and only if there exists an infimum and a supremum for every finite subset of X. The infimum of $Y \subseteq X$ is denoted by the symbol $\bigwedge Y$. Alternatively, we write $\bigwedge_{j \in J} y_j$ instead of $\bigwedge Y$ if $Y = \{y_j : j \in J\}$ for some index set J. Similar notations are used to denote the supremum of Y. We speak of a *complete lattice* X if every (finite or infinite) subset has an infimum and a supremum in X. From now on, we denote complete lattices by the symbols $\mathbb{L}$ and $\mathbb{M}$.

The elementary operations of mathematical morphology are erosion, dilation, anti-dilation, and anti-erosion [5]. In the general complete lattice setting, an *erosion* is an operator $\varepsilon : \mathbb{L} \to \mathbb{M}$ that commutes with the infimum operation [18, 37]. In other words, the operator ε represents an erosion if and only if the following equality holds for every subset $Y \subseteq \mathbb{L}$:

$$\varepsilon(\bigwedge Y) = \bigwedge_{y \in Y} \varepsilon(y)\,. \tag{8.1}$$

Similarly, an operator $\delta : \mathbb{L} \to \mathbb{M}$ that commutes with the supremum operation is called a *dilation*. In other words, the operator δ represents a dilation if and only if the following equality holds for every subset $Y \subseteq \mathbb{L}$:

$$\delta(\bigvee Y) = \bigvee_{y \in Y} \delta(y)\,. \tag{8.2}$$

Apart from erosions and dilations, we will also consider the elementary operators anti-erosion and anti-dilation that are defined as follows [5, 18]. An operator $\bar{\varepsilon}$ is called an anti-erosion if and only if (8.3) holds for every $Y \subseteq \mathbb{L}$ and an operator $\bar{\delta}$ is called a anti-dilation if and only if (8.4) holds for every subset $Y \subseteq \mathbb{L}$.

$$\bar{\varepsilon}(\bigwedge Y) = \bigvee_{y \in Y} \bar{\varepsilon}(y)\,, \tag{8.3}$$

$$\bar{\delta}(\bigvee Y) = \bigwedge_{y \in Y} \bar{\delta}(y)\,. \tag{8.4}$$

Erosions, dilations, anti-erosions, and anti-dilations exemplify the concept of morphological operator, i.e., an operator that arises in the context of mathematical morphology [18]. Banon and Barrera showed that every mapping f between complete lattices $\mathbb{L}$ and $\mathbb{M}$ can be represented either as the supremum of pair-wise infimums of erosions and anti-dilations or as the infimum of pair-wise supremums of dilations and anti-erosions [5].

In minimax algebra, we define certain algebraic structures called *belts* and *blogs* ("bounded lattice ordered groups"). The set of extended real numbers

$\mathbb{R}_{\pm\infty} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$, exemplifies a belt as well as a blog. Specifically, we have that $(\mathbb{R}_{\pm\infty}, \vee, +)$ and $(\mathbb{R}_{\pm\infty}, \wedge, +')$ represent belts and that $(\mathbb{R}_{\pm\infty}, \vee, \wedge, +, +')$ represents a blog. The operations "$+$" and "$+'$" act like the usual sum operation and are identical on $\mathbb{R}_{\pm\infty}$ with the following exceptions:

$$(-\infty) + (+\infty) = (+\infty) + (-\infty) = -\infty\,, \tag{8.5}$$

$$(-\infty) +' (+\infty) = (+\infty) +' (-\infty) = +\infty\,. \tag{8.6}$$

If $(\mathbb{E}, \oplus, \otimes)$ and $(\mathbb{F}, \oplus', \otimes')$ are belts then a *belt homomorphism* is a function $f : \mathbb{E} \to \mathbb{F}$ that is compatible with the operations. We refer to f as a *belt isomorphism* if f is bijective. In this article, we employ the belt isomorphism of *conjugation*, denoted by a "*" symbol, between the belts $(\mathbb{R}_{\pm\infty}, \vee, +)$ and $(\mathbb{R}_{\pm\infty}, \wedge, +')$. This isomorphism is given as follows.

$$x^* = \begin{cases} -x & \text{if } x \in \mathbb{R}\,, \\ -\infty & \text{if } x = +\infty\,, \\ +\infty & \text{if } x = -\infty\,. \end{cases} \tag{8.7}$$

We say that $(\mathbb{R}_{\pm\infty}, \vee, +)$ is the conjugate of $(\mathbb{R}_{\pm\infty}, \wedge, +')$ or simply that the blog $(\mathbb{R}_{\pm\infty}, \vee, \wedge, +, +')$ is self-conjugate. Note that $(\mathbb{Z}_{\pm\infty}, \vee, \wedge, +, +')$ also represents a self-conjugate blog.

A matrix $A \in \mathbb{R}_{\pm\infty}^{m\times n}$ corresponds to a conjugate matrix $A^* \in \mathbb{R}_{\pm\infty}^{n\times m}$. Each entry $a_{ij}^* = [A^*]_{ij}$ of A^* is given by

$$a_{ij}^* = (a_{ji})^*\,. \tag{8.8}$$

Obviously, $(A^*)^* = A$ for all $A \in \mathbb{R}_{\pm\infty}^{m\times n}$, and thus the isomorphism of conjugation $\mathbb{R}_{\pm\infty}^{m\times n} \to \mathbb{R}_{\pm\infty}^{n\times m}$ is involutive. We say that a matrix $A \in \mathbb{R}_{\pm\infty}^{m\times n}$ is *finite* if every row vector and every column vector has at least one finite entry. In particular, a vector $\mathbf{x} \in \mathbb{R}_{\pm\infty}^{n}$ is finite if and only if $\mathbf{x} \in \mathbb{R}^n$.

The maximum and the minimum of two matrices are performed elementwise. For matrices $A, B \in \mathbb{R}_{\pm\infty}^{m\times n}$, we have,

$$(A \vee B)^* = A^* \wedge B^* \quad \text{and} \quad (A \wedge B)^* = A^* \vee B^*\,. \tag{8.9}$$

There are two types of matrix products with entries in $\mathbb{R}_{\pm\infty}$. For an $m \times p$ matrix A and a $p \times n$ matrix B with entries from $\mathbb{R}_{\pm\infty}$, the matrix $C = A \boxtimes B$, also called the *max product* of A and B, and the matrix $D = A \boxtimes B$, also called the *min product* of A and B, are defined by

$$c_{ij} = \bigvee_{k=1}^{p} (a_{ik} + b_{kj}) \quad \text{and} \quad d_{ij} = \bigwedge_{k=1}^{p} (a_{ik} +' b_{kj})\,. \tag{8.10}$$

Suppose that A is an arbitrary matrix in $\mathbb{R}_{\pm\infty}^{m\times n}$. Consider operators ε_A and δ_A such that $\varepsilon_A(\mathbf{x}) = A \boxtimes \mathbf{x}$ and $\delta_A(\mathbf{x}) = A \boxtimes \mathbf{x}$. Note that ε_A and δ_A associate elements of the complete lattice $\mathbb{R}_{\pm\infty}^{n}$ with elements of the complete lattice

$\mathbb{R}^m_{\pm\infty}$. Clearly, we have that ε_A is an erosion and δ_A is a dilation. Morphological associative memories employ erosions and dilations of this form.

For appropriately sized matrices A and B with entries in $\mathbb{R}_{\pm\infty}$, we obtain the following equalities that will also be useful for describing MAMs.

$$(A \boxtimes B)^* = B^* \boxtimes A^* \quad \text{and} \quad (A \boxtimes B)^* = B^* \boxtimes A^* \,. \tag{8.11}$$

Note that the second halves of (8.9) and (8.11) are the duals of the first halves. As another example for this duality relationship, the reader may find a true statement of minimax algebra as well as the corresponding dual statement in (8.12) and (8.13) [13]. The matrices A, B, C are assumed to be appropriately sized.

$$A \boxtimes (B \wedge C) \leq (A \boxtimes B) \wedge (A \boxtimes C) \qquad \forall\, A, B, C \,. \tag{8.12}$$

$$A \boxtimes (B \vee C) \geq (A \boxtimes B) \vee (A \boxtimes C) \qquad \forall\, A, B, C \,. \tag{8.13}$$

Finally, note that (8.9) and (8.11) imply that every statement in minimax algebra induces a dual statement which simply arises by replacing each "$\wedge$" symbol with a "$\vee$" symbol and vice versa, and by reversing each inequality. Taking advantage of this fact, we only need to present primal statements on MAMs and we may omit the corresponding dual statements.

8.3 A Brief Review of Morphological Associative Memories

Morphological associative memories (MAM) were originally conceived as simple matrix memories endowed with recording recipes that are similar to correlation recording. Suppose that we want to record k vector pairs $(\mathbf{x}^1, \mathbf{y}^1), \ldots, (\mathbf{x}^k, \mathbf{y}^k)$ using a *morphological associative memory* [35]. Let X denote the matrix in $\mathbb{R}^{n \times k}$ whose column vectors are the vectors $\mathbf{x}^\xi \in \mathbb{R}^n$ and let Y denote the matrix in $\mathbb{R}^{m \times k}$ whose column vectors are the vectors $\mathbf{y}^\xi \in \mathbb{R}^m$, where $\xi = 1, \ldots, k$. For simplicity, we write $X = [\mathbf{x}^1, \ldots, \mathbf{x}^k]$ and $Y = [\mathbf{y}^1, \ldots, \mathbf{y}^k]$. The first recording scheme consists in constructing an $m \times n$ matrix W_{XY} as follows:

$$W_{XY} = Y \boxtimes X^* \,. \tag{8.14}$$

In other words, the entry w_{ij} of the matrix W_{XY} is given by the equation

$$w_{ij} = \bigwedge_{\xi=1}^{k} (y_i^\xi - x_j^\xi) \,. \tag{8.15}$$

The second, dual scheme consists in constructing an $m \times n$ matrix M_{XY} of the form $M_{XY} = Y \boxtimes X^*$. Note that the identity $(W_{XY})^* = M_{YX}$ can be deduced from (8.14) and (8.11).

If the matrix W_{XY} receives a vector $\mathbf{x}$ as input then the product $W_{XY} \boxtimes \mathbf{x}$ is formed. Dually, if the matrix M_{XY} receives a vector $\mathbf{x}$ as input then the product $M_{XY} \boxtimes \mathbf{x}$ is formed.

We speak of a binary MAM if $X \in \{0,1\}^{n \times k}$ and $Y \in \{0,1\}^{m \times k}$. In the special case that $X = Y$, we obtain the *autoassociative morphological memories* (AMMs) W_{XX} and M_{XX} [43]. If $X \neq Y$, we have a *heteroassociative morphological associative memory*.

From now on, we will focus on the MAM W_{XY}. Results concerning the dual model M_{XY} can be obtained in a similar fashion by applying the duality relationship given by (8.9) and (8.11).

Example 8.1

$$X = \begin{pmatrix} 0 & 3 & -3 & 1 \\ 8 & 6 & -4 & 5 \\ 6 & 3 & -3 & -4 \end{pmatrix}, \tag{8.16}$$

$$W_{XX} \boxtimes X = X = M_{XX} \boxtimes X\,. \tag{8.17}$$

Note that although the number of stored patterns exceeds the length of the patterns in this example, we have perfect recall for undistorted patterns. We will see in the next section that the absolute storage capacity for autoassociative morphological memories is unlimited, i.e., as many patterns as desired can be stored in an AMM with perfect recall.

Example 8.2 *Let X be as in Example 8.1 and let Y be as follows.*

$$Y = \begin{pmatrix} 3 & 3 & -1 & 1 \\ -1 & -4 & -2 & -5 \\ 0 & 1 & 4 & -1 \\ -1 & -4 & -5 & -5 \end{pmatrix}, \tag{8.18}$$

$$W_{XY} \boxtimes X = \begin{pmatrix} 3 & 3 & -3 & 1 \\ -1 & -4 & -10 & -5 \\ 0 & 1 & -5 & -1 \\ -1 & -4 & -10 & -5 \end{pmatrix}. \tag{8.19}$$

Note the difference between Y and $W_{XY} \boxtimes X$ and the fact that $Y \geq W_{XY} \boxtimes X$. Unlike AMMs, heteroassociative morphological memories (HMMs) are not capable of storing an arbitrary number of patterns.

8.4 Fundamental Results on Gray-scale Autoassociative Morphological Memories

Autoassociative models represent an important special case for associative memories [9, 19, 47]. In this section, we review some fundamental results concerning gray-scale autoassociative morphological memories that have recently

appeared in the literature [43]. We only need to formulate the results for the AMM W_{XX} since similar results for the dual model M_{XX} can be obtained by applying the relationship of duality that was discussed at the end of Sect. 8.2.

We begin by providing a powerful theorem that yields a complete characterization of the fixed points and basins of attraction of gray-scale AMMs. We say that a pattern $\mathbf{x} \in \mathbb{R}^n_{\pm\infty}$ is a *fixed point* of the AMM W_{XX} if and only if $W_{XX} \boxtimes \mathbf{x} = \mathbf{x}$. Similarly, we say that a pattern $\mathbf{x} \in \mathbb{R}^n_{\pm\infty}$ is a *fixed point* of the AMM M_{XX} if and only if $M_{XX} \boxtimes \mathbf{x} = \mathbf{x}$. We denote *the set of finite fixed points* of W_{XX} using the symbol $F(W_{XX})$ and we denote the set of finite fixed points of M_{XX} using the symbol $F(M_{XX})$.

Theorem 8.1 *For $X \in \mathbb{R}^{n\times k}$, the sets $F(W_{XX})$ and $F(M_{XX})$ coincide. If $\mathcal{F}$ denotes this set then $\mathcal{F}$ consists exactly of the following expressions:*

$$\bigvee_{i=1}^{n} \bigwedge_{\xi=1}^{k} (a_i^\xi + \mathbf{x}^\xi)\,, \quad \text{where} \quad a_i^\xi \in \mathbb{R}\,. \tag{8.20}$$

Alternatively, the set $\mathcal{F}$ can be characterized as the set of all expressions of the form

$$\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi)\,, \quad \text{where} \quad c_j^\xi \in \mathbb{R} \quad \text{and} \quad r \in \mathbb{N}\,. \tag{8.21}$$

Moreover, given an arbitrary pattern $\mathbf{x} \in \mathbb{R}^n$, we have

$$W_{XX} \boxtimes \mathbf{x} = \hat{\mathbf{x}} \text{ and } M_{XX} \boxtimes \mathbf{x} = \check{\mathbf{x}}\,, \tag{8.22}$$

where $\hat{\mathbf{x}}$ is the supremum of $\mathbf{x}$ in $\mathcal{F}$ and where $\check{\mathbf{x}}$ is the infimum of $\mathbf{x}$ in $\mathcal{F}$.

Corollary 8.1 *If $\mathcal{F}$ denotes $F(W_{XX}) = F(M_{XX})$ where $X \in \mathbb{R}^{n\times k}$ then we have*

$$\mathcal{F} = \{ \bigwedge_{j=1}^{n} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi) \,:\, c_j^\xi \in \mathbb{R} \}\,. \tag{8.23}$$

The proof of Theorem 8.1 involves some theorems on eigenvectors and eigenvalues in minimax algebra [43]. Another, different proof of the first statement of this theorem was independently presented in [33].

Theorem 8.1 has several important consequences, most notably the unlimited absolute storage capacity and one-step convergence of AMMs. These facts are formally expressed in the following corollaries.

Corollary 8.2 *For all $X \in \mathbb{R}^{n\times k}$, the fixed points of W_{XX} include the patterns $\mathbf{x}^1, \ldots, \mathbf{x}^k$.*

Corollary 8.3 *Let $X \in \mathbb{R}^{n\times k}$. The set of finite fixed points of W_{XX} consists of all $W_{XX} \boxtimes \mathbf{x}$ such that $\mathbf{x} \in \mathbb{R}^n$. Moreover, if $\mathbf{x}$ is attracted to $\mathbf{x}^\xi$ for some $\xi \in \{1, \ldots k\}$ under an application of W_{XX} then $\mathbf{x}$ is an eroded version of $\mathbf{x}^\xi$, i.e., $\mathbf{x} \leq \mathbf{x}^\xi$.*

Theorem 8.1 also induces necessary and sufficient conditions for the perfect recall of an original pattern $\mathbf{x}^\gamma$ [43]. These conditions are formulated in Theorem 8.2.

Theorem 8.2 *Let $X \in \mathbb{R}^{n\times k}$ and let $\mathbf{x} \in \mathbb{R}^n$. The equality $W_{XX} \boxtimes \mathbf{x} = \mathbf{x}^\gamma$ holds if and only if $\mathbf{x} \leq \mathbf{x}^\gamma$ and there is no "linear combination" $\mathbf{l} = \bigvee_{\xi=1}^{k}(c_\xi + \mathbf{x}^\xi) \neq \mathbf{x}^\gamma$ such that $\mathbf{x} \leq \mathbf{l} \leq \mathbf{x}^\gamma$.*

Example 8.3 *Figure 8.1 depicts four images of size 64×64 with 256 gray levels (these images represent downsized versions of images contained in the database of the Computer Vision Group, University of Granada, Spain). For each of these image, we generated a vector $\mathbf{x}^\xi$ of length 4096. We synthesized the weight matrices W_{XX} and $M_{XX} = -W_{XX}^t$ of size 4096×4096, applied them to the original patterns $\mathbf{x}^\xi$, and we confirmed that perfect recall was achieved as we had pointed out in Corollary 8.2. We also stored the vectors $\mathbf{x}^\xi$, $\xi = 1, \ldots, 4$, using the* optimal linear associative memory *(OLAM) [23],* kernel associative memory *(KAM) [47] the* generalized BSB *model of Costantini* et al. *[12], and the* complex-valued Hopfield net *of Müezzinoğlu* et al. *[28].*

Example 8.4 *In this experiment, we probed the associative memory models under consideration with incomplete patterns which arose from leaving away substantial parts of the original images. The outcome of this experiment is visualized in Fig. 8.2. Table 8.1 lists the resulting NMSEs produced by the morphological memory W_{XX}, the OLAM, and the generalized BSB model of Costantini* et al. *for each partial image. Since the KAM model performed very poorly in this experiment, we refrained from displaying the output of the KAM in Fig. 8.2.*

We would like to clarify that we did not conduct this experiment using the complex-valued Hopfield net for the following reasons. Due to computational limitations, the complex-valued Hopfield net can only store small segments of the images. In this experiment, we may have an input segment that contains no information at all, making it impossible to recover the desired image segment.

Fig. 8.1. Original images that were used in constructing the memories W_{XX} and M_{XX}. Presenting the corresponding patterns as inputs to either one of W_{XX} or M_{XX} results in perfect recall

Fig. 8.2. The images in the top row represent severely incomplete versions of original face images. The following rows show - from top to bottom - the corresponding recalled patterns using the morphological memory W_{XX}, the OLAM, and the generalized BSB model

Table 8.1. NMSEs produced by AM models in applications to incomplete patterns of Fig. 8.4

	AMM W_{XX}	OLAM	KAM	Generalized BSB
Tree	0.0017	0.2302	1.0000	0
Lena	0.0137	0.4588	1.0000	0.0449
Cameraman	0.0088	0.6017	1.0000	0.1721
Church	0.0030	0.7448	1.0000	0.2332

Example 8.5 *Theorem 8.1 and the dual version of Theorem 8.2 indicate that the AMM M_{XX} exhibits tolerance with respect to dilative noise. In order to exemplify this type of noise tolerance, we added the absolute value of gaussian noise with zero mean and with variance* 0.1 *to the original patterns $\mathbf{x}^\xi$, $\xi = 1, \ldots, 4$. We compared $\mathbf{x}^\xi$ with the patterns that were retrieved by the M_{XX} memory and the other associative memory models. Table 8.2 displays the resulting NMSEs for each model in* 100 *experiments for each pattern $\mathbf{x}^\xi$, $\xi = 1, \ldots, 4$. Figure 8.3 provides for a visual interpretation of this simulation.*

Table 8.2. NMSEs produced by AM models in applications to patterns that were corrupted by adding the absolute value of gaussian noise with zero mean and variance 0.1

	M_{XX}	OLAM	KAM	Gen. BSB	Compl. Hopf.
Dil. Gauss.	0.0247	0.4185	0.9950	0.4538	0.5460

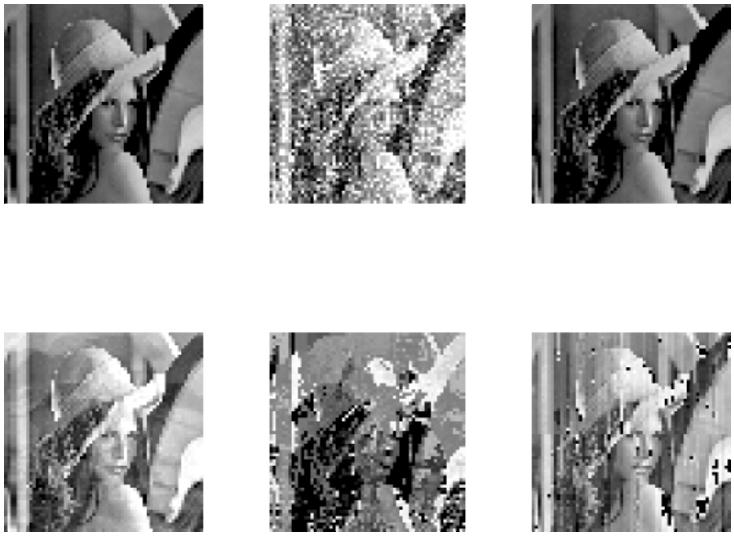

Fig. 8.3. The images in the top row display the original Lena image, a corrupted image that was generated by adding dilative gaussian noise, and the output of the morphological memory M_{XX}. The images in the bottom row show the corresponding recalled patterns using - from left to right - the OLAM, the generalized BSB, and the complex-valued Hopfield model

Since the KAM model performed very poorly in this experiment, we refrained from displaying the output of the KAM in Fig. 8.3.

Theorems 8.1 and 8.2 imply that the AMMs W_{XX} and M_{XX} are not suited for dealing with arbitrary noise, i.e., noise that is neither (mostly) erosive nor (mostly) dilative. To overcome these limitations of the original AMM models, we introduced modified versions of W_{XX} and M_{XX} that we denoted using the symbols $W_{XX} + \nu$ and $M_{XX} + \mu$. The AMMs $W_{XX} + \nu$ and $M_{XX} + \mu$ exhibit a much better tolerance with respect to arbitrary noise compared to W_{XX} and M_{XX} while maintaining the properties of optimal absolute storage capacity and one-step convergence. For further details, we refer the reader to a recent journal article [43].

8.4.1 Applications of gray-scale autoassociative morphological memories to classification problems

Autoassociative memory models such as the morphological models W_{XX} and M_{XX} can be applied to solve multi-class classification problems. Suppose that X^j represents the matrix that consists of all training patterns belonging to class j. For a given test pattern $\mathbf{x}$, we compute the Chebyshev distance $\zeta(\mathbf{x}, \mathbf{x}^{(j)})$ where $\mathbf{x}^{(j)}$ denotes $W_{X^j X^j} \boxtimes \mathbf{x}$. The smallest error $\zeta(\mathbf{x}, \mathbf{x}^{(j)})$ indicates the class that corresponds to $\mathbf{x}$. Obviously, the same principle of classification can be applied to other AMM models. For instance, we can employ autoassociators such as the OLAM and the KAM together with the Euclidean distance.

Example 8.6 *Let us consider the image segmentation problem that is available from the UCI Repository of Machine Learning Databases [1]. In this problem, we have 19 continuous attributes concerning a* 3×3 *region of a hand-segmented image. The instances were drawn randomly from a database of 7 outdoors images, namely, brickface, sky, foliage, cement, window, path, and grass. The data set contains 30 instances per class for training and 300 instances per class for testing. The data was standardized before processed.*

Table 8.3 displays the errors of classification that we obtained using the autoassociative morphological memories W_{XX} *and* M_{XX}*, the OLAM, and the KAM model. The classifiers based on* W_{XX} *and* M_{XX} *yield the same result. Table 8.3 also includes the image segmentation results obtained by a* fuzzy lattice neural network *(FLNN) [31] and a* support vector machine *(SVM) with gaussian kernel and one-against-one method. Note that the AMMs outperformed the other classifiers except for the SVM model which succeeded in providing the correct classification for* 4 *more patterns than the AMM models* W_{XX} *and* M_{XX}*.*

The implementation of the SVM model that we used in this experiment is available on the Internet [7]. We chose to adopt the default parameters. Finally, recall that a particular FLNN model or FLR classifier depends on the choice of a positive valuation function [3, 20]. We employed the linear positive valuation function $v(x) = (x - x_{min})/(x_{max} - x_{min})$*, where* $x \in [x_{min}, x_{max}]$*. We also conducted experiments considering the sigmoid valuation function* $v(x) = 1/(1 - \exp[-\lambda(x - x_{med})])$*, where* $x_{med} = (x_{max} - x_{min})/2$*. The error of classification obtained for* $\lambda = 1$ *and* $\lambda = 0.1$ *were* 12.14% *and* 12.52%*, respectively.*

Table 8.3. Results of the image segmentation problem

Classifier	Error Rate	Classifier	Error Rate
AMM W_{XX}	11.05%	AMM M_{XX}	11.05%
SVM	10.86%	FLNN	12.00%
KAM	82.52%	OLAM	82.95%

Table 8.4. Results of the glass classification problem

Classifier	Error Rate	Classifier	Error Rate
W_{XX}	34.8 ± 4.1	M_{XX}	34.8 ± 4.1
KAM	37.9 ± 4.2	KAA-2	37.4 ± 5.4
MLP	56.9 ± 6.1	SVM	42.2 ± 5.9

Example 8.7 *Let us consider the Glass Recognition problem, another classification problem that can be found in the UCI Repository of Machine Learning Databases [1]. The data set consists of six types of glass. Each type has* 70*,* 17*,* 76*,* 13*,* 9*, or* 27 *instances. The goal is to determine the glass type from nine attributes.*

Zhang et al. *have considered this problem in a recent paper [48]. Several classifiers such as multi-layer perceptrons and support vector machines were tested using two-fold cross-validation. The data were normalized to the range* $[-1, 1]$ *in order to remove the scale effect, and each network was fine tuned. Table 8.4 shows the results of the experiment. The acronym KAA-2 denotes an extension of the KAM that was introduced by H. Zhang* et al. *The error rates concerning the KAM, the KAA-2, the MLP, and the SVM model were taken from [48].*

Table 8.4 also displays the results obtained via applications of the morphological autoassociative memories. Note that the AMM based models outperformed the other classifiers including the SVM model. Moreover, an application of the AMM model does not require any fine tuning of the network.

8.5 Heteroassociative and Fuzzy MAMs

Heteroassociative morphological memories (HMMs) naturally extend the autoassociative models that we reviewed in the previous section. Heteroassociative morphological memories have proved to be useful in several applications [16, 17, 32]. Additional motivation for discussing HMMs can be drawn from the fact that the following theorems on HMMs have some important consequences for the fuzzy domain.

In contrast to binary HMMs [41], gray-scale HMMs have yet to be studied extensively. In fact, only two theorems on gray-scale HMMs concerned with conditions for perfect recall and tolerance to noise are known [35]. Unfortunately, these conditions are rather complicated and hard to understand. Therefore, we consider it timely to present new results which offer considerable insight into the functionality of HMMs. To this end, we will generalize the fundamental results on gray-scale AMMs that we presented in the previous section.

First, let us introduce a few pertinent notations. The symbol $O(W_{XY})$ denotes the set of all $W_{XY} \boxtimes \mathbf{x}$ such that $\mathbf{x} \in \mathbb{R}^n$. Similarly, the symbol $O(M_{XY})$ denotes the set of all $M_{XY} \boxtimes \mathbf{x}$ such that $\mathbf{x} \in \mathbb{R}^n$. Theorem 8.1 and

Corollary 8.3 imply that the set $O(W_{XX})$ consists of the expressions given by (8.20) or by (8.21). The first two statements of Theorem 8.3 generalize this result to include the heteroassociative case.

Theorem 8.3 *For $X \in \mathbb{R}^{n\times k}$ and $Y \in \mathbb{R}^{m\times k}$, the sets $O(W_{XY})$ and $O(M_{XY})$ coincide. If $\mathcal{O}$ denotes this set then $\mathcal{O}$ consists exactly of the following expressions:*

$$\bigvee_{i=1}^{n} \bigwedge_{\xi=1}^{k} (a_i^\xi + \mathbf{y}^\xi), \quad \text{where} \quad a_i^\xi \in \mathbb{R}. \tag{8.24}$$

Alternatively, the set $\mathcal{O}$ can be characterized as the set of the following expressions:

$$\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi), \quad \text{where} \quad c_j^\xi \in \mathbb{R} \quad \text{and} \quad r \in \mathbb{N}. \tag{8.25}$$

Moreover, for arbitrary $\mathbf{x} \in \mathbb{R}^n$, the pattern $W_{XY} \boxtimes \mathbf{x}$ equals the smallest expression given by (8.24) such that $\bigvee_{i=1}^{n} \bigwedge_{\xi=1}^{k} (a_i^\xi + \mathbf{x}^\xi)$ is the supremum of $\mathbf{x}$ in $\mathcal{F}$. In addition, the pattern $W_{XY} \boxtimes \mathbf{x}$ equals the smallest expression given by (8.25) such that $\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi)$ is the supremum of $\mathbf{x}$ in $\mathcal{F}$.

Proof. Let $\mathbf{x} \in \mathbb{R}^n$ be an arbitrary input pattern. Consider the following matrix $Z \in \mathbb{R}^{p\times k}$ and the following vector $\mathbf{z} \in \mathbb{R}^p_{\pm\infty}$, where $p = n + m$.

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix} \text{ and } \mathbf{z} = \begin{pmatrix} \mathbf{x} \\ -\boldsymbol{\infty} \end{pmatrix}. \tag{8.26}$$

Here $-\boldsymbol{\infty}$ denotes the constant vector $(-\infty, \ldots, -\infty)^t$ of length m.

Note that $W_{ZZ} \in \mathbb{R}^{p\times p}$ can be written in block matrix form as follows.

$$W_{ZZ} = \begin{pmatrix} W_{XX} & W_{YX} \\ W_{XY} & W_{YY} \end{pmatrix}. \tag{8.27}$$

Computing the max product $W_{ZZ} \boxtimes \mathbf{z}$ yields

$$W_{ZZ} \boxtimes \mathbf{z} = \begin{pmatrix} W_{XX} \boxtimes \mathbf{x} \vee W_{YX} \boxtimes (-\boldsymbol{\infty}) \\ W_{XY} \boxtimes \mathbf{x} \vee W_{YY} \boxtimes (-\boldsymbol{\infty}) \end{pmatrix} = \begin{pmatrix} W_{XX} \boxtimes \mathbf{x} \\ W_{XY} \boxtimes \mathbf{x} \end{pmatrix}. \tag{8.28}$$

The resulting vector is finite since W_{XX}, W_{XY}, and $\mathbf{x}$ are finite. Moreover, the vector $W_{ZZ} \boxtimes \mathbf{z}$ represents a fixed point of W_{ZZ} because $W_{ZZ} \boxtimes W_{ZZ} = W_{ZZ}$ and because the max product is associative [14]. Identity $W_{ZZ} \boxtimes W_{ZZ} = W_{ZZ}$ follows from Theorem 11 of [43] and the fact that W_{ZZ} has a zero diagonal [40].

Let $\mathbf{f}$ denote $W_{ZZ} \boxtimes \mathbf{z} \in \mathcal{F}_Z = F(W_{ZZ}) = F(M_{ZZ})$. By Theorem 8.1, the elements of $\mathcal{F}_Z$ are described by (8.20) and (8.21) (where $\mathbf{z}^\xi$ replaces $\mathbf{x}^\xi$). Comparing these equations with (8.28) reveals that $W_{XY} \boxtimes \mathbf{x}$ can be expressed in terms of (8.24) and (8.25).

Regarding the proof of the last two statements, we observe that $\mathbf{f}$ represents the supremum of $\mathbf{f}$ in $\mathcal{F}_Z$. As mentioned before, the set $\mathcal{F}_Z$ consists of the expressions given by (8.20) and (8.21) with $\mathbf{z}^\xi$ replacing $\mathbf{x}^\xi$. Recall that $\mathbf{f}$ can also be written in the form $(W_{XX} \boxtimes \mathbf{x}, W_{XY} \boxtimes \mathbf{y})^t$. Therefore, we have

$$\mathbf{f} = \bigvee_{i=1}^{n} \bigwedge_{\xi=1}^{k} (a_i^\xi + \mathbf{z}^\xi) = \bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{z}^\xi)\,, \tag{8.29}$$

where $\bigvee_{i=1}^{n} \bigwedge_{\xi=1}^{k} (a_i^\xi + \mathbf{x}^\xi) = \bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi)$ is the supremum of $\mathbf{x}$ in $\mathcal{F} = F(W_{XX}) = F(M_{XX})$. Therefore, $W_{XY} \boxtimes \mathbf{x}$ is as stated in the theorem.

Theorem 8.4 *If $\mathcal{O}$ denotes $O(W_{XY}) = O(M_{XY})$ where $X \in \mathbb{R}^{n\times k}$ and $Y \in \mathbb{R}^{m\times k}$ then we have*

$$\mathcal{O} = \{ \bigwedge_{j=1}^{n} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi) \,:\, c_j^\xi \in \mathbb{R} \}\,. \tag{8.30}$$

For arbitrary $\mathbf{x} \in \mathbb{R}^n$, the pattern $W_{XY} \boxtimes \mathbf{x}$ equals the smallest expression given by (8.30) such that $\bigwedge_{j=1}^{n} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi)$ is the supremum of $\mathbf{x}$ in $\mathcal{F}$.

Proof. The proof of this theorem employs Corollary 8.1 and resembles the proof of Theorem 8.3.

Corollary 8.4 *For $X \in \mathbb{R}^{n\times k}$, $Y \in \mathbb{R}^{m\times k}$, and let $\mathbf{x} \in \mathbb{R}^n$ be such that $W_{XX} \boxtimes \mathbf{x} = \mathbf{x}^\gamma$. We have $W_{XY} \boxtimes \mathbf{x} = \mathbf{y}^\gamma$ if and only if the following implication holds for all $c^\xi \in \mathbb{R}$ where $\xi = 1, \ldots, k$.*

$$\mathbf{x}^\gamma \leq \bigvee_{\xi=1}^{k} c^\xi + \mathbf{x}^\xi \;\Rightarrow\; \mathbf{y}^\gamma \leq \bigvee_{\xi=1}^{k} c^\xi + \mathbf{y}^\xi\,. \tag{8.31}$$

Proof. Let us again consider the matrix Z and the vector $\mathbf{z}$ represented in (8.26). Let $\mathbf{x}$ be as stated above.

Suppose that $W_{XY} \boxtimes \mathbf{x} = \mathbf{y}^\gamma$ which – in view of (8.28) – implies that $W_{ZZ} \boxtimes \mathbf{z} = \mathbf{z}^\gamma$. In other words, we have $\hat{\mathbf{z}} = \mathbf{z}^\gamma$, where $\hat{\mathbf{z}}$ denotes the supremum of $\mathbf{z}$ in $\mathcal{F}_Z$. By Corollary 8.2, $\mathcal{F}_Z$ consists of all patterns $\bigwedge_{j=1}^{n} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{z}^\xi)$ such that $c_j^\xi \in \mathbb{R}$.

Consider arbitrary scalars c^ξ, where $\xi = 1, \ldots, k$, such that the left hand side of (8.31) holds. If $\mathbf{u}$ denotes $\bigvee_{\xi=1}^{k} (c^\xi + \mathbf{z}^\xi)$ then $\mathbf{u} \in \mathcal{F}_Z$ represents an upper bound of $\mathbf{z} = (\mathbf{x}, -\infty)^t$. We obtain the following relationships.

$$\begin{pmatrix} \mathbf{x}^\gamma \\ \mathbf{y}^\gamma \end{pmatrix} = \mathbf{z}^\gamma = W_{ZZ} \boxtimes \mathbf{z} \leq W_{ZZ} \boxtimes \mathbf{u} = \mathbf{u} = \bigvee_{\xi=1}^{k} \left[c^\xi + \begin{pmatrix} \mathbf{x}^\xi \\ \mathbf{y}^\xi \end{pmatrix} \right]. \tag{8.32}$$

In particular, we recognize that $\mathbf{y}^\gamma \leq \bigvee_{\xi=1}^{k} (c^\xi + \mathbf{y}^\xi)$ which shows that the (8.31) is a necessary condition for $W_{XY} \boxtimes \mathbf{x} = \mathbf{y}^\gamma$.

Let us now prove that (8.31) implies $W_{XY} \boxtimes \mathbf{x} = \mathbf{y}^\gamma$. Let $\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi)$ denote the smallest pattern such that $\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi) = \hat{\mathbf{x}}$, the supremum of $\mathbf{x}$ in $\mathcal{F} = F(W_{XX})$.

By Theorem 8.3, we have $W_{XY} \boxtimes \mathbf{x} = \bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi)$. Theorem 8.1 states that $\hat{\mathbf{x}} = W_{XX} \boxtimes \mathbf{x}$ which equals $\mathbf{x}^\gamma$ by assumption. Thus, we obtain the following identity

$$\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi) = \mathbf{x}^\gamma \,. \tag{8.33}$$

Equation (8.33) implies that the patterns $\bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi)$ are bounded from below by $\mathbf{x}^\gamma$ for all $j = 1, \ldots, n$. From (8.31) we infer that

$$\bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi) \geq \mathbf{y}^\gamma \quad \forall\, j = 1, \ldots, r \,. \tag{8.34}$$

$$\Leftrightarrow \bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi) \geq \mathbf{y}^\gamma \tag{8.35}$$

$$\Leftrightarrow \bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi) \wedge \mathbf{y}^\gamma = \mathbf{y}^\gamma \,. \tag{8.36}$$

Obviously, we have $\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi) \wedge \mathbf{y}^\gamma \leq \bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi)$ and $\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi) \wedge \mathbf{x}^\gamma = \hat{\mathbf{x}} \wedge \mathbf{x}^\gamma = \hat{\mathbf{x}}$ since $\hat{\mathbf{x}} = \bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi) = W_{XX} \boxtimes \mathbf{x} = \mathbf{x}^\gamma$. Therefore, $\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi) \wedge \mathbf{y}^\gamma$ and $\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi)$ coincide since the former is bounded from above by the latter pattern and since the latter pattern $\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi)$ represents the smallest pattern such that $\bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{x}^\xi) = \hat{\mathbf{x}}$. In view of (8.36), we are able to finish the proof of the theorem as follows:

$$W_{XY} \boxtimes \mathbf{x} = \bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi) = \bigwedge_{j=1}^{r} \bigvee_{\xi=1}^{k} (c_j^\xi + \mathbf{y}^\xi) \wedge \mathbf{y}^\gamma = \mathbf{y}^\gamma \,. \tag{8.37}$$

Theorem 8.3 and Corollary 8.4 completely characterize the output of HMMs for any input pattern $\mathbf{x}$. The rest of the article is concerned with the implications of these results with respect to fuzzy morphological associative memory (FMAM) models.

Note that the fuzzy domain $[0,1]^n$ represents a complete lattice. An erosion $\varepsilon : [0,1]^n \to [0,1]^m$ is called a *fuzzy erosion*. In a similar vein, we speak of *fuzzy dilations*, *fuzzy anti-dilations*, and *fuzzy anti-erosions*. In *fuzzy mathematical morphology*, a fuzzy erosion can be defined in terms of a fuzzy inclusion measure and a fuzzy dilation can be defined in terms of a fuzzy intersection

measure [29, 42]. We consider fuzzy morphological neural networks to be models of artificial neural networks that calculate an elementary operation of mathematical morphology such as fuzzy erosion or fuzzy dilation at each node. FMAMs belong to this class of models.

Let us consider a particular FMAM model. Let $X \in [0,1]^{n\times k}$, $Y \in [0,1]^{m\times k}$, and let W_{XY} be defined as before. The symbols $\mathbf{0}_{m\times n}$ and $\mathbf{0}_m$ stand for the zero matrix of size $m \times n$ and the zero vector of length m. If W denotes the $m\times n$ matrix $W_{XY} \wedge \mathbf{0}_{m\times n}$ then the *Lukasiewicz FMAM* is defined in terms of the following relationship between an input pattern $\mathbf{x} \in [0,1]^n$ and an output pattern $\mathbf{y} \in [0,1]^m$.

$$\mathbf{y} = (W \boxtimes \mathbf{x}) \vee \mathbf{0}_m \,. \tag{8.38}$$

The Lukasiewicz FMAM performs a fuzzy dilation $[0,1]^m \to [0,1]$ at every node. This fuzzy dilation can also be expressed in terms of a supremum of fuzzy Lukasiewicz conjunctions, hence the name Lukasiewicz FMAM. Furthermore, it can be shown that (8.38) yields the Lukasiewicz IFAM model [44, 45].

Recall that $W_{XY} = \bigvee_{\xi=1}^{k} [\mathbf{y}^\xi \boxtimes (\mathbf{x}^\xi)^*]$. Hence, the probability that $W_{XY} \in [-1,0]^{m\times n}$ increases as more and more fundamental memories $(\mathbf{x}^\xi, \mathbf{y}^\xi)$ are stored in the network. If $W_{XY} \in [-1,0]^{m\times n}$ then W_{XY} equals W, the weight matrix of the Lukasiewicz FMAM, and we may apply Theorem 8.3 to Lukasiewicz FMAMs which leads to the following corollary.

Corollary 8.5 *Let $X \in [0,1]^{n\times k}$ and $Y \in [0,1]^{m\times k}$ such that $W_{XY} \in [-1,0]^{m\times n}$. For an arbitrary input pattern $\mathbf{x} \in [0,1]^n$, an application of the Lukasiewicz FMAM produces the maximum of $\mathbf{0}_m$ and the smallest expression given by (8.24) such that $\bigvee_{i=1}^{n} \bigwedge_{\xi=1}^{k} (a_i^\xi + \mathbf{x}^\xi)$ is the supremum of $\mathbf{x}$ in $\mathcal{F}$.*

Corollary 8.6 *Let $X \in [0,1]^{n\times k}$ and $Y \in [0,1]^{m\times k}$ be such that $W_{XY} \in [-1,0]^{m\times n}$. Suppose that $\mathbf{x} \in [0,1]^n$ satisfies $W_{XX} \boxtimes \mathbf{x} = \mathbf{x}^\gamma$. We have $(W \boxtimes \mathbf{x}) \vee \mathbf{0}_m = \mathbf{y}^\gamma$ if the following implication holds for all $c^\xi \in \mathbb{R}$ where $\xi = 1, \ldots, k$.*

$$\mathbf{x}^\gamma \leq \bigvee_{\xi=1}^{k} c^\xi + \mathbf{x}^\xi \;\Rightarrow\; \mathbf{y}^\gamma \leq \bigvee_{\xi=1}^{k} c^\xi + \mathbf{y}^\xi \,. \tag{8.39}$$

Proof. Let W_{XY} and $\mathbf{x} \in [0,1]^n$ be as stated above. By Corollary 8.4, the assumption that (8.31) holds for all $c^\xi \in \mathbb{R}$, where $\xi = 1, \ldots, k$, implies that $W_{XY} \boxtimes \mathbf{x} = \mathbf{y}^\gamma$ which belongs to $[0,1]^m$. Therefore, we have $(W_{XY} \boxtimes \mathbf{x}) \vee \mathbf{0}_m = \mathbf{y}^\gamma$. Finally, note that W_{XY} can be replaced by W since $W_{XY} \in [-1,0]^{m\times n}$.

8.5.1 Applications of FMAMs in prediction

Fuzzy associative memories such as the FMAM can be used to implement mappings of fuzzy rules. In this case, a set of rules in the form of human-like

IF-THEN conditional statements are stored. In this subsection, we present two applications of the FMAM model to the problem of forecasting time-series.

Example 8.8 *Let us consider the problem presented in [10] and discussed latter in [44, 45]. This problem consists of assessing the manpower requirement in steel manufacturing industry in the state of West Bengal, India. Initially, we have five linguistic values representing concepts such as "the requirement in manpower is* large*". A set of fuzzy conditional statements such as "If the manpower requirement of year n is* large*, then that of year n+1 is* very large*" is obtained from the past values. We converted these conditional statements into the set of input-output pairs that are represented in Table 8.5 and we generated the following matrix W_{XY}.*

$$W_{XY} = \begin{bmatrix} -0.5 & -1.0 & -1.0 & -1.0 & -1.0 \\ 0 & -0.5 & -1.0 & -1.0 & -1.0 \\ -0.5 & -0.5 & -0.5 & -1.0 & -1.0 \\ -1.0 & -1.0 & -1.0 & -.5 & -.5 \\ -1.0 & -1.0 & -1.0 & -0.5 & 0 \end{bmatrix}. \tag{8.40}$$

Since $W_{XY} \in [-1, 0]^{5\times 5}$, the matrices W and W_{XY} coincide. Therefore, Corollaries 8.5 and 8.6 can be applied. For example, if $\mathbf{x} = [1, 0.6, 0, 0, 0]^T$ then the supremum of $\mathbf{x}$ in $\mathcal{F}$ is $W_{XX} \boxtimes \mathbf{x} = [1, 0.6, 0.1, 0, 0]^T = \mathbf{x}^1 \vee [(0.1 + \mathbf{x}^1) \wedge \mathbf{x}^2]$. Note that the latter expression represents a meet of joins that can be easily brought into the form $\bigvee_{i=1}^{5} \bigwedge_{\xi=1}^{5} (a_i^\xi + \mathbf{x}^\xi)$ by adding some coefficients $a_i^\xi \geq 1$. By Corollary 8.5, we have $(W_{XY} \boxtimes \mathbf{x}) \vee \mathbf{0}_5 \leq \mathbf{y}^1 \vee [(0.1 + \mathbf{y}^1) \wedge \mathbf{y}^2] \vee \mathbf{0}_5 = [0.5, 1, 0.5, 0, 0]^T$. In fact, we calculate $(W_{XY} \boxtimes \mathbf{x}) \vee \mathbf{0}_5 = [0.5, 1, 0.5, 0, 0]^T$.

Now, let us consider Corollary 8.6. On one hand, Corollary 8.2 implies that $W_{XX} \boxtimes \mathbf{x}^2 = \mathbf{x}^2$. On the other hand, an application of the Lukasiewicz FMAM to $\mathbf{x}^2$ yields $[0, 0.5, 0.5, 0, 0]^T \neq \mathbf{y}2$. Therefore, there are some c^ξ such that the implication in (8.39) does not hold. We have, for example, that $\mathbf{x}^2 \leq \mathbf{x}^3$ whereas $\mathbf{y}^2 \not\leq \mathbf{y}^3$.

Table 8.5. Set of input and output pairs used in the forecasting application

ξ	$\mathbf{x}^\xi$	$\mathbf{y}^\xi$
1	$[1.0, 0.5, 0, 0, 0]^T$	$[0.5, 1.0, 0.5, 0, 0]^T$
2	$[0.5, 1.0, 0.5, 0, 0]^T$	$[0.5, 1.0, 0.5, 0, 0]^T$
3	$[0.5, 1.0, 0.5, 0, 0]^T$	$[0, 0.5, 1.0, 0.5, 0]^T$
4	$[0, 0.5, 1.0, 0.5, 0]^T$	$[0.5, 1.0, 0.5, 0, 0]^T$
5	$[0, 0.5, 1.0, 0.5, 0]^T$	$[0, 0.5, 1.0, 0.5, 0]^T$
6	$[0, 0.5, 1.0, 0.5, 0]^T$	$[0, 0, 0.5, 1.0, 0.5]^T$
7	$[0, 0, 0.5, 1.0, 0.5]^T$	$[0, 0, 0.5, 1.0, 0.5]^T$
8	$[0, 0, 0.5, 1.0, 0.5]^T$	$[0, 0, 0, 0.5, 1.0]^T$
9	$[0, 0, 0, 0.5, 1.0]^T$	$[0, 0, 0, 0.5, 1.0]^T$

Table 8.6. Average errors in forecasting manpower

Method	Average Error
Lukasiewicz FMAM	2.29%
Kosko's FAM	2.67%
Lukasiewicz GFAM	2.67%
Gödel IFAM	2.73%
Max-min FAM with threshold	2.73%
Goguen IFAM	2.99%
AM of Wang and Lu	2.99%
ARIMA2	5.48%
ARIMA1	9.79%

If $\mathbf{x}$ *represents a fuzzy set corresponding to the manpower of year* n *then (8.38) can be used to forecast the manpower of year* $n+1$ *by means of the Lukasiewicz FMAM. Specifically, a defuzzification of the output pattern* $\mathbf{y}$ *according to the rule described in [10] yields the prediction for year* $n+1$*. Table 8.6 displays the average errors in the predictions that were obtained by means of the Lukasiewicz FMAM and several other methods that can be found in the literature [11, 24, 25, 44, 45, 46]. Figure 8.4 plots the manpower data of the years 1984 through 1995. The actual values are compared to the predictions obtained by some of these methods. Note that the Lukasiewicz FMAM outperformed the other models.*

Example 8.9 *In this example, we applied the Lukasiewicz FMAM to the problem of forecasting the average monthly streamflow of a large hydroelectric plant called Furnas, that is located in southeastern Brazil. This problem was previously discussed in [26, 27].*

Note that the seasonality of the monthly streamflow suggests the use of 12 different models, one for each month of the year. Let s_ξ*, for* $\xi = 1, \ldots, q$*, be samples of a seasonal streamflow time series. The goal is to estimate the value of* s_γ *from a subsequence of* $(s_1, s_2, \ldots, s_{\gamma-1})$*. Here, we employ subsequences that correspond to a vector of the form*

$$\mathbf{p}^\gamma = (s_{\gamma-h}, \ldots, s_{\gamma-1})^T, \tag{8.41}$$

where $h \in \{1, 2, \ldots, \gamma - 1\}$*. In this experiment, our FMAM based model only uses a fixed number of three antecedents. For example, the values of January, February, and March were considered for predicting the streamflow of April.*

The uncertainty that is inherent in hydrological data suggests the use of fuzzy sets to model the streamflow samples. For $\xi < \gamma$*, a fuzzification of* $\mathbf{p}^\xi$ *and* s^ξ *using Gaussian membership functions yields fuzzy sets* $\mathbf{x}^\xi : \mathcal{U} \to [0, 1]$ *and* $\mathbf{y}^\xi : \mathcal{V} \to [0, 1]$ *respectively, where* $\mathcal{U}$ *and* $\mathcal{V}$ *represent finite universes of discourse. A subset* S *of the resulting input-output pairs* $\{(\mathbf{x}^\xi, \mathbf{y}^\xi), \xi < q\}$ *is implicitly stored in the Lukasiewicz FMAM (we only construct the parts of the*

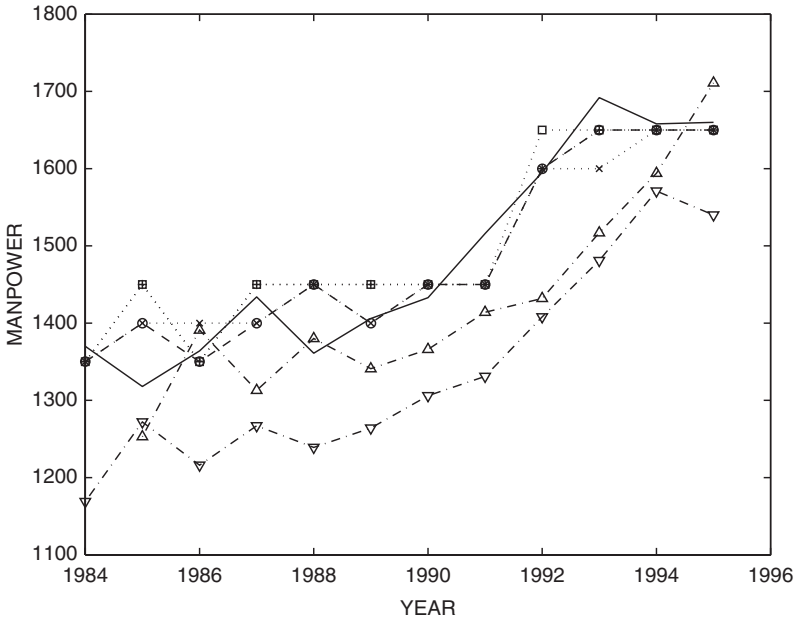

Fig. 8.4. Predictions of manpower. The continuous line represents the actual manpower. The dashed line marked by '$\circ$' corresponds to the Lukasiewicz FMAM model, the dotted line marked by '$\times$' corresponds to Kosko's FAM model and the Lukasiewicz Generalized FAM, the dotted line marked by '$+$' corresponds to Max-min FAM with threshold and Gödel IFAM, and the dotted line marked by '$\square$' corresponds to the Associative Memory model of Wang and Lu and the Goguen IFAM. The lines marked by '$\triangle$' and '∇' represent ARIMA1 and ARIMA2

weight matrix that are actually used in the recall phase). We employed the subtractive clustering method *to determine the set S [8]. Feeding the pattern* $\mathbf{x}^\gamma$ *into the FMAM model, we retrieved the corresponding output pattern* $\mathbf{y}^\gamma$*. For computational reasons,* $\mathbf{x}^\gamma$ *is modeled as a discrete Dirac-*δ *(impulse) function. A defuzzification of* $\mathbf{y}^\gamma$ *using the mean of maximum yields* s_γ*.*

Figure 8.5 shows the forecasted streamflows estimated by the prediction model based on the FMAM for the Furnas reservoir from 1991 to 1998. Table 8.7 compares the errors that were generated by the FMAM model and several other models [26, 27]. In contrast to the FMAM-based model, the MLP, NFN, and FPM-PRP models were initialized by optimizing the number of the parameters for each monthly prediction. For example, the MLP considers 4 *antecedents to predict the streamflow of January and* 3 *antecedents to predict the streamflow for February. Moreover, the FPM-PRP model also takes into account slope information which requires some additional "fine tuning". We experimentally determined a variable number of parameters (including slopes) for the FMAM model such that* $\text{MSE} = 0.88 \times 10^5$*,* $\text{MAE} = 157$*, and* $\text{MPE} = 15$*.*

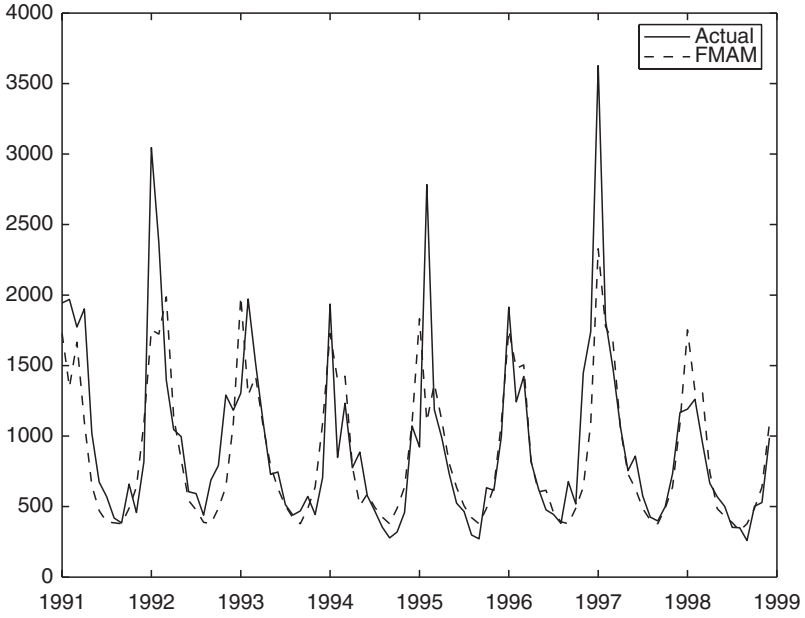

Fig. 8.5. The streamflow prediction for the Furnas reservoir from 1991 to 1998. The continuous line corresponds to the actual values and the dashed line corresponds to the predicted values

Table 8.7. Mean square, mean absolute, and mean relative percentage errors produced by the prediction models

Methods	MSE ($\times 10^5$)	MAE (m^3/s)	MPE (%)
FMAM	1.42	226	22
PARMA	1.85	280	28
MLP	1.82	271	30
NFN	1.73	234	20
FPM-PRP	1.20	200	18

References

1. UCI Repository of Machine Learning Databases. University of California, Irvine, Dept. of Information and Computer Sciences. Available at: `http://www.ics.uci.edu/~mlearn/MLRepository.html`
2. Araújo RA, Madeiro F, Sousa RP, Pessoa LFC, Ferreira TAE (2006) An evolutionary morphological approach for financial time series forecasting. In: Proc IEEE Congress on Evolutionary Computation pp 2467–2474
3. Athanasiadis I, Kaburlasos VG (2006) Air quality assessment using fuzzy lattice reasoning (FLR). In: Proc IEEE International Conference on Fuzzy Systems pp 231–236

4. Banon GJF, Barrera J (1991) Minimal representation for translation invariant set mappings by mathematical morphology. SIAM J Appl Math 51(6): 1782–1798
5. Banon GJF, Barrera J (1993) Decomposition of mappings between complete lattices by mathematical morphology: part I - general lattices. Signal Processing 30:299–327
6. Birkhoff G (1993) Lattice Theory. American Mathematical Society, 3 ed.
7. Canu S, Grandvalet Y, Guigue V, Rakotomamonjy A (2005) SVM and Kernel Methods Matlab Toolbox. Perception Systèmes et Information, INSA de Rouen, Rouen, France
8. Chiu S (1994) Fuzzy model identification based on cluster estimation. Journal of Intelligent and Fuzzy Systems 2(3)
9. Chiueh T, Goodman R (1991) Recurrent correlation associative memories. IEEE Trans on Neural Networks 2(2):275–284
10. Choudhury J, Sarkar B, Mukherjee S (2002) Forecasting of engineering manpower through fuzzy associative memory neural network with arima: a comparative study. Neurocomputing 47(1-2):241–257
11. Chung F, Lee T (1996) On fuzzy associative memory with multiple-rule storage capacity. IEEE Transactions on Fuzzy Systems 4(3):375–384
12. Costantini G, Casali D, Perfetti R (2003). Neural associative memory storing gray-coded gray-scale images. IEEE Transactions on Neural Networks 14(3):703–707
13. Cuninghame-Green R (1979) Minimax Algebra. Lecture Notes in Economics and Mathematical Systems 166. Springer-Verlag, New York
14. Cuninghame-Green R (1995) Minimax algebra and applications. In: Hawkes P (ed) Advances in Imaging and Electron Physics 90:1–121. Academic Press, New York
15. Gader PD, Khabou M, Koldobsky A (2000) Morphological regularization neural networks. Pattern Recognition (special issue on Mathematical Morphology and Its Applications) 33(6):935–945
16. Graña M, Gallego J, Torrealdea FJ, D'Anjou A (2003a) On the application of associative morphological memories to hyperspectral image analysis. Lecture Notes in Computer Science 2687:567–574
17. Graña M, Sussner P, Ritter G (2003b) Innovative applications of associative morphological memories. Mathware and Softcomputing 10(3):155–168
18. Heijmans H (1994) Morphological Image Operators. Academic Press, New York
19. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. In: Proc National Academy of Sciences 79:2554–2558
20. Kaburlasos V, Petridis V (2000) Fuzzy lattice neurocomputing (FLN) models. Neural Networks 13(10):1145–1170
21. Khabou M, Gader P (2000) Automatic target detection using entropy optimized shared-weight neural networks. IEEE Transactions on Neural Networks 11(1):186–193
22. Khabou M, Gader PD, Keller JM (2000) LADAR target detection using morphological shared-weight neural networks. Machine Vision and Applications 11(6):300–305
23. Kohonen T (1984) Self-Organization and Associative Memory. Springer-Verlag
24. Kosko B (1992) Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence. Prentice Hall, Englewood Cliffs, New Jersey

25. Liu P (1999) The fuzzy associative memory of max-min fuzzy neural networks with threshold. Fuzzy Sets and Systems 107(2):147–157
26. Magalhães M (2004) Redes neurais, metodologias de agrupamento e combinação de previsores aplicados a previsão de vazões naturais. Master's Thesis, State University of Campinas
27. Magalhães M, Ballini R, Gonçalves R, Gomide F (2004) Predictive fuzzy clustering model for natural streamflow forecasting. In: Proc IEEE International Conference on Fuzzy Systems pp 390–394
28. Müezzinoğlu M, Güzeliş C, Zurada J (2003) A new design method for the complex-valued multistate Hopfield associative memory. IEEE Trans on Neural Networks 14(4):891–899
29. Nachtegael M, Kerre E (2001) Connections between binary, gray-scale and fuzzy mathematical morphologies. Fuzzy Sets and Systems 124(1):73–85
30. Pessoa L, Maragos P (2000) Neural networks with hybrid morphological/rank/linear nodes: a unifying framework with applications to handwritten character recognition. Pattern Recognition 33(6):945–960
31. Petridis V, Kaburlasos V (1998) Fuzzy lattice neural network (FLNN): a hybrid model for learning. IEEE Transactions on Neural Networks 9(5):877–890
32. Raducanu B, Graña M, Albizuri XF (2003) Morphological scale spaces and associative morphological memories: results on robustness and practical applications. Journal of Mathematical Imaging and Vision 19(2):113–131
33. Ritter GX, Gader PD (2006) Fixed points of lattice transforms and lattice associative memories. In: Hawkes P (ed) Advances in Imaging and Electron Physics 144:165–242. Elsevier, Amsterdam, The Netherlands
34. Ritter GX, Sussner P (1996) An introduction to morphological neural networks. In: Proc 13th International Conference on Pattern Recognition pp 709–717
35. Ritter GX, Sussner P, de Leon JLD (1998) Morphological associative memories. IEEE Transactions on Neural Networks 9(2):281–293
36. Ronse C (1990) Why mathematical morphology needs complete lattices. Signal Processing 21(2):129–154
37. Serra J (1988) Image Analysis and Mathematical Morphology, vol. 2: Theoretical Advances. Academic Press, New York
38. Simpson P (1992) Fuzzy min-max neural networks – part 1: classification. IEEE Transactions on Neural Networks 3(5):776–786
39. Simpson P (1993) Fuzzy min-max neural networks – part 2: clustering. IEEE Transactions on Neural Networks 1(1):32–35
40. Sussner P (2000) Fixed points of autoassociative morphological memories. In: Proc International Joint Conference on Neural Networks pp 611–616
41. Sussner P (2005) New results on binary auto- and heteroassociative morphological memories. In: Proc International Joint Conference on Neural Networks pp 1199–1204
42. Sussner P, Valle ME (2005) A brief account of the relations between gray-scale mathematical morphologies. In: Proc Brazilian Symposium on Computer Graphics and Image Processing pp 79–86
43. Sussner P, Valle ME (2006a) Grayscale morphological associative memories. IEEE Transactions on Neural Networks 17(3):559–570
44. Sussner P, Valle ME (2006b) Implicative fuzzy associative memories. IEEE Transactions on Fuzzy Systems 14(6):793–807

45. Valle ME, Sussner P, Gomide F (2004) Introduction to implicative fuzzy associative memories. In: Proc IEEE International Joint Conference on Neural Networks pp 925–931
46. Wang S, Lu H (2004) On new fuzzy morphological associative memories. IEEE Transactions on Fuzzy Systems 12(3):316–323
47. Zhang BL, Zhang H, Ge SS (2004) Face recognition by applying wavelet subband representation and kernel associative memory. IEEE Transactions on Neural Networks 15(1):166–177
48. Zhang H, Huang W, Huang Z, Zhang B (2005) A kernel autoassociator approach to pattern classification. IEEE Transactions on Systems, Man and Cybernetics, Part B 35(3):593–606