An Overview of Morphological Neural Networks

Peter Sussner

Mathematical Imaging and Computational Intelligence Group
Department of Applied Mathematics, IMECC
University of Campinas

16th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems
Organization of this talk

1. Introduction
2. Basic Concepts of MM and Lattice Theory
3. General Concepts of MNNs
4. Some Examples of MNNs
5. Hybrid Models
6. Conclusions and Perspectives for the Future
Introduction

Some Remarks on Morphological Neural Networks (MNNs)

Introductory Remarks:

- MNNs incorporate concepts of mathematical morphology (MM) into artificial neural networks;
- The aggregation functions of MNNs compute (elementary) operations of MM;
- MNNs are approaches towards lattice computing, specifically towards computational intelligence based on lattice theory;
- First attempts at formulating MNNs for special applications appeared in the early 1990’s;
- The first general framework for computing with MNNs was presented in 1996;
- An artificial intelligence approach towards morphological image processing can be found in M. Schmitt’s doctoral thesis (1989).
Purpose of this Talk

This keynote reviews the following issues related to MNNs:

- Origins and basic concepts;
- Foundations in mathematical morphology (MM);
- Lattice-theoretical background of MM for MNNs;
- Types of MNNs and examples;
- Typical applications.
Mathematical Morphology (MM) is a theory for the processing and analysis of images using structuring elements (SEs).

Applications of MM include:

1. noise removal;
2. skeletonizing;
3. edge detection;
4. automatic target recognition;
5. image segmentation;
6. image restauration.
Two Perspectives on MM

MM from two different (but not mutually exclusive) points of view:

- MM in the geometrical or topological sense: employs SEs as well as inclusion and intersection measures;
- MM in the algebraic sense: usually defined in a complete lattice setting, recently extended to complete semilattices.

MM in the geometrical or topological sense

- **Erosion**: yields the (crisp or fuzzy) degree of inclusion of the translated SE at every pixel;
- **Dilation**: yields the (crisp or fuzzy) degree of intersection of the image with the (reflected and) translated SE at every pixel.
Some Basic Notions of Binary and Fuzzy MM

Let $X$ be either $\mathbb{R}^d$ or $\mathbb{Z}^d$ and let $\mathcal{F}(X) = [0, 1]^X$ denote the class of fuzzy sets over the universe $X$.

- **Binary MM** deals with images $A$ and SEs $S$ such that $A, S \subseteq X$. The translation of $S$ by $x \in X$ is $S_x = \{s + x : s \in S\}$.

- **Fuzzy MM** deals with images $a$ and SEs $s$ such that $a, s \in \mathcal{F}(X)$. The translation of $s_x$ is given by $s_x(y) = s(y - x)$ for all $y \in X$.

The concepts of inclusion and intersection of sets lie at the root of MM. In particular, we have

- the **binary erosion** of image $A$ by $S$ yields the set of points for which a translation of $S$ is contained in the input image.

- The **fuzzy erosion** of an image $a \in \mathcal{F}(X)$ by $s \in \mathcal{F}(X)$ at $x$ is the degree of inclusion of $s_x$ in $a$. 

Illustration of Binary Erosion

Figure:

Binary image $A$, SE $S$, and binary erosion of $A$ by $S$. 
Illustration of Fuzzy Erosion

Figure:

Fuzzy image \(a\), SE \(s\), and fuzzy erosion of \(a\) by \(s\).
Some Remarks on LT and MM

- LT has found applications in many areas such as:
  - mathematical morphology;
  - fuzzy set theory;
  - computational intelligence;
  - automated decision making;
  - formal concept analysis.

- Complete lattices are generally accepted as the appropriate mathematical framework for MM;

- Morphological operators in the complete lattice framework always come in dual pairs, e.g., erosion/dilation;

- Morphological neural networks (MNNs) are lattice computing approaches towards computational intelligence.
Complete Lattices

- A lattice is a partially ordered set $\mathbb{L} \neq \emptyset$ such that every finite, non-empty subset $Y \subseteq \mathbb{L}$ has an infimum, denoted by $\bigwedge Y$ and a supremum, denoted by $\bigvee Y$ in $\mathbb{L}$. If $\mathbb{L}$ is totally ordered (i.e., $x \leq y$ or $y \leq x \ \forall x, y \in \mathbb{L}$) then $\mathbb{L}$ is a chain.

- A lattice is complete if $\bigwedge Y$ and $\bigvee Y$ exist in $\mathbb{L}$ for every $Y \subseteq \mathbb{L}$.

- Examples of complete lattices include $\mathbb{R}_{\pm \infty} = \mathbb{R} \cup \{+\infty, -\infty\}$, $\mathbb{Z}_{\pm \infty} = \mathbb{Z} \cup \{+\infty, -\infty\}$, and $[0, 1]$.

- The class of $\mathbb{L}$-sets over the universe $\mathbb{X}$, denoted using the symbol $\mathcal{F}_{\mathbb{L}}(\mathbb{X})$, is the class of all functions $\mathbb{X} \rightarrow \mathbb{L}$. If $\mathbb{L}$ is a complete lattice then $\mathcal{F}_{\mathbb{L}}(\mathbb{X})$ and $\mathbb{L}^n$ with the component-wise partial orderings are complete lattices as well.

- Important special cases of complete lattices are given by $\mathbb{R}^n_{\pm \infty} = (\mathbb{R}_{\pm \infty})^n$, $\mathbb{Z}^n_{\pm \infty} = (\mathbb{Z}_{\pm \infty})^n$, and $\mathcal{F}(\mathbb{X})$, the class of fuzzy sets over the universe $\mathbb{X}$. 
Elementary Operators of MM on Complete Lattices

- From now on, the symbols $L$ and $M$ denote complete lattices.
- Consider operators $\varepsilon, \delta, \bar{\delta}, \bar{\varepsilon}: L \to M$.
  - $\varepsilon$ is called an (algebraic) erosion if
    $$\varepsilon \left( \bigwedge Y \right) = \bigwedge_{y \in Y} \varepsilon(y), \quad \forall Y \subseteq L.$$  
  - $\delta$ is called an (algebraic) dilation if
    $$\delta \left( \bigvee Y \right) = \bigvee_{y \in Y} \delta(y), \quad \forall Y \subseteq L.$$  
  - $\bar{\delta}$ and $\bar{\varepsilon}$ are resp. called anti-dilation and anti-erosion if
    $$\bar{\delta} \left( \bigvee Y \right) = \bigwedge_{y \in Y} \bar{\delta}(y), \quad \bar{\varepsilon} \left( \bigwedge Y \right) = \bigvee_{y \in Y} \bar{\varepsilon}(y) \quad \forall Y \subseteq L.$$
Adjuncions and (Algebraic) Dilations and Erosions

Definition
Consider $\delta : L \to M$ and $\varepsilon : M \to L$. The pair $(\varepsilon, \delta)$ is called an adjunction from $L$ to $M$, in other words $\varepsilon$ and $\delta$ are adjoint if

$\delta(x) \leq y \Leftrightarrow x \leq \varepsilon(y) \ \forall x \in L, y \in M$.

Proposition
Let $\delta : L \to M$ and $\varepsilon : M \to L$.

- If $(\varepsilon, \delta)$ is an adjunction then $\delta$ is a dilation and $\varepsilon$ is an erosion.
- For every dilation $\delta$ there is a unique erosion $\varepsilon$ such that $(\varepsilon, \delta)$ is an adjunction.
- For every erosion $\varepsilon$ there is a unique dilation $\delta$ such that $(\varepsilon, \delta)$ is an adjunction.
Decomposition of Operators between Complete Lattices

**Theorem**

For every \( \psi : \mathbb{L} \rightarrow \mathbb{M} \) there exist an index set \( I \) and erosions \( \varepsilon^i \) as well as anti-dilations \( \bar{\delta}^i \) such that

\[
\psi = \bigvee_{i \in I} (\varepsilon^i \land \bar{\delta}^i). \tag{1}
\]

Similarly, there exist an index set \( J \) and dilations \( \delta^j \) as well as anti-erosions \( \bar{\varepsilon}^j \) such that

\[
\psi = \bigwedge_{j \in J} (\delta^j \lor \bar{\varepsilon}^j). \tag{2}
\]
Lattice-Ordered Groups

- A lattice that also represents a group such that every group translation $x \mapsto a + x + b$ is isotone is called an $l$-group.
- An $l$-group $F$ such that $F$ is a conditionally complete lattice is called a conditionally complete $l$-group.
- A complete lattice $G$ such that $F = G \setminus \{\bigvee G, \bigwedge G\}$ forms an $l$-group is called a complete $l$-group extension.

Examples

1. $\mathbb{R}^n$ and $\mathbb{Z}^n$ are conditionally complete $l$-groups $\forall n \in \mathbb{N}$.
2. $\mathbb{R}_{\pm\infty} = \mathbb{R} \cup \{-\infty, +\infty\}$ and $\mathbb{Z}_{\pm\infty} = \mathbb{Z} \cup \{-\infty, +\infty\}$ represent complete $l$-group extensions. In fact they are totally ordered group extensions.
Max Product, Min Product, and Conjugate

Let $G$ be a complete $l$-group extension and $F = G \setminus \{\vee G, \wedge G\}$. Let $A \in F^{m \times n}$ and $B \in G^{n \times p}$.

- $M = A \boxdot B$ - max product of $A$ and $B$: $m_{ij} = \vee_{k=1}^{n}(a_{ik} + b_{kj})$.
- $W = A \boxcap B$ - min product of $A$ and $B$: $w_{ij} = \wedge_{k=1}^{n}(a_{ik} + b_{kj})$
- $A^*$ - conjugate of $A$: $A^* = -A^T$

An (algebraic) erosion is given by

$$\varepsilon_A : G^n \rightarrow G^m$$

$$x \mapsto A \boxcap x$$

An (algebraic) dilation is given by

$$\delta_A : G^n \rightarrow G^m$$

$$x \mapsto A \boxdot x$$
Max-C Product and Min-D Product

Let $A \in [0, 1]^{m \times n}$ and $B \in [0, 1]^{n \times p}$.

- Consider a fuzzy conjunction $C$ and a fuzzy disjunction $D$. Recall that $C$, $D$ are increasing operators $[0, 1]^2 \rightarrow [0, 1]$ s.t.:
  - $C(0, 0) = C(0, 1) = C(1, 0) = D(0, 0) = 0$;
  - $C(1, 1) = D(0, 1) = D(1, 0) = D(1, 1) = 1$;

- $M = A \circ B$ - max-C product of $A$ and $B$: $m_{ij} = \bigvee_{k=1}^{n} C(a_{ik}, b_{kj})$.
- $W = A \bullet B$ - min-D product of $A$ and $B$: $w_{ij} = \bigwedge_{k=1}^{n} D(a_{ik}, b_{kj})$.

If $C(a, \cdot)$ is a dilation for every $a \in [0, 1]$ then we have an (algebraic) dilation

$$\delta^{\mathcal{F}}_A : [0, 1]^n \rightarrow [0, 1]^m$$

$$x \mapsto A \circ x$$

If $D(a, \cdot)$ is an erosion for every $a \in [0, 1]$ then we have an (algebraic) erosion

$$\varepsilon^{\mathcal{F}}_A : [0, 1]^n \rightarrow [0, 1]^m$$

$$x \mapsto A \bullet x$$
Other Operators Based on Max-C and Min-D Products

Theorem

The following operator $D_X$ represents a dilation for every $X \in [0, 1]^{n \times p}$ if and only if $C(\cdot, x)$ is a dilation for every $x \in [0, 1]$.

$$D_X : [0, 1]^{m \times n} \rightarrow [0, 1]^{m \times p}$$

$$A \mapsto A \circ X$$

The following operator $E_X$ represents a erosion for every $X \in [0, 1]^{n \times p}$ if and only if $D(\cdot, x)$ is an erosion for every $x \in [0, 1]$.

$$E_X : [0, 1]^{m \times n} \rightarrow [0, 1]^{m \times p}$$

$$A \mapsto A \bullet X$$
Some Types of Morphological Neurons

Additive Max and Min Neurons

Let $\mathcal{G}$ be a complete $l$-group extension and $\mathcal{F} = \mathcal{G} \setminus \{\vee \mathcal{G}, \wedge \mathcal{G}\}$. For an input vector $\mathbf{x} \in \mathcal{G}^n$ and a vector of synaptic weights $\mathbf{w} \in \mathcal{F}^n$, the output $y \in \mathcal{G}$ is computed as follows:

- Additive max neuron: $y = \bigvee_{j=1}^{n} (w_j + x_j) = \mathbf{w}^T \bigotimes \mathbf{x}$;
- Additive min neuron: $y = \bigwedge_{j=1}^{n} (w_j + x_j) = \mathbf{w}^T \bigodot \mathbf{x}$.

Max-$C$ and Min-$D$ Neurons:

Let $C$ and $D$ be resp. a fuzzy conjunction and disjunction. For an input vector $\mathbf{x} \in [0, 1]^n$ and a vector of synaptic weights $\mathbf{w} \in [0, 1]^n$, the output $y \in [0, 1]$ is computed as follows:

- Max-$C$ neuron: $y = \bigvee_{j=1}^{n} C (w_j, x_j) = \mathbf{w}^T \circ \mathbf{x}$;
- Min-$D$ neuron: $y = \bigwedge_{j=1}^{n} D (w_j, x_j) = \mathbf{w}^T \bullet \mathbf{x}$.
Morphological Neurons in Complete Lattices

Observations

- Additive max and min neurons as well as max-$\mathcal{C}$ and min-$\mathcal{D}$ neurons for continuous $\mathcal{C}$ and $\mathcal{D}$ yield elementary morphological operators between complete lattices.

- The aggregation functions of max-$\mathcal{C}$ and min-$\mathcal{D}$ neurons can also be viewed as fuzzy dilations and erosions of $\mathbf{x}$ by SEs (determined by $\mathbf{w}$).
Another Type of a Morphological Neuron

Let \( S \) be a fuzzy subsethood measure. Given a weight vector \( \mathbf{w} \in [0, 1]^n \) and input \( \mathbf{x} \in [0, 1]^n \), compute \( y \in [0, 1] \) as follows:

\[
y = S(\mathbf{w}, \mathbf{x}) .
\]

We have a fuzzy erosion of \( \mathbf{x} \) by the SE \( \mathbf{w} \) but not an algebraic erosion. Recall that \( S : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1] \) satisfies:

1. If \( A \subseteq B \) then \( S(A, B) = 1 \);
2. \( S(X, \emptyset) = 0 \);
3. If \( A \subseteq B \subseteq C \) then \( S(C, A) \leq S(B, A) \) and \( S(C, A) \leq S(C, B) \).

If \( t \) is a \( t \)-norm and \( S \) is such that \( S(A, A^c) = 0 \) for all \( A \subseteq X \) then a similarity measure \( SM : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1] \) is given by

\[
SM(A, B) = S(A, B) \ t \ S(B, A) \ \forall A, B \in \mathcal{F}(X) . \quad (3)
\]
Examples of MNNs

Morphological Associative Memories (MAMs)

Comments on Original MAM Models

1. The MAMs $W_{XY}$ and $M_{XY}$ employ “$\Box$” and “$\vee$”.

2. $W_{XY}$, $M_{XY}$ yield functions $\mathbb{R}^n_{\pm\infty} \rightarrow \mathbb{R}^m_{\pm\infty}$. Replacing $\mathbb{R}^n_{\pm\infty}$ by any complete $l$-group extension $G$, let $F = G \setminus \{\lor G, \land G\}$.

Definitions of $W_{XY}$ and $M_{XY}$

For $X = [x^1, \ldots, x^k] \in F^{n \times k}$ and $Y = [y^1, \ldots, y^k] \in F^{m \times k}$, let

$$W_{XY} = Y \Box X^*, \quad M_{XY} = Y \lor X^* \in F^{m \times n}.$$  \hspace{1cm} (4)

Given $x \in G^n$, the outputs of $W_{XY}$ and $M_{XY}$ are resp. calculated in terms of a dilation and an erosion:

$$y = W_{XY} \Box x, \quad z = M_{XY} \Box x.$$  

If $X = Y$, we speak of an auto-associative MAM (AMM).
Properties of AMMs

Some Advantages:
1. Unlimited absolute storage capacity;
2. One-step convergence if employed with feedback.

Some Disadvantages:
1. Both $W_{XX}$ and $M_{XX}$ are unable to deal with arbitrary noise;
2. Large number of spurious memories.

Alternatives
1. Use modified MAMs $W_{XX} + \nu$ or $M_{XX} + \mu$;
2. Substitute the complete lattice $(\mathbb{G}^n, \leq)$ with a cisl of the form $(\mathbb{F}^n, \preceq_r)$ and define new AM in this setting.
Some Applications of MAMs

The MAM models $W_{XY}$ and $M_{XY}$ have been applied in diverse areas such as:

- hyperspectral image analysis;
- color image segmentation;
- image compression;
- robot vision;
- face localization;
- a variety of other pattern recognition problems.
An Application of MAMs in Classification

- Let $X^j$ denote the matrix consisting of all class $j$ training patterns.
- Given a test pattern $x$, compute the Chebyshev distances $\zeta(x, x^{(j)})$ where $x^{(j)}$ denotes $W_{X^jX^j} \Box x$ (or $M_{X^jX^j} \Box x$).
- The smallest error $\zeta(x, x^{(j)})$ indicates the class corresponding to $x$.
- The same principle of classification can be applied to other auto-associative models together with the Euclidean distance.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Error Rate</th>
<th>Classifier</th>
<th>Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_{XX}$</td>
<td>34.8 ± 4.1</td>
<td>$M_{XX}$</td>
<td>34.8 ± 4.1</td>
</tr>
<tr>
<td>KAM</td>
<td>37.9 ± 4.2</td>
<td>KAA-2</td>
<td>37.4 ± 5.4</td>
</tr>
<tr>
<td>MLP</td>
<td>56.9 ± 6.1</td>
<td>SVM</td>
<td>42.2 ± 5.9</td>
</tr>
</tbody>
</table>

Table: Results of the glass classification problem.
Examples of MNNs

Fuzzy Morphological Associative Memories (FMAMs)

Definition of Max-C FMAM:

Let $X = [x^1, \ldots, x^p] \in [0, 1]^{n \times p}$ and $Y = [y^1, \ldots, y^p] \in [0, 1]^{m \times p}$. A max-C FMAM model $\mathcal{W}$ is given by

$$y = \mathcal{W}(x) = W \circ x,$$

where $W \in [0, 1]^{m \times n}$, $x \in [0, 1]^n$, and $y \in [0, 1]^m$. Ideally, we have

- $\mathcal{W}(x^\xi) = y^\xi \ \forall \xi = 1, \ldots, p$;
- $\mathcal{W}(\tilde{x}^\xi) = y^\xi$ for noisy or incomplete versions $\tilde{x}^\xi$ of $x^\xi$ ($\xi = 1, \ldots, p$).

Examples of Max-C FMAMs:

- Kosko’s max-min and max-product FAMs;
- the generalized FAM of Chung & Lee;
- impliciative FAMs (IFAMs).
Adjunction-Based Learning for Max-C FMAMs

Synthesis of Weight Matrix $W$ for Max-C FMAM $\mathcal{W}$:

- Let $C(\cdot, x)$ be a dilation $\forall x \in [0, 1]$;
- Consider the dilation $D_X : [0, 1]^{m \times n} \rightarrow [0, 1]^{m \times p}$ given by $D_X(A) = A \circ X$ $\forall A \in [0, 1]^{m \times n}$;
- Let $\mathcal{E}^D_X : [0, 1]^{m \times p} \rightarrow [0, 1]^{m \times n}$ be the unique erosion that forms an adjunction with $D_X$;
- Define $W = \mathcal{E}^D_X(Y)$. 
Adjunction-Based Learning for Min-D FMAMs

Definition of Min-D FMAM:
Let $X = [x^1, \ldots x^p] \in [0, 1]^{n \times p}$ and $Y = [y^1, \ldots y^p] \in [0, 1]^{m \times p}$. A min-D FMAM model $\mathcal{M}$ is given by

$$y = \mathcal{M}(x) = M \cdot x,$$

where $W \in [0, 1]^{m \times n}$, $x \in [0, 1]^n$, and $y \in [0, 1]^m$.

Synthesis of Weight Matrix $M$ for Min-D FMAM $\mathcal{M}$:
- Let $D(\cdot, x)$ be an erosion $\forall x \in [0, 1]$;
- Consider the erosion $\mathcal{E}_X : [0, 1]^{m \times n} \rightarrow [0, 1]^{m \times p}$ given by $\mathcal{E}_X(A) = A \cdot X \ \forall A \in [0, 1]^{m \times n}$;
- Let $\mathcal{D}_{X}^\mathcal{E} : [0, 1]^{m \times p} \rightarrow [0, 1]^{m \times n}$ be the unique dilaton that forms an adjunction with $\mathcal{E}_X$;
- Define $M = \mathcal{D}_{X}^\mathcal{E}(Y)$.
Observations on FMAMs

Links to Classical MAMs

- The MAMs $W_{XY}$ and $M_{XY}$ are closely related to the Lukasiewicz FMAM and the Lukasiewicz dual FMAMs;
- We can show that $W_{XY}$ and $M_{XY}$ use adjunction-based learning;
- FMAMs and classical MAMs have similar properties, e.g., optimal absolute storage capacity and one-step convergence if $X = Y$.

Applications of FMAMs

FMAMs can be used to implement fuzzy rule-based systems for applications such as prediction and control.

Possible Extensions of FMAMs

Since logical operators can be defined on any complete lattice $\mathbb{L}$ (e.g. the classes of interval-valued or general type-2 fuzzy sets), we can generalize FMAMs to $\mathbb{L}$-fuzzy MAMs;
An Example of an Application to Time-Series Prediction in Industry

- Problem of forecasting the average monthly streamflow of Furnas, a large hydroelectric plant in southern Brazil;
- Seasonality of the monthly streamflow suggests the use of 12 different predictor models;
- We standardized the monthly data by subtracting the mean and dividing by the standard deviation and estimated the monthly streamflow $s_\gamma$ from a subset of past values;
- Using fuzzy c-means, the training data $(p_\gamma, s_\gamma)$, where $p_\gamma = [s_{\gamma-3}, s_{\gamma-2}, s_{\gamma-1}]$, yield the centers and standard deviations of Gaussian membership functions $x_\xi$ and $y_\xi$.
- $\{(x_\xi, y_\xi) : \xi = 1, \ldots, k\}$ can be stored implicitly in $W$;
- The estimated value is obtained by defuzzifying $y_\gamma = \mathcal{W}(x_\gamma)$ (using the centroid method);
### Performance of Several Predictors on Test Data

<table>
<thead>
<tr>
<th>Model</th>
<th>MRE (%)</th>
<th>RMSE ($m^3/s$)</th>
<th>MAE ($m^3/s$)</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAR</td>
<td>18.08</td>
<td>266.13</td>
<td>154.44</td>
<td>0.88</td>
</tr>
<tr>
<td>ANFIS</td>
<td>20.12</td>
<td>262.21</td>
<td>166.31</td>
<td>0.88</td>
</tr>
<tr>
<td>C-FSM</td>
<td>20.19</td>
<td><strong>260.82</strong></td>
<td>163.48</td>
<td><strong>0.89</strong></td>
</tr>
<tr>
<td>A-FSM</td>
<td>19.08</td>
<td>278.42</td>
<td>167.77</td>
<td>0.87</td>
</tr>
<tr>
<td>FMAM $\mathcal{W}_F$</td>
<td><strong>18.8</strong></td>
<td>278.08</td>
<td>167.33</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Table: Comparison of the prediction errors produced by the max-$C_F$ FMAM $\mathcal{W}_F$ ($C_F$ denotes the cross-ratio uninorm), an online adaptive (first order Takagi-Sugeno) fuzzy system model (A-FSM), an offline constructive (first order Takagi-Sugeno) fuzzy system model (C-FSM), the adaptive network-based fuzzy inference system (ANFIS) of Jang, and a periodic autoregressive (PAR) model using the data from 1996-2005.
Examples of MNNs

Predictions Produced by the Max-$C_F$ FMAM $\mathcal{W}_F$

Figure: The streamflow prediction for the Furnas reservoir from 1996 to 2005.
Subsethood Measure FAM (S-FAM)

Definition and Some Properties

1. Let $S : \mathcal{F}(\mathbf{X}) \times \mathcal{F}(\mathbf{X}) \to [0, 1]$ be an arbitrary subsethood measure;
2. Given a set of associations $\{(A^\xi, B^\xi) : \xi = 1, \ldots, p\}$, define
   \[
   I(A) = \{j \in \{1, \ldots, p\} : S(A^j, A) = \bigvee_{\xi=1}^p S(A^\xi, A)\}, \quad \forall A \in \mathcal{F}(\mathbf{X}).
   \]
3. A subsethood FAM (S-FAM) is given by $S$, where
   \[
   S(A) = \bigcup_{j \in I(A)} B^j, \quad \forall A \in \mathcal{F}(\mathbf{X}).
   \]
4. A dual S-FAM arises by inverting the order of $A^\xi$ and $A$.
5. If $S(A^\xi, A^\gamma) < 1 \quad \forall \xi \neq \gamma$ then $S(A^\xi) = B^\xi \quad \forall \xi = 1, \ldots, p$.
6. If $S$ is continuous with respect to a metric $d$, then $\exists \delta_S > 0$ such that
   \[
   d(A, A^\xi) < \delta_S \Rightarrow S(A) = B^\xi \quad \forall \xi = 1, \ldots, p.
   \]
Some Subsethood Measures

If $\nu$ is a non-negative function such that $\nu(A) = 0 \iff A = \emptyset$ then we have the following subsethood measures:

(a)  $\bar{S}(A, B) = \begin{cases} 1, & \text{if } A = \emptyset, \\ \frac{\nu(A \cap B)}{\nu(A)}, & \text{if } A \neq \emptyset, \end{cases}$

(b)  $\hat{S}(A, B) = \begin{cases} 1, & \text{if } A = B = \emptyset, \\ \frac{\nu(B)}{\nu(A \cup B)}, & \text{otherwise}. \end{cases}$

For a finite universe $X$, Kosko’s subsethood measure is given using $\nu_K(A) = \sum_{x \in X} \mu_A(x)$ and the S-FAM based on Kosko’s subsethood measure is known as KS-FAM.
Examples of MNNs

Similarity Measure FAM (SM-FAM)

**Definition and Properties:**

- Let $SM : \mathcal{F}(X) \times \mathcal{F}(X) \to [0, 1]$ be an arbitrary similarity measure;
- Given a set of associations $\{(A^\xi, B^\xi) : \xi = 1, \ldots, p\}$, define
  
  $$K(A) = \{j \in \{1, \ldots, p\} : SM(A, A^j) = \sqrt[p]{\prod_{\xi=1}^{p} SM(A, A^\xi)}\}, \forall A \in \mathcal{F}(X).$$

- A similarity measure FAM (SM-FAM) is given by $S.M$, where
  
  $$S.M(A) = \bigcup_{j \in K(A)} B^j, \forall A \in \mathcal{F}(X).$$

- Properties of SM-FAMs are similar to the ones of S-FAMs, e.g.:
  
  - If $SM(A^\xi, A^\gamma) < 1 \ \forall \xi \neq \gamma$ then $SM(A^\xi) = B^\xi \ \forall \xi = 1, \ldots, p$.
  - If $SM$ is continuous with respect to a metric $d$, then $\exists \delta_{SM} > 0$ such that $d(A, A^\xi) < \delta_{SM} \Rightarrow S.M(A) = B^\xi \ \forall \xi = 1, \ldots, p$. 
Examples of MNNs

Relationship of S-FAMs and SM-FAMs to MM

Classification of S-FAMs as MNNs:
- An S-FAM is given by a two-layer MNN;
- The aggregation functions of the hidden neurons compute components of fuzzy erosions of $A$ by the SEs $A^\xi$;
- The output values are computed using dilations (in both senses).

Classification of SM-FAMs as MNNs:
- Suppose that $SM$ is based on a similarity measure $SM$ given by $SM(A, B) = S(A, B) \times S(B, A)$ for some a subsethood measure $S$ and a t-norm $t$.
- Since $S(B, A)$ and $S(A, B)$ yield the degrees of inclusion of the SE $B$ in $A$ and of $A$ in the SE $B$, $SM$ can be viewed as a component of a morphological operator known as hit-or-miss transform;
- Therefore, this type of SM-FAM can also be considered an MNN.
Examples of MNNs

Some Comments on Applications of S-FAMs, dual S-FAMs, and SM-FAMs

Some Experimental Results:

- **Storage and recall of gray-scale images**: the KS-FAM outperformed several competitive models such as the Hamming net, the complex-valued Hopfield net, the OLAM, and the KAM. Only the KS-FAM was able with different types of noise as well as incomplete patterns and variations in brightness and orientation.

- **Vision-based self-localization in robotics**: the KS-FAM exhibited a significantly better performance than other approaches from the literature;

- **Text-independent, closed-set automatic speaker identification**: Parametrized S-FAMs, dual S-FAMs, and SM-FAMs yielded competitive results but were outperformed by an SVM-based approach.
Some Simulations with Gray-Scale Images

Figure: Incomplete, distorted, and corrupted versions of the original images.
Examples of MNNs

Some Simulations with Gray-Scale Images

**Table:** NRMSEs Produced by AM Models in Applications to Patterns Exhibiting Variations in Brightness and Orientation

<table>
<thead>
<tr>
<th>Model</th>
<th>Brightness</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.3 (-77)</td>
<td>+0.3 (+77)</td>
</tr>
<tr>
<td>KS-FAM</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Kosko’s FAM</td>
<td>0.5351</td>
<td>0.8197</td>
</tr>
<tr>
<td>Junbo’s FAM</td>
<td>0.3317</td>
<td>0.5494</td>
</tr>
<tr>
<td>IFAM of Luk.</td>
<td>0.4773</td>
<td>0.5494</td>
</tr>
<tr>
<td>$W_{XX}$</td>
<td>0.5242</td>
<td>0.5498</td>
</tr>
<tr>
<td>$W_{XX} + \nu$</td>
<td>0.5850</td>
<td>0.5032</td>
</tr>
<tr>
<td>complex Hopfield</td>
<td>0.8422</td>
<td>0.7620</td>
</tr>
<tr>
<td>OLAM</td>
<td>0.5215</td>
<td>0.5449</td>
</tr>
<tr>
<td>KAM</td>
<td>0.4664</td>
<td>0.3554</td>
</tr>
<tr>
<td>Hamming net</td>
<td>0.4434</td>
<td>0.6137</td>
</tr>
</tbody>
</table>
Examples of MNNs

An Application of the KS-FAM in Vision-Based Self-Localization

Figure: The landmark images corresponding to the positions selected to build the map.
An Application of the KS-FAM in Vision-Based Self-Localization

Table: Comparison of the results produced by the KS-FAM using 9 associations per position with the best results previously obtained for each walk using off-line mapping by LICA, MF-ICA and MS-ICA, and endmember selection for SLAM.

<table>
<thead>
<tr>
<th></th>
<th>Walk 3</th>
<th>Walk 4</th>
<th>Walk 5</th>
<th>Walk 6</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>KS-FAM</td>
<td>0.82</td>
<td>0.65</td>
<td>0.81</td>
<td>0.79</td>
<td>0.76</td>
</tr>
<tr>
<td>LICA</td>
<td>0.75</td>
<td>0.66</td>
<td>0.73</td>
<td>0.75</td>
<td>0.72</td>
</tr>
<tr>
<td>MF-ICA</td>
<td>0.62</td>
<td>0.54</td>
<td>0.65</td>
<td>0.53</td>
<td>0.58</td>
</tr>
<tr>
<td>MS-ICA</td>
<td>0.69</td>
<td>0.62</td>
<td>0.74</td>
<td>0.69</td>
<td>0.68</td>
</tr>
<tr>
<td>SLAM</td>
<td>0.76</td>
<td>0.6</td>
<td>0.69</td>
<td>0.64</td>
<td>0.67</td>
</tr>
</tbody>
</table>
Examples of MNNs

Morphological Perceptron with Competitive Learning (MP/CL)

- The MP/CL is a feedforward artificial neural network for classification problems in $\mathbb{L}^n$, where $\mathbb{L}$ is a totally ordered group (extension);
- The MP/CL consists of $K$ modules where $K$ is the number of classes. Given an input pattern $x \in \mathbb{L}^n$, the $k$th module computes:

$$y_k = \bigvee_{j=1}^{m_k} (\varepsilon_{v_j^k}(x) \land \bar{\delta}_{w_j^k}(x)),$$

where $\bar{\delta}_{w_j^k}(x) = x^* \square w$. The number $m_k$ of pairs $(\varepsilon_{v_j^k}, \bar{\delta}_{w_j^k})$ as well as the weight vectors $v_j^k$ and $w_j^k$ are determined by the MP/CL learning algorithm.

- Then the MP/CL produces the class label $y = \arg \max_k y_k$. 
Module of an MP/CL as well as an MP/CL corresponding to the $s$th class.

Figure:
Examples of MNNs

Individual Step of MP/CL Algorithm for Binary Problem

Figure:

Step of the MP/CL Algorithm for Binary Classification.

Peter Sussner (Unicamp)
Illustration of a Decision Surface in Binary Problem

Figure:
The difference in shading corresponds to the decision surface produced by the MP/CL model for Ripley’s synthetic problem.
Characteristics of the MP/CL

MP/CL Learning Algorithm

This supervised and constructive algorithm can be applied to training data in $\mathbb{L}_1 \times \mathbb{L}_n$ for chains $\mathbb{L}_i$. We proved:

1. Convergence in a finite number of steps;
2. Perfect separation of the training data according to their class labels;
3. No areas of indecision after training;
4. Independence of the order of the training patterns.

Experimental Results

1. Fast Convergence of learning algorithm;
2. Very satisfactory classification results;
3. Low computational effort.
Fuzzy Lattice

A fuzzy lattice is a pair \((\mathbb{L}, \mu)\) consisting of a lattice \(\mathbb{L}\) and a function \(\mu : \mathbb{L} \times \mathbb{L} \rightarrow [0, 1]\) such that \(\mu(x, y) = 1\) if and only if \(x \leq y\).

Inclusion Measure

For a complete lattice \(\mathbb{L}\), an “inclusion measure” \(\sigma\) is a function \(\mathbb{L} \times \mathbb{L} \rightarrow [0, 1]\) such that \(\forall u, w, x, y \in \mathbb{L}\):

1. \(\sigma(x, O) = 0\) if \(x \neq O = \wedge \mathbb{L}\);
2. \(\sigma(x, x) = 1\) \(\forall x \in \mathbb{L}\);
3. \(u \leq w \Rightarrow \sigma(x, u) \leq \sigma(x, w)\);
4. \(x \wedge y < x \Rightarrow \sigma(x, y) < 1\),

In this case, \((\mathbb{L}, \sigma)\) is a fuzzy lattice.
Examples of MNNs

Some Details on FLR Classifiers

Training Phase

Given a set of training data \{\((E_1, c_1), \ldots, (E_L, c_L)\)\}, where \(E_i\) are information granules such as fuzzy interval’s numbers or hyperboxes in \(\mathbb{R}^n\) and \(c_i\) are class labels, a clustering algorithm is performed resulting in a set \{\((\bar{E}_1, \bar{c}_1), \ldots, (\bar{E}_M, c_M)\)\} with elements of the same type.

Recall Phase

Upon presentation of a granule \(F\), compute \(\sigma(F, \bar{E}_j)\) \(\forall j = 1, \ldots, M\). Assign \(F\) to the class label \(c_J\) such that \(J = \arg\max_j \sigma(F, \bar{E}_j)\).

Remarks

- \(\sigma(F, E)\) is a component of an erosion of \(E\) by the SE \(F\);
- FLR training depends on the order of the training data;
- FLR models have been proposed for more general types of data than MP/CLs and (dual) S(M)-FAMs;
Hybrid Models having morphological as well as other types of neurons include:

- Morphological shared-weight neural networks (NNs);
- Morphological regularization NNs;
- Morphological-Rank-Linear (MRL) NNs.

Morphological shared-weight and regularization NNs

- Two stages: feature extraction stage and conventional linear stage;
- Morphological operators known as hit-or-miss transforms are used in the feature extraction stage;
- Morphological regularization NNs include a regularizing term in the objective function.
- Applications include automatic target recognition, in particular landmine detection, and handwriting recognition.
Morphological-Rank-Linear (MRL) NNs

Characteristics:

- MRL NNs are feedforward artificial NNs;
- Each layer consists of a convex linear combination of a linear and a morphological/rank module followed by the application of an activation function $f$;
- The morphological/rank module employs the $r$th rank function $\mathcal{R}_r$, where $\mathcal{R}_r(x)$ equals the $r$th largest element of $\{x_1, \ldots, x_n\}$.
- MRL NNs have been trained using either modified versions of backpropagation or evolutionary algorithms.
- Applications of MRL NNs include noise removal in image processing, handwritten character recognition, and financial time series prediction. (The latter application involves learning an additional weight corresponding to the phase of the signal.)
Single-Layer MRL NN

Example

A single-layer MRL NN with a single output neuron and $f = id_{\mathbb{R}}$ computes the following output $y$ for an input $x \in \mathbb{R}^n$:

$$y = \lambda \alpha + (1 - \lambda) \beta, \quad \lambda \in [0, 1],$$

where

$$\alpha = R_r(a + x) \quad \text{and} \quad \beta = b^T \cdot x = \sum_{i=1}^{n} b_i x_i$$

represent resp. the linear and morphological/rank modules.

- Here the weight vector consists of $\lambda \in \mathbb{R}$, $r \in \{1, \ldots, n\}$, and $a, b \in \mathbb{R}^n$.
- Note that for fixed (SE) $a \in \mathbb{R}^n$ and inputs $x \in \mathbb{R}^n$, the function $R_r(a + \cdot)$ generalizes both $\varepsilon_a$ and $\delta_a$. 
Concluding Remarks

- MNNs can be seen as approaches towards lattice computing or computational intelligence based on lattice theory.
- MNNs perform morphological operations in the lattice-algebraic or geometrical/topological sense.
- Most MNN models have strong theoretical foundations in MM on complete lattices (note that we recently introduced associative memories based on MM on complete inf-semilattices).
- MNNs have been used by researchers from Europe, the Americas, Japan, and China for a variety of applications such as pattern recognition, image and signal processing, computer vision, approximate reasoning, and prediction.
Perspectives for the Future of MNNs

- Several types of MNNs allow for extensions to more general lattice structures. In particular, there is an interest in disparate types of data.

- Since many classes of information granules are lattice ordered, there is a need to explore the potential of MNNs for granular computing. In particular, there is an interest in processing disparate types of data.

- In particular, the advent of \( \mathbb{L} \)-fuzzy MM provides an access to MNNs on \( \mathbb{L} \)-fuzzy sets such as interval-valued, intuitionistic, and bipolar fuzzy sets as well as (general) Type-2 fuzzy sets, that have become increasingly important in rule-based systems for applications in engineering and computing with words as well as in approximate reasoning.

Thanks for your interest!

Peter Sussner (Unicamp)