Keynote talk: Machine Learning in Grey-Based Soft Computing techniques

Keynote speaker: Prof. Jose L. Salmeron
Universidad Pablo de Olavide (Seville, Spain)

September 23, 2015
About me

Current position

- Prof. Jose L. Salmeron, Ph.D. CS (AI), Ph.D. BA, MSc. CS, MBA, BSc. CS, BBA
  Professor of Information Systems.
  University Pablo de Olavide at Seville (Spain)

- **Home page**: [http://eps.upo.es/salmeron/](http://eps.upo.es/salmeron/)

- **e-mail**: salmeron@acm.org

Research

- **Research interests**: Decision Making in Autonomous Systems, Fuzzy Cognitive Maps, Decision Support Systems, Forecasting, Soft computing, Computational Intelligence, Knowledge-Based Systems, and always researching in something new...

- Member of societies as ACM, Association for Logic Programming, ENBIS and others.

- **Consulting** for Getronics, KPMG, Everis, GEA21, UHK, DDTel, and others.
Outline

1. FCM fundamentals
   - Basics
   - Construction
   - Analysis

2. Fuzzy Grey Cognitive Maps
   - Grey Systems Theory
   - FGCM fundamentals
   - Analysis
   - Example: A FGCM-based intelligent system

   - Unsupervised Learning
   - Supervised Learning
Outline

1. FCM fundamentals
   - Basics
   - Construction
   - Analysis

2. Fuzzy Grey Cognitive Maps
   - Grey Systems Theory
   - FGCM fundamentals
   - Analysis
   - Example: A FGCM-based intelligent system

   - Unsupervised Learning
   - Supervised Learning
Fuzzy Cognitive Maps components

Representation

- FCM is a soft computing technique, closely to a **one-layer recurrent dynamic neural network**.
- FCMs consist of concepts, that illustrate different aspects in the system’s behavior and these concepts interact with each other showing the dynamics of the system.
- FCMs are represented by directed graphs capable of modelling relationships or casualities existing between concepts. Concepts \((c_i)\) are represented by nodes and edges \((e_{ij})\) represent relationships between them.
Introduction

Definitions

A FCM can be represented as a 4-tuple

$$\Gamma = \langle N, E, f, r \rangle$$

where $N$ is the set of nodes, $E$ are the set of edges between nodes, $f(\cdot)$ the activation function and $r$ the nodes’ range, $r = \{[0, +1]|[-1, +1]\}$, $N = \{< n_i >\}$ where $n_i$ are the nodes.

$E$ is represented as

$$E = \{< e_{ij}, w_{ij} > | n_i, n_j \in V\}$$

where $e_{ij}$ is the edge from node $n_i$ to node $n_j$, and $w_{ij}$ is the weight of the edge $e_{ij}$. 
Keynote talk: Machine Learning in Grey-Based Soft Computing techniques

FCM fundamentals

- Basics

Topology

One-layer recurrent dynamical NN

- FCM models a system as an one-layer NN.
- Note that it is not a neural network. Figure shows just an analogy.

Main difference with NN

- Each node has a meaning in FCM
Expert based

**Experts issues**

- Expert-based construction is strongly dependent on the experts’ selection and its knowledge.

**Augmented Weighted FCM**

- Each expert could have a credibility weight \( e_i \).
- \[ A^* = \sum_{k=1}^{E} \left( \frac{e_k}{\sum_{m=1}^{E} e_k} \right) \cdot A_k \]

**Defuzzification**

- It is necessary to convert the fuzzy quantities into crisp quantities for further FCM processing. For instance, Center of Gravity:

\[
COG_y = \frac{\int_{x \in X} x \cdot \mu_A(x) \, dx}{\int_{x \in X} \mu_A(x) \, dx}
\]
Delphi/Augmented FCM

Delphi method

- It is just for expert-based FCMs.
- Experts design FCM in several rounds.
- After the first one, each expert knows the overall data and he/she can adjust his/her previous judgement.

Augmented FCM

- It doesn’t need that experts change their initial judgement for consensus as Delphi methodology.
- The augmented adjacency matrix \( A^* = [w_{ij}^*]_{m \times m} \) is built adding the adjacency matrix of each data source.
- If there are common nodes within the adjacency matrices, the element \( w_{ij}^* \) would be
  \[
  w_{ij}^* = \frac{1}{n} \cdot \sum_{k=1}^{n} w_{ijk}
  \]
Keynote talk: Machine Learning in Grey-Based Soft Computing techniques

FCM fundamentals

Construction

Data-driven

Goal

- Improve the FCM model

<table>
<thead>
<tr>
<th>Alg.</th>
<th>Learning goal</th>
<th>Human involvement</th>
<th>Data type</th>
<th>Activation function</th>
<th>Learning type</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHL</td>
<td>Adj. mat.</td>
<td>No</td>
<td>Single</td>
<td>N/A</td>
<td>Hebbian</td>
</tr>
<tr>
<td>BDA</td>
<td>Adj. mat.</td>
<td>No</td>
<td>Single</td>
<td>Binary</td>
<td>modified Hebbian</td>
</tr>
<tr>
<td>NHL</td>
<td>Adj. mat.</td>
<td>Initial</td>
<td>Single</td>
<td>Continuous</td>
<td>modified Hebbian</td>
</tr>
<tr>
<td>AHL</td>
<td>Adj. mat.</td>
<td>Initial</td>
<td>Single</td>
<td>Continuous</td>
<td>modified Hebbian</td>
</tr>
<tr>
<td>GS</td>
<td>Adj. mat.</td>
<td>No</td>
<td>Multiple</td>
<td>Continuous</td>
<td>Genetic</td>
</tr>
<tr>
<td>PSO</td>
<td>Adj. mat.</td>
<td>No</td>
<td>Multiple</td>
<td>Continuous</td>
<td>Swarm</td>
</tr>
<tr>
<td>GA</td>
<td>Initial vector</td>
<td>N/A</td>
<td>N/A</td>
<td>Continuous</td>
<td>Genetic</td>
</tr>
<tr>
<td>RCGA</td>
<td>Adj. mat.</td>
<td>No</td>
<td>Single</td>
<td>Continuous</td>
<td>Genetic</td>
</tr>
</tbody>
</table>

AHL/NHL human involvement

- Initial human intervention is necessary, but later when applying the algorithm there is no human intervention needed
### Expert-based vs. data-driven approaches

<table>
<thead>
<tr>
<th></th>
<th>Expert-based</th>
<th>Data-driven</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of modeling</strong></td>
<td><em>Deductive</em></td>
<td><em>Inductive</em></td>
</tr>
<tr>
<td><strong>Main objective</strong></td>
<td>To create a model that is structurally understandable</td>
<td>To create a model that provides accurate simulations</td>
</tr>
<tr>
<td><strong>Main application</strong></td>
<td>Static analysis</td>
<td>Dynamic analysis</td>
</tr>
<tr>
<td><strong>Main shortcoming</strong></td>
<td>Dynamic analysis could be inaccurate</td>
<td>Static analysis could be inaccurate and more difficult</td>
</tr>
</tbody>
</table>
Static

Concept

Static analysis study the characteristics of the FCM weighted directed graph that represent the model, using graph theory techniques.

Density

Density ($D$) is the relation between the nodes ($N$) and the edges ($E$) of the model.

$$D = \frac{E}{N \cdot (N - 1)}$$

High density indicates increased complexity in the model and respectively to the problem that the model represents.
## Static

### Some measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-degree</td>
<td>$d^+(c_j) = \sum_{i=1}^{n}</td>
</tr>
<tr>
<td>Out-degree</td>
<td>$d^-(c_i) = \sum_{j=1}^{n}</td>
</tr>
<tr>
<td>Centrality</td>
<td>$c(c_i) = d^+(c_i) + d^-(c_i)$</td>
</tr>
<tr>
<td>Weighted in-degree</td>
<td>$d^+<em>w(c_j) = \sum</em>{i=1}^{n}</td>
</tr>
<tr>
<td>Weighted out-degree</td>
<td>$d^-<em>w(c_i) = \sum</em>{j=1}^{n}</td>
</tr>
<tr>
<td>Weighted centrality</td>
<td>$c_w(c_i) = d^+_w(c_i) + d^-_w(c_i)$</td>
</tr>
<tr>
<td>Input node</td>
<td>$n_i</td>
</tr>
<tr>
<td>Output node</td>
<td>$n_i</td>
</tr>
</tbody>
</table>

![Diagram of a network with labeled nodes and edges](image)

Legend:
- **Blue** Input node
- **Red** Output node
- **Regular** Regular node

Keynote talk: Machine Learning in Grey-Based Soft Computing techniques
- FCM fundamentals
- Analysis
FCM Dynamics

Concept

FCM is a deterministic model. It predicts future state evolution deterministically as shown in the updating rules with the expectation that the initial state vector is converged finally to a fixed point.

Updating nodes

\[ c_i(t + 1) = f\left( \sum_{j=1}^{n} w_{ji} \cdot c_j(t) \right) \]

\[ c_i(t + 1) = f\left( \sum_{j=1}^{n} w_{ji} \cdot c_j(t) + c_i(t) \right) \]

\[ c_i(t + 1) = f\left( k_1 \cdot \sum_{j=1}^{n} w_{ji} \cdot c_j(t) + k_2 \cdot c_i(t) \right) \]

1st iteration

\[ c(0) = (c_1(0) \cdot \cdots \cdot c_4(0)) \rightarrow c(1) = (f(\cdot) \cdot \cdots \cdot f(\cdot)) \]

2nd iteration

\[ c(1) = (c_1(1) \cdot \cdots \cdot c_4(1)) \rightarrow c(2) = (f(\cdot) \cdot \cdots \cdot f(\cdot)) \]

\[ \cdots \cdots \cdots \]

Last iteration

\[ \sqrt{\sum_{i=1}^{N} (c_i(t) - c_i(t - 1))^2} < \text{tolerance} \]

\[ \epsilon \]
Activation functions


**Concept**

The FCM inference process finish when the stability is reached. The FCM reaches either one of the following states following the iterations.

- Fixed-point attractor \( c(t - 1) = c(t) \)
- Limited cycle \( \exists t, k | (c(t - k) = c(t)) \land (c(t - k + 1) \neq c(t)) \)
- Chaotic attractor

**Main activation functions**

- **Bivalent**
  \[
  f(c_i) = \begin{cases} 
  0 & \text{if } c_i \leq 0 \\
  1 & \text{if } c_i > 0 
  \end{cases}
  \]

- **Trivalent**
  \[
  f(c_i) = \begin{cases} 
  -1 & \text{if } c_i \leq -0.5 \\
  0 & \text{if } -0.5 < c_i < 0.5 \\
  +1 & \text{if } c_i \geq -0.5 
  \end{cases}
  \]

- **Unipolar sigmoid**
  \[
  f(c_i) = \frac{1}{1 + e^{-\lambda \cdot c_i}}
  \]

- **Hyperbolic tangent**
  \[
  f(c_i) = \frac{e^{2 \lambda \cdot c_i} + 1}{e^{2 \lambda \cdot c_i} - 1}
  \]
What-if analysis

- **Goal**: Impact of \(c(0)\) over \(c(t)\)
- **Method**: Running FCM dynamics

Population-based analysis

- **Method**: Evolving a population of initial vector states \(\{c(0)\}_{i=1}^{n}\)
- **Fitness**: \(\min(err)\) where \(err_i = |c(t)_i - c(t)_i^*|\)
- **Forward analysis**
  - Finding the optimum final vector state \((c(t))\)
- **Backward analysis**
  - Finding \(c(0)\) that generates a specific \(c(t)\)
Ranking FCM scenarios with TOPSIS


FCM consensus
This proposal includes Delphi

FCM dynamics
Dynamic analysis is run for each scenario

TOPSIS-based rank
The closer scenario to the positive-ideal scenario is the best solution.
FCM extensions

- Fuzzy Grey Cognitive Maps
- Rule-Based Fuzzy Cognitive Maps
- Probabilistic Fuzzy Cognitive Maps
- Intuitionistic Fuzzy Cognitive Maps
- Dynamical Cognitive Networks
- Belief Degree-Distributed Fuzzy Cognitive Maps

- Rough Cognitive Maps
- Dynamic Random Fuzzy Cognitive Maps
- Fuzzy Cognitive Networks
- Evolutionary FCMs
- Fuzzy Time Cognitive Maps
- Fuzzy Rules Incorporated with FCMs
- Timed Automata-based Fuzzy Cognitive Maps
Outline

1. FCM fundamentals
   - Basics
   - Construction
   - Analysis

2. Fuzzy Grey Cognitive Maps
   - Grey Systems Theory
   - FGCM fundamentals
   - Analysis
   - Example: A FGCM-based intelligent system

   - Unsupervised Learning
   - Supervised Learning
Overview

Introduction

- A very effective solving problem set of techniques within environments with high uncertainty.
- GST have been designed to analyze small data samples with poor information.

Comparison between GST, and Fuzzy theory

- Fuzzy theory holds some previous information (usually based on experience); while GST deal with objective data, they do not need any previous information other than the data sets that need to be disposed.

- The key difference between fuzzy and grey systems concepts is intension and extension of the analyzed objects. While GST focus on objects with clear extension and ambiguous intension, fuzzy theory mostly analyze the objects with clear intension and ambiguous extension.

Comparison between Black, Grey and White systems

<table>
<thead>
<tr>
<th></th>
<th>Black</th>
<th>Grey</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>Unknown</td>
<td>Incomplete</td>
<td>Known</td>
</tr>
<tr>
<td>Conclusion</td>
<td>No result</td>
<td>Multiple solution</td>
<td>Unique solution</td>
</tr>
</tbody>
</table>
Modelling grey uncertainty with Fuzzy Grey Cognitive Maps

Jose L. Salmeron

School of Engineering, University Pablo de Olavide, Ctra. de Utrera, km. 1, 41013 Seville, Spain
Introduction


Fuzzy Grey Cognitive Maps

- A FGCM models unstructured knowledge through causalities through imprecise terms and grey relationships between them based on Fuzzy Cognitive Maps.
- FGCMs are a generalization of FCMs, since a FGCM with all the relations’ intensities represented by white numbers would be a FCM.
- FGCM represents the human intelligence better than FCM, because it faces unclear relations between factors and incomplete information better than FCM.

Comparison

<table>
<thead>
<tr>
<th>Requirements</th>
<th>Systems Dynamics</th>
<th>Bayesian Networks</th>
<th>Neural Networks</th>
<th>FCM</th>
<th>FGCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Considers all possible connections</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Do not ignore the uncertainty</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Directed graph with cycles</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>The propagation does not carry on an es-</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>tablished pattern</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assume information is scarce</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Process the data source’s uncertainty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Consider multi-meaning problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Introduction

Definitions

A FGCM can be represented as a 4-tuple

\[ \Psi = \langle N, E, f, r \rangle \]

where \( N \) is the set of nodes and \( E \) are the set of edges between nodes, \( f(\cdot) \) the grey activation function and \( r \) the limits nodes' state range and \( \{c, \overline{c}\} \in \{[0, +1]|[-1, +1]\} \). Note that \( r_\ominus \) and \( \overline{r} \) must have the same range.

As FCMs, \( N \) is represented as a tuple

\[ N = \{< n_i >\} \]

where \( n_i \) are the nodes. \( E \) is represented by

\[ E = \{< e_{ij}, \otimes w_{ij} > | n_i, n_j \in V \} \]

where \( e_{ij} \) is the edge from node \( n_i \) to node \( n_j \), and \( \otimes w_{ij} \) is the grey weight of the edge \( e_{ij} \). Note that FGCMs include grey uncertainty within grey weights and grey nodes' states.
Keynote talk: Machine Learning in Grey-Based Soft Computing techniques

---

Fuzzy Grey Cognitive Maps

FGCM fundamentals

Grey weights $⊗g$ limits and measures

Grey numbers $⊗g \in [g, \bar{g}]$ in FGCM

1. $⊗g_1 \in [\underline{g}_1, \bar{g}_1] \mid \underline{g}_1 < g_1$
2. $⊗g_2 \in [\underline{g}_2, \bar{g}_2] \mid g_2 < +\infty, \bar{g}_2 = +\infty$
3. $⊗g_3 \in [\underline{g}_3, \bar{g}_3] \mid g_3 = -\infty, \bar{g}_3 > -\infty$
4. $⊗g_4 \in [\underline{g}_4, \bar{g}_4] \mid g_4 = -\infty, \bar{g}_4 = +\infty$
5. $⊗g_5 \in [\underline{g}_5, \bar{g}_5] \mid \underline{g}_5 = \bar{g}_5$

Whitenization

The transformation process of grey numbers in white ones

$\hat{⊗g} = \delta \cdot \underline{g} + (1 - \delta) \cdot \bar{g} \mid \delta \in [0, 1]$

Greyness

GST includes greyness as an uncertainty measurement. With higher values of greyness then a higher uncertainty degree.

$\phi(⊗g) = \frac{|\ell(⊗g)|}{\ell(⊗ψ)}$
Grey weight with base value and symmetric / asymmetric vibration

\[ \otimes w(e) \in [w, \bar{w}] \]

\[ \otimes w(e) \in \begin{cases} 
[e - \delta_e, e + \delta_e] & \text{if } (-1 \leq e - \delta_e \leq +1) \land (-1 \leq e + \delta_e \leq +1) \\
[e - \delta_e, +1] & \text{if } (-1 \leq e - \delta_e \leq +1) \land (+1 < e + \delta_e) \\
[-1, e + \delta_e] & \text{if } (e - \delta_e < -1) \land (-1 \leq e + \delta_e \leq +1) \\
[-1, +1] & \text{if } (e - \delta_e < -1) \land (+1 < e + \delta_e) 
\end{cases} \]
Keynote talk: Machine Learning in Grey-Based Soft Computing techniques

Fuzzy Grey Cognitive Maps

FGCM fundamentals

Grey weight with same base value and different (+|−) vibration

Vibration illustrative example

- Same base value \( e \) with a couple of different vibrations \( \delta^a_e \) and \( \delta^b_e \)

\[ \otimes w^a(e) \]

- Base value positive and \( e \pm \delta^a_e > 0 \)

\[ \otimes w^b(e) \]

- Base value positive and \( w^b < 0 \wedge w^b > 0 \)
- Note that base value is the same in both cases \( (\otimes w^a(e) \) and \( \otimes w^b(e) ) \)
- The vibration includes a negative zone within the grey number \( [e - \delta^b_e, 0] \)
FGCM uncertainty measures

Greyness

Furthermore, FGCM includes greyness as an uncertainty measurement. Higher values of greyness mean that the results have a higher uncertainty degree. It is computed as follows

\[ \phi(\otimes c_i) = \frac{|\ell(\otimes c_i)|}{\ell(\otimes \psi)} \]

where \( |\ell(\otimes c_i)| \) is the absolute value of the length of grey node \( \otimes c_i \) state value, and \( \ell(\otimes \psi) \) is the absolute value of the range in the information space, denoted by \( \psi \).

FGCM maps the nodes’ states within an interval \([0, 1]\) or \([-1, +1]\) if negative values are allowed. In this sense,

\[ \ell(\otimes \psi) = \begin{cases} 1 & \text{if } \{\otimes c_i, \otimes w_i\} \subseteq [0, 1] \quad \forall \otimes c_i, \otimes w_i \\ 2 & \text{if } \{\otimes c_i, \otimes w_i\} \subseteq [-1, +1] \quad \forall \otimes c_i, \otimes w_i \end{cases} \]
Keynote talk: Machine Learning in Grey-Based Soft Computing techniques

- Fuzzy Grey Cognitive Maps
- Analysis

Static

Grey measures

- FGCM includes new uncertainty measures

| Greyness | \( \phi^+(\otimes c_j) = \sum_{i=1}^{n} \frac{|\ell(\otimes w_{ij})|}{\ell(\otimes \psi)} \) |
|---------|-------------------------------------------------|
| in-degree | \( \phi^-(\otimes c_j) = \sum_{j=1}^{n} \frac{|\ell(\otimes w_{ij})|}{\ell(\otimes \psi)} \) |
| Greyness | Overall greyness \( \phi^{\oplus}(\otimes c_i) = \phi(\otimes c_i) + \phi^+(\otimes c_j) + \phi^-(\otimes c_j) \) |
**FCM vs. FGCM**

**Fuzzy Cognitive Maps**

1. Just one kind of relationship between nodes \( w_{13} \)

**Fuzzy Grey Cognitive Maps**

1. White relationship
   \( \otimes w_{13} \in [\overline{w}_{13}, \overline{w}_{13}] \ | \overline{w}_{13} = \overline{w}_{13} \)

2. Grey relationship
   \( \otimes w_{12} \in [w_{12}, \overline{w}_{12}] \ | w_{12} < \overline{w}_{12} \)

3. Black relationship
   \( \otimes w_{23} \in [\overline{w}_{23}, w_{23}] \ | w_{23} = -1 \land \overline{w}_{23} = +1 \)
FGCM dynamics

Initial grey vector state

FGCMs dynamics begins with the design of the initial grey vector state $\otimes \vec{c}(0)$, which represents a proposed initial grey stimuli.

Grey activation function

The updated nodes’ states are computed in an iterative inference process with a grey activation function.

Grey unipolar sigmoid

The unipolar sigmoid function is the most used one in FCM and FGCM when the concept value maps in the range $[0, 1]$.

Grey hyperbolic tangent

When the concepts’ states map in the range $[-1, +1]$ the function used would be the hyperbolic tangent.
FGCM dynamics

Initial grey vector state

FGCMs dynamics begins with the design of the initial grey vector state $\vec{c}(0)$, which represents a proposed initial grey stimuli.

Grey activation function

The updated nodes' states are computed in an iterative inference process with a grey activation function.

Grey unipolar sigmoid

The unipolar sigmoid function is the most used one in FCM and FGCM when the concept value maps in the range $[0, 1]$.

Grey hyperbolic tangent

When the concepts' states map in the range $[-1, +1]$ the function used would be the hyperbolic tangent.

FGCM stability

- The FGCM inference process finish when the stability is reached.
- After its inference process, the FGCM reaches either one stable state following a number of iterations.
  1. It settles down to a fixed pattern of node states, the so-called grey hidden pattern or grey fixed-point attractor.
  2. Moreover, the state could keep cycling between several fixed states, known as a limit grey cycle.
  3. Using a continuous activation function, a third state would be a grey chaotic attractor. It happens when, instead of stabilizing, the FGCM continues to produce different grey vector states for each iteration.
FCM vs. FGCM Construction process

**Grey Weights determination**

1. Base value definition. It could be assign as the same way than in FCM. Anyway, the base value is located within the FGCM nodes/weights' range $w_{ij}^e \in \{[0, +1] | [-1, +1]\}$.

2. Selection of the vibration value of the base value $\delta_e$. If construction is automatic, vibration could represent noise or another measurements distortions. In the experts based construction process the base value determine the $\delta_e$ value.

   - If the expert has the whole trust on the base value $e$, the vibration’s value would be $\delta_e = 0$.
   - If the expert has not any trust on the base value, the vibration’s value would be as $w_{ij}^e \pm \delta_e = \pm 1$.
   - Otherwise, $(w_{ij}^e + \delta_e \leq +1) \land (w_{ij}^e - \delta_e \geq -1)$.
Applications

A Fuzzy Grey Cognitive Maps-based Decision Support System for radiotherapy treatment planning

Jose L. Salmeron a,*, Elpiniki I. Papageorgiou b

a University Pablo de Olavide, School of Engineering, 1st km. Ulterra Road, 41013 Seville, Spain
b Technological Educational Institute of Lamia, Department of Informatics and Computer Technology, 3rd km. Old National Road Lamia-Athens, TR 35100 Lamia, Greece

Fuzzy Grey Cognitive Maps in reliability engineering

Jose L. Salmeron a, Ester Gutierrez b,*

a University Pablo de Olavide, 1st km. Ulterra Road, 41013 Seville, Spain
b University of Seville, C. Descubrimientos s/n, 41092 Seville, Spain
Keynote talk: Machine Learning in Grey-Based Soft Computing techniques

Fuzzy Grey Cognitive Maps

Example: A FGCM-based intelligent system

Model

Problem

- A synthetic case study for showing the added value of FGCMs.
- The simulation consists in an Intelligent Security System controlling several surveillance assets.
Keynote talk: Machine Learning in Grey-Based Soft Computing techniques
---
- Fuzzy Grey Cognitive Maps
- Example: A FGCM-based intelligent system

## Model

### Problem
- A synthetic case study for showing the added value of FGCMs.
- The simulation consists in an Intelligent Security System controlling several surveillance assets.

### Table: Initial grey vector states \( \{ \otimes c(0)_i \}_{i=1}^3 \)

<table>
<thead>
<tr>
<th>( \otimes c_1(0) )</th>
<th>( \otimes c_1(0)_1 )</th>
<th>( \otimes c_1(0)_2 )</th>
<th>( \otimes c_1(0)_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \otimes c_2(0) )</td>
<td>[.0, .0]</td>
<td>[.0, .0]</td>
<td>[.0, .0]</td>
</tr>
<tr>
<td>( \otimes c_3(0) )</td>
<td>[.4, .4]</td>
<td>[.0, .0]</td>
<td>[.0, .0]</td>
</tr>
<tr>
<td>( \otimes c_4(0) )</td>
<td>[.0, .0]</td>
<td>[.0, .0]</td>
<td>[.0, .0]</td>
</tr>
<tr>
<td>( \otimes c_5(0) )</td>
<td>[−.5, −.5]</td>
<td>[−.5, −.5]</td>
<td>[−.8, −.8]</td>
</tr>
<tr>
<td>( \otimes c_6(0) )</td>
<td>[.6, .6]</td>
<td>[−.4, −.4]</td>
<td>[.0, .0]</td>
</tr>
</tbody>
</table>

### Solution
- \( \otimes c_1(s)_2 \succ \otimes c_1(s)_3 \succ \otimes c_1(s)_1 \)

\( \otimes c_1(s)_1 \in [−.9420, −.8909] \)
\( \otimes c_1(s)_2 \in [.8240, .9043] \)
\( \otimes c_1(s)_3 \in [.7720, .8848] \)
Outline

1. FCM fundamentals
   - Basics
   - Construction
   - Analysis

2. Fuzzy Grey Cognitive Maps
   - Grey Systems Theory
   - FGCM fundamentals
   - Analysis
   - Example: A FGCM-based intelligent system

   - Unsupervised Learning
   - Supervised Learning
Learning Fuzzy Grey Cognitive Maps using Nonlinear Hebbian-based approach

Elpiniki I. Papageorgiou\textsuperscript{a,}\ast, Jose L. Salmeron\textsuperscript{b}

\textsuperscript{a} Technological Educational Institute of Lamia, Department of Informatics and Computer Technology, 3rd km. Old National Road Lamia-Athens, TK 35100 Lamia, Greece

\textsuperscript{b} University Pablo de Olavide, School of Engineering, 1st km. Utrera Road, 41013 Seville, Spain

**NHL-FGCM**

- NHL learning rule for FGCMs introduces a learning rate parameter, the determination of input and output nodes, and the termination conditions.

\[
\Delta \otimes w_{ji}(k) = \eta_k \cdot \otimes c_j(k - 1) \cdot (\otimes c_i(k - 1) - \otimes c_j(k - 1) \cdot \otimes w_{ji}(k - 1))
\]

where the coefficient \( \eta_k \) is a very small positive scalar factor called learning parameter.

- This simple rule states that if \( c_i(k) \) is the value of node \( c_i \) at iteration \( k \), and \( c_j \) is the value of the triggering node \( c_j \) which triggers the node \( c_i \), the corresponding grey weight \( \otimes w_{ji} \) from node \( c_j \) towards the node \( c_i \) is increased proportional to their product multiplied with the learning rate parameter minus the grey weight decay at iteration step \( k \). As a result

\[
\otimes w_{ji}(k) = \frac{w_{ji}(k)}{w_{ij}(k - 1) + \eta_k \cdot \otimes c_j \cdot (\otimes c_i(k) - \otimes c_j \cdot \otimes w_{ji}(k - 1))}
\]

\[
\overline{w}_{ij}(k - 1) + \eta_k \cdot \otimes c_j \cdot (\otimes c_i(k) - \otimes c_j \cdot \otimes w_{ji}(k - 1)) \]

\[
\overline{w}_{ji}(k)
\]

**NHL-FGCM termination criteria**

These criteria determine when the iterative process of the learning algorithm terminates. Through this process and when the termination conditions are met, the final weight matrix updated $\otimes \mathbf{w}^{updated}$ is obtained.

1. Maximization of the objective function $J$, which has been defined by Hebb’s rule.

   $\max J = E\{z^2\}$

   subject to: $||\mathbf{w}|| = 1$

   where $z = f(y)$, and $f$ is the sigmoid function. The objective function $J$ is defined as

   $J = \sum_{i=1}^{l}(OC_i)^2$, where $l$ is the number of $OC$s.

2. Minimization of the variation of two subsequent values of $OC$s.

   $|OC_j^{k+1} - OC_j^k| < e$

   where the term $e$ is a tolerance level keeping the variation of values of $OC$'s as low as possible and it is proposed as $e = 0.001$.


### NHL-FGCM Supervisor

<table>
<thead>
<tr>
<th>Node</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC1</td>
<td>Tumor Localization.</td>
</tr>
<tr>
<td>SC2</td>
<td>Dose prescribed from Treatment Planning (TPD).</td>
</tr>
<tr>
<td>SC3</td>
<td>Machine factors.</td>
</tr>
<tr>
<td>SC4</td>
<td>Human factors.</td>
</tr>
<tr>
<td>SC5</td>
<td>Patient positioning and immobilization.</td>
</tr>
<tr>
<td>SC6</td>
<td>Quality Assurance (QA).</td>
</tr>
<tr>
<td>SC7</td>
<td>Final Dose given to the target volume (FD).</td>
</tr>
</tbody>
</table>
Evolutionary learning of fuzzy grey cognitive maps for the forecasting of multivariate, interval-valued time series

Wojciech Froelich a,*, Jose L. Salmeron b, **

a Institute of Computer Science, University of Silesia, ul. Bedzinska 39, Sosnowiec, Poland
b Computational Intelligence Lab, University Pablo de Olavide, 1st km. Utrera road, 41013 Seville, Spain
RCGA-FGCM


Population

- Data sources (final states’ set vs. time series)
- The objective is to optimize the matrix $\otimes A_{n \times n}$ with respect to the forecasting accuracy.

Fitness

- Fitness function depends on raw data

Algorithm 1: RCGA-based FGCM pseudo-code

<table>
<thead>
<tr>
<th>Data: Domain’s raw data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result: Automatic built FCM</td>
</tr>
<tr>
<td>1 Choose initial FGCM population (random);</td>
</tr>
<tr>
<td>2 Design fitness function;</td>
</tr>
<tr>
<td>3 Evaluate each FGCM’s fitness;</td>
</tr>
<tr>
<td>while Termination is false do</td>
</tr>
<tr>
<td>5 Select best-ranking FGCM to reproduce;</td>
</tr>
<tr>
<td>6 Mate pairs at random;</td>
</tr>
<tr>
<td>7 Prune FGCM population;</td>
</tr>
<tr>
<td>8 Crossover operator;</td>
</tr>
<tr>
<td>9 Mutation operator;</td>
</tr>
<tr>
<td>10 Evaluate each FGCM’s fitness;</td>
</tr>
<tr>
<td>11 Check termination criteria;</td>
</tr>
<tr>
<td>end</td>
</tr>
<tr>
<td>13 Select best-fitness FGCM</td>
</tr>
</tbody>
</table>
More ideas

- FGCM learning and automatic construction with new algorithms
- FGCM in control systems
- FGCM in biomedical engineering
- FGCM in environmental control
- FGCM synaptic plasticity
- and so on
- . . .
International Network on Grey Systems and its Applications

Keynote talk: Machine Learning in Grey-Based Soft Computing techniques

Keynote speaker: Prof. Jose L. Salmeron
Universidad Pablo de Olavide (Seville, Spain)

September 23, 2015