Chapter 4
Extending Neuro-fuzzy Techniques with Grey-Based Hybridisation

Jose L. Salmeron

4.1 Introduction

Real-world problems are commonly composed by interrelated components in many complex ways. They are usually dynamic, that is, they change with time through a series of interactions among related components (Salmeron, 2012). Classical decision-making techniques cannot support these kinds of challenges. For that reason, this paper focused on the extension of neuro-fuzzy techniques as Fuzzy Cognitive Maps with Grey Systems Theory.

Fuzzy Cognitive Maps (FCMs) constitute neuro-fuzzy systems, which are able to include experts' knowledge (Kosko, 1986; Lee et al., 2002; Salmeron, 2009a, 2009b). From an Artificial Intelligence point of view, FCMs are supervised learning neural systems, whereas more and more data is available to model the problem, the system becomes more accurate at adapting itself and reaching a solution (Papageorgiou & Groumpos, 2005; Rodriguez-Repiso et al., 2007). FCM is a modelling technique for complex systems generated from mixing fuzzy logic and artificial neural networks (Salmeron, 2009a, 2009b).

FCMs have some benefits over classical modelling methods; they offer more information in the relationships between variables or concepts. Moreover, the FCM model is dynamic, and customizable and capture nonlinear relationships. The final fuzzy model is worthy for simulation, analysis and check the influence of variables and forecast the behaviour of the whole modelled system.

FCM have been used in many different fields as medicine (Papageorgiou, 2011; Salmeron, 2012), failure modes effects analysis (Pelaez & Bowles, 1995; Salmeron &

J. L. Salmeron (✉)
CUNEF Universidad, Madrid, Spain
e-mail: joseluis.salmeron@cunef.edu

Y. Yang and S. Liu (eds.), *Emerging Studies and Applications of Grey Systems*,
Series on Grey System, https://doi.org/10.1007/978-981-19-3424-7_4

Gutierrez, 2012) simulation (Fu, 1991), affective computing (Salmeron, 2012), engineering (Salmeron & Papageorgiou, 2014; Stylios & Groumpos, 2000), space exploration (Furfaro et al., 2010), relationship management in airline services (Kang et al., 2004), web-mining inference amplification (Lee et al., 2002) and others. For a deep review of the FCM research see (Papageorgiou & Salmeron, 2013).

4.2 Fuzzy Cognitive Maps

4.2.1 Theoretical Background

Fuzzy Cognitive Maps (Kosko, 1996) are a worthy technique for modelling and analysing the systems and people's behaviour. FCMs are a set of nodes linked by edges. The nodes model concepts relevant to a given problem. The causal links between these concepts are modelled by the edges showing the direction of the influence. The other attribute of an edge is its sign, which can be positive (a promoting effect) or negative (an inhibitory effect).

The main aim of a FCM around a problem is to be able to forecast the outcome by letting the critical issues interact with one another. These forecasts can be used for checking whether an initial scenario is consistent with the entire collection of stated causal assertions (Bueno & Salmeron, 2009). Causal relationships between the nodes can have different weights, represented in a range [−1, +1] (or [0, +1]). FCM substitutes the signs by a fuzzy value between −1 and +1 where the zero value indicates the absence of causality. Moreover, it involves feedback, where the impact of the change in a node may affect other nodes (Lopez & Salmeron, 2013) (Fig. 4.1).

The FCM nodes (c_i) would model such concepts such as heat, radiation, temperature or marketing strategy, among others. An edge linking two nodes models the

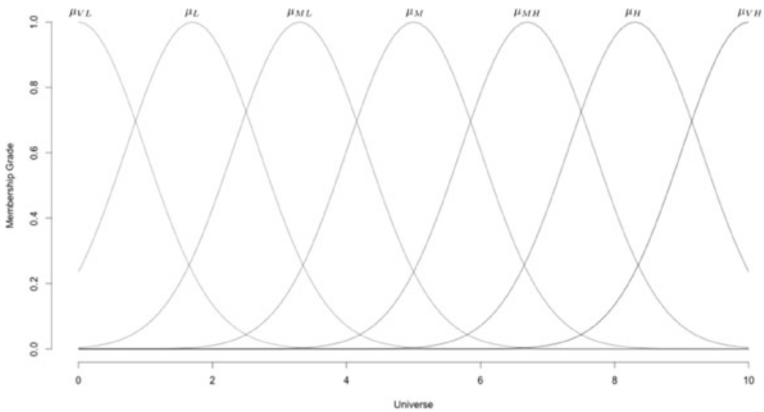

Fig. 4.1 Gaussian membership functions

causal influence of the causal node on the effect node (e.g.: the influence of the temperature to failure).

In a more formal way, an FCM can be represented as a 4-tuple;

$$\Omega = (\mathbf{N}, \mathbf{A}, f, r) \tag{4.1}$$

where N is the set of nodes $N = \{n_i\}_{i=1}^{m}$ with m the number of them, $A = [w_{ij}]_{n\times n}$ is the adjacency matrix modelling the edges between nodes as $\{w_{ij}\}_{i,j=1}^{m}$, $f(\cdot)$ the activation function, and r the nodes' range.

An adjacency matrix A represents the FCM nodes relationships. FCMs measure the strength of the causal relation between two nodes and if no causal relation be it is denoted by 0 in the adjacency matrix.

$$A = \begin{array}{c} C_1 \\ \vdots \\ C_n \end{array} \overset{\begin{array}{ccc} C_1 & \dots & C_n \end{array}}{\begin{pmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nn} \end{pmatrix}} \tag{4.2}$$

4.2.2 FCM Dynamic Analysis

FCMs are dynamical systems involving feedback, where the effect of change in a node may affect other nodes, which in turn can affect the node initiating the change. The analysis starts with the design of the initial vector state (c(0)), which models the initial value of each variable or concept (node) (Salmeron & Lopez, 2012). The initial vector state c(0) with n nodes is denoted as

$$c(0) = (c1(0),\ c2(0),\ \dots,\ cn(0)) \tag{4.3}$$

where $c_1(0)$ is the value of the concept $i = 1$ at instant $t = 0$.

The new states of the nodes are computed in an evolutive updating process with an activation function, which is applied to map monotonically the node state into a normalized range [0, 1] or [−1, +1]. The sigmoid function is the most used one (Bueno & Salmeron, 2009) when the node value maps in the range [0, 1]. The vector state $c(t + 1)$ at time $t + 1$ would be

$$\begin{aligned} c(t+1) &= f(c(t)\cdot A) \\ &= f(c_1(t)\cdot A), \dots, f(c_n(t)\cdot A)) \\ &= c_1(t+1)), \dots, (c_n(t+1) \end{aligned} \tag{4.4}$$

The element i of the vector state $c_i(t+1)$ at time $t+1$ can be computed as shown in Eqs. 4.5 and 4.6. Equation 4.5 computes $c_i(t+1)$ just with pre-synaptic computation.

$$\overbrace{C_i(t+1)}^{post-synaptic} = f\left(\overbrace{\sum_{j=1}^{n} W_{ji}.C_j(t)}^{pre-synaptic}\right) \tag{4.5}$$

Equation 4.6 computes the pre-synaptic influence and aggregate the state of the post-synaptic node in the previous iteration. In that situation, the node has memory of its former state.

$$\overbrace{C_i\ (t+1)}^{post-synaptic} = f\left(\overbrace{C_i(t)}^{memory} + \overbrace{\sum_{j=1}^{n} W_{ji}\ .C_j(t)}^{pre-synaptic}\right) \tag{4.6}$$

If $f(\cdot)$ is the unipolar sigmoid function and the updating rule is as Eq. 4.6, then the element i of the vector state c(t) at time t would be

$$C_i(t+1) = \frac{1}{1+e^{-\lambda.(C_i(t)+\sum_{j+1}^{n} W_{ji}.C_j(t))}} \tag{4.7}$$

where λ is the constant modelling the degree of normalization (function slope). If $f(\cdot)$ is the hyperbolic tangent function and Eq. 4.6 the updating rule, then the element i of the vector state c(t) at time t would be as follows

$$C_i(t+1) = \frac{e^2.\lambda.\left(c_i(t) + \sum_{j=1}^{n} w_{ji}.c_j(t)\right) - 1}{e^2.\lambda.\left(c_i(t) + \sum_{j=1}^{n} w_{ji}.c_j(t)\right) + 1} \tag{4.8}$$

After an inference process, the FCM reaches either one of two states after the iterations. It settles down to a fixed pattern of node values, the so-called hidden pattern or fixed-point attractor (as $c(t) \approx c(t-1)$).

Alternatively, it keeps cycling between several fixed states, known as a limit cycle. Using a continuous activation function, a third possibility known as a chaotic attractor exists. This occurs when, instead of stabilizing, the FCM keeps producing different state vector values for each iteration (Salmeron & Froelich, 2016).

4.2.3 FCM Consensus

It is possible to build FCMs from real-world raw data (e.g.: with evolutionary techniques) or from a panel of experts (Salmeron, 2009a, 2009b). Delphi methodology it is a methodology applied to manage the communication of a panel of experts in order to reach a consensus regarding a problem. One of the main features of Delphi method is when the experts get feedback reports, they could modify their own judgement based on the feedback (Bueno & Salmeron, 2009).

A more analytical way is the Augmented FCM. It doesn't require that experts change their initial judgement for consensus as Delphi method needs. The augmented adjacency matrix is built adding the adjacency matrix of each expert (Salmeron, 2009a, 2009b). Let k FCMs with common nodes. The augmented adjacency matrix (A*) would be computed as follows

$$A* = \begin{pmatrix} w^*_{11=\frac{1}{k}} \cdot \sum_{m=1}^{k} w_{11}.m & \cdots & \frac{1}{k} \cdot \sum_{m=1}^{k} w_{1n}.m \\ \vdots & \ddots & \vdots \\ w^*_{n1=\frac{1}{k}} \cdot \sum_{m=1}^{k} w_{1}.m & \cdots & \frac{1}{k} \cdot \sum_{m=1}^{k} w_{nn}.m \end{pmatrix} \quad (4.9)$$

If the FCMs have not common nodes, then the adjacency matrix would be added including zero rows and columns. Let a couple of FCMs ($\{w'_{ij}\}$ and $\{w''_{ij}\}$) without common nodes, the augmented adjacency matrix would be

$$A* = \begin{pmatrix} w^*_{11=\frac{1}{2}} . w'_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & & w^*_{nn=} \ \frac{1}{2} . w''_{nn} \end{pmatrix} \quad (4.10)$$

Figure 4.2c shows an example of Augmented FCM building process of a couple of FCMs with common and not common nodes. Figure 4.2a, b detail the FCMs from experts A and B and Fig. 4.2c is the Augmented FCM.

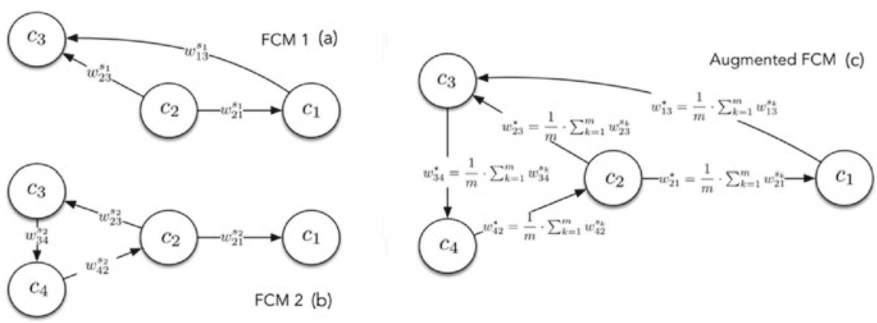

Fig. 4.2 Augmented FCM

4.3 Fuzzy Grey Cognitive Maps

Fuzzy Grey Cognitive Maps (FGCMs) are a FCM generalization (proposed in Salmeron (2010)) for modelling complex systems with high uncertainty, under discrete incomplete and small data sets and it is based on Grey Systems Theory (Deng, 1989).

4.3.1 Theoretical Background

4.3.1.1 Grey Systems Theory

Grey Systems Theory (GST) is a useful approach for solving problems with high uncertainty and discrete small and incomplete datasets (Liu & Lin, 2006). GST has been successfully applied in fields such as military science, agriculture, medicine, meteorology, business, industry, energy, transportation, geology and so on (Yamaguchi et al., 2007).

GST categorizes the information according to the degree of the available knowledge in three-fold. The system is white if the knowledge of that system is fully known (whole understanding), but it is black if the system is completely unknown. Finally, a system with partial knowledge known and partial knowledge unknown is called a grey system (Deng, 1989).

GST includes data fuzziness because it can consider it (Liu & Lin, 2006). Furthermore, the fuzzy theory holds information usually based on experience; while GST is concerned with objective data, they do not need any information different than the datasets to be disposed.

In a more formal way, let Θ be the universal set, then a grey set $\Phi \subset \Theta$ is defined by its both mappings. Note that $\left\{\mu_{\Phi}^{+}(\cdot), \mu_{\Phi}^{-}(\cdot)\right\} \in [0, 1]$, where $\mu_{\Phi}^{-}(\cdot)$ is the lower membership function, $\mu_{\Phi}^{+}(\cdot)$ is the upper one and $\mu_{\Phi}^{-}(\cdot) \leq \mu_{\Phi}^{+}(\cdot)$. Also, GST extends fuzzy logic, since the grey set Φ becomes a fuzzy set when $\mu_{\Phi}^{-}(\cdot) = \mu_{\Phi}^{+}(\cdot)$. The crisp value of a grey number is unknown, but the range in which the value is found is known.

An interval grey number is denoted as $\mathrm{x}^{\pm} \in [\mathrm{x}^{-}, \mathrm{x}^{+}] \mid \mathrm{x}^{-} \leq \mathrm{x}^{+}$ and it has an upper (x^{+}) and a lower (x^{-}) limit (Yang & John, 2012). Both limits are crisp numbers in first order interval grey numbers.

When the grey number $\mathrm{x}^{\pm}$ has just an upper limit is as follows $\mathrm{x}^{\pm} \in (-\infty, \mathrm{x}^{+}]$ and when it has just a lower limit is $\mathrm{x}^{\pm} \in [\mathrm{x}^{-}, +\infty)$. A black number has both unknown limits $\mathrm{x}^{\pm} \in (-\infty, +\infty)$ and it becomes a white number when both limits are the same $\mathrm{x}^{-} = \mathrm{x}^{+}$. The transformation of grey numbers in crisp ones $\widehat{x}$ is known as whitenisation (Liu & Lin, 2006) and it is computed as follows

$$\widehat{x} = \mathrm{x} + \cdot\xi + (1 - \xi) \cdot \mathrm{x} - |\xi \in [0, 1] \tag{4.11}$$

where ξ is a parameter to control the position of the crisp values according to the limits. If $\xi = 0.5$, then it is called equal mean whitenisation. The length of a grey number is computed as $l(x^{\pm}) = |x^{+} - x^{-}|$. In that sense, if the length of the grey number is zero ($l(x^{\pm}) = 0$), then it is a white number.

Despite the length of a grey number with only one limit known (lower or upper), $x^{\pm} \in [x^{-}, +\infty)$ or $x^{\pm} \in (-\infty, x^{+}]$ being infinite, the grey number is not a black number because it is possible to know one of the two limits.

Equations 4.12–4.15 compute the grey arithmetic operations

$$x\frac{+}{1} + x\frac{+}{2} \in \left[x_1^{-} + x_1^{-}, x_1^{+} + x_2^{+}\right] \tag{4.12}$$

$$x\frac{+}{1}\_x\frac{+}{2} \in \left[x_1^{-} + x_2^{-}, x_1^{+} + x_2\right] \tag{4.13}$$

$$x\frac{+}{1} \cdot x\frac{+}{2} \in [min\left\{x_1^{-} \cdot x_2^{-}, x_1^{+} \cdot x_2^{+}, x_1^{-} \cdot x_2^{+}, x_1^{+} \cdot x_2^{-}\right\}, \\ max\left\{x_1^{-} \cdot x_2^{-}, x_1^{+} \cdot x_2^{+}, x_1^{-} \cdot x_2^{+}, x_1^{+} \cdot x_2^{-}\right\}] \tag{4.14}$$

$$\frac{x\frac{+}{1}}{x\frac{+}{2}} \in \left[min\left\{\frac{x_1^{-}}{x_2^{-}}, \frac{x_1^{+}}{x_2^{+}}, \frac{x_1^{-}}{x_2^{+}}, \frac{x_1^{+}}{x_2^{+-}}\right\}, max\left\{\frac{x_1^{-}}{x_2^{-}}, \frac{x_1^{+}}{x_2^{+}}, \frac{x_1^{-}}{x_2^{+}}, \frac{x_1^{+}}{x_2^{+-}}\right\}\right] \left|\left\{x_i^{-}, x_i^{+}\right\}_{i=1}{}^{2} Y \neq 0\right| \tag{4.15}$$

4.3.1.2 FGCM Fundamentals

The FGCM nodes are concepts or variables and the relationships between them are modelled by grey weighted and directed edges (Nápoles et al., 2021; Salmeron, 2010). Formally, a FGCM is denoted as a 4-tuple

$$\Omega = \left(\zeta,\ W,\ f^{\pm}(\cdot),\ l(\psi)\right) \tag{4.16}$$

where $\zeta = \{\alpha_i^{\pm}\}_{i=1}^{n}$ is the set of nodes with grey numbers as grey states, $W = \{w_{ij}^{\pm}\}$ is the set of edges linking the nodes weighted with grey weights, $f^{\pm}(\cdot)$ the grey activation function and $l(\psi)$ the range of the information space.

An interval grey weight between the nodes $\alpha_i^{\pm}$ and $\alpha_j^{\pm}$ is denoted as $w_{ij}^{\pm} \in [w_{ij}^{-}, w_{ij}^{+}]$ and it has a lower (w_{ij}^{-}) and an upper (w_{ij}^{+}) limit.

FGCM includes greyness as an uncertainty score (Salmeron, 2010; Salmeron & Gutierrez, 2012; Salmeron & Papageorgiou, 2014). If a grey state (or weight) has a high score of greyness, then it means that the grey state have a high uncertainty associated. It is computed as follows

$$\emptyset\left(\alpha\frac{+}{i}\right) = \frac{\left|l(\alpha\frac{+}{i})\right|}{l(\psi)} \tag{4.17}$$

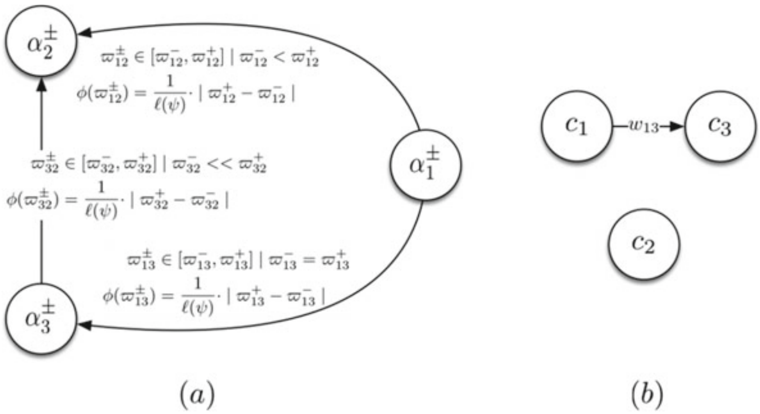

Fig. 4.3 Potential relationships of the nodes in FGCM (**a**) and FCM (**b**)

where $|l(\alpha_i^{\pm})| = |\alpha_i^{+} - \alpha_i^{-}|$ is the length of grey node state $\alpha_i^{\pm}$ in absolute value and $l(\psi)$ is the length of the information space, denoted by ψ. It is computed as follows

$$l(\psi) = \begin{cases} 1 if \left\{\alpha_i^{\pm}, \varpi_i^{\pm}\right\} \subseteq [0, 1] \\ 2 if \left\{\alpha_i^{\pm}, \varpi_i^{\pm}\right\} \subseteq [-1, +1] \end{cases} \tag{4.18}$$

where $l(\psi) = 1$ for unipolar FGCMs and $l(\psi) = 2$ for bipolar FGCMs.

FGCMs model the human reasoning in a more realistic way than FCM does, because it is able to manage the uncertainty between the concepts and the incomplete information about the modelled system as FCMs cannot do (Salmeron & Palos-Sanchez, 2019).

In Fig. 4.3a, the relationship ($\varpi_{13}^{\pm}$) between $\alpha_1^{\pm}$ and $\alpha_3^{\pm}$ is a white one, between $\alpha_1^{\pm}$ and $\alpha_2^{\pm}$ is a grey one ($w_{12}^{\pm}$), and between $\alpha_3^{\pm}$ and $\alpha_1^{\pm}$ is a black one ($w_{32}^{\pm}$).

FCMs are just able to model white (crisp) relationships (Fig. 4.3b). As a result, a huge amount of information is lost. Another advantage of FGCMs over FCM is related with the uncertainty associated to the elements of the problem (or system) modelled (Salmeron & Papageorgiou, 2012).

4.3.2 FGCM Dynamic Analysis

FGCMs are dynamical systems involving feedback just as FCMs are, where the influence of change in a grey node may affect other grey nodes, which in turn can impact the grey node starting the dynamics. The dynamic analysis begins with an initial grey vector state ($\alpha^{\pm}(0)$), which models the initial grey value of each concept (grey node). The initial grey vector state $\alpha \pm (0)$ with n grey nodes is denoted as

$$\alpha^{\pm}(0) = \left(\alpha^{\pm}(0),\ \alpha^{\pm}(0), \ldots, \alpha^{\pm}(0)\right) \tag{4.19}$$

where $\alpha^{\pm}(0)$ is the state of the concept i = 1 at time t = 0.

The grey state values of the grey nodes are updated in an evolutive process with a grey activation function, which is applied to map monotonically the grey node state into the information space (Salmeron, 2010). The grey state is updated as follows

$$\begin{aligned}\alpha^{\pm} &= f^{\pm}\left(\alpha^{\pm}(t) \cdot A^{\pm}\right)\\ &= \left(f^{\pm}\left(\alpha_1^{\pm}(t) \cdot A^{\pm}\right), \ldots, f^{\pm}\left(\alpha_n^{\pm}(t) \cdot A^{\pm}\right)\right)\\ &= \left(\alpha_1^{\pm}(t+1), \ldots, \alpha_n^{\pm}(t+1)\right)\end{aligned} \tag{4.20}$$

where $A^{\pm}$ is the grey adjacency matrix and $f^{\pm}(\cdot)$ the grey activation function. Mostly, the grey activation function used to be the grey unipolar sigmoid (Eq. 4.21)

$$\alpha\frac{+}{i}(t+1) = \left[\frac{1}{1+e^{-\lambda.\left(a\frac{+}{i}(t)+\sum_{j=1}^{n} w\frac{+}{ji}.\alpha\frac{+}{j}(t)\right)}}, \frac{1}{1+e^{-\lambda.\left(a\frac{+}{i}(t)+\sum_{j=1}^{n} w\frac{+}{ji}.\alpha\frac{+}{j}(t)\right)}}\right] \tag{4.21}$$

or the grey hyperbolic tangent (Eq. 4.22) as follows

$$\alpha\frac{+}{i}(t+1) = \left[\frac{e^{2.\lambda.\left(a\frac{+}{i}(t)+\sum_{j=1}^{n} w_{ji}.\alpha\frac{+}{j}(t)\right)}-1}{e^{2.\lambda.\left(a\frac{+}{i}(t)+\sum_{j=1}^{n} w_{ji}.\alpha\frac{+}{j}(t)\right)}+1}, \frac{e^{2.\lambda.\left(a\frac{+}{i}(t)+\sum_{j=1}^{n} w_{ji}.\alpha\frac{+}{j}(t)\right)}-1}{e^{2.\lambda.\left(a\frac{+}{i}(t)+\sum_{j=1}^{n} w_{ji}.\alpha\frac{+}{j}(t)\right)}+1}\right] \tag{4.22}$$

4.3.3 FGCM Consensus

It is possible to build FGCMs from real-world raw data (e.g.: using evolutionary techniques) or from a panel of experts. As far as the author knows, there have been not used Delphi methodology for building FGCMs yet, but augmented FGCMs have been used. The augmented adjacency grey matrix is built adding the adjacency grey matrix of each expert (Salmeron & Palos-Sanchez, 2019).

Let k FGCMs with common grey nodes. The augmented adjacency grey matrix ($A^{\pm *}$) would be computed as follows

$$A^{\pm *} = \begin{pmatrix} \varpi_{11}^{*} = \frac{1}{k}\cdot\sum_{m=1}^{k}\varpi\frac{+}{n1}.m & \cdots & \varpi\frac{*}{1n} = \frac{1}{k}\cdot\sum_{m=1}^{k}\varpi\frac{+}{1n}.m \\ \vdots & \ddots & \vdots \\ \varpi_{n1}^{*} = \frac{1}{k}\cdot\sum_{m=1}^{k}\varpi\frac{+}{n1}.m & \ldots & \varpi\frac{*}{nn} = \frac{1}{k}\cdot\sum_{m=1}^{k}\varpi\frac{+}{nn}.m \end{pmatrix} \tag{4.23}$$

If the FGCMs have not common grey nodes, then the adjacency grey matrix would be added including null rows and columns. Let a couple of FGCMs $\{w_{ij}{}^{\pm(a)}\}$ and $\{w_{ij}{}^{\pm(b)}\}$ without common grey nodes, the augmented adjacency grey matrix would be

$$A^{\pm *} = \begin{pmatrix} \varpi_{11}^{+(a)*} = \frac{1}{2}.\varpi_{11}^{+(a)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \varpi_{nn}^{+(b)*} = \frac{1}{2}.\varpi_{nn}^{+(b)} \end{pmatrix} \tag{4.24}$$

4.4 Conclusions

This work shows the extension of neuro-fuzzy techniques with Grey Systems Theory. FGCMs are a generalization for modelling systems with high uncertainty, under discrete incomplete and small data sets.

FGCMs model the human reasoning in a more realistic way than FCM does, because it is able to process the uncertainty between the nodes and the incomplete information about the modelled problem as FCMs cannot do.

References

Bueno, S., & Salmeron, J. L. (2009). Benchmarking main activation functions in fuzzy cognitive maps. *Expert Systems with Applications, 36*(3 Part 1), 5221–5229.

Deng, J. (1989). Introduction to grey system theory. *The Journal of Grey System, 1*(1), 1–24.

Fu, L. (1991). Causim: A rule-based causal simulation system. *Simulation, 56*(4).

Furfaro, R., Kargel, J. S., Lunine, J. I., Fink, W., & Bishop, M. P. (2010). Identification of cryovolcanism on Titan using fuzzy cognitive maps. *Planetary and Space Science, 58*(5), 761–779.

Kang, I., Sangjae, L., & Choi, J. (2004). Using fuzzy cognitive map for the relationship management in airline service. *Expert Systems with Applications, 26*, 545–555.

Kosko, B. (1986). Fuzzy cognitive maps. *International Journal on Man-Machine Studies, 24*, 65–75.

Kosko, B. (1996). *Fuzzy engineering*. Prentice-Hall.

Lee, K. C., Kim, J. S., Chung, H. N., & Kwon, S. J. (2002). Fuzzy cognitive map approach to web-mining inference amplification. *Expert Systems with Applications, 22*, 197–211.

Liu, S., & Lin, Y. (2006). *Grey information*. Springer.

Lopez, C., & Salmeron, J. L. (2013). Dynamic risks modelling in erp maintenance projects with fcm. *Information Sciences, 256*, 25–45.

Nápoles, G., Salmeron, J. L., & Vanhoof, K. (2021). Construction and supervised learning of long-term grey cognitive networks. *IEEE Transactions on Cybernetics, 51*(2), 686–695.

Papageorgiou, E. (2011). A new methodology for decisions in medical informatics using fuzzy cognitive maps based on fuzzy rule-extraction techniques. *Applied Soft Computing, 11*(1), 500–513.

Papageorgiou, E., & Groumpos, P. (2005). A weight adaptation method for fine-tuning fuzzy cognitive map causal links. *Soft Computing Journal, 9*, 846–857.

Papageorgiou, E., & Salmeron, J. L. (2013). A review of fuzzy cognitive map research at the last decade. *IEEE Transactions on Fuzzy Systems, 21*(1), 66–79.

Pelaez, C., & Bowles, J. (1995). Applying fuzzy cognitive maps knowledge representation to failure models effects analysis. In *IEEE Annual Reliability and Maintainability Symposium.*

Rodriguez-Repiso, L., Setchi, R., & Salmeron, J. L. (2007). Modelling it projects success with fuzzy cognitive maps. *Expert Systems with Applications, 32*, 543–559.

Salmeron, J. L. (2009a). Augmented fuzzy cognitive maps for modelling LMS critical success factors. *Knowledge-Based Systems, 22*(4), 275–278.

Salmeron, J. L. (2009b). Supporting decision makers with fuzzy cognitive maps. *Research-Technology Management, 52*(3), 7581–7588.

Salmeron, J. L. (2010). Modelling grey uncertainty with fuzzy grey cognitive maps. *Expert Systems with Applications, 37*(12), 7581–7588.

Salmeron, J. L. (2012). Fuzzy cognitive maps for artificial emotions forecasting. *Applied Soft Computing, 12*(12), 3704–3710.

Salmeron, J. L., & Froelich, W. (2016). Dynamic optimization of fuzzy cognitive maps for time series forecasting. *Knowledge-Based Systems, 105*, 29–37.

Salmeron, J. L., & Gutierrez, E. (2012). Fuzzy grey cognitive maps in reliability engineering. *Applied Soft Computing, 12*(12), 3818–3824.

Salmeron, J. L., & Lopez, C. (2012). Forecasting risk impact on erp maintenance with augmented fuzzy cognitive maps. *IEEE Transactions on Software Engineering, 38*(2), 439–452.

Salmeron, J. L., & Palos-Sanchez, P. R. (2019). Uncertainty propagation in Fuzzy Grey Cognitive Maps with Hebbian-like learning algorithms. *IEEE Transactions on Cybernetics, 49*(1), 211–220.

Salmeron, J. L., & Papageorgiou, E. L. (2012). A fuzzy grey cognitive maps-based decision support system for radiotherapy treatment planning. *Knowledge-Based Systems, 30*(1), 151–160.

Salmeron, J. L., & Papageorgiou, E. (2014). Fuzzy grey cognitive maps and nonlinear Hebbian learning in process control. *Applied Intelligence, 41*(1), 223–234.

Stylios, C. D., & Groumpos, P. P. (2000). Fuzzy cognitive maps in modeling supervisory control systems. *Journal of Intelligent & Fuzzy Systems, 8*(2), 83–98.

Yamaguchi, D., Li, G., Chen, L., & Nagai, M. (2007). Reviewing crisp, fuzzy, grey and rough mathematical models. In IEEE (Ed.), *Proceedings of the IEEE International Conference on Grey Systems and Intelligent Services* (pp. 547–552).

Yang, Y., & John, R. (2012). Grey sets and greyness. *Information Sciences, 185*(1), 249–264.