

A First Study of Fuzzy Cognitive Maps Learning Using Particle Swarm Optimization

Konstantinos E. Parsopoulos

Department of Mathematics,
University of Patras Artificial Intelligence
Research Center (UPAIRC),
University of Patras,
GR-26110 Patras, Greece
kostasp@math.upatras.gr

Elpiniki I. Papageorgiou

Department of Electrical and Computer Engineering,
University of Patras Artificial Intelligence
Research Center (UPAIRC),
University of Patras,
GR-26500 Patras, Greece
epapageo@ee.upatras.gr

Petros P. Groumpos

Department of Electrical and Computer Engineering,
University of Patras Artificial Intelligence
Research Center (UPAIRC),
University of Patras,
GR-26500 Patras, Greece
groumpos@ee.upatras.gr

Michael N. Vrahatis

Department of Mathematics,
University of Patras Artificial Intelligence
Research Center (UPAIRC),
University of Patras,
GR-26110 Patras, Greece
vrahatis@math.upatras.gr

Abstract- In this paper a new algorithm for Fuzzy Cognitive Maps learning is introduced. The proposed approach is based on the Particle Swarm Optimization method and it is used for the detection of proper weight matrices that lead the Fuzzy Cognitive Map to desired steady states. For this purpose a properly defined objective function that incorporates experts' knowledge is constructed and minimized. The application of the proposed methodology to an industrial control problem supports the claim that the proposed technique is efficient and robust.

1 Introduction

Fuzzy Cognitive Maps (FCMs) are a soft computing methodology developed by Kosko as an expansion of cognitive maps which are widely used to represent social scientific knowledge [4, 20]. They belong to the class of neuro-fuzzy systems, which are able to incorporate human knowledge and adapt it through learning procedures [22]. FCMs are designed by experts through an interactive procedure of knowledge acquisition [15], and they have a wide field of application, including modelling of complex and intelligent systems [16], decision analysis [19], and extend graph behavior analysis [15]. They have also been used for planning and decision-making in the fields of international relations and social systems modelling [43, 44], as well as in management science, operations research and organizational behavior [9]. Dicherson and Kosko have used FCMs to construct virtual worlds [10]. Furthermore, FCMs have been proposed for modelling supervisory systems [14] and for decision-making in radiation therapy planning systems [30, 31].

The wide recognition of FCMs as a promising modelling and simulation methodology for complex systems, characterized by abstraction, flexibility and fuzzy reasoning, promoted the research on new concepts in this area. How-

ever, the established developments still require enhancement, stronger mathematical justification, and further testing on systems of higher complexity. Moreover, the elimination of deficiencies, such as the abstract estimation of the initial weight matrix and the dependence on the subjective reasoning of experts' knowledge, will significantly improve FCMs' functionality. In this context, the development of learning algorithms is a stimulating research topic.

A few algorithms have been proposed for FCM learning [22, 29]. The main task of the learning procedure is to find a setting of the FCM's weights, that leads the FCM to a desired steady state. This is achieved through the minimization of a properly defined objective function. Established algorithms are heavily dependent on the initial weight matrix approximation, which is provided by the experts. Recently, preliminary results on a different approach, based on Evolution Strategies, have been reported [23].

This paper proposes, a new approach for FCM learning, which is based on the Particle Swarm Optimization (PSO) method. PSO is used for the determination of proper weight matrices for the system, through the minimization of a properly defined objective function. PSO is selected due to its efficiency and effectiveness on a plethora of applications in science and engineering [1, 2, 7, 13, 24, 25, 27, 28, 32, 34, 35, 36, 37, 38, 45], and its straightforward applicability. The proposed approach is illustrated on an industrial process control problem, with promising results.

The rest of the paper is organized as follows: in Section 2 the main principles underlying FCMs are described. In Section 3, the PSO algorithm is briefly presented; Section 4 is devoted to the description and analysis of the proposed learning algorithm. The process control problem, on which the proposed algorithm is tested, is described in Section 5, while the obtained results are reported and discussed in Section 6. Section 7 closes the paper, with conclusions and ideas for future research.

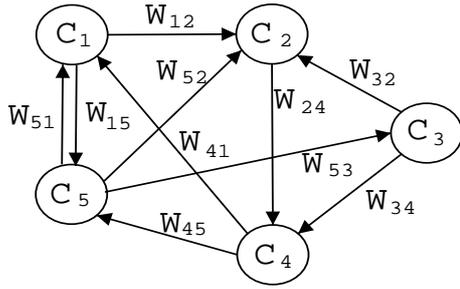


Figure 1: A simple Fuzzy Cognitive Map.

2 Overview of Fuzzy Cognitive Maps

FCMs have been introduced by Kosko in 1986 [20] as signed directed graphs for representing causal reasoning and computational inference processing, exploiting a symbolic representation for the description and modelling of a system. Concepts are utilized to represent different aspects of the system, as well as, their behavior. The dynamics of the system are implied by the interaction of concepts. FCM structures are used to represent both qualitative and quantitative data. The construction of an FCM requires the input of human experience and knowledge on the system under consideration. Thus, FCMs integrate the accumulated experience and knowledge concerning the underlying causal relationships amongst factors, characteristics, and components that constitute the system.

An FCM consists of nodes—concepts, C_i , $i = 1, \dots, N$, where N is the total number of concepts. Each node—concept, represents one of the key—factors of the system, and it is characterized by a value $A_i \in [0, 1]$, $i = 1, \dots, N$. The concepts are interconnected through weighted arcs, which imply the relations among them. A simple FCM with five nodes and ten weighted arcs is illustrated in Fig. 1. Each interconnection between two concepts C_i and C_j , has a weight W_{ij} , which is proportional to the strength of the causal link between C_i and C_j . The sign of W_{ij} indicates whether the relation between the two concepts is direct or inverse. The direction of causality indicates whether the concept C_i causes the concept C_j or vice versa. Thus, there are three types of weights:

$$\begin{cases} W_{ij} > 0, & \text{expresses positive causality,} \\ W_{ij} < 0, & \text{expresses negative causality,} \\ W_{ij} = 0, & \text{expresses no relation.} \end{cases}$$

Human knowledge and experience on the system determines the type and the number of nodes, as well as the initial weights of the FCM. The value A_i , of a concept C_i , expresses the quantity of its corresponding physical value and is derived by the transformation of the fuzzy values assigned by the experts, to numerical values. Having assigned values to the concepts and the weights, the FCM converges to a steady state, through the interaction process subsequently described.

At each step, the value A_i of a concept is influenced by the values of concepts—nodes connected to it, and is updated

according to the scheme [22]:

$$A_i(k+1) = f \left(A_i(k) + \sum_{\substack{j=1 \\ j \neq i}}^n W_{ji} A_j(k) \right), \quad (1)$$

where k stands for the iteration counter; and W_{ji} is the weight of the arc connecting concept C_j to concept C_i . The function f is the sigmoid function:

$$f(x) = \frac{1}{1 + e^{-\lambda x}}, \quad (2)$$

where $\lambda > 0$ is a parameter that determines its steepness in the area around zero. In our approach, the value $\lambda = 1$ has been used. This function is selected since the values A_i of the concepts, by definition, must lie within $[0, 1]$. The interaction of the FCM results after a few iterations in a steady state, i.e. the values of the concepts are not modified further. Desired values of the output concepts of the FCM guarantee the proper operation of the simulated system.

The design of an FCM is a process that heavily relies on the input from experts [42]. At the beginning, experts are pooled to determine the relevant factors that will be represented in the map as concepts. Then, each expert describes the causal relationships among the concepts using a linguistic notion. First, experts determine the influence of a concept on another, as “negative”, “positive” or “no influence”. Then, linguistic weights, such as “strong”, “weak”, etc. are assigned to each arc. The linguistic variables that describe each arc, for each expert, are defined in [8]. The linguistic variables are combined, and the aggregated linguistic variable is transformed to a single linguistic weight, through the SUM technique [26]. Finally, the Center of Area (CoA) defuzzification method [21, 26], is used for the transformation of the linguistic weight to a numerical value within the range $[-1, 1]$. This methodology has the advantage that experts are not required to assign directly numerical values to causality relationships, but rather to describe qualitatively the degree of causality among the concepts. Thus, an initial matrix $W^{\text{initial}} = [W_{ij}]$, $i, j = 1, \dots, N$, with $W_{ii} = 0$, $i = 1, \dots, N$, is obtained. Using the initial concept values, A_i , which are also provided by the experts, the matrix W^{initial} is used for the determination of the steady state of the FCM, through the application of the rule of Eq. (1).

The critical dependence on the opinions of the experts and the potential convergence to undesired steady states, are the two most significant weaknesses of FCMs. Learning procedures constitute means to increase the efficiency and robustness of FCMs, by updating the weight matrix so as to avoid convergence to undesired steady states. Up-to-date, there are just a few FCM learning algorithms and they are mostly based on ideas coming from the field of artificial neural networks training [3, 10, 29]. Such algorithms start from an initial state and an initial weight matrix, W^{initial} , of the FCM, and adapt the weights, in order to compute a weight matrix that leads the FCM to a desired steady state. The desired steady state is characterized by values of the FCM’s output concepts accepted by the experts, *ex post*.

The main drawback of this approach is the heavy dependence of the final weights on the initial weight matrix.

A novel learning procedure that alleviates the problem of the potential convergence to an undesired steady state, is proposed in this paper. This approach is based on a swarm intelligence algorithm which is briefly presented in the next section.

3 The Particle Swarm Optimization Method

Particle Swarm Optimization (PSO) is a stochastic optimization algorithm. More specifically, it belongs to the class of *Swarm Intelligence* algorithms, which are inspired from the social dynamics and emergent behavior that arise in socially organized colonies [5, 17, 18, 35].

PSO is a population based algorithm, i.e., it exploits a population of individuals to probe promising regions of the search space. In this context, the population is called *swarm* and the individuals (i.e., the search points) are called *particles*. Each particle moves with an adaptable velocity within the search space, and retains a memory of the best position it ever encountered. In the *global* variant of PSO, the best position ever attained by all individuals of the swarm is communicated to all the particles. In the *local* variant, each particle is assigned to a topological neighborhood consisting of a prespecified number of particles. In this case, the best position ever attained by the particles that comprise the neighborhood is communicated among them [12, 18].

Assume a D -dimensional search space, $S \subset \mathbb{R}^D$, and a swarm consisting of N particles. The i -th particle is in effect a D -dimensional vector $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})^T \in S$. The velocity of this particle is also a D -dimensional vector, $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})^T \in S$. The best previous position encountered by the i -th particle is a point in S , denoted by $P_i = (p_{i1}, p_{i2}, \dots, p_{iD})^T \in S$. Assume g_i to be the index of the particle that attained the best previous position among all the particles in the neighborhood of the i -th particle, and t to be the iteration counter. Then, the swarm is manipulated by the equations [6, 11, 39, 40]:

$$V_i(t+1) = \chi \left[wV_i(t) + c_1 r_1 (P_i(t) - X_i(t)) + c_2 r_2 (P_{g_i}(t) - X_i(t)) \right], \quad (3)$$

$$X_i(t+1) = X_i(t) + V_i(t+1), \quad (4)$$

where $i = 1, \dots, N$; c_1 and c_2 are two parameters called *cognitive* and *social* parameters respectively; r_1, r_2 , are random numbers uniformly distributed within $[0, 1]$; and g_i is the index of the particle that attained either the best position of the whole swarm (global version), or the best position in the neighborhood of the i -th particle (local version). The parameters χ and w are called *constriction factor* and *inertia weight* respectively, and they are used as mechanisms for the control of the velocity's magnitude, corresponding to the two main PSO versions.

The value of the constriction factor is derived analytically [6, 46]. On the other hand, the inertia weight, w , is computed empirically, taking into consideration that large values encourage global exploration, while small values

promote local exploration. According to a rule of thumb, an initial value of w around 1.0 and a gradual decline towards 0 is considered a proper choice [39, 40].

In general, the constriction factor version of PSO is faster than the one with the inertia weight, although in some applications its global variant suffers from premature convergence. Regarding the social and cognitive parameter, the default values $c_1 = c_2 = 2$ have been proposed. The initialization of the swarm and the velocities, is usually performed randomly and uniformly in the search space, although more sophisticated initialization techniques can enhance the overall performance of the algorithm [33].

4 The Proposed Approach

The present work focuses on the development of an FCM learning procedure based on PSO. The purpose is to determine the values of the cause–effect relationships among the concepts, i.e. the values of the weights of the FCM, that produce a desired behavior of the system. The determination of the weights is of major significance and it contributes towards the establishment of FCMs as a robust methodology. The desired behavior of the system is characterized by output concept values that lie within desired bounds prespecified by the experts. These bounds are in general problem dependent.

The learning procedure is, to some extent, similar to that of neural networks training. Let C_1, \dots, C_N , be the concepts of an FCM, and let $C_{out_1}, \dots, C_{out_m}$, $1 \leq m \leq N$, be the output concepts, while the remaining concepts are considered input, or interior, concepts. The user is interested in restricting the values of the output concepts in strict bounds:

$$A_{out_i}^{\min} \leq A_{out_i} \leq A_{out_i}^{\max}, \quad i = 1, \dots, m,$$

predetermined by the experts, which are crucial for the proper operation of the modelled system. Thus, the main goal is to detect a weight matrix, $W = [W_{ij}]$, $i, j = 1, \dots, N$, that leads the FCM to a steady state at which, the output concepts lie in their corresponding bounds, while the weights retain their physical meaning. The latter is attained by imposing constraints on the potential values assumed by weights. To do this, we consider the following objective function:

$$F(W) = \sum_{i=1}^m H(A_{out_i}^{\min} - A_{out_i}) |A_{out_i}^{\min} - A_{out_i}| + \sum_{i=1}^m H(A_{out_i} - A_{out_i}^{\max}) |A_{out_i}^{\max} - A_{out_i}|, \quad (5)$$

where H is the well-known Heaviside function

$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0, \end{cases}$$

and A_{out_i} , $i = 1, \dots, m$, are the steady state values of the output concepts, that are obtained through the application of the procedure of Eq. (1), using the weight matrix W . Obviously, the global minimizers of the objective function F ,

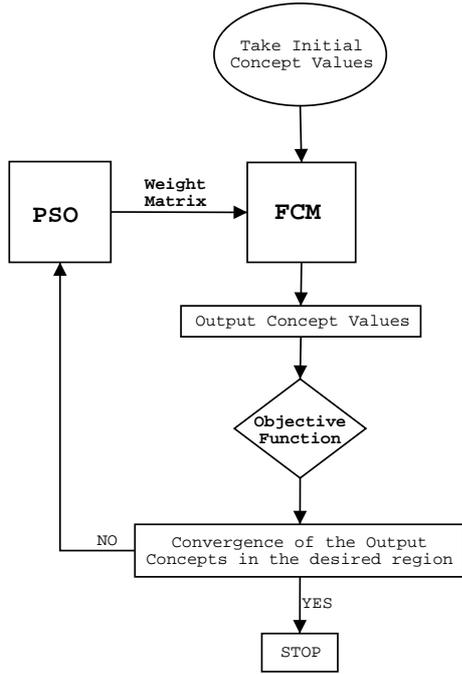


Figure 2: Flowchart of the proposed learning procedure.

are weight matrices that lead the FCM to a desired steady state, i.e. all output concepts are bounded within the desired regions. The objective function F suits straightforwardly the problem, however, it is non-differentiable and, thus, gradient-based methods are not applicable for its minimization. On the other hand, in the proposed approach, PSO is used for the minimization of the objective function defined by Eq. (5). The non-differentiability of F poses no problems in our approach since PSO, like all evolutionary algorithms, requires function values solely, and can be applied even on discontinuous functions. The weight matrix W is represented by a vector which consists of the rows of W in turn, excluding the elements of its main diagonal, $W_{11}, W_{22}, \dots, W_{NN}$, which are by definition equal to zero:

$$X = \left[\underbrace{W_{12}, \dots, W_{1N}}_{\text{row 1}}, \underbrace{W_{21}, \dots, W_{2N}}_{\text{row 2}}, \dots, \underbrace{W_{N1}, \dots, W_{N,N-1}}_{\text{row N}} \right].$$

Thus, an FCM with N fully interconnected concepts (i.e. each concept interacts with all other concepts), corresponds to a $N(N-1)$ -dimensional minimization problem. If some interconnections are missing, then their corresponding weights are zero and they can be omitted, reducing the dimensionality of the problem. This is most often the case, since the FCMs provided by experts are rarely fully connected.

Each interconnection of an FCM has a specific physical meaning, and, thus, several constraints are posed by the experts on the values of the weights. Constraints are provided in the form of negative or positive relations between two concepts. So, if two concepts C_i and C_j are negatively related, then the weight $W_{ij} \in [-1, 0]$, while if they are positively related, it takes values within $[0, 1]$. More strict constraints may be additionally posed on some weights, either by the experts, or by taking into consideration the conver-

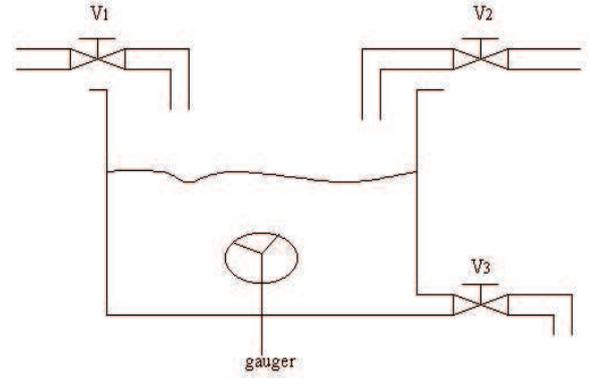


Figure 3: Illustration of a process control problem from industry.

gence regions obtained through the application of the learning algorithm, as illustrated in Section 6. Such constraints may enhance the overall performance of the algorithm.

The application of PSO for the minimization of the objective function F , starts with an initialization phase, where a swarm, $\mathbb{S} = \{X_1, \dots, X_M\}$, of size M , is generated randomly, and it is evaluated using F . Then, the Eqs. (3) and (4) are used to evolve the swarm. As soon as a weight configuration that globally minimizes F is reached, the algorithm is terminated. A flowchart of this procedure is depicted in Fig. 2.

There is, in general, a plethora of weight matrices that lead to convergence of the FCM to the desired regions of the output concepts. PSO is a stochastic algorithm, and, thus, it is quite natural to obtain such suboptimal matrices which differ in subsequent experiments. All these matrices are proper for the design of the FCM and follow the constraints of the problem, though, each matrix may have different physical meaning for the system. Statistical analysis of the obtained weight matrices may help in the better understanding of the system's dynamic, as it is implied by the weights, as well as in the selection of the most appropriate suboptimal matrix. Any information available *a priori*, may be incorporated to enhance the procedure, either by modifying the objective function in order to exploit the available information, or by imposing further constraints on the weights. The proposed approach has proved to be very efficient in practice. In the following section, its operation on an industrial process control problem, is illustrated.

5 An Industrial Process Control Problem

A simple process control problem encountered in chemical industry, is selected to illustrate the workings of the proposed learning algorithm [41]. The process control problem, illustrated in Fig. 3, consists of one tank and three valves that influence the amount of a liquid in the tank. Valve 1 and Valve 2 pour two different liquids into the tank. During the mixing of the two liquids, a chemical reaction takes place in the tank, and a new liquid is produced. Valve 3 empties the tank when the new liquid produced reaches a specific level. A sensor is placed inside the tank to measure

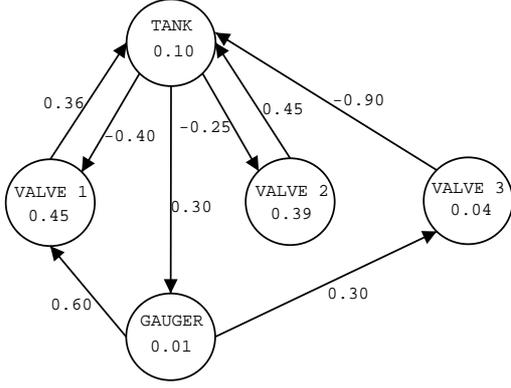


Figure 4: The FCM that corresponds to the problem of Fig. 3.

the specific gravity of the produced liquid. When the value, G , of the specific gravity lies in a range $[G_{\min}, G_{\max}]$, the desired liquid has been produced. There is also a limit on the height, T , of the liquid in the tank, i.e. it cannot exceed a lower limit T_{\min} and an upper limit T_{\max} . The control target is to keep the two variables, T and G , within their bounds:

$$\begin{aligned} T_{\min} &\leq T \leq T_{\max}, \\ G_{\min} &\leq G \leq G_{\max}. \end{aligned}$$

A group of experts constructs the FCM for the simulation of this system, following the procedure described in Section 2. The FCM that models and controls the specific system is depicted in Fig. 4. It consists of five concepts which are defined as:

- Concept 1 – the amount of the liquid in the tank. It depends on the operational state of Valves 1, 2 and 3;
- Concept 2 – the state of Valve 1 (closed, open or partially opened);
- Concept 3 – the state of Valve 2 (closed, open or partially opened);
- Concept 4 – the state of Valve 3 (closed, open or partially opened);
- Concept 5 – the specific gravity of the produced liquid in the tank.

There is a consensus among the experts regarding the direction of the arcs among the concepts. For each weight, the overall linguistic variable and its corresponding fuzzy set, are also determined by the experts. The ranges of the weights implied by the fuzzy regions, are:

$$-0.50 \leq W_{12} \leq -0.30, \quad (6)$$

$$-0.40 \leq W_{13} \leq -0.20, \quad (7)$$

$$0.20 \leq W_{15} \leq 0.40, \quad (8)$$

$$0.30 \leq W_{21} \leq 0.40, \quad (9)$$

$$0.40 \leq W_{31} \leq 0.50, \quad (10)$$

$$-1.0 \leq W_{41} \leq -0.80, \quad (11)$$

$$0.50 \leq W_{52} \leq 0.70, \quad (12)$$

$$0.20 \leq W_{54} \leq 0.40, \quad (13)$$

and the initial weight matrix, derived through the CoA defuzzification method, is:

$$W^{\text{initial}} = \begin{pmatrix} 0 & -0.4 & -0.25 & 0 & 0.3 \\ 0.36 & 0 & 0 & 0 & 0 \\ 0.45 & 0 & 0 & 0 & 0 \\ -0.90 & 0 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0.3 & 0 \end{pmatrix}.$$

All experts agreed on the same range for the weights W_{21} , W_{31} , and W_{41} , and most of them agreed on the same range for the weights W_{12} and W_{13} . However, there was no such agreement on the cases of the weights W_{15} , W_{52} , and W_{54} , where their opinions varied significantly.

PSO is applied to update the eight nonzero weight values of the FCM. To avoid physically meaningless weights, the bounds $[-1, 0]$ or $[0, 1]$, implied by the directions of the corresponding arcs of the FCM, are imposed on each weight.

The output concepts for this problem are the concepts C_1 and C_5 . The desired regions for the two output concepts, which are crucial for the proper operation of the modelled system, have been defined by the experts:

$$0.68 \leq C_1 \leq 0.70, \quad (14)$$

$$0.78 \leq C_5 \leq 0.85. \quad (15)$$

In the next section, the simulation results are reported and analyzed.

6 Simulation Results

The behavior of the system accepting all the weights' constraints imposed by the experts has been investigated. The obtained results are very interesting and provide insight regarding the appropriateness of the experts' suggestions as well as suboptimal weight matrices that lead the FCM to the desired steady state.

A total of 100 independent experiments have been performed using the local variant of the constriction factor PSO version, with neighborhood size equal to 3. This version was selected due to its fast convergence rates and efficiency. Swarm size has been set equal to 20 for all experiments, since it proved sufficient to detect global minimizers of the objective function effectively and efficiently. Moreover, further experiments with larger swarms and different PSO versions did not result in significantly different convergence rates, in terms of the required number of function evaluations. The constriction factor as well as the cognitive and the social parameters have been set to their optimal values, $\chi = 0.729$, $c_1 = c_2 = 2.05$ [6, 46]. The accuracy for the determination of the global minimizer of the objective function, has been equal to 10^{-8} .

Regarding the weights, the constraints defined by Eqs. (6)–(13), which are derived by the fuzzy regions proposed by the experts, have been initially used. However, no

	W_{12}	W_{13}	W_{15}	W_{21}	W_{31}	W_{41}	W_{52}	W_{54}
Mean	-0.4027	-0.2016	0.8991	0.3999	0.5000	-0.8000	0.9659	0.1043
Median	-0.4329	-0.2000	0.9050	0.4000	0.5000	-0.8000	0.9837	0.1000
St.Dev.	0.0487	0.0056	0.0909	0.0011	0.0003	0.0002	0.0420	0.0090
Min	-0.4500	-0.2291	0.7156	0.3889	0.4971	-0.8014	0.8685	0.1000
Max	-0.3500	-0.2000	1.0000	0.4000	0.5000	-0.8000	1.0000	0.1363

Table 1: Statistical analysis of the results for the first scenario.

solution was detected, indicating that the suggested ranges for the weights, as well as the initial weight matrix, W^{initial} , provided by the experts are not proper and do not lead the FCM to the desired steady state. The best weight matrix detected in these regions, in terms of its objective function value (i.e. the matrix that corresponds to the smallest objective function value) is:

$$W = \begin{pmatrix} 0 & -0.35 & -0.20 & 0 & 0.40 \\ 0.40 & 0 & 0 & 0 & 0 \\ 0.50 & 0 & 0 & 0 & 0 \\ -0.80 & 0 & 0 & 0 & 0 \\ 0 & 0.75 & 0 & 0.20 & 0 \end{pmatrix},$$

which led the FCM to the steady state:

$$\begin{aligned} C_1 &= 0.6723, & C_2 &= 0.7417, & C_3 &= 0.6188, \\ C_4 &= 0.6997, & C_5 &= 0.7311, \end{aligned}$$

that clearly violates the constraints for both C_1 and C_5 , defined in Eqs. (14) and (15).

Since the consideration of all eight constraints on the weights prohibits the detection of a suboptimal matrix, some of the constraints were omitted. Specifically, the constraints for the three weights W_{15} , W_{52} , and W_{54} , for which the experts' suggestions regarding their values varied widely, were omitted, one by one at the beginning, and subsequently in pairs. The corresponding weights were allowed to assume values in the range $[-1, 0]$ or $[0, 1]$, in order to avoid physically meaningless weight matrices. Despite this, no solutions were detected in these cases. However, suboptimal matrices were detected after omitting all three constraints. The statistics of the weights' values for this case are reported in Table 1 and depicted in the boxplot of Fig. 5. As shown in the figure, the weights W_{21} , W_{31} , and W_{41} , converged to almost the same value in each experiment; a value which is close to the bounds defined by the experts. The weights W_{13} and W_{54} converged also in very small ranges, while the remaining weights assumed values in wider regions. Moreover, the three unconstrained weights W_{15} , W_{52} , and W_{54} , converged in regions significantly different than those determined by the experts. The mean number of PSO iterations required in the experiments was 40.

The ranges of the output concepts' values for the obtained suboptimal matrices are depicted in Fig. 6. The output concept C_1 converges to almost the same value for each suboptimal matrix, while C_5 takes a wide range of values, always within the desired bounds. Regarding the remaining concepts, C_3 and C_4 , they converge to almost the same values, while the values of C_2 vary slightly. The obtained

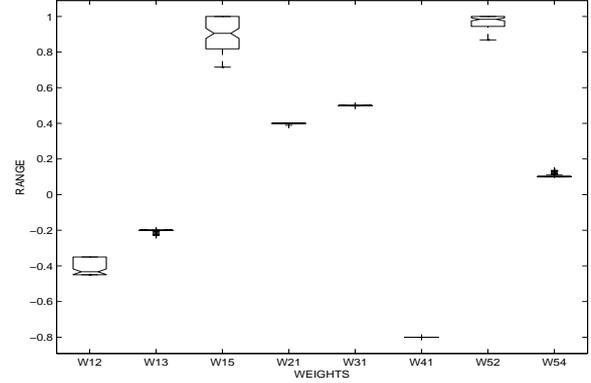


Figure 5: Boxplot of the obtained results for the weights for the first scenario.

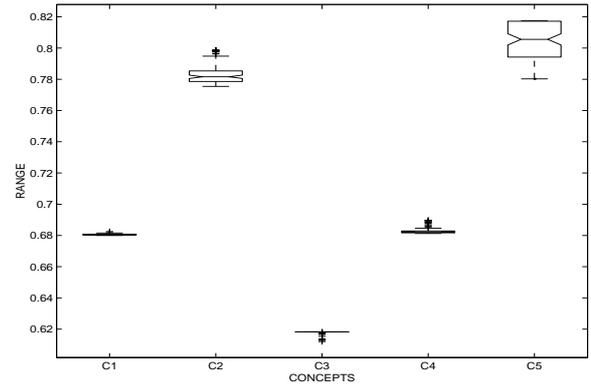


Figure 6: Boxplot of the obtained results for the concepts for the first scenario.

values for these three concepts are physically meaningful and appropriate for the operation of the system.

One of the obtained suboptimal matrices is the following:

$$W = \begin{pmatrix} 0 & -0.45 & -0.20 & 0 & 0.84 \\ 0.40 & 0 & 0 & 0 & 0 \\ 0.50 & 0 & 0 & 0 & 0 \\ -0.80 & 0 & 0 & 0 & 0 \\ 0 & 0.99 & 0 & 0.10 & 0 \end{pmatrix},$$

which leads the FCM to the desired steady state:

$$\begin{aligned} C_1 &= 0.6805, & C_2 &= 0.7798, & C_3 &= 0.6176, \\ C_4 &= 0.6816, & C_5 &= 0.7967. \end{aligned}$$

It is clear from the obtained results that there is a significant divergence of some weights from the initial weight

values suggested by the experts. The weights W_{21} , W_{31} , and W_{41} , take almost identical values in every experiment, near the initial bounds suggested by the experts. Finally, the weight W_{12} deviates slightly from its initial region. Thus, the proposed learning algorithm is able to provide proper weight matrices for the design of the FCM, efficiently and effectively, alleviating deficiencies caused by deviation in the experts' suggestions. Exploiting a priori information, such as constraints posed by the experts on weights, enhances its performance. Moreover, a primitive statistical study of the obtained results provides an intuition on the operation and the dynamics of the modelled system.

7 Conclusions

Fuzzy Cognitive Maps (FCMs) are widely used to successfully model and analyze complex systems. The need to improve the functional representation of FCMs has been outlined. A new learning algorithm for determining suboptimal weight matrices for Fuzzy Cognitive Maps with fixed structures, in order to reach a desired steady state, is introduced. The proposed approach is based on the minimization of a properly defined objective function through the Particle Swarm Optimization algorithm. The new learning approach for the determination of the FCM's weight matrix is formulated and explained.

An industrial process control problem is used for the illustration of the proposed learning algorithm. The results appear to be very promising, verifying the effectiveness of the learning procedure. The physical meaning of the obtained results is retained. The proposed approach also provides a robust solution in the event of divergent opinions of the experts concerning the system.

Future work will consider further investigation of the proposed approach considering different scenarios of the employed industrial problem as well as applications on systems of higher complexity.

Acknowledgment

We thank the anonymous reviewers for their helpful remarks and comments.

Bibliography

- [1] M.A. Abido. Optimal design of power system stabilizers using particle swarm optimization. *IEEE Trans. Energy Conversion*, 17(3):406–413, 2002.
- [2] D.K. Agrafiotis and W. Cedeno. Feature selection for structure–activity correlation using binary particle swarms. *Journal of Medicinal Chemistry*, 45(5):1098–1107, 2002.
- [3] J. Aguilar. Adaptive random fuzzy cognitive maps. In F.J. Garijo, J.C. Riquelme Santos, and M. Toro, editors, *IBERAMIA*, volume 2527 of *Lecture Notes in Computer Science*, pages 402–410. Springer, 2002.
- [4] R. Axelrod. *Structure of Decision: the Cognitive Maps of Political Elites*. Princeton University Press, Princeton, NJ, 1976.
- [5] E. Bonabeau, M. Dorigo, and G. Théraulaz. *From Natural to Artificial Swarm Intelligence*. Oxford University Press, New York, 1999.
- [6] M. Clerc and J. Kennedy. The particle swarm–explosion, stability, and convergence in a multidimensional complex space. *IEEE Trans. Evol. Comput.*, 6(1):58–73, 2001.
- [7] A.R. Cockshott and B.E. Hartman. Improving the fermentation medium for *Echinocandin B* production part ii: Particle swarm optimization. *Process Biochemistry*, 36:661–669, 2001.
- [8] E. Cox. *The Fuzzy Systems Handbook*. Academic Press, Cambridge, MA, 1999.
- [9] P. Craiger and M.D. Coover. Modelling dynamic social and psychological processes with fuzzy cognitive maps. In *Proc. 3rd IEEE Conf. Fuzzy Systems*, 1994.
- [10] J.A. Dickerson and B. Kosko. Adaptive cognitive maps in virtual worlds. In *Annual Meeting World Congress Neural Networks*, 1994.
- [11] R.C. Eberhart and Y. Shi. Comparison between genetic algorithms and particle swarm optimization. In V.W. Porto, N. Saravanan, D. Waagen, and A.E. Eiben, editors, *Evolutionary Programming*, volume VII, pages 611–616. Springer, 1998.
- [12] R.C. Eberhart, P. Simpson, and R. Dobbins. *Computational Intelligence PC Tools*. Academic Press, 1996.
- [13] P.C. Fourie and A.A. Groenwold. The particle swarm optimization algorithm in size and shape optimization. *Struct. Multidisc. Optim.*, 23:259–267, 2002.
- [14] P.P. Groumpos and C.D. Stylios. Modelling supervisory control systems using fuzzy cognitive maps. *Chaos, Solutions and Fractals*, 11:329–336, 2000.
- [15] M. Hagiwara. Extended fuzzy cognitive maps. In *Proc. IEEE Int. Conf. Fuzzy Systems*, pages 795–801, 1992.
- [16] J.S. Jang, C.T. Sun, and E. Mizutani. *Neuro–Fuzzy and Soft Computing*. Prentice Hall, Upper Saddle River, NJ, 1997.
- [17] J. Kennedy and R.C. Eberhart. Particle swarm optimization. In *Proceedings IEEE International Conference on Neural Networks*, volume IV, pages 1942–1948, Piscataway, NJ, 1995. IEEE Service Center.
- [18] J. Kennedy and R.C. Eberhart. *Swarm Intelligence*. Morgan Kaufmann Publishers, 2001.
- [19] S. Khan, A. Chong, and T.A. Gedeon. Methodology for developing adaptive fuzzy cognitive maps for decision support. In *Proc. 3rd Australia–Japan Evolutionary Systems*, pages 93–100, Canberra, 1999.

- [20] B. Kosko. Fuzzy cognitive maps. *Int. J. Man–Machine Studies*, 24:65–75, 1986.
- [21] B. Kosko. *Neural Networks and Fuzzy Systems*. Prentice Hall, New Jersey, 1992.
- [22] B. Kosko. *Fuzzy Engineering*. Prentice Hall, New York, 1997.
- [23] D.E. Koulouriotis, I.E. Diakoulakis, and D.M. Emiris. Learning fuzzy cognitive maps using evolution strategies: A novel schema for modeling and simulating high-level behavior. In *Proc. 2001 IEEE Cong. Evol. Comp.*, Seoul, Korea, 2001.
- [24] E.C. Laskari, K.E. Parsopoulos, and M.N. Vrahatis. Particle swarm optimization for integer programming. In *Proceedings of the IEEE 2002 Congress on Evolutionary Computation*, pages 1576–1581, Hawaii (HI), USA, 2002. IEEE Press.
- [25] E.C. Laskari, K.E. Parsopoulos, and M.N. Vrahatis. Particle swarm optimization for minimax problems. In *Proceedings of the IEEE 2002 Congress on Evolutionary Computation*, pages 1582–1587, Hawaii (HI), USA, 2002. IEEE Press.
- [26] C.T. Lin and C.S. Lee. *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*. Prentice Hall, Upper Saddle River, N.J., 1996.
- [27] W.Z. Lu, H.Y. Fan, A.Y.T. Leung, and J.C.K. Wong. Analysis of pollutant levels in central hong kong applying neural network method with particle swarm optimization. *Environmental Monitoring and Assessment*, 79:217–230, 2002.
- [28] C.O. Ourique, E.C. Biscaia, and J. Carlos Pinto. The use of particle swarm optimization for dynamical analysis in chemical processes. *Computers and Chemical Engineering*, 26:1783–1793, 2002.
- [29] E.I. Papageorgiou, C.D. Stylios, and P.P. Groumpos. Activation hebbian learning rule for fuzzy cognitive maps. In *Proc. 15th IFAC World Congress*, Barcelona, Spain, 2002.
- [30] E.I. Papageorgiou, C.D. Stylios, and P.P. Groumpos. Decision making in external beam radiation therapy based on fuzzy cognitive maps. In *Proc. 1st Int. IEEE Symp. Intelligent Systems*, Varna, Bulgaria, 2002.
- [31] E.I. Papageorgiou, C.D. Stylios, and P.P. Groumpos. An integrating two-level hierarchical system for decision making in radiation therapy using fuzzy cognitive maps. *IEEE Transactions on Biomedical Engineering*, 2003. accepted for publication.
- [32] K.E. Parsopoulos, E.C. Laskari, and M.N. Vrahatis. Particle identification by light scattering through evolutionary algorithms. In *Proc. 1st International Conference for Mathematics and Informatics for Industry*, pages 97–108, Thessaloniki, Greece, 2003.
- [33] K.E. Parsopoulos and M.N. Vrahatis. Initializing the particle swarm optimizer using the nonlinear simplex method. In A. Grmela and N.E. Mastorakis, editors, *Advances in Intelligent Systems, Fuzzy Systems, Evolutionary Computation*, pages 216–221. WSEAS Press, 2002.
- [34] K.E. Parsopoulos and M.N. Vrahatis. Particle swarm optimization method for constrained optimization problems. In P. Sincak, J. Vascak, V. Kvasnicka, and J. Pospichal, editors, *Intelligent Technologies—Theory and Application: New Trends in Intelligent Technologies*, volume 76 of *Frontiers in Artificial Intelligence and Applications*, pages 214–220. IOS Press, 2002.
- [35] K.E. Parsopoulos and M.N. Vrahatis. Recent approaches to global optimization problems through particle swarm optimization. *Natural Computing*, 1(2–3):235–306, 2002.
- [36] K.E. Parsopoulos and M.N. Vrahatis. Computing periodic orbits of nondifferentiable/discontinuous mappings through particle swarm optimization. In *Proceedings of the IEEE Swarm Intelligence Symposium*, pages 34–41, Indianapolis (IN), USA, 2003.
- [37] T. Ray and K.M. Liew. A swarm metaphor for multi-objective design optimization. *Engineering Optimization*, 34(2):141–153, 2002.
- [38] A. Saldam, I. Ahmad, and S. Al-Madani. Particle swarm optimization for task assignment problem. *Microprocessors and Microsystems*, 26:363–371, 2002.
- [39] Y. Shi and R.C. Eberhart. A modified particle swarm optimizer. In *Proceedings IEEE Conference on Evolutionary Computation*, Anchorage, AK, 1998. IEEE Service Center.
- [40] Y. Shi and R.C. Eberhart. Parameter selection in particle swarm optimization. In V.W. Porto, N. Saravanan, D. Waagen, and A.E. Eiben, editors, *Evolutionary Programming*, volume VII, pages 591–600. Springer, 1998.
- [41] C.D. Stylios and P.P. Groumpos. The challenge of modelling supervisory systems using fuzzy cognitive maps. *J. Intelligent Manufacturing*, 9:339–345, 1998.
- [42] C.D. Stylios and P.P. Groumpos. Fuzzy cognitive maps in modeling supervisory control systems. *J. Intelligent & Fuzzy Systems*, 8(2):83–98, 2000.
- [43] R. Taber. Knowledge processing with fuzzy cognitive maps. *Expert Systems With Applications* 2, pages 83–87, 1991.
- [44] R. Taber. Fuzzy cognitive maps model social systems. *AI Experts*, 9:8–23, 1994.
- [45] V. Tandon, H. El-Mounayri, and H. Kishawy. Nc end milling optimization using evolutionary computation. *Int. J. Machine Tools & Manufacture*, 42:595–605, 2002.
- [46] I.C. Trelea. The particle swarm optimization algorithm: Convergence analysis and parameter selection. *Information Processing Letters*, 85:317–325, 2003.