1st International Conference “From Scientific Computing to Computational Engineering”
1st IC-SCCE
Athens, 8-10 September, 2004

RECENT ADVANCES IN FUZZY COGNITIVE MAPS LEARNING USING EVOLUTIONARY COMPUTATION TECHNIQUES

K.E. Parsopoulos[1,3], E.I. Papageorgiou[2,3], P.P. Groumpos[2,3] and M.N. Vrahatis[1,3]

[1] Department of Mathematics, University of Patras, GR-26110 Patras, Greece
e-mail: {kostasp,vrahatis}@math.upatras.gr
[2] Department of Electrical and Computer Engineering, University of Patras, GR-26500 Patras, Greece
e-mail: {epapageo,groumpos}@ee.upatras.gr
[3] University of Patras Artificial Intelligence Research Center (UPAIRC)

Keywords: Fuzzy Cognitive Maps, Particle Swarm Optimization, Evolutionary Computation

Abstract. *A recently proposed swarm intelligence technique for Fuzzy Cognitive Map learning is described. The technique employs the Particle Swarm Optimization algorithm to minimize a proper objective function, whose global minimizers correspond to suboptimal weight matrices of the Fuzzy Cognitive Map. New instances of an industrial test problem are studied, justifying the usefulness of the technique as an efficient and effective algorithm for Fuzzy Cognitive Map learning.*

1 INTRODUCTION

Fuzzy Cognitive Maps (FCMs) are used for modeling, simulating and controlling systems in engineering applications. Kosko introduced FCMs for the representation of causal relationships among concepts, as well as for the analysis of inference patterns[1,2]. The flexibility of FCMs renders them a promising methodology with applications in many scientific fields, such as bioinformatics, manufacturing, organization behavior, political science, and decision-making. However, there are also several deficiencies of FCMs, which are mainly related to their heavy dependence on human knowledge, as well as to the lack of enhanced and robust learning algorithms. Therefore, the development of new learning techniques is a field of increasing interest.

A recently proposed approach for FCMs learning, which is based on a swarm intelligence algorithm, is presented. The technique is based on the minimization of a properly defined objective function through the Particle Swarm Optimization (PSO) algorithm[3,4], to determine suboptimal weight matrices of the FCM. Preliminary results of this technique on different problems from industry and bioinformatics are very promising[3,4,5]. We extend these studies, providing results on new instances of an industrial control problem, to justify the usefulness of the new learning technique in a wide range of different scenarios.

The paper is organized as follows: the PSO algorithm is briefly presented in Section 2, while Section 3 is devoted to the description of both the basic principles of FCMs and the learning procedure. In Section 4, the process control problem is described, and experimental results are reported and discussed in Section 5. The paper closes with a synopsis in Section 6.

2 PARTICLE SWARM OPTIMIZATION

Particle Swarm Optimization (PSO) is a stochastic optimization algorithm that belongs to the class of *swarm intelligence* algorithms. Although it is closely related to the Evolutionary Computation field, the inspiration behind the development of PSO lies rather in social dynamics and emergent behavior that arise in organized groups of individuals[6,7,8].

PSO is a population-based algorithm, i.e., it exploits a population of search points that probe the search space concurrently. The population is called a *swarm* and the individuals (i.e., the search points) are called *particles*. Each particle moves with an adaptable velocity within the search space, retaining a memory of the best position it ever visited. Also, it exchanges information with the rest of the particles, regarding the most promising regions of the search space. There are two main variants of PSO with respect to the information that is communicated among the particles. In the *global* variant, the best position ever attained by all individuals of the swarm is communicated to all the particles, while, in the *local* variant, each particle is assigned to a neighborhood that consists 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[7,8,9].

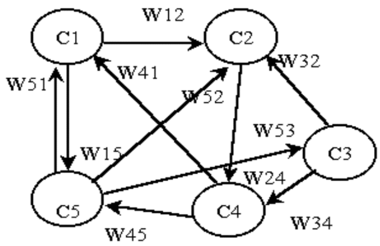

Figure 1. A Fuzzy Cognitive Map.

Assume an n-dimensional search space, $S \subset \mathbf{R}^n$, and a swarm consisting of N particles, $X_1,\ldots,X_N$. Each particle is an n-dimensional vector, $X_i = (x_{i1}, x_{i2},\ldots,x_{in})^{\mathrm{T}}$, in S, and it moves with an adaptable velocity, $V_i = (v_{i1}, v_{i2},\ldots,v_{in})^{\mathrm{T}}$. The best previous position encountered by X_i is a point in S, denoted by $P_i = (p_{i1}, p_{i2},\ldots, p_{in})^{\mathrm{T}}$. Let g_i denote the index of the particle that attained the best previous position among all the particles in the neighborhood of X_i, and t to represent the iteration counter. Then, the swarm is manipulated by the equations[7,9,10,11],

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

$$X_i(t+1) = X_i(t) + V_i(t+1), \tag{2}$$

where $i = 1,\ldots,N$; c_1 and c_2 are two parameters called *cognitive* and *social* parameters, respectively; and r_1, r_2, are random numbers uniformly distributed within [0,1]. The parameter χ is called the *constriction factor* and it is used as a mechanism for controlling the magnitude of the velocity.

The value of the constriction factor depends on the values c_1 and c_2, and it is derived analytically by the formula[10],

$$\chi = \frac{2\kappa}{\left|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}\right|}, \tag{3}$$

with $\kappa = 1$, $\varphi = c_1 + c_2$, following the stability analysis of PSO[10,11]. Usually, the default values $c_1 = c_2 = 2.05$ are used, resulting in $\chi = 0.729$. The initialization of the swarm and velocities is usually performed randomly and uniformly in the search space, although more sophisticated initialization techniques can enhance the overall performance of the algorithm[12].

3 BASIC CONCEPTS OF FUZZY COGNITIVE MAPS

Kosko has introduced FCMs in 1986 as signed directed graphs for representing causal reasoning and computational inference processing[1]. FCMs comprise a soft computing modeling methodology that follows a method similar to human reasoning. They exploit a symbolic representation for the description and modeling of the system, and *concepts* are utilized to represent different aspects of the system as well as their behavior. The dynamic of the system is simulated by the interaction of concepts, which is expressed using weighted arcs that interconnect them.

Let M be the total number of concepts of an FCM. The concepts are denoted by C_i, $i = 1,\ldots,M$, and they represent the key-factors of the system. Each concept is characterized by a value A_i in [0,1], $i = 1,\ldots,M$. 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 arc between two concepts, C_i and C_j, has a weight W_{ij}, which is analogous 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, a positive weight expresses positive causality, a negative weight expresses negative causality, and a zero weight expresses the lack of any relation between the interconnected concepts.

FCMs are designed and configured by a group of human experts, thus, integrating their experience and knowledge on the characteristics and components of the simulated system. The experts determine the type and the number of concepts, as well as the initial weights of the FCM. The values determined by the experts are fuzzy, in order to avoid deficiencies caused by the direct assignment of numerical values. At the beginning, the experts determine the relevant factors that will be represented as concepts. Then, they individually describe the causal relationships among the concepts, using a linguistic notion. The influence of a concept on another is characterized as "negative", "positive" or "no influence", and the linguistic weights on each arc are characterized as "strong", "weak", etc[13]. The linguistic variables are aggregated in a single linguistic variable, which is transformed to a single linguistic weight, through the SUM technique[14]. Finally, the Center of Area (CoA) defuzzification method[2,14] is used for the transformation of the linguistic to a numerical weight within [-1,1]. Thus, an initial weight matrix,

$$W^{initial} = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1M} \\ W_{21} & W_{22} & \cdots & W_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MM} \end{pmatrix}, \tag{4}$$

with $W_{ii} = 0$, $i = 1,\ldots,M$, is obtained.

The value, A_i, of a concept, C_i, expresses the quantity of its corresponding physical value and it is determined by transforming the fuzzy values to numerical values. Having assigned values to concepts and weights, the FCM is let to converge to a steady state, through the iterative scheme[2],

$$A_i(k+1) = f\left(A_i(k) + \sum_{\substack{j=1 \\ j \neq i}}^{M} W_{ji} A_j(k) \right), \tag{5}$$

where k stands for the iteration counter; and W_{ji} is the weight of the arc connecting concepts C_j and C_i. The function f is the well-known sigmoid function,

$$f(x) = \frac{1}{1+\exp(-\lambda x)}, \tag{6}$$

with $\lambda > 0$. In this work, we used $\lambda = 1$, since the values A_i must lie, by definition, within [0, 1]. The interaction of the FCM results after a few iterations in a steady state, where the values of the concepts are not modified further by the application of Eq. (5). If the steady state values of the output concepts of the FCM lie within desired ranges, then the proper operation of the simulated system is guaranteed.

The critical dependence on the opinions of the experts, as well as 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 such that convergence to undesired steady states is alleviated. There are just a few established FCM learning algorithms and they are mostly based on ideas coming from the field of artificial neural networks training[15,16,17]. These algorithms start from an initial state and an initial weight matrix 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 that are accepted by the experts. The main drawback of this approach is the heavy dependence of the final weights' configuration on the initial weight matrix. Wrong initial estimation of the weights or large deviation among the suggestions of the experts may lead in reduced efficiency of the algorithms, as well as undesired steady states of the system.

A different learning procedure that alleviates the problem of the potential convergence to an undesired steady state was proposed in[3,4,5]. This approach is based on PSO. The main goal 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 desired behavior is characterized by output concept values that lie within desired, problem dependent bounds that are prespecified by the experts. The learning procedure is, to some extent, similar to that of neural networks training. More specifically, let $C_1,\ldots,C_M$, be the concepts of an FCM, and let $C_{out1},\ldots,C_{outm}$, be the output concepts. The remaining concepts are considered input or interior concepts. The user is interested in restricting the values of these output concepts in strict bounds,

$$A_{outi}^{\min} \leq A_{outi} \leq A_{outi}^{\max}, \tag{7}$$

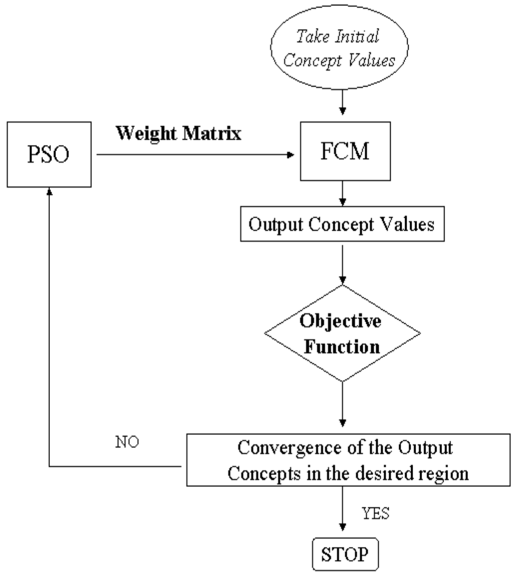

Figure 2. Flowchart of the proposed learning methodology.

which are predetermined by the experts, and they are crucial for the proper operation of the modeled system. Thus, the main goal is to detect a weight matrix, $W = [W_{ij}]$, $i,j=1,\ldots,M$, 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. Imposing constraints on the potential values assumed by weights attains the latter. To do this, we consider the objective function[3,4,5],

$$F(W)=\sum_{i=1}^{M} H\left(A_{outi}^{\min}-A_{outi}\right)\left|A_{outi}^{\min}-A_{outi}\right|+\sum_{i=1}^{M} H\left(A_{outi}-A_{outi}^{\max}\right)\left|A_{outi}^{\max}-A_{outi}\right|, \tag{8}$$

where H is the well-known Heaviside function,

$$H(x)=\begin{cases}0, & x<0,\\ 1, & x\geq 0,\end{cases} \tag{9}$$

and A_{outi}, $i = 1,\ldots,$m, are the steady state values of the output concepts that are obtained through the application of the iterative scheme of Eq. (5), using the weight matrix W. Obviously, the global minimizers of the objective function F are weight matrices that lead the FCM to a desired steady state, i.e., all output concepts lie within the desired regions. The objective function F is non-differentiable and, thus, gradient-based methods are not applicable for its minimization. For this purpose, PSO is used for the minimization of F. The non-differentiability of F poses no problems in our approach since PSO, like all evolutionary algorithms, requires function values solely, and it can be applied even on discontinuous functions. An FCM with M, fully interconnected concepts corresponds to an $M(M\text{-}1)$-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. A flowchart of the learning procedure is depicted in Fig.2.

The physical meaning of each interconnection is retained through constraints posed by the experts on the FCM's weights. Constraints are provided in the form of negative or positive relations between two concepts. For example, if two concepts, C_i and C_j, are positively related, then the corresponding weight, W_{ij}, takes values in [0,1], while, if they are negatively related, it takes values within [-1,0]. More strict constraints may be additionally posed on some weights, either by the experts or by taking into consideration the convergence regions obtained through the application of the learning algorithm, enhancing the overall performance of the algorithm.

There is, in general, a plethora of weight matrices that lead the FCM to convergence to desired regions of the output concepts. PSO is a stochastic algorithm, and, therefore, different suboptimal matrices can be obtained 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

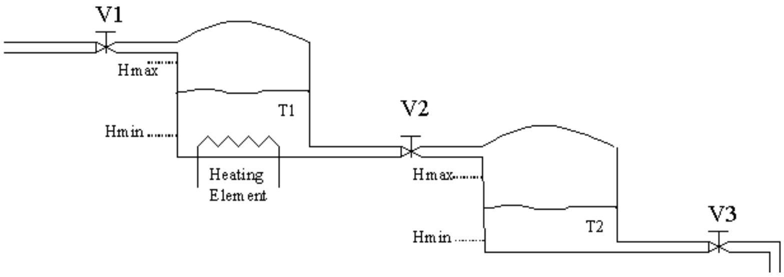

Figure 3. The industrial process control problem.

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. Further information on the problem at hand may be incorporated to enhance the procedure, either by modifying the objective function or by imposing further constraints on the weights. The proposed approach has proved to be very efficient in practice[3,4,5]. In the following section, its operation on an industrial process control problem is illustrated.

4 AN INDUSTRIAL PROCESS CONTROL PROBLEM

A process control problem from the field of chemical industry has been used to illustrate the workings of the proposed learning algorithm[3,4]. The system consists of two tanks, three valves, one heating element and two thermometers, one in each tank, as depicted in Fig. 3. Each tank has an inlet valve and an outlet valve. The outlet valve of the first tank is the inlet valve of the second tank. The main goal is, on the one hand, to keep the height of the liquid in both tanks within some limits, $H_{\min}$ and $H_{\max}$, and, on the other hand, to keep the temperature of the liquid in both tanks within limits, $T_{\min}$ and $T_{\max}$. The temperature of the liquid in Tank 1 is regulated through the heating element. The temperature of the liquid in Tank 2 is measured through a sensor thermometer. When the temperature of the liquid in Tank 2 decreases, valve V2 must open, in order to have hot liquid poured from Tank 1 into Tank 2. Thus, the control objective is to ensure that the following relations hold:

$$\begin{aligned} &H^1_{\min} \le H^1 \le H^1_{\max}, \quad T^1_{\min} \le T^1 \le T^1_{\max}, \\ &H^2_{\min} \le H^2 \le H^2_{\max}, \quad T^2_{\min} \le T^2 \le T^2_{\max}. \end{aligned} \tag{10}$$

An FCM model has been constructed for this process control problem. The height and temperature of the liquid in each tank as well as the state of the valves have been represented as concepts of the FCM. More specifically, the following eight concepts have been used:

Concept 1: Height of liquid in Tank 1. It is dependent on the state of valves V1 and V2.
Concept 2: Height of liquid in Tank 2. It is dependent on the state of valves V2 and V3.
Concept 3: State of the valve V1 (open, closed or partially open).
Concept 4: State of the valve V2 (open, closed or partially open).
Concept 5: State of the valve V3 (open, closed or partially open).
Concept 6: Temperature of the liquid in Tank 1.
Concept 7: Temperature of the liquid in Tank 2.
Concept 8: Heating element, which is used to increase the temperature of the liquid in Tank 1.

These concepts are interconnected with each other. The signs and the weight of each connection have been determined in[18,19]. Three experts were pooled to construct the FCM, following the methodology described in[19]. The resulted FCM is depicted in Fig. 4. The values of the concepts correspond to real measurements of the actual physical process. The output concepts of the problem, as proposed by the experts, are C_1, C_2, C_6 and C_7. All experts agreed regarding the direction of the interconnections between the concepts, and they determined the overall linguistic variable as well as the corresponding fuzzy set for each weight. The final ranges for the weights, as implied by the fuzzy regions, are

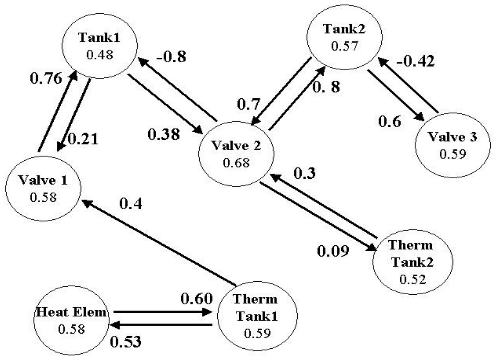

Figure 4: The FCM model for the process control problem.

$$\begin{array}{lllll} 0.0 \le w_{13} \le 0.5, & 0.0 \le w_{14} \le 0.5, & 0.5 \le w_{24} \le 0.75, & 0.25 \le w_{25} \le 0.75, & 0.5 \le w_{31} \le 1.0, \\ -1.0 \le w_{41} \le -0.5, & 0.5 \le w_{42} \le 0.75, & 0.0 \le w_{47} \le 0.25, & -0.75 \le w_{52} \le 0.0, & 0.0 \le w_{63} \le 0.5, \\ 0.25 \le w_{68} \le 0.7, & 0.0 \le w_{74} \le 0.5, & 0.25 \le w_{86} \le 0.7. & & \end{array} \tag{11}$$

The initial weights of the FCM (in the order of appearance in Eq. (11)) are

$$W^{initial} = [0.21, 0.38, 0.70, 0.60, 0.76, -0.8, 0.8, 0.09, -0.42, 0.4, 0.53, 0.3, 0.6]. \tag{12}$$

The experts determined that the values of the concepts must change asynchronously. They also specified the desired regions for the output concepts, which reflect the proper operation of the modeled system.

Preliminary results on the process control problem described above are reported in[4]. The learning procedure described in Section 3 succeeded in determining proper weight matrices of the FCM that result in desired output concept values. In the next section, two different cases for the output concept values are investigated. The most important component in developing the FCM is the determination of the concepts that best describe the system and the direction and grade of causality between concepts.

5 EXPERIMENTAL RESULTS

Two scenarios for the weight ranges and two different sets of constraints for the output concepts were considered. For each scenario, 100 independent experiments have been performed using the global variant of a constriction factor PSO. The swarm size was set to 5. The constriction factor, as well as the cognitive and the social parameter have been set to their default values, $\chi = 0.729$, $c_1 = c_2 = 2.05$ [10]. The first set of output concepts values was

$$0.5 \le C_1 \le 0.7, \quad 0.75 \le C_2 \le 0.8, \quad 0.6 \le C_6 \le 0.7, \quad 0.6 \le C_7 \le 0.8, \tag{13}$$

while the second set was

$$0.5 \le C_1 \le 0.7, \quad 0.75 \le C_2 \le 0.8, \quad 0.73 \le C_6 \le 0.81, \quad 0.65 \le C_7 \le 0.75. \tag{14}$$

In previous work[4], it was derived that the weights w_{47}, w_{52} and w_{86} take values in smaller regions than the regions proposed by the experts. More specifically they lied within the bounds[4]

$$-0.25 \le w_{47} \le 0.25, \quad -0.5 \le w_{52} \le 0.2, \quad 0 \le w_{86} \le 0.4. \tag{15}$$

The first scenario considered the constraints of Eq. (15), while the rest weights take values in the ranges given in Eq. (11). Experiments were performed for the two different cases of the output concept values, (Eqs. (13) and (14)), using the proposed learning technique. The obtained convergence regions for the weights are depicted in boxplots in Fig. 5.

A suboptimal weight vector for the first set of output concept values was

$$W = [0.406, 0.29, 0.717, 0.965, 0.841, -0.765, 0.681, 0.151, -0.453, 0.227, 0.356, 0.126, 0.18],$$

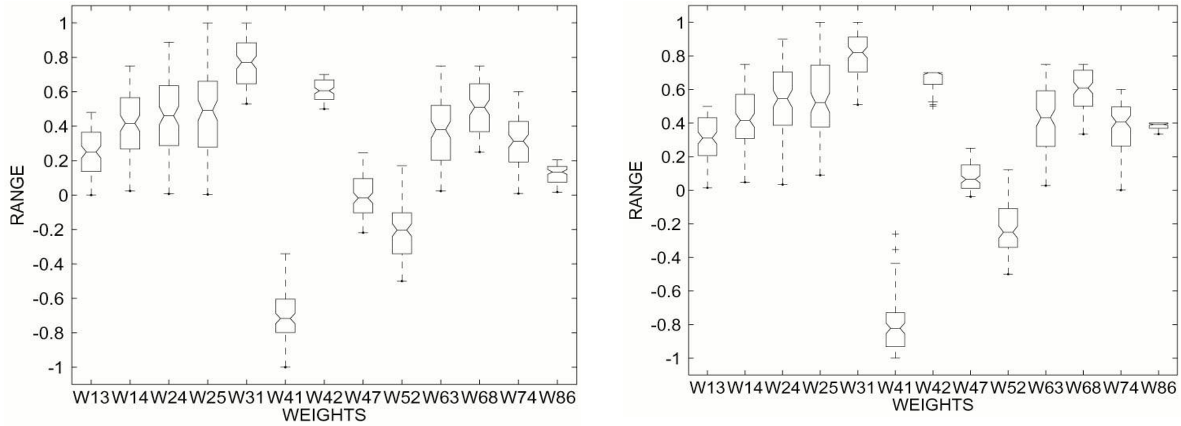

Figure 5. Weight ranges for the first scenario, and for the two sets of output concept values.

and the corresponding values of the output concepts were $C_1 = 0.66$, $C_2 = 0.713$, $C_6 = 0.696$, $C_7 = 0.694$. A suboptimal weight vector for the second set of output concept values was

$$W = [0.20, 0.612, 0.162, 0.316, 0.686, -0.807, 0.7, 0.215, -0.22, 0.657, 0.75, 0.48, 0.4],$$

and the corresponding values for the output concepts were $C_1 = 0.62$, $C_2 = 0.768$, $C_6 = 0.742$, $C_7 = 0.709$. Comparing the derived convergence regions of the weights with the corresponding bounds provided by the experts, it is observed that weights w_{42} and w_{86} take values within even smaller regions. Also, the weight w_{86} has almost a constant value, approximately equal to 0.4, in the second set of the output concept values.

In the second scenario, we took into consideration the results of the first scenario and considered the weights w_{14}, w_{24} and w_{25}, to lie in the ranges

$$0.25 \le w_{14} \le 0.5, \quad 0.25 \le w_{24} \le 0.6, \quad 0.25 \le w_{25} \le 0.7. \tag{16}$$

The results obtained by keeping the rest weights within the ranges provided in Eqs. (11) and (15), and using the proposed learning technique, are depicted in Fig. 6. A suboptimal weight vector for the first set of output concept values was

$$W = [0.275, 0.5, 0.6, 0.57, 0.95, -0.78, 0.596, 0.105, -0.134, 0.442, 0.685, 0.104, 0.171],$$

and the corresponding values of the output concepts were $C_1 = 0.682$, $C_2 = 0.762$, $C_6 = 0.696$, $C_7 = 0.684$. A suboptimal weight vector for the second set of output concept values was

$$W = [0.29, 0.389, 0.441, 0.616, 0.855, -0.64, 0.522, -0.022, 0.17, 0.502, 0.601, 0.285, 0.352],$$

and the corresponding values for the output concepts are $C_1 = 0.699$, $C_2 = 0.797$, $C_6 = 0.731$, $C_7 = 0.653$.

It is clear that the learning algorithm provides proper weight matrices and weight ranges for the FCM model, efficiently. The boxplots in Figs. 5 and 6 provide indications regarding the quality of the weights' bounds determined by the experts, and it can be used in the future as a mechanism of the evaluation of the experts' suggestions. In all cases, the mean number of PSO iterations was 5.

6 SYNOPSIS

An FCM learning technique using the PSO algorithm was described. This learning technique determines the appropriate cause-effect relationships (weights) among the concepts of FCMs, based on the minimization of a properly defined objective function through PSO. New instances of an industrial process control problem were studied, and the effectiveness of the methodology was investigated using two different scenarios for two different sets of desired output concepts values. The proposed learning procedure proved to be a promising approach, improving the effectiveness of the FCM operation mode.

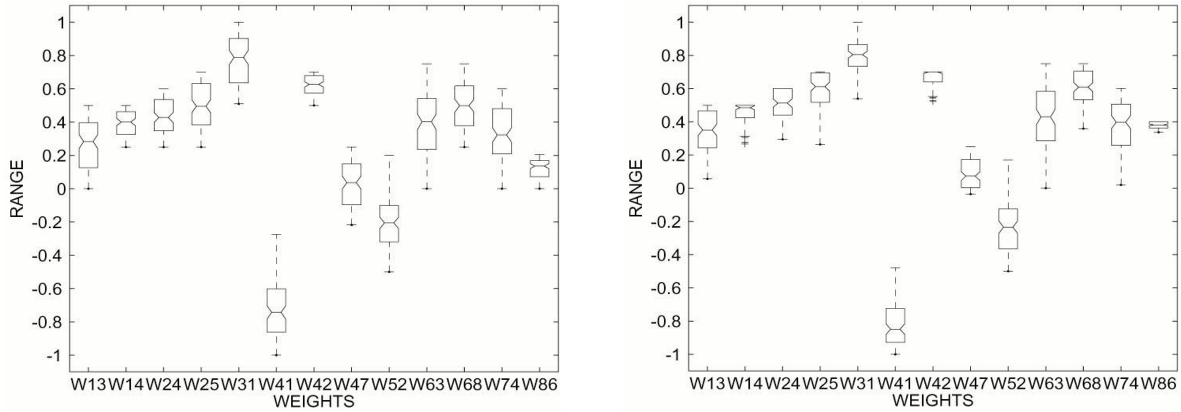

Figure 6. Weight ranges for the second scenario, and for the two sets of output concept values.

ACKNOWLEDGEMENT

This work was partially supported by the “Pythagoras” research grant awarded by the Greek Ministry of Education and Religious Affairs and the European Union.

REFERENCES

[1] Kosko, B., (1986), “Fuzzy Cognitive Maps”, Int. J. Man-Machine Studies, Vol. 24, pp. 65-75.

[2] Kosko, B., (1997), *Fuzzy Engineering*, Prentice-Hall, New Jersey,.

[3] Parsopoulos, K.E., Papageorgiou, E.I., Groumpos, P.P. and Vrahatis, M.N. (2003), “A First Study of Fuzzy Cognitive Map Learning Using Particle Swarm Optimization”, *Proceedings of IEEE 2003 Congress on Evolutionary Computation*, Canberra, Australia, IEEE Press, pp. 1440-1447.

[4] Papageorgiou, E.I., Parsopoulos, K.E., Groumpos, P.P. and Vrahatis, M.N. (2004), “Fuzzy Cognitive Maps Learning Through Swarm Intelligence”, *Lecture Notes in Computer Science (LNAI)* 3070, pp. 344-349, Springer Verlag.

[5] Parsopoulos, K.E., Papageorgiou, E.I., Groumpos, P.P. and Vrahatis, M.N. (2004), “Evolutionary Computation Techniques for Optimizing Fuzzy Cognitive Maps in Radiation Therapy Systems”, *Lecture Notes in Computer Science* 3102, pp. 402-413, Springer Verlag.

[6] Kennedy, J. and Eberhart, R.C. (1995), “Particle Swarm Optimization”, *Proceedings IEEE International Conference on Neural Networks*, Piscataway, NJ, IEEE Service Center, Volume IV, pp. 1942-1948.

[7] Kennedy, J. and Eberhart, R.C. (2001), *Swarm Intelligence*, Morgan Kaufmann Publishers.

[8] Parsopoulos, K.E., Vrahatis, M.N. (2002), “Recent Approaches to Global Optimization Problems Through Particle Swarm Optimization”, *Natural Computing,* Vol. 1 (2-3), pp. 235-306.

[9] Parsopoulos, K.E., Vrahatis, M.N. (2002), “On the Computation of All Global Minimizers Through Particle Swarm Optimization”, *IEEE Transactions Evolutionary Computation*, in press.

[10] Clerc, M. Kennedy, J. (2002), “The Particle Swarm - Explosion, Stability and Convergence in a Multidimensional Complex Space”, *IEEE Transactions Evolutionary Computation*, Vol. 6, pp.58-73.

[11] Trelea, I.C. (2003), “The Particle Swarm Optimization Algorithm: Convergence Analysis and Parameter Selection”, *Information Processing Letters* Vol. 85, pp. 317-325.

[12] Parsopoulos, K.E., Vrahatis, M.N. (2002), “Initializing the particle swarm optimizer using the nonlinear simplex method”, In: Grmela, A., Mastorakis, N., eds.: *Advances in Inteligent systems, Fuzzy Systems, Evolutionary Computation*, WSEAS Press, pp. 216-221.

[13] Cox, E. (1999), *The Fuzzy Systems Handbook*, Academic Press, Cambridge, MA.

[14] Lin, C.T., and Lee, C.S. G. (1996), *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*, N.J., Prentice Hall.

[15] Dickerson, J., Kosko, B. (1994). “Fuzzy Virtual Worlds”, *AI Expert*, pp. 25-31.

[16] Papageorgiou, E.I., Stylios, C.D., Groumpos, P.P. (2004), “Active Hebbian Learning Algorithm to Train Fuzzy Cognitive Maps”, *International Journal of Approximate Reasoning*, Elsevier, in press.

[17] Hebb, D.O. (1949), *The Organization of Behaviour: A Neuropsychological Theory*, Wiley, New York.

[18] Stylios, C.D., Groumpos, P.P. (1999), “Fuzzy Cognitive Maps: a model for intelligent supervisory control systems”, *Computers in Industry,* Vol. 39, pp. 229-238.

[19] Stylios, C.D. Groumpos, P.P. (2000), “Fuzzy Cognitive Maps in Modeling Supervisory Control Systems”, *Journal of Intelligent & Fuzzy Systems,* Vol. 8, pp. 83-98.