Evolutionary Computation Techniques for Optimizing Fuzzy Cognitive Maps in Radiation Therapy Systems

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

Abstract. The optimization of a Fuzzy Cognitive Map model for the supervision and monitoring of the radiotherapy process is proposed. This is performed through the minimization of the corresponding objective function by using the Particle Swarm Optimization and the Differential Evolution algorithms. The proposed approach determines the cause–effect relationships among the concepts of the supervisor–Fuzzy Cognitive Map by computing its optimal weight matrix, through extensive experiments. Results are reported and discussed.

1 Introduction

Several pathological illness cases can be addressed by eliminating the infected cells through the application of ionizing radiation to the patient. This procedure is widely known as *radiotherapy*. In the case of cancer cells, the radiation consists mainly of photons or electrons. Healthy cells are also affected by the radiation. Clearly, the determination of the dosage distribution of radiation, as well as information regarding the affection of the tumor by irradiation and the affection of the healthy tissues, are of major importance [1].

Radiotherapists–doctors must take into consideration many different (complementary, similar or conflicting) factors that influence the selection of the radiation dose and, consequently, the final result of the therapy. All these factors are usually incorporated in an optimization process, where the main objectives are to minimize the total amount of radiation at which the patient is exposed, maximize the minimum final radiation dose received by the tumor, minimize the radiation to critical structure(s) and healthy tissues, and produce acceptable dosage distributions with the smallest computational effort [1].

Several algorithms have been proposed and used for the optimization of radiation therapy treatment plans [2,3]. Dose calculation algorithms [4,5], dose–volume feasibility search algorithms [6], and biological objective algorithms [7]

K. Deb et al. (Eds.): GECCO 2004, LNCS 3102, pp. 402–413, 2004.

have been employed for the determination of dosage distributions in treatment planning systems under multiple criteria and dose–volume constraints [3]. Different algorithms have been proposed for the optimization of beam weights and beam directions [8]. Gradient–descent methods have been used to optimize the objective functions as well as the intensity distributions [9]. Moreover, methods related to knowledge–based expert systems and neural networks, have been proposed for the optimization of treatment variables and the support of decisions during radiotherapy planning [10,11].

The kind, the nature, as well as, the number of the parameters–factors that are taken into consideration for the determination of the radiation therapy treatment, give rise to a highly complex, uncertain and fuzzy overall model. Fuzzy Cognitive Maps (FCMs) have been applied for the modeling of the decision–making process of radiation therapy, with promising results [12]. FCMs can model complex systems that involve different factors, states, variables, and events, integrating the influence of several controversial factors in a decision–making process [13]. In FCMs, the causal effects among different factors are taken into consideration in the calculation of the values of all causal concepts that determine the radiation dose, so as to keep the dose at a minimum level, while destroying the tumor and inflicting the minimum injuries to healthy tissues and organs [1].

In this paper, two different algorithms, Particle Swarm Optimization (PSO) and Differential Evolution (DE), coming from the fields of Swarm Intelligence and Evolutionary Computation, respectively, are employed for the optimization of the supervisor–FCM used in an established radiation therapy treatment planning system. Both methods have proved to be very efficient in a plethora of applications in science and engineering. Also, PSO has recently proved to be very efficient algorithm for FCMs learning in an industrial problem [14].

The rest of this article is organized as follows: the PSO and DE algorithms are briefly presented in Sections 2 and 3, respectively. A review of the basic concepts and notion of FCMs, as well as a description of the FCM model for the supervision of the radiation therapy process, are given in Section 4. The proposed approach and experimental results are analyzed in Section 5. The paper concludes in Section 6.

2 The Particle Swarm Optimization Algorithm

Particle Swarm Optimization (PSO) is a stochastic optimization algorithm. It belongs to the class of *Swarm Intelligence* algorithms, which are inspired from, and based on the social dynamics and emergent behavior in socially organized colonies [15,16,17]. PSO is a population based algorithm, i.e. it exploits a population of individuals to synchronously probe promising regions of the search space. In this context, 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, 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 at each iteration. In the *local* variant, each particle is assigned to a neighborhood consisting of a prespecified number of particles. In this case, the best position ever attained by the particles that comprise a neighborhood is communicated among them [16, 18].

Assume a D–dimensional search space, $S \subset \mathbb{R}^D$, and a swarm consisting of N particles. Let $X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})^\top \in S$, be the i–th particle and $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})^\top \in S$, be its velocity. Let also the best previous position encountered by the i–th particle in S be denoted by $P_i = (p_{i1}, p_{i2}, \ldots, p_{iD})^\top$. 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 G to be the iteration counter. Then, the swarm is manipulated by the equations [19]:

$$V_i(G+1) = \chi \Big[V_i(G) + c_1 r_1 \big(P_i(G) - X_i(G)\big) + c_2 r_2 \big(P_{g_i}(G) - X_i(G)\big)\Big], \quad (1)$$

$$X_i(G+1) = X_i(G) + V_i(G+1), \quad (2)$$

where $i = 1, \ldots, N$; χ is a parameter called *constriction factor*; 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 $[0, 1]$.

Alternatively, a different version of the velocity's update equation, which incorporates a parameter called *inertia weight*, has been proposed [20,21]:

$$V_i(G+1) = w V_i(G) + c_1 r_1 \big(P_i(G) - X_i(G)\big) + c_2 r_2 \big(P_{g_i}(G) - X_i(G)\big), \quad (3)$$

where w is the *inertia weight.*

Both the constriction factor and the inertia weight are mechanisms for controlling the magnitude of velocities. However, there are some major differences regarding the way these two are computed and applied. The constriction factor is derived analytically through the formula [19],

$$\chi = \frac{2\kappa}{|2 - \phi - \sqrt{\phi^2 - 4\phi}|}, \quad (4)$$

for $\phi > 4$, where $\phi = c_1 + c_2$, and $\kappa = 1$. Different configurations of χ, as well as a thorough theoretical analysis of the derivation of (4), can be found in [19,22]. On the other hand, experimental results suggest that it is preferable to initialize the inertia weight w to a large value, giving priority to global exploration of the search space, and gradually decrease it, so as to obtain refined solutions [20,21]. This finding is intuitively very appealing. In conclusion, an initial value of w around 1.0 and a gradual decline towards 0 is considered a proper choice for w.

Regarding the social and cognitive parameter, although the default values $c_1 = c_2 = 2$ have been proposed and usually used, experimental results indicate that alternative configurations, depending on the problem at hand, may produce superior performance [17,19,23]. 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 [24].

3 The Differential Evolution Algorithm

The Differential Evolution (DE) algorithm has been developed by Storn and Price [25]. It utilizes N, D–dimensional vectors $X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})^\top$, $i = 1, \ldots, N$, as a population for each iteration (generation), G, of the algorithm. The initial population is taken to be uniformly distributed in the search space. At each generation, the *mutation* and *crossover* (recombination) operators are applied on the individuals, and a new population arises. Then, the selection phase starts, where the two populations compete each other, and the next generation is formed [25].

According to the *mutation* operator, for each vector $X_i(G)$, $i = 1, \ldots, N$, a *mutant vector*, $V_i(G+1) = (v_{i1}, v_{i2}, \ldots, v_{iD})^\top$, is determined through the equation:

$$V_i(G+1) = X_{r_1}(G) + F\left(X_{r_2}(G) - X_{r_3}(G)\right), \tag{5}$$

where $r_1, r_2, r_3 \in \{1, \ldots, N\}$, are mutually different random indexes, and, $F \in (0, 2]$. The indexes r_1, r_2, r_3, also need to differ from the current index, i. Consequently, N must be greater than or equal to 4, in order to apply mutation.

Following the mutation phase, the *crossover* operator is applied on the population. Thus, a *trial vector*, $U_i(G+1) = (u_{i1}, u_{i2}, \ldots, u_{iD})^\top$, is generated, where,

$$u_{ij} = \begin{cases} v_{ij}, & \text{if } \ (\mathrm{randb}(j) \leqslant CR) \ \text{ or } \ j = \mathrm{rnbr}(i), \\ x_{ij}, & \text{if } \ (\mathrm{randb}(j) > CR) \ \text{ and } \ j \neq \mathrm{rnbr}(i), \end{cases} \tag{6}$$

where, $j = 1, 2, \ldots, D$; $\mathrm{randb}(j)$, is the j–th evaluation of a uniform random number generator in the range $[0, 1]$; CR is the (user specified) crossover constant in the range $[0, 1]$; and, $\mathrm{rnbr}(i)$ is a randomly chosen index from the set $\{1, 2, \ldots, D\}$.

To decide whether or not the vector $U_i(G+1)$ should be a member of the population comprising the next generation, it is compared to the initial vector $X_i(G)$. Thus,

$$X_i(G+1) = \begin{cases} U_i(G+1), & \text{if } \ f\left(U_i(G+1)\right) < f\left(X_i(G)\right), \\ X_i(G), & \text{otherwise.} \end{cases}$$

The procedure described above is considered as the standard variant of the DE algorithm. Different mutation and crossover operators have been applied with promising results [25]. In order to describe the different variants, the scheme DE/x/y/z, is used, where x specifies the mutated vector ("rand" for randomly selected individual or "best" for selection of the best individual); y is the number of difference vectors used; and, z denotes the crossover scheme (the scheme described here is due to independent binomial experiments, and thus, it is denoted as "bin") [25]. According to this description scheme, the DE variant described above is denoted as DE/rand/1/bin. One highly beneficial scheme that deserves special attention is the DE/best/2/bin scheme, where,

$$V_i(G+1) = X_{\mathrm{best}}(G) + F\left(X_{r_1}(G) + X_{r_2}(G) - X_{r_3}(G) - X_{r_4}(G)\right). \tag{7}$$

The usage of two difference vectors seems to improve the diversity of the population, if N is high enough. A parallel implementation of DE is reported in [26].

4 A Fuzzy Cognitive Map Model for Supervision of the Radiation Therapy Process

Fuzzy Cognitive Maps (FCMs) have been introduced by Kosko [27] to describe a cognitive map model with the following characteristics:

1. Causal relationships between nodes are fuzzified, i.e. instead of using only signs to indicate positive or negative causality, a number is associated with each relationship to express the degree of causality between two concepts.
2. The system is dynamic and has feedback, i.e. the effect of change in a concept node also affects other nodes, which in turn can affect the node initiating the change. The presence of feedback introduces temporality to the operation of FCMs.

The concepts reflect attributes, characteristics, and qualities of the system. The interconnections among the concepts signify the cause and effect relationships among the concepts. Let us denote by C_i, $i = 1, \ldots, M$, the nodes–concepts of an FCM. Each concept represents one of the key–factors of the system, and it takes a value $A_i \in [0, 1]$, $i = 1, \ldots, M$. Each interconnection between two concepts C_i and C_j, has a weight $w_{ij} \in [-1, 1]$, 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. There are three types of causal relationships among concepts: positive causality ($w_{ij} > 0$), negative causality ($w_{ij} < 0$), and no relation ($w_{ij} = 0$). So the FCM provides qualitative as well as quantitative information regarding the relationships among concepts [11].

In general, the value of each concept is calculated by aggregating the influence of the other concepts to the specific one [10], by applying the following rule:

$$A_i^{(t)} = f\left(A_i^{(t-1)} + \sum_{\substack{j=1 \\ j \neq i}}^{M} w_{ji} A_j^{(t)} \right), \tag{8}$$

where $A_i^{(t)}$ is the value of C_i at time t, and f is a sigmoid threshold function.

The methodology for developing FCMs primarily draws on a group of experts who are asked to define the concepts and describe the relationships among them. IF–THEN rules are used to describe the cause and effect relationships among the concepts, and infer a linguistic weight for each interconnection [10]. Each expert describes independently every interconnection with a fuzzy rule; the inference of the rule is a linguistic variable, which describes the relationship and determines the grade of causality between the corresponding concepts. Subsequently, the inferred fuzzy weights suggested by the experts, are aggregated to a single linguistic weight, which is transformed to a numerical weight, $w_{ij} \in [-1, 1]$, using the Center of Area (CoA) defuzzification method [13]. This weight represents the aggregated suggestion of the whole experts' group. Thus, an initial weight matrix, $W^{\text{initial}} = [w_{ij}]$, with $w_{ii} = 0$, $i = 1, \dots, M$, 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 defined in (8).

The most significant weaknesses of FCMs are the critical dependence on the experts' opinions, and the potential convergence to undesired steady states. 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. The desired steady state is characterized by values of the FCM's output concepts that are accepted by the experts [14].

Radiation therapy is a complex process involving a large number of treatment variables. The objective of radiotherapy is to deliver the highest radiation dose to the smallest possible volume that encloses the tumor, while retaining at a minimum the exposure of healthy tissues and critical organs to radiation. Treatment planning is another complex process that takes place before the final treatment execution. The performance criteria for this process include the maximization of the final dose received by the target volume (tumor), the maximization of the dose derived from the treatment planning within the target region, and dose minimization for the surrounding critical organs and normal tissues. To achieve these goals, several factors need to be taken into consideration [12,13].

In [12], an FCM with 33 concepts (factor–concepts, selector–concepts and output–concepts) has been developed, to model the aforementioned treatment planning and determine the dose distribution for the target volume, the healthy tissues and the critical organs. A different, more abstract FCM model is needed to supervise the whole radiotherapy process. This model must consist of more abstract concepts that represent the final parameters before the treatment execution, simulating, thus, the doctor's decision–making. In the proposed model [12], the supervision process is modeled with another FCM (supervisor–FCM) that models, monitors, and evaluates the whole process of radiation therapy. The supervisor–FCM is based on the knowledge of experts that supervise the actual process, and it consists of the following six concepts:

1. C_1–*Tumor localization*: It depends on the patient's contour, sensitive critical organs and tumor volume. It embodies the values and influences among these factor–concepts.
2. C_2–*Dose prescribed from the treatment planning*: This concept describes the prescribed dose and it depends on the concepts of the delivered dose to the target volume, normal tissues and critical organs, which are determined by the treatment planning model of the first level's FCM.
3. C_3–*Machine factors*: This concept describes the equipment characteristics.
4. C_4–*Human factors*: This is a general concept describing the experience and knowledge of the medical staff.
5. C_5–*Patient positioning and immobilization*: This concept describes the cooperation of the patient with the doctors and his willingness to follow their instructions.

6. C_6–*Final dose received by the target volume*: A measurement of the radiation dose received by the target tumor.

The supervisor–FCM has been developed following the methodology described previously in this section. Three oncologists were independently asked to describe the relationships among the concepts $C_1, \ldots, C_6$, using IF–THEN rules, and infer a linguistic weight for each interconnection [28]. The degree of influence is represented by a member of the fuzzy set,

$$\Big\{\text{ positive very high, positive high, positive medium, positive weak,}$$
$$\text{zero, negative weak, negative medium, negative low, negative very low }\Big\}.$$

The following connections among the concepts of the supervisor–FCM were suggested:

1. *Linkage 1*: Connects C_1 with C_6. It relates the tumor localization with the delivered final dose.
2. *Linkage 2*: Relates C_2 with C_1; when the dose derived from treatment planning is high, the value of tumor localization increases by a small amount.
3. *Linkage 3*: Connects C_2 with C_6; when the dose from treatment planning is high, the final dose given to the patient will be also high.
4. *Linkage 4*: Relates C_3 with C_2; when the machine parameters increase, the dose from treatment planning decreases.
5. *Linkage 5*: Connects C_3 with C_6; any change to machine parameters influences negatively the final dose given to the target volume.
6. *Linkage 6*: Relates C_4 with C_6; the human factors cause decrease in the final dose.
7. *Linkage 7*: Connects C_4 with C_5; the presence of human factors causes a decrease in the patient's positioning.
8. *Linkage 8*: Relates C_5 with C_4; any change on the patient positioning influences negatively the factors related to humans.
9. *Linkage 9*: Connects C_5 with C_6; when the patient positioning increases, the final dose also increases.
10. *Linkage 10*: Connects C_6 with C_5; when the final dose reaches an upper value, the patient positioning is influenced positively.
11. *Linkage 11*: Connects C_6 with C_1; any change in final dose causes change in tumor localization.
12. *Linkage 12*: Connects C_6 with C_2; when the final dose increases to an acceptable value, the dose from treatment planning also increases to a desired value.

After the determination of the linkages among concepts, experts suggested fuzzy values for the weights of the linkages. The fuzzy values were defuzzified and transformed in numerical weights, resulting in the following weight matrix:

$$W^{\text{supervisor}} = \begin{pmatrix} 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.43 \\ 0.28 & 0.00 & 0.00 & 0.00 & 0.00 & 0.57 \\ 0.00 & -0.30 & 0.00 & 0.00 & 0.00 & -0.39 \\ 0.00 & 0.00 & 0.00 & 0.00 & -0.32 & -0.43 \\ 0.00 & 0.00 & 0.00 & -0.37 & 0.00 & 0.68 \\ 0.22 & 0.67 & 0.00 & 0.00 & 0.54 & 0.00 \end{pmatrix}.$$

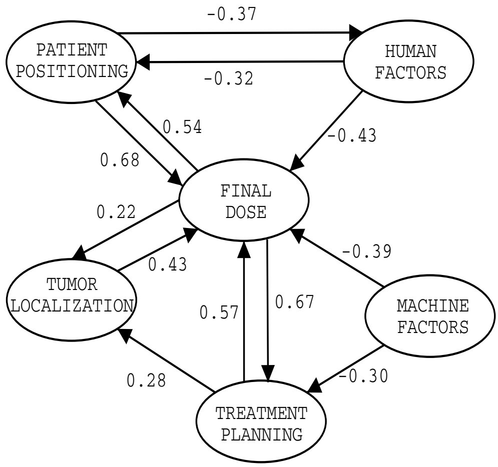

Fig. 1. The supervisor–FCM.

The final obtained supervisor–FCM is illustrated in Fig. 1.

The control objectives for the supervisor–FCM are to keep the amount of the final dose, FD, which is delivered to the patient, as well as the dose, D, prescribed by the treatment planning, within prespecified ranges:

$$FD_{\min} \leqslant FD \leqslant FD_{\max}, \tag{9}$$

$$D_{\min} \leqslant D \leqslant D_{\max}. \tag{10}$$

These objectives are defined by the related AAPM and ICRP protocols [1,10, 11], for the determination of accepted dose levels for each organ and part of the human body. The supervisor–FCM evaluates the success or failure of the treatment by monitoring the value of the "Final Dose" concept. Successful treatment corresponds to values of the final dose that lie within the desired bounds. The value of final dose identifies the supervisor model [29]. The supervisor–FCM has been incorporated to an integrated two–level hierarchical decision making system for the description and determination of the specific treatment outcome and for scheduling the treatment process before its treatment execution [12]. Thus, optimizing the supervisor–FCM, i.e. detecting the weights that correspond to the maximum values of the concepts FD and D, within their prespecified ranges, results in an enhanced control system which models the radiotherapy procedure more accurately and makes decision–making more reliable.

5 The Proposed Approach and Results

It has already been mentioned that the optimization of the supervisor–FCM described in Section 4, enhances the simulation ability of the system, resulting in more reliable decision–making. For this purpose, the PSO and DE algorithms have been used for the optimization of the supervisor–FCM, through the minimization of an appropriate objective function. The selection of these algorithms was based solely on their superior performance on a diverse field of applications, as well as due to the minimum effort required for their implementation.

The objective function can be straightforwardly defined as:

$$f(W) = -FD(W) - D(W), \tag{11}$$

where $FD(W)$ and $D(W)$ are the values of the final dose and the dose prescribed from the treatment planning, respectively, that correspond to the weight matrix W. The minus signs are used to transform the maximization problem to its equivalent minimization problem. Thus, the main optimization problem under consideration is the minimization of the objective function f, such that the constraints (9) and (10) hold.

The weight matrix W can, in general, be represented by a vector which consists of the rows of W in turn, excluding the elements of its main diagonal, $w_{11}, w_{22}, \ldots, w_{MM}$, which are by definition equal to zero. In the supervisor–FCM, the experts determined only 12 linkages, as described in Section 4, and thus, the corresponding minimization problem is 12–dimensional. Moreover, the bounds determined for the parameters FD and D, have been determined by the three radiotherapy oncologists (experts) in the following ranges:

$$0.90 \leqslant FD \leqslant 0.95, \tag{12}$$
$$0.80 \leqslant D \leqslant 0.95. \tag{13}$$

Taking into consideration the fuzzy linguistic variables that describe the cause–effect relationships among the concepts, as they were suggested by the three experts, the following ranges for the weight values were determined:

$$\begin{array}{lll} 0.3 \leqslant w_{16} \leqslant 0.5, & -0.4 \leqslant w_{36} \leqslant -0.1, & 0.5 \leqslant w_{56} \leqslant 0.8, \\ 0.2 \leqslant w_{21} \leqslant 0.4, & -0.5 \leqslant w_{45} \leqslant -0.2, & 0.2 \leqslant w_{61} \leqslant 0.4, \\ 0.5 \leqslant w_{26} \leqslant 0.7, & -0.5 \leqslant w_{46} \leqslant -0.2, & 0.6 \leqslant w_{62} \leqslant 0.9, \\ -0.4 \leqslant w_{32} \leqslant -0.2, & -0.6 \leqslant w_{54} \leqslant -0.1, & 0.5 \leqslant w_{65} \leqslant 0.9. \end{array} \tag{14}$$

These ranges were incorporated as constraints on the parameter vector in the experiments conducted.

The two most common variants of PSO and DE were used in the experiments. Specifically, the local versions of both the constriction factor and the inertia weight PSO variant, as well as the DE/rand/1/bin and DE/best/2/bin DE variants were used. Default values of the PSO parameters were used: $\chi = 0.729$, $c_1 = c_2 = 2.05$, and w decreasing from 1.2 to 0.1 [19,22]. Regarding the DE parameters, the values $F = 0.5$ and $CR = 0.5$ were selected after a trial and error process. The swarm (or population) size was always equal to 50. For each algorithm variant, 100 independent experiments were performed, to enforce the reliability of the results. Each algorithm was allowed to perform 1000 iterations (generations) per experiment.

The same solution was obtained by all algorithms in every experiment:

$$W^* = \begin{pmatrix} 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.5 \\ 0.4 & 0.0 & 0.0 & 0.0 & 0.0 & 0.7 \\ 0.0 & -0.2 & 0.0 & 0.0 & 0.0 & -0.1 \\ 0.0 & 0.0 & 0.0 & 0.0 & -0.2 & -0.2 \\ 0.0 & 0.0 & 0.0 & -0.6 & 0.0 & 0.8 \\ 0.4 & 0.9 & 0.0 & 0.0 & 0.9 & 0.0 \end{pmatrix},$$

which corresponds to the following final state of the concepts after the convergence of the FCM:

$$C_1^* = 0.819643, \quad C_2^* = 0.819398, \quad C_3^* = 0.659046,$$

$$C_4^* = 0.501709, \quad C_5^* = 0.824788, \quad C_6^* = 0.916315.$$

Obviously, the values of C_2 and C_6 lie within the desired regions defined by the relations (12) and (13), while the weights fulfill the constraints (14) posed by the experts. This result supports the claim that the obtained solution seems to be the true optimal solution.

An interesting remark is the high positive influence of the concept C_6 (final dose) to the concept C_2 (dose prescribed from the treatment planning) as well as to the concept C_5 (patient positioning). This means that if we succeed to deliver the maximum dose to the target volume, then the initial calculated dose from treatment planning is the desired and the same happens with patient positioning. Another interesting fact is that the estimated weights assume their optimum values at the edges of the suggested fuzzy sets. This behavior has been also identified by other researchers [30,31].

The optimal values of “Final Dose” and “Dose Prescribed from the Treatment Planning” are acceptable according to the ICRU protocols [32,33], optimizing the whole treatment process. This supports the claim that the proposed approach is efficient and useful for the FCM–controlled radiation therapy process.

6 Conclusions

A Fuzzy Cognitive Map model, which supervises and monitors the radiotherapy process, resulting in a sophisticated decision support system, is optimized using the Particle Swarm Optimization and the Differential Evolution algorithms. The objective of the radiation treatment procedure is to give the acceptable–optimum amount of delivered dose to the target volume. The proposed methods determine the cause–effect relationships among concepts, determining the optimal weight matrix for the supervisor–FCM model.

Extensive experiments were performed, using different variants of the two stochastic optimization algorithms, always resulting in the same solution, which satisfies all optimality criteria and constraints imposed by the experts. This numerical evidence supports the claim that the obtained solution can be considered as an optimal by the user. The results contribute towards the direction of more reliable decision support system for radiation therapy.

Future work will focus on the optimization of the generic hierarchical system in all levels of radiation therapy, taking also into consideration the treatment planning (low–level) model of the system.

Acknowledgment. We thank the anonymous referees for their useful comments and suggestions. Also, we acknowledge the partial support by the "Pythagoras" research grant, awarded by the Greek Ministry of Education and Religious Affairs, as well as, the European Union.

References

1. Khan, F.: The Physics of Radiation Therapy. Williams & Wilkins, Baltimore (1994)
2. Brahme, A.: Optimization of radiation therapy and the development of multileaf collimation. Int. J. Radiat. Oncol. Biol. Phys. **25** (1993) 373–375
3. Brahme, A.: Optimization of radiation therapy. Int. J. Radiat. Oncol. Biol. Phys. **28** (1994) 785–787
4. Gibbons, J.P., Mihailidis, D.N., Alkhatib, H.A.: A novel method for treatment plan optimisation. In: Proc. 22–nd Ann. Int. Conf. IEEE Engin. in Med. and Biol. Soc. Volume 4. (2000) 3093–3095
5. Mageras, G.S., Mohan, R.: Application of fast simulated annealing to optimization of conformal radiation treatments. Med. Phys. **20** (1993) 639–647
6. Starkschall, G., Pollack, A., Stevens, C.W.: Treatment planning using dose–volume feasibility search algorithm. Int. J. Radiat. Oncol. Biol. Phys. **49** (2001) 1419–1427
7. Brahme, A.: Treatment optimization using physical and biological objective functions. In Smith, A., ed.: Radiation Therapy Physics. Springer, Berlin (1995) 209–246
8. Rowbottom, G., Khoo, V.S., Webb, S.: Simultaneous optimization of beam orientations and beam weights in conformal radiotherapy. Med. Phys. **28** (2001) 1696–1702
9. Soderstrom, S.: Radiobiologically Based Optimization of External Beam Radiotherapy Techniques Using a Small Number of Fields. PhD thesis, Stockholm University, Stockholm, Sweden (1995)
10. Wells, D., Niederer, J.: A medical expert system approach using artificial neural networks for standardized treatment planning. Int. J. Radiat. Oncol. Biol. Phys. **41** (1998) 173–182
11. Willoughby, T., Starkschall, G., Janjan, N., Rosen, I.: Evaluation and scoring of radiotherapy treatment plans using an artificial neural network. Int. J. Radiat. Oncol. Biol. Phys. **34** (1996) 923–930
12. Papageorgiou, E.I., Stylios, C.D., Groumpos, P.P.: 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.
13. Papageorgiou, E.I., Stylios, C.D., Groumpos, P.P.: Decision making in external beam radiation therapy based on fuzzy cognitive maps. In: Proc. 1st Int. IEEE Symp. Intelligent Systems, Varna, Bulgaria (2002)
14. Parsopoulos, K.E., Papageorgiou, E.I., Groumpos, P.P., Vrahatis, M.N.: A first study of fuzzy cognitive maps learning using particle swarm optimization. In: Proceedings of the IEEE 2003 Congress on Evolutionary Computation, Canberra, Australia, IEEE Press (2003) 1440–1447

15. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: Proceedings IEEE International Conference on Neural Networks. Volume IV., Piscataway, NJ, IEEE Service Center (1995) 1942–1948
16. Kennedy, J., Eberhart, R.C.: Swarm Intelligence. Morgan Kaufmann Publishers (2001)
17. Parsopoulos, K.E., Vrahatis, M.N.: Recent approaches to global optimization problems through particle swarm optimization. Natural Computing **1** (2002) 235–306
18. Eberhart, R.C., Simpson, P., Dobbins, R.: Computational Intelligence PC Tools. Academic Press (1996)
19. Clerc, M., Kennedy, J.: The particle swarm–explosion, stability, and convergence in a multidimensional complex space. IEEE Trans. Evol. Comput. **6** (2002) 58–73
20. Shi, Y., Eberhart, R.C.: A modified particle swarm optimizer. In: Proceedings IEEE Conference on Evolutionary Computation, Anchorage, AK, IEEE Service Center (1998) 69–73
21. Shi, Y., Eberhart, R.C.: Parameter selection in particle swarm optimization. In Porto, V.W., Saravanan, N., Waagen, D., Eiben, A.E., eds.: Evolutionary Programming. Volume VII. Springer (1998) 591–600
22. Trelea, I.C.: The particle swarm optimization algorithm: Convergence analysis and parameter selection. Information Processing Letters **85** (2003) 317–325
23. Parsopoulos, K.E., Plagianakos, V.P., Magoulas, G.D., Vrahatis, M.N.: Objective function "stretching" to alleviate convergence to local minima. Nonlinear Analysis, Theory, Methods & Applications **47** (2001) 3419–3424
24. Parsopoulos, K.E., Vrahatis, M.N.: Initializing the particle swarm optimizer using the nonlinear simplex method. In Grmela, A., Mastorakis, N., eds.: Advances in Intelligent Systems, Fuzzy Systems, Evolutionary Computation. WSEAS Press (2002) 216–221
25. Storn, R., Price, K.: Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. J. Global Optimization **11** (1997) 341–359
26. Plagianakos, V.P., Vrahatis, M.N.: Parallel evolutionary training algorithms for "hardware–friendly" neural networks. Natural Computing **1** (2002) 307–322
27. Kosko, B.: Fuzzy cognitive maps. Int. J. Man–Machine Studies **24** (1986) 65–75
28. Jang, J.S., Sun, C.T., Mizutani, E.: Neuro–Fuzzy and Soft Computing. Prentice Hall, Upper Saddle River, NJ (1997)
29. Dechlich, F., Fumasoni, K., Mangili, P., Cattaneo, G.M., Iori, M.: Dosimetric evaluation of a commercial 3–d treatment planning system using report 55 by aapm task group 23. Radiotherap. Oncol. **52** (1999) 69–77
30. Tsadiras, A.K., Margaritis, K.G.: Cognitive mapping and certainty neuron fuzzy cognitive maps. Information Sciences **101** (1997) 109–130
31. Tsadiras, A.K., Margaritis, K.G.: An experimental study of the dynamics of the certainty neuron fuzzy cognitive maps. Neurocomputing **24** (1999) 95–116
32. Determination of absorbed dose in a patient irradiated by beams of x or gamma rays in radiotherapy procedures. Technical Report 24, International Commission on Radiation Units and Measurements, Washington, USA (1976)
33. Prescribing, recording and reporting photon beam therapy. Technical Report 50, International Commission on Radiation Units and Measurements, Washington, USA (1993)