

Automated Radiotherapy Treatment Planning using Fuzzy Inference Systems

Joana Dias^{1,2}, Humberto Rocha^{1,2}, Tiago Ventura³, Brígida Ferreira⁴, Maria do Carmo Lopes^{3,5}

¹CeBER and Faculdade de Economia, Universidade de Coimbra, 3004-512 Coimbra, Portugal

²Inesc-Coimbra, Rua Sílvio Lima, Pólo II, 3030-290 Coimbra, Portugal

³Serviço de Física Médica, IPOC-FG, EPE, 3000-075 Coimbra, Portugal

⁴School for Allied Health Technologies, Porto, Portugal

⁵I3N, Departamento de Física, Universidade de Aveiro, 3810-193 Aveiro, Portugal

joana@fe.uc.pt, hrocha@mat.uc.pt,
tiagoventura@ipocoimbra.min-saude.pt, brigida@ua.pt,
mclopes@ipocoimbra.min-saude.pt

Abstract. Radiotherapy is one of the treatments available for cancer patients, aiming to irradiate the tumor while preserving healthy structures. The planning of a treatment is a lengthy trial and error procedure, where treatment parameters are iteratively changed and the delivered dose is calculated to see whether it complies with the desired medical prescription. In this paper, a procedure based on fuzzy inference systems (FIS) for automated treatment planning is developed, allowing the calculation of high quality treatment plans without requiring human intervention. The procedure is structured in two different phases, incorporating the automatic selection of the best set of equidistant beam irradiation directions by an enumeration procedure. The developed method is extensively tested using ten head-and-neck cancer cases.

Keywords: Fuzzy Inference Systems, Radiotherapy Planning, IMRT, Optimization, Decision Problem

1 Introduction

Radiation therapy is, along with surgery and chemotherapy, one of the main treatments for cancer patients. Radiation has the capability of damaging and killing cells. The main objective is to be able to kill all cancer cells (that have more difficulties recovering from radiation), and at the same time spare as most as possible healthy cells. The treatment is usually performed applying a linear accelerator, mounted on a gantry that can rotate along a central axis parallel to the couch where the patient lays immobilized. The rotation of the couch along with the rotation of the gantry allows radiation to be delivered from almost any direction (angle) around the tumor. Nevertheless, most of the times, it is the equidistant coplanar angle configuration that is considered (only angles that lay in the plane of rotation of the gantry are considered). There are different radiotherapy treatment modalities: 3D-conformal, Intensity Modu-

lated Radiation Therapy (IMRT), Arc-Therapy, for instance. In this work, our focus is on IMRT treatments, although the developed methodology can also be extended to other radiotherapy modalities. In IMRT it is possible to modulate the radiation intensity due to the use of a multileaf collimator, that is composed of pairs of individual leaves, moving independently and blocking the radiation beam. This allows not only the conformal shaping of the treatment beams to the tumor shape, but also the possibility of achieving radiation intensity maps: each radiation beam can be interpreted as a set of individual beamlets. Before a radiotherapy treatment can be delivered, CT-scans are taken allowing a 3D discretization of the patient into voxels, and a delineation of all important structures (Planning Target Volumes –PTV– and healthy organs that are meant to be spared known as Organs at Risk – OAR). A radiation dose prescription defines minimum dose coverage constraints for the PTVs and maximum allowed radiation doses to OARS. Having as input the 3D discretization of the patient and the desired prescription, the treatment planner will try to generate a treatment plan. The treatment planning procedure uses software that simulates the treatment and calculates the radiation dose that is being deposited in each voxel (Treatment Planning System – TPS). The planner needs to define *a priori* the radiation beam angles and a set of parameters that are TPS dependent, but are usually weights and lower and/or upper bounds. With these parameters fixed, the TPS will run an optimization procedure (IMRT Fluence Map Optimization – FMO) that will generate the optimal radiation intensity associated with each beamlet from each of the angles to be used in the treatment (fluence maps), so that the dose that is being deposited in each voxel (measured in Gy) is calculated. The commonly used trial-and-error treatment planning procedure is a lengthy and cumbersome process, taking from several hours to several days for each patient. The quality of the treatment plan is highly dependent on the planner, and it is not possible to know whether more time and effort will be rewarded with a better quality treatment plan. Automating the treatment planning procedure will constitute an important breakthrough in this area. Probably the best known and general developed approach so far is *iCycle* [1] and other related works [2, 3]. In this paper, we present an approach where the trial-and-error procedure is replaced by an automated procedure that optimizes fluence maps by using FIS. The way in which the planner drives the TPS aiming at calculating better plans is based on his own reasoning and it is not possible to describe this process in a rigorous mathematical form. Moreover, two different planners will most probably have different reasoning. Nevertheless, it is possible to think of some simple rules that will certainly be followed. If, for instance, structure *A* is an OAR and it is receiving a higher dose than what is desired, then the importance of this structure in the FMO should be increased. This kind of simple rules can be defined for each structure and can be the basis of an automated procedure that guides the search process towards an acceptable treatment plan. Fuzzy logic has been applied before to radiotherapy planning. Li and Yin [4] apply fuzzy logic to determine the best prescription for the normal tissue. Yan et al [5, 6] apply a fuzzy inference system (FIS) composed of eight rules that changes the *a priori* defined weights assigned to each structure. This work is further extended in [7] by a neuro-fuzzy inference system that uses a trained neural network to determine the parameters of the fuzzy inference system. Our approach is different from the cited

works since it mimics the planner's trial-and-error procedure, and the implicit rules that the planner uses, in an automated way and automatically changing all the parameters available and not only weights. Furthermore, if a feasible treatment plan is found, the algorithm will try to automatically improve this plan by increasing PTV coverage and/or achieving better organ sparing. If a feasible plan is not found, then it will calculate a plan that complies as much as possible with the medical prescription. The present approach builds on the work of Dias et al. [8], adapting it so that it accommodates the choice of the best equidistant beam angle geometry, without increasing the planner's workload. Furthermore, extensive computational tests were performed to assess the effect of using different parameters in the fuzzy inference system. The proposed methodology is applied to ten retrospectively treated head-and-neck cancer cases at the Portuguese Oncology Institute of Coimbra (IPOCFG). In Section 2, the IMRT treatment planning procedure is explained. In section 3, the FMO is detailed. The algorithm is described in section 4. In section 5, computational results are shown, and discussed in section 6.

2 IMRT Treatment Planning

The aim of IMRT treatment planning is to achieve a treatment plan that satisfies all dosimetry treatment constraints. The dosimetry constraints can take a variety of forms, depending on whether they relate to OAR or PTV. Considering OAR, for instance, there are generally constraints that limit the maximum or mean dose that the OAR can receive. The constraints to consider will depend on the organs' functionality. OARs can be classified as being serial or parallel organs. For serial organs even if only a small percentage of the organ is over-irradiated then the whole organ's functionality is jeopardized (spinal cord, for instance). It is thus necessary to guarantee that the maximum dose received anywhere within the organ is upper bounded. Other organs can still perform their function even if a small percentage of the organ is damaged (lung, for instance). In this case, a dose volume condition has to be preserved and can be expressed, for instance, by guaranteeing that the mean dose received is not above a given threshold. Regarding PTV, the medical prescription defines a desired dose and also dose-volume constraints. As an example, consider a patient that has different PTV structures that can be clustered in two groups: PTVs where the desired dose is 70Gy (PTV70) and PTVs where the desired dose is 59.4Gy (PTV59). Consider also that structures belonging to PTV70 are inside structures belonging to PTV59. For each of the PTV structures, a constraint could state that at least 95% of the PTV volume ($D_{95\%}$) receives at least 95% of the prescribed dose. For PTVs, maximum doses should not be exceeded. No PTV voxel should receive, for instance, more than 107% of the prescribed dose ($v_{107\%}$). If PTV70 is inside PTV59, then it is not possible to consider that no voxel in PTV59 will receive more than 63.6Gy since we aim at achieving 70Gy for PTV70. In this case, the percentage of voxels in PTV59 that are allowed to receive more than the maximum desired dose is limited. Table 1 presents an example of such dosimetry constraints. According to Table 1, a treatment plan will be considered in accordance with the medical prescription if and only if:

- There is no voxel belonging to the spinal cord receiving more than 45Gy;
- There is no voxel belonging to the brainstem receiving more than 54Gy;
- The mean dose in the parotids does not exceed 26Gy;
- 95% of the voxels in PTV₇₀ are receiving at least 66.5Gy;
- No voxel belonging to PTV₇₀ receives more than 74.9Gy;
- 95% of the voxels in PTV₅₉ are receiving at least 56.4Gy;
- The percentage of voxels in PTV₅₉ that are allowed to receive more than 107% of the prescribed dose are limited to the percentage of PTV₇₀ volume inside PTV₅₉ plus a 10% margin.

In the planning of every single treatment, the questions that need to be answered are: *Is this medical prescription attainable? If it is, what is the treatment plan that complies with all the constraints? If it is not, how can we comply as much as possible with the medical prescription?*

Table 1. Prescribed doses for the structures considered for FMO

Structure	Type of constraint		Limit
Spinal cord	Maximum dose	Lower than	45 Gy
Brainstem	Maximum dose	Lower than	54 Gy
Left parotid	Mean dose	Lower than	26 Gy
Right parotid	Mean dose	Lower than	26 Gy
PTV ₇₀	$D_{95\%}$	Greater than	66.5 Gy
PTV ₇₀	Maximum dose	Lower than	74.9 Gy
PTV ₅₉	$D_{95\%}$	Greater than	56.4 Gy
PTV ₅₉	$V_{107\%}$	Lower than	Percentage of PTV ₇₀ volume inside PTV ₅₉ plus a 10% margin
Body	Maximum dose	Lower than	80 Gy

The objective of a treatment planning procedure is thus to reach a solution that satisfies all the inequalities defined by the dose-volume constraints. It is not an optimization problem that needs to be solved, in the sense that no objective function needs to be defined. It is rather a feasibility problem, in the sense that what the planner really needs is to find a feasible solution. One of the difficulties encountered with this feasibility problem is the fact that the dose-volume constraints that have to be considered are not simple to represent in tractable mathematical programming models. To tackle this problem we will consider a procedure that iteratively solves an unconstrained optimization problem with parameters being changed automatically.

3 Fluence Map Optimization

The therapeutic radiation dose is delivered through a set of radiation beams usually determined *a priori*. Each beam can be interpreted as being discretized in a set of beamlets. It is then necessary to determine the intensity of each of these beamlets, defining the fluence map (FMO problem). FMO has been mainly tackled by constrained optimization models such that an objective function is optimized subject to a

set of dose requirements [9-14]. The objective function usually considers a weighted sum of deviations from prescribed doses (underdose for PTVs and overdose for OARs). It is possible to find examples of linear models [15], mixed integer linear models [16], nonlinear models [17, 18], multiobjective models [19] and parallel programming approaches [20, 21]. In this paper, a voxel-based convex penalty non-linear model for FMO is used. The objective function considers the minimization of the sum of the penalties associated with each voxel, calculated as the square difference between the amount of dose received by the voxel and a given upper and/or lower bound. There are only linear nonnegativity constraints on the intensity (fluence) values. Considering that beam angles have already been fixed, let V represent the number of voxels, N the number of beamlets and D the dose matrix, such that D_{ij} represents the contribution of beamlet j to the total dose deposited in voxel i . The total dose received by voxel i can be calculated as $\sum_{j=1}^N D_{ij} w_j$ with w_j representing the intensity of beamlet j . Let U_i be the upper bound associated with voxel i , L_i the lower bound associated with voxel i , $\underline{\lambda}_i$ and $\bar{\lambda}_i$ the penalty weights of underdose and overdose of voxel i , respectively. The FMO model can be defined as follows, where $(\bullet)_+ = \max\{0, \bullet\}$:

$$f(w) = \text{Min}_w \sum_{i=1}^V \left[\underline{\lambda}_i \left(L_i - \sum_{j=1}^N D_{ij} w_j \right)_+^2 + \bar{\lambda}_i \left(\sum_{j=1}^N D_{ij} w_j - U_i \right)_+^2 \right] \quad (1)$$

$$\text{s.t. } w_j \geq 0, j = 1, \dots, N \quad (2)$$

Although this formulation allows unique weights and unique upper/lower bounds for each voxel, in the current approach weights are assigned by structure only. Considering a given structure S we can thus define $\bar{\lambda}_S$, $\underline{\lambda}_S$, U_S and L_S such that $\bar{\lambda}_i = \bar{\lambda}_S, \forall i \in S$, $\underline{\lambda}_i = \underline{\lambda}_S, \forall i \in S$, $U_i = U_S, \forall i \in S$ and $L_i = L_S, \forall i \in S$. This nonlinear formulation implies that a very small amount of deviation from the established bounds may be accepted, but larger deviations are decreasingly tolerated. We have chosen to work with this model for two main reasons: (1) the fact that it is an unconstrained problem guarantees that it will always be possible to find an admissible solution for the optimization problem (although it may not be feasible for the original treatment planning problem); (2) it is a convex problem that can be solved by known optimization algorithms. The main disadvantage of this model consists in the difficulty of predicting the impact that changes in the model parameters (namely weights and bounds) will have in the optimal solution calculated. Although this FMO model was chosen, the developed methodology could be extended to deal with different FMO models. It should be stressed that the value of the objective function has absolutely no clinical meaning whatsoever. The quality of the solution is not assessed by looking at the FMO objective function value but rather by looking at the dosimetry values. This means that the FMO objective function should be interpreted as nothing more than a

technical tool that guides the optimization algorithm to interesting regions of the solution space, where feasible solution of the treatment planning problem can be found. If a planner were to work with such a FMO through a TPS, he would be asked to define *a priori* all the model's parameters: weights and upper and lower bounds. Looking at the objective function (1), and considering a given current solution w , how would the planner change the model's parameters to obtain a better solution? Let's imagine that for a given OAR S the maximum allowed dose in the current solution is not being respected. Then, in order for the optimization algorithm to produce a solution that spares S more than it is being spared, two things can be done: either one increases the upper weight associated with S ($\bar{\lambda}_S$) or one changes (decreases) the corresponding upper bound (U_S). Both options will increase $\bar{\lambda}_i \left(\sum_{j=1}^N D_{ij} w_j - U_i \right)_+^2$ for all voxels $i \in S$ that fail to respect the prescribed dose. Although it is easy enough to see the direction into which the model's parameters should be changed in order to reach solutions that are more in accordance with the desired medical prescription, there is not a straightforward way of automatically determining the best change in the parameters. It is expected that larger deviations between the dosimetry results and the desired doses should correspond to more pronounced changes in the parameters. The main idea in the developed methodology is to drive this iterative procedure by resorting to FIS, that will mimic the reasoning of the human planner.

4 Automated FMO by Fuzzy Inference Systems

The proposed methodology will automatically change the FMO model's parameters in each iteration, by considering the deviations between the current and the desired dosimetry results. These changes will drive the FMO algorithm to search more promising regions of the searchable surface, until a plan complying as much as possible with the desired prescription is attained.

4.1 Initialization

The initial FMO model's parameters are defined by the dosimetry constraints. OARs, for instance, will only have upper bounds and will not have any lower bounds. Upper bounds are equal to the maximum dose or to the mean dose allowed, depending on the type of OAR. Regarding PTVs, both upper and lower bounds are defined. Upper bounds are usually considered as being equal to the maximum allowed dose (107% of the prescribed dose, for instance). Lower bounds are equal to the minimum required dose. Regarding the weights to assign to each structure, every structure begins by having a weight equal to 1. The only exception to this rule is when PTVs are inside other PTVs. Actually, the inner PTVs are subject to contradicting constraints, since each voxel will belong to more than one structure. In this case, and if the inner PTV is much smaller than the outer PTV, it is better to give an increased weight to the inner PTV. It should be pointed out that these weights will be automatically updated by the

algorithm, so the initial weights will not have an impact in the quality of the final outcome (they can, however, have an impact in the total computational time, and this initialization will simply give the algorithm a better starting point). Weight initialization different than 1 could also be thought for special cases of very tiny OARs that are in close proximity to other OARs or PTVs (the case of crystalline lens, for instance), or for OARs that the planner knows that will be especially difficult to spare. It should be stressed that all weights could be initialized to 1: the algorithm will update these weights if necessary, at the cost of increased computational time.

4.2 Fuzzy Inference Systems

At each iteration, the FMO is optimized and dosimetry results are calculated. For each structure, the algorithm will check whether the defined constraints are being satisfied or not, and changes the model parameters in accordance. Upper/lower bounds are updated first, and weights are only changed if updating bounds proves not to be sufficient to attain the desired goals. Actually, changing bounds will potentially affect a greater number of voxels, therefore the penalty associated with the structure will be greater than changing only weights. When only weights are changed there will be a smaller number of voxels contributing to the objective function value. For a given structure S , if there is a deviation equal to d between the desired dosimetry metric and the current one, the corresponding bounds U_s and L_s are changed according to very simple common sense rules: if d is *large* then increase (decrease) L_s (U_s) by a *large* amount; if d is *medium* then increase (decrease) L_s (U_s) by a *medium* amount; if d is *small* then increase (decrease) L_s (U_s) by a *small* amount. As it is not possible to define in a crisp way the notion of *large*, *medium* or *small*, these linguistic concepts are represented through the use of fuzzy sets and FIS is used to determine how much U_s and L_s should be changed. These concepts are represented resorting to triangular and trapezoidal membership functions, where d is measured as a percentage in relation to the desired value and the change in the bound is also measured in percentage. Linguistic concepts related to the current deviation are defined in equations (3) to (5), and are also depicted in Fig. 1 (considering $a=5$). Equations (6) to (8) define the membership functions representing the linguistic concepts associated with the change in the bounds (depicted in Fig. 3, for $b=10$). A Mamdani type implication rule was considered [22]. This means that the Mamdani implication operator (min operator) is applied between the resulting antecedent membership function and the consequent membership function for all three fuzzy rules. The defuzzification technique is the

Centroid Defuzzification Technique (center of gravity), calculated as
$$\frac{\int \mu(z) \cdot z \partial z}{\int \mu(z) \partial z},$$

where $\mu(z)$ represents the function that results from the fuzzy inference system.

$$\mu_{d_small}(x) = \begin{cases} \frac{a-x}{a}, & x < a\% \\ 0, & x \geq a\% \end{cases} \quad (3)$$

$$\mu_{d_medium}(x) = \begin{cases} 0, & x \leq 2\% \\ \frac{x-(a-3)}{4}, & (a-3)\% < x \leq (a+1)\% \\ \frac{(a+5)-x}{4}, & (a+1)\% < x < (a+5)\% \\ 0, & x \geq (a+5)\% \end{cases} \quad (4)$$

$$\mu_{d_large}(x) = \begin{cases} 0, & x \leq (a+2)\% \\ \frac{x-(a+2)}{5}, & (a+2)\% < x \leq (a+7)\% \\ 1, & x > (a+7)\% \end{cases} \quad (5)$$

$$\mu_{c_small}(x) = \begin{cases} \frac{b-x}{b}, & x < b\% \\ 0, & x \geq b\% \end{cases} \quad (6)$$

$$\mu_{c_medium}(x) = \begin{cases} 0, & x \leq (b-5)\% \\ \frac{x-(b-5)}{5}, & (b-5)\% < x \leq b\% \\ \frac{(b+5)-x}{5}, & b\% < x < (b+5)\% \\ 0, & x \geq (b+5)\% \end{cases} \quad (7)$$

$$\mu_{c_large}(x) = \begin{cases} 0, & x \leq b\% \\ \frac{x-b}{10}, & b\% < x \leq (b+10)\% \\ 1, & x > (b+10)\% \end{cases} \quad (8)$$

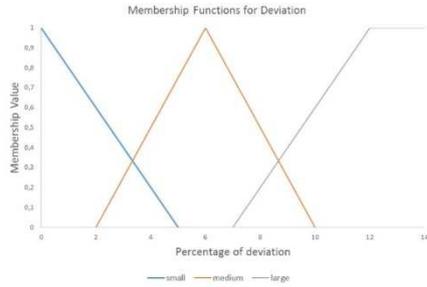


Fig. 1. Percentage of Deviation Membership Functions

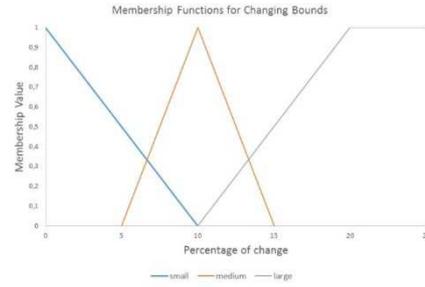


Fig. 2. Percentage of Change in the Bound Membership functions

The chosen membership functions result in an algorithm behaviour that will change the bounds in a continuous and smooth way. If different thresholds are considered (different values for a and b), more or less accentuated changes in bounds will be achieved that could lead to either similar results in less computational time or difficulties in achieving convergence towards a feasible solution. In many situations, it is possible to achieve admissible treatment plans by changing upper and lower bounds only. However, in more complicated cases, it will also be necessary to change structures' weights, namely when an upper bound is reaching very low values, or when a

lower bound is increasing towards very large values. In this case, the algorithm will change the corresponding structure's weights, also resorting to a set of fuzzy rules, and following the same reasoning as before. Weights $\bar{\lambda}_s$ are changed whenever U_s reaches less than $l\%$ of its initial value. Weights $\underline{\lambda}_s$ are changed whenever L_s reaches more than u times its original value. Whenever the weights are updated, the upper and lower bounds take their initial values. If we increase (decrease) the l value (u value) for updating $\bar{\lambda}_s$ ($\underline{\lambda}_s$), then the weights will be updated more often. The weights are changed by using fuzzy rules similar to the ones already described: the current deviation is characterized as being *small*, *medium* or *large* according to (3)-(5). The change in the weights is given as a nominal change determined by the membership functions depicted in equations (9)-(11). Parameter c can be used to change the behaviour of the FIS regarding the magnitude of changes in the weights.

$$\mu_{w_small}(x) = \begin{cases} 1, & x \leq c \\ \frac{c-x}{2}, & c < x \leq (c+2) \\ 0, & x \geq (c+2) \end{cases} \quad (9)$$

$$\mu_{w_medium}(x) = \begin{cases} 0, & x \leq (c-2) \\ \frac{x-(c-2)}{3}, & (c-2) < x \leq (c+2) \\ \frac{(c+5)-x}{3}, & (c+2)\% < x < (c+5) \\ 0, & x \geq (c+5) \end{cases} \quad (10)$$

$$\mu_{w_large}(x) = \begin{cases} 0, & x \leq (c+2) \\ \frac{x-(c+2)}{2}, & (c+2) < x \leq (c+4) \\ 1, & x > (c+4) \end{cases} \quad (11)$$

4.3 The whole procedure

After initialization, the algorithm will run for a predetermined number of iterations ($Nmax$) or until a plan complying with all the desired dosimetry results is reached. At each iteration, the bounds and weights are changed, in accordance with the deviations between the current and the desired solution, and the FMO is solved again. If $Nmax$ is reached without an admissible treatment plan being found, then the algorithm will choose to relax some of the dosimetry constraints, trying to look for a plan that will not comply completely with the desired specifications, but that will comply as much as possible. The way the algorithm relaxes the dosimetry constraints is guided by the planner's preferences. The planner can choose, from the existing constraints, which of them he is more willing to relax. He can also assign a priority to each of these constraints, so that the highest the priority, the less relaxed the constraint will be. The relaxation of a given constraint is considered by calculating an admissibility slack for that constraint, such that if the deviation between the desired and the current dosimetry value is within that slack the constraint is considered as being satisfied. The value of this slack is calculated as a percentage value by using a FIS that takes as input the priority of the corresponding structure (the higher the priority the lowest the slack).

The set of rules that are considered are as follows: if the priority is *high* then the slack is *small*; if the priority is *low* then the slack is *large*; if the priority is *medium* then the slack is *medium*. Linguistic concepts *high*, *low*, *medium*, *small*, *large* are defined in a similar manner as similar concepts already described. If the algorithm was able to comply with all the defined constraints before reaching the maximum number of iterations, then it will try to improve even further the treatment plan. This means that a new feasibility problem is considered, with constraints that are harder to satisfy. This can be done in two phases, also considering the planner's preferences. If the planner thinks it is important to try to obtain a better PTV coverage, then he can signalize the PTVs as being structures to improve. The algorithm will be more demanding regarding PTV coverage than initially defined. The path that the algorithm takes in attempting to increase PTV coverage can be configured by the planner. As an example: the PTV constraint of having 95% of the volume receiving at least 95% of the prescribed dose is changed to 96% of the volume receiving 95% of the prescribed dose, and so on, as long as it is possible to improve PTV coverage. When the algorithm is no longer capable of improving PTV coverage, it considers the best admissible plan found so far, fixes the desired PTV coverage as being equal to what was achieved, and tries to improve OARs in the second phase of the algorithm. In this second phase, the OARs' constraints are more demanding regarding their allowed maximum or mean doses (the desired values will be decreased). The decrease that is considered is given, once again, by a FIS that looks at how far the current dose received by an OAR is from the admissibility tolerance dose. If the dose received by the OAR is very close to its upper admissible limit, then the algorithm will not be very demanding. If the OAR is satisfying its constraint by a large amount, then the algorithm will be more demanding. Whenever the desired value is changed, a slack is considered so that the algorithm has some flexibility in the search for better solutions (nevertheless guaranteeing that the initial desired values are always attained). Slacks are calculated through a FIS system, similar to the one described above. The procedure is repeated until it is not possible to find an admissible solution or a maximum number of iterations is reached.

The algorithm will test different sets of equidistant beam angles, trying to find out the best set of angles. At the present moment, different beam angle sets are considered one at a time, in sequential executions of the algorithm. However, only the first phase of the algorithm is run for every set of angles to be considered. OAR improvement is run for the set of angles that achieved better PTV coverage only. Actually, when the OAR constraints are being satisfied, the preferred treatment plan is usually the one that better irradiates the PTVs. The whole procedure is described next.

1. Choose a set of radiation beam angles. $improve \leftarrow 0$; $IT \leftarrow 0$.
2. Initialize all the model's parameters; $it \leftarrow 0$.
3. Solve the FMO with the current parameters; $it \leftarrow it + 1$; $IT \leftarrow IT + 1$.
4. Do the dosimetry calculations. $Admissible \leftarrow true$.
5. For each structure S
 - (a) If there is a deviation between the current dosimetry metrics and the desired ones, change the upper/lower bounds associated with S according to the FIS. $Admissible \leftarrow false$.

- (b) If the upper/lower bound has reached a predetermined threshold, then change the corresponding weight according to the FIS. Reset all upper and lower bounds to their initial values.
- 6. If *Not Admissible* go to 7, else
 - (a) If *improve* then change the desired dosimetry constraints for selected OARs, trying to spare these OARs more. Go to 3.
 - (b) Else go to 10.
- 7. If $it \leq Nmax$ then go to 3, else go to 8.
- 8. If *improve* then stop. If $IT = MaxIter$ then go to 11, else go to 9.
- 9. Relax some of the defined constraints according to the planner's preferences. $it \leftarrow 0$. Go to 3.
- 10. Improve the current PTV coverage, by being more demanding regarding the dosimetry constraints. Go to 2.
- 11. If every set of angles was already considered, then go to 12. Else, select a different set of angles and go to 2.
- 12. If *improve* then stop, else $improve \leftarrow -1$ and go to 13.
- 13. Consider the beam angle solution that resulted in better PTV coverage. Starting from the best solution found so far, improve the OARs sparing by being more demanding with the defined dosimetry constraints for selected structures. $it \leftarrow 0$. Go to 2.

5 Computational Results

This algorithm was applied to ten head-and-neck cancer cases identified as cases where proper PTV coverage and OAR sparing was difficult to obtain in clinical practice. The OARs considered were the spinal cord, the brainstem and the parotid glands. There are two or more PTVs with different prescribed doses. The desired dosimetry constraints are the ones depicted in Table 1. In clinical practice, most of the times, these cases are treated with 5 up to 11 beam angles. In this paper 9 beam angles plans are considered. It is also common to choose the equidistant solution starting at angle 0° . In these computational experiments, every equidistant beam angle solution was tried, considering a 5° discretization. Tests were performed on an Intel Core i7 CPU 2.8 GHz computer with 4GB RAM and Windows 7. CERR 3.2.2 version [23] and MATLAB 7.4.0 (R2007a) were used. The dose was computed using CERR's pencil beam algorithm (QIB), with corrections for heterogeneities. For each of the ten head-and-neck cases, the sample rate used for Body was 32 and for the remaining structures was 4 (meaning that each set of 32 Body voxels was considered as one voxel in the optimization procedure and, for all other structures, one out of 4 voxels was used in the optimization procedure). To address the convex non-linear formulation of the FMO problem we used a trust-region-reflective algorithm (fmincon) of MATLAB 7.4.0 (R2007a) Optimization Toolbox. FIS were implemented resorting to the Fuzzy Sets Toolbox. The termination criteria for the algorithm are as follows: for each of the beam angles' set, the algorithm will stop after 20 iterations without being able to find an admissible solution ($Nmax=20$), or if 200 iterations in total are reached ($Max-$

Iter=200). The weights are being initialized as follows: if the inner PTV volume is less than 5% of the outer PTV volume, then the weight of the inner PTV will be equal to 50; if the volume is greater than 5% but less than or equal to 10%, then the weight will be equal to 10; in all other cases, it will be equal to 5. The choice of these weights can be seen as arbitrary, and it is indeed. These weights are as good as any others that represent the same idea: the smaller the inner PTV in relation with the outer PTV, the greater the weight we should assign to it. The algorithm was run considering different values for the parameters a , b , c , l and u to assess whether the algorithm is or is not very sensible to these parameters. Simultaneously decreasing a and increasing b will make the update of lower and upper bounds more accentuated. Increasing l and decreasing u will make the weights be updated more often, and increasing c will make the updates steeper. Eight different configurations were tested. Table 2 shows the number of iterations needed to reach a first feasible treatment plan. The minimum number of iterations for each case is shown in squares. It can be seen that it is possible to achieve a feasible treatment for all combinations tested except for one case and one configuration. The algorithm presents the expected behavior: for higher values of a and lower values of b it takes longer to reach feasibility. However, it may present problems converging if it tries to adjust too quickly. The update of weights has little influence in the algorithm's behavior. To analyze computational time and treatment plan quality, we fixed the parameters to $a=5$, $b=10$, $l=30\%$, $u=1.5$ and $c=5$. Table 3 shows the computational time in minutes that the algorithm took to find the first treatment plan complying with all the dosimetry constraints, and also the total computational time (comprising the OAR improvement phase for the best equidistant beam angle set chosen). It is possible to see that the algorithm reaches a first admissible treatment plan in a very short period of time. The total computational time is compatible with clinical practice. The equidistant solution starting from 0° was the best equidistant angle solution for only 3 out of the 10 cases. This reinforces the idea that beam angle optimization is of the utmost importance.

Table 2. Number of iterations until the first feasible solution is found

Cases	$l=15\%, u=2, c=3$										$l=30\%, u=1.5, c=5$									
	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
$a=10, b=5$	16	23	11	14	36	23	12	20	12	30	16	23	11	14	36	23	12	20	12	30
$a=5, b=10$	7	10	4	4	16	10	4	7	7	69	7	10	4	4	15	10	4	7	7	26
$a=3, b=12$	68	6	4	5	14	44	5	5	14	35	23	6	4	5	13	19	5	5	15	39
$a=3, b=15$	18	6	4	5	15	40	5	5	17	---	18	6	4	5	27	37	5	5	26	45

Table 3. Computational time in minutes

Time (minutes)	1	2	3	4	5	6	7	8	9	10
First admissible solution	12	25	8	7	15	21	8	9	4	25
Total computational time	539	717	618	155	95	236	363	405	358	407

For all cases, it was possible to find treatment plans complying with the desired PTV coverage. Fig. 3 and Fig. 4 depict the dose received for at least 95% of the PTV

volume, for each one of the cases. An admissible plan should attain at least 95% of the desired dose, meaning that all points in the charts should be above the horizontal line which is indeed guaranteed. For all patients that have several different PTV70 and/or PTV59 structures, a weighted average considering the number of voxels in each structure was calculated, so that a single number could be shown in the chart, for ease in the presentation of results. It is, however, important to stress that the dosimetry constraints are fulfilled for all individual PTV structures. For OARs, it is important to guarantee that maximum admissible doses are not exceeded. The algorithm produced plans that guarantee all the dosimetry constraints for all OARs (Fig. 5 to Fig. 8). The depicted lines represent the maximum admissible value, so all values should be under these lines. This is exactly what is happening in all cases.

6 Discussion and Conclusions

A methodology based on FIS that is able to attain admissible IMRT treatment plans in an automated way is described. The inputs to the algorithm are the patient's CT-scans with information regarding all the structures of interest and the desired dosimetry constraints. The algorithm will iteratively solve FMO problems, changing in an automated way the model's parameters, and driving the optimization process into interesting regions of the searchable surface. As the process is totally automated, it is possible to repeat the process for different sets of beam angles, increasing the probability of reaching high quality plans. The main advantage of the proposed methodology is the fact that it releases the human planner from a long period of trial-and-error iterations. It has shown to be capable of producing high quality plans, complying with all the desired dosimetry constraints. In clinical practice, it is very difficult to obtain plans in which all structures have their corresponding constraints satisfied through the trial and error procedure. One very interesting feature of this algorithm is that it is capable of dealing with dose-volume restrictions that are usually considered as being very difficult to include in FMO problems, namely the constraints guaranteeing that at most a given volume of the structure receives more than a given dose. Some authors have developed models using dose-volume histogram constraints [24-26], but these constraints have the drawback of creating a non-convex feasibility space, with many local minima. It can also be useful to consider the mean-tail-dose rather than conventional dose-volume constraints [27] (mean dose of either a hottest or coldest specified fractional volume). The algorithm is able to deal quite well with this kind of constraints. It has also been shown that the algorithm's behavior is not very sensitive to changes in the FIS parameters. In this paper a very simple equidistant beam angle optimization is proposed, by enumerating the possible equidistant solutions and finding the best treatment plan for each of them. Further developments will consider embedding FIS into beam angle optimization algorithms [28-33] in order to further contribute to automated treatment planning. The presented methodology lends itself naturally to parallel implementations, which will help tackle the computational burden. We think that using FIS for radiotherapy treatment planning will prove to be an interesting research direction with important applications and expected impacts.

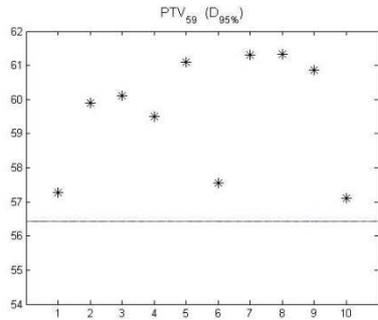


Fig. 3. Target irradiation metrics for PTVs

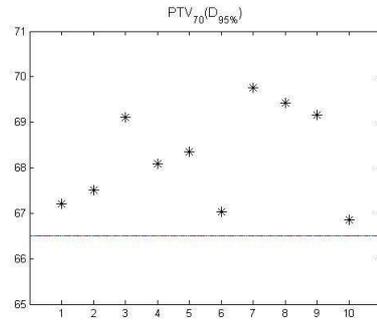


Fig. 4. Target irradiation metrics for PTVs

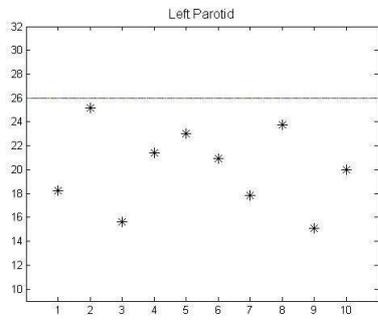


Fig. 5. Left Parotid Mean Dose

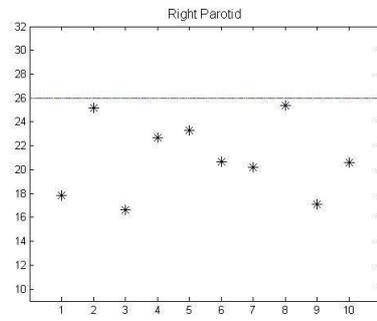


Fig. 6. Right Parotid Mean Dose

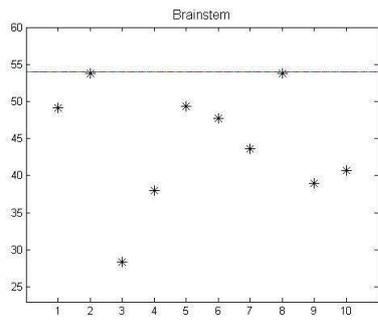


Fig. 7. Brainstem maximum dose

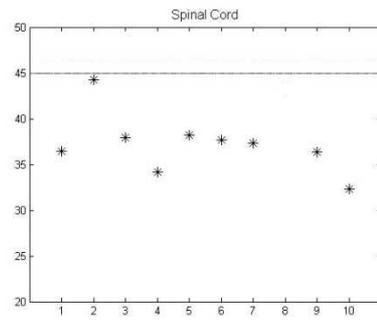


Fig. 8. Spinal Cord Maximum Dose

Acknowledgments

This work has been supported by the Fundação para a Ciência e a Tecnologia (FCT) under project grant UID/MULTI/00308/2013.

References

1. Breedveld, S., Storchi, P.R.M., Voet, P.W.J., Heijmen, B.J.M.: iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. *Medical physics*. 39, 951-963 (2012)
2. Breedveld, S., Storchi, P.R.M., Keijzer, M., Heemink, A.W., Heijmen, B.J.M.: A novel approach to multi-criteria inverse planning for IMRT. *Physics in medicine and biology*. 52, 6339 (2007)
3. Voet, P.W.J., Dirkx, M.L.P., Breedveld, S., Fransen, D., Levendag, P.C., Heijmen, B.J.M.: Toward fully automated multicriterial plan generation: a prospective clinical study. *International Journal of Radiation Oncology* Biology* Physics*. 85, 866-872 (2013)
4. Li, R.-P., Yin, F.-F.: Optimization of inverse treatment planning using a fuzzy weight function. *Medical physics*. 27, 691-700 (2000)
5. Yan, H., Yin, F.F., Guan, H., Kim, J.H.: Fuzzy logic guided inverse treatment planning. *Medical physics*. 30, 2675-2685 (2003)
6. Yan, H., Yin, F.-F., Willett, C.: Evaluation of an artificial intelligence guided inverse planning system: clinical case study. *Radiotherapy and oncology*. 83, 76-85 (2007)
7. Stieler, F., Yan, H., Lohr, F., Wenz, F., Yin, F.-F.: Development of a neuro-fuzzy technique for automated parameter optimization of inverse treatment planning. *Radiation Oncology*. 4, 39 (2009)
8. Dias, J., Rocha, H., Ventura, T., Ferreira, B., Lopes, M.d.C.: Automated fluence map optimization based on fuzzy inference systems. *Medical physics*. 43, 1083-1095 (2016)
9. Lim, G.J., Choi, J., Mohan, R.: Iterative solution methods for beam angle and fluence map optimization in intensity modulated radiation therapy planning. *OR Spektrum*. 30, (2008)
10. Lodwick, W.A., McCourt, S., Newman, F., Humphries, S.: Optimization Methods for Radiation Therapy Plans. In: C. Börgers, F. Natterer (eds) *Computational Radiology and Imaging: Therapy and Diagnostics*. vol. pp. (1999)
11. Shepard, D.M., Ferris, M.C., Olivera, G.H., Mackie, T.R.: Optimizing the Delivery of Radiation Therapy to Cancer Patients. *SIAM Review*. 41, 721-744 (1999)
12. Holder, A.: Designing Radiotherapy Plans with Elastic Constraints and Interior Point Methods. *Health Care Management Science*. 6, 5-16 (2003)
13. Ólafsson, A., Wright, S.J.: *Linear Programming Formulations and Algorithms for Radiotherapy Treatment Planning*. Optimization Methods and Software. (2006)
14. Preciado-Walters, F., Langer, M.P., Rardin, R.L., Thai, V.: Column generation for IMRT cancer therapy optimization with implementable segments. *Annals of Operations Research*. 148, 65-79 (2006)
15. Romeijn, H.E., Ahuja, R.K., Dempsey, J.F., Kumar, A., Li, J.G.: A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning. *Physics in Medicine and Biology*. 48, 3521-3542 (2003)
16. Lee, E.K., Fox, T., Crocker, I.: Integer Programming Applied to Intensity-Modulated Radiation Therapy Treatment Planning. *Annals of Operations Research*. 119, 165-181 (2003)
17. Cheong, K., Suh, T., Romeijn, H., Li, J., Dempsey, J.: Fast Nonlinear Optimization with Simple Bounds for IMRT Planning. *Medical Physics*. 32, 1975-1975 (2005)

18. Aleman, D.M., Kumar, A., Ahuja, R.K., Romeijn, H.E., Dempsey, J.F.: Neighborhood search approaches to beam orientation optimization in intensity modulated radiation therapy treatment planning. *Journal of Global Optimization*. 42, 587-607 (2008)
19. Craft, D.L., Halabi, T.F., Shih, H.A., Bortfeld, T.R.: Approximating convex Pareto surfaces in multiobjective radiotherapy planning. *Medical physics*. 33, 3399-3407 (2006)
20. Men, C., Gu, X., Choi, D., Majumdar, A., Zheng, Z., Mueller, K., Jiang, S.B.: GPU-based ultrafast IMRT plan optimization. *Physics in medicine and biology*. 54, 6565 (2009)
21. Ziegenhein, P., Kamerling, C.P., Bangert, M., Kunkel, J., Oelfke, U.: Performance-optimized clinical IMRT planning on modern CPUs. *Physics in medicine and biology*. 58, 3705-3715 (2013)
22. Mamdani, E.H., Assilian, S.: An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of human-computer studies*. 51, 135-147 (1999)
23. Deasy, J.O., Blanco, A.I., Clark, V.H.: CERR: A computational environment for radiotherapy research. *Medical Physics*. 30, 979-985 (2003)
24. Deasy, J.O.: Multiple local minima in radiotherapy optimization problems with dose-volume constraints. *Medical Physics*. 24, 1157 (1997)
25. Zarepisheh, M., Shakourifar, M., Trigila, G., Ghomi, P.S., Couzens, S., Abebe, A., Noreña, L., Shang, W., Jiang, S.B., Zinchenko, Y.: A moment-based approach for DVH-guided radiotherapy treatment plan optimization. *Physics in medicine and biology*. 58, 1869-1887 (2013)
26. Scherrer, A., Yaneva, F., Grebe, T., Küfer, K.H.: A new mathematical approach for handling DVH criteria in IMRT planning. *Journal of Global Optimization*. 1-22 (2014)
27. Romeijn, H.E., Ahuja, R.K., Dempsey, J.F., Kumar, A.: A new linear programming approach to radiation therapy treatment planning problems. *Operations Research*. 54, 201-216 (2006)
28. Rocha, H., Dias, J., Ferreira, B.C., Lopes, M.C.: Selection of intensity modulated radiation therapy treatment beam directions using radial basis functions within a pattern search methods framework. *Journal of Global Optimization*. 57, 1065-1089 (2013)
29. Rocha, H., Dias, J., Ferreira, B.C., Lopes, M.C.: Beam angle optimization for intensity-modulated radiation therapy using a guided pattern search method. *Physics in Medicine and Biology*. 58, 2939-2953 (2013)
30. Rocha, H., Dias, J., Ferreira, B.C., Lopes, M.C.: Pattern search methods framework for beam angle optimization in radiotherapy design. *Applied Mathematics and Computation*. 219, 10853-10865 (2013)
31. Dias, J., Rocha, H., Ferreira, B.C., Lopes, M.C.: A genetic algorithm with neural network fitness function evaluation for IMRT beam angle optimization. *Central European Journal of Operations Research*. 22, 431-455 (2014)
32. Dias, J., Rocha, H., Ferreira, B.C., Lopes, M.C.: IMRT Beam Angle Optimization Using DDS with a Cross-Validation Approach for Configuration Selection. In: M.e. al (eds) *Computational Science and Its Applications. Lecture Notes in Computer Science*, vol. 8580, pp. 1-16. Springer, (2014)
33. Dias, J., Rocha, H., Ferreira, B., Lopes, M.C.: Simulated annealing applied to IMRT beam angle optimization: A computational study. *Physica Medica: European Journal of Medical Physics*. 30, 747-756 (2014)