Knowledge-light adaptation approaches in case-based reasoning for radiotherapy treatment planning

Sanja Petrovic, Gulmira Khussainova, Rupa Jagannathan

Operations Management and Information System Division, Nottingham University Business School, Nottingham NG8 1BB, UK

Abstract

Objective: Radiotherapy treatment planning aims at delivering a sufficient radiation dose to cancerous tumour cells while sparing healthy organs in the tumour-surrounding area. It is a time-consuming trial-and-error process that requires the expertise of a group of medical experts including oncologists and medical physicists and can take from 2-3 hours to a few days. Our objective is to improve the performance of our previously built case-based reasoning (CBR) system for brain tumour radiotherapy treatment planning. In this system, a treatment plan for a new patient is retrieved from a case base containing patient cases treated in the past and their treatment plans. However, this system does not perform any adaptation, which is needed to account for any difference between the new and retrieved cases. Generally, the adaptation phase is considered to be intrinsically knowledge-intensive and domain-dependent. Therefore, an adaptation often requires a large amount of domain-specific knowledge, which can be difficult to acquire and often is not readily available. In this study, we investigate approaches to adaptation that do not require much domain knowledge, referred to as knowledge-light adaptation.

Methodology: We developed two adaptation approaches: adaptation based on machine-learning tools and adaptation-guided retrieval. They were used to adapt the beam number and beam angles suggested in the retrieved case. Two machine-learning tools, neural networks and naive Bayes classifier, were used in the adaptation to learn how the difference in attribute values between the retrieved and new cases affects the output of these two cases. The adaptation-guided retrieval takes into consideration not only the similarity between the new and retrieved cases, but also how to adapt the retrieved case.

Results: The research was carried out in collaboration with medical physicists at the Nottingham University Hospitals NHS Trust, City Hospital Campus, UK. All experiments were performed using real-world brain cancer patient cases treated with three-dimensional (3D)-conformal radiotherapy. Neural networks-based adaptation improved the success rate of the CBR system with no adaptation by 12%. However, naive Bayes classifier did not improve the current retrieval results as it did not consider the interplay among attributes. The adaptation-guided retrieval of the case for beam number improved the success rate of the CBR system by 29%. However, it did not demonstrate good performance for the beam angle adaptation. Its success rate was 29% versus 39% when no adaptation was performed.

Conclusions: The obtained empirical results demonstrate that the proposed adaptation methods improve the performance of the existing CBR system in recommending the number of beams to use. However, we also conclude that to be effective, the proposed adaptation of beam angles requires a large number of relevant cases in the case base.

Keywords: Case-based reasoning, adaptation-guided retrieval, machine-learning tools, radiotherapy treatment planning

Email address: sanja.petrovic@nottingham.ac.uk, psxgk@nottingham.ac.uk, rupa.jagannathan@nottingham.ac.uk

1. Introduction

Radiotherapy is a type of cancer treatment that uses ionising radiation beams to control cancer growth. Radiotherapy treatment planning (RTP) aims to deliver a sufficient dose of radiation to kill tumour cells while minimising damage to healthy tissues and critical organs in the vicinity of the tumour. This type of treatment is based on the different abilities of healthy and cancerous cells to recover from radiation damage. Ideally, the radiation dose to the healthy tissue should be as low as possible, and the shape of the dose should be conformal to the tumour and avoid healthy surrounding tissues to reduce radiation side effects.

The process from diagnosis to treatment contains a number of steps. Medical images are taken using computed tomography (CT) and magnetic resonance imaging (MRI) to detect and find the location of the tumour. These images consist of three-dimensional (3D) image data of the affected region along with textual information about the patient. Oncologists then determine the required radiation dose to treat the patient and pass that information to the medical physicists, who design a detailed treatment plan that includes radiation beam number, beam angles, dose fraction in each treatment session, beam wedges, and so on. A detailed treatment plan has to be created to determine exactly how the patient will be irradiated to achieve the required dose distribution [1]. In this study, we consider 3D-conformal radiotherapy. RTP is usually a time-consuming trial-and-error process and requires the expertise of one or more experienced medical physicists and oncologists. Planning can take up to 3-4 days depending on the complexity of the plan. Developing a treatment plan depends highly on the expertise and experience of the medical staff. When determining suitable values of planning parameters such as the number of beams, beam angles, number of wedges and their configuration and dose in different phases of treatment, various methodologies can be used. The literature on RTP offers a variety of methods that can be classified into optimisation and knowledge-based methods [2–5].

Case-based reasoning (CBR) has been widely used for developing knowledge-based decision support systems in the medical domain. CBR is based on using experience gained when solving similar problems in the past to understand and solve new problems. Kolodner [6] introduced the term CBR and demonstrated its application in the everyday activities of human life. The CBR process comprises of the steps of reasoning, which take into consideration all previous similar events and features that could help in solving a new problem.

When creating a treatment plan for a new patient, medical staff try to remember/retrieve similar cases that they have encountered in the past along with their treatment plans and try to map them to the new case. Therefore, using CBR seems to be a natural decision in designing an application that will help medical staff make a treatment plan. The main purpose of an automated software application system is to save time when developing a treatment plan.

The research presented in this paper builds on the CBR system for the RTP of brain cancer, which was developed in collaboration with the Nottingham University Hospitals Trust, NHS, Nottingham City Hospital Campus and is described by Jagannathan et al. [7]. The experiments showed promising results but suggested that an adaptation of the retrieved solution needs to be performed to better address the requirements of the new case. Here, we investigate adaptation mechanisms to improve the results of the current retrieval system. Our focus will be on the adaptation that is not knowledge-intensive but instead makes use of domain knowledge available in the case base. Two approaches will be implemented and compared: a machine learning-based adaptation that employs neural networks and naive Bayes classifier and an adaptation-guided retrieval. These methods were chosen to acquire knowledge from the case base without interaction with medical physicists.

The developed adaptation methods were evaluated using the case base containing real-world brain cancer patient cases and their treatment records from the Nottingham City Hospital. Although only the knowledge contained in the case base was used for adaptation, the obtained results are promising and will initiate further research.

This paper is organised as follows. An overview of the relevant literature and the background of this work is provided in Section 2. Section 3 provides an overview of the existing CBR system. Section 4

(Sanja Petrovic, Gulmira Khussainova, Rupa Jagannathan)

introduces machine-learning tools used for adaptation followed by Section 5, which discusses the implemented adaptation-guided retrieval method. The experimental results on real-world brain cancer patients records are presented in Section 6. Section 7 provides the conclusion and future research directions.

2. Background and related work

This section provides a review of the approaches to RTP found in the literature followed by a description of CBR systems for cancer treatment. The section ends with a discussion of adaptation approaches in CBR.

2.1. Approaches to radiotherapy planning

Radiotherapy planning can be approached from two angles: forward and inverse planning. In inverse planning, the required dose distribution is specified, and a treatment plan is then created so that it fits the prescribed dose distribution. In forward treatment planning, the planning parameters are adjusted iteratively to achieve an acceptable dose distribution of irradiation. The literature on radiotherapy planning covers both angles.

The methodologies that were developed to support RTP can be divided into two types: optimisation methods and knowledge-based methods. Optimisation methods developed for RTP include linear and non-linear programming, quadratic programming, and heuristic optimisation methods such as simulated annealing, genetic algorithms and ant colony optimisation.

A number of papers provide a survey of optimisation methods for particular types of radiotherapy treatment. A survey on the use of mathematical optimisation in intensity modulated radiotherapy treatment was given by Ehrgott et al. [1]. Holder and Salter [8] examined the shortcomings of linear, non-linear and quadratic models in optimisation for radiation oncology. Oldham et al. [9] reviewed applications of optimisation methods for two types of brain cancer RTP. Their results demonstrated that optimisation algorithms improved the treatment planning when beam orientations, beam weights and wedge angles were optimised. However, these algorithms had no effect on the treatment planning when only beam weights and wedge angles were considered for optimisation. A binary probabilistic Tabu search was developed for inverse treatment planning in intensity-modulated radiation therapy, in which each radiation beam consists of many beamlets, each of different intensity formed by a multileaf collimator [10]. Ehrgott and Johnston [11] used heuristics in the optimisation of beam directions in intensity-modulated RTP. De Neve et al. [12] used optimisation heuristics methods in radiation dose planning in head and neck cancer. A review of the models of **?**], **?**] and **?**], which address different aspects of radiotherapy treatment design, showed that they generally do not sufficiently address the clinical issues that arise in treatment planning as they fell short when compared to the design process conducted by a physician.

In general, these treatment planning systems use complex mathematical models that neglect to capture the expertise and experience gained from physicists. Experienced planners often know what configurations are suitable for a cancer patient, even when the underlying theory is not thoroughly understood [7].

Many knowledge-based methods for radiotherapy planning have been developed to decide upon parameters for treatment plans, including the radiation dose plan, beam configuration, beam weight and wedge configuration [2–5, 7]. Such methods aim at utilising knowledge available in the domain and used by the domain end-user to simulate the natural solving process as much as possible. These methods do not use mathematical algorithms, but instead make use of domain knowledge or past experience in determining the parameters of the treatment plan. Artificial intelligence methods such as rule-based reasoning, CBR and hierarchical organisation of knowledge have been employed to automate the radiotherapy planning process [13, 14].

There are also systems in which optimisation methods and knowledge-based methods are combined to enhance the performance of each other; Teodorovic et al. [15] developed such a system in which CBR was enhanced with bee colony optimisation to determine the thyroid cancer radioiodine dose. Bee colony optimisation was used to determine attribute weights in the similarity measure of the CBR system.

2.2. Case-based reasoning in cancer treatment

CBR solves new problems by retrieving similar previously solved problems and their solutions from a database of cases referred to as a case base. The solutions can be applied directly to the new problem for certain tasks such as classification. However, tasks that require design or planning usually need revision or adaptation of the retrieved solution in order to meet the exact needs of the new problem. A CBR system usually comprises retrieval, reuse, revise and retain phases. In the retrieval phase, a problem similar to the new one is retrieved from the case base. The solution of the retrieved case is then reused for the new case and possibly revised to meet all of the requirements of the new case. If the case is assessed to be useful for future problem solving, it is stored with its solution in the case base for future retrieval during the retain phase.

There are many advantages of using CBR in the medical domain, such as:

(a) CBR can be used to find solutions to problems when the complete understanding of the domain is not captured or is very complex.

(b) It allows the reuse of past solutions, which helps to avoid solving problems from scratch and thus saves time.

(c) It allows learning from mistakes of retrieving incorrect cases [16].

(d) Cases can be used for a number of different problem-solving purposes as they contain more information, while rules can be generally reused only for the purpose for which they were defined [17].

CBR has been used in medical applications in diverse areas such as clinical diagnosis and treatment in psychiatry and epidemiology. Holt et al. [18] discussed the medical applications of CBR in the last two decades and grouped them into the following categories: systems for classification, diagnosis, planning and tutoring. CBR has been successfully used in many applications. However, the authors pointed out that adaptation in CBR was a problem in many domains, including the medical domain, because knowledge is expensive to gain; thus, only a few applications have been used in clinical settings. CBR systems for cancer treatment have been reported in the literature, although to a limited extent. Some examples of CBR systems developed for radiotherapy planning are listed in Table 1.

Liber and Bresson [5] made use of both rule-based and CBR in their system for breast cancer treatment planning. The CBR part employed the experience and knowledge of a therapy committee in adjusting a treatment plan suggested by the rule-based reasoning part.

System implemented	Cancer site	References
Casimir/CBR	Breast cancer	Liber and Bresson [5]
ROENTGEN	Lung cancer	Berger [14]
CBR in robotic radiosurgery	Prostate cancer	Schlaefer and Dieterich [4]
Dempster-Shafer rule-based CBR system	Prostate cancer	Song et al. [13], Petrovic et al. [19]
Trade-off planning	Prostate cancer	Mishra et al. [2]
2-phase retrieval CBR system	Brain cancer	Jagannathan et al. [7], Jagannathan and Petrovic [20]

Table 1. Examples of CBR systems used in the healthcare literature on cancer treatment planning

The first CBR system for RTP for thorax cancer, Roentgen, was introduced by Berger [14]. Roentgen retrieved cases for RTP based on the location and geometrical measures of relevant organs and the tumour along with the treatment constraints. The dose distribution in treatment planning was determined according to the patient's internal geometry; hence, a good match had to be similar with respect to the geometry of the new patient. In addition, the dose delivered to the tumour and healthy tissues of the patient in the retrieved plan had to meet the constraints of the new patient. However, there are no published results of this study.

CBR was used in robotic radiosurgery for beam generation planning by Schlaefer and Dieterich [4]. The results indicated a considerable reduction of almost 50% in the runtime of the algorithm for beam generation compared to the conventional one.

Song et al. [13] and Petrovic et al. [19] used CBR to determine the dose plan for prostate cancer patients in RTP. The dose plan was determined based on the trade-off between the expected benefit of cancer control by delivering a high radiation dose and the risk of side effects caused by the radiation of the surrounding organs at risk (OARs). A CBR approach using fuzzy set theory in the similarity measure was proposed in [13]. Dempster-Shafer theory was applied in the retrieval phase to fuse plans from multiple retrieved similar cases to strengthen agreements and decrease conflicts in situations where oncologists might have different opinions of the same case [19]. Mishra et al. [2] developed a CBR system for dose planning in prostate cancer to determine the extent to which dose limits set for an OAR can be exceeded to better fight cancer cells.

A CBR system for brain cancer RTP was developed by Jagannathan et al. [7]. They focused mainly on the retrieval process of the CBR system. Different similarity measures were proposed, including a fuzzy similarity measure that used fuzzy sets to represent the contribution of each individual attribute to the similarity between two cases along with a non-linear similarity measure, in which the similarity between two values of an attribute of two cases contributed in a non-linear way to the aggregate similarity between two cases. The similarity measure that used local weights assigned to case attributes that depended on the values of other case attributes was described in [20]. The research presented in this paper builds upon this work and investigates the extent to which the adaptation can improve the plan proposed by the retrieval phase of the CBR system.

2.3. Adaptation in case-based reasoning

In many real-world applications, an adaptation phase is required after retrieval; in this phase, the solution of the retrieved case is modified to meet the requirements of the new case.

Fuchs et al. [21] reviewed studies of adaptation in CBR and concluded that a general model of adaptation has not yet been developed due to the large amount of domain knowledge usually required by adaptation; this knowledge is not always readily available or easy to acquire. In addition, as the knowledge is highly domain-dependent, the authors argued that it was difficult to design a generic model. In practical applications, the adaptation step is commonly left out, and the retrieved case solution is applied to the new problem without any modification; alternatively, adaptation is left to the end user.

Wilke and Bergmann [22] presented three types of adaptation: transformational, generative and compositional adaptation. In *transformational adaptation*, the retrieved case solution is transformed into a new solution for the new problem. This type of adaptation is done by modifying, adding or deleting some elements of the solution. *Generative adaptation* implies the insertion of a problem-solver into the system. Such a problem-solver requires knowledge that helps to construct a solution from scratch rather than an explanation of differences between problem descriptions that lead to differences in the solution. *Compositional adaptation* is the combination of both transformational and generative adaptation.

Methodologies that have been used in the adaptation phase can be divided into four groups: rule-based adaptation, case-based adaptation, adaptation-guided retrieval and machine learning-based adaptation. While the former two are knowledge-intensive approaches, adaptation-guided retrieval and machine learning-based adaptation can be classed as knowledge-light approaches.

Adaptation rules can be derived after reviewing and revising the retrieved cases to specify how to modify the value of a feature under a certain situation, or how to insert or delete certain features in the case representation in order to adapt a solution to the new problem. Such an approach to rule-based adaptation was employed in the CBR system for RTP for prostate cancer by Mishra et al. [2], in which experts' personal opinions were modelled by rules.

The majority of adaptation approaches currently reported in the literature are based on acquiring adaptation knowledge from domain experts and building an adaptation case base. This approach requires interaction with domain experts who adapt the case suggested by the CBR system. The adapted solution along with the adaptation procedure are then stored in the adaptation case base for the future reference. The adapted case will be used for the next adaptation when the difference between the new problem and the retrieved case is similar to the difference between the case that generated the proposed adaptation case and its corresponding retrieved case. Hennessy and Hinkle [23] developed a CBR system for autoclave loading that

employed case-based adaptation. The solution adapted by the expert was added to the case base. Hence, the case base expanded gradually with the use of the system. Vong et al. [24] claimed that integrating experts into the adaptation process was the most natural way of acquiring domain knowledge and developed a CBR system in the design of a hydraulic production machine in which adaptation was performed with the help of domain experts. D'Aquin et al. [25] went a step further and introduced an adaptation knowledge acquisition process for learning adaptation rules from experts. However, the process also required communication with experts.

Smyth and Keane [26] suggested that it was unwarranted to assume that the most similar case was also the most appropriate for adaptation. They proposed adaptation-guided retrieval, in which the retrieval mechanism retrieved a case that was easy to adapt. This is further stressed in the work of D'Aquin et al. [25] and Fuchs et al. [21]. In adaptation-guided retrieval, the retrieved case should be as similar as possible to the new case. However, the retrieved case was considered to be similar if it was possible to adapt its solution for the new case [21].

It has been claimed that to perform adaptation, a large amount of domain-specific knowledge is required (i.e., it is knowledge-intensive). However, the knowledge needed for adaptation is not always available and can be difficult to obtain. Contrary to knowledge-intensive approaches, knowledge-light approaches do not rely on the use of domain-specific knowledge. In knowledge-light approaches, the knowledge available in the case base is used as-is with no or very little additional knowledge. Knowledge-light approaches can be useful when sources of additional knowledge for adaptation are not available.

The knowledge-light approach to adaptation introduced by McDonnell and Cunningham [27] is a type of adaptation-guided retrieval. This approach was developed for non-linear systems where the influence of each attribute on the output value is not linear throughout the range of values for that attribute. Adaptation was performed using the knowledge contained in the case base itself. The retrieval phase searched for both a case similar to the new case and cases appropriate for adaptation, and they were all used in solving the problem given in the new case. Their knowledge-light adaptation approach was used to predict the values of numeric variables, particularly in the housing domain. The price of a house was adapted based on the difference in values of attributes such as the number of bedrooms and locations of neighbouring houses. The knowledge-light adaptation approach used differences in attribute values between cases as a tool for adaptation purposes. In RTP, a knowledge-light approach to adaptation was applied in the CBR system for dose planning in prostate cancer [28]. In particular, the calculation of how much to violate the constraint on the radiation dose limit imposed to OARs was performed by using the retrieved case and a pair of cases with a similar difference in the input attributes as between the new and the retrieved case.

Machine-learning tools have been used in the retrieval phase of CBR; however, to our knowledge, there are only two approaches that employ neural networks (NNs) in adaptation. In the medical domain, Henriet et al. [29] developed an adaptation tool in a CBR system for the numerical representation of human organs to make a phantom in 3D using an NN model. The retrieved solution was suggested to the experts, who decided whether to use it or whether a new solution was required. To adapt a case, experts passed input descriptors to the CBR system, and the NN was used to adapt the retrieved solution using these descriptors. The research suggested that when the adaptation was guided by experts, the obtained results were well-suited for the new cases. Corchardo et al. [30] retrieved a few similar cases during the CBR retrieval phase and then used NNs for weighting these cases in the adaptation phase to determine the best solution in oceanographic water temperature forecasting. They reported that the CBR system performed better with the NN-based adaptation compared to the CBR system with no adaptation. The advantage of using machine-learning tools is their ease of application.

3. Overview of the CBR system for radiotherapy treatment planning

In this section, a brief description of the CBR system for the RTP of brain cancer developed by Jagannathan et al. [7] is given. The case base consists of brain cancer patient cases containing a description of the patient and the plan used for the treatment. At the current stage of development, the solution suggested by the CBR system includes two treatment plan parameters: the number of beams applied to the planning treatment volume (tumour) and the angles of those beams. Beam angles are defined by the gantry angle

and the patient couch angle. Figure 1 illustrates the beam angle; the red lines crossing the head represent the beams targeting the tumour at different angles.
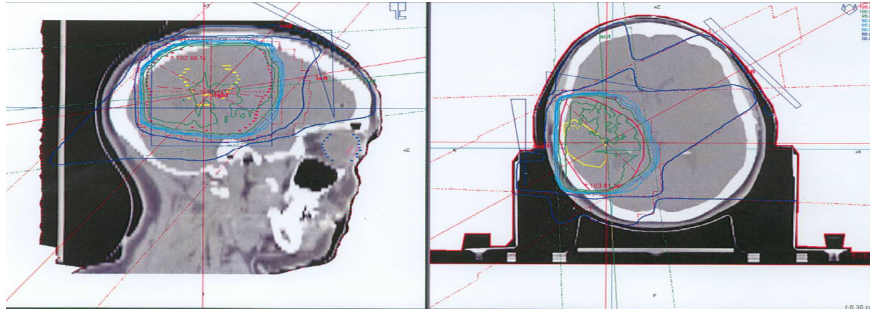

Figure 1. Beam angles (red lines) projected on PTV [7]

Patient data were received in the Digital Imaging and Communications in Medicine (DICOM) file format introduced by the National Electrical Manufacturers Association [31]. DICOM is a standard for handling images in clinical practice. DICOM images contain image data such as CT and MRI scan files along with textual information about a patient.

In the Nottingham City Hospital, the oncologist examines the image files of patients and outlines the planning treatment volume (PTV) and OARs in the treatment planning system. The treatment plan is then generated by medical physicists who place beams on the patient image and determine the acceptable dose distribution in a trial-and-error fashion. In the proposed CBR system, the attributes that describe a case have been chosen in consultation with medical physicists. The case attributes must be carefully selected so that the similarity calculation between two cases accurately represents the relative applicability of the solution of a case to the new case. Since treatment planning aims to focus the prescribed radiation on the PTV while avoiding the OAR, the treatment plan parameters are largely determined based on the location and dimensions of the PTV and the spatial relationship between the PTV and the OAR. These attributes determine the geometry of a patient. We have identified the following case attributes that describe the patient in terms of attributes relevant to the treatment plan; these attributes are graphically presented in Figure 2.

Vol - volume of PTV. Value in the range [24.87-729.82 mm^3]

A - angle between the line connecting the origin of the DICOM patient-based coordinate system in the image (O) with the centre of the PTV and with the centre of the OAR. This attribute gives information about the location of the tumour in relation to the OAR. Value in the range [0 °-360 °]

E - minimum distance between the edge of the PTV outline and OAR outline. This attribute is important since a large distance allows for more flexibility when placing the beams, whereas a smaller distance between PTV and OAR might require a more complicated plan and possibly a larger number of beams. Value in the range [2.04-674.87 mm].

R - the ratio of the PTV volume to the volume of the patient's body. This attribute is relevant because medical physicists have noted that similar treatment plans frequently have similar PTV and body ratios, even if the actual PTV or body volumes differ. Value in the range [6.40-255.92 ratio]

Dt - minimum distance between the body edge and PTV outline edge. This attribute provides positional information about the tumour and describes how close the tumour is to the surface of the body. Value in the range [0-53.66 mm].

P - the relative position of the PTV with respect to the OAR. This attribute is relevant because medical physicists use different beam angles based on the location of the tumour in the head. This attribute contains six labels: left, right, posterior, anterior, superior and inferior.

The following OARs are considered: the right and left lenses of the eyes, right and left optical nerves, spinal cord, optic chiasm and brainstem.

The retrieval process starts with the filtering of the case base, leaving only cases with the same OARs as the new case in the case base. A weighted nearest neighbour (wNN) similarity measure was defined with local weights identified for each of the attributes depending on the values of other attributes [20].
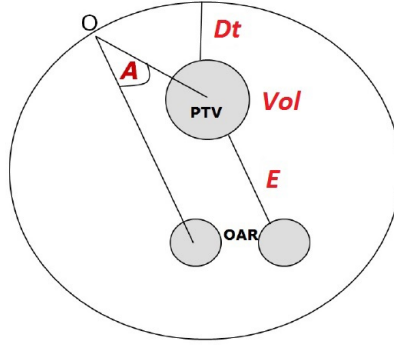

Figure 2. Attributes of a case [7]

.

The retrieval consists of two phases. The first phase of the retrieval determines the number of beams to be used in the treatment plan. The proposed number of beams is stored and used in the second phase, in which only cases in which treatment plans have the proposed number of beams are available for retrieval. Among these cases, the case most similar to the new case is retrieved and suggests the beam angles. The reason for finding the correct beam numbers first is that medical physicists consider this parameter to be one of the most important elements in determining a treatment plan. During treatment planning, it is easier to adjust beam angles than the number of beams; therefore, it is more important that the CBR system is able to correctly determine the beam number. Furthermore, if the retrieved case has an unsuitable beam number, the missing beam angles or extra beams need to be considered in the adaptation. Therefore, in this study, we first perform the beam number adaptation followed by beam angle adaptation.

The similarity between two cases is calculated using the wNN similarity measure, where the weights of the case attributes represent their relative contributions to the similarity. The weights of the attributes are different in the retrieval of a case to suggest the beam number (suggested in phase 1 of the CBR system) from the retrieval for beam angles (suggested in phase 2). An algorithm was developed that learns from the data-appropriate rules to assign a local weight to each attribute based on the values of the new case attributes [20].

The similarity measure used in the retrieval is as follows. Let C_n and C_{cb} denote a new case and a case from the case base, respectively. The distance between the cases is calculated using the Euclidean distance:

$$d(c_n, c_{cb}) = \sqrt{\sum_{l=1,\ldots,6} w_l (V_{n,l} - V_{cb,l})^2} \tag{1}$$

where V_n and $V_{cb,l}$ denote the attribute values of attribute l, where $l = 1,\ldots,6$ for the attributes A, E, Vol, R, Dt and P of the new case and a case from the case base, respectively, and w_l is the weight of attribute l. All the attributes are normalised to take values in the range of 0 to 1.

The similarity between C_n and C_{cb} is calculated using:

$$s(c_n, c_{cb}) = 1 - d(c_n, c_{cb}) = 1 - \sqrt{\sum_{l=1,\ldots,6} w_l (V_{n,l} - V_{cb,l})^2} \tag{2}$$

The performance of a CBR system is commonly evaluated using test cases obtained directly from the

case base [32]. The treatment plans of the cases contained in the case base consist of patient cases that have been successfully treated at the Nottingham City Hospital. Their treatment plans were generated by senior medical physicists and approved by a consultant oncologist. The plans are assumed to have the best configuration of treatment plan parameters for that patient. The CBR system is therefore evaluated based on the differences between the treatment plan suggested by the CBR system and the known treatment plan of that patient using leave-one-out cross validation, in which each case in the case base is consecutively made the new case, and the remaining cases constitute the case base. The error in the beam number is defined as the difference between the beam number of the retrieved case and the actual beam number used in the treatment plan. The retrieval is deemed to be successful if $E_{BN} = 0$, where E_{BN} is the beam number error. The error in the beam angle is defined as the average difference between the beam angles in the treatment plan suggested by the CBR system and the beam angles in the treatment plan generated by medical physicists. The retrieval is successful if $E_{BA} < 30°$, where E_{BA} is the beam angle error. This threshold of $30°$ was chosen based on discussion with medical physicists. The case base consisting of 80 real-world brain cancer cases was used in leave-one-out cross validation. The success rate achieved was 61% with respect to beam number error; that is, in 49 out of 80 cases, the system retrieved a case with the same beam number suggested by the medical physicists. The success rate for the beam angles retrieval was 39%; i.e., in 29 cases out of 80, the retrieved E_{BA} was less than $30°$. In this study, we use the same validation technique to measure the success rate of the proposed adaptation.

4. Adaptation

In this section, three approaches to adaptation are described: two based on machine learning and one adaptation-guided retrieval.

4.1. Adaptation based on machine learning

We considered two well-known machine learning tools, NN and naive Bayes classifier, for adaptation. NNs are computational models that have been successfully used in machine learning in a variety of problem domains including classification. The cases in the case base have two, three or four beam numbers. Therefore, an adaptation of beam number suggested by the retrieved case can be to add or subtract one or two beams or not to change it. Such an adaptation can be formulated as a classification problem. This motivated us to investigate NNs as an adaptation tool in CBR.

NNs consist of interconnected input neurons, hidden neurons, which are organised in layers, and output neurons that serve as decision variables. Connections between neurons are assigned adaptive weights, which are activated and modified during the training phase of NN development. The learned weights are then used in the classification.

The aim of the developed NN is to learn how the differences in attribute values of the retrieved and new cases affect the outputs of these two cases. In the training phase, we use the beam number data recorded in the cases. Given a pair of training cases, the NN is trained to discover the pattern of differences between the attributes of two cases and the difference in their beam numbers. Once the NN is trained, it is used to suggest the adaptation of the beam number of the retrieved case based on the differences between the attributes of the new and retrieved cases. The input neurons contain the attribute differences between the new and the retrieved cases, while the output neuron is the class that denotes the required change to the beam number of the retrieved case. In this study, we only consider cases with three or four beams because only one case in the case base has two beams. Experiments have shown that having only one case with two beams is not sufficient for training, resulting in poor results. Consequently, our NN model contains three classes that correspond to three possible changes: 0 if no change is needed and 1 or -1 if the suggested beam number has to be increased or decreased by 1, respectively. However, the proposed NN model is general and can handle any number of beams in the cases on the condition that there are a reasonable number of cases with different numbers of beams in the case base. In such settings, if we identify beam number change as (`BN_Change`), the output neuron will have the value `-BN_Change` or `+BN_Change` to denote the required change to the beam number of the retrieved case.

Naive Bayes classifier [33] is often used when attributes are assumed to be independent of each other (i.e., the value of a particular attribute determines a class independently of the presence or absence of any other attribute). The naive Bayes method is different from NN; all attribute differences are considered together as one input, and their overall effect on the outcome is observed. Furthermore, naive Bayes classifiers are considered to be particularly suitable when the range of the input attribute values is highly varied, which is characteristic of radiotherapy cases. Despite the simplicity of naive Bayes classifier, there are examples of its successful use in practice and also in CBR systems. For example, Yeow et al. [34] used it in a CBR system to learn weights of the case attributes in the retrieval phase.

Experiments have shown that in RTP, some attributes play a more considerable role than others in determining the output. Therefore, we investigate if naive Bayes classifier can detect such attributes (i.e., whether considering each attribute difference value separately allows the classifier to correctly suggest the change in beam number).

Naive Bayes classifier uses the input training set and classifies each case in the classes −1, 0 and +1, which have the same meaning as in NN. It then calculates the prior probabilities of class membership. The prior probabilities are recalculated with respect to the actual output of the training set. The input values of a new testing set of cases are classified based on the probability values.

4.2. Adaptation-guided retrieval of beam number

Many real-world problems do not exhibit global linearity in the sense that differences in attribute values have different impacts on the output in different regions of the domain space. RTP belongs to such a class of non-linear domains. The knowledge-light approach to adaptation introduced by McDonnell and Cunningham [27], which was developed for non-linear domains, has been modified to fit the RTP problem.

The standard CBR practice would be to retrieve a case from the case base that is the most similar to the new case and afterwards adapt the beam number in the retrieved case to address the differences between the retrieved and a new case. However, in our retrieval, we consider adaptation as well. The aim of the developed adaptation-guided retrieval of beam number (AGRBN) is to retrieve cases from the case base that suggest a correct change in the beam number of the retrieved case for treating a new patient. This approach assumes some degree of local linearity (i.e., in a small region within which local linearity is assumed), and the similar differences in attribute values result in similar differences in the output value. Apart from searching the case base for a case most similar to the new case, the retrieval algorithm searches for a pair of cases whose attribute differences have similar impacts on the beam number as the attribute differences between the new and retrieved case. The difference in the beam number of the treatment plans for that pair of cases is used to adapt the beam number suggested for the new case by the retrieved case.

Let C_n denote the new case and C_r denote the case retrieved by the CBR system. In order to retrieve an appropriate case C_r from the case base, which can be adapted to meet the requirements of the new case, the retrieval algorithm searches for three cases (C_r, C_1, C_2) such that the impact of the attribute differences between C_1 and C_2 on the beam number (output) of C_1 and C_2 is similar to the impact of the attribute differences between C_n and C_r on their beam numbers. This approach is different to the one presented in [28], where the most similar case is retrieved, and the retrieval then searches for a pair of cases whose difference in attribute values is the most similar to the difference between the new and retrieved case.

Two vectors are defined to be used in the knowledge-light adaptation: the case differences vector, which represents the differences between the corresponding attributes of the new case and a retrieved case, and the case gradient vector, which represents the rate of change of the output with respect to the attribute differences. Their mathematical formulation is given below.

Each element of the case differences vector, $\triangle(C_n, C_r)$, represents the difference between the values of attribute l, l=1,...,6 for attributes A, E, Vol, R, Dt and P, respectively:

$$\triangle_l(C_n, C_r) = V_{n_l} - V_{r_l}, \quad l = 1, 2, 3, 4, 5, 6 \tag{3}$$

where V_{n_l} and V_{r_l} are values of attribute l in cases C_n and C_r, respectively.

The case differences vector is defined as the vector of all attribute differences:

$$\triangle(C_n, C_r) = [\triangle_1(C_n, C_r), \triangle_2(C_n, C_r), \triangle_3(C_n, C_r), \triangle_4(C_n, C_r), \triangle_5(C_n, C_r), \triangle_6(C_n, C_r)] \tag{4}$$

Let Y be the output of function f, which depends on the attributes a_1, a_2,...,a_m. The gradient of function $Y = f(a_1, a_2, ..., a_m)$ is defined as the rate of change of output Y with respect to input attributes a_1, a_2, ...,a_m, as shown in Expression (5):

$$\nabla f = (\frac{\partial Y}{\partial a_1}, \frac{\partial Y}{\partial a_2}, ... \frac{\partial Y}{\partial a_m}) \tag{5}$$

In our adaptation procedure, function f describes the change in beam number (BN), and its analytical form is not known. If we denote C_{sim} as the case that is the most similar to case C with beam number BN_{Csim}, retrieved using Expression (2), then the gradient vector of case C with beam number BN_C is calculated as follows:

$$\nabla C = [\frac{\partial BN}{\partial V_1}, ..., \frac{\partial BN}{\partial V_6}] = [\frac{BN_C - BN_{C_{sim}}}{V_{C_1} - V_{C_{sim1}}}, ..., \frac{BN_C - BN_{C_{sim}}}{V_{C_6} - V_{C_{sim6}}}] \tag{6}$$

This is a general formula for the gradient of case C. This will be used to calculate the gradient of cases C_r and C_2.

The purpose of using the gradient is to investigate how large the change in the attributes has to be in a given region of the case base to cause the beam number to change. In our domain, the majority of cases have three or four beams. Consequently, the most similar case retrieved can have the same beam number as the case for which the gradient is calculated, resulting in a value of 0 for the gradient. Therefore, the case retrieved for the gradient calculation should have a different beam number. If a small difference in attribute values triggers the change in the beam number, then the gradient is large; otherwise, it is small. The retrieval algorithm searches the cases sorted by their similarity with a case for which gradient is calculated until the most similar case with a different beam number is found.

The retrieval can be stated as an optimisation problem. The task is to find a set of three cases to be used in the adaptation, (C_r, C_1, C_2), such that the sum of the difference between C_r and C_n, the difference between the product of attribute differences between C_1 and C_2, the gradient of C_2, the product of attribute differences between C_n and C_r, and the gradient of C_r is minimised:

$$R_{C_r, C_1, C_2} = d(C_n, C_r) + d(R_{C_n, C_r}, R_{C_1, C_2}) \tag{7}$$

where

$$R_{C_n, C_r} = \triangle(C_n, C_r) * \nabla C_r \tag{8}$$

and

$$R_{C_1, C_2} = \triangle(C_1, C_2) * \nabla C_2. \tag{9}$$

The pseudo-code for the retrieval of the triplet of cases (C_r, C_1, C_2) is given in Algorithm 1.

Algorithm 1 Algorithm for retrieval of the set of three cases for the beam number adaptation

---

1: BEGIN
2: $M_{NxN} = [\triangle(C_j, C_k)]$ holds the case differences vectors of cases C_j, and C_k in the case base
3: $m_{j,k} = \triangle(C_j, C_k)$ calculate using Expression (4)
4: $R_{min} = 0$
5: **for** $r = 1, ..., N$ **do**
6:     $d_1 = d(C_n, C_r)$ calculate using Expression (1)
7:     $R_{11} = \triangle(C_n, C_r)$ calculate using Expression (4)
8:     $R_{12} = \nabla C_r$ calculate using Expression (6)
9:     $R_1 = R_{11} * R_{12}$ calculate using Expression (8)
10:     **for** $r_1 = 1, ..., N \quad r_1 \neq r$ **do**
11:       **for** $r_2 = 1, ..., N \quad r_2 \neq r_1 \neq r$ **do**
12:         $R_{21} = \triangle(C_{r_1}, C_{r_2})$ (take from matrix M)
13:         $R_{22} = \nabla C_{r_2}$ calculate using Expression (6)
14:         $R_2 = R_{21} * R_{22}$ calculate using Expression (9)
15:         $d_2 = d(R_1, R_2)$ calculate using Expression (1)
16:         $R = d_1 + d_2$ calculate using Expression (7)
17:         **if** $R <= R_{min}$ **then**
18:             $R_{min} = R$
19:         **end if**
20:       **end for**
21:     **end for**
22: **end for**
23: Output triplet (C_r, C_{r_1}, C_{r_2}) for which $R_{C_r, C_{r_1}, C_{r_2}}$ has minimum value R_{min}
24: $C_1 = C_{r_1}, C_2 = C_{r_2}$
25: END

First, the algorithm calculates the case differences vector $\triangle(C_j, C_k)$ for each pair of cases (C_j, C_k) and forms a matrix $M = [\triangle(C_j, C_k)]_{NxN}$. The algorithm searches for a pair (C_1, C_2) for which the relation between the case differences vector $\triangle(C_1, C_2)$ and gradient of C_2 is similar to the case differences vector $\triangle(C_n, C_r)$ and gradient of C_r, respectively (the remaining lines of the algorithm).

Two different situations can occur in which the considered pair of cases (C_1, C_2) yields an appropriate change to the beam number of C_r. First, there is a good match between the case differences vectors $\triangle(C_n, C_r)$ and $\triangle(C_1, C_2)$; that is, both case differences vectors have small or large values, and both gradients ∇C_r and ∇C_2 have either large or small values. This implies that C_1 and C_2 are in the region of similar characteristics as C_n and C_r, and, consequently, the relation between the beam numbers of C_n and C_r should be the same as that between C_1 and C_2. The second situation occurs when the case differences vector $\triangle(C_1, C_2)$ does not match $\triangle(C_1, C_2)$ well; i.e., $\triangle(C_n, C_r)$ is small, while $\triangle(C_1, C_2)$ is large or vice versa. If the corresponding differences vector and gradient offset each other, the pair (C_1, C_2) can then be used to change the beam number of C_r. For example, if $\triangle(C_n, C_r)$ is small, and the gradient of C_r is large, a small difference in attribute values triggers the change in BN. If $\triangle(C_1, C_2)$ is large, and the gradient of C_2 is small, a large difference in attribute values is needed to cause a change in BN. This means that even, if the differences between the attribute values of C_n and C_r and between C_1 and C_2 do not match, if the gradients of C_r and C_2 are opposite (small versus large) they affect the BN change in the corresponding regions of the domain space accordingly.

In the case base, not all OARs are present in all patient cases. The absence of a certain OAR in a case means that no radiation risk is imposed for that particular OAR (i.e., it is too far from the tumour to consider). Note that in the search for C_r, only the cases that have the same OAR as the new case are considered. Similarly, the (C_1, C_2) pair should have the same OARs. However, (C_1, C_2) and (C_n, C_r) pairs do not need to have the same OARs.

The beam number of the retrieved case is adapted based on the beam numbers of cases C_1 and C_2 in the

obtained triplet of cases. If cases C_1 and C_2 have the same beam number in their treatment plans, then the beam number in the treatment plan of the retrieved case C_r is not changed. If C_2 has a different number of beams than C_1, the beam number of the retrieved case is then changed by adding or subtracting a beam accordingly. The formula for adaptation is as follows:

$$BN_{C_r} = BN_{C_r} + (BN_{C_2} - BN_{C_1}) \qquad (10)$$

Examples of the retrieved beam number adaptation for the new case based on the difference of beam numbers in C_1 and C_2 is given in Table 2.

Case	BN(adaptation 1)	BN (adaptation 2)	BN (adaptation 3)
C_r	4	3	4
C_1	3	4	3
C_2	3	3	4
Suggested change BN_{C_2}-BN_{C_1}	3-3=0	4-3=+1	3-4=-1
$BN_{C_n} = BN_{C_r} + (BN_{C_2}$-$BN_{C_1})$	4 + 0 = 4	3 + 1 = 4	4 - 1 = 3

Table 2. Three examples of the retrieved beam number adaptation

In Adaptation 1, the retrieved case C_r suggests four beams. Since the difference in beam numbers between the retrieved pair of cases (C_1, C_2) is equal to 0, the retrieved beam number for the new case is applied without change (4 + 0 = 4). The remaining adaptations are to be interpreted similarly.

4.3. Adaptation-guided retrieval of beam angles

Experiments have demonstrated that the adaptation of beam number performed by the adaptation-guided retrieval method was more successful than the adaptation by NN. Therefore, an adaptation-guided retrieval approach to beam angle adaptation was developed and investigated.

A beam angle consists of two angles that determine the position of the patient during radiation: patient couch and gantry. However, in the majority of cases, the patient couch angle is set to $0\,°$ or $90\,°$, which would make the calculation of gradient difficult. Therefore, we use only the gantry angle as beam angle. In contrast to the beam number retrieval, which has to suggest a single value for the beam number, the beam angle retrieval deals with multiple values (i.e., an angle value for each beam). Furthermore, the order in which the beams are applied is not fixed; therefore, in the comparison of two cases, the best correspondence between beam angles must be established. The matching of the beams of two cases with the minimum difference between the corresponding angles determines the correspondence. The pseudo-code of the retrieval of the triplet of cases for beam angle adaptation is the same as for the beam number adaptation with a modification of the calculation of gradient given by Expression (6):

$$\nabla C = [\frac{\partial BA}{\partial V_1}, ..., \frac{\partial BA}{\partial V_6}] = [\frac{BA_C - BA_{C_{sim}}}{V_{C_1} - V_{C_{sim1}}}, ..., \frac{BA_C - BA_{C_{sim}}}{V_{C_6} - V_{C_{sim6}}}] \qquad (6')$$

where $BA_C = \{angle_{C_1}, ..., angle_{C_{BN}}\}$, and $BA_{Csim} = \{angle_{C_{sim1}}, ..., angle_{C_{simBN}}\}$.

The adaptation-guided retrieval of beam angles is illustrated using an example. Note that in our description, {} denotes a set, while [ ] denotes a vector.

Let us suppose that case C_n is described by the following attributes:

$C_n = [Vol = 0.0840, R = 0.4104, A = 0.4655, Dt = 0.5475, E = 0.1568, P = 0.0823]$.

The CBR system with adaptation-guided retrieval of beam number suggests that four beams be used in the treatment. Therefore, in the second phase of the retrieval, only cases with four beams are considered. We will illustrate one iteration in the adaptation-guided retrieval of beam angles. Let us suppose that in the current iteration, the retrieval considers case C_r with attributes $[0.0737, 0.1869, 0.2314, 0.1494, 0.1428, 0.0423]$ and beam angles $[300\,°, 306\,°, 235\,°, 90\,°]$.

The difference between C_n and C_r, $d(C_n, C_r)$, is calculated to be 0.0654 using Expression 1. The case differences vector $\triangle(C_n, C_r) = [0.0103, 0.2235, 0.3161, 0.0465, 0.0140, 0.0400]$ is calculated using Expression (4).

To calculate the gradient of case C_r, the most similar case to C_r (C_{sim}, with the same OAR and four beams) is retrieved with beam angles $[355°, 315°, 270°, 51°]$,

The gradient of case C_r is calculated using Expression (6'):

$$\nabla C_r = \begin{bmatrix} 0.0178 & 0.0460 & 1.7730 & 1.4467 & 0.2582 & 0.0540 \\ -0.0773 & -0.1994 & -7.6832 & -6.2691 & -1.1188 & -0.2340 \\ 0.0694 & 0.1789 & 6.8952 & 5.6261 & 1.0041 & 0.2100 \\ 0.0109 & 0.0281 & 1.0835 & 0.8841 & 0.1578 & 0.0330 \end{bmatrix}$$

Using Expression (8), R_{C_n,C_r} is calculated as follows:

$$R_{C_n,C_r} = \begin{bmatrix} 0.1842 & 10.2812 & 560.5443 & 67.2692 & 3.6232 & 2.1600 \\ -0.0008 & -0.0446 & -2.4290 & -0.2915 & -0.0157 & -0.0094 \\ 0.0007 & 0.0400 & 2.1799 & 0.2616 & 0.0141 & 0.0084 \\ 0.0011 & 0.0628 & 3.4255 & 0.4111 & 0.0221 & 0.0132 \end{bmatrix}$$

All combinations of (C_1, C_2) from the filtered case base are considered for adaptation of the beam angles suggested by C_r.

Let us consider a pair C_1 $[0.2832, 0.4042, 0.5408, 0.5743, 0.0134, 0.3145]$ with beam angles $[325°, 270°, 136°, 59°]$ and C_2 $[0.1000, 0.1800, 0.2700, 0.0743, 0.0140, 0.2311]$ with beam angles $[270°, 348°, 312°, 52°]$.

The case differences vector $\triangle(C_1, C_2) = [0.1832, 0.2242, 0.2708, 0.0653, 0.0040, 0.1111]$.

The most similar case to C_2 is case C_{sim2} with beam angles corresponding to C_2 $[0°, 88°, 310°, 330°]$. It is used to calculate the gradient of C_2 using Expression (6'):

$$\nabla C_2 = \begin{bmatrix} -0.1984 & -0.0888 & -1.2488 & -1.5966 & -0.2448 & -0.1566 \\ 0.2052 & 0.0918 & 1.2919 & 1.6517 & 0.2533 & 0.1620 \\ 0.2281 & 0.1020 & 1.4354 & 1.8352 & 0.2814 & 0.1800 \\ 102.6249 & 45.9147 & 645.9253 & 825.8396 & 126.6419 & 81.0001 \end{bmatrix}$$

Subsequently, R_{C_1,C_2} is calculated as follows:

$$R_{C_1,C_2} = \begin{bmatrix} -0.3636 & -0.1990 & -3.3818 & -1.0423 & -0.0098 & -0.1740 \\ 37.6116 & 20.5839 & 349.8448 & 107.8233 & 1.0139 & 18.0000 \\ 41.7907 & 22.8710 & 388.7164 & 119.8036 & 1.1266 & 20.0000 \\ 18.8058 & 10.2920 & 174.9224 & 53.9116 & 0.5069 & 9.0000 \end{bmatrix}$$

In order to calculate the difference $d(R_{C_n,C_r}, R_{C_1,C_2})$, the correspondence between the beam angles of C_r $[300°, 306°, 235°, 90°]$ and C_2 $[270°, 348°, 312°, 52°]$ has to be established. The angles of C_r and C_2 are matched in pairs, and the sum of angle differences is calculated. The matching, which results in the minimum sum, establishes the following correspondence between the beam angles of C_r and C_2:

C_r $[306°, 90°, 235°, 300°]$ and C_2 $[348°, 52°, 270°, 312°]$

Using Expression (7), we obtain:

$d(R_{C_n,C_r}, R_{C_1,C_2}) = 1.0935 \times 10^{+03}$

Let us suppose that this triplet of cases (C_r, C_1, C_2) has the minimum value of R_{C_r,C_1,C_2} among all other combinations of pairs (C_1, C_2) from the case base. Then, the angles in C_r will be adapted as follows:

$C_n = C_r + C_1 - C_2$

This will lead to the following beam angles suggested for the case C_n:

$C_n = [306°, 90°, 235°, 300°] + [325°, 270°, 136°, 59°] - [348°, 52°, 270°, 312°] = [283°, 308°, 101°, 47°]$.

5. Experiments and results

The case base used for the evaluation of the proposed adaptation methods consists of 80 cases collected from the Nottingham City Hospital. Out of 80 cases, 56 were treated with four beams, 23 cases with three

beams and one case with two beams.

5.1. Machine learning-based adaptation results

Based on suggestions for NN training found in the literature [35] and also our empirical results, the NN was trained with 80% of cases (65 cases), while 20% of cases (15 cases) were used for testing. The larger training set ensures that the NN model can capture the generalisation more efficiently, thus making a better prediction for the testing set.

We ran 50 experiments with one and two layers and a different number of neurons for each layer, ranging from nine to 34. This range was determined based on the number of input (seven attribute differences) and output (beam number change) neurons [?] and gives the best result in the testing set. In each run, the training and testing set cases were chosen arbitrarily from the case base. This ensured that a large number of combinations of cases was considered. The experiments have shown that the single layer with 15 hidden neurons produced the best result. The results obtained with a different number of neurons in the NN models with one hidden layer can be seen in Table 3. With 15 hidden neurons, the suggested beam number change was correct in 73.3% of the test cases. The leave-one-out cross-validation method described in Section 3 was used to measure the success rate of the CBR system, which used NNs for beam number adaptation.

NN hidden nodes	14	15	16	17	21
Correct *BN* change (in 15 test cases)	8	11	7	9	6
Average accuracy (%)	53.3	73.3	46.7	60	40

Table 3. NN results with different numbers of hidden neurons

The initial retrieval success rate of the CBR without adaptation was 61%. Including NNs increased it to 73.3%, which demonstrates that NNs can be useful for beam number adaptation.

Prior to using machine-learning tools, random guessing was used to suggest the number of beams; the results were worse than when NNs were used, with an average success rate of 41% based on 50 runs.

The naive Bayes classifier was evaluated in a slightly changed experimental setting. The classifier was trained on 70% of the cases in the case base (56) and tested on the remaining 30% of the cases (24) based on the generally accepted observation that an advantage of this classifier is that it only requires a small amount of training data [?]. Training the classifier with a larger number of cases (e.g., 90% of the case base (72 cases)) in the hope of learning the pattern better did not lead to any improvements. The obtained results are shown in Table 4.

Adapted BN	Correct BN change		
	0	**1**	**-1**
0	7	1	2
1	0	0	1
-1	10	3	0

Table 4. Naive Bayes classifier results

The naive Bayes classifier could adapt the beam number correctly only when no change to the beam number of the retrieved case was required. In all other cases, the suggested change was incorrect. For example, for 10 new cases, the correct beam number was suggested by the retrieved case, but the adaptation by the naive Bayes classifier suggested a decrease of 1 in the beam number. As the naive Bayes classifier considers every attribute's individual effect on the outcome, its low success rate indicates that the interplay between attributes should be considered when determining the correct beam number for the new case.

5.2. Adaptation-guided retrieval results

The results of experiments using the adaptation-guided retrieval method for beam number and beam angle adaptation are reported in this section.

5.2.1. Beam number

Comprehensive experiments were carried out to evaluate the performance of the adaptation-guided retrieval of beam number. The comparison of the results of the CBR with no adaptation with those of the CBR with AGRBN are presented in Table 5. The leave-one-out cross validation explained previously is used to measure the success rate. First, using AGRBN, in 80% of retrievals, the cases retrieved to suggest the beam number, C_r, are different from those retrieved in the original CBR system with no adaptation. The retrieved case in AGRBN is often not the most similar case to the new case. This suggests that the case most similar to the new case might not be the best to adapt. The experiments demonstrate that the performance of the CBR system with AGRBN is improved compared to the CBR with no adaptation. The success rate of CBR with AGRBN is 90%, which represents an improvement compared to the success rate of the CBR with no adaptation (61%).

Methodology	CBR with no adaptation	CBR with AGRBN
Correct *BN*	49	72
Incorrect *BN*	31	8
Success rate	61%	90%

Table 5. Comparison of CBR retrieval with and without the proposed AGRBN adaptation

The detailed results of the AGRBN adaptation are presented in Table 6. AGRBN suggested the correct adaptation (i.e., the beam number is adapted correctly) in 72 out of 80 test cases; that is, the beam number was increased by one, decreased by one or not changed. In 13 cases in which the retrieved beam number had three beams, AGRBN correctly suggested to add one beam to the retrieved case. In eight cases, AGRBN correctly adapted the retrieved beam number (i.e., decreased it by one). In 51 cases, no change was performed, which was correct because the beam number in the retrieved case was equal to the actual beam number of the new case.

		CBR with AGRBN (80 cases)
Correct BN	No change to *BN*	51
	Add beam to *BN*	13
	Subtract beam to *BN*	8
	Number of cases	**72**
Incorrect *BN*	Add beam to *BN*	5
	Subtract beam to *BN*	3
	No change to *BN*	0
	Number of cases	**8**
	Success rate	90%

Table 6. Results obtained using the proposed AGRBN adaptation

The experiments have shown that the CBR system with the AGRBN adaptation performed better than the original CBR system and the CBR system with machine learning-based adaptation.

5.2.2. Beam angle

Initially, there were 80 cases in the case base. However, one case that had two beams had to be discarded because at least three cases with the same beam number have to be present in the case base to form a triplet of cases (C_r,C_1,C_2) required for AGRBA. Hence, 79 cases were used. AGRBA adaptation did not produce any results for five cases because the system could not find similar cases with the same OARs or cases that could be used to calculate the gradient. AGRBA produced results for 74 cases, among which 20 cases were treated by medical physicists with three beams, and 54 cases were treated with four beams. The success rate of the proposed adaptation method was measured by calculating the beam angle error, E_{BA}, as explained in Section 3.

The success rate of AGRBA was compared with the success rate obtained with the CBR system with no adaptation. The comparison of the performances of the two methods is presented in Table 7. The

cases retrieved in the two systems may have different numbers of beams; namely, CBR with no adaptation retrieved 21 cases with three beams and 53 cases with four beams, while CBR with AGRBA retrieved 18 and 56 cases with three and four beams, respectively. The success rate of CBR without adaptation for beam angles is 39% (i.e., the suggested beam angles had beam angle error, E_{BA}, less than $30°$ in 29 out of 74 cases). However, the success rate of AGRBA is lower than that of CBR with no adaptation (i.e., the beam angle error, E_{BA}, was less than $30°$ in only 21 cases). AGRBA correctly performed the adaptation in only three retrievals of cases that suggested three beams. The success rate was higher (18 cases) when the retrieved case had four beams. This higher success rate for retrievals of cases with four beams is likely due to the larger number of cases with four beams in the case base. When a treatment requires three beams, the case base does not contain enough cases for the beam angle adaptation, which requires a triplet of cases as opposed to the retrieval of a single most similar case.

Case base	CBR with no adaptation	CBR with AGRBA
Retrieved cases with 3 beams	7/21	3/18
Retrieved cases with 4 beams	22/53	18/56
Total	29	21
Success rate	39%	29%

Table 7. Success rates of CBR with no adaptation and CBR with AGRBA

The next step in the analysis of the obtained results is to take a closer look at the beam angles suggested by CBR with no adaptation and CBR with AGRBA (i.e., examine the beam angle errors, E_{BA}, produced by the two methods). The following question arises: for how many new cases did CBR with AGRBA produce a smaller E_{BA} compared to CBR with no adaptation, even if the retrieval was not successful (i.e., E_{BA} was greater than $30°$). The results are presented in Table 8.

Case base	Number of cases in which CBR with AGRBA performed better than CBR with no adaptation
20 cases with 3 beams	6
54 cases with 4 beams	19
Total	25

Table 8. Improvement in E_{BA} using CBR with AGRBA compared to using CBR with no adaptation

In six cases with three beams and 19 cases with four beams, CBR with AGRBA resulted in a smaller E_{BA} than CBR with no adaptation. However, in the remaining majority of cases, CBR with AGRBA produced worse results.

Later, in each retrieval, we measure the difference in E_{BA} produced by the two CBR systems and count the number of cases in which one CBR system outperformed the other. Figure 3 presents the distribution of beam angle error improvement by CBR with AGRBA over CBR with no adaptation and vice versa. The difference in E_{BA} between the two systems is shown in intervals of $10°$. In higher ranges of E_{BA} difference, where the E_{BA} difference is greater than $60°$, both CBR with AGRBA and CBR with no adaptation have similar performances. That is, in two cases, AGRBA performs better by improving the E_{BA} compared to that of CBR with no adaptation, while the opposite occurred in two other cases. In a lower range of E_{BA} difference, where the difference in E_{BA} is smaller than $10°$, both systems perform similarly, with AGRBA exhibiting a better performance over CBR with no adaptation in seven cases and CBR with no adaptation having the better performance in eight cases. However, in the majority of the remaining cases, CBR with no adaptation had a smaller E_{BA} than CBR with AGRBA.

6. Conclusions

In this paper, we have described three adaptation methods, which can be referred to as knowledge-light approaches, to improve the results obtained by the retrieval phase of the previously developed CBR system.
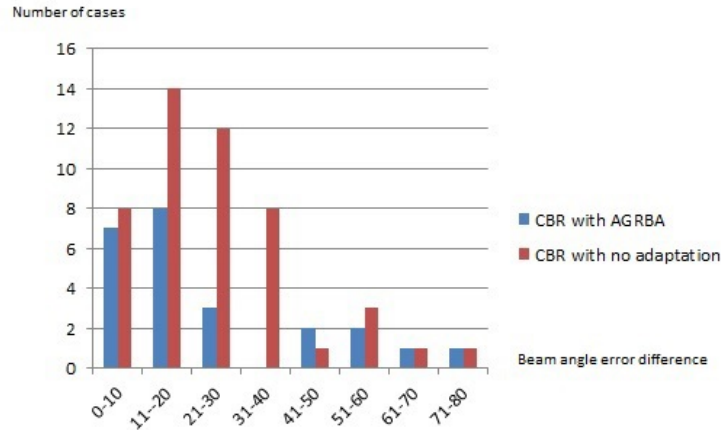
Figure 3. Difference in E_{BA} between CBR with AGRBA and CBR with no adaptation
.

First, existing machine-learning tools, neural networks and naive Bayes classifier were used to adapt the retrieved beam number for a new case. The naive Bayes classifier did not improve the current retrieval results as it did not consider the interplay among attributes. However, adaptation based on neural networks improved the success rate by 12%. This method could also adapt the beam number by stating the direction of the change (i.e., either increase or decrease the number of beams retrieved). Second, a new adaptation-guided retrieval method was developed to suggest changes to the beam number and beam angle suggested by a retrieved case.

The adaptation-guided retrieval approach for beam number improved the success rate of the CBR system by 29%. In most retrievals, the cases retrieved after implementing the adaptation-guided retrieval were different from the cases obtained by the CBR system with no adaptation because not only was the similarity of the retrieved case considered, but also how it can be adapted for the new case.

The beam angle adaptation is formulated as a more complex optimisation problem in terms of decision variables to consider. Its performance did not improve the CBR with no adaptation. In particular, this approach performed the worst for the cases with three beams as the relevant case base was too small.

In the current evaluation of the CBR system, we compare the beam number and beam angles suggested by the retrieval/adaptation process with the beam number and beam angles initially suggested by the medical physicists and stored in the cases in the case base. The disadvantage of this approach is that it only considers the retrieval to be successful when the retrieved cases have the same beam number or differ within a pre-defined threshold as suggested by the medical physicists. However, it may be that the CBR system suggests a better treatment plan in terms of the radiation dose distribution with a different beam number and/or beam angles. In reality, medical physicists evaluate the treatment plan based on the homogeneity of dose administered to cancer cells and the dose administered to OARs and non-tumour tissue (i.e., healthy tissue) by looking at the dose distribution lines (isodose lines) visualised in a treatment planning system. In our future research work, we plan to involve medical physicists in the evaluation of the developed CBR system.

Our next research step is to investigate adaptation that is more knowledge-intensive and simulates the reasoning process used by medical staff when adapting a retrieved treatment plan. The adaptation procedure will be documented, and a case base of adaptation cases will be built. Using domain knowledge from the medical physicists could lead to a better learning process of the CBR adaptation and hence may improve the success rate of the CBR system.

The number of cases we currently have in the case base is satisfactory considering the difficulties in collecting and processing cases for the case base. On the other hand, this number of cases is insufficient for any profound statistical analyses. Therefore, more cases will be added to the case base as they become

available, enabling us to evaluate the performance of the adaptation-guided retrieval using a larger case base.

We will also investigate the application of the developed CBR system to other radiotherapy treatment planning modalities, particularly intensive modulated radiotherapy treatment (IMRT).

Acknowledgements

References

[1] Matthias Ehrgott, Çiğdem Güler, Horst W Hamacher, and Lizhen Shao. Mathematical optimization in intensity modulated radiation therapy. *4OR*, 6(3):199–262, 2008.

[2] N. Mishra, S. Petrovic, and S. Sundar. A self-adaptive case-based reasoning system for dose planning in prostate cancer radiotherapy. *Medical Physics*, 38(12):6528, 2011.

[3] R. Bergmann and W. Wilke. Towards a new formal model of transformational adaptation in case-based reasoning. In *ECAI*, pages 53–57, 1998.

[4] A. Schlaefer and S. Dieterich. Feasibility of case-based beam generation for robotic radiosurgery. *Artificial Intelligence in Medicine*, 2011.

[5] J. Liber and B. Bresson. Case-based reasoning for breast cancer treatment decision helping. *Lecture Notes in Computer Sceince*, pages 173–185, 1999.

[6] J.L. Kolodner. An introduction to case-based reasoning. *Artificial Intelligence Review*, 6(1):3–34, 1992.

[7] R. Jagannathan, S. Petrovic, A. McKenna, and L. Newton. A novel two phase retrieval mechanism for a clinical case based reasoning system for radiotherapy treatment planning. *International Journal on Artificial Intelligence Tools*, 21(04):46–70, 2012. doi: 10.1142/S0218213012400179. URL http://www.worldscientific.com/doi/abs/10.1142/S0218213012400179.

[8] Allen Holder and Bill Salter. A tutorial on radiation oncology and optimization. In *Emerging Methodologies and Applications in Operations Research, (Chapter 4)*, pages 4–1. H. Greenberg (Ed.), Kluwer Academic Press,Boston, MA, 2004.

[9] M. Oldham, VS Khoo, CG Rowbottom, JL Bedford, and S. Webb. A case study comparing the relative benefit of optimizing beam weights, wedge angles, beam orientations and tomotherapy in stereotactic radiotherapy of the brain. *Physics in medicine and biology*, 43:2123, 1998.

[10] H. Rocha, J.M. Dias, B.C. Ferreira, and M.C. Lopes. Combinatorial optimization for an improved transition from fluence optimization to fluence delivery in imrt treatment planning. *Optimization*, 61(8):969–987, 2012. doi: 10.1080/02331934.2011.607498. URL http://dx.doi.org/10.1080/02331934.2011.607498.

[11] Matthias Ehrgott and Rick Johnston. Optimisation of beam directions in intensity modulated radiation therapy planning. *OR Spectrum*, 25(2):251–264, 2003.

[12] W De Neve, Carlos De Wagter, K De Jaeger, M Thienpont, C Colle, S Derycke, and J Schelfhout. Planning and delivering high doses to targets surrounding the spinal cord at the lower neck and upper mediastinal levels: static beam-segmentation technique executed with a multileaf collimator. *Radiotherapy and oncology*, 40(3):271–279, 1996.

[13] X. Song, S. Petrovic, and S. Sundar. A case-based reasoning approach to dose planning in radiotherapy. In *Workshop Proceedings, The 7th International Conference on Case-Based Reasoning ICCBR*, volume 7, pages 13–16. Citeseer, 2007.

[14] J. Berger. Roentgen: Radiation therapy and case-based reasoning. In *Artificial Intelligence for Applications, 1994., Proceedings of the Tenth Conference on*, pages 171–177. IEEE, 1994.

[15] Dusan Teodorovic, Milica Selmic, and Ljiljana Mijatovic-Teodorovic. Combining case-based reasoning with bee colony optimization for dose planning in well differentiated thyroid cancer treatment. *Expert Systems with Applications*, 40(6): 2147–2155, 2013.

[16] Kristian J. Hammond. Chef: A model of case-based planning. In *AAAI*, pages 267–271, 1986.

[17] D.W. Aha. The omnipresence of case-based reasoning in science and application. *Knowledge-Based Systems*, 11(5): 261–273, 1998.

[18] A. Holt, I. Bichindaritz, R. Schmidt, and P. Perner. Medical applications in case-based reasoning. *The Knowledge Engineering Review*, 20(03):289–292, 2005.

[19] S. Petrovic, N. Mishra, and S. Sundar. A novel case based reasoning approach to radiotherapy planning. *Expert Systems with Applications*, 38(9):10759–10769, 2011.

[20] Rupa Jagannathan and Sanja Petrovic. *A Local Rule-Based Attribute Weighting Scheme for a Case-Based Reasoning System for Radiotherapy Treatment Planning*, volume 7466 of *Lecture Notes in Computer Science*, pages 167–181. Springer Berlin Heidelberg, 2012. URL http://www.seolinks.fr.

[21] Beatrice Fuchs, Jean Lieber, Alain Mille, and Amedeo Napoli. Differential adaptation: An operational approach to adaptation for solving numerical problems with cbr. *Knowledge-Based Systems*, 68 (0):103 – 114, 2014. ISSN 0950-7051. doi: http://dx.doi.org/10.1016/j.knosys.2014.03.009. URL http://www.sciencedirect.com/science/article/pii/S0950705114000926.

[22] W. Wilke and R. Bergmann. Techniques and knowledge used for adaptation during case-based problem solving. *Tasks and Methods in Applied Artificial Intelligence*, pages 497–506, 1998.

[23] D. Hennessy and D. Hinkle. Applying case-based reasoning to autoclave loading. *IEEE Expert*, 7(5):21–26, 1992.

[24] CM Vong, TP Leung, and PK Wong. Case-based reasoning and adaptation in hydraulic production machine design. *Engineering Applications of Artificial Intelligence*, 15(6):567–585, 2002.

[25] M. D'Aquin, J. Lieber, and A. Napoli. Adaptation knowledge acquisition: a case study for case-based decision support in oncology. *Computational intelligence*, 22(3-4):161–176, 2006.

[26] Barry Smyth and Mark T. Keane. Adaptation-guided retrieval: Questioning the similarity assumption in reasoning. *Artificial Intelligence*, 102:249–293, 1998.

[27] Neil McDonnell and Pádraig Cunningham. A knowledge-light approach to regression using case-based reasoning. In ThomasR. Roth-Berghofer, MehmetH. Göker, and H.Altay Güvenir, editors, *Advances in Case-Based Reasoning*, volume 4106 of *Lecture Notes in Computer Science*, pages 91–105. Springer Berlin Heidelberg, 2006. ISBN 978-3-540-36843-4. doi: 10.1007/11805816. URL `http://dx.doi.org/10.1007/11805816`.

[28] N. Mishra, S. Petrovic, and S. Sundar. A knowledge-light nonlinear case-based reasoning approach to radiotherapy planning. In *Tools with Artificial Intelligence, 2009. ICTAI'09. 21st International Conference on*, pages 776–783. IEEE, 2009.

[29] Julien Henriet, Pierre-Emmanuel Leni, Ramy Laurent, and Michel Salomon. Case-based reasoning adaptation of numerical representations of human organs by interpolation. *Expert Systems with Applications*, 41(02):260–266, 2013. ISSN 0957-4174. doi: http://dx.doi.org/10.1016/j.eswa.2013.05.064. URL `http://www.sciencedirect.com/science/article/pii/S095741741300362X`.

[30] J.M. Corchardo, Brian Lees, Colin Fyfe, Nigel Rees, and Jim Aiken. Neuro-adaptation method for a case-based reasoning system. In *Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence. The 1998 IEEE International Joint Conference on*, volume 1, pages 713–718. IEEE, 1998.

[31] NEMA. Dicom - digital imaging and communications in medicine. Technical report, National Electrical Manufacturers Association, 1998.

[32] A.J. Gonzalez, Lingli Xu, and U.M. Gupta. Validation techniques for case-based reasoning systems. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 28(4):465–477, Jul 1998. ISSN 1083-4427. doi: 10.1109/3468.686707.

[33] Nir Friedman, Dan Geiger, and Moises Goldszmidt. Bayesian network classifiers. *Machine learning*, 29(2-3):131–163, 1997.

[34] Wei Liang Yeow, Rohana Mahmud, and Ram Gopal Raj. An application of case-based reasoning with machine learning for forensic autopsy. *Expert Systems with Applications*, 41(7): 3497–3505, 2013. ISSN 0957-4174. doi: http://dx.doi.org/10.1016/j.eswa.2013.10.054. URL `http://www.sciencedirect.com/science/article/pii/S0957417413008713`.

[35] Brian D. Ripley. *Pattern recognition and neural networks*. Cambridge university press, 1996.