Knowledge Acquisition and Design Support in a Medical Domain*

Jeffrey Berger

Artificial Intelligence Laboratory

University of Chicago

1100 East 58th Street

Chicago, IL 60637

Email: berger@cs.uchicago.edu

Abstract

ROENTGEN is a case-based system which provides design support for a human designer of radiation therapy plans. It learns both from specific problem-solving experiences and from direct instruction from the user. The first sort of learning is the normal case-based method of storing problem solutions so that they can be reused. The second sort is necessary because ROENTGEN does not, initially, have an internal model of the physics of its problem domain. This dependence on explicit user instruction brings to the forefront representational questions regarding indexing, failure definition, failure explanation and repair. This paper presents the techniques used to support ROENTGEN in its knowledge acquisition and design activities.

Introduction

Some AI research efforts are launched by the desire to find general answers to important questions about cognition. Others are the result of wanting to discover the implications of applying a set of ideas to a particular practical problem. The ROENTGEN project was born out of the second set of circumstances.

The aim of the ROENTGEN(Berger et al. 1990; Berger and Hammond 1991) project is to examine the issues which arise when case-based(Hammond 1986; Schank 1982; Lebowitz 1980; Kolodner 1984; Rissland and Ashley 1986; Sycara 1987; Bareiss 1989; Koton 1988) techniques are used to provide design support in a real-world domain. Our approach is to build a system which learns both from specific problem-solving experiences and from direct instruction from the user. The first sort of learning is the standard case-based method of storing problem-solving experiences in a case memory so their results can be used in similar, future sit-

uations. The second sort of learning is necessary because ROENTGEN does not, initially, have an internal model of the physics of its problem domain. When ROENTGEN is unable to complete a step in the design of an RT plan because of missing domain physics knowledge, the user is asked to supply the missing knowledge. The new knowledge then becomes part of the system's knowledge-base.

A consequence of this approach is that ROENTGEN's capabilities develop through use. In the initial stage, it is primarily concerned with putting radiation therapy (RT) plans for patients in its case memory and adding pieces of domain knowledge to its knowledge-base. Soon, it is able to use this stored information to assist the human user in the design effort. It can make initial suggestions for plans to treat patients, detect problems when the effects of a plan are simulated, suggest possible explanations for the problems it detects, and recommend repairs to a faulty design to correct problems. The endpoint of this evolution is a competent, stand-alone RT plan designer.

ROENTGEN depends on explicit, user-instruction as the source of its knowledge regarding indexing, failure definition, failure explanation and repair. Consequently, these four areas have been important research concerns:

- Indexing of past cases. What features of a case are important for retrieving useful suggestions from memory? Can a language of primitives and operators be developed that supports the definition of new features?

- Failure detection. What constitutes a failure or undesirable condition in the simulated results of RT plan application? Can a failure definition language be developed so that the user can point out new sorts of failures to the system as they arise?

- Failure explanation. What sort of information is necessary to provide an explanation for observed failures so that they can be repaired in the current case and avoided in future ones?

- Plan repair. What are the primitive design repairs for an RT plan and what methods of combining them

are required to fix the sorts of design failures which actually occur in the domain?

We present our answers to these concerns in this paper.

Case-based Reasoning

The primary source of knowledge for a case-based reasoning (CBR) system is its memory of past problem-solving experiences. When solving a new problem, the CBR system retrieves a case from this memory which closely resembles the new problem. The system uses the solution for the retrieved case as an initial solution for the new problem. It then checks to see if the solution actually works. If it does, the problem-solving process is succesfully concluded. If it doesn't, the system uses repair knowledge to fix the attempted solution so that it correctly solves the new problem. Finally, the system stores the new solution and problem in its case memory for future use. In this way, the system learns from its own problem solving experience.

The motivation behind the CBR paradigm is that it is often easier to repair a nearly correct solution than to build a fresh one from first principles. The paradigm takes advantage of the regularity which exists in many natural problem areas: small changes in the features of a problem usually result in small changes to the solution. ROENTGEN uses a case-based approach in order to take advantage of this regularity in the domain of radiotherapy treatment planning (RTP).

ROENTGEN differs from previous work(Paluszyński 1990; Kalet and Paluszyński 1985) in this problem domain in several ways. Its knowledge is memory-based rather than rule-based. Hence, it will learn new plans through its experience and become more capable at its task over time. ROENTGEN is an interactive system as well as a stand-alone one. It can support the human designer with suggestions throughout the design task and ask the designer to supply knowledge it needs to solve a problem. This knowledge then becomes a permanent part of the system's knowledge. Finally, ROENTGEN doesn't attempt to produce an "optimum" plan. The question of whether or not a plan is acceptable is left to the human designer. This is a strength, not a weakness. Researchers in RTP have yet to agree on an objective definition of RT plan optimality(Goitein and Niemierko 1989).

The RT Planning Problem

The radiotherapy designer—the *dosimetrist*—has the job of designing acceptable RT treatment plans. The crux of the RT planning problem is that some vital tissues have lower radiation tolerances than cancer cells(Bentel *et al.* 1989). Acceptable plans must deliver the prescribed dose, sufficient to destroy the cancer cells, while ensuring that vital tissues do not receive more than they can tolerate.
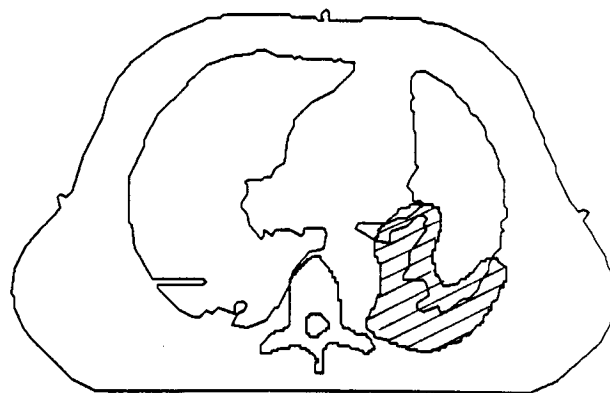

Figure 1: The treatment plane cross section, as seen from below, for Cassius. The outermost polygon is the body outline. The large, distorted crescent-shapes on the right and left are lungs. The small circle is the spinal cord. The shaded polygon is the tumor.

In external photon beam therapy, multiple beams of very high energy x-rays are directed into the patient from different locations. The dosimetrist's task is to determine the number of beams, their arrangement, relative weights, shape, energy and determine the values of other plan parameters required to produce an acceptable result.

As presented to ROENTGEN and to the dosimetrist, a new patient consists of two elements: an image of an anatomical cross section through the patient (see figure 1) and dose information.

The main piece of information defining a new case is the image of an anatomical cross section through the patient. The plane of the cross section is the plane on which all the treatment beams in the plan to be developed will lie.[1] Outlines of the tissues in the cross section are described by polygons. The cross section is presented graphically to the dosimetrist during the design of a treatment plan.

The dose information consists of the prescribed dose and the dose history. The prescribed dose is the minimum acceptable dose to the target volume. The dose history consists of the previous dose to the target and other tissues of interest. Frequently, radiotherapy proceeds in stages. The effects of radiation are cumulative, so doses from previous stages must be taken into account.

Dosimetrists use a *generate, test and debug* design strategy. When a problem specification is first pre-

---

[1]Most external beam treatment planning is done assuming coplanar beams.

sented to the dosimetrist, she decides on the general sort of plan to use—the *generate* phase. Then, looking at a simulation of the plan's effects, the dosimetrist repairs the plan to correct undesirable effects. This is the *test-and-debug* phase. The dosimetrist repeats the latter phase until an acceptable plan results.

roentgen's Architecture

ROENTGEN provides support to the dosimetrist during both the generate and test-and-debug phases. It contributes to plan generation by retrieving the most similar case it can find in its *Patient Memory* and then fitting the retrieved solution to the new patient. During the test-and-debug phase, ROENTGEN uses its knowledge of explanations for past failures to explain and repair failures currently being experienced. ROENTGEN stores knowledge it obtains during either phase to use in future cases.

Plan generation

RT plans can be categorized in terms of the number of beams in the plan and their geometric relationship to eachother. Within each category, plans can be characterized by their *orientation angle*. Figure 3 is an example of an *opposed beam* plan with an orientation angle of 10 degrees.

To capture this distribution of plans, ROENTGEN organizes *Patient Memory* on the basis of *plantype indices* which describe the basic beam arrangement and the orientation angles at which each plan is likely to succeed. Thus, ROENTGEN's first task is to construct plantype indices for the new case, and use those indices to retrieve past cases from memory.

In general, multiple cases will be retrieved. ROENTGEN ranks the past cases by several criteria: match of orientation angle, strictness of dose limits, and plan simplicity. Cases are also ranked on the basis of *modification level*, a measure of plan refinement. A plan with a higher modification level has been fine-tuned to satisfy more constraints than a plan of the same type with lower modification level. In the end, the RETRIEVER returns the first plan in the list.

Next, the MODIFIER tailors the plan from the retrieved case to fit the new patient. It uses critics in *Critic Memory* to look for and correct problems ROENTGEN has learned about from past failures. ROENTGEN's critics are the main way the system generalizes its failure knowledge from specific to more general settings.

The MODIFIER produces an untested treatment plan for the new patient and a *plantype index*. The plantype index is a composite feature which describes the orientations at which the tailored plan is expected to be acceptable (see Section). It becomes the first item on the *Explanation Stack*. At the end of the design of an acceptable plan, this stack will also contain the explanations for any failures which were corrected to arrive at the final plan. Thus, it will trace the design
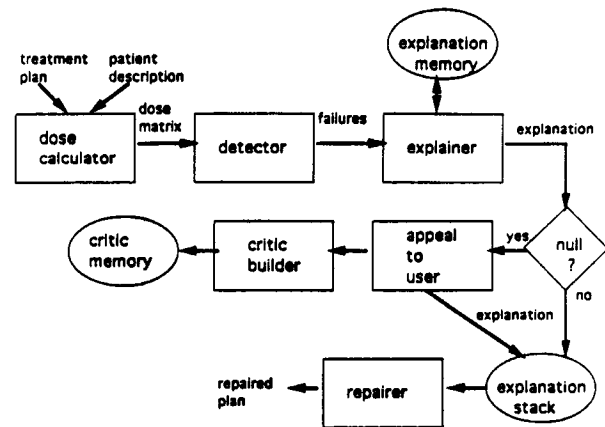

Figure 2: The ROENTGEN modules which test and debug a treatment plan.

process of the plan from the starting point to the finish. This information will be used later to store acceptable plans in *Patient Memory*.

Test and debug

The second stage of RT design is the test-and-debug stage. Testing is done using a DOSE CALCULATOR. This module is not a part of ROENTGEN, but an independent process which provides its results to both the dosimetrist and ROENTGEN. It produces a 40 by 40 array—the dose matrix—which represents the amount of radiation deposited by the proposed treatment plan at each point of a grid superimposed on the patient cross section.

ROENTGEN analyzes the dose matrix in the context of the tissues defined by the patient cross section and the dose requirements for the patient to produce a list of failures. It then looks for an explanation in *Explanation Memory* which accounts for the greatest number of the failures. An explanation links three things: the failures being explained, the features in the geometry of the patient cross section which cause the failures, and repair techniques to fix the failed plan. The explanation is pushed onto the *Explanation Stack* for later use.

If ROENTGEN is unable to find an explanation for the failures, it appeals to the dosimetrist to provide a new explanation knowledge structure with the failures being explained, the appropriate repair techniques, and supporting causal features.

The explanation is also used to construct a new critic. The critic will be used in the future to change

retrieved plans to avoid this failure.

Finally, the REPAIRER gets the list of repair techniques from the explanation just pushed on the *Explanation Stack*. The prefered repairs are placed at the front of this list. Each repair has an associated applicability test. The REPAIRER applies the first repair it encounters whose test is satisfied to the failed plan. The test-and-debug cycle is then repeated on the repaired plan.

Storing acceptable plans

Although ROENTGEN can detect undesirable conditions in the effects of a plan, it does not have the ability to determine whether or not a plan is acceptable for a given patient. In fact, sometimes plans with undesirable effects must be accepted because no better plan can be constructed. The attending physician decides which plan to accept based on the alternatives prepared by the dosimetrist.

Once a plan has been accepted, the plan and patient description may be stored. If the plan was the result of using a critic to correct a retrieved plan, or involved a repair stemming from a new, user-provided explanation, ROENTGEN will store the new case. The case will also be stored if its plantype index enumerates a different set of angles than the cases already in memory stored under that index. In this way, if a similar problem is seen in the future, no recourse to critics or repairs will be needed as the required plan will already be present in *Patient Memory*.

Before storing a new case, ROENTGEN builds a new index from the information in the *Explanation Stack*—the plantype index of the generate phase and explanations from the test-and-debug phase. The index from the generate phase enumerates the angles for which the retrieved plan should have been acceptable in the new patient. Each explanation in the stack contains a list of features in the geometry of the current patient cross section which *cause* the failures being explained. These features, together with the original index used in the generate phase, form a new plantype index of the orientation angles at which the new plan should work.

The STORER places the new case in memory using this new index. Then the index, and its method of computation, are added to the list of plantype indices ROENTGEN knows.

A Vocabulary for RT Planning

Because ROENTGEN depends so heavily on its memory of past cases, a good indexing vocabulary is essential. In particular, the vocabulary must be capable of representing the geometric relations of the patient cross section which are important for RTP. ROENTGEN's vocabulary is constructed in steps from *radial characteristics*, through *radial features*, to *plantype indices*:

- A radial characteristic represents a fact about the cross section geometry, and is computed directly

from the cross section's polygon vertices.

- Radial features tell us for which angles some property of interest holds. Radial features can be built from radial characteristics or from already existing radial features.

- Plantype indices enumerate the plan orientation angles at which basic necessary conditions should hold for the likely success of an associated therapy plan. These indices are the primary distinguishing features of cases in memory.

Defining Failures

A failure in a therapy plan is an undesirable condition in the dose distribution which the plan produces. Examples of failures are spinal cord overdose, target underdose and large, healthy-tissue, high-dose area.

To fix such failures, ROENTGEN must specify where they occur in the body cross section. For repair purposes, beam-relative locational information is more useful than absolute locations. ROENTGEN's locational terms divide the patient cross section into regions with respect to a specified beam. For example, the path of a beam through the body is divided into *entry*, *mid*, and *exit* regions, with the mid region containing the tumor.

Rather than attempt to predefine and thus restrict all the sorts of failures the can be recognized, we have chosen to develop a failure definition language. Using this language, the user can specify which problems to attend to. (This approach is similar in spirit to (Bareiss 1989).)

The basic data structure on which this language is built is a raster produced by merging the dose matrix, computed by the DOSE CALCULATOR, with the patient cross section contours. An individual cell in the raster corresponds to a small area of the cross section and contains the dose deposited in the area and the body tissues—spinal cord, target, etc.—which occur there. The language consists of functions which can be applied to subsets of the raster to produce other subsets, operators which map subsets to real numbers, and predicates which take raster subsets and/or other arguments to produce boolean values.

Dose-greater-than-or-equal-to is an example of the first sort. It takes as arguments a subset of the raster and a number and returns all the cells in the subset which receive a dose greater than or equal to that number.

At-beam-relative is another example of the first sort and is the way the beam-relative coordinates of a region of interest are specified. It takes a subset of the raster, a beam number and a beam-relative location specifier as arguments. It returns all the cells in the subset at the location specified with respect to the beam indicated.

Cell-area is an operator which computes the area of a collection of raster cells.

144

Finally, *ratio-greater-than-or-equal-to* is a predicate which takes three number—a quotient, a divisor, and a limit—and returns a true value iff the ratio of the quotient to the divisor is greater than or equal to the limit.

The user defines the failures ROENTGEN attends to with this vocabulary. Once a definition has been made, ROENTGEN's DETECTOR will be able to detect the failure when it occurs.

Example

Our example is designed to illustrate the use of plantype indices in memory access, how the dosimetrist provides the information to build explanations and critics, and what role the retrieval vocabulary plays in the process.

Suppose the description of the patient Regan with associated plan (figure 3) is in memory along with other cases. The next case to be planned is Cassius (figure 1). For each plantype index known to the system, the RETRIEVER computes the values of that index using Cassius' cross section as input. This yields a list of plantype indices. Included in the list is the opposed-beam-index under which Regan is stored. The index enumerates several angles for Cassius indicating that an opposed-beam-plan at any of the enumerated angles may be acceptable. For the purposes of our example, we assume that the Regan case is the prefered opposed-beam plan and its plan is selected to be the initial plan for Cassius.

When Regan's plan is applied to Cassius, however, there is a problem. Regan's target is well away from the body surface. Cassius' target is much closer. When the effects of the plan are simulated, the part of Cassius' target nearest Beam 1 receives too low a dose. ROENTGEN's DETECTOR notes the failure, but the EXPLAINER can't find an explanation for this general type of failure in *Explanation Memory*. The system asks the user for help.

This failure occurs because an entering photon beam deposits little radiation near the skin surface. There is a buildup region between the skin surface and the point inside the body where a beam deposits the maximum amount of radiation. The width of the buildup region depends, in part, on the beam's energy. For 24 megavolt (*mv*) beams, it is approximately 5 *cm*. A portion of Cassius' target lies within this region for Beam 1 of the plan.

The dosimetrist must define a new radial feature which can be used to predict when the failure is likely to occur in future cases. She builds this feature from the radial characteristic *skin-target-depth* which lists the distances from the skin to the target. The method of computing the feature is to apply the feature constructor *measure-underlimit* with the limit 5.0 *cm*. This returns a radial feature which, given a patient cross section, enumerates the angles at which a 24 *mv* beam is likely to underdose the target due to insuffi-
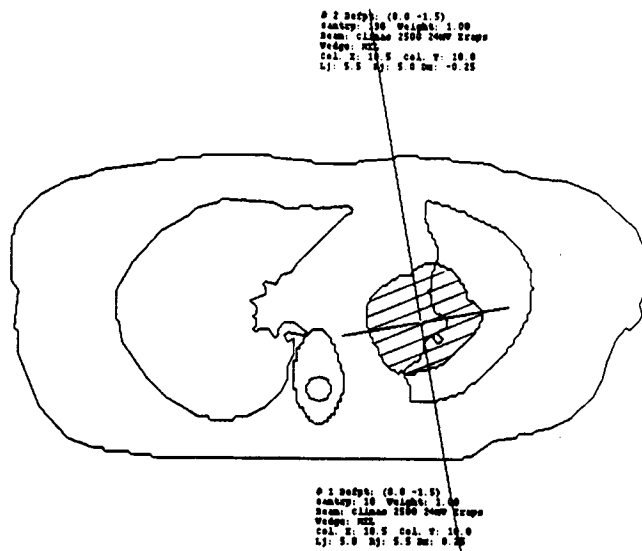

Figure 3: The treatment plane cross section, as seen from below, showing an opposed beam plan for Regan. The plan uses two 24 *mv* beams. The longer vertical line corresponds to the center-line of the two beams. Beam number 1 is directly below the patient; beam number 2, directly above. The shorter horizontal line indicates the width of both beams as they pass through the target. Note that the beams are wide enough to cover the target with a margin to spare.

cient buildup margin. The dosimetrist names the feature "insufficient-24mv-buildup-margin", and enters it as a feature predictive of this failure.

Next, a repair technique must be defined. Since the buildup region is narrower for lower beam energies, the dosimetrist can repair this problem by lowering the implicated beam's energy. So the repair technique is to lower Beam 1's energy from 24 *mv* to 6 *mv* (and make other minor adjustments).

Together, the predictive feature and the repair form the *explanation* of the failure. The completed explanation is placed in *Explanation Memory*. If ROENTGEN detects the same sort of failure in future plan applications, it will be able to explain it without appeal to the user. The explanation is pushed on the *Explanation Stack* as well, to record this failure/repair episode in the trace of the current plan's development. A new critic is also built and placed in *Critic Memory*. This

critic will be used in all future cases where an opposed beam plan (regardless of modification level) is passed to the MODIFIER, unless the plan already takes this problem into account.

Returning to Cassius, the REPAIRER now has the necessary information to proceed. Beam 1 in the plan for Cassius has its energy lowered to 6 *mv*. When the plan is simulated again, no failures are found. The plan is acceptable; and the dosimetrist indicates to ROENTGEN that the case should be stored.

The STORER must first construct a new plantype index from the design history contained in the *Explanation Stack*. The stack holds the opposed-beam-index used to retrieve the Regan plan and the explanation for the "Target underdosed on beam 1's side" failure detected during plan design. The new index is formed by applying AND to the opposed-beam-index and the insufficient-24mv-buildup-margin feature which caused the failure. The modification level of the new index is 1. Using the new index, the STORER places Cassius and the repaired treatment plan in *Patient Memory*.

Finally, the method of computation for the new index is added to the system's repertoire of plantype indices. ROENTGEN will henceforth retrieve this plan when future patients have a cross section which is amenable to an opposed beam plan and in which the target is too close to the skin surface to allow using a plan with two 24 *mv* beams. Regan may be retrieved as well, but that won't matter. Since the plantype index for Regan has a lower modification level (0) than Cassius, Cassius with its plan involving one 6 *mv* beam will be prefered over Regan.

Conclusion

ROENTGEN learns from its problem-solving experiences and from explicit instruction from the user. To support these activities, we have developed an architecture which combines classic, case-based elements with opportunities for user intervention. Declarative representational forms have been found which enable user definition of methods for computing memory indices and plan failures, and for capturing user-provided failure explanations.

While the complete system is still under construction, we are encouraged by the progress made so far.

References

Bareiss, R. 1989. *Exemplar-Based Knowledge Acquisition*, volume 2 of *Perspectives in Artificial Intelligence*. Academic Press, San Diego, CA.

Bentel, G. C., Nelson, C. E., & Noell, K. T. 1989. *Treatment Planning and Dose Calculation in Radiation Oncology*. Pergamon Press, New York and Oxford, fourth edition.

Berger, J. & Hammond, K. J. 1991. Roentgen: A memory-based approach to radiation therapy treatment design. In Bareiss, R., editor, *Proceedings: Case-Based Reasoning Workshop*, San Mateo, California. Morgan Kaufmann Publishers, Inc.

Berger, J., Hammond, K., & Chen, G. T. Y. 1990. Roentgen: A case-based approach to radiation therapy planning. In *The Proceedings of the 32nd Annual Scientific Meeting of the American Society for Therapeutic Radiology and Oncology*.

Berger, J. 1992. Roentgen: Case-based reasoning and radiation therapy planning. In *The Proceedings of the 16th Annual Symposium on Computer Applications in Medical Care*, New York. McGraw-Hill.

Goitein, M. & Niemierko, A. 1989. Biologically based models for scoring treatment plans. In Zink, S., editor, *Future directions of computer-aided radiotherapy*, San Antonio, TX. National Cancer Institute.

Hammond, K. 1986. Chef: A model of case-based planning. In *The Proceedings of the Fifth National Conference on Artificial Intelligence*, Philadelphia, PA. AAAI.

Kalet, I. J. & Paluszyński, W. 1985. A production expert system for radiation therapy planning. In Levy, A. H. & Williams, B. T., editors, *AAMSI Congress*, pages 315–319, Washington, D.C. American Association for Medical Systems and Informatics.

Kolodner, J. 1984. *Retrieval and Organizational Strategies in Conceptual Memory*. Lawrence Erlbaum Associates, Hillsdale, N.J.

Koton, P. 1988. *Using Experience in Learning and Problem Solving*. PhD thesis, MIT.

Lebowitz, M. 1980. *Generalization and Memory in an Integrated Understanding System*. PhD thesis, Yale University. Technical Report 186.

Paluszyński, W. 1990. *Designing Radiation Therapy for Cancer: An Approach to Knowledge-based Optimization*. PhD thesis, University of Washington.

Rissland, E. & Ashley, K. 1986. Hypotheticals as heuristic device. In *The Proceedings of the Fifth National Conference on Artificial Intelligence*. AAAI.

Schank, R. 1982. *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*. Cambridge University Press.

Sycara, E. P. 1987. *Resolving Adversarial Conflicts: An Approach Integrating Case-based and Analytic Methods*. PhD thesis, School of Information and Computer Science, Georgia Institute of Technology.