Optimization problems for radiation therapy treatment planning

H. Edwin Romeijn

Department of Industrial and Operations Engineering
The University of Michigan, Ann Arbor, Michigan

3rd Nordic Optimization Symposium
KTH, Stockholm, Sweden
March 13–14, 2009

Outline

Radiation Therapy

- Each year, more than 10 million people worldwide are newly diagnosed with cancer
  - about 50-65% of these will be treated by some form of radiotherapy
  - about half of these may benefit from *external beam conformal radiation therapy*
- We will discuss optimization problems dealing with the design of *effective* and *efficiently deliverable* treatment plans
  - in particular for *Intensity Modulated Radiation Therapy* (IMRT)

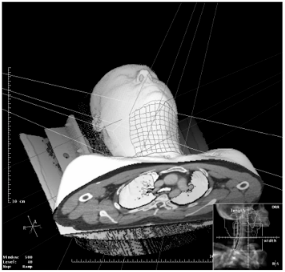
Radiation Therapy Delivery

- Patients are generally treated using a radiation source generating high-energy photons which is mounted on a gantry that can rotate around the patient
  - linear accelerator: X-rays
  - Cobalt source: γ rays

- The radiation source has constant output (intensity)

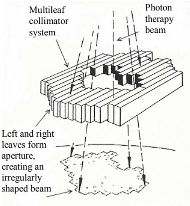
Radiation Therapy Delivery

- During radiotherapy, beams of radiation pass through a patient, killing both normal and cancerous cells
- Patients are therefore generally treated with beams from *multiple directions*



Intensity Modulated Radiation Therapy (IMRT)

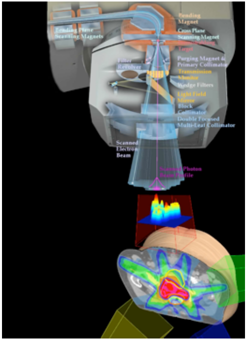
- In IMRT, a *multi-leaf collimator* (MLC) system allows modulation of the intensity (fluence) of the beams
  - leaves can dynamically block part of a beam to form different *apertures*:



- This allows for the creation of treatment plans that yield complex dose distributions that
  - adequately cover *targets*
  - preserve the functionality of *critical structures*

Intensity Modulated Radiation Therapy (IMRT)

- The superposition of the set of apertures and their intensities corresponds to an *intensity profile* for each beam direction



IMRT treatment planning

- IMRT treatment planning is often performed in three phases
  1. Beam orientation selection
     - Usually performed manually
  2. Fluence map optimization
     - Discretize each beam into small *beamlets* or *bixels*
     - Determine an *intensity profile* for each beam (i.e., an intensity for each bixel) that yields a high quality treatment plan
  3. Leaf sequencing
     - Decompose the intensity profile of each beam into a collection of deliverable apertures and corresponding intensities

Optimization problems

- The remainder of this talk will consist of two components:

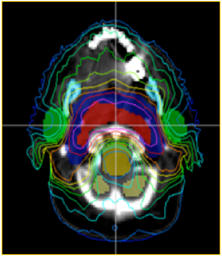
1. *Treatment plan design*
    - Quantifying the quality of a treatment plan
2. *Treatment plan delivery*
    - Optimizing the delivery efficiency of a treatment plan

Radiotherapy side effects

- Head-and-neck cancer
  - Delivering too much dose across the spinal cord has the same effect as cutting it
  - Preserving salivary gland function is very important to the quality of life of the patient as well
    - this function allows eating, speaking, maintaining oral hygiene
- Prostate cancer
  - Bowel complications (bleeding, inflammations) should be avoided
- Lung cancer
  - Overdosing the lung may cause a fatal buildup of fluid

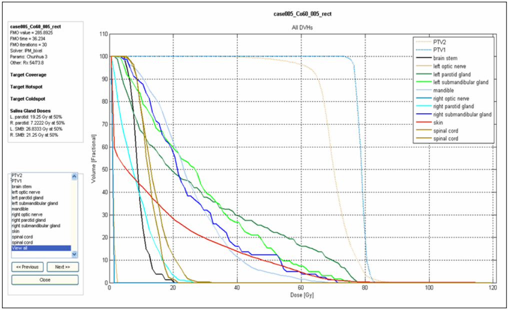
Radiotherapy goals

- The goal is to design a treatment plan that
  - delivers a prescribed dose to *targets*
  - while sparing, to the greatest extent possible, *critical structures*
- Radiation therapy therefore seeks to conform the geometric shape of the delivered *dose distribution* to the targets



Evaluation of a dose distribution

- A physician typically considers the dose distribution received by each individual structure
  - The *dose-volume histogram (DVH)* is an important tool that specifies, for each dose value, the fraction of a structure that receives at least that amount of dose



Evaluation of a dose distribution

- Rephrasing the goal of treatment plan design: we wish to identify a treatment plan that has
  - a "desirable" dose distribution in the targets
  - an "acceptable" dose distribution in the critical structures

- Representing a particular structure as a subset $\mathcal{V}$ of the patient, define a random variable D that represents the dose (rate) at a uniformly generated point in $\mathcal{V}$
  - Letting F_D denote the cumulative distribution function at that point, the DVH is simply the function $1 - F_D$

- This suggests an interesting connection between (financial) *risk management* and *treatment planning*
  - In both fields, we wish to control the shape of the probability distribution of one or more random variables

Criteria

- Broadly speaking, we wish to penalize
  - overdosing of both target and critical structure voxels
  - underdosing of target voxels
- Physical criteria are therefore often of the form

$$\mathcal{E}\left(u(D)\right)$$

  for an appropriately chosen function u

  - In the context of risk management the function u would be a utility function, its shape depending on risk preferences
  - In radiation therapy treatment planning the function u depends on the biological properties of the underlying structure being evaluated

Criteria for underdosing

- *Target*, underdosing
    - u decreasing, usually convex
    - when $u(d) = e^{-\alpha d}$ (with $\alpha > 0$) the expected utility is a monotone transformation of a measure of *tumor control probability* (TCP)

$$\mathrm{TCP} = \exp\left(-N\mathcal{E}\left(e^{-\alpha D}\right)\right)$$

- $N =$ number of clonogen cells in the target
- $\alpha =$ rate of cell kill per unit dose

Criteria for overdosing

- *Target or critical structure*, overdosing:
  - u increasing
  - shape depends on biological response of tissue to radiation
    - serial: high dose to a small fraction of the structure can destroy its functionality
    - parallel: sparing a part of the structure will preserve its functionality
  - u convex
    - when $u(d) = d^k$ (with $k \geq 1$) the expected utility is a monotone transformation of the so-called *equivalent uniform dose* (EUD)

$$\text{EUD} = \left(E(D^k)\right)^{1/k}$$

  - u "S-shaped"

Criteria

- Interesting special cases are
  - mean excess or mean shortfall criteria:

$$u(z) = \max\{0, z - T\} \quad \text{or} \quad u(z) = \max\{0, T - z\}$$

    - A related measure is the so-called *Conditional Value-at-Risk*
    - i.e., the upper or lower tail average of the dose distribution
  - *DVH-criteria* that evaluate points on the DVH

$$u(z) = 1_{[0,\delta]}(z) \quad \text{and } u_z = 1_{(\delta,\infty)}(z)$$

    - A related measure is the so-called *Value-at-Risk*
    - i.e., the dose level that is exceeded by (or not exceeded by) a given fraction of the structure

Optimization model: objective

- The dose distribution is evaluated over a discretization of the irradiated area into a finite set of cubes (*voxels*), V
- In particular, we consider a collection of treatment plan evaluation criteria:

$$G_1(z), \ldots, G_L(z)$$

  expressed as a function of the dose distribution, i.e., the vector of voxel doses $z \in \mathbb{R}^{|V|}$
  - where smaller values are preferred to larger values
  - Optimization models now typically contain a
    - single objective, by assigning appropriate weights to the different criteria
    - multi-criteria objective

Optimization model: variables

- Each beam is decomposed into a collection of *beamlets (bixels)*
  - We then optimize the intensity of each of these bixels, the *intensity profile*: x_i ($i \in N$)

- We can predetermine the dose distribution generated by each individual beamlet at unit intensity
  - The optimization model then becomes

$$\text{minimize } \{G_1(z), \ldots, G_L(z)\}$$

subject to

$$z_j = \sum_{i \in N} D_{ij} x_i \qquad j \in V$$

$$x_i \geq 0 \qquad i \in N$$

Decomposition of apertures

- An intensity profile for a given beam is *rounded* and expressed as the product of
  - a real (nonnegative) number
  - a (nonnegative) integer matrix
- We should then find a decomposition of this rounded intensity profile into
  - a number of apertures
  - with associated intensities

Representation of apertures

- Recall that
  - apertures are formed using pairs of leaves
  - a beam is discretized into a matrix of bixels
- Therefore, any aperture can be represented by a binary matrix having the *consecutive ones property*:
  - In either all *rows* or all *columns* of the matrix, the ones appear in consecutive elements
  - 0 represents that the corresponding bixel is covered while 1 represents that it is exposed

Decomposition of intensity profile

- *Example*:

  The intensity profile $\begin{bmatrix} 3 & 6 & 4 \\ 2 & 1 & 5 \end{bmatrix}$ can be decomposed as:

$$1 \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} + 2 \times \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} + 4 \times \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

  or, alternatively, as

$$1 \times \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} + 1 \times \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} +$$

$$2 \times \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} + 2 \times \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Objective: beam-on-time

- Minimize *beam-on-time*: sum of aperture intensities
- This problem is very well-studied and generally easy
  - The problem decomposes by beamlet row
  - For each row, the problem can be solved in linear time
  - See e.g. Bortfeld et al. (1994); Siochi (1999); Kamath et al. (2003); Ahuja & Hamacher (2005)

Objective: total treatment time

- Minimize *total treatment time*
  - This can be well-approximated by the sum of
    - beam-on-time
    - aperture setup time $\times$ number of apertures
  - If the actual beam-on-time is insignificant as compared to the setup time we can approximate the total treatment time by
    - number of apertures
- It may be desirable to constrain the beam-on-time to be minimal

Objective: total treatment time

- Minimizing the number of apertures (or total treatment time) is NP-hard (Baatar et al., 2005)
  - The problem *does not* decompose by beamlet row
  - Even for a single beamlet row the problem is NP-hard
- First attempt to solve the problem to optimality was using integer programming (e.g. Langer et al., 2001)
  - However, standard optimization solvers are unable to solve nontrivial problem instances in a reasonable amount of time
- Much of the research on this problem has therefore focused on heuristics
  - Xia & Verhey (1998); Siochi (1999–2007); Baatar et al. (2005); Engel (2005); Kalinowski (2005); Lim & Choi (2007)

Exact methods

- Recently, significant progress has been made on exact algorithms for solving different variants the problem
  - Dynamic programming
    - Kalinowski (2004)
  - Integer programming
    - Baatar et al. (2005); Mak (2007); Taşkın et al. (2009); Wake et al. (2009)
  - Constraint programming
    - Baatar et al. (2005); Ernst et al. (2009)
- It is now possible to solve problems of clinical dimension to optimality in reasonable time

Summary

- Optimization is an important tool in the design of radiation therapy treatments
- *Treatment plan optimization*:
  - Discussed ways to quantify the quality of a treatment plan
  - Established a link with risk management that provides an alternate motivation for the use of established criteria
- *Treatment plan delivery*:
  - Discussed a new algorithm for optimizing *total treatment time* that can solve clinical problem instances with reasonable computational effort

Future research

- The problem of efficiently finding a high-quality *static* treatment plan based on a *single image set* and assuming *stationarity* of the patient is relatively well-solved
- However, we need to better deal with *uncertainties* and *nonstationarities*
- Operations research can help by:
  - making optimal use of available technology
  - help assess which potential future technological enhancements are most valuable to the patient