# An Automated Personalized Health Coach

## A class experiment in health persuasive technologies

Chrysanne DiMarco

David R. Cheriton School of Computer Science
and
Department of English
University of Waterloo

January 22, 2010

# Outline



# The problem with most health information

- In healthcare situations, 'computerspeak' is a deterrent to people's willingness to accept computer-generated materials:
  - *You should always use sunscreen. You should always wear a hat.*
  - A little better: *You should always wear sunscreen and it's good to wear a hat as well.*
- Our aim: Assisting physicians to provide personalized information to patients about their health and treatments.
- Goals: improved patient satisfaction, outcomes; improved use of physician time and hospital resources.

# Why we need personalized health information (1)

- "One size fits all" generic health information either says too little or too much...
- ...and so isn't useful: patients don't get the advice they need.
- A brochure with only information relevant to **all** readers:
  - *Nine hundred different things could go wrong during your surgery.*
  - ...is too superficial and doesn't tell you what you want to know.

# Why we need personalized health information (2)

- A brochure with all information relevant to **any** reader:
  - Complications of the surgery could include pain, scarring, swelling, infection, and discoloration. Numbness or neurologic problems may occur. The exact nature and duration of problems may not be determinable and may be irreversible. Other complications may involve inflammation of a vein, cardiovascular problems, injury to surrounding tissue, bone fractures, delayed healing, allergic reactions to drugs or medications, even **DEATH**. In those cases where bone cement is used to secure implants in place, adverse reactions, including deep and surgical wound infection and a temporary lowering of blood pressure, may occur.
  - Too much information wont even be read!

# Experiment: Generating personalized anti-smoking letters

- Patients answered survey; later were sent personalized letter.
  - Letters were produced by combining snippets of text:
  - *You decided to quit smoking in order to make your clothes and breath smell better. By now everyone has probably noticed what a difference quitting has made. You should feel great about yourself now that your clothes and breath don't smell like cigarette smoke and other people will enjoy being around you more. Smelling better is only one of many good things that will happen now that you have quit.*
- Results: The recipients of personalized letters...
  - Read more of the document. And were three times more likely to quit smoking.

# But mass personalization is hard for people

- All possible combinations of text snippets must make sense.
- All possible personalized brochures must sound well-written.
- This quickly becomes much too hard for humans to manage:
  - In the experiment on anti-smoking, human patient educators had to manage **55,000** possible combinations of text snippets!
- Our research:
  - Automating the large-scale generation of finely tailored persuasive health education materials.

# The promise of computational rhetoric

- Most current Natural Language systems deal only with restricted language or use simplified methods of analysis, such as shallow parsing.
- While computational approaches to language have used the term "rhetoric", the deep resources of classical rhetoric have only been tapped to a small degree.
- This tradition studies three attributes of texts we can formalize and use:
  - Style (lexical choice, syntactic structure, modes of address);
  - Purpose (description, persuasion, instruction);
  - Affect (trust, deference, anger).

# How AI can help: "Generation-by-selection-and-repair"

- Our approach:
  - Author a repository (a tailorable "Master Document") ahead of time of reusable content variants.
  - Contains all info needed for any patient.
  - Content variants annotated according to patient features.
- Text snippets from Master Document selected then combined for a specific patient.
- But combined text might sound repetitive or awkward.
- So will automatically "repair" text to restore fluency.

# What a "repair engine" might do

- Suppose the following sentences are selected and turn up adjacent:
  1. **Non-insulin-dependent diabetes** is the most common type of diabetes.
  2. **Non-insulin-dependent diabetes** usually develops in adults over age forty, especially those who are overweight.
     - Highlighted NPs have coreference links to same object, so second occurrence is marked for pronominalization.
  3. **Non-insulin-dependent diabetes** is the most common type of diabetes. **It** usually develops in adults over age forty, especially those who are overweight.

# Potentially many kinds of repairs...

- Suppose the following sentences are selected:
  1. The condition that you have is **non-insulin-dependent diabetes**.
  2. **Non-insulin-dependent diabetes** is the most common type of diabetes.
     - Highlighted phrases indicate redundancies that can be repaired by combining—aggregating—the sentences:
  3. The condition you have is **non-insulin-dependent diabetes**, **which** is the most common type of diabetes.

# Aside: But maybe some repairs shouldn't be made

1. Smoking can lead to cancer.
2. Smoking can aggravate asthma.
3. Smoking can cost you financially.
   - A series of short repetitions might be more forceful than one long sentence...
4. Smoking can lead to cancer, aggravate asthma, and cost you financially.

# First we have to author the content (and this is hard)

- Author determines "workflow" of physician-patient encounter:
  - In surgical intervention, physician informs patient about each stage: pre-op, operation summary, post-op, complications, discharge plan, rehabilitation, and more.
  - Potentially many different patient "types" for each stage.
- Author considers purpose of information, attitudes to convey to patient, factors to vary on, then creates content variants:
  - In diabetes prevention, health "coach" encourages, empathizes, but aims for patient to be accountable for own wellness.
  - Dozens of possible rhetorical dimensions.
- Overall, potentially many thousands of different combinations of content.

# Authoring the Master Document: Levels of content management

- Knowledge level:
  - Knowledge engineer models concerns for each "stakeholder".
- Rhetorical level:
  - Linguist models the document rhetorical "workflow".
  - Physician-patient interaction of conversational "moves".
- Content level:
  - Health-educator writer "populates" Master Document with content variants for different combinations of patient features.
- And what about deep-linguistic level for complex text repairs?
  - Computational linguist.

# A model of information design for tailorable documents

- Goal: Multipurpose multi-user tailorable "Master" document.
- Physician starts with blocks of raw "medicalese" content.
- Rhetorician shapes content using models of argumentation and health persuasion.
- Authoring of a Master Document:
  - High-level document architecture: By the physician.
  - Rhetorical structure: By the health rhetorician.
  - Content variants and conditions: By the patient educator.
- What we needed:
  - Sophisticated authoring environment to assist providers.
- What we did:
  - Designed a text-editor–based authoring tool based on our paradigm of selection-and-repair and formal models of rhetoric.

# A generic health rhetoric model

- In wellness advisory intervention, health educator delivers sequence of conversational "moves", like acts in a play:
  - Establish coach role.
  - Evaluate patient progress.
  - Address barriers and give solutions.
  - Restate patient goals.

# Specific genre: Health coach interventions

- Analyzed (very) small corpus of personalized health-coach interventions.
- Defined rhetorical structure of health-coach genre:
  - Discourse components (**"part-genres"**).
    - Part-genre: Section of a full genre with own specific pattern of organization.
    - Sequence of part-genres provides high-level discourse structure for a Master Document architecture.
  - Rhetorical moves.
    - Characteristic pattern of conversational "moves" that make up the presentation of the information.
  - Rhetorical steps.
    - Components of moves.

# A sample human-authored health coach note

- Male/early 40s/265lbs.
- Goals: Lose 75–100lbs, keep up with kids, get healthy.

*Hi John,*
*You have had a busy week! Congratulations on a job well done. This week you continued your weight loss trend by adding another 2 lbs to the total. You are losing weight at a nice steady pace which has been shown to help maintain weight loss. I had a chance to review your food and exercise logs for last week. You have done a great job reducing your overall calories however there is still some room for improvement when it comes to decreasing your fat grams. Can you think of an alternative breakfast choice that is lower in fat and higher in protein? . . .*

# A sample health coach note (cont'd)

*. . . I want to applaud your efforts for diversifying your exercise choices. Walking around the block every day can get boring. The more variation you are able to include in your physical activity choices the more likely you will be to stick with a long term exercise plan. According to your estep pedometer you are walking about 5300 steps per day. This is a significant increase from your baseline activity but I want to challenge you to increase it to 6000.*

*Congratulations again on all your hard work. Your efforts are paying off. You will be out playing with your kids in no time.*

*To your health, <COACH>*

# Coach Note genre structure: From coaching goals to rhetorical goals

- Establish positive relation with client.
- Help set goals and monitor progress.
- Provide reinforcement and champion success.
- Plan for lapses, prevent relapses: Create strategies.
- Challenge negative thought and increase self-efficacy.
- Create accountability and maintain open contact for advice.

# From rhetorical goals to genre structure

| Component | Move | Step |
|---|---|---|
| Introduction (Frame) | Greeting<br>Connection | |
| Intervention (Coaching) | Monitoring<br>Reinforcement | <br>Appraise behaviour<br>Reinforce results, congratulate<br>Discourage negative thinking |
| | Action | Anticipate progress next week<br>Identify barriers<br>Create strategic steps |
| | Goal-setting | Restate overall goals<br>Encourage progress/self-efficacy |
| Conclusion (Frame) | Connection<br><br>Greeting | Open contact for next week<br>Reinforce positive relation |

# The HealthDoc toolkit: People, software, and knowledge

- Students.
- Linguists, rhetoricians, health educators, software developers.
- A knowledge representation for tailorable multimedia documents.
- A text-based authoring tool.
- A tailoring engine.

# The class experiment in a nutshell

1. Class of third-year rhetoric students in Information Design.
2. "How to be a health coach" in one lesson.
3. Explain health coach rhetorical model.
4. Provide simulated patients.
5. Iterate:
   - Write, write, write tailored coach notes.
   - Evolve the patients.
6. Merge class notes into one big Coach Note Master Document.
7. Turn rhetoricians loose.
8. Get developers to make it all work.
9. Here's what we did:
   - http://healthdoc.cs.uwaterloo.ca

# Conclusions: Automating generation of personalized text

- "Text-to-text" generation by pre-authoring and reuse of text avoids hard problems of generation from scratch.
- But fine-grained personalization requires linguistically rich content embodying models of rhetoric.
- Fully automated solution will require solving open problems in intelligent authoring tools, computational rhetoric, automated text (and multimedia) revision and repair.
- What's next:
  - Which factors matter for tailoring?
  - Can we develop a process for authoring tailorable content?
  - Does the medium of delivery matter?

# The HealthDoc Project (1994–present)

- Academic:
  - Chrysanne DiMarco (Computer Science and English).
  - Eduard Hovy (Information Sciences Institute/USC).
  - Neil Randall, Randy Harris (Waterloo, English)
  - David Wiljer (Princess Margaret Cancer Centre, Toronto).
- Industry:
  - Desire2Learn (learning management systems).
  - DPS Health (chronic disease management systems).
  - Nortel Networks (telecommunications).