Generating monologue and dialogue to present personalised medical information to patients

Conference or Workshop Item

Generating monologue and dialogue to present personalised medical information to patients

Sandra Williams, Paul Piwek and Richard Power
Department of Computing Science
The Open University
Walton Hall, Milton Keynes, MK7 6AA, U.K.
{s.h.williams, p.piwek, r.power}@open.ac.uk

Abstract

Medical information is notoriously difficult to convey to patients because the content is complex, emotionally sensitive, and hard to explain without recourse to technical terms. We describe a pilot system for communicating the contents of electronic health records (EHRs) to patients. It generates two alternative presentations, which we have compared in a preliminary evaluation study: the first takes the form of a monologue, which elaborates the information taken from the patient's EHR by adding explanations of some concepts and procedures; the second takes the form of a scripted dialogue, in which the content is recast as a series of questions, answers and statements assigned to two characters in the dialogue, a senior and a junior nurse. Our discourse planning method designs these presentations in tandem, first producing a monologue plan which is then elaborated into a dialogue plan.

1 Introduction

Increasingly, health service providers are storing patient medical histories in machine-usable form – i.e., in databases of Electronic Health Records (EHRs) – and are obliged by legislation to make this information accessible to patients. We explore here an application in which material from EHRs is selected and organised for presentation to patients either as a monologue or as a script for a dialogue. Emphatically these are not intended to replace face-to-face consultation with a specialist. Rather, the aim is to help patients use consultations to better effect.

We have constructed a natural language generation (NLG) system that generates monologue or dialogue scripts which are spoken by synthetic voices. If a patient were to listen to a review of her recent record (but not, of course, containing any new diagnoses of which she was previously unaware) it could potentially help her in her next consultation in a number of ways:

- by reminding her of her own case history;
- by giving her practical examples of the meaning and usage of medical terms relating to her case;
- by demonstrating how to ask practical medical questions relating to her case (dialogues only).

We generate dialogue because it opens up new ways of making rhetorical and argumentative information explicit (Piwek et al., 2005). Also, as Craig et al. (2000) and Cox et al. (1999) found, it would help the patient recall the current state of her treatment, it would encourage her to ask questions during the subsequent consultation by reminding her of the exact meaning of technical terms, and would give her more confidence in expressing herself both in technical language and informal language. It could also save doctors' time by providing an additional source of explanation for patients.

The CLEF system (Hallett and Scott, 2005) summarises EHRs of breast cancer patients for *clinical staff*. We have implemented a pilot system for *patients* that selects and describes events from the same EHR repository. At present the wording of utterances is somewhat stilted, since our initial focus is on designing a discourse planner for both monologue and dialogue. Below is an extract from a monologue generated by the current system.

> On November 27th you had an examination. The doctor found that you had lymphadenopathy in your right axillary lymphnodes. Lymphadenopathy is swelling of the lymph nodes which the doctor can feel when you are examined. The commonest cause of lymphadenopathy is infection. Lympadenopathy can also be caused by a build up of cancer cells within the lymph nodes. Axillary lymphnodes are rounded masses of tissue under the arm containing white blood cells.

A dialogue script generated from the same data follows. The characters are two nurses, one junior and one senior, where the junior is an eager but not very knowledgeable character while the senior is an experienced, trustworthy instructor whose words carry weight. The junior reads from a report which contains some terms that she does not understand; the senior nurse explains them.

Junior Nurse: On November 27th she had an examination.

Senior Nurse: What did the examination find?

Junior Nurse: The doctor found that she had lymphadenopathy. What is lymphadenopathy?

Senior Nurse: Lymphadenopathy is swelling of the lymph nodes which the doctor can feel when you are examined. The commonest cause of lymphadenopathy is infection. Lympadenopathy can also be caused by a build up of cancer cells within the lymph nodes.

Junior Nurse: The lymphadenopathy was in her right axillary lymphnodes. What are axillary lymphnodes?

Senior Nurse: Axillary lymphnodes are rounded masses of tissue under the arm containing white blood cells.

We have piloted the system with colleagues to find out whether monologue and dialogue would be successful at conveying information in the EHR. Many useful ideas for improving the system emerged, as can be seen in Section 4.

2 Generating monologues and dialogues

2.1 Input

The system's input comes from a relational database representing medical histories of simulated breast cancer patients. The patient simulator was developed by Rogers et al. (2006) for the CLEF project. Database entries describe medical interventions, investigations, problems, drugs and patients. Primary keys provide relational links between database tables in the usual manner; additionally, a *Relations* table specifies explicit rhetorical relations – e.g., the evidential relation *HAS_FINDING* which indicates that an investigation provided evidence for diagnosis of a medical problem. Twenty types of relation are present in the *Relations* table, although our current prototype only describes a subset of them.

2.2 Discourse planning

Document planning has two or four stages depending on whether monologue or dialogue output is required. First, we select content and arrange it into a discourse structure. Second, we add explanation relations. At this stage, the discourse structure is complete for monologue generation and the final task is to recast it into a series of sentence templates. Discourse planning for dialogue output has two further stages: the third adds questions and answers, and the last allocates portions of the discourse structure to each conversational partner and recasts them into a series of utterance templates from which dialogue turns are to be generated. These stages are described in more detail below.

Content selection (stage 1a) queries the database for records on a particular simulated patient. These are scanned for recent medical investigations and interventions that are linked to medical problems via explicit evidential and motivational relations present in the *Relations* table. Selection of EHR data is the same for dialogue and monologue.

Initial discourse structures for medical episodes are built from the selected content (stage 1b). Fig. 1 (A) shows a *medical episode* where *Investigation* and *Problem* concepts are linked to each other by a *HAS_FINDING* relation and to a *Locus* concept by *HAS_TARGET* and *HAS_LOCUS* relations. The nodes in the diagram represent instances of concepts, while the arcs represent relations. Instances generally represent records found in database tables; here they are records from the *Investigation*, *Problem* and *Locus* tables; the nodes contain all the fields that were present in the record, e.g. Fig. 1 (A) shows the "Name" fields "examination", "lymphadenopathy" and "axillary lymphnodes". The arcs in the diagram represents relations present in the *Relations* table mentioned above. *Medical episodes* like the one in Fig. 1 (A) are ordered sequentially according to the date fields of the retrieved data.

Explanations are added (stage 2) as shown in Fig. 1 (B). The program consults a table of term definitions and introduces *EXPLANATION* relations linking instances of concepts to glosses that explain these concepts. At present, we do not have an automatic procedure for deciding which concepts require explanations; concepts like "examination" are left unexplained because they seem intuitively well-known, while less familiar concepts like "axilla" are explained. A search in the BNC confirmed that raw BNC frequencies correspond well with our intuitions: e.g., "examination" has high frequency, whereas "axilla" has low frequency. Our technical term definition glossary currently has 73 entries comprising canned phrases and sentences from trusted sites such as Cancer Research UK patient information pages (www.cancerresearchuk.org). Where appropriate these have been simplified slightly and their tenses have been changed.

Questions and Answers are added (stage 3) through the relations *RAISES_Q* (i.e., a concept raises a question) and *HAS_ANS* (i.e., a given question has an answer), as shown in Fig. 1 (C). Two question types are illustrated: firstly "What is it?" questions, which are subgraphs spanning *EXPLANATION* relations; and secondly a "What was found?" question asking about a specific argument in a specific relation (in this case about the finding of the *HAS_FINDING* relation, i.e., the *Problem*

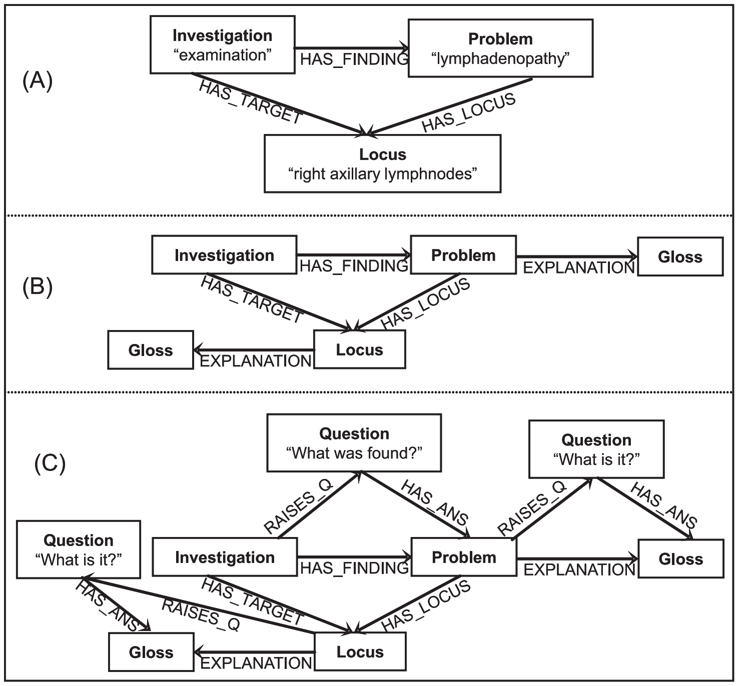

Figure 1: Stages in discourse planning.

that the evidence suggested). Note that questions ask about both concepts and relations.

Assignment to dialogue turns (stage 4) is achieved by mapping parts of the discourse structure to utterance specifications (i.e. questions, answers and statements) and by using dialogue templates to assign these to the dialogue partners. Below is a list of utterance specifications generated from Fig. 1 (C). The resulting dialogue script is given in section 1 where, for example, “On November 27th she had an examination” realises the utterance specification *inform(Investigation)*.

Speaker	**Specification**
Junior	inform(Investigation)
Senior	question(HAS_FINDING)
Junior	answer(Problem)
Junior	question(Problem.Name)
Senior	answer(Gloss)
Junior	inform(HAS_LOCUS)
Junior	question(Locus.Name)
Senior	answer(Gloss)

Note that a single turn can contain several moves and that both characters ask questions, thus avoiding ‘conversational ping-pong’ (Davis, 1998) — i.e., exchanges where turns degenerate into nothing more than a sequence of alternating questions and answers.

2.3 Microplanning, realisation and output

These are achieved through simple string processing with reference to a discourse history which maintains a list of entities mentioned so far, prohibits questions and explanations being generated more than once, and allows determiners such as “another” to be generated when an action is performed repeatedly. The output is formatted as web pages with embedded Active X commands to control text-to-speech. Examples of the system’s output were given in section 1.

3 Pilot evaluation

As a preliminary test of our system, we piloted the effectiveness of communicating medical information through scripted dialogue and monologue by conducting a small evaluation with eleven colleagues from The Open University, divided randomly into two groups, Group A (five people) and Group B (six people).

Materials: We generated a monologue and dialogue from the same simulated patient’s EHR. Loquendo text-to-speech voice Kate read the monologue, while in the dialogue Simon read the junior nurse and Kate the senior. Section 1 illustrates fragments of the materials that were actually used in the pilot. A practice monologue was also produced by hand, with both voices reading alternate paragraphs from a cancer patient web site.

Method: Both groups listened to the practice monologue and answered questions about the voices and a practice multiple-choice comprehension question. Group A then listened to the monologue and answered multiple-choice questions. Afterwards, this group listened to the dialogue, stated their preferences, and were asked for comments. Group B listened to the dialogue first, answered questions, listened to the monologue, stated their preferences, and added comments.

Results: Participants' comments were very informative. People particularly did not like pronunciations of technical terms and thought that the synthetic voices spoke too fast; these are problems that can be addressed by configuring the text-to-speech system. Furthermore, the general impression was that too much information was given too quickly, although one person commented that the dialogue was better because information was presented in shorter "chunks".

Both groups performed well on comprehension (their mean scores out of 10 were 8.6 for Group A and 7.5 for Group B), and an independent samples t-test showed no significant difference between them. A few people told us that they already knew some of the medical concepts queried by the comprehension questions; however, performance on questions for which the answers could not have been known in advance was also slightly (but not significantly) higher in Group A (monologue).

A Pearson chi-square test showed no difference between groups' monologue/dialogue preferences. Despite that, none of the participants were neutral about the choice — i.e., none selected the option "I have no preference".

Discussion: The most important outcome is that the system ought to communicate information at a slower rate and present it in smaller chunks. This could be achieved by adding conversational padding that would enable the density of new information to be finely tuned. Piwek and van Deemter (forthcoming) have proposed a possible solution to this problem.

4 Conclusion

Patients accessing their EHRs have very different needs from doctors: configuring a suitable presentation is not merely a question of replacing technical language. Doctors have to treat many patients every day, under severe time pressure; what they want is a condensed version of the facts relating to each individual case. Patients have plenty of time to view a single presentation of intense personal interest, and so are likely to prefer a gentler information flow with informal touches (i.e., dialogue rather than monologue) and some easily assimilated instruction. We have not tried to eliminate technical terms altogether, but to give patients some mastery of technical concepts that are highly relevant to their case. This is an original communication genre with subtle aims: the patient learns by viewing a conversation in which she is not addressed directly.

Technically, the planning of the patient communications is achieved through a process in which an initial plan, similar to that for generating condensed reports for doctors (Hallett and Scott, 2005), is progressively *overlaid* with further rhetorical and semantic content — first, by adding explanations, secondly by recasting assertions as conversational response pairs, and thirdly by grouping moves into turns. Our pilot evaluation suggests that the resulting dialogues are still too dense: they lack the reassurance that comes from a lighter, more discursive style, and some repetition of information that is *already familiar* — the equivalent of small talk. There are obvious more superficial ways in which our demonstrator could be improved (e.g., using more pronouns instead of ponderous repetition of nouns like 'lymphadenopathy'), but the crucial challenge lies in pushing our approach to discourse planning much further, so that the content from the EHR is folded into a plan that meets the special needs of this new genre.

References

Cox, R., McKendree, J., Tobin, R., Lee, J. and Mayes, T. (1999) Vicarious learning from dialogue and discourse: a controlled comparison. *Instructional Science* 27: 431-458.

Craig, S D., Gholson, B., Ventura, M., Graesser, A (2000) Overhearing Dialogues and Monologues in Virtual Tutoring Sessions: Effects on Questioning and Vicarious Learning. *International Journal of Artificial Intelligence in Education*, 11, 242-25.

Davis, R. (1998) *Writing Dialogue for Scripts*, A & C Black Ltd., London, ISBN 0-7136-4802-3.

Hallett, C., and Scott, D. (2005). Structural variation in generated health reports. *Proceedings of the 3rd International Workshop on Paraphrasing*.

Piwek, P., Power, R., Scott, D., and van Deemter, K. (2005) Generating Multimedia Presentations: From Plain Text to Screen Play. In O. Stock and M. Zancanaro, eds., *Multimodal Intelligent Information Presentation*. Springer.

Piwek, P., van Deemter, K. (forthcoming). Generating under Global Constraints: the Case of Scripted Dialogue. *Research on Language and Computation*. Kluwer.

Rogers, J., Puleston, C., Rector, A. (2006) The CLEF Chronicle: Patient Histories derived from Electronic Health Records. *Proc. of the 22nd International Conference on Data Engineering Workshops*, IEEE.