In Claude Frasson, Gilles Gauthier and Alan Lesgold, eds., *Intelligent Tutoring Systems: Third International Conference (ITS '96), Proceedings*, Montreal. (Springer-Verlag Lecture Notes in Computer Science, no. 1086.) Berlin: Springer-Verlag, 1996, pp. 632-640.

Generating and Revising Hierarchical Multi-turn Text Plans in an ITS

Reva Freedman
Department of EECS
Northwestern University
freedman@delta.eecs.nwu.edu

Martha W. Evens
Department of CSAM
Illinois Institute of Technology
mwe@schur.math.nwu.edu

Abstract. CIRCSIM-Tutor v. 3 is a natural-language based ITS for cardiac physiology. In this paper, we describe TIPS, a new text planning engine for CIRCSIM-Tutor based on current research in text generation. Since conversations cannot be completely planned in advance, TIPS plans and executes iteratively. It maintains a goal hierarchy for the tutor while carrying on a conversation with the student. It can handle multi-turn plans on the part of the tutor, and it can back up and replan when the student gives an unexpected answer. In this paper we sketch the design of TIPS using an analysis of human-to-human tutoring transcripts to shape the requirements.

1 Introduction

The negative feedback loop which maintains a steady blood pressure in the human body is one of the more difficult topics for first-year medical students to master. CIRCSIM-Tutor v. 3 is the latest in a series of CAI systems intended to help students master the concepts involved. This paper describes TIPS, a new text planner for CIRCSIM-Tutor v. 3. TIPS stands for "Text generation Interactively, a Planning System". We chose this acronym because of the pun it gives rise to: after all, one of the main goals of CIRCSIM-Tutor is to generate verbal hints to help the student, i.e. tips.

TIPS views responding to the student in an ITS as a text planning problem. It searches a library of plan operators to determine possible responses. The text controlled by plan operators ranges in size from the whole conversation down to primitive speech acts. Constraints on the plan operators can be used to take other knowledge sources into account, such as the domain knowledge base, the dialogue history, and the student model.

We start by analyzing previously collected human-to-human tutoring sessions from a text planning perspective. We characterize the aspects of the tutor's productions which are the most significant to model in a computer-based tutoring system. Then

---

This research was conducted at the Illinois Institute of Technology as part of the CIRCSIM-Tutor project. It was supported by the Cognitive Science Program, Office of Naval Research under Grant No. N00014-94-1-0338, to Illinois Institute of Technology. The content does not reflect the position or policy of the government and no official endorsement should be inferred.

we sketch the design of the new planner and demonstrate that it satisfies the design goals. Finally, we describe the relationship between TIPS and other text planning systems which handle dialogues.

2 A Look at Some Naturalistic Data

2.1 Background

Over the last four years, the CIRCSIM-Tutor project has collected over 75 hours of transcripts of human-to-human tutoring sessions in order to model the pedagogical and linguistic strategies needed for CIRCSIM-Tutor. The sessions were conducted as keyboard-to-keyboard sessions instead of verbally in order to collect text samples which were as close as possible to potential CIRCSIM-Tutor input and output. Our domain experts, who are highly experienced professors of physiology, served as tutors. The students were volunteer medical students.

In their beginning physiology course, the students are given a simplified qualitative model of the heart. Then they are given problems to solve. In each problem, something happens to change the processing of the heart. The student is then asked to predict the direction of change of seven core variables at three points in time: the DR or *direct response* phase immediately after the disruption, the RR or *reflex response* phase, which shows the effect of the negative feedback loop controlled by the nervous system, and the SS or *steady state* phase after a new steady state has emerged.

In the experimental tutoring sessions, the students solve the same type of problem. After each phase, the tutor conducts a dialogue with the student. The tutor discusses each incorrect variable with the student until both are satisfied that the student understands.

2.2 Major Features of the Transcripts with Respect to Planning

Previous studies of the CIRCSIM-Tutor transcripts have concentrated on the syntactic structure of our human tutors' productions [Seu et al. 1991] and their pedagogical goals [Hume et al. 1996, Evens et al. 1993]. We reanalyzed the corpus of texts using a text planning model in order to determine which aspects of the human tutors' productions were the most significant in terms of text planning. The following points summarize our major conclusions.

1) The tutor maintains a hierarchy of goals, visible via the hierarchical structure of the resulting conversation, while at the same time carrying on a conversation with an agent whose responses can't be predicted in advance.

2) The tutor can teach the same thing in more than one way. If the student doesn't understand one method, the tutor has several alternatives:

a) Explain the concept, then continue.
b) Use a nested method to go into more detail.
c) Drop that attempt and try a different method.

3) The tutor sometimes attempts to correct the student's understanding using a multi-turn process. If the student doesn't respond as expected, the tutor can shift smoothly to a different explanation.

4) The tutor can say the same thing in more than one fashion.

2.3 Conversation as a Hierarchical Structure

The high-level organization of the collected dialogues can be succinctly described as follows:

- Within each stage, the text is divided into segments, one for each incorrect core variable.
- The student must explicitly state the correct value of each variable before the tutor proceeds to the next one.
- The variables are discussed in a partially ordered sequence which corresponds to the solution trace of the problem.

Within the basic outline, tutors can follow an idiosyncratic pattern, depending on the student's initial skill at solving the problems, the student's ability to respond to the tutor's help, the tutor's personal style, and a certain amount of free variation. Thus the tutoring session cannot be implemented as a simple algorithmic process. Instead, we have identified twenty or so "correction mechanisms" in the transcripts. These discourse mechanisms form the basis of the TIPS planner. As the planner runs, it produces a hierarchically organized discourse with correction attempts nested inside variables for each physiological stage.

2.4 Making Multiple Attempts To Teach The Same Concept

The correction of each variable is divided into attempts. Once the student gets the right answer for the value of the variable, the tutor goes on to the next variable:

Correct a variable:
  Introduce variable (explicitly or implicitly)
  Make first attempt at correction
  Make another attempt at correction
  …

The following excerpt from the transcripts shows an incorrect prediction by the student followed by two attempts at correction. Each attempt ends with the tutor asking for the value of the variable to see if the student understands yet. After the second attempt, the student gives the correct answer.

(1) S: CC increases.
T: Yes, that's the effect of increased sympathetic stimulation on the myocardium. However, what happens to CC in the DR period?
S: CC increases.
T: Reminder, the DR occurs before there are any reflex (neural) changes. What happens to CC in the DR?
S: CC remains constant.

2.5 Use of Multi-Turn Processes

Some of the tutors' most frequently used tutorial mechanisms are recurrent two- and three-turn sequences. In the following excerpt, the tutor points out a contradiction to the student. The variables MAP, CO and TPR are connected by the equation MAP = CO ∗ TPR.

(2) T: … What are the determinants of MAP?
S: CO and TPR.
T: Correct. And you have predicted CO increases and TPR increases. So how can you say MAP decreases?

If the student had not "played along" with the tutor by correctly answering the first question, the tutor would have been unable to continue as planned. In fact, that's exactly what happens in the next example:

(3) T: Now look at your predictions: MAP decreases, TPR increases, CO doesn't change. Is that possible?
S: Yes.
T: Let's try again. MAP = CO ∗ TPR. CO doesn't change. TPR increases…

In this example, the tutor recovers by trying the same plan again but wording it differently. The cue phrase "let's try again" is used to mark a repetition. The tutor could also have chosen any of the alternatives in Section 2.2 for the second attempt. We call these multi-turn plans *interactive explanations* or *directed lines of reasoning*.

2.6 Conversation as a Process Controlled by Two Agents

In the preceding sections, we have described the correction of the core variables as a hierarchical process. Within each attempt, an individual goal, for example the tutor's desire to get the student to state a particular fact, might engender subprocesses which could themselves spawn multiple attempts. At the same time the tutor must maintain a dialogue. In the typical dialogue pattern, predicted by the Conversation Analysis school [Sinclair & Coulthard 1975, Stenström 1994] and observed in our transcripts, every turn has the following basic structure:

Turn:
- Response to student's previous statement (optional)
- New material
- Question for the student (or an imperative)

The response can take many forms. Here is a sampling:

- A response with no physiology content
  - A simple *yes* or *no*
  - A more extensive acknowledgment, such as "Yes, you're right"
- A restatement of the student's last statement
  - Using the student's words: "Yes, <statement>"
  - Using more precise or otherwise preferred language
- A denial of the student's last statement or some aspect of it
  - "No, <contradictory statement>"

In the transcripts the closing question is sometimes left implicit. For example, if the tutor gives the student a hint, the student will usually spontaneously make another attempt to solve the problem. CIRCSIM-Tutor always terminates its turns with a question because people expect computers to be explicit about their requests for data.

The response at the beginning of the tutor's turn usually closes out a tutoring goal, and the new material which follows is generated by a different goal. Even though the parts of a turn are derived from unrelated goals, they must still fit together to form cohesive text. Furthermore, a single sentence may contain material from multiple tutorial goals. For these reasons, we treat turn planning as a parallel process which fashions turns from material provided by the tutorial planner. We accumulate tutorial goals for a turn and pass them to the turn planner as a unit.

2.7 Multiple Ways of Saying the Same Thing

In addition to using a variety of argumentative forms to tutor each concept, the tutor can also express the communicative acts which comprise the forms in more than one way. The sentences below show several locutions which the human tutors use to start the dialogue. Although each of these sentences has a slightly different semantic range, they serve the same purpose within CIRCSIM-Tutor. In addition to the pedagogical value of variety, multiple realizations may be necessary in order to have one which fits into the turn being planned.

(4) T: … Now let's review some of your predictions…
(5) T: … Let's take a look at some of your predictions…
(6) T: … let's talk about your predictions…
(7) T: … There are some errors here. Let's start with this issue…

3 Design of the TIPS Planner

3.1 Requirements for the Planner

From the discussion above, we can deduce the following criteria which we want our planner to satisfy. The first four goals deal with the tutor's ability to make and carry out plans while maintaining a dialogue with the student. The final two goals are concerned with the expressive power of the tutor.

1) The tutor must be able to handle interspersed planning and execution, since we can't and don't want to plan the whole conversation in advance.

A game-playing program is a good analogy. A chess program may employ multi-turn strategies, but it cannot predict its opponent's response to intermediate steps of those strategies. Additionally, since we may want to revise the plan later on, we do not want to waste time working out details of later plan steps which may not be needed.

2) The tutor must be able to maintain the tutor's goal hierarchy while carrying on a natural-sounding conversation with the student.

3) The tutor must be able to back up and replan when the student gives a wrong or unexpected answer.

4) The tutor must be able to handle multi-turn plans.

The final two goals ensure that the tutor can generate text for concepts in a variety of ways both at the pedagogical and linguistic levels. Variety at the pedagogical level is important if the tutor is to be able to help the greatest number of students. At the linguistic level, variety is important to ensure that students continue to read the tutor's output and do not resort to solving the problems by rote.

5) The tutor must be able to teach the same concept in multiple ways.

6) The tutor must have the ability to say the same thing in more than one fashion.

3.2 Syntax of the Plan Operators

The TIPS planner is an enhanced "classical" (i.e. STRIPS-style) planner. The plan operators have four fields:

- Name
- Effects
- Constraints
- Decomposition

The intended interpretation of the plan operator is as follows:

1) The *name* field identifies the communicative act which the rule is intended to implement.

2) The *effects* field identifies any state changes which can be assumed to be true as a result of the operator.

3) The *constraints* field refers to *a priori* preconditions which must be true before the operator will be considered.

4) The *decomposition* field describes how to implement the communicative act represented by the *name*, or equivalently how to obtain the state described by the *effects*, in terms of lower-level operators or primitive speech acts. If it contains more than one entry, the entries must be satisfied sequentially. We also have a notation which means "as many times as possible" so that indefinite repetition, such as iterating over each incorrect variable, can be expressed in a simple fashion.

A plan operator may be selected for use if its name (including the arguments) or its effects can be matched against the current entry in the decomposition of the plan operator currently being expanded. The *effects* field permits the specification of the result of an action instead of specifying how the action is to be carried out.

The planning literature contains three kinds of preconditions for plan operators: those which must be true *a priori* (the planner will not try to make them true), those which the planner will try to satisfy if they are not already true, and those which the planner must instantiate even if they are already true. The TIPS planner can handle all three types of preconditions.

We use the *constraints* field of the plan operator to implement the first type of precondition, which is useful for reducing the search space. The second and third types of preconditions both occur in the decomposition field of the plan operator. Since we allow plan operators to match either on the operator name or on the effects, we adopt the convention that entries which contain the name of an operator must be implemented, while entries stated in terms of a desired effect are skipped if the effect is already true. The intention is that if the operator calls for a specific communicative act on the part of the tutor, the tutor must perform that act. If the operator simply requests that a state be true, no action need be taken if the state is already true.

4 Relationship to Previous Work

4.1 Explanation Systems

Other large-scale interactive generation systems have been implemented by Cawsey [1992], Maybury [1992], and Moore [1995]. Of these, Cawsey's EDGE system is the closest in spirit to the work described here, as it teaches concepts from a causal model and embeds the content in a naturally structured conversation. However, EDGE always considers an explanation as a unit. In contrast, CIRCSIM-Tutor needs the ability to drop an incomplete tutoring plan if the student cannot answer intermediate questions.

Maybury's TEXPLAN system illustrates the ability to have multiple ways to explain a concept, including several types of definitions and descriptions. One aspect of TEXPLAN which resembles our system is that plan operators may match on either the *header* field, which identifies the speaker's communicative act, or the *effects* field, which contains the intended effect on the hearer. Moore's PEA system focuses on the intended effect on the student of the tutor's statements. The goal of Moore's work is to be able to ask more intelligent follow-up questions by understanding the meaning of the student's response in context. CIRCSIM-Tutor obtains similar knowledge about the student's understanding by asking further questions.

As each of these systems is primarily an explanation system which accepts follow-up questions rather than a tutoring system, none of them needs to generate the variety of tutoring mechanisms which CIRCSIM-Tutor requires. While explanation is one of the methods which CIRCSIM-Tutor uses to enhance the student's understanding, it is not a major focus of our work. In general, we only give an explanation after the failure of a method such as hinting, which requires more active participation on the part of the student. Furthermore, in each of these systems, the dialogue terminates when the student is satisfied. In contrast, CIRCSIM-Tutor does not terminate the dialogue until the student has answered every question correctly.

4.2 Planning Systems

Wilkins' SIPE system [1988] is an advanced classical planning system used for activities such as robot planning. Wilkins' work is relevant to ours because it is one of the few planning systems using the classical model which can cope with unexpected events happening during plan execution. For example, the robot might arrive at a room and find the door locked. Wilkins presents a taxonomy of actions which the plan execution module can attempt in order to correct the plan, such as grafting new steps onto a plan in progress.

Text planning differs from robot planning is that the tutor cannot "unsay" anything. If the robot drops a screw, it can pick up another screw and try again with few or no untoward consequences. If the student does not understand an idea and the tutor decides to try again, the second attempt must be different in some way, and the first attempt remains in the conversation.

5 Conclusions

This paper describes a new planner called TIPS which will be used to generate natural language dialogue for v. 3 of CIRCSIM-Tutor, and could be used by any similar dialogue-based ITS. TIPS views the generation of a response as primarily a text planning problem. We have shown that TIPS can generate discourse structures similar to those used by expert human tutors. We expect that CIRCSIM-Tutor v. 3 will generate text which is more natural, more varied and more cohesive than previous ITSs, and thus contribute to greater student understanding of cardiac physiology.

6 Acknowledgments

This work could not have been completed without our co-investigators and domain experts, Professors Allen A. Rovick and Joel A. Michael of Rush Medical College. Reva Freedman would also like to thank Professor Gilbert K. Krulee of Northwestern University for his support.

7 References

Cawsey, Alison. *Explanation and Interaction: The Computer Generation of Explanatory Dialogues.* Cambridge, MA: MIT Press, 1992.

Evens, Martha W., John Spitkovsky, Patrick Boyle, Joel A. Michael and Allen A. Rovick. "Synthesizing Tutorial Dialogues," *Proceedings of the 15th Annual Conference of the Cognitive Science Society*, Boulder, 1993. Hillsdale, NJ: Lawrence Erlbaum, 1993.

Hume, Gregory D., Joel A. Michael, Allen A. Rovick and Martha W. Evens. "Hinting as a Tactic in One-on-One Tutoring," *Journal of the Learning Sciences*, 5(1): 32–47, 1996.

Maybury, Mark T. "Communicative Acts for Explanation Generation," *International Journal of Man-Machine Studies* 37(2): 135–172, 1992.

Moore, Johanna D. *Participating in Explanatory Dialogues: Interpreting and Responding to Questions in Context.* Cambridge, MA: MIT Press, 1995.

Seu, Jai, Ru-Charn Chang, Jun Li, Martha W. Evens, Joel A. Michael and Allen A. Rovick. "Language Differences in Face-to-Face and Keyboard-to-Keyboard Tutoring Sessions," *Proceedings of the 13th Annual Conference of the Cognitive Science Society*, Chicago, 1991. Hillsdale, NJ: Lawrence Erlbaum, 1991.

Sinclair, John M. and Richard M. Coulthard. *Towards an Analysis of Discourse: The English Used by Teachers and Pupils.* London: Oxford University Press, 1975.

Stenström, Anna-Brita. *An Introduction to Spoken Interaction.* London: Longman, 1994.

Wilkins, David E. *Practical Planning: Extending the Classical AI Planning Paradigm.* San Mateo, CA: Morgan Kaufmann, 1988.