Proceedings of the 11th International Florida Artificial Intelligence Research Symposium, Sanibel Island, 1998, pp. 153-157.

# Responding to Unexpected Student Utterances in CIRCSIM-Tutor v. 3: Analysis of Transcripts

Jung Hee Kim, Reva Freedman†, and Martha W. Evens
Department of CSAM
Illinois Institute of Technology
10 W. 31st Street 236–SB
Chicago, IL 60616
{janice, freedman}@steve.csam.iit.edu, csevens@minna.cns.iit.edu

**Abstract**

CIRCSIM-Tutor is a dialogue-based intelligent tutoring system that conducts dialogues with medical students about blood pressure regulation. To obtain models for computer-generated dialogues, we analyzed dialogues involving expert human tutors. In this paper we describe some of the interesting and complex patterns we isolated from the human tutorial dialogues in cases where the student gave erroneous or otherwise unexpected results.

## Introduction

CIRCSIM-Tutor is an intelligent tutoring system for cardiovascular physiology designed to help first-year medical students solve problems involving the negative feedback system that controls blood pressure. Students are asked to predict the direction of change of seven core variables, the most important cardiovascular parameters. The tutor analyzes these predictions, finds the errors, and chooses a method to correct each one. The tutor then embarks on a remedial dialogue.

To understand the behavior of human tutors and to produce a detailed knowledge representation for the tutorial planner in the next version of CIRCSIM-Tutor v. 3, we started with human tutoring dialogues. We have approximately fifty transcripts of professors of physiology tutoring medical students, with the tutor and student in different rooms communicating keyboard-to-keyboard. Out of the more than 5000 turns we have collected, we have analyzed over 270 turns of dialogue. Using an SGML-based annotation system, we marked up the tutorial goal structure in the transcripts to use as a basis for plan-based text generation (Freedman and Evens, 1996). Our analysis produces multiple nested annotations showing both global goals for tutoring and local goals for maintaining a coherent conversation and providing appropriate responses to the student.

In this paper we show how this method of analysis can be used to model the complex dialogues generated by expert tutors when faced with unexpected student utterances.

This work 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.

† Effective 1/1/98 Reva Freedman has moved to the Learning Research and Development Center of the University of Pittsburgh.



## Methods of Tutoring

CIRCSIM-Tutor v. 3 requires a set of tutorial and conversational goal schemata in order to produce coherent conversations. The analysis in this paper is an extension of the one introduced by Freedman (1996). It is based on approximately 350 instances of global tutoring goals and 50 instances of local goals.

Tutorial goals are expanded in a hierarchy. Two sections of dialogue are generated for each variable that the student did not predict correctly. ***T-introduces-variable*** introduces the variable as a referent in the conversation and ***T-tutors-variable*** does the actual tutoring. For example, a dialogue for correcting RAP might start like this:

**<T-corrects-variable var=RAP>**
  **<T-introduces-variable>**
  tu: You made some errors here. Let's start with RAP.
  **</T-introduces-variable>**
  **<T-tutors-variable>**
    **<T-tutors-via-determinant>**
      **<T-tutors-relationship from-var=CO to-var=RAP>**
        **<T-elicits>**
  tu: How would a change in CO influence RAP?
        **</T-elicits>**
          . . .
    **</T-tutors-via-determinant>**
  **</T-tutors-variable>**
**</T-corrects-variable>**

Tutoring requires at least three levels of goals below the variable level: the *method* level, the *topic* level, and the *primitive* level. The method level shows how to teach about a variable. It can be used to express various types of deductive reasoning, interactive questioning and exploration of anomalies. The topic level represents each item that must be taught. These content items largely involve domain content. The primitive level shows how this information is communicated to the student. In the above example, ***T-tutors-via-determinants*** is a method, ***T-tutors-relationship*** is a topic, and ***T-elicits*** is a primitive level goal.

To refine ***T-tutors-variable*** the tutor chooses a method depending on a number of factors, including domain knowledge (e.g., the mechanism of action of a variable), dialogue history (e.g., the student's previous utterance) and the student model (e.g., how well the student is doing). For example, if the variable is controlled by the nervous system, the tutor often chooses the question and answer style method ***T-does-neural-DLR***. (DLR stands for directed line of reasoning, a form of Socratic dialogue.)

***T-does-neural-DLR***

tu: … Can you tell me how TPR is controlled?
st: Autonomic nervous system.
tu: [Yes.] And the predictions that you are making are for the period before any neural changes take place.
tu: So what about TPR?
st: No change.

For non-neural variables the most common schema is ***T-tutors-via-determinants***. With this method the tutor corrects the value of a variable by invoking a relationship with another core variable.

***T-tutors-via-determinants***

tu: What parameter determines the value of RAP?
st: CO.
tu: What relationship do they have?
st: Inverse.
tu: Right, then what is the value of RAP?

Both of these methods are based on the domain reasoning used by the tutor to solve the problem. Our tutors also use other methods which are less directly based on domain reasoning. ***T-moves-forward*** is similar to ***T-tutors-via-determinants*** but it applies when the determinant has already been mentioned in the conversation.

***T-moves-forward***

tu: [Since CO goes up early in the response, that will cause RAP to fall.] Now what will happen to SV?

In ***T-shows-contradiction***, the tutor corrects the student's error by pointing out a physiological inconsistency in the student's answers.

***T-shows-contradiction***

tu: You said that RAP goes up but earlier you said that CO went up. How is that possible?

***T-explores-anomaly*** is superficially similar, but is used in cases where the reported facts only appear inconsistent. Its goal is to ensure that the student really understands the deeper qualitative relationships among the variables.

***T-explores-anomaly***

tu: So, we have HR down, SV up and CO down. How is that possible?

Although our main goal in annotating transcripts is to collect data for text generation (Reiter and Dale, 1997), a secondary goal is to learn about the tutoring strategies of human tutors. Thus we occasionally annotate strategies which a computerized tutor may not be able to handle in the same way that human tutors do. For example, we use the term ***T-diagnoses-error*** when the tutor wants to identify the student's problem.

***T-diagnoses-error***

tu: Why do you think that TPR will decrease?

Methods like ***T-tutors-via-deeper-concepts*** are used to give more detailed explanations to the student after failing to get a correct answer from the student using only the seven core variables. This method gives information to the student (or elicits it from the student) in terms of a more detailed physiological model.

***T-tutors-via-deeper-concepts***

tu: The central venous compartment is a compliant structure that contains a certain volume of blood ...

## Topics

A method consists of a series of topic operators. For example, the following topic operators could be used to build the ***T-tutors-via-determinants*** form mentioned above.

***T-tutors-determinant***

tu: What are the determinants of RAP?

***T-tutors-relationship***

tu: How does the value of CO affect the value of RAP?

***T-tutors-value***

tu: So, what would happen to RAP?

To build the ***T-does-neural-DLR*** form, one would need the following topic operators, followed by ***T-tutors-value***.

***T-tutors-mechanism***

tu: Can you tell me how TPR is controlled?

***T-tutors-DR-info***

tu: The predictions that you are making are for the period before any neural changes take place.

Whenever a deeper conceptual model has been introduced, the tutor must eventually return to the core variable which started the discussion. The topic ***T-tutors-PT-entry*** can be used for this purpose:

***T-tutors-PT-entry***

tu: What parameter in the prediction table reflects the filling of the left ventricle?

## Primitives

The topics share the primitive operators ***T-elicits*** and ***T-informs***. The ***T-elicits*** operator is used when we want the student to participate actively by answering a question. With ***T-informs*** the tutor gives some information to the student. At any level operators can have arguments such as variable name or information desired. Other arguments refer to interpersonal aspects of an utterance (attitude) or textual aspects (narrative mode). Arguments are also inherited from higher level, enclosing goals.

***T-elicits info=var-value***

tu: What is the value of cardiac output [CO]?

***T-informs info=DR-info attitude=remind***

tu: Remember, we are dealing with the period before any change in nervous activity occurs.

***T-informs narrative-mode=summary***

tu: So HR increases and that makes CO go up ...

## Interacting with the Student

The transcripts show several kinds of student answers:

- correct
- clearly incorrect
- near miss answer, which is pedagogically useful but not the desired answer
- don't know answer, where the student said something equivalent to "I don't know"
- partially correct answer, meaning some part of the answer is correct and the rest is incorrect

Depending upon the category of student answer, the tutor may continue with the correct strategy or choose a new one. If the student answer is correct, the tutor moves to the next goal, sometimes giving an acknowledgment such as "good" or "right". In response to a student answer that is clearly incorrect, the tutor may change to a new method which will build on the student's answer. Other possibilities are to ask the question again in a different way, or to give the student the answer so that tutoring can continue. In any of these cases, the tutor may address the student's wrong answer before continuing with hierarchical expansion of the tutoring plan. This feature is a way of tailoring our responses to the student's needs.

A common motivation for changing to a new method is to refer to a more detailed physiological model. This is often triggered by the student's use of a term coming from a deeper model, one type of near miss. (The deeper model can also be introduced by the tutor, in the case where it has not been possible to explain a concept using only the core variables.)

"Don't know" answers are treated in a similar fashion to incorrect answers, but the tutor has less information available to fashion a specific response.

In response to a partially correct answer, the tutor usually responds to the erroneous part first, then the correct part. This usually results in correcting the student's error, then returning to the original tutorial plan.

## Examples

Figure 1 shows an attempt to teach the student about the value of RAP (right atrial pressure) using the ***T-tutors-via-determinants*** method described above. The student gives an incorrect answer about the relationship between CO

**<T-tutors-via-determinants var=RAP>**
. . .
**<T-tutors-relationship from-var=CO to-var=RAP>**
**<T-elicits>**
tu: How does CO affect RAP?
*<S-ans catg=incorrect>*
st: An increase in CO causes an increase in CVP.
*</S-ans>*
**</T-elicits>**
**</T-tutors-relationship>**
**</T-tutors-via-determinants>**

**<T-tutors-via-deeper-concepts from-var=CO to-var=RAP>**
**<T-tutors-relationship from-var=CO to-var=CBV>**
**<T-informs>**
tu: Increased CO decreases CBV.
**</T-informs>**
**</T-tutors-relationship>**
**<T-tutors-relationship from-var=CBV to-var=CVP>**
**<T-elicits info=det-value>**
tu: If CBV decreases what happens to CVP?
*<S-ans catg=correct>*
st: It decreases.
*</S-ans>*
**</T-elicits>**
**</T-tutors-relationship>**
**<T-tutors-value var=RAP>**
**<T-elicits>**
tu: So what would happen to RAP?
*<S-ans catg=correct>*
st: It decreases.
*</S-ans>*
**</T-elicits>**
**</T-tutors-value>**
**</T-tutors-via-deeper-concepts>**

Figure 1. Response to incorrect student answer

(cardiac output) and RAP. Boldface forms represent the tutorial hierarchy while italic ones represent responses to the student's immediately previous utterance. In this example, the tutor replaces the current method with ***T-tutors-deeper-concepts***, which will attempt to teach the same information in a different way. The new method goes step-by-step from CO to CBV (central blood volume), CBV to CVP (central venous pressure), and CVP to RAP.

Figure 2 shows an example of the tutor responding to a near miss. The student mentions CVP, which is not a core variable. The tutor responds by attempting to lead the student from CVP, which is a step on the right path, to the desired answer, CO. In this case this procedure must be followed twice in order to get to CO. Once the desired answer is obtained, the ***T-tutors-via-determinants*** method continues as before, with the topic ***T-tutors-value***.

In Figure 1, since CO was mentioned in the first method, the tutor knows at the beginning of the new method that all the steps from CO to RAP will be covered. In Figure 2, the tutor moves backward from RAP to CO, using the student's responses to determine how many steps must be mentioned. The sequence terminates when the student mentions CO.

**<T-tutors-via-determinants var=RAP>**
**<T-tutors-determinant>**
**<T-elicits>**
tu: What parameter determines RAP?
*<S-ans catg=near-miss>*
st: CVP.
*</S-ans>*
**</T-elicits>**
**<T-moves-toward-PT>**
**<T-tutors-determinant var=CVP>**
**<T-elicits>**
tu: What determines CVP?
*<S-ans catg=near-miss>*
st: Blood volume [CBV].
*</S-ans>*
**</T-elicits>**
**</T-tutors-determinant>**
**</T-moves-toward-PT>**
**<T-moves-toward-PT>**
**<T-tutors-determinant var=CBV>**
**<T-elicits>**
tu: What determines CBV?
*<S-ans catg=correct>*
st: CO.
*</S-ans>*
**</T-elicits>**
**</T-tutors-determinant>**
**</T-moves-toward-PT>**
**</T-tutors-determinant>**
**<T-tutors-value>**
**<T-elicits>**
tu: How would RAP change?
. . .

Figure 2. Response to "near miss" student answer

Figure 3 shows an example of a response to a "don't know" answer. In the near miss example of Figure 2, the student gave the tutor something to build on but this is not the case with the "don't know" answer in Figure 3. Therefore the tutor must teach the knowledge from scratch, as in Figure 1. Figure 3 differs from Figure 1 in two ways. First, the tutor is working backward along the concept map because the desired endpoint, CO, has not been mentioned yet. Second, the new method is subordinate to the current topic, namely ***T-tutors-determinant***, because the goal of the new method is only to answer the question "what determines the value of RAP?" Thus the last topic of the original method, i.e. ***T-tutors-value***, is still required whereas in Figure 1 it is not.

In all three examples the student's answer triggers a change in the tutor's plan, thus personalizing the response in accordance with the student's knowledge.

## Conclusion

Transcript analysis has given us new insights into a hierarchical plan-based understanding of the behavior of human tutors. It has enabled us to be more precise about many aspects of operator structure and selection. We illustrated this fact by showing several examples of tutors responding to different student errors.

In addition, by using standard, machine-readable SGML, we can perform computerized analyses of the corpus. Starting from a detailed and rigorous markup which corresponds to the plan operators needed by the system reduces the amount of knowledge engineering required to add the operators to our knowledge base. This is an important step toward enabling CIRCSIM-Tutor v. 3 to generate correct and natural tutoring dialogues.

## Acknowledgments

Joel A. Michael and Allen A. Rovick, professors in the Department of Physiology at Rush Medical College, are responsible for the pedagogical and domain knowledge in CIRCSIM-Tutor. They also served as expert tutors for the transcripts. Michael Glass contributed in many ways to the development of the ideas expressed here.

## References

Freedman, R. 1996. Interaction of Discourse Planning, Instructional Planning and Dialogue Management in an Interactive Tutoring System. Ph.D. diss., Dept. of EECS, Northwestern Univ.

Freedman, R and Evens, M. W. 1996. Generating and

Revising Hierarchical Multi-Turn Text Plans in an ITS. In Intelligent Tutoring Systems: Third International Conference (ITS ’96), Montreal, pp. 632–640.

Reiter, E. and R. Dale (1997). Building Applied Natural-Language Generation Systems. *Journal of Natural-Language Engineering* 3: 57–87.

**<T-tutors-via-determinants var=RAP>**
  **<T-tutors-determinant>**
    **<T-elicits>**
  tu: Do you know what determines the value of RAP?
*<S-ans catg=don’t-know>*
  st: I don’t know.
*</S-ans>*
    **</T-elicits>**
    **<T-tutors-via-deeper-concepts-back from-var=CO to-var=RAP>**
      **<T-tutors-determinant-deep var=RAP>**
        **<T-informs>**
  tu: RAP is essentially the same as CVP.
        **</T-informs>**
      **</T-tutors-determinant-deep>**
      **<T-tutors-determinant-deep var=CVP>**
        **<T-informs>**
  tu: CVP is determined by the compliance of the system and the volume of blood in the central venous compartment.
        **</T-informs>**
      **</T-tutors-determinant-deep>**
      **<T-tutors-determinant var=CBV>**
        **<T-elicits>**
  tu: What determines that volume?
*<S-ans catg=correct>*
  st: CO.
*</S-ans>*
        **</T-elicits>**
      **</T-tutors-determinant>**
    **</T-tutors-via-deeper-concepts-back>**
  **</T-tutors-determinant>**
  **<T-tutors-value var=RAP>**
    **<T-elicits>**
  tu: So what would happen to RAP?
*<S-ans catg=correct>*
  st: It decreases.
*</S-ans>*
    **</T-elicits>**
  **</T-tutors-value>**
**</T-tutors-via-determinants>**

Figure 3. Response to “don’t know” student answer