An Analysis of Multiple Tutoring Protocols[1]

Byung-In Cho[1], Joel A. Michael[2], Allen A. Rovick[2], Martha W. Evens[1]

[1] Department of Computer Science, Illinois Inst. of Technology,
10 W. 31st Street 236–SB, Chicago, IL 60616, U.S.A.
chobyun@charlie.iit.edu. evens@iit.edu
http://www.csam.iit.edu/~circsim
[2] Department of Physiology, Rush Medical College
1750 W. Harrison St., Chicago, IL 60612, U.S.A.
{jmichael,arovick}@rush.edu

Abstract. The tutoring protocol controls the interaction between the tutor and the student in a tutoring session. Our goals are to understand human tutoring so that we can emulate it better and to discover which tutoring protocol gives the best results in teaching causal reasoning. We used C5.0 to analyze a set of human tutoring transcripts to discover how and when human tutors switch protocols. In order to understand which students prefer which protocols, we compared the students' performance using CIRCSIM-Tutor with their responses to a questionnaire about the program.

1 Introduction

Determining effective tutoring strategies may be the most important and hardest issue in intelligent tutoring systems. The tutoring protocol controls the interaction between the tutor and the student in a tutoring session. Moore [9] identified three types of interaction: student-content, student-teacher, and student-student. In traditional classroom teaching (student-teacher) interaction is normally immediate. Much educational research supports the belief that immediate feedback increases a sense of excitement and spontaneity [2], [10], [15]. Our colleagues, Joel Michael (JAM) and Allen Rovick (AAR) believe, however, that immediate feedback is not always the best choice. They feel that they can do a better job of tutoring if they ask the student to make predictions first, because the improved student model allows them to plan a tutorial strategy that targets the student's misconceptions [8].

1

We are building an intelligent tutoring system called CIRCSIM-Tutor designed to help medical students learn to solve cardiovascular problems. Our system is based on the study of human tutoring sessions carried out by JAM and AAR. The tutors had decided to use the following protocol they had designed for our intelligent tutoring system in their human tutoring sessions: collect predictions first and tutor afterwards, in order to provide us with examples of the kind of tutoring they wanted the system to produce. We discovered in the analysis described here, however, that the tutors did not always follow the protocol. Sometimes the tutor did not wait until the student finished the predictions. If the student started out with poor predictions then the tutor immediately began to guide the student in the right track with hints or explanations. So, in fact, they changed the tutoring protocol to best fit the student's needs at the time.

In this paper we analyze a subset of those human tutoring transcripts to discover how and when the tutors switch protocols. We used the machine learning program C5.0 [13] to find more rules that govern this change in tutoring. We also wanted to find out how the students feel about the issue of immediate feedback. To do so we analyze the students' performance using CIRCSIM-Tutor and their responses to a questionnaire about their view of CIRCSIM-Tutor. Our goals are to understand human tutoring so that we can emulate it better and to discover which tutoring protocol gives the best results in teaching causal reasoning.

1.1 Planning

The first and most important capability of an Intelligent Tutoring System is dynamic planning. The planner must be able to decide what and how to teach next. It must have a dynamic planning capability; it must be able to generate plans, monitor the execution of the plans, and generate new plans. It must be able to replan when necessary [17]. Finally, the planner must be adaptive. It must customize its tutoring plans for each student [5], [16], [17].

Planners select and sequence the subject matter. Curriculum Planning is concerned with selecting the next problem [1]. Instructional Planning selects and sequences the material to be tutored. Discourse Planning controls the actual presentation of material to the students [4].

1.2 CIRCSIM-Tutor

The domain of CIRCSIM-Tutor is cardiovascular physiology. CIRCSIM-Tutor assists students to reason about the qualitative, causal responses of the human circulatory system when the blood pressure is perturbed. The system asks the student to enter predictions in the Prediction Table [12] indicating how the perturbation affects seven important physiological variables at three different stages of the response, and then it initiates a tutorial dialogue to remedy any errors. Table 1 shows a prediction table for

2

the perturbation "Increase Venous Resistance to 200% of normal." The three stages are the Direct Response (DR): the immediate change in the variables induced by the perturbation; the Reflex Response (RR): the change induced by the response of the central nervous system to the change in blood pressure; and the Steady State (SS): the long term balance between the effects of the perturbation and the effects of the negative feedback. We ask the student to predict the qualitative change from the values before the perturbation to the new steady state. The primary variable is the first variable in the DR column of the Prediction Table that is affected by the current perturbation. Therefore the student should identify and predict the primary variable first.

Table 1. The Prediction Table of the Procedure "Increase Venous Resistance to 200% of Normal" with Correct Answers

Physiological Variable	DR	RR	SS
Inotropic State (IS)	0	+	+
Central Venous Pressure (CVP)	−	−	−
Stroke Volume (SV)	−	−	−
Heart Rate (HR)	0	+	+
Cardiac Output (CO)	−	+	−
Total Peripheral Resistance (TPR)	0	+	+
Mean Arterial Pressure (MAP)	−	+	−

(+: Increased, −: Decreased, 0: unchanged)

2   Tutoring Protocols

Instructional planning determines the content and sequence of the subject matter to be taught in a single procedure. One of the important features of the tutorial planning process is the tutoring protocol. The tutoring protocol defines the overall communication between the tutor and the student. We want to be able to compare the effects of different protocols or to change the protocol during a session.

2.1   Tutoring Protocol 3

Khuwaja described three tutoring protocols that he found used in human tutoring sessions [6]. In Tutoring Protocol 1 the tutor ignores the sequence of the student's predictions and explores the student's response at each point in problem solving. Here the tutor provides immediate feedback for each student prediction and response. In Tutoring Protocol 2 the tutor insists that the student follow the preferred prediction sequence but does not correct the values of the variables until all predictions have

been made. In Tutoring Protocol 3 the tutor makes sure that the student chooses the primary variable (DR) first and predicts its change correctly before asking the student to predict the remaining variables in any order. In RR and SS the students are free to start with any variable and to make predictions in whatever sequence they choose.

2.2 A new tutoring protocol

We have accumulated over 60 transcripts of human tutoring sessions. Our domain experts JAM and AAR carry out keyboard-to-keyboard tutoring using a program, CDS [7], that establishes communications between two computers using modems. In this study we made a detailed analysis of a set of tutoring sessions to see how the tutors used the tutoring protocols. Most sessions contain only one procedure and are one hour long. We chose to study the nine sessions that involved two procedures and lasted up to two hours, so we could observe changes in behavior over time. In four sessions the tutor started with the Centrifuge procedure in which the primary variable is CVP. In five sessions the tutor started with the Alpha-adrenergic procedure in which the primary variable is TPR.

```
Tutor (Problem)
   Tutor (DR)
      Prediction & Tutoring (Primary Variable)
         Prediction (Primary Variable)
         Tutoring (Primary Variable)
      Prediction & Tutoring (Rest of the prediction
                   table variables)
         If not "sequence violation"
             Then Prediction (Variable X)
                 Tutoring (Variable X)
             Else give a sequence hint
   Tutor (RR)
      Prediction & Tutoring (Prediction table variables)
         If not "sequence violation"
             Then Prediction (Variable X)
                 Tutoring (Variable X)
             Else give a sequence hint
   Tutor (SS)
      Prediction & Tutoring (Prediction table variables)
         Prediction (Variable X)
         Tutoring (Variable X)
```

Figure 1. Tutoring Protocol 4

The results of the analysis of protocol use in the tutoring sessions were a complete surprise to us. The tutors planned to use Protocol 3. This means that the tutor analyzed the student's prediction results and then planned the tutoring strategy based on these results. Sometimes, however, the tutor does not wait until the student finishes the predictions. If the student starts with poor predictions then the tutor starts

4

Gauthier, Gilles, Claude Frasson, and Kurt VanLehn, eds. 2000. *Intelligent Tutoring Systems: 5th International Conference, ITS 2000, Montreal*, Springer, Lecture Notes in Computer Science no. 1839, pp. 212–221.

to guide the student in the right track with hints or explanations. We named this new protocol, Protocol 4. According to our analysis of the transcripts, when the tutor used immediate feedback they also enforced a particular tutoring sequence. In Protocol 4 (see Figure 1) the tutor considers the student's prediction sequences. The student must follow a prediction sequence that the tutor thinks correct. The tutor explores the student's response at each point in problem solving. Therefore the tutor provides immediate feedback for each student prediction and response.

Each tutoring session in our transcripts can be divided into prediction phases and tutoring phases. The structure of the Prediction Table, shown in Figure 1, divides the problem into three stages - DR, RR, and SS. In each stage the tutor performs two common operations. During the first operation "Prediction", the student predicts whether a physiology variable will increase (go up, +, up,…), decrease (go down, −, down, …), or stay the same (unchanged, 0, stay, …). During the second operation "Tutor", the tutor starts a dialogue to remedy any prediction errors.

In Protocol 4, like other protocols, the primary variable is predicted and taught first. The rest of the variables should be predicted and taught in the sequence defined by the problem. If the student does not follow the sequence, the tutor gives a sequence hint about the prediction order based on the causal reasoning to be followed. Otherwise, the tutor gives instant feedback for the predicted variable.

3   Analysis of Human Tutoring Transcripts

We used C5.0 [13], which is an upgraded version of the decision tree induction program C4.5 [11], to produce the rules that describe when our domain experts switched tutoring protocols. In this experiment we had 44 cases (the number of tutoring phases recorded in the 18 procedures studied), each with 11 attributes.

3.1   Attributes

The first three attributes in Table 2 are related to the discussion about the basic concept. The basic concept involves the effects of the centrifuge in the Centrifuge procedure, or the Alpha-adrenergic Receptors in the Alpha-adrenergic procedure.

Gauthier, Gilles, Claude Frasson, and Kurt VanLehn, eds. 2000. *Intelligent Tutoring Systems: 5th International Conference, ITS 2000, Montreal*, Springer, Lecture Notes in Computer Science no. 1839, pp. 212–221.

Table 2. Attributes for Rule Extraction

attribute	value	remark
Discussion Type	T, S	Tutor-Primary / Student-Explanation
Discussion Success	S, U	Satisfactory / Unsatisfactory
Discussion Length	continuous	How many turns in the discussion
Primary Error	continuous	Wrong answers for a primary variable
Prediction Score	continuous	(right - wrong) prediction
Remediation	continuous	Correct answers in the total answers
Sequence Error	continuous	Sequence error in Protocol 4
Pre-Prediction Score	continuous	Previous stage, (right - wrong) prediction
Pre-Remediation	continuous	Previous stage, correct answers in the total answers
Pre-Sequence Error	continuous	Previous stage, sequence error in Protocol 4
Current Stage	dr, rr, ss	Current stage

The students often had difficulty in determining the primary variable. Therefore tutoring frequently began with a discussion of the relationship between the basic concept and the primary variable. Thirteen of the eighteen procedures started with a discussion of the basic concepts. The Discussion Type (DT) in Table 2 is T if the tutor started the discussion to remedy a wrong primary variable prediction. It is S if the student began the discussion with a request for an explanation. The Discussion Success (DS) indicates whether the discussion was successful or not, that is, whether the tutor is satisfied with the student's responses at least 50% of the time. The Discussion Length (DL) indicates the number of turns in the discussion counting from the start to the turn in which the student gave the right answer for the primary variable. The Primary Error (PE) is a count of the number of wrong answers entered for the primary variable. The number of primary errors reflects the comprehension of the procedure. The Prediction Score (PS) indicates how many right or wrong answers were made in the prediction phase. The score we used was the number of right answers minus the number of wrong answers. The value of the Remediation (RM) attribute is the percentage of correct answers among the total answers given by the student in that stage. The Sequence Error (SE) attribute represents the number of sequence errors during Protocol 4. The Pre-Prediction Score, Pre-Remediation, and Pre-Sequence Error: These attributes represent the Prediction Score, Remediation, and Sequence Error from the previous stage. The Current Stage (CS) indicates the stage on which the student is now working.

Table 3. The Summary of Human Tutoring Sessions K30 - K38

Tutor	Session	DR							RR			SS		
		DT	DS	DL	PE	PS	RM	SE	PS	RM	SE	PS	RM	SE
AAR	K30 CVP	-	-	-	0	1	0	-	3	0.22	-	1	0.25	-
	K30 TPR	T	U	19	3	3	0.44	-	5	0	-	3	0	-
	K31 TPR	S	U	28	5	-1	0.3	-	-3	-	-		=	
	K31 CVP	-	-	-	2	-	0.2	0		=			=	
	K32 TPR	S	S	4	1	5	0.28	-	3	0.25	-	5	0	-
	K32 CVP	T	S	13	2	7	0	-	7	0	-	5	0.4	-

Gauthier, Gilles, Claude Frasson, and Kurt VanLehn, eds. 2000. *Intelligent Tutoring Systems: 5th International Conference, ITS 2000, Montreal*, Springer, Lecture Notes in Computer Science no. 1839, pp. 212–221.

	K33 CVP	-	-	-	0	7	0	-	1	0	-	5	0.5	-
	K33 TPR	T	S	13	1	1	0.5	-	3	0.5	-	7	0	-
	K34 CVP	T	U	5	1	1	0.125	-	-1	-	-	=		
	K34 TPR	S	U	4	1	-	0.45	0	=			=		
	K35 CVP	-	-	-	1	1	0	-	3	0.4	-	1	-	-
	K35 TPR	S	U	14	1	-	0.17	0	-	0.25	0	-	0	0
	K36 TPR	S	U	14	1	1	-	-	=			=		
JAM	K36 CVP	T	S	7	1	7	0	0	1	0.33	-	=		
	K37 CVP	-	-	-	0	-1	0	-	7	1	-	3	0	-
	K37 TPR	S	S	10	0	3	0.1	-	5	0	-	7	0	-
	K38 TPR	T	U	17	1	-2	-	-	-	0.25	2	=		
	K38 CVP	T	S	11	1	-	0.17	1	1	0.75	-	3	0.33	-

CVP: Centrifuge Procedure
TPR: Alpha-adrenergic Procedure

Table 3 summarizes sessions K30 - K38, which are the input to the rule induction program. "=" means the transcript does not have the stage data. "-" means that the data is not available. For example, the "K30 CVP" procedure did not include a discussion about the basic concepts. A white cell indicates that Protocol 3 was in use and a shaded cell indicates that Protocol 4 was in use in that stage.

3.2 Switching Rules

The rules extracted by C5.0 do not classify all cases correctly; there is an error rate of 9.1 %. The switching rules are:

```
If Discussion Success = S
   If Remediation <= 0.5
     If Current Stage = ss
        If Prediction Score <= 1
        then switch from Protocol 3 to Protocol 4

If Discussion Success = U
   If Primary Error > 2
     If Prediction Score <= -2
     then switch from Protocol 3 to Protocol 4

If Discussion Success = U
   If Primary Error <= 2
     If Pre-Remediation > 0.35
        If Current Stage <> rr
        then switch from Protocol 3 to Protocol 4

If Discussion Success = U
   If Primary Error <= 2
     If Pre-Remediation <= 0.35
        If Discussion Type = T
```

```
                        If Current Stage = rr
                        then switch from Protocol 3 to Protocol 4

  If Discussion Success = S
     If Remediation > 0.5
     then switch from Protocol 4 to Protocol 3

  If Discussion Success = S
     If Remediation <= 0.5
        If Current Stage = dr
           If Pre-Sequence Eerror <= 1
              If Pre-Prediction Score <= 1
              then switch from Protocol 4 to Protocol 3
```

Examing these rules we see that two important factors determined whether the tutor switches the protocol from Protocol 3 to Protocol 4. The occurrence of a discussion about the basic concepts of the procedure at the very beginning of a procedure is an important factor in protocol switches. For example, the protocol switch is likely if the student asked for some explanation before the prediction and did not understand that explanation right away, or if the tutor asked some question about the basic concepts to remedy the student's wrong primary variable prediction, but the student replied with unsatisfactory answers. The other important factor, which makes the tutor switch the protocol from Protocol 3 to Protocol 4, is the student's performance scores. One score is the prediction score and the other is the number of primary variable errors. However, if the student performed well in the previous stage then the tutor did not switch but gave some hints about the primary variable.

On the other hand, the tutor went back from Protocol 4 to Protocol 3 under the opposite conditions. If the student made a good performance in the previous stage and there was a satisfying discussion about the basic procedure concepts then the tutor switched back to Protocol 3.

4   Some other important characteristics of tutoring protocols

We found some by-product rules during the analysis of human transcripts. We wondered how long the tutor keeps a given protocol. The tutor always starts a procedure with Protocol 3. However, after switching the protocol, the tutor sticks with the Protocol 4 to the next stage or next procedure.

Student initiatives also affect protocol switches. A student initiative means any student contribution to the dialogue that is not an answer to a question asked by the tutor [3], [14]. In Protocol 3, sometimes our tutor met with a simple student initiative that requires only a short response from the tutor. In the prediction phase, for example, the student may ask a simple question [3] or the student could not answer in more than one minute. In these cases, the tutor gives a simple hint [18] and sticks with the current protocol. In particular, if the student performed well in the previous

Gauthier, Gilles, Claude Frasson, and Kurt VanLehn, eds. 2000. *Intelligent Tutoring Systems: 5th International Conference, ITS 2000, Montreal*, Springer, Lecture Notes in Computer Science no. 1839, pp. 212–221.

stage, but starts with poor predictions, then the tutor does not switch protocols but gives hints about the current stage.

5   Which students prefer immediate feedback?

Who prefers immediate feedback? Which students feel a lack of instant feedback during the prediction phase in Protocol 3? We can imagine the situation intuitively. Sometimes the student may want to know whether the current prediction is correct or not. The variables in the prediction table are causally related. Therefore if the student is not sure of one variable then the uncertainty may affect the following variables. In order to discover how the students feel about the protocol issue we compared the students' performance using CIRCSIM-Tutor with their answers to a questionnaire asking about their view of CIRCSIM-Tutor.

This data came from an experiment that was performed by forty-eight first year students at Rush Medical College, in November 1998. The system presented four cardiovascular procedures to be solved. The system was designed to use Protocol 3, which means no immediate feedback and discard the prediction sequence except for the primary variable.

After using CIRCSIM-Tutor the students answered a questionnaire that asks the students' view of the system. The questionnaire had ten questions and employed a five point Likert scale. Question 8 "I would prefer that the system always tell me about my mistakes immediately" asked the student's opinion about the tutoring protocol.

Seven students answered that they would prefer immediate feedback and six of them (86%) made poor predictions while using the program. We defined a poor/good prediction result to mean that the student's prediction score on four procedures was under/over the average prediction result of all students. Which students do not ask for immediate feedback? Twelve students answered that they prefer the current protocol, Protocol 3, which does not give feedback immediately, in the questionnaire and eight of them (67%) made good predictions when using the program. The analysis results suggest that the students who made poor predictions are eager to know their mistakes immediately.

6   Conclusion

In this study we carried out a detailed analysis of tutoring sessions to see how the tutors used the tutoring protocols. Two important factors that determined whether the tutor switches the protocol are the discussion about the basic concepts of the procedure and the student's performance scores. If the student did not perform well then the tutor responded immediately. The tutor did not wait until the student finished the predictions but started to tutor the student in the right track. However, when the

9

Gauthier, Gilles, Claude Frasson, and Kurt VanLehn, eds. 2000. *Intelligent Tutoring Systems: 5th International Conference, ITS 2000, Montreal*, Springer, Lecture Notes in Computer Science no. 1839, pp. 212–221.

tutor met a simple student initiative the tutor did not change the protocol but gave some hints. If the student performed well in the previous stage, giving good predictions and a satisfactory discussion about the basic procedure concepts, then the tutor switched back to Protocol 3.

In order to find out which protocol the students prefer we also analyzed the result of the CIRCSIM-Tutor experiment and the questionnaire. The analysis results say that the students' protocol preference is closely related to their performance using CIRCSIM-Tutor. Students who are doing well are comfortable waiting for feedback while most students who are doing badly want immediate feedback.

We plan to enable CIRCSIM-Tutor to switch protocols. Then we will carry out a controlled experiment to find out which protocol improves the students' learning outcomes.

References

1. Cho, B., Michael, J., Rovick, A., and Evens, M.: A Curriculum Planning Model for an Intelligent Tutoring System. Proceedings of 12th International Florida Artificial Intelligence Research Symposium. Orlando, FL. (1999) 197–201
2. Cuffman, D. and MacRae, N.: Faculty Development Programs in Interactive Television. Proceedings of the 1996 Mid-South Instructional Technology Conference. Track 2: Distance Learning. (download:1999.12.30.) http://www.mtsu.edu/~itconf/proceed96.html
3. Freedman, R.: Degrees of Mixed-Initiative Interaction in an Intelligent Tutoring System. In: Haller, S. and McRoy, S. (eds.): Working Notes of AAAI97 Spring Symposium on Mixed-Initiative Interaction, Stanford, CA. (1997) 44-49
4. Freedman, R.: Atlas: A Plan Manager for Mixed-Initiative, Multimodal Dialogue. AAAI '99 Workshop on Mixed-Initiative Intelligence, Orlando. FL. (1999) 107-114
5. Katz, S., Lesgold, A., Eggan, G., Girdin, M., and Greenberg, L.: Self-adjusting Curriculum Planning in Sherlock II, Lecture Notes in Computer Science: Proceedings of the Fourth International Conference on Computers in Learning (ICCAL '92). Berlin: Springer Verlag. (1992) 343-355
6. Khuwaja, R. Rovick, A., Michael, J., and Evens, M.: A Tale of Three Tutoring Protocols: The Implications for Intelligent Tutoring Systems. Proceedings of Golden West, Las Vegas, Nevada, June 9-12 (1994) 109-118
7. Li, J., Seu, J., Evens, M., Michael, J., and Rovick A.: Computer dialogue system (CDS): A system for capturing computer-mediated dialogue. Behavior Research Methods, Instruments, & Computers, 24(4) (1992) 535-540
8. Michael, J., Rovick A., Evens, M., Shim, L., Woo, C., and Kim, N.: The uses of multiple student inputs in modeling and lesson planning in CAI and ICAI programs. In I. Tomek (editor), Computer Assisted Learning. Berlin: Springer Verlag. (1992) 441-452
9. Moore, M.: Three types of interaction. American Journal of Distance Education. Volume 3 Number 2 (1989) 1-6
10. Moore, M. and Kearsley, G.: Distance Education: A Systems Perspective. Wadsworth, Bermont, CA. (1996)
11. Quinlan, J.: C4.5: Programs for Machine Learning. Los Altos, CA: Morgan Kaufmann (1993)

Gauthier, Gilles, Claude Frasson, and Kurt VanLehn, eds. 2000. *Intelligent Tutoring Systems: 5th International Conference, ITS 2000, Montreal*, Springer, Lecture Notes in Computer Science no. 1839, pp. 212–221.

12. Rovick, A. and Michael, J.: The predictions table: a tool for assessing students' knowledge. American Journal of Physiology 263 (Advances in Physiology Education 8) (1992) S33-S36
13. RuleQuest. RuleQuest Research Pty Ltd. (download:1999.9.8.)
   http://www.rulequest.com/download.html
14. Shah, F, and Evens, M.: Student Initiatives and Tutor Responses in a Medical Tutoring System. In: Haller, S. and McRoy, S. (eds.): Working Notes of AAAI97 Spring Symposium on Mixed-Initiative Interaction, Stanford, CA. (1997) 138-144
15. Travers, A. and Decker, E.: New Technology and Critical Pedagogy. Radical Pedagogy, Volume 1: Issue 2, Summer (1999)
16. Wilensky, R., Chin, D., Luria, M., Martin, J., Mayfield, J. and Wu, D.: The Berkely UNIX Consultant Project (CSD-89-520). (download:1998.6.4.)
   http://sunsite.berkeley.edu:80/Dienst/UI/2.0/Describe/ncstrl.ucb%2fCSD-89-520?abstract.
17. Woo, C.: Instructional Planning in an Intelligent Tutoring System: Combining Global Lesson Plans with Local Discourse Control, Ph. D. Dissertation, Illinois Institute of Technology (1991)
18. Zhou, Y., Freedman, R. Glass, M. Michael, J., Rovick, A., and Evens, M.: Delivering Hints in a Dialogue-Based Intelligent Tutoring System, Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI-99), Orlando. FL. (1999) 128-134