Proc. Sixteenth Midwest AI and Cognitive Science Conf. MAICS2005, Dayton, April 2005 pp. 25-30

Effective Behaviors in a Comparison Between Novice and Expert Algebra Tutors

Hyeun Me Chae[1], Jung Hee Kim[1], and Michael Glass[2]

[1]Department of Computer Science
North Carolina A&T State University
Greensboro, NC 27411
hchae@ncat.edu, jungkim@ncat.edu

[2]Mathematics and Computer Science Department
Valparaiso University
Valparaiso, IN 46383
michael.glass@valpo.edu

Abstract

The algebra tutorial dialogue transcripts collected at North Carolina A&T State University afford an opportunity for a head-to-head comparison of an expert and a novice tutor teaching instances of the same algebra problem. In this paper we concentrate our comparison on quantifiable phenomena that have proved fruitful in the analysis of tutoring. By comparing expert to novice we discover what behaviors constitute expertise. In our transcripts a striking difference is that experts engage in many more collaborative goal-setting moves, especially in situations where they are not addressing a student impasse. We also note increases in eliciting vs. informing dialogue moves, as well as in addressing student errors immediately.

Introduction

The North Carolina A&T State University Algebra Tutorial Dialogue Project has collected over 50 one-hour transcripts of tutoring of college-level remedial algebra problems. These sessions are computer-mediated, typed conversation between tutor and student, using tutors with several different levels of experience (Kim and Glass, 2004). The data from these transcripts affords us an opportunity to study quantitative differences in the styles of expert and novice tutors in this domain, thus highlighting the kinds of behaviors the expert tutors have acquired that sets them apart from novices.

This study augments our earlier results (Glass et al., 1999) using transcripts of medical students being tutored in physiology. In that study, the most marked difference was that experienced tutors engaged in considerably more eliciting activity relative to informing activity than the novices did. This study reinforces that conclusion. In addition, we studied different dialog patterns and events in tutoring that have been postulated to correlate with learning (VanLehn et al., 2003). To that end we marked up our transcripts according to dialog events such as goal setting and impasses, and we counted the number of times these phenomena occurred singly and in combinations.

Environment and Procedure

Our tutoring dialogues consist of a tutor and a student working problems collaboratively, with the typed dialogue, the problem statement, and the equation being worked on all visible to both parties. Typed communication is somewhat artificial. The advantage from the point of view of our studies is we can capture all the communication between student and tutor, there are no gaze, gesture, prosodic features, and so on to capture and annotate. Further, this is the mode most tractable for computer-based instruction, an eventual goal of this project. We use software designed for supporting and capturing algebra tutoring (Patel, Glass & Kim, 2003), e.g. by supporting a mathematical equation-building toolbar.

An extract from a tutorial dialogue is illustrated in Figure 1. During these conversations the tutor and student can both update the equation under construction in the manner of collaborative problem-solving observed by Fox (1993). For analysis and display purposes, the equation is interpolated into the verbal transcript whenever it changes.

The students are all recruited from the first year basic algebra classes. Problems solved during these tutoring sessions include both symbolic manipulation problems and word problems, viz:

1. Please factor 3x2 – 11x + 6.
2. Bob drove “m” miles from Denver to Fargo. Normally this trip takes “n” hours, but on Tuesday there was good weather and he saved 2 hours. Write an equation for his driving speed “s”.

Students solve an average of between 5 and 6 problems in an hour session.

Prior to each tutoring session the student take a pre-test of twelve different types of problems. The tutor chooses the selection of topics to teach based on the results of this pre-test. The post-test covers only those types problems that were taught. Learning gains are thus measured using only those types of problems that were taught during the session.

Proc. Sixteenth Midwest AI and Cognitive Science Conf. MAICS2005, Dayton, April 2005 pp. 25-30

	`Problem:` **`(3x-2) (x+5)`**
Tu:	*What operation do you need to apply here?*
St:	*Factoring.*
Tu:	*No, Not factoring.*
Tu:	*The expression is already factored.*
Tu:	*What is the opposite of factoring?*
St:	*Distributing.*
Tu:	*You can say distributing.*
Tu:	*But we are multiplying here.*
Tu:	*Now, what process do we use when we are multiplying two binomials.*
St:	*Foil.*
Tu:	*Excellent, You use FOIL.*
Tu:	*What does FOIL stand for?*
St:	*First Outer Inner Last.*
Tu:	*Good, Now go ahead and use FOIL t o multiply the two binomials together.*
St:	$3x^2 + 15x - 3x^2 - 10$
St:	$3x^2 + 15x - 2x - 10$
St:	$3x^2 + 13x - 10$
Tu:	*Excellent.*

Figure 1: Typical tutoring dialogue.

The expert tutor in this study is Dr. Kathy Cousins-Cooper, a professor in the Mathematics Department at NC A&T State University who has taught and tutored basic algebra for many years. The novice is an upper-level undergraduate mathematics student who prior to these sessions had the typical student tutoring experiences of a mathematics major.

The two problems we study in this paper are illustrated in Figure 2. We have nine examples of problem 1 taught by the expert and six examples taught by the novice. For problem 2 we have eight expert examples and four novice.

Annotation of the dialogues was performed by the first author of this paper and vetted by the second. We did not compute any reliability measures.

Phenomena Measured

Gross Measures of the Text

We analyzed the time to tutor a single problem and the number of dialogue turns needed. In this paper, one person talking is one turn. We also analyzed learning gains for the two individual problems studied, and for the whole sessions of expert and novice transcripts in order to discover whether our expert tutors are indeed producing better outcomes.

Problem 1: subtract	$\frac{x}{x^2 - x - 6} - \frac{2}{x^2 - 7x + 12}$
Problem 2: solve for x	$5x = 2x^2 + 1$

Figure 2: The two problems studied in this paper.

Eliciting and Informing

We counted the number of times the tutor engaged in an eliciting or an informing dialogue move. In our previous study (Glass et al. 1999) this showed up as the most salient difference between expert and novice tutors. An elicting dialogue move is a turn where the tutor attemped to elicit some information from the student, for example a tutor turn that attempts to elicit from the student the name of the next operation to be performed. This is counted as an elicit whether or not the student responds with the desired information. If the tutor tells the student the some information, that is counted as an inform. It will count as an inform whether or not the tutor previously elicited the same information. Simple prompts, e.g. when the tutor says "go ahead," are not counted.

Goal setting

Following VanLehn et al. (2003), we counted and categorized instances of *goal setting*. Goal setting is a hinting activity in task-oriented dialogues. The motivation to study goal setting comes from cognitive task analysis, and consists of one party verbalizing the next small sub-task to accomplish. Occasionally students set the goals during the dialogue, more usually it is the tutor. Goal setting hints make sure that the tutor and the student are on the same road to reach the answer and keep the student focused on the problem. VanLehn et al. also point out that they serve as retrieval and selection cues, reminding the student of the principles and activities involved.

Sometimes one goal setting hint is enough for the student. McArthur et al. (1990) document one tutor who uses goal setting hints in a scaffolding and fading style, reducing their use as the student becomes more proficent. In the algebra tutoring we have observed, goal setting is often exhibited throughout the solving of a problem, starting from right after the problem statement. In our analysis, we suspect that collaborative goal setting hints done by both parties could help learning more than informed goal setting hints by a tutor.

When counting total goal setting hints we categorized them as follows.

1) Collaborative goal setting vs. informed goal setting. If the tutor elicits a goal from the student, or in some other way both parties collaborate in the setting of the next task goal, we count that as collaborative. Otherwise the goal is produce by one party (the tutor or the student), a behavior we label informed.

2) Goal setting introduced by tutor or goal setting introduced by student, annotating who starts the process.

An example of collaborative goal setting introduced by the tutor is as follows.

T: What is your first step?
S: To get everything on one side and equal to zero
T: Well, remember when we complete the square we must have what as a leading coefficient?
S: So I divide by 5 first?
T: We should probably get our x-terms on one side and the constant on the other.
Then maybe you can recognize what are leading coefficient is.
S Ok

T21:105-110

Here is an example of collaborative goal setting introduced by the student:

S: Do you mean solve as in factor the equation or find out what is "x"?
T: You will have to solve for x.
In most cases when you do factor, you are solving for solutions of x.

T52:702-703

Here is informed goal setting introduced by the tutor:

T: Ok, please simplify the expression.
S: $\frac{x-2}{-8x}$
T: No, that isn't right.
Look at your denominators, you can factor those because for this problem, you must find a Least Common Denominator (LCD).

T41:401-403

Impasses

VanLehn et al. (2003) defines an impasse operationally: "an impasse occurs when a student gets stuck, detects an error, or does an action correctly but expresses uncertainty about it."

We consider that an impasse occurs whenever either the tutor or student signals it. Errors that are not caught are not impasses. In our algebra tutoring, the tutor often immediately intervenes when the student gets off track while solving the problem, meaning that most errors result in instant impasse and tutor action. This is congruent with Person and Graesser's (2003) observation that tutors rarely provide delayed feedback, they usually respond immediately.

Our markup and analysis includes both when an impasse occurs and how a tutor reacts after. We categorize impasses by how they are first observed, as follows:

1. Student asks for help.
2. Student pauses for a long time, and the tutor intervenes.
3. Student applies a procedure incorrectly, and the tutor intervenes.
4. Student asks for confirmation of the next procedural step.
5. Student asks for confirmation of whether an answer is correct.
6. Miscommunication occurs.

The main categories of tutor behaviors after impasses are as follows:

1. Tutor tries to correct an error.
2. Tutor tries to explain.
3. Tutor tries to give a goal setting hint.
4. Tutor repairs miscommunication.

An example of an a tutor-detected impasse followed by an error correction:

T: Now what is the LCD?
S: x-3
T: No, the LCD = (x - 3)(x + 2)(x - 4)

T7:517-519

An example of the student asking for procedural help followed by an explanation:

T: No, just distribute the X and the 2 in each term.
Like this…
$x^2 - 4x - 2x - 4$
S: so what happens to the numbers at the bottom?
T: $\frac{x^2 - 4x - 2x - 4}{(x+2)(x-3)(x-4)}$

$$\frac{x^2 - 6x - 4}{(x+2)(x-3)(x-4)}$$

This is the correct answer.

T41:425-427

An example of an impasse followed by none of the above, conversational repair in this case:

T: How are you going to solve it?
S: Get x's on one side.
T: Okay, What I should have asked you was what method are you going to use to solve this?

T2:403-405

Results and Discussion

Learning Gains

We calculated learning gains for problems (1) and (2) separately by:

(Posttest - pretest) / (1-pretest)

where the test scores range from 0.0 to 1.0. We also applied the same formula to the whole-session test scores.

Table 1 shows the difference in average learning gains between novice and expert transcripts.

	Expert	Novice
Problem 1	0.58	0.20
Problem 2	0.27	0.13
All sessions	0.47	0.24

Table 1 Average Learning Gains

The expert tutor produced higher mean learning gains for both problems. Using the two-tailed t-test to test the hypothesis that the mean learning gains are the same for the two tutors, the difference is significant at the $p<0.1$ level for problem 1 and not significant for problem 2. The whole-session learning gains for all expert (n=24) and novice (n=10) tutoring sessions shows that the expert is produces markedly better results, with weak significance $p<0.1$ measured by the t-test.

Duration for One Problem

Table 2 shows average time and average number of turns per tutoring one problem for the expert and novice tutors. Time per problem is about the same, but turns per problem is significantly higher ($p<0.01$) for the expert. This is an overall broad measure that the expert tutor has a more interactive, and potentially more cognitively engaging, style. The particular differences in behavior are studied below.

	Expert	Novice
Average time/problem	13:27	14:45
Average turns/problem	17.5	12.4

Table 2. Time and turns per problem, both questions

Elicits and Informs

In Table 3 we see the difference in counts of eliciting and informing dialogue acts for the two problems. The expert elicits considerably more often. The chi-squared test on the 2x2 table of elicit and inform counts shows the difference is significant at the $p<0.001$ level.

	Expert		Novice	
	Avg.	Total	Avg.	Total
Elicit	7.26	214	5.29	45
Inform	9.30	167	3.21	74
E/(E+I)	0.56		0.38	

Table 3. Eliciting vs. Informing Acts

The most pronounced difference between novice and expert tutors in our earlier study in another domain (Glass et al., 1999) is that the ratio of eliciting to (eliciting + informing) acts was 0.2 for novices and 0.5 for experts. Those ratios are not as markedly different here, but still they are quite different and quite significant. Experts ask more questions, they elicit more material from the students, and thus probably keep the students more engaged.

Goal-Setting Behavior

Looking at instances of goal setting in Table 4 we see that the expert tutors exhibit significantly more instances of both collaborative and informed goal setting events ($p<0.001$ for each type, using a two-tailed t-test.). The distribution of instances of goal-setting between collaborative and informed is not significantly different as measured by the chi-squared test.

	Expert		Novice	
	Avg./Prob	n	Avg./Prob	n
Collab.	3.88	66	1.00	10
Informed	1.59	27	0.50	5
Total	5.47	93	1.50	15

Table 4. Collaborative vs. Informing Goal-Setting

This would seem to indicate that goal-setting is a learned behavior that might account for the expert's better performance. The experts set approximately 5.5 goals per problem solved, vs. 1.5 for the novices. Considering that there are 17.5 turns per problem in an expert session, it is clear that setting goals is a dominant activity. VanLehn et al. (2003) also observe that this is prominent. It would seem that goal-setting is a way to communicate procedural knowledge in what is mostly a problem-solving activity.

In our transcripts we very often observe the interaction between procedural and principled knowledge in action at the very outset of tutoring one problem. The most common pattern is for the tutor to elicit (if possible) from the student 1) a statement of what operation is being requested (principled knowledge) and 2) what is the first goal needed in solving it (procedural knowledge). This behavior fairly explicitly links the two ways of thinking of the task at hand. Goal-setting is a necessary part of that opening dialogue.

Goal-setting does not show up as a category among the possible tutor responses discussed by Person and Graesser (2003), it seems to be sometimes a hint and sometimes an elaboration in their typology. They document both as occurring frequently.

The large overall difference in the ratio between eliciting and informing behaviors does not, curiously, show up as a significant difference in the ratio between collaborative and informing goal-setting .

There is a small difference between novice-run and expert-run tutoring sessions in who engages in goal-setting behavior, the tutor or the student, as seen in Table 5. That the student initiates more goal setting episodes with novice tutors was not significant. The difference between tutor initiated and student initiated goal setting is significant at $p<0.001$ for both expert and novice tutors. However in either regime, the fraction of student-initiated cases is small.

	Expert	Novice
	Avg/Prob	Avg/Prob
Tutor initiates	5.1	1.1
Student initiates	0.2	0.4

Table 5. Who Initiates Goal-Setting, Avg/Problem

Impasses

Among impasse events, we see in Table 6 distribution of how they are recognized. The numbers are average per problem tutored, combined for problems 1 and 2.

	Expert	Novice
St requests help	0.47	1.30
St time delay	0.06	0.30
Tu intervenes	4.47	1.70
St has procedural question	0.06	0.30
St asks for confirmation	0.53	1.40
Miscommunication	0.35	0.10
Total	4.71	3.90

Table 6. How Impasses Occur

We notice that tutor intervention is by far the most common way an impasse occurs. This is consistent with the findings of both VanLehn et al. (2003) and Person and Graesser (2003). Again, we see that this behavior is far more pronounced in the expert tutoring transcripts. However it is suggestive, though not statistically significant, that the students initiate more impasses when dealing with the novice tutors. Perhaps this is because the less agressive tutor behavior gives the students more opportunity to discover their own errors.

Table 7 shows the tutor responses to impasses as average per problem in all Problem 1 and Problem 2 transcripts.

	Expert	Novice
Error Correction	3.94	2.20
Explanation	0.88	1.60
Goal Setting	0.65	1.00
Other	0.41	0.20

Table 7. Tutor Responses to Impasses

What is noteworthy here is not a difference, but a similarity between expert and novice. The expert and the novice produce a similar number of goal-setting moves per problem tutored in response to impasses. In fact, most of the novice's 1.5 goal setting moves per problem are explained by impasse responses, only a small fraction of the expert's are. The conclusion is that the expert learned that setting goals unilaterally, without need for an immediately preceding impasse. This does not contradict the observation that most goal-setting is collaborative, it says only that the tutor initiated it. This is, at least partially, the start-of-problem behavior we commented on above, where the tutor asks the student to name the problem and identify the first procedural step. It seems like the tutor's greater effectiveness might be explained in part by learning to set goals even when the student has not yet exhibited a problem.

Notice also that the novice is much more likely than the expert to engage in explanation behavior. This is consistent with the observation that novices engage in eliciting behaviors relatively less often than experts do.

Conclusions

Expert algebra tutors indeed behave differently than novices. They ask more questions relative to statements, they correct more errors. Most notably, they engage in far more goal-setting behaviors.

The expert's use of goal setting, furthermore, is quite different than the novice's. The novice uses goal setting mostly in response to a student impasse, the expert uses it more often unilaterally, when no impasse has occurred, frequently at the start of tutoring the problem. In fact, since goal setting incidents are approximately one-fourth as common as conversation turns, and the expert tutors heavily favor collaborative goal setting (requiring several turns), it can be inferred that a large fraction of the expert's dialogue turns are devoted to goal setting as a means of tutoring. Explicit tutoring of principles, as opposed to procedures, therefore takes a back seat.

We do not know to what behavior to attribute the expert's better learning outcomes. But it would seem that the conclusion of Person and Graesser's (2003) and Graesser et al. (1995) that tutors largely react to the immediate previous turn, and eschew sophisticated tutoring behaviors, is at least somewhat contradicted by our expert's behavior. Although we did not look for multi-turn schemata in this study, we notice that the expert tutors, unlike the novices, seem to be engaged in teaching the problem-solving procedure as a sequence of collaboratively-set goals.

Acknowledgments

This work was supported by the Cognitive Science Program, Office of Naval Re-search, under grant N00014-02-1-0164, to North Carolina A&T State University. The content does not reflect the position or policy of the government and no official endorsement should be inferred.

We are also indebted to the former students who worked on the NC A&T State University Algebra Dialogue project for creating the software and collecting the dialogues that made our study possible.

References

Fox, Barbara. 1993. *The Human Tutorial Dialogue Project*. Hillsdale, NJ: Erlbaum.

Glass, Michael, Jung Hee Kim, Martha W. Evens, Joel A. Michael, and Allen A. Rovick. 1999. Novice vs. Expert Tutors, a Comparison of Style. In Tenth Midwest Artificial Intelligence and Cognitive Science Conference (MAICS-99), Bloomington, IN.

Graesser, Arthur C., Natalie K. Person, and Joseph P. Magliano. 1995. Collaborative Dialogue Patterns in Naturalistic One-to-One Tutoring, *Applied Cognitive Psychology*, vol. 9, pp. 495-522.

Kim ,Jung Hee and Michael Glass. 2004. Evaluating Dialogue Schmata with the Wizard of Oz Computer-Assisted Algebra Tutor. In Lester, J., R. M. Vicari, and F. Paraguacu, eds., *Intelligent Tutoring Systems: 7th International Conference (ITS 2004), Maceio, Brazil*. Berlin: Springer. Published as LNCS 3220.

McArthur, David, Cathleen Stasz, and Mary Zmuidzinas. 1990. Tutoring Techniques in Algebra, *Cognition and Instruction*, vol. 7, pp. 197-244.

Patel, Niraj, Michael Glass, and Jung Hee Kim. 2003. Data Collection Applications for the NC A&T State University Algebra Tutoring Dialogue (Wooz Tutor) Project. In Ralescu, A. ed., Fourteenth Midwest Artificial Intelligence and Cognitive Science Conference (MAICS-2003), Cincinnati, 2003.

Person, Natalie K. and Arthur Graesser. 2003. Fourteen Facts About Human Tutoring: Food for Thought for ITS Developers, Artificial Intelligence in Education, Proceedings of the 2003 conference (AI-ED 2003), Sidney.

VanLehn, Kurt, Stephanie Siler, Charles Murray, Takashi Yamauchi, and William Baggett. 2003. Why do only some events cause Learning During Human Tutoring? Cognition and Instruction 21(3), pp. 209-249.