

Year 1 Annual Report

Improving Science Learning through Tutorial Dialogs

Submitted to IES, May 2008

Wayne Ward, Ron Cole

Boulder Language Technologies

Year 1 activities consisted of (a) putting infrastructure in place, (b) staffing and training, (c) planning, designing, testing and refining dialogs, (d) corpus development, and (e) developing and integrating software components into the intelligent tutoring system. Knowledge about the project was disseminated to the scientific community through invited talks at scientific meetings and colloquia at universities and through development of a project web site. Six papers by Rodney Nielsen and Wayne Ward, both currently at Boulder Language Technologies, and James Martin at CU have been accepted for presentation and publication. These papers represent work related to the current project that was begun at CU and is being continued at BLT.

Infrastructure

Boulder Language Technologies acquired a Beowulf computing cluster from Penguin Computer, a leading supplier of such systems. The cluster has been loaded with all software required for the project and project staff have been trained in its use. Required licenses and subcontracts are in place. Subcontracts to CU, University of Pittsburgh and Boulder Valley School district will be re-issued after pending transfer of the award from CU to BLT.

Staffing and Training Tutors

In order to develop an intelligent tutoring system to improve science learning, it is necessary to collect and annotate a significant amount of data during dialogs between human tutors and students. The transcribed data are used to train a speech recognizer to recognize the words that students produce during tutoring sessions, to develop natural language processing system to interpret spoken utterances, and to develop dialog models to interpret students' utterances in the context of the ongoing conversation to produce responses by the virtual tutor consistent with learning objectives incorporated into the dialog model. A significant amount of transcribed speech data, collected from at least 50 and preferably 100 students for each science investigation, is needed to develop an initial tutoring system that can be tested and refined in Wizard of Oz (WoZ) experiments. Audio and video data collected during tutoring sessions is also necessary for training and providing feedback to human tutors on their performance using the Questioning the Author (QtA) approach.

Because of the large amounts of data that must be collected and transcribed in a relatively short period of time to meet project milestones, it was necessary to hire and train between 8 and 12

tutors. We advertised for tutors on Craigs List in Denver and Boulder, requiring applicants to have expertise in teaching and tutoring science. We interviewed over 25 applicants, and hired an expert team of project tutors. Each project tutor was either a former science teacher or a science graduate student at the University of Colorado specializing in science education. Eleven tutors were hired and trained, of which 9 are still with the project. Brief bios of project tutors can be found on the BLT web site (www.bltek.com).

All project staff participated in initial meetings and training sessions. These included (a) a kickoff meeting in September 2007 with presentations by senior project personnel on each key component of the project (e.g., project overview, the FOSS science program, Questioning the Author, the process for developing dialogs, the stages of developing, testing and refining the intelligent tutoring system, and assessing outcomes), (b) a two day workshop by Margaret (Moddy) Mckeown explaining the Questioning the Author approach to classroom instruction and how to adapt the approach to individualized tutoring, and (c) two one-day training sessions by Kelly Armitage on classroom instruction using FOSS science investigations for Magnetism and Electricity and for Measurement. Following these initial training sessions, research and development activities followed two major paths: (a) Developing and refining tutoring dialogs and data collection, and (b) Developing the intelligent tutoring system. Activities in each of these areas are summarized below.

In addition to hiring project tutors, we hired a project liaison through the Boulder Valley School District (BVSD) through a subcontract from the grant. The project liaison plays a critical role in the project; this person is responsible for recruiting schools, for working with school principals and teachers to schedule tutoring sessions, for delivering and collecting consent forms, for scheduling project tutors for tutoring sessions, for organizing and inputting the audio and video data recorded during tutoring sessions and for scheduling and administering surveys to teachers and students. Because of delays in implementing the subcontract between the University of Colorado (CU) and BVSD, it was not possible for us to advertise and interview for this position until September 2007. Fortunately, we hired an excellent candidate, Jennifer Borum, who has done an exceptional job at all of the above tasks.

Developing and Refining Tutoring Dialogs

Infrastructure for Developing Dialogs

Boulder Valley School District purchased 8 laptop computers for displaying visuals to students and for collecting audio data from students during tutoring sessions. Boulder Language Technologies purchased 8 video cameras for taping tutoring sessions on DVDs.

Boulder Language Technologies developed software for recording and managing student's speech data during tutoring sessions.

Tutorial Development Process

We identified four people from the set of trained tutors to comprise a core development group. The core group is responsible for developing learning objectives and dialog maps for each science investigation, and for communicating these to the larger tutoring group. The general tutorial development process is:

- The BVSD liaison recruits a set of schools and teachers that will be teaching the target FOSS modules and agree to participate in the tutoring process, and organizes a schedule for tutoring at the schools.
- The core group develops learning objectives, dialog maps and supplementary materials for an investigation. They then meet with the full tutor group to go over the tutorial.
- The tutors go into the schools and tutor students using the materials developed. The student's speech is recorded on a laptop computer and the entire session is videotaped on a DVD.
- The tutor group reviews the session tapes and critiques the presentation, offering suggestions for improvement. A subset of the sessions is sent to Dr. McKeown who reviews them and annotates the transcripts with comments. The tutorial presentations are revised based on the collective feedback.
- The previous 2 steps are iterated for the duration of the presentation of the investigation.

Examples of dialog maps and other supporting materials are displayed in Appendix 1.

Summary of in-school tutoring activities in Year 1

Tutorials were developed and presented for 2 modules as of April 30 2008. The table shows the data collected (tutorial sessions presented) to this point.

	Number of Schools	Number of Students	Number of tutorial sessions
Module 1 (M&E)	3	96	336
Module 2 (Measurement)	1	78	78

Surveys were given to students to give feedback on their experience. Results of these surveys are shown in Appendix 2.

Based on our experiences during year 1, we added a second phase of human tutoring to the development cycle. Our analysis of dialogs demonstrated many instances of misunderstandings and “sticking points” shared by several students that would benefit from the introduction of illustrations, pictures and animations that could be used to “ground” the dialogs. A complete set of animations were designed and refined by the Boulder team in collaboration with Kathy Long at Lawrence Hall of Science. These are now being tested in a second phase of tutoring for Magnetism and Electricity. Our year 2 development plan reflects the addition of a second phase of human tutoring for each FOSS module that incorporates “visuals” into the dialog design. These dialogs will be used to inform the integration of illustrations and animations into the intelligent tutoring system.

Proposed Timelines and Milestones for Tutorial Development

trimester	2	3	1	2	3	1	2
	2007-2008	Spring '08	Fall '08	2008-2009	spring '09	Fall '09	2009-2010
M and E	M&E T1	M&ET2	M&E Woz				
Measurement		M T1	M T2	M Woz			
New Unit 1			Unit1 T1	Unit1 T2	Unit1 Woz		
New Unit 2				Unit2 T1	Unit2 T2	Unit2 Woz	
New Unit 3					Unit3 T1	Unit3 T2	Unit3 Woz

T1 = QtA tutoring with hard copies of visuals

T2 = Tutoring with visuals

Woz = Wizard of Oz

Virtual Tutor System Development

Virtual Human Toolkit Development

In this project, we are building on software developed at CU, called the Virtual Human Toolkit. The VHT had to be extended to support the type of dialog interaction we are using.

- We have extended the system to be able to present flash animations concurrently with having conversational interactions with the student. For example, the system can be presenting an animation illustrating a concept while the student is explaining what is going on in the visual and the speech recognition and dialog management system are decoding what is being said by the student.
- An entirely new dialog manager is being developed that allows a much more conversational interaction about concepts. The original dialog manager was oriented toward gathering facts in a form-filling paradigm. The new model requires representing target propositions and comparing what users say to them in order to generate follow-up actions by the system. The modifications are scheduled to be completed by the end of May 2008.

Corpus Development

The Speech Recognition, Semantic Parsing and Dialog Management components of the system all require user data to develop. The required data comes from the audio and video recordings made of the human tutoring sessions, data collected in Wizard Of Oz sessions and data collected during live use of the system. Audio recordings contain only the student's voice while videotapes have both sides of the dialog, student and tutor. The data collected in the human tutoring sessions is used to create an initial WOZ system. Audio recordings are transcribed and used to train acoustic models and language models for the speech recognizer. The transcripts are also used to develop grammars for the semantic parser. The videotapes are also transcribed, but at a coarser level, not suitable for acoustic training. These transcripts are used to train the dialog model. The WOZ system is used to gather data that more closely models the target interaction, students with virtual tutors. These data are then used to tune the system for fully automatic operation. Data is logged continually and used to adapt the system from time to time.

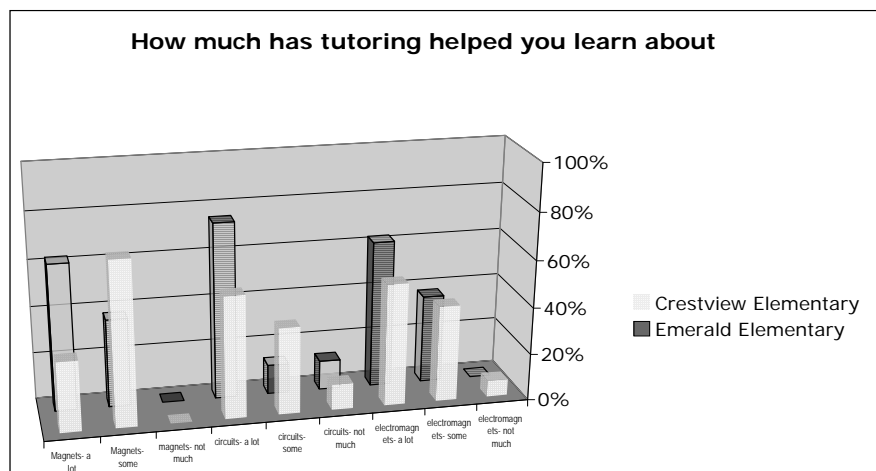
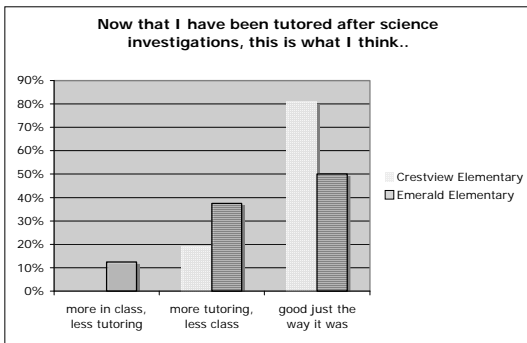
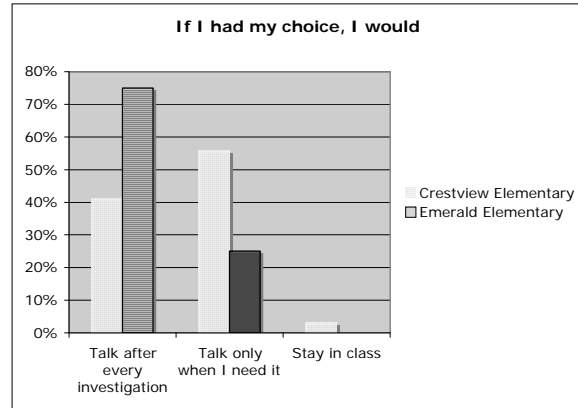
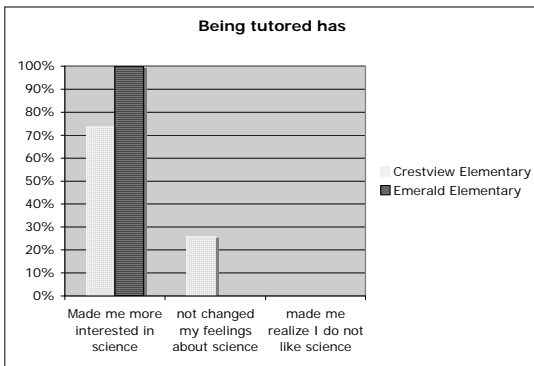
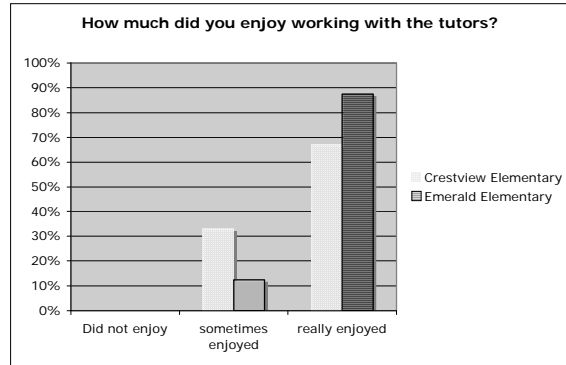
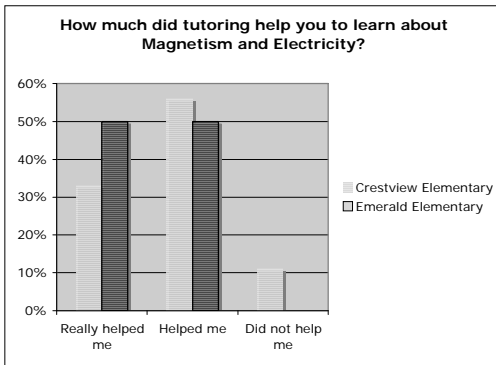
To date we have transcribed 40 sessions from module 1, which is approximately 10 hours of speech data. An additional subset of videotapes has also been transcribed.

System Development

Using the updated Virtual Human Toolkit, we have created an initial prototype for our virtual tutor, including speech recognition input, dialog management, speech output, video output and presentation of flash animations. The actual content is not present yet, just examples to test system operation.

We anticipate completion of the initial Wizard of Oz system at the end of May 2008. The WOZ requires taking the operational tutor system and adding hooks so that a tutor can control operation from a second laptop in cases where the tutor will not generate an appropriate action. The WOZ system will be tested in the lab over the summer and deployed in schools in the fall.

Appendix 2 - Student Survey results



	CRESTVIEW n=27	EMERALD n=8	CRESTVIEW percentage	EMERALD
How much did tutoring help you learn about M and E?				
really helped me	9	4	33%	50%
helped me some	15	4	56	50
did not help	3	0	11	0
How much did you enjoy working with the tutors?				
Did not enjoy	0	0	0	0
sometimes enjoyed	9	1	33	12.5
really enjoyed	18	7	67	87.5
If I had my choice I would:				
Talk after every investigation	11	6	41	75
Talk only when I need it	15	2	56	25
Stay in class	1	0	3	0
Now that I have been tutored after doing science investigations, this is what I think:				
more in class, less tutoring	0	1	0	12.5
more tutoring, less class	5	3	19	37.5
good just the way it was	22	4	81	50

Being tutored has:

Made me more interested in science	20	8	74	100
not changed my feelings about science	7	0	26	0
made me realize I do not like science	0	0	0	0

How much has the tutoring helped you learn about:

Magnets- a lot	8	5	30	62.5
Magnets- some	18	3	70	37.5
magnets- not much	0	0	0	0
circuits- a lot	14	6	52	75
circuits- some	10	1	37	12.5
circuits- not much	3	1	11	12.5
electromagnets- a lot	14	5	52	62.5
electromagnets- some	11	3	41	37.5
electromagnets- not much	2	0	7	0