

IES and NSF Projects: Tutorial Dialogs for Improving Science Learning

Similarities, Synergies and Differences

The purpose of this summary is to outline important similarities, differences, and synergies between two FY07 research grants awarded separately to Boulder Language Technologies (BLT) and to the Center for Spoken Language Research (CSLR) at the University of Colorado – Boulder.

The Grants

IES Grant

In July 2007 the Department of Education's Institute of Education Sciences (IES) Cognition and Student Learning Program awarded a four year grant entitled *Improving Science Learning through Tutorial Dialogs* to the University of Colorado's Center for Spoken Language Research (CSLR). The goal of the project is to improve science learning by 3rd, 4th and 5th grade students in Boulder Valley schools. In May 2008 the grant was transferred from the University of Colorado to BLT, with subcontracts from BLT to CSLR (James Martin, PI), University of Pittsburgh (Margaret Mckeown, PI) and the Boulder Valley School District.

NSF Grant

In September 2007 the National Science Foundation's Division of Education and Human Resources awarded five year collaborative research grants through its K-12 Knowledge Discovery Program entitled *Improving Science Learning in Inquiry-based Programs* to Boulder Language Technologies (Wayne Ward, PI) and the University of Colorado (Sarel van Vuuren, PI). Subcontractors include the University of Pittsburgh (Margaret Mckeown, PI) and Boulder Valley School District.

Similarities: Underlying FOSS Curriculum and General Use of Tutorial Dialogs

The projects share the overarching goal of improving science achievement by elementary school students. The same staff, to a large extent, will carry out the studies.

While the projects have important differences discussed below, the similarities between the projects promise synergies that are likely to produce much greater returns for each sponsoring agency than if only one or the other of the awards succeeded in their respective competitions.

For example, each project involves developing tutorial dialogs in which students interact with a human or a computer tutor. The dialogs will be designed to

- (a) assess students' knowledge and misconceptions about the science being taught in classroom science investigations using the FOSS program;
- (b) help students learn key concepts; and
- (c) use these concepts to reason about, explain and predict scientific phenomena.

Also, in both studies, students will receive tutoring sessions 2 or 3 times per week following classroom science investigations. The overall R & D plan is similar across the projects: develop tutorial dialogs by a team of computer and cognitive scientists, education researchers and teachers, refine the dialogs by observing and analyzing and tutoring sessions between human tutors and students, develop a computer tutor that emulates the behaviors of an effective human tutor, conduct Wizard of Oz experiments to refine the computer dialogs, field test the computer dialogs, and then assess the effectiveness of the tutoring sessions by comparing science achievement of students in classroom control, human tutoring and computer tutoring conditions. Finally, in both projects the expected outcomes will be new knowledge about the effectiveness of specific tutoring approaches (which differ across the projects) to science achievement and the development of intelligent tutoring systems (which differ across the projects) that are designed to improve science learning.

Differences: Specific Dialog Forms, Tutoring Approach, and Assessment

The two projects differ in several ways. The NSF project will develop tutorial dialogs aligned to science investigations in four FOSS modules, two in fourth grades and two in fifth grades. The IES project will develop tutorial dialogs for one third grade science module, two fourth grade science modules and one fifth grade science module.

Dialogs

Although three of the science modules for which we will develop dialogs overlap, the actual dialogs will differ because the NSF supported dialogs will entail one-on-one tutoring, and IES-supported dialogs will involve group tutoring sessions. For these three modules we will be able to examine learning processes and gains in one-on-one tutoring sessions (NSF) versus small group tutoring sessions (IES) conducted by human tutors and intelligent tutoring systems, and examine learning using two different approaches to tutoring.

Tutoring Approach

The tutorial development, structure and data collection accordingly will differ between the two projects. The NSF project divides each 20 minute tutoring session into two 10 minute sessions: acquisition of conceptual knowledge and acquisition of procedural knowledge. Following Bybee, conceptual knowledge refers to developing accurate mental models of the key constructs and concepts related to an area of science. Procedural knowledge refers to acquiring the ability to use these concepts to reason about scientific phenomena to explain observations and generate hypotheses to predict outcomes of experiments. Initial tutoring sessions in the NSF project will focus on assessing each student's understanding of concepts by developing concept inventories, instruments that identify the different mental models associated with concepts. These include accurate mental models, and specific misconceptions (inaccurate mental models) that prevent deep learning of science. Concept inventories are developed by presenting students with questions (e.g., Why is it warmer in summer?) and then analyzing and clustering explanations to identify the 2 or 3 most common misconceptions. Concept inventories are presented as multiple choice questions, with answer choices indicating accurate and inaccurate mental models; answer choices are then used to guide the subsequent dialog, which may include multimedia presentations (e.g., narrated animations that illustrate concepts) followed by question – answer dialogs designed to help students understand the inadequacy of inaccurate mental models to explain specific scientific phenomena. The second half of each tutoring session consists of conversational question and answer dialogs designed to make students think and reason about the concepts they have learned to generate explanations and predictions using the language of science. Answers to questions during this phase of tutoring can also reveal

lingering misconceptions that are addressed through additional review of concepts and focused dialogs.

The tutoring in the IES project takes a qualitatively different approach, by which it consists of a continuous question –answer dialog during the entire 20 minute tutoring session, with multimedia presentations invoked when needed based on the student’s responses to questions. Students are asked open ended questions and their responses are analyzed to determine if their answers are correct and complete, or whether they reveal possible gaps in knowledge or misconceptions. When the latter are encountered, the human or virtual tutor asks another question designed to stimulate the student to recognize missing or inaccurate knowledge and generate an answer that is more accurate and complete. Hints or multimedia presentations are provided when the student is thrashing, followed by additional open ended questions. Questions become more sophisticated as the dialog continues, moving from questions that require complete and accurate answers, to explanations and predictions of new phenomena using key concepts. In the NSF project, all tutoring sessions are conducted with one tutor (human or virtual) working with one student. In the IES project, one-on-one tutoring sessions are conducted only during the initial development of the tutorial dialogs. Tutoring sessions during field testing and assessment of the dialogs will occur in groups of three to four students. Students will discuss answers to questions, with students taking turns speaking to the tutor.

Assessing Effectiveness of Tutoring Dialogs

The differences in dialog structure and data collection naturally lead to significantly different assessment strategies in the respective projects. In the NSF study, where all tutoring is one-on-one, the assessment compares performance of students in individual classrooms randomly assigned to three conditions: a) classroom controls, b) individual tutoring with a human tutor, and c) individual tutoring with a virtual tutor. Additional details about the NSF Project’s assessment design can be found in the NSF proposal on the BLT Web site.

The IES project focuses on developing tutorial dialogs using the Questioning the Author (QtA) approach. This approach was developed to improve comprehension of texts through classroom discussions, and has been demonstrated effective in independent evaluations. QtA uses open ended questions and follow up queries

to stimulate students to think about what the author is saying and to make connections with other portions of the text, prior classroom discussions and other life experiences. Thus, dialogs in the IES study are designed to maintain the key features of QtA dialogs using open ended questions and follow up queries that stimulate thinking and reasoning, but to adapt the approach to science investigations, where the “author” is the science investigation (or perhaps Mother Nature) and the “text” is the set of observations that students must interpret. Since QtA dialogs were designed by Beck and McKeown as a classroom discussion technique, with ideas contributed by many students as integral to the learning process, our assessment of the tutorial dialogs developed in the IES project for each FOSS module compares learning by students in large and small groups. The four conditions are a) a classroom control condition, b) in classroom groups of 16 students with discussions facilitated by a trained QtA tutor, c) in small group dialogs of 3 or 4 students with a trained QtA tutor, and d) in small group dialogs of 3 to 4 students with the virtual tutor. The experimental design uses random assignment of students in paired classrooms in the same school to these groups. This design enables us to compare tutoring in large (N =16) versus small (N=3) groups, and to compare human tutoring to dialogs with a virtual tutor for small groups. Learning gains are measured using FOSS benchmark assessment and state tests administered to fifth grade students.

Synergies of Complementary Projects

These projects have sufficient similarity and variation to promise benefits for each agency and for the research community that were not originally envisioned. The most immediate and obvious benefit is that, given the common goals of developing time-intensive tutorial dialogs in both studies, significantly more effort can be devoted to planning and developing them. Although the dialogs will differ, the issues involved in creating them are similar in the two projects. This allows us to increase the amount of effort devoted to research, leading to greater effectiveness and knowledge-generation returns for both projects. Not only do we expect the research to be more effective, but also more efficient. For example, project tutors will be hired as a single group and used efficiently across the two projects. There will also be some savings in infrastructure costs, and more focused effort devoted to infrastructure development, resulting in more effective tools.

There are many ways in which synergies will be realized. For example, the same virtual human toolkit will be used in both projects, as well as some commonality in the software for Wizard of Oz studies. Since there is overlap on three of the five science modules, videotaped data of tutoring sessions can be managed, transcribed and analyzed more efficiently and less expensively. Because much more speech and dialog data will be collected across the two studies, the speech recognition, natural language understanding and dialog modeling and concept inventories will be more accurate.

In addition to the savings from the points of overlap in the two projects, there are perhaps even greater benefits to be derived from the differences. For example, the multimedia presentations designed to present concepts in optimal ways developed in the NSF study can be inserted into tutorial dialogs as needed in the IES study. The development of concept inventories in the NSF study produce instruments (that look like multiple choice questions) with answer choices that reveal mental models associated with common misconceptions in science. These concept inventories will be invaluable in developing effective dialogs in both NSF and IES studies. The differences between the studies will also reveal significant new knowledge about the benefits and challenges of different tutoring approaches and the benefits and challenges of tutoring one-on-one verses small groups.

In short, both the overlaps and differences between the two projects combine to increase the likelihood of successful outcomes, the acquisition of new knowledge and the development of new technologies distributed in new toolkits to the research community.