Robot Facilitation as Dynamic Support for Collaborative Learning

Naomi Miyake, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan, nmiyake@p.u-tokyo.ac.jp
Sandra Y. Okita, Teachers College, Columbia University, 525 West 120th Street, New York, NY 10027, okita@tc.columbia.edu

Abstract: Automated as well as remotely controlled robots have high potential to expand our research and practice in collaborative learning environments. The area of computer supported collaborative learning (CSCL) is one in which advanced technologies have been used effectively to automatically monitor and dynamically structure group discussions for learning. Traditional desktop environments for collaborative learning have included interactions between groups and computer agents with conversational abilities. Robots can take the collaboration out of the computer and into the three dimensional, real world environments of students. This symposium raises the important question of how advanced technologies, such as robotics, may enable us to take the same insights from the area of scripted collaboration and use them to support groups of learners in these environments. Other topic areas include design choices in robotic features that may influence scripted interactions, social behavior, meaningful engagement, and conditions that dynamically support situational factors and generate collaborative learning among students.

Introduction

As collaborative learning has proven to be effective in helping students learn (Chi et. al, 2001; Graesser, Person, & Magliano, 1995; Palincsar & Brown, 1984), it is only natural to ask how sophisticated artifacts can be used as facilitating tools to dynamically support collaborative learning. Conversational agents, avatars, and humanoid robots are highly directable, allowing researchers to capitalize on a dynamic nature that enables humans to share knowledge and ideas. Similar to the use of conversational agent technologies in desktop CSCL environments, having a robot participate in collaborative learning allows researchers to control some part of the collaborative activities, which enables us to better understand the basic mechanisms of collaborative learning. This work has built on theories and findings from a wide range of areas of the Learning Sciences as well as the technical fields of Machine Learning, Language Technologies, and Human-Computer Interaction. Some of the most cutting edge work in this area has been the use of robots to support collaborative learning.

The symposium takes a close look at three areas that may contribute to the unique value of robots in assisting collaborative learning and assessment. One area that is critical in the design of meaningful engagement is scripted collaboration. In the past decade, the area of scripted collaboration has produced tremendous gains in terms of insight into how to elevate the occurrence of valuable discussion-based learning behaviors from students in online and offline settings (Dillenbourg, 2002; Kollar, Fischer, & Slotta, 2005). This insight has also led to advanced technologies, such as robotics, that may enable us to take these findings and use them to support groups of learners in the three-dimensional world. More recently, advances in the area of automatic collaborative learning process analysis have enabled the first attempts at dynamic support for collaborative learning, which builds on the insights learned from static forms of scripted collaboration (Rosé, Wang, Cui, Arguello, Stegmann, Weinberger, & Fischer, 2008).

The second area takes from the theories and findings of Human-Computer Interaction (HCI) and Human-Robot Interaction (HRI). This area examines how specific features in robots may influence; how interactive scenarios are delivered to humans (e.g., with an angry voice, with gesture), how behavioral scripts based on familiar narratives and routines can be generated from humans (e.g., common children play routines like tea party, play doctor), and how robotic features may influence the quality of interaction (e.g., testing the robot or meaningful engagement). In examining the situational factors that promote collaborative learning, much attention has been given in making machines more responsive and sophisticated. The advancement in sensors and audio-visual tools has helped robots detect human behavior (e.g., facial expressions), while the sophisticated automation and expressive tools have helped robots provide dynamic and expressive responses toward humans (e.g., gesture, tone of voice, attention). Identifying the delicate balance between detection and response has improved the overall quality of human-robot interaction. However, little research has been done to examine how these features when combined with specific scripts and scenarios can assist collaborative learning and meaningful engagement.

The final area of interest in this symposium is taking the insights from the first two areas (i.e., scripted collaboration and human-robot interaction) and exploring conditions that dynamically support and generate the optimal level of collaboration among groups of students. Robots and other agent technologies can participate in
and support reflection in active ways, for example by providing learners with precise replay of past learning experiences upon request, which has not been as feasible with human support. In addition to these capabilities, however, the use of robots goes beyond that, allowing us to implement basic design principles of group activities in more tangible ways. They offer the unique opportunity to support groups of students in real world activities. Automated as well as remotely controlled robots have high potential to expand our research and practice in collaborative learning. As a learning collaborator, robots can present their prepared explanations to their human partners, comment on others’ ideas, and exchange questions and answers, at least to the extent those exchanges could be either prepared prior to the classes or provided by human operators. As a research partner, robots can run controlled experiments during a collaborative learning process in classroom-like settings. One example of such control is to have a robot in each discussion group and to have it deliver the same information to different groups, so that we can explore the effects of a particular discourse in collaboration.

Reflectively it is illuminating that the various effects of this kind of dynamic and interactive support have not been carefully researched in the past. Collaborative learning has been practiced widely in the community of learning scientists, with relative to profound change in the quality of learning achievements (Bransford et al., 1999; Sawyer, 2006). In order to better design such practices within a wider variety of cultural settings, we need stronger scientific evidence for such success, as well as a more precise explanation of the mechanisms of how and why it works (cf., Miyake, 2008). One of our motivational factors for proceeding in this direction is the need for us to better equip ourselves with stronger sets of evidence of the power of collaborative, learner-centered orientation in education, to work effectively with students in the collaborative setting.

**Target Audience**

Our audience would include a wide range of learning science researchers who seek to evolve research on collaboration. We particularly aim this symposium at practice-oriented, classroom-based research, from which we could learn how to build sustainable, effective communities, to foster impacts on real world learning. We are also keen to invite engineers and robotics researchers seeking focused research fields to foresee what is needed for the creation of symbiotic robots. Collaborating with them would open up a stronger promise for us to reach out to real world classrooms.

**Symposium Presentations**

The symposium presentations will focus on three topic areas: 1) dynamic collaborative learning support techniques with scripted collaboration, 2) influence of robotic features on interactive scenarios and facilitating engagement, and 3) exploring conditions that dynamically support situational factors that generate the optimal level of collaboration among groups of students.

The research in the first presentation covers interactions that are text-based in a two-dimensional chat environment with group interactions, but the architecture and dynamic collaborative learning support techniques are highly applicable to physical robots environment. The research in the second presentation involves a close examination of physical humanoid robots and humans in a one-to-one (robot-to-human) interaction (not in groups), but the findings has helped identify important perceptual cues and responses that need to be recognized and handled respectively that are applicable to collaborative groups settings. The research in the final presentation involves multiple physical humanoid robots spread across multiple student groups and addresses many of the common insights and issues from the first two presentations. To help highlight the three interest areas, we discuss them separately, but all three are at play in each presentation.

The symposium consists of two parts. The first part involves presentations of recent work that shows promising results, strengths of, and challenges in robot facilitation. The second part of the session leads into a discussion organized by discussants that engage presenters and audiences to explore practical use of robots and agents/avatars in collaborative learning (e.g., how different situational factors require different dynamic support, and how robots may facilitate in each setting).

**Presentation 1**

**Basilica 3D: Towards Architecture for Operating Robotic Accountable Talk Facilitation Agents**

Carolyn Penstein Rosé  
Language Technologies Institute and Human-Computer Interaction Institute  
School of Computer Science, Carnegie Mellon University  
cprose@cs.cmu.edu
While theories of discussion-based learning both within the computer supported collaborative learning (CSCL) community as well as in the classroom discourse community are many, there is consensus about what types of group interactions are desirable. For example, the value placed on taking ownership of reasoning, articulation of reasoning, evaluation of reasoning, and building on the reasoning of self and others are almost ubiquitous. Yet it is widely acknowledged that groups do not operate at an ideal level without support. Thus, researchers in the area of scripted collaboration have worked to develop design principles to guide development of support that elicits the kind of group behaviors that are valued within the CSCL community. However, while much of that work was developed and evaluated with static forms of support, there has been a growing interest in making that support dynamic, both in the sense of triggering based on real time analysis of the ongoing collaboration (Rosé et al., 2008; Mu et al., 2011), as well as in the sense of support that is itself interactive (Kumar et al., 2006).

Conversational agents have a long history of successful support for individual learning with technology (Rosé et al., 2001; Rosé & Van Lehn, 2005; Kumar et al., 2006). A series of results offer hope that they can be used productively to offer support for collaborative learning, especially in chat environments (Kumar & Rosé, 2011; Kumar et al., 2010; Kumar et al., 2011; Chaudhuri et al., 2009; Ai et al., 2010; Howely, Mayfield, & Rosé, 2011). The Basilica architecture has been developed for the purpose of enabling efficient development of intelligent agents with rich conversational behaviors to participate with groups of students in online chat environments and to play a facilitative role (Kumar & Rosé, 2011). Early work in this area simply imported the same conversational agents used in individual learning with technology into group learning settings. However, more recent work has directly employed techniques from the classroom discourse community to develop conversational agents that can employ to some limited extent the practices of classroom facilitation techniques.

In early work on dynamic support for collaborative learning, agent behaviors were triggered through analysis of text-based interactions in the chat environment. Recent work on monitoring group interactions through speech (Gweon et al., 2011) gives hope that this dynamic collaborative learning support technique can be adapted from the two-dimensional text-based chat environment to the three-dimensional, face to face collaboration setting where the facilitation may be conducted by robots controlled through this same architecture. This talk takes a visionary look at how this Basilica work may be extended for this purpose.

**Presentation 2:**
**Multimodal Approach to Facilitating Affective Human-Robot Interactions**

Sandra Y. Okita  
Communication, Computing, and Technology in Education  
Teachers College, Columbia University  
okita@tc.columbia.edu

The advancement in sensors and audio-visual tools has helped robots detect human behavior (e.g., facial expression), while sophisticated automation and expressive tools have helped robots provide dynamic and expressive responses toward humans (e.g., gesture, tone of voice, attention). Identifying the delicate balance between detection and response has improved the overall quality of human-robot interaction. However, little research has been done to examine how these features when combined with specific interactive scenarios can facilitate meaningful and affective engagement.

The research shares some preliminary findings on how specific features in robots may influence how interactive scenarios are delivered to humans (e.g., with an angry voice, exaggerated gesture), how familiar scenarios and behavioral scripts can be generated from humans (e.g., hide-and-seek), and how robotic features may influence the quality of interaction (e.g., child testing the robot but not communicating with the robot), affective engagement (e.g., robotic voice scares child) and learning (e.g., peer-like robot vs. teacher-like robot). Having physical humanoid robots in the collaborative learning setting provides researchers with a range of design choices (e.g., tone of voice, attention, gesture, detection and response timing) on how interactive scenarios can be delivered. For example, is it the content of the interaction or the timing of feedback that is more important? If the interactive scenario is accompanied with robotic gestures (Ng-Throw-Hing, Luo, & Okita, 2010; Okita, Ng-Throw-Hing, & Sarvadevabhatla, 2009), will the interpretation be different? Will interaction be more affective if the robot maintains longer eye contact, pay more attention, mimic human gestures, or stand more closely to students? As some features may be common across different platforms (e.g. computer desktop, virtual reality environments, robot) some features are unique to the physical environment (e.g., proxemics). Examining these characteristics can be potentially useful when incorporating the findings into designing interactive and dynamic scenarios for collaborative learning.

In the first line of work, we applied familiar play routines when interacting with robots (e.g., turn-taking scenario, setting up the dinner table). Often times, the child’s ability to pretend or engage in
collaborative play was constrained by what the robot could do in response (e.g., timing and limitation of response). In another line of work we found that when the robot activated a familiar schema (e.g., going to the zoo, story-telling time, The Three Little Pigs), the child sustained a social relationship with the robot (Okita, Ng-Thow-Hing, & Sarvadevabhatla, 2011). Children seemed to start out by drawing on prior knowledge through which they decide the “can and cannot do interactions” with robots (Okita & Schwartz, 2006). The studies helped identify important perceptual cues and responses that needed to be recognized and handled respectively to improve the quality of interaction. Until robots have the intelligence to flexibly respond to a wide range of interactive bids, following a well-known script or scenario seemed quite successful in guiding the interaction. In the third line of work, we examined whether familiar game scenarios (e.g., children's game "Captain May I?") could influence physical distance (e.g., proxemics) between human and robots (Okita, Ng-Thow-Hing, & Sarvadevabhatla, 2012). Current limitations in sensor technology made physical distance a crucial factor to detect and respond to situations (e.g., avoid collision during collaborative tasks), user state (e.g., facial expression, speech recognition), and better choice of expressive communication (e.g., verbal or gesture).

Presentation 3:
Robots facilitate “constructive listening” for strengthening individualized learning in collaborative learning situations

Naomi Miyake
Consortium for Renovating Education of the Future,
School of Education
The University of Tokyo
nmiyake@p.u-tokyo.ac.jp

Remotely operable robots in collaborative classrooms can serve at least two roles. One is to work with children as “a learning friend” who helps members in the class enjoy and stay engaged in dynamic learning, where each individual member promotes her/his own understanding. The second role is to work with researchers as a data-collector and a reflector of our own facilitating behavior so that we can learn both the basic mechanisms and the design principles of productive activities in collaborative classrooms. In this presentation I will focus on the former, to illustrate what kind of new research can be developed by using robots, and then lead the way toward the second role.

In order to establish the research context for this study, we have devised a strongly scripted yet dynamically collaborative learning situation based on the Jigsaw method. We call this framework the Knowledge Constructive Jigsaw, where it is emphasized that each individual student is responsible for integrating perspectives given by the learning materials and from other students with their own understanding. The class design involves a shared question to be answered and some relevant learning materials from different perspectives distributed among the different groups first in expert groups, to be later exchanged and integrated to answer the question in the jigsaw groups (Miyake, 2011). In accordance to some basic mechanisms of constructive interaction (e.g., Miyake, 1986), the design naturally requires each student to become a task-doer in the jigsaw group. It also provides each student with a chance, or chances, to be a monitor who infers what the other students say and why they say that, to integrate others’ ideas with their own. Yet the proportion of being the doer could differ from individual to individual. We now have data from nearly thirty classes of different grade levels on different topics. The analyses of these classes of middle schools on different science topics have revealed that there is little correlation between the achievement levels and the proportion of the role exchange. Rather, the monitors, who could spend almost the entire class without “speaking up,” learned a lot from just attending to others’ talks and inwardly working to integrate such inputs into their own understandings. This result suggests the importance of “constructive monitoring,” similar to the notion of “constructive listening” proposed by Greeno and van de Sande (2007), for better analyses and design of the interactive learning processes (Damsa et al., 2010).

In a series of follow-up studies of this construct, we used remotely operable robots so that we could control the levels of soliciting constructive monitoring. We used the established learning plans and their associated materials with new groups of children and a robot, to see whether we could recreate similarly successful classes with robots as facilitators, and also to see what kind of conversational cues the monitoring children would use to start taking responsibility of their own learning. The basic analyses of these classes have revealed that children around 10 year olds have a keen sense of distinguishing whether the robot “knows” the answer but does not say it or the robot is “just like the other kid who does not know the answer, but sincerely working to know the answer.” When the robots are accepted as the latter, the children had a better chance to get involved in the constructive conversation, resulted in better learning (Miyake, 2011; Miyake, et al., 2011; Oshima, et al., 2011).
Discussion Topics (Tentative)

1. Comparison with alternative technologies.
How robots can enhance spontaneous, creative collaborative discourse among learners without interrupting or distracting from the inter-personal interaction (Gerry Stahl)

How can robots create values beyond existing devices for learning and collaboration?
Can we use existing measures for collaborative learning to evaluate this type of learning partnership?

2. Pragmatic considerations of deployment to formal and informal learning environments.
What needs to be considered when taking robots out of the lab and into real world (e.g., user support).

Discussant 1:

Gerry Stahl
College of information Science and Technology
Drexel University
Gerry.Stahl@drexel.edu

Area of Expertise:
Dr. Stahl is trained in computer science, artificial intelligence, social philosophy, cognitive science, and learning science. Since 2002, he has taught human-computer interaction (HCI), computer-supported collaborative learning (CSCL), computer-supported cooperative work (CSCW) and social informatics (SI) at Drexel University. Dr. Stahl’s research approach includes theory building, system development and empirical studies of software usage. He has developed software systems and prototypes to explore support for collaborative learning, design rationale, perspectives and negotiation. His theory combines various sources from philosophy, education, sociology, communication and anthropology. He has developed a methodology of fine-grained empirical investigation into how groups of people learn to use artifacts like groupware systems in real-world settings such as school classrooms and virtual math teams. Dr. Stahl is a world-class researcher in CSCL, having organized international conferences, founded an international journal, published a volume on Group Cognition in MIT Press and one on Studying Virtual Math Teams in Springer Press and written over 150 professional papers. His website and blog are major resources for the CSCL research community.
Retrieved excerpts from Drexel University faculty www site (November 7, 2011).

Discussant 2:

Frank Fischer
Munich Center of the Learning Sciences
Ludwig-Maximilians-Universität München
Frank.fischer@psy.lmu.de

Area of Expertise:
Frank Fischer earned his Doctorate in Psychology in 1997, and his Habilitation (professorial dissertation) in Psychology and Educational Science in 2002, both from the University of Munich. He is a Full Professor of Educational Science and Educational Psychology at the University of Munich. Since 2008 he is serving as the Director of the Department of Psychology at this university. Since 2009, he is the speaker of the Munich Center of the Learning Sciences, an interdisciplinary collaboration of more than 30 research groups focusing on advancing research on learning “From Cortex to Community”. In this context, he is also directing the Doctoral Training Program “Learning Sciences”. He is currently a member of the Board of Directors of the International Society of the Learning Sciences. His own research revolves around learning and instruction, with projects on collaborative learning, problem- based learning as well as inquiry and simulation- based learning. An overarching question is how technology-enhanced learning environments can advance knowledge and skills of learners in school, higher and further education. With respect to methodology, he has been contributing to the development of use- inspired basic research approach in the field of learning and instruction. He has published more than 100 journal articles and book chapters, and co- edited 6 books and special issues of scientific journals.
Implications

In order to better design such practices within a wider variety of cultural settings, we need stronger scientific evidence for such success, and more precise explanation of the mechanisms of how and why things works. The studies in the presentations provide some portrayal on how students interact, engage in conversation, respond to, and work with robots and agents in a collaborative learning setting. This has some practical importance, as we hope to examine whether or not advanced technologies have promising roles for students in collaborative learning.

The recent findings in the presentations, and the insight and expertise from the discussants, is sure to generate a strong discussion on the pros and cons, challenges and promises in using such artifacts in formal and informal learning environments. What is most crucial in such an endeavor is have an opportunity to gather, exchange ideas, rack brains, and plan future steps that will lead to successful research and implementation.

References


