Engaging learning groups using Social Interaction Strategies

Rohit Kumar **Carolyn P. Rosé**
Language Technologies Institute
Carnegie Mellon University, Pittsburgh, PA, 15213
rohitk@cs.cmu.edu cprose@cs.cmu.edu

Abstract

Conversational Agents have been shown to be effective tutors in a wide range of educational domains. However, these agents are often ignored and abused in collaborative learning scenarios involving multiple students. In our work presented here, we design and evaluate interaction strategies motivated from prior research in small group communication. We will discuss how such strategies can be implemented in agents. As a first step towards evaluating agents that can interact socially, we report results showing that human tutors employing these strategies are able to cover more concepts with the students besides being rated as better integrated, likeable and friendlier.

1 Introduction

Conversational Agents (CAs) are autonomous interfaces that interact with users via spoken or written conversation. One of the applications of CAs is tutoring. Various research groups have developed tutoring agents in domains like reading, algebra, geometry, calculus, physics, computer literacy, programming, foreign languages, research methods and thermodynamics. Many of the evaluations show that CAs can be effective tutors (Arnott et. al., 2008; Kumar et. al., 2007; Graesser et. al., 2005).

Most systems that use CAs as tutors have been built for learning scenarios involving one student. Evaluation of learning technologies involving students working in groups with interactive agents has shown that learners are helped both by learning as a group and receiving tutorials from agents (Kumar et. al., 2007). However, some previous studies have reported that students learning in groups ignore the tutor's messages, unlike the case where students are individually tutored. Groups are more likely to abuse tutors than individual students.

We reason that the presence of other students in collaborative learning scenarios causes the agents to compete for the attention of the students. Since the agents are not adept at performing social interactive behavior, which makes up the bulk of formative communication in a group, they are quickly pushed to the periphery of the group.

Research on small group communication has identified twelve interaction categories that are commonly observed in small groups (Bales, 1950). These categories are broadly classified into task and social-emotional categories. Content presented by most current CAs mostly classifies under the task categories. In section 2, we will list the conversational strategies motivated from the three positive social-emotional interaction categories. Thereafter, the implementation and evaluation of a CA that interleaves these social interaction strategies while executing a task plan will be described.

2 Social Interaction Strategies

Balesian methodology (Bales, 1950) identifies three positive social-emotional interaction categories: showing solidarity, showing tension release and agreeing. Participants contribute turns of these categories to address the problems of re-integration, tension release and decision respectively. We have mapped these categories to practically implementable conversational strategies. This mapping is shown in table 1 ahead.

Each strategy is implemented as an instantiation of a conversational behavior. Most of the strategies listed in Table 1 are realized as prompts, triggered by rules based on agent plan, discourse and context features. For example, strategy 1e is triggered

Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL, pages 677–680,
Los Angeles, California, June 2010. ©2010 Association for Computational Linguistics

when one or more students in the group are found to be inactive for over 5 minutes. In this event, the tutor chooses to raise the status of the inactive students by eliciting contributions from them through a prompt like: *Do you have any suggestions Mike?* More implementation details of these strategies and triggers are discussed in the following section.

1. Showing Solidarity *Raises other's status, gives help, reward*
1a. Do Introductions *Introduce and ask names of all participants*
1b. Be Protective & Nurturing *Discourage teasing*
1c. Give Re-assurance *When student is discontent, asking for help*
1d. Complement / Praise *To acknowledge student contributions*
1e. Encourage *When group or members are inactive*
1f. Conclude Socially
2. Showing Tension Release *Jokes, laughs, shows satisfaction*
2a. Expression of feeling better *After periods of tension, work pressure*
2b. Be cheerful
2c. Express enthusiasm, elation, satisfaction *On completing significant steps of the task*
3. Agreeing *Shows passive acceptance, understands, concurs, complies*
3a. Show attention *To student ideas as encouragement*
3b. Show comprehension / approval *To student opinions and orientations*

Table 1. Social Interaction Strategies for three social-emotional interaction categories

3 WrenchTalker: Implementation

WrenchTalker is a CA we have built to employ the social interaction strategies listed in section 2. It helps teams of engineering students learn and apply basic concepts of mechanical stress while they participate in a freshmen lab project to design an aluminum wrench. Students can interact with this agent using a text-based chat environment.

The agent is built using the Basilica architecture (Kumar and Rosé, 2009). Under this architecture, CAs are modeled as a network of behavioral components. There are three types of components: actors (actuators / performers), filters (perceptors / annotators / cordinators) and memories. Figure 1 below shows a simplified depiction of the WrenchTalker component network.

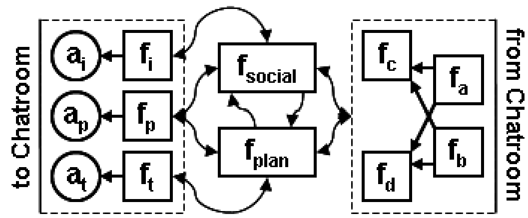

Figure 1. Component Network of WrenchTalker

Three of the actor and filter components correspond to three observable behaviors of the tutor, i.e., Introducing (a_i, f_i), Prompting (a_p, f_p) and Tutoring (a_t, f_t). Most of the other filter components form a sub-network that annotates turns with applicable semantic categories, accumulates them to identify inactive students and generates events that regulate the controllers.

The plan controller (f_{plan}) is responsible for executing the agent's interaction plan, which is comprised of 37 steps. The plan is executed largely sequentially; however the plan controller can choose to skip some steps in the interest of time. In the experiment described in section 5, the same plan controller is used in all three conditions. The social controller (f_{social}) implements the 12 strategies listed earlier. The strategies are triggered by rules based on combinations of three conditions: the last executed plan step, semantic categories associated with the most recent student turns and the ratio of tutor turns generated by f_{social} to f_{plan}. The first two conditions attempt to ensure that social behavior is suitable in the current conversational context and the third condition regulates the amount of social behavior by the CA.

The plan and social controllers are connected so that they regulate each other. For instance, when the plan controller is working, it blocks f_{social}. Upon completion of the blocking step, f_{social} is given control, which can then choose to perform a strategy by blocking f_{plan} before it progresses to the next step. Reflex strategies like 1b are not blocked.

Once the controllers determine a step or a strategy that is to be generated, the actors generate their turns. For example, strategy 1a is generated by actor a_i after it is triggered by the social controller.

We note that Basilica provides the flexibility to build complicated pipelines, as demonstrated in this case by the use of two controllers.

4 Related Work

To contextualize our research with other work on CAs, we classify agents with the social interaction strategies listed in Table 1 as *social interfaces* following the taxonomy proposed by Isbister (2002). Within this class of CAs, researchers have investigated the technical challenges and effects of conversational behavior that are similar in motivation to the ones we are exploring. Bickmore et. al. (2009) report that users found agents with autobiographies, i.e., back stories in first person more enjoyable and they completed more conversations with such agents. Dybala et. al. (2009) found that agents equipped with humor were evaluated as more human-like, funny and likeable. In a multi-party conversational scenario, Dohsaka et. al. (2009) found that an agent's use of emphatic expressions improved user satisfaction and user rating of the agent. We note that use of CAs as social interfaces has been found to have effects on both performance and perception metrics.

5 Experimental Design

In order to evaluate the effect of social interaction strategies listed in Table 1, we designed an experiment with three conditions. In the experimental condition (Social), students interacted with an agent that was equipped with our social interaction strategies, unlike the control condition (Task). In the third condition, a human tutor was allowed to intervene while the students interacted with a Task agent. In all three conditions, students go through the same task plan. However, the degree of social performance is varied from minimal (Task) to ideal (Human). We hypothesize that the human and social agents will be rated better than the Task agent.

We conducted a between subjects experiment during a freshmen computer aided engineering lab. 98 students participated in the experiment, which was held over six sessions spread evenly between two days. The two days of the experiment were separated by two weeks. Students were grouped into teams of three to four individuals. Students were grouped so that no two members of the same team sat next to each other during the lab, to ensure all communication was recorded. The teams were distributed between the three conditions.

Each session started with a follow-along tutorial of computer-aided analysis where the students analyzed a wrench they had designed earlier. The experimental manipulation happened during a collaborative design competition after the tutorial. Students were asked to work as a team to design a better wrench considering three aspects: ease of use, cost and safety. Students were instructed to make three new designs and calculate success measures of each of the three considerations. They were also told that a tutor will help them with two designs so that they are well-prepared to do the final design. No additional details about the tutor were given. The students communicated with each other and with the tutors using ConcertChat, an on-line environment that provides text-based instant messaging and workspace sharing facilities.

After spending 30-35 minutes on the design competition, each student filled out a questionnaire. It was comprised of eighteen questions on a seven point Likert-scale ranging from Strongly Disagree (1) to Strongly Agree (7). The questions were designed to elicit four types of ratings.

- Ratings about the tutor
- Ratings about the other team members
- Ratings about the design task
- Ratings about the team functioning

The questions in the first two classes elicited perceived liking and integration and checked whether the students noticed the tutor's display of the social interaction strategies. Task related questions asked about satisfaction, perceived legitimacy and discussion quality.

6 Results

Table 2 below shows the mean values for questionnaire categories apart from ratings about team members, since there were no significant effects related to those questions.

	D1	D2		T	S	H
Integration	3.85	3.94		3.03	*3.94*	**<u>4.77</u>**
Liking	3.68	3.63		2.78	3.53	**<u>4.73</u>**
Friendly	5.13	5.43		4.47	**5.56**	**<u>5.83</u>**
T.Releasing	4.49	4.63		3.84	4.61	**<u>5.27</u>**
Agreeing	4.30	4.45		3.97	4.44	4.73
Satisfaction	4.66	**<u>5.77</u>**		5.09	4.75	5.97

Table 2. Mean outcomes per condition ((T)ask,(S)ocial, (H)uman) and per day (Day1, and Day2)

The means are highlighted appropriately (**<u>p<0.001</u>**, **p<0.05**, *p<0.08*) to indicate significant differences from Day1 to Day2 and between the Task condition and each of the other two using a pairwise Tukey comparison.

First of all, we note that there is a significant difference in task satisfaction between the two days. We fine-tuned the timing parameters of the plan controller after day 1 so that the students had sufficient time to follow along with each of the steps. This was particularly useful for the task condition where the steps would be executed rapidly due to lack of regulation by the social controller.

On the right side of Table 2, we notice that the human tutors (H) were rated higher on being part of the team (Integration), being more liked, being friendlier and keeping the group more socially comfortable (T.Releasing). On the other hand, the social tutors (S) were rated to be friendlier and were only marginally better at being seen as part of the team.

	Strategy	**Social**	**Human**
Introducing	**1a**	2.67	3.80
Friendly	**1b-1e**	5.61	8.10
Concluding	**1f**	0.97	1.80
T.Releasing	**2a-2c**	5.81	1.77
Agreeing	**3a-3b**	1.78	4.90
Sum		16.83	22.17

Table 3. Mean counts of social turns by tutor

Note that human tutors were restricted to exhibit only social behaviors, which were displayed in addition to the same task related content given to students in the other two conditions. Clearly, the human tutors were better at employing the social interaction strategies. To further investigate this, we compare the number of turns corresponding to the broad categories of strategies in Table 3. Human tutors performed significantly more ($p<0.001$) social turns than the automated tutors in all strategies except showing tension release.

7 Conclusions

In order to make CAs that can participate in multi-party conversational scenarios, the agents must be able to employ Social Interaction Strategies. Here we have shown that the human tutors that use these strategies are better integrated into the group, and are considered more likeable and friendlier. These tutors also cover more steps and concepts and take less time to tutor the concepts, suggesting that the students are more engaged and responsive to them. On the other hand, automated tutors that employ these strategies in our current implementation do not show significant differences compared to task tutor.

We note a contrast between the performance of the human and the automated tutors with respect to the frequency with which they employ these strategies. Besides the frequent use of these strategies, we believe human tutors were better at identifying opportunities for employing these strategies, and they are able to customize the prompt to better suit the discourse context.

Acknowledgments

The research was supported by NSF grant number DUE 837661

References

Elizabeth Arnott, Peter Hastings and David Allbritton, 2008, Research Methods Tutor: Evaluation of a dialogue-based tutoring system in the classroom, *Behavior Research Methods*, 40 (3), 694-698

Robert F. Bales, 1950, *Interaction process analysis: A method for the study of small groups*, Addison-Wesley, Cambridge, MA

Timothy Bickmore, Daniel Schulman and Langxuan Yin, Engagement vs. Deceit: Virtual Humans with Human Autobiographies, 2009, *IVA*, Amsterdam

Kohji Dohsaka, Ryoto Asai, Ryichiro Higashinaka, Yasuhiro Minami and Eisaku Maeda, Effects of Conversational Agents on Human Communication in Though Evoking Multi-Party dialogues, 2009, *10th Annual SigDial*, London, UK

Pawel Dybala, Michal Ptaszynski, Rafal Rzepka and Kenji Araki, Humoroids: Conversational Agents that induce positive emotions with humor, 2009, *AAMAS*, Budapest, Hungary

Arthur C. Graesser, Patrick Chipman, Brian C. Haynes, and Andrew Olney, 2005, AutoTutor: An Intelligent Tutoring System with Mixed-initiative Dialogue, *IEEE Transactions in Education*, 48, 612-618

Katherine Isbister and Patrick Doyle, Design and Evaluation of Embodied Conversational Agents: A Proposed Taxonomy, 2002, *AAMAS Workshop: Embodied Conversational Agents*, Bologna, Italy

Rohit Kumar, Carolyn Rosé, Mahesh Joshi, Yi-Chia Wang, Yue Cui and Allen Robinson, Tutorial Dialogue as Adaptive Collaborative Learning Support, *13th AIED 2007*, Los Angeles, California

Rohit Kumar, Carolyn Rosé, Building Conversational Agents with Basilica, 2009, *NAACL*, Boulder, CO