Modeling Efficient Grounding in Chat-based CSCL: an Approach for Adaptive Scripting?

Michael Oehl, Hans-Ruediger Pfister, Leuphana University of Lueneburg, Wilschenbrucher Weg 84a, 21335 Lueneburg, Germany,
Email: oehl@uni.leuphana.de, pfister@uni.leuphana.de

Abstract: Adaptive or intelligent scripting is a promising new approach to computer-supported collaborative learning (CSCL). Such collaboration scripts need mechanisms to adapt to the learners' progress and to assess the learners' evolving interactions. However, researchers are still facing many challenges how to model this process effectively. In our current study we applied a probabilistic approach to model efficient and inefficient communicative processes in chat-based CSCL.

Introduction

Adaptive or intelligent scripting is a promising new approach to computer-supported collaborative learning (CSCL). Such collaboration scripts need mechanisms to adapt to the learners' progress and to assess the learners' evolving interactions. However, researchers are still facing many challenges how to model this process effectively. In our current study we applied a probabilistic approach to model efficient and inefficient communicative processes in chat-based CSCL based on the grounding theory in linguistics according to Clark (1996). Especially chat-based CSCL often suffers from grounding deficiencies. Collaboration scripts in general vary according to the degree of flexibility of the implemented tools. With respect to the optimal degree of coercion the designer is confronted with a dilemma: a lack of coercion may affect the functionality of the script, whereas a surplus could lead to a phenomenon called over-scripting resulting in sterile, constrained interactions and an increased cognitive load for the contributors. An approach recently discussed is the development of highly flexible and *adaptive scripts*, providing support in terms of visualizations and feedback mechanisms exclusively in situations, when it is necessary. Prospectively it might be possible to implement a computer-supported detection of structural elements and patterns characteristic for deficient communication processes, resulting in an intelligent and adaptive script to foster grounding and thus learning processes. Such collaboration scripts need mechanisms to adapt to the learners' progress and to assess the learners' evolving interactions.

A promising approach to the automatic identification of suboptimal communication and grounding processes in CSCL are probabilistic models such as *Hidden Markov Models (HMM)*, used e.g. by Soller (2004) in a rather coercive script. For each sequence of tokens (e.g., a specific sequence of speech acts) a model can be computed consisting of a fixed number of states, initial probabilities, transition probabilities and output probabilities for the tokens (Rabiner, 1989). Additionally, the probability of a specific sequence of tokens being generated by this model can be calculated as well.

The long-term objective of this research is an adaptive CSCL script automatically promoting and facilitating grounding processes and thereby group cognition and learning in chat-based learning groups. For now in a first step the aim of the current study is to model efficient and inefficient chat discourses in a rather free chat environment.

Method

Students (N = 118; 70.10% female; age: M = 25.03 years, SD = 4.07) volunteered as participants and were randomly assigned to 33 learning groups, consisting of three or four members. Participants were instructed to learn about the mechanisms and causes of the Chernobyl nuclear power plant disaster using the ConcertChat interface (Mühlpfordt, 2006). We reported the learning results and preliminary grounding analyses in Pfister and Oehl (2009). For the current detailed process analyses of the *chat discourse structures* and the *grounding processes* with HMM, each contribution was rated according to a coding scheme of grounding activities based on Beers, Boshuizen, Kirschner and Gijselaers (2007). The original scheme by Beers et al. (2007) was specially developed to capture the negotiation of common ground in CSCL. However, to fit the characteristics of our study, some modifications had to be made resulting in a coding scheme of five categories with regard to grounding activities:

(1) Initiation: A new topic in form of a statement or a question is introduced.
(2) Verification: Information is requested about a previous contribution.
(3) Response: A content-related reaction to an initiation or verification.
(4) Positive feedback: A positive reaction (e.g., expression of intelligibility or agreement).
(5) Negative feedback: A negative reaction (e.g., expression of unintelligibility or disagreement).

It was evaluated and optimized in a validation process and achieved a good inter-coder reliability (Cohen's kappa) of r = .85, p < .001 between three coders.

The best and worst quartile (25%) of the groups with regard to the learning results (Pfister & Oehl, 2009), each eight groups, was used for further discourse analyses and to generate a HMM on the one hand for efficient learning discourses and on the other hand for inefficient chat discourses.

Results and Conclusion

Thread-based analyses showed that efficient groups' discourses contained significantly less different threads (*M* = 10.08, *SD* = 2.78) than discourses of inefficient groups (*M* = 12.93, *SD* = 4.13) $t_{(409)} = 8.22$, $p < .001$. But these discourse threads in the efficient groups were more in-depth (*M* = 2.75, *SD* = 0.72) than in the inefficient groups (*M* = 2.21, *SD* = 0.49) $t_{(409)} = -8.88$, $p < .001$. The occurrence of the five grounding activities in efficient versus inefficient discourses was significantly different $\chi^2_{(4)} = 16.41$, $p < .005$. Based on these differences of grounding activities the modeling of the HMM revealed some noticeable differences regarding the discourse structure (efficient vs. inefficient) as shown in figure 1.

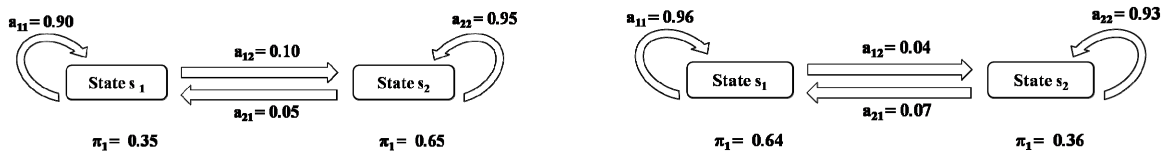

Figure 1. HMM for Efficient (Left) and Inefficient (Right) Grounding Discourses.

In order to specify the differences between the models, the output probabilities of grounding activities as displayed in table 1 have to be considered. The different initial probabilities (π_1) of the two HMM discourse models were salient, i.e., the initial probabilities in the model for the efficient discourses had a high tendency that the group starts its grounding discourse in state s_2 ($\pi_1 = .65$), remains in this state ($a_{22} = .95$) or gets in the discourse process from state s_1 into s_2 ($a_{12} = .10$). Especially the probabilities of grounding activities in state s_2 within the *efficient chat* discourses (table 1) might be interpreted as an almost prototypical grounding process, i.e., an initiation is followed by a response and again grounded by positive or negative feedback. In contrast to this, the more likely state s_1 within the *inefficient chat* discourses resembles rather incomplete grounding with lacking feedback in communication (table 1).

Table 1: HMM probabilities of grounding activities in efficient and inefficient group discourses.

	Efficient		Inefficient	
Code	State 1	State 2	State 1	State 2
Initiation	.25	.53	.40	.49
Verification	.14	.01	.05	.00
Response	.55	.21	.50	.37
Positive Feedback	.00	.17	.00	.10
Negative Feedback	.06	.76	.05	.03

Taking together, even though HMM are a promising approach with respect to assess the learners' evolving interactions for intelligent or adaptive scripting in chat-based CSCL, they raise methodical questions. We will discuss these issues of modeling grounding processes as well as we will outline implications for further research and application.

References

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. H. (2007). The analysis of negotiation of common ground in CSCL. *Learning and Instruction, 17*, 427-435.

Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.

Mühlpfordt, M. (2006). *ConcertChat* [Computer software]. Retrieved February 8th, 2011, from http://www.ipsi.fraunhofer.de/concert/index_en.shtml?projects/chat.

Pfister, H.-R., & Oehl, M. (2009). The impact of goal focus, task type, and group size on synchronous net-based collaborative learning discourses. *Journal of Computer Assisted Learning, 25*, 161-176.

Rabiner, L. R. (1989). A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE, 77*(2), 257-286.

Soller, A. (2004). Understanding knowledge-sharing breakdowns: A meeting of the quantitative and qualitative minds. *Journal of Computer Assisted Learning, 20*(3), 212-223.