A Semi-Automated Wizard of Oz Interface for Modeling Tutorial Strategies

Paola Rizzo[1], Hyokyeong Lee[2], Erin Shaw[2], W. Lewis Johnson[2], Ning Wang[2], and Richard E. Mayer[3]

[1] Dept. of Computer Science, University of Rome "La Sapienza"
Via Salaria 113, 00198 Rome, Italy
paola.rizzo@uniroma1.it
[2] Information Sciences Institute, University of Southern California
4676 Admiralty Way, Marina del Rey, CA 90292 USA
{hlee, shaw, johnson, ning}@isi.edu
[3] Dept. of Psychology, University of California, Santa Barbara
Santa Barbara, CA, 93106-9660 USA
mayer@psych.ucsb.edu

Abstract Human teaching strategies are usually inferred from transcripts of face-to-face conversations or computer-mediated dialogs between learner and tutor. However, during natural interactions there are no constraints on the human tutor's behavior and thus tutorial strategies are difficult to analyze and reproduce in a computational model. To overcome this problem, we have realized a Wizard of Oz interface, which by constraining the tutor's interaction makes explicit his decisions about why, how, and when to assist the student in a computer-based learning environment. These decisions automatically generate natural language utterances of different types according to two "politeness" strategies. We have successfully used the interface to model tutorial strategies.

1 Introduction

The typical approach to modeling human teaching strategies for realizing Intelligent Tutoring Systems (ITSs) and Interactive Learning Environments (ILEs) is twofold. First, one records the interactions taking place between the tutor and student in a natural setting or computer-mediated interface; then, the transcripts are analyzed to find effective teaching patterns, which are reproduced in a computational model that constitutes the basis of an artificial tutor. This approach has some shortcomings: analyzing videotaped interactions is difficult and time consuming, and the results depend on the reliability of the raters. Furthermore, the only perceivable outputs from the tutors are their utterances and non-verbal behaviors: one cannot access their teaching strategies other than by interviewing, a method that shares the typical shortcomings of indirect analyses of cognitive processes. This makes it hard to reproduce tutoring strategies within a computational model.

In order to efficiently model human tutoring strategies, we propose that the tutors use a semi-automated Wizard of Oz interface which forces them to take explicit and visible pedagogical decisions, and that automatically outputs those decisions as natural language utterances. This approach has three advantages: (a) the tutor's behavior can be analyzed much faster than transcripts and videotapes; (b) the pedagogical decisions can be easily correlated with student performance; and (c) the natural language generator can be tested for coverage and robustness.

We have realized a Wizard of Oz interface (WozUI, Fig. 1) that enables a human tutor to communicate with a student by selecting a pedagogical goal, choosing an object of discourse, and applying a communicative act to it. As a learning environment we use the Virtual Factory Teaching System (VFTS) [5], a web-based factory modeling and simulation system. The tutor can view the learner's screen activities in the VFTS, using MS NetMeeting, and exchange messages with the learner. An animated puppet speaks the comments sent by the experimenter, using text-to-speech software.

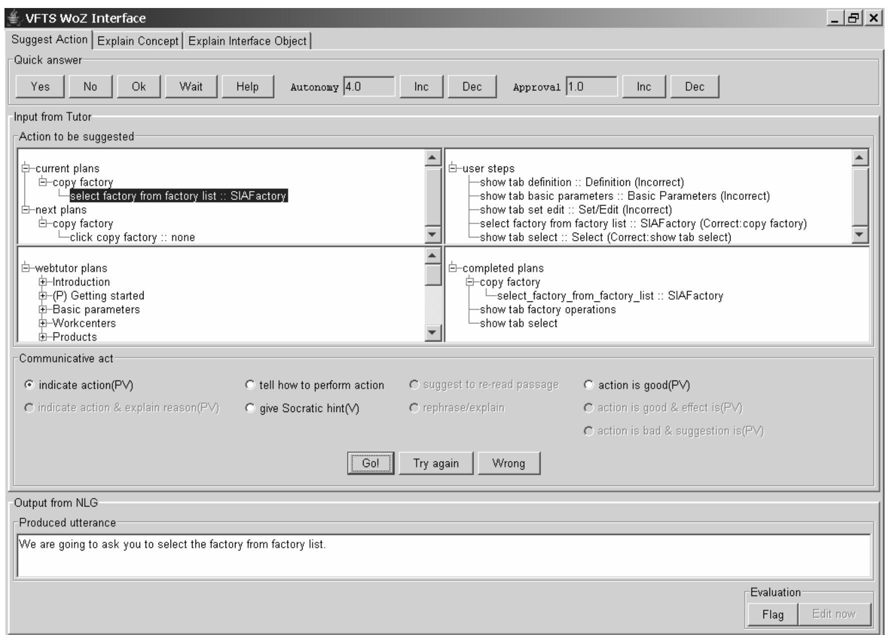

Fig 1. Wizard of Oz interface

In order to track tutor decisions about why, what, and when to communicate with the student in the VFTS, the WozUI: (a) models the student, thanks to a plan recognizer that tracks the student's actions, and a "Web tutor" that traces the tutorial paragraphs visible to the student at any time [12], as shown in the four upper windows; and (b) embodies a model of the strategies that the tutor performs. In fact, the lower part of the WozUI enables the tutor to select a communicative act categorized according to one of the following pedagogical goals, taken from the taxonomy of pedagogical objectives proposed by Bloom and coworkers [3, 1]. *Application*: indicate an action to perform, indicate an action to perform and explain its reason, tell how to perform an action, give a Socratic hint (i.e. a cue about an action to perform). *Knowledge*: suggest to re-read a passage of the tutorial, explain a paragraph of the tutorial, provide an example for a paragraph of the tutorial. *Motivation*: tell the student that the action he has performed is good, tell the student that the action he has performed is good and describe its effect.

Together with other authors [2, 10], we are particularly interested in modeling human tutors' ability to take not only cognitive, but also affective student goals into account, according to the Politeness Theory proposed by Brown & Levinson [4]. Therefore we have realized a natural language generator (NLG), coupled with a Politeness Module, that transforms tutor communicative acts into utterances with several levels of politeness [8]. For example, a suggestion to perform an action, such as saving the current factory description, can be stated either directly (e.g., "Save the factory now"), or as a hint, ("Do you want to save the factory now?"), as a suggestion of a joint action ("Why don't we save our factory now?"), etc.

An example of how the WozUI works is shown in Figure 1: the tutor is pursuing the goal of suggesting an action; he chooses "select factory from factory list" as argument and "indicate action" as communicative act, and sends these decisions to the system by pressing the "Go!" button. The resulting utterance (after being processed by the Politeness Module), shown at the bottom, is: "We are going to ask you to select the factory from factory list".

2 Using the Wizard of Oz Interface: An Experiment

The experiment compared two tutoring modes: "direct", in which no politeness strategy was used; and "polite", where the NLG automatically applied the politeness strategies described above. The experiment involved 10 graduate and undergraduate students in technical and engineering domains: 5 of them were assigned to the "direct" condition, while the other 5 were assigned to the "polite" condition. After the experiment, the log files produced by the WozUI were analyzed for understanding which pedagogical criteria were used by the tutor.

2.1 Experimental Results: Timing of Tutorial Interventions

First of all, we analyzed the average timing of the tutorial interventions across subjects, by observing whether they were more proactive (spontaneous) or reactive (in response to student questions), both in the direct vs polite conditions and in general ("Direct + Polite"). There were slightly more proactive than reactive interventions, but the difference is not statistically significant ($t(8) = 0.58$, $p = 0.57$).

2.2 Experimental Results: Types of Tutorial Interventions

Secondly, we analyzed the average number of tutor interventions across subjects, according to the pedagogical strategies described above. The results showed that the tutor using the WozUI was pursuing the same pedagogical goals pursued by the other human tutor in a natural setting as described in [7]: in fact, the Wizard tutor provided suggestions for actions more often than providing positive feedback, and more often than explaining concepts. The difference between the number of "Application" interventions and the number of "Knowledge" interventions is statistically significant ($t(8) = 3.59$, $p = 0.007$).

2.3 Experimental Results: Types of Communicative Acts

We also analyzed the communicative acts selected within the most frequently chosen pedagogical goal (Application), in both the polite and direct mode. The two most used communicative acts were "Indicate action" and "Tell how to perform action", which are more short and explicit in telling the student what to do with respect to the "Socratic hint" and the "Indicate action & explain reason" acts. The difference between interventions of type "Indicate action" and "Tell how to perform action" is statistically significant ($t(8) = 5.33$, $p = 0.0007$). A possible reason for the larger use of concise communicative acts might be that the learning task is not very difficult, and therefore the student just needs some simple help when he seems unsure about how to operate the VFTS interface. However, another hypothesis might explain this result simply as a preference of the tutor towards given types of communicative act.

3 Related Work

While most WozUIs require the Wizard to type free text in computer-based dialogs with learners, a few examples exist where the experimenter has to produce utterances by using a limited set of options that produce canned text [6, 9]. However, a number of differences between their and our approach can be found: (a) in the other cases a domain-dependent script predefines the possible interactions with the student, while in our case the interactions between tutor and student are much less constrained; (b) the goals of the interactions, when modeled, are usually domain dependent, rather than related to a more general pedagogical taxonomy like Bloom's one; and (c) the other WozUIs do not allow to parameterize the tutor's utterances according to features like politeness.

Regarding affect, motivation and politeness in learning environments and in human-computer interaction, a number of other works have been referenced and compared with our approach in [8, 12].

4 Conclusion and Future Work

The semi-automated Wizard of Oz technique made it possible to analyze the experimental results very rapidly and systematically. If we had not employed this approach, it would have been necessary to annotate the dialogs manually. In comparison, we have found in the past that post-hoc dialog markup is time consuming, and raises the possibility of inter-rater reliability problems. We are currently analyzing the data coming from another experiment with more subjects, so as to see observe the robustness and generalizability of our current findings.

Regarding the usability of the WozUI, apparently the tutor did not find the interface difficult to use, nor overly constraining: before running the experiment, a pilot study showed that the human tutor was able to monitor the student's activities and quickly react to them. The experiment results also seem to show that working through the interface did not bias the tutor's behavior: in fact the frequencies of tutor interventions were analogous to those observed in the natural setting. Anyway, we also plan to carry out a more formal usability analysis of the WozUI.

References

1. Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C.: A taxonomy of learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman, New York (2001)
2. André, E. Rehm, M., Minker, W. & Bühner, D.: Endowing spoken language dialogue systems with emotional intelligence. In E André et al. (Eds.): Affective Dialogue Systems: Tutorial and Research Workshop, ADS 2004. Springer, Berlin (2004) 178-187
3. Bloom, B.S. (ed.): Taxonomy of educational objectives: The classification of educational goals. Handbook I: cognitive domain. Longman, London (1956)
4. Brown, P. & Levinson, S.C.: Politeness: Some universals in language use. Cambridge University Press, New York (1987)
5. Dessouky, M.M., Verma, S., Bailey, D., & Rickel, J.: A methodology for developing a Web-based factory simulator for manufacturing education. IEE Trans. 33 (2001) 167-180
6. Fiedler A.., Gabsdil M., Horacek H.: A tool for supporting progressive refinement of Wizard of Oz experiments in natural language. In: Lester, J.C., Vicari, R. M., Paraguacu, F. (eds.): Intelligent Tutoring Systems: 7th International Conference, ITS 2004. Springer, Berlin (2004) 325-335
7. Johnson, W.L., Rizzo, P.: Politeness in tutoring dialogs: 'Run the factory, that's what I'd do'. In: Lester, J.C., Vicari, R. M., Paraguacu, F. (eds.): Intelligent Tutoring Systems: 7th International Conference, ITS2004. Springer (Lecture Notes in Computer Science, Vol. 3220), Berlin (2004) 67-76
8. Johnson, W.L., Rizzo, P., Bosma W., Kole S., Ghijsen M., Welbergen H.: Generating Socially Appropriate Tutorial Dialog. In In E André et al. (Eds.): Affective Dialogue Systems: Tutorial and Research Workshop, ADS 2004. Springer, Berlin (2004) 254-264
9. Kim J. K. & Glass M.: Evaluating dialogue schemata with the Wizard of Oz computer-assisted algebra tutor. In: Lester, J.C., Vicari, R. M., Paraguacu, F. (eds.): Intelligent Tutoring Systems: 7th International Conference, ITS2004. Springer (Lecture Notes in Computer Science, Vol. 3220), Berlin (2004) 358-367
10. Porayska-Pomsta, K.: Influence of Situational Context on Language Production. Ph.D. thesis, University of Edinburgh (2004)
11. Qu, L., Wang, N., & Johnson, W.L.: Pedagogical Agents that Interact with Learners. In: Proceedings of the Workshop on Embodied Conversational Agents: Balanced Perception and Action, Third International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2004). New York (2004) 42-50
12. Wang N., Johnson L.W., Rizzo P., Shaw E., Mayer R.E.: Experimental Evaluation of Polite Interaction Tactics for Pedagogical Agents. In: International Conference on Intelligent User Interfaces. ACM Press, New York (2005) 12-19