The Potential for Chatbots in Negotiated Learner Modelling: A Wizard-of-Oz Study

Alice Kerly and Susan Bull

Electronic, Electrical and Computer Engineering, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK
{alk584, s.bull}@bham.ac.uk

Abstract. This paper explores the feasibility of using conversational agents, or chatbots, in negotiated learner modelling. This approach aims to combine the motivational, intuitive and domain-independent benefits of natural language dialogue using a chatbot, with the opportunities for learner reflection and increased model accuracy that can be achieved through negotiation of the learner model contents. A Wizard-of-Oz paradigm allowed investigation into the interactions between learners and their learner model in order to highlight key issues for the design of a chatbot for this purpose. Users appreciated interacting with a chatbot, and found it useable and an aid to negotiation. The study suggested many avenues for future investigation of the role of conversational agents in facilitating user-system dialogue about learner understanding.

1 Introduction

It has been argued that making visible ("opening") the learner model of an intelligent tutoring system (ITS) to its users can provide opportunities for learner reflection that strengthen and enhance the learning experience (e.g. [1], [2], [3], [4]). Mr. Collins [5] and STyLE-OLM [2] proposed and implemented the learner's participation in the construction and maintenance of the learner model through a process of negotiation of the model contents. They allowed learners to discuss their beliefs with the system, and to argue against the system's assessment of their knowledge if they disagreed with it, with the aim of enhancing learner reflection and improving learner model accuracy.

Mr. Collins was developed with the open learner model (OLM) central to the system. In order to discuss beliefs, the system must have knowledge of the student's real beliefs (in case these differed from the system inferences), as well as its own beliefs about the student's knowledge. Thus, two belief measures were maintained, each of which the system used to determine appropriate adaptive interactions [5]. This allowed both student and system to retain autonomy over their individual beliefs. With this ethos, the student is able to alter their self-assessment as they wish (though the system will challenge them if it disagrees), and must engage in negotiation if they wish to alter the system's belief. Baker's [6] notion of interaction symmetry was applied by making all negotiation moves available to both student and system. The Mr. Collins system utilised menu-selection for negotiation. This provided an efficient and usable interaction (as demonstrated in a laboratory study [5]), but may have restricted some users by being an unnatural method for communicating beliefs.

M. Ikeda, K. Ashley, and T.-W. Chan (Eds.): ITS 2006, LNCS 4053, pp. 443 – 452, 2006.

STyLE-OLM [2] also provided negotiation of the learner model, through graphical representations of conceptual graphs. Both learner and system formed propositions about learner understanding by constructing conceptual graphs and selecting an appropriate dialogue statement. In a laboratory trial STyLE-OLM was successful in engaging learner reflection and in improving the resultant learner model.
However, the use of a graphical approach may be difficult for some learners, and may sometimes detract from learning of the domain in question.

Given the positive reactions to negotiated learner modelling in Mr. Collins and STyLE-OLM, it would be useful to consider allowing negotiation in a more flexible, intuitive and naturalistic way than previously implemented. A method that does not require learning additional communication skills, and that does not detract from the target domain, is required. Natural language negotiation may be a way to enable this.

Conversational agents, or chatbots, provide a natural language interface to their users. Their design has become increasingly complex, and modern commercial chatbots, such as Lingubot technology [7] offer sophisticated development environments, allowing the building of intelligent conversational agents with complex, goal driven behaviour. In 'Lingubots' both the words and grammatical structure of user input are analysed. This allows the development of a user model, which is used in conjunction with conversational context and dialogue content to determine the chatbot's response. This seems particularly appropriate for student-system collaborative construction of the learner model.

2 A Wizard-of-Oz Study

This paper describes a Wizard-of-Oz experiment to explore the feasibility of using a chat-based interface to an OLM system. Participants interacted with the system to view their learner model and discuss or negotiate over the contents with the system 'chatbot'. The role of the 'chatbot' was taken by a human 'Wizard', a fact not revealed to the participants. All interactions were logged. Participants completed questionnaires indicating their opinions on the interaction with the 'chatbot'.

2.1 The Wizard-of-Oz Paradigm

The Wizard-of-Oz method allowed the collection of data on the human-computer interactions that arose in negotiating the learner model. Computer mediated human-human interaction data can be an unreliable source of information for some important aspects of dialogue design as humans and computers can be expected to perform differently in conversation [8]. The dialogues required to negotiate over the learner model are expected to be complex and varied in terms of language, tasks, and domain content. Wizard-of-Oz studies have been shown to be an appropriate approach in collecting data about dialogues in complex domains such as this [9].

In this study, the 'Wizard' (experimenter taking the role of the chatbot) followed a protocol designed to ensure that responses to students remained consistent between users, and that the 'chatbot' was believable to users. The Wizard-of-Oz paradigm necessarily involves deception of the participants about their true activities in the experiment. Work was conducted within British Psychological Society Ethical Principles for Research with Human Participants [10].

2.2 The Learner Modelling System

This study employed a modified version of Flexi-OLM [11], which operates in the domain of C programming. The system's beliefs are inferred from students' answers to multiple-choice and short-answer questions. For this study Flexi-OLM was extended to also allow learners to record their own beliefs about their knowledge. These two belief sets make up the learner model contents, viewable in seven formats.

2.3 Participants, Materials and Methods

The participants were 11 final year undergraduates from the University of Birmingham Electronic, Electrical and Computer Engineering Department. All had taken courses in educational technology and C programming. All were competent English language speakers, though in some cases English was not their first language.

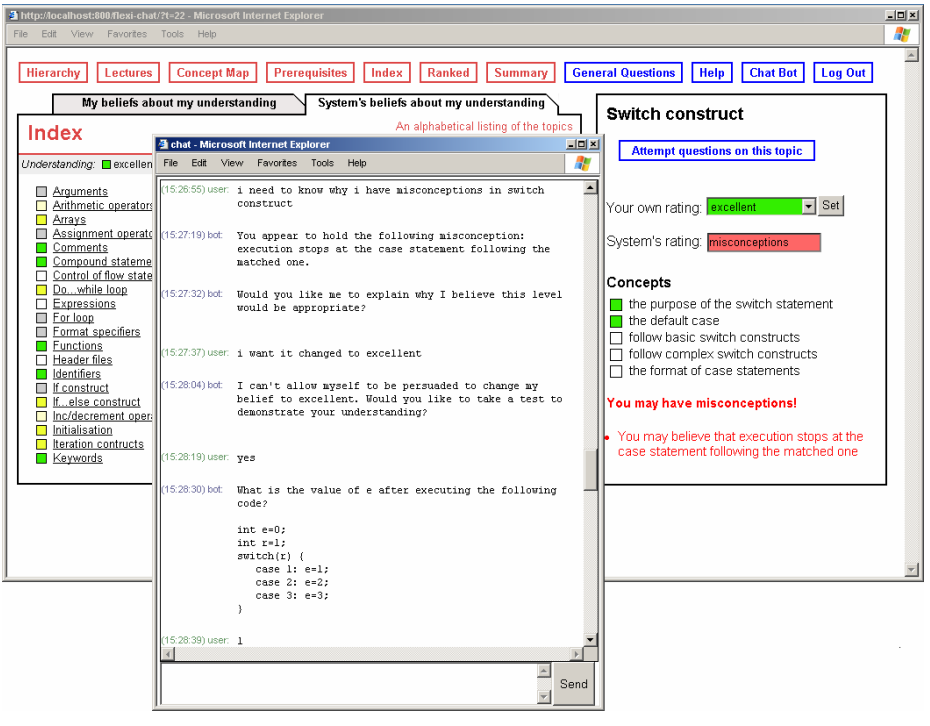

Fig. 1. User's screen in Index view of System beliefs on student knowledge & chat interface

Figure 1 shows a screenshot of user activity in Flexi-OLM. Buttons at the top of the screen allow the user to view their learner model contents as either: lecture structure, topic hierarchy, concept map, pre-requisites, index, ranked or textual summary views. Although these views differ in their presentation, the same model information is available in each. Below these buttons are tabs allowing the user to switch between their own or the system's beliefs about their knowledge.

Topics group together several smaller concepts. For each topic or concept a learner's knowledge is represented at one of four levels by a node, coloured between white (very limited), through yellow and green to bright green (excellent). If the learner holds a misconception about a concept, the parent topic is changed to red.

The current topic (here, "Switch construct") is shown on the right. A button allows answering questions on this topic, and the user's and system's beliefs are shown for ease of comparison. To allow investigation of negotiation through a chatbot interface, a chat window was provided. Students viewed the 'chatbot's' conversation in the top part of the window, and typed their responses to the 'chatbot' in the bottom section.

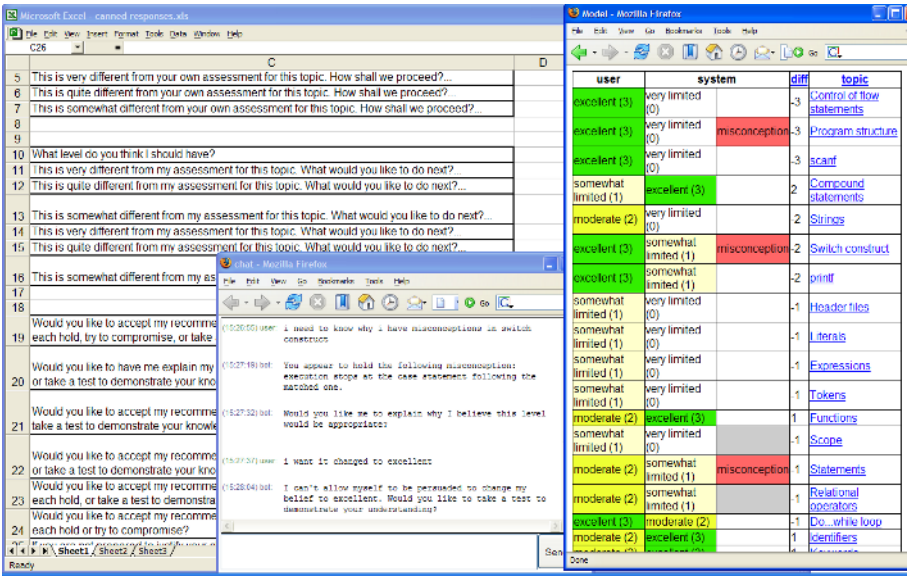

Fig. 2. Wizard's view showing (left) part of the canned responses; (centre) Chat interface - excerpt reproduced below; (right) Learner Model (part)

- Chatbot: Would you like me to explain why I believe this level would be appropriate?
- User: I want it changed to excellent.
- Chatbot: I can't allow myself to be persuaded to change my belief to excellent. Would you like to take a test to demonstrate your understanding?

Mr. Collins [1] allowed negotiation by a range of strategies: ask user if they wish to accept system view; offer compromise; ask user to justify their belief (by taking a test to demonstrate knowledge); system justify its belief; or offer student opportunity to view learner model. These strategies formed the conversational basis of the 'chatbot'.

Table 1 shows the key structure of the decision tree constructed to assist the Wizard in providing appropriate responses. To enact the protocol, the Wizard was provided with some 350 pre-authored 'chatbot' negotiation initiations and responses to user inputs. These allowed the Wizard to respond to user inputs quickly and consistently, and to avoid betraying the fact that the 'chatbot' was controlled by a

human. These can be seen in the left section of Figure 2. The learner model (right of Figure 2) was presented to enable the Wizard to compare user and system beliefs for each topic at a glance. User and system beliefs were colour coded using the previously described scheme. Unanswered questions were available to allow the student to demonstrate their knowledge in a further test. The window provided for 'chatbot' responses was identical to that of the users.

Table 1. Key actions from the Wizard's decision protocol

If:	User has challenged system belief	
Then:	Respond with an appropriate strategy (view model, compromise, user accept, system justify, user justify)	
	System can be persuaded without evidence to move one level up or down to match user belief, except if user wishes to change system belief to excellent, insist on test to demonstrate knowledge	
If:	Discrepancies between system and user beliefs AND user has not already initiated negotiation AND user is not currently answering questions OR user has changed own belief, causing discrepancy	
Then:	Initiate interaction on topics with largest discrepancy/misconception first	
	Offer the five strategies to proceed for user to choose between:	
	User accepts system view	Update user belief
	System justify itself	Show user's previous answers
	User justify self	Give test questions
	Compromise	Set user's and system's beliefs to mid-point
	View learner model	Remind user how to view
If:	All options exhausted for topic, but no resolution to negotiation	
Then:	Agree to differ	

2.3.1 Experimental Setup

Participants were instructed that they would be able to use a chatbot feature of Flexi-OLM to discuss their learning of C programming and which could lead to modification of the learner model. They were told that they could answer questions; view the system's model of their knowledge, or their own belief model in seven formats; edit their own beliefs for a topic (though the system may challenge this); and negotiate over the model by challenging the system's beliefs, seeking justification for the system's beliefs, offering their own justifications of their knowledge, taking additional tests on a topic, attempting to compromise with the system, or by discussing the differences between the belief models. Participants were told that they could initiate discussion at any stage, or that the chatbot may do so. Participants were requested to be natural in their language, but precise and clear.

Each participant was shown how to use the chatbot and asked to interact with it for at least 20 minutes. Prior to using the system, each participant completed a self-assessment, rating their knowledge of each concept, providing the initial content of the user's beliefs in the learner model. Nine users had previously used Flexi-OLM and data from this usage allowed construction of the system's beliefs in the learner model. Two users interacted with the system before beginning the experiment to enable the system to model their understanding.

A post-experiment questionnaire comprised statements requiring participants to indicate their level of agreement (strongly agree, agree, neutral, disagree, strongly disagree) and open-ended questions to elicit further comments or suggestions.

2.4 Results

This section presents results relating to negotiation and interaction with the 'chatbot'.

2.4.1 Negotiation Issues Findings

The logs of interactions show there were 20 occasions when the system initiated discussion of the learner model, a mean of 1.8 per user session. The user initiated negotiation on 29 occasions, a mean of 2.6 per user session. The greatest number of initiations by a single user was eight. Two users did not attempt to initiate negotiation.

Table 2. User opinions of the negotiation facilities

		<strongly agree....strongly disagree>					Mean score
		(1)	(2)	(3)	(4)	(5)	
I was frequently convinced by the Chatbot's arguments		1	6	3	1	0	2.4
The negotiation changed my view of my understanding		2	5	4	0	0	2.2
In negotiation, I wanted to discuss	my knowledge	7	3	1	0	0	1.5
	my knowledge level	7	3	1	0	0	1.5
	my misconceptions	7	3	0	1	0	1.5
	the gaps in my knowledge	4	5	0	2	0	2.0
I wanted to be able to negotiate at a topic overview level		3	6	0	2	0	2.1
I wanted to be able to negotiate at a more detailed (concept) level		3	3	1	4	0	2.5
I used the Chatbot to help me understand my learner model		2	4	3	1	1	2.5
I always challenged the Chatbot when I disagreed with it (or I would)		7	3	0	1	0	1.5

As shown in Table 2, in the questionnaire seven users agreed or strongly agreed that they were frequently convinced by the chatbot's arguments; one user disagreed. User comments supported this, including indicating that the chatbot was "right most of the time, and never unreasonable". Another user said, "I could see *why* it was disagreeing".

Users were also asked to comment on what they thought of the chatbot when it agreed or disagreed with them. When there was disagreement, users made comments such as "I saw it as an opportunity to put it right if it was wrong and correct myself if it proved I was wrong". Users also wanted to understand the reasons for any disagreement, variously stating, "I could see why it was disagreeing so it wasn't too bad" and "when it disagreed it was justified".

When users found the chatbot in agreement with their views, opinions were mainly positive. Users' comments included "it only agreed when the difference was marginal so it was good" and "I liked that it wasn't just agreeing with me all the time". As with situations when the chatbot disagreed with the user, users wanted to understand the reasons for the chatbot's belief, for example, "[I] could see why it was agreeing, so I was pleased", and requested "feedback of how many answers got right or wrong".

Regarding when and what the users wanted to discuss, most users indicated that they would always challenge the chatbot when its belief differed from their own, that they wanted to discuss each of their knowledge, their knowledge level or their misconceptions, and that they wanted to discuss the gaps in their knowledge. Nine students wanted to be able to negotiate the learner model at a topic overview level, with six wanting to negotiate at a more detailed (concept) level.

Users suggested that the chatbot was "an interesting approach to learning – more enjoyable than just using the original Flexi-OLM" and that it was a "good way of negotiation and justification".

One of the principal aims of opening the learner model is to promote user reflection. Seven users stated that the negotiation had changed their view of their understanding while four remained neutral on this. Six users indicated that they used the negotiation to understand their learner model, while three remained neutral, and two disagreed.

2.4.1 Chatbot Interaction Findings

As Table 3 shows, all users stated they liked the chatbot. Most indicated that they enjoyed interacting with it.

More users liked the chatbot when it agreed with them than when it disagreed. Only three users stated they liked the chatbot when it disagreed with them, while seven liked the chatbot when it was in agreement with their beliefs.

Table 3. User opinions of the chatbot

	<strongly agree...strongly disagree>					Mean score
	(1)	(2)	(3)	(4)	(5)	
I liked the Chatbot	2	9	0	0	0	1.8
I liked the Chatbot when it disagreed with me	1	2	6	1	1	2.9
I liked the Chatbot when it agreed with me	2	5	4	0	0	2.2
I enjoyed interacting with the Chatbot	5	5	1	0	0	1.6
I was able to achieve the negotiation tasks I wished with the Chatbot	4	5	2	0	0	1.8
The Chatbot made negotiation easy	4	6	0	1	0	1.8
The Chatbot was a convenient way to provide my opinions to the system	5	4	2	0	0	1.7

An indicator of user enjoyment of the chatbot is the users' language that shows their level of relaxation in the interaction. Examples of inputs include "lol" (meaning laugh out loud, indicating amusement), and "lol, never mind then :-)". Users also commented that the interaction "felt more relaxed" and even that they would like to

be able to "discuss other stuff, e.g. weather, just for a little break from working!". However, a few students reported that the interaction seemed rigid.

A number of users felt that they were not alone in the use of the system, and it appears that the presence of the chatbot was felt to be companionable. For example, users commented, "I liked the fact that there was an interaction going on. I wasn't in the learning area by myself", "it makes learning system active, such as chatting with real people", and "it was very good to use chatbot because you can interact with [it]".

Given the intention of the chatbot is to facilitate communication, it was important to discover whether this was the user experience. Nine users said that the chatbot was a convenient way to provide their opinions to the system. Their comments included, "The chatbot is a fast and effective yet enjoyable way of creating an accurate learner model", and "A more effective way of giving information into the learner model".

Most users agreed that the chatbot made negotiation easy and that they were able to satisfactorily complete their desired tasks.

3 Discussion

The central concern of this study was whether a chatbot might effectively provide or enhance negotiation in an open learner model, to explore user reactions to the 'chatbot', and to investigate the technical challenges for implementation.

3.1 Issues Raised by Study

The results of this study suggest that users found the negotiation easy and useful, and appreciated the facilitation of the chatbot. One user commented that she saw disagreements with the system as "an opportunity to put it right if it was wrong and correct myself if it proved I was wrong". This, and the fact that users were often convinced by the chatbot's arguments, demonstrates a willingness to revise personal beliefs that is essential for the success of negotiated learner modelling.

When the system and user disagreed, users were clear in their wish to see why this was. This indicates users' need to understand the reasoning behind system beliefs. A capacity for explanation could prove a valuable addition to OLM systems.

The finding that most users would always challenge the system if they disagreed with it is an important indicator that users were sufficiently concerned about the chatbot's opinion to make the effort to challenge it. This accords with findings for the menu-based negotiation of Mr. Collins [5].

3.2 Implications for the Development of Chatbots in Negotiated Learner Modelling

The user comments indicate support for the inclusion of a chatbot in a negotiated OLM system. However, several points indicated areas for particular consideration for an implementation to be successful in user acceptance and fitness for purpose.

A number of users commented that the chatbot sometimes appeared rigid or formal. This is not surprising given the limited scope of the Wizard's responses, but should be rectified in future work. Smalltalk (conversation in areas unrelated to the

subject in hand) may be useful in helping students build a rapport with the chatbot. When combined with effective methods to draw the user back on topic, incorporation of smalltalk gives a far more rounded and naturalistic conversation, more akin to a discussion and less of a clever question answering system [12].

The user interactions in this study were relatively short, and were single sessions. Although the limited canned responses were necessary for the structure of this experiment, if users are to remain engaged with the chatbot, and to find its comments both useful and interesting, a richer range of chatbot responses should be provided.

It is also worth noting that many types of interaction are potentially valuable to the learning experience, and it is therefore reasonable to expect the chatbot to be able to deal with these. Although negotiation of the learner model is the primary task, it would be reasonable to expect the chatbot to cope with smalltalk, information requests, being told when its response is wrong, and clarification when things are not understood. Future development of the chatbot may enable it to extract and provide teaching material or point users to it, perhaps by opening relevant web pages.

4 Summary

This paper has described a Wizard-of-Oz experiment exploring the potential of using natural language conversation to negotiate an OLM system. Eleven participants used a simple 'chatbot' interface to engage with an existing OLM system.

Users were happy to engage with the system, found the 'chatbot' easy to interact with, and saw it as an effective way of providing information to the system. Most found the 'chatbot' friendly and helpful, and were happy to accept the 'chatbot's' arguments, revising their beliefs where necessary. Users' comments suggested a strong need to understand the reasoning behind system beliefs. A chatbot provides excellent facility for this type of explanation.

The study demonstrated the potential for using a 'chatbot' to take a central role in negotiating the learner model. Users indicated building rapport with the chatbot, which could be improved through adding smalltalk capabilities to core negotiation.

This work has provided support for implementation of a real chatbot, although limitations of the participant sample size are acknowledged. Future work will address shortcomings identified by users, and consider how negotiation can be enhanced through the inclusion of a wider range of natural language dialogue strategies. The results showed the potential extension a natural language interface could provide an OLM system, a useful first step in the investigation of the potential of chatbots in negotiated learner modelling. Work has commenced to consolidate the findings and, in collaboration with Elzware, a leading UK developer of Lingubot chatbots, development of a prototype OLM system with embedded chatbot negotiation.

Acknowledgement

We thank Andrew Mabbott for extending Flexi-OLM [11] for this study. The first author is funded by the UK Engineering and Physical Sciences Research Council.

References

1. Bull, S., Pain, H., & Brna, P.: Mr. Collins: A collaboratively constructed, inspectable student model for intelligent computer assisted language learning. Instructional Science 23. Kluwer Academic Publishers. Netherlands (1995) 65-87
2. Dimitrova, V.: STyLE-OLM: Interactive Open Learner Modelling. International Journal of Artificial Intelligence in Education, Vol 13 (2003) 35-78
3. Morales, R.: Exploring participative learner modelling and its effects on learner behaviour, unpublished PhD Thesis, University of Edinburgh (2000)
4. Zapata-Rivera, J-D. & Greer, J.: Externalising Learner Modelling Representations, Proc. of Workshop on External Representations in AIED: Multiple Forms and Multiple Roles. Intl Conference on Artificial Intelligence in Education (2001) 71-76
5. Bull, S. & Pain, H.: 'Did I say what I think I said, and do you agree with me?': Inspecting and Questioning the Student Model. Proceedings of World Conference on Artificial Intelligence in Education, Charlottesville, VA (1995) 501-508
6. Baker, M. J.: Negotiated Tutoring, and Approach to Interaction in Intelligent Tutoring Systems, unpublished PhD thesis, Open University, Milton Keynes, UK (1990)
7. Lingubot technologies UK distributor. http://www.creativevirtual.com/ accessed 14/12/05
8. Dahlbäck, N., Jönsson, A. & Ahrenberg, L.: Wizard of Oz Studies – Why and How. Proceedings of Intelligent User Interfaces (1993) 193-200
9. Bernsen, N. O., Dybkjaer, H., & Dybkjaer, L. Designing Interactive Speech Systems – From First Ideas to User Testing. Springer. (1998), cited in: Fiedler, A., & Gabsdil, M.: Supporting Progressive Refinement of Wizard-of-Oz Experiments. Proceedings of the ITS 2002 - Workshop on Empirical Methods for Tutorial Dialogue Systems, San Sebastian, Spain (2002) 62-69
10. British Psychological Society Ethical Principles for Conducting Research with Human Participants. Available: http://www.bps.org.uk/the-society/ethics-rules-charter-code-of-conduct/code-of-conduct/ethical-principles-for-conducting-research-with-human-participants.cfm accessed 18/10/05
11. Mabbott, A & Bull, S.: Student Preferences for Editing, Persuading, and Negotiating the Open Learner Model. 8th International Conference of Intelligent Tutoring Systems, Springer-Verlag, Berlin Heidelberg (in press)
12. Hall, P.: Director, Elzware (http://www.elzware.com) Personal communication 20/12/05