

## Generation of Dialogs Adapted to the Student Model for Pedagogic Conversational Agents

Cristóbal Hermida-Portales<sup>1</sup>, Diana Pérez-Marín<sup>1</sup>, Ismael Pascual-Nieto<sup>2</sup>

<sup>1</sup> Computing Languages and Systems I Department, Universidad Rey Juan Carlos

<sup>2</sup> Computer Science Department, Universidad Autónoma de Madrid, Spain  
[c.hermida@alumnos.urjc.es](mailto:c.hermida@alumnos.urjc.es), [diana.perez@urjc.es](mailto:diana.perez@urjc.es), [ismael.pascual@uam.es](mailto:ismael.pascual@uam.es)

**Abstract.** Conversational agents can be defined as computer programs which can have an animated face and/or body, understand natural language and respond in natural language to a user request. Pedagogic Conversational Agents are a subset specifically designed for educational purposes. Some of the existing conversational agents provide predefined answers irrespectively of the student model. In this paper, our hypothesis is that the use of Natural Language Processing techniques would allow the generation of dialog templates adapted to each student model and that it would improve the perceived quality and trustworthiness of the Pedagogic Conversational Agent. In order to support that hypothesis, an initial module of a more complete system to generate dialog templates has been implemented and it is described with some initial results.

**Keywords:** generation of dialog templates, adaptation to the student model.

### 1 Motivation

Conversational agents can be defined as computer programs which can have an animated face and/or body (embodied conversational agent), understand natural language and respond in natural language to a user request [1].

ELIZA was the first conversational agent, based on a simple pattern matching technique [2]. Since then, more and more conversational agents have appeared based on different techniques and implementations [3].

Regarding the interface, the simplest form of conversational agent could be a text area in which users type the sentences and read the written answer of the agent. While the most sophisticated form could be an agent with an animated face and body able to make gestures according to the dialogue in natural language (e.g. a smile on the face when the mood of the conversation is happy).

Conversational agents can be applied to multiple domains such as e-commerce, web assistance to help users navigating among the pages or retrieving information, training and education.

However, currently many of the existing conversational agents provide predefined sentences. Therefore, the generated dialogue is quite similar for different users, irrespectively of their preferences and needs. For instance, conversational agents that

answer just the same when asked about where some information in a web page is or to book a travel.

The reason for not generating a more adapted dialogue in natural language could be found in a perceived difficulty of using Natural Language Processing (NLP) tools, or the fear that NLP tools still make many mistakes that could lead to a more confusing dialogue instead of a more sophisticated interaction.

However, we believe that not generating and using student models for conversational agents supposes a main drawback since it means that they loss believability and trustworthiness. Both of them important factors to create realistic dialogues.

This is particularly relevant for the training and education domain providing benefits both for teachers and students. For teachers, the task of creating a course would not be so complex, and for students, the questions would not only be presented in different order but with different sentences adapted to them.

For instance, if instead of asking teachers to type the questions for a course, teachers would only need to type information about the course and the questions were generated from that information, it would be possible that the agent could generate questions about different concepts and their relationships depending on the course information and adapted to each student.

That way, if the agent asks the student about a certain concept and the student gives a correct answer, the agent could use that concept to create more complex questions taking into account that the previous concept has been assimilated according to the Meaningful Learning Theory of Ausubel [4].

It is therefore expected that the interaction would not only be more natural, but also more efficient, based on a dialogue more focused on the specific conceptual difficulties of each student.

In this paper, we propose an ongoing work on the creation and implementation of a procedure to automatically generate dialogs from the information of a course domain and a student model to adapt and personalize the interaction and contents to each student. That way, it is expected to increase the perceived quality and trustworthiness of the Pedagogic Conversational Agent.

The organization of the paper is as follows: Section 2 briefly reviews related work; Section 3 focuses on the description of the proposal and its current implementation; Section 4 provides some preliminary results attained; and, finally Section 5 ends the paper with some lines of future work.

## 2 Related Work

Natural Language Interaction is a multidisciplinary research area that combines techniques from Natural Language Processing [5] and Human-Computer Interaction [6]. One of its main benefits is that it can help bridging the digital gap by letting users without computer knowledge interact with computers in natural language.

However, it is still not possible to allow users to “talk” to the computer in the same way that they talk to other human because of the difficulty of automatically processing natural language (not only understanding natural language but generating

it) [5]. On the other hand, the progress made in NLP in the last decades and the creation of more resources have made possible to create conversational agents able to keep a dialogue on a certain domain with a specific goal [3].

In particular, Pedagogic Conversational Agents (PCAs) can be defined as virtual characters which can teach or be taught by students in a domain, trying to focus the attention of the student to the topic under study rather than deviating their attention to how to use the e-b/learning platform. Three main types of PCAs can be distinguished according to their role with the students [7]:

- *PCAs that serve as tutors*. For instance, AutoTutor [8] that has successfully been used with university students of literature and physics. Furthermore, AutoTutor keeps a student model to improve the didactics of the tutoring process [9].
- *PCAs that serve as students*. For instance, Betty [10] that has been used with school science students. The pedagogic strategy here is different because the agent is not supposed to be the source of knowledge but to be taught by the student.
- *PCAs that serve as companions*. For instance, MyPet [11] that is an animated dog or cat whose aim is just to be in the system to motivate the student.

### 3 Proposal

Following the work reported in [9], our hypothesis is that the use of Natural Language Processing techniques would allow the generation of dialog templates adapted to each student model and that it would improve the perceived quality and trustworthiness of the Pedagogic Conversational Agent.

In order to test that hypothesis, we are in the process of building a complete Pedagogic Conversational Agent from the core engine of the Willow free-text scoring system [12]. Willow is able to keep a student model with an estimation of the level of knowledge that a student has on the concepts of a course. Currently, the questions of the course and the key concepts have to be specified by the teacher of the course in the Willow's authoring tool, but our aim is that Willow will be able to generate its own questions to dialogue with the student and focus on the less known concepts.

The architecture of Willow would change as shown in Figure 1. In this paper, a first implementation of the module 'Generator of dialogue templates from the student model' is presented. The module is shadowed in the architecture.

As can be seen it gets as input the student model in the form a conceptual model such as the one shown in Figure 2 and currently produces a question in the form 'Is X a kind of Y?' where X and Y are two related concepts in the student's conceptual model. Furthermore, X is a precedent concept according to the Meaningful Learning Theory of Ausubel and it should be learnt before Y. For instance, in an English Learning Course, a teacher could have identified 'apple' and 'fruit' as two key concepts and the module could generate the question 'Is an apple a kind of fruit?'

That way, teachers would not need either to type those taxonomic relationships questions in their course or to type their answers. It is because from the information in the domain model, it is trivial to automatically identify that those concepts are related and thus, Willow would expect an affirmative question for them and vice versa.



## 4 Results

The generator for the taxonomic questions has already been implemented in Java. An independent web system from Willow has been deployed to test the module. The input files provided to the module has been extracted from the students' conceptual and domain models in Willow. In particular, Table 1 shows 5 runs tested with 5 different students' conceptual models ranging from 2 up to 15 concepts. As can be seen, the number of different questions that can be automatically generated is quite high.

**Table 1.** Number of questions generated according to the number of concepts.

Number of concepts	Number of questions
2	1
5	6
11	30
13	36
15	40

## 5 Discussion and Future Work

Given the preliminary nature of the results in this paper, it is still early to provide evidence to support any hypothesis. Nevertheless, the simple fact that 40 questions can be generated from a student conceptual model with 15 questions without the involvement of the teacher and adapting the relationships between the concepts to the level of knowledge automatically is encouraging enough to keep working in this line of research.

It is our intention to keep working in the generator module so that it is not limited just to taxonomic relationships questions, but it is also able to generate more complex questions from the information extracted in the students' answers provided to Willow and the conceptual model kept by the system.

Furthermore, once the implementation is finished, we will carry out an experiment with university students to test whether the use of Willow (being able not only to automatically evaluate the students' answer but also to generate the questions) produces a statistically significant higher perception of quality in the dialogue and trustworthiness to use the system.

Other line of future work that is to be researched with Willow is in the field of personalization. Currently, Willow is represented by an owl. In the last experience with Willow we introduced a first change in that respect and asked the students to use a different avatar for Willow from a set provided.

However, many students did not use that feature. We think that it could be due to the limitation of choosing just among the avatars provided. Therefore, in the next experiment, we will separate the test students who will use the agent Willow into two subgroups: subgroup A will be able to use any figure from Internet or their local disk to personalize Willow's avatar and their own, while subgroup B will not have those possibilities and will remain with the typical Willow's avatars as shown in Figure 3.

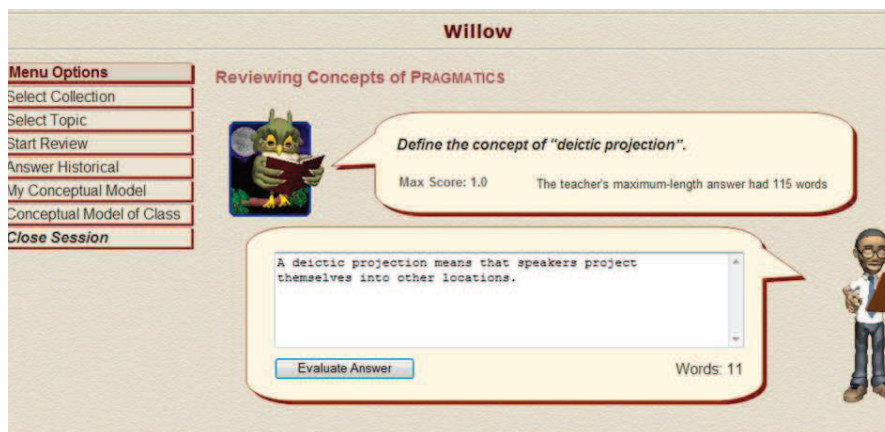


Fig. 3. Snapshot of the Willow interface

**Acknowledgements.** This work has been sponsored by Spanish Ministry of Science and Technology, project TIN2007-64718.

## References

1. Macskassy, S., Stevenson, S.: A conversational agent. Rutgers University (1996)
2. Weizenbaum, J.: Eliza, a computer program for the study of natural language communication between man and machine. *Communications of the ACM* 9, 26-45 (1966)
3. Lester, J., Brandy, K., Mott, B.: *The Practical Handbook of Internet Computing*. Chapman & Hall, chapter Conversational Agents, 220-241 (2004).
4. Ausubel, D.: *The Psychology of Meaningful Verbal Learning*. NY: Grune & Stratton (1963)
5. Mitkov, R.: *The Oxford Handbook of Computational Linguistics*, Oxford Press (2003).
6. Baecker, R.M. Buxton, W.: *Readings in human-computer interaction: A multidisciplinary approach*. Morgan Kaufmann Publishers, CA (1987)
7. Kerly, A., Ellis, R., & Bull, S.: *Conversational Agents in E-Learning, Applications and Innovations in Intelligent Systems XVI - Proceedings of AI*, Springer, London (2008).
8. Graesser, A., Person, N.; Harter, D. Teaching tactics and dialog in AutoTutor, *International Journal of Artificial Intelligence in Education* 12, 3, 23-29 (2001).
9. Jackson, T., Mathews, E., Lin K., Olney, A., Graesser, A.: *Modeling Student Performance to Enhance the Pedagogy of AutoTutor*. LNCS 2702, Springer, UM (2003)
10. Wagster, J., Kwong, H., Segedy, J., Biswas, G., Schwartz, D.: Bringing CBLEs into Classrooms: Experiences with the Betty's Brain System. *IEEE ICALT*, 252-256 (2008).
11. Chen, Z., Liao, C., Chien, T. & Chan, T.: Animal Companion Approach to Fostering Students' Effort-Making Behaviors, *in 'Artificial Intelligence in Education'* (2009).
12. Pérez-Marín, D., Alfonseca, E., Rodríguez, P. & Pascual-Nieto, I. Willow: Automatic and adaptive assessment of students free-text answers, *in 'Proceedings of the 22nd International Conference of the Spanish Society for the Natural Language Processing'* (2006).