Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence

Extending Computer Assisted Assessment Systems with Natural Language Processing, User Modeling and Recommendations Based on Human Computer Interaction and Data Mining

Ismael Pascual-Nieto[1], Olga C. Santos[1], Diana Perez-Marin[2], Jesus G. Boticario[1]
[1] aDeNu Research Group. Computer Science School. UNED
[2] Laboratory of Information Technologies in Education (LITE). Universidad Rey Juan Carlos
ismael.pascual@bec.uned.es, ocsantos@dia.uned.es, diana.perez@urjc.es, jgb@dia.uned.es

Abstract

Willow is a free-text Adaptive Computer Assisted Assessment system, which supports natural language processing and user modeling. In this paper we discuss the benefits coming from extending Willow with recommendations. The approach combines human computer interaction methods to elicit the recommendations with data mining techniques to adjust their definition. Following a scenario-based approach, 12 recommendations were designed and delivered in a large scale evaluation with 377 learners. A statistically significant positive impact was found on indicators dealing with the engagement in the course, the learning effectiveness and efficiency, as well as the knowledge acquisition. We present the overall system functionality, the interaction among the different subsystems involved and some evaluation findings.

1 Introduction

Assessment is essential to learning [Berry, 2003]. For this reason, Computer Assisted Assessment (CAA) systems were developed to support the assessment process through objective testing, such as multiple choice questions and fill-in-the-blank exercises. However, evaluating the learners' learning progress in this way is not enough to measure higher cognitive skills [Birenbaum et al, 1993; Mitchell *et al.*, 2003] and limits the feedback that can be provided to the learners in terms of self-regulation features [Chang 2005]. Natural Language Processing (NLP) can be used to automatically assess the learners' answers in free-text [Valenti *et al.*, 2003; Mitkov, 2003]. In fact, there are more than 30 CAA systems relying on different techniques and applied to several domains [Pérez-Marín *et al.*, 2009].

A few of them, the so-called free-text Adaptive CAA (ACAA) systems are also able to track some information in the learners' answers to automatically generate learner models (i.e. user modeling – UM), and use that information to provide feedback adapted to each learner [Aguilar & Kaijiri, 2007]. Moreover, if the learner model is open for scrutability purposes [Kay, 1999] and shows users the information known about them and enables them to manage this information [Bull and Kay, 2008], it can increase self-learning awareness and even the learner's motivation [Hummel *et al.*, 2005].

In this paper, we present the approach followed to extend a free-text ACAA system with recommendation possibilities combining artificial intelligence (AI) techniques from different areas. The integrated system has been evaluated in a large scale evaluation with 377 learners, and the analysis of the evaluation has shown a statistically significant positive impact on indicators dealing with the engagement in the course, the learning effectiveness and efficiency, as well as the knowledge acquisition.

First, we introduce the research background. The next section introduces the approach and presents a scenario to illustrate the overall system functionality. It identifies the different subsystems involved, explaining their focus as well as the interactions among them. After that, some findings are shown to discuss on the benefits and improvement gained. Finally, some conclusions are compiled.

2 Background

Willow is a free-text ACAA system [Pérez-Marín *et al.*, 2009]. As any ACAA systems, it is able to automatically process short learner answers or essays written in natural language and provide feedback on the responses given, which are adapted to each learner's preferences and level of knowledge. Learner's preferences are gathered from Willow user interface, and can be, for instance, the topics that she wants to get questions of. Willow uses a combination of statistic techniques and Latent Semantic Analysis (LSA). The statistic techniques are focused on processing the style of the answer, dealing with synonyms and word sense disambiguation. The LSA is more focused on processing the answer content, so that words are related not because they are next to each other, but because they are similar in semantics. By using a corpus of more than one hundred questions and answers, and computing the Pearson correlation between the automatic scores and a human teacher scores for the same set of questions, the combination of the statistic and LSA has reached values around 70% depending on the type of question (definitions usually are more easily evaluated than questions asking for examples, which are the hardest to automatically evaluate).

Moreover, Willow follows an open learner model strategy showing the evolving learner model to the learners and the teachers. In this way, it allows learners to track their evolution over the learning process, and teachers to find out up to which point the learners understand the lessons.

Willow has been widely used in blended learning contexts to support learners at home when practicing the theory given in the face to face sessions by reviewing the concepts through questions [Pérez-Marín *et al.*, 2009]. In the research reported here the purpose was to evaluate the usage of ACAA in a full e-learning context, where the physical presence of the teacher is not available.

To achieve our goal, we researched on existing alternatives from the AI field to improve the system functionality by using educational criteria to provide a personalized guidance to the learners while using the ACAA system. To this end, we followed the Semantic Educational Recommender Systems –SERS– approach [Santos and Boticario, 2010a], which differs from traditional Educational Recommender Systems (ERS) in the semantic characterization of the recommendations. While ERS focus on applying similar techniques from e-commerce to retrieve learning objects which can be of interest to a learner [Manouselis *et al.*, 2010], SERS apply human computer interaction (HCI) techniques to involve educators in identifying recommendation needs for their e-learning scenarios. These HCI techniques consist in the application of usability methods [Bevan, 2003] along the user-centered design (UCD) cycle as defined in the ISO-9241-210 standard [ISO, 2010], such as the Rosson and Carrol scenario-based approach [Rosson and Carrol, 2001] or the card sorting techniques [Spencer, 2009]. To guide the recommendations' elicitation process with HCI techniques, we proposed the TORMES methodology [Santos and Boticario, 2011a], which describes how this UCD methods can be used to help the educator identifying appropriate recommendations. Very briefly, according to TORMES, the educator observes her current scenario -where she wants to identify recommendations needs to improve the learning experience of her learners. For this, UCD methods support her in the process. Once the educator identifies the recommendations, she can model them in terms of a semantic recommendations model –SRM– [Santos and Boticario, 2010b]. Designed recommendations can be complemented with findings coming from the application of data mining (DM) techniques, which can suggest, for instance, that learners which a non-collaborative profile should receive an encouraging recommendation. Once the recommendations are designed by the educator, these recommendations descriptions are tested by other educator and learners, thus helping their refinement. Once ready, they can be delivered in the system to be offered in the runtime scenario and their impact evaluated. Learners are involved in the process of the recommendations design, but are not leading it.

Involving the educator in the recommendations elicitation process is essential to obtain qualitative information from an educational perspective, which can be considered in the recommendation description. This information is managed in the SRM and covers issues such as what to recommend and in which situation (i.e., applicability conditions), as well as to characterize the recommendation with some metadata (e.g. category, relevance, origin) to support the reasoning process. These qualitative descriptions can be complemented with DM, which can be used to tune the educators' design work with additional attributes or specific values for the applicability conditions by analyzing previous interactions in the learning environment. Once the recommendations are designed, the information stored in the SRM is used by the SERS at runtime to automatically deliver the appropriate recommendations to a given learner in a specific learning situation. Recommendation are defined in a general way, and then instantiated for the specific object that is to be recommended

The goal of extending Willow with the SERS approach is to support learners to review concepts by answering free-text questions in a full e-learning scenario guided by educational recommendations. To the best of our knowledge [Santos and Boticario, 2011b], no free-text CAA or ACAA system has ever incorporated the possibilities of a SERS (or an ERS).

3 Extended system

This paper shows how a free-text ACAA system can be extended with recommendation possibilities combining AI techniques from different areas. As introduced above, traditionally, free-text ACAA systems have just focused on the assessment of open-ended questions, providing feedback to the students' answers in free-text. However, the feedback was limited to the content and the style of the answers. As for the content, it was usually indicated missing information or wrong data. Regarding the style, it was usually indicated the mistakes in grammar or in the use of terminology and syntax [Valenti *et al.*, 2003; Mitkov, 2003]. On the other hand, ERS (or SERS) can focus on providing hints to learners about what they should study next, or to assist them in their interaction with the learning environment. By combining both approaches (i.e. free-text ACAA and SERS) the extended system would not only automatically assess learners' free-text answers, but it would advise learners about what to do next, and how to solve doubts in the use of the system in case they have any.

Let us suppose the following scenario. Maria is a Primary Degree student who is using for the first time a free-text ACAA system. She thinks that it could be useful for her to review the concepts studied in class. However, she is afraid of not knowing how to correctly use the system because it is the first time that she uses a system like that by her own. In case that it were just a free-text ACAA system, she would be presented questions and she would not receive any help or hint to guide her with educational criteria during the interaction. When the free-text ACAA system integrates the SERS approach, right after Maria logs into the system, she can take into account the actions suggested, which aim to provide a personalized guidance based on educational criteria. Thus, she is told what the most appropriate actions for her to take in the system are at any moment to support her learning process and facilitate her to grasp the full functionality of the learning environment. Moreover, when she answers the open-ended questions, she is not only provided with feedback about the style or the content, but she is also recommended some actions to encourage her to keep studying the concepts that have been identified as less understood at the same time that she is pointed to appropriate sections in the learning environment where useful material (e.g. concepts definitions) can be found. In this way, the overall system combines two independent approaches (ACAA and SERS) which involve several subsystems. To provide the overall functionality (i.e. to support learners to review concepts by answering free-text questions in a full e-learning scenario guided by educational recommendations), each of the subsystems deals with the following functionality:

- **CAA** (Computer Assisted Assessment): automatically scores the learners answers and gives feedback to them according to their responses.
- **NLP** (Natural Language Processing): assesses the free-text answers provided by the learners in response to the questions formulated by the system.
- **UM** (User Modeling): dynamically adjusts the difficulty level of the questions provided to the learners according to their knowledge.
- **RS** (Recommender System): selects and delivers the appropriate recommendation to a given learner in a given context to provide a personalized guidance based on educational criteria following a rule-based approach.
- **HCI** (Human Computer Interaction): applies UCD methods to involve the educator in identifying recommendation needs in learning scenarios and describes them in terms of the SRM.
- **DM** (Data Mining): extracts knowledge from past interactions of the learners in order to discover relevant attributes as well as specific values for the selected attributes to be used as applicability conditions by the rule engine.

The interaction process among the above subsystems covers design and runtime issues. At **design time**, educators prepare the questions to be offered by the CAA to the learners as well as the reference answers that will be compared to the learners' answers with NLP techniques. These questions are categorized in terms of the difficulty level so that the UM process at runtime can select the appropriate one for the current learner. Moreover, following the scenario based approach (i.e. HCI methods), educators identify appropriate recommendations to be delivered by the RS that can support learners during their interaction in the given learning scenario. In turn, educators can be supported in terms of DM methods to past interaction data, which aim at identifying relevant features to be considered in the recommendation process as well as specifying appropriate values for the applicability conditions. For instance, a feature (or attribute) for the applicability conditions can be the learning style of the learner, and its value, one of the dimensions according to Felder's and Silverman's, theory [Felder and Silverman, 1988], e.g. reflective.

At **runtime,** the CAA presents learners questions about the course concepts to reinforce their learning. The learner is required to provide free-text answers to those questions. Following the dialogue metaphor in HCI, the system offers a user friendly interaction paradigm to the learners. Figure 1 shows the question posed by the system (represented by the avatar at the top) and the text field provided to the learner to write her answer in natural language (at the bottom).

Once the learner provides her answer, the system evaluates the text written by the learner using NLP, scores the learner's answer and provides feedback to her through the CAA. Moreover, it takes into account the conceptual model of the user, which is derived from UM techniques, to decide which question give her next.

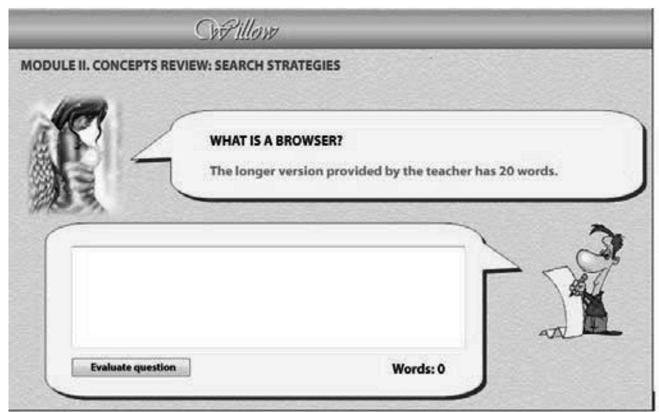

Figure 1. Dialogue in Willow to practice concepts

In terms of the dialogue metaphor, the learner can receive educational oriented recommendations delivered by the RS to guide her during the interaction with the system, as shown in Figure 2. In this case, the system (at the top of Figure 2) offers her a list of suggested actions that can be appropriate to be carried out by the learner according to her user model and course context. Recommendations are defined in a general way, and then instantiated for the specific object that is to be recommended. Each recommendation shows *what action* is recommended on a *given object* in Willow. An *object* is any element available in Willow, and can be simple elements such as a message in the forum, a content item, the definition of a concept, or more complex elements such as the user model. A detailed description for each recommendation (which includes the justification and the semantic information proposed in [Santos and Boticario, 2010b]) is provided by clicking in the right icon that follows each recommendation. The learner has the freedom to follow each recommendation. When followed, the next time she accesses the page with the recommendation list, she can provide feedback on the utility of the recommendation (at the bottom of Figure 2) as three possible values (useful, not needed, not given in the appropriate moment). This information is to be used by DM to adjust the design of the recommendations, in particular, the i) applicability conditions required to deliver the recommendation and ii) its relevance.

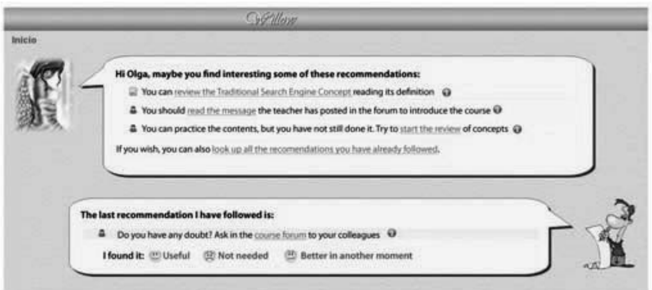

Figure 2. Dialogue in Willow to show the recommendations

The implementation of the different components that make up the extended system has been done in J2EE and the integration follows the web services approach. Moreover, Weka suite has been used for the DM process. The overall system is deployed using Tomcat application server. The MySQL database is used to store the data.

4 Benefits from the approach

The benefits of the ACAA approach compared to CAA are described elsewhere [Pérez-Marín *et al*., 2006]. These authors reported that when the learners in their experiment were asked about the system's features one by one, most preferred the ACAA system because it fitted better their needs, the order of the questions was more adequate, and they felt more satisfied as the system controlled their progress. In particular, the learners who used Willow found its use more amusing and they felt more engaged to keep answering questions.

In turn, the usefulness of the SERS approach has also been described elsewhere [Santos and Boticario, 2010b]. These authors carried out an experiment with 40 learners and reported that i) learners are already demanding a personalized support in terms of recommendations, ii) they trust the recommendations coming from the tutor and appreciate those that take into account their own preferences, and iii) learners are interested in being informed about the recommendation features.

Drawing on the lessons learnt from the aforesaid two experiences, we have designed a large scale evaluation to assess the improvement gained by combining both approaches, the ACAA and the SERS. Thus, the extended system is able to support the learners in reviewing concepts by answering free-text questions in a full e-learning scenario guided by educational recommendations.

4.1 Course design

We prepared a course on 'Search strategies in the Web with Educational Goals' addressed to teachers who wanted to improve their teaching skills with information and communication technologies. Two modules were designed with an estimated dedication time of 20 hours per module to be worked over a two-week period. Each module comprised 3 lessons. For each lesson, a set of 40 slides was prepared. Moreover, for the reviewing process in Willow, 15 questions with their correct answers were defined for each lesson. 20 key concepts were covered in each module.

A forum was available on each module, where 3 threads were created, one per lesson. Learners were not allowed to create new threads, but they could only replay to existing ones. A welcome message was posted to the forum to explain the course operation and included some documents on how to use Willow. Information about the course was given to the learners in an e-mail message sent when enrolled in the course.

Following the scenario-based approach, 12 recommendations were elicited by the 2 educators involved in the process. Table 1 gathers them.

ID	Recommendation
R1	Choose a lesson to review
R2	Start the review of the concepts
R3	Review the concept estimated as less known
R4	Use the forum to share a doubt
R5	Read a thread of the forum with many posts
R6	Read the educators' welcome message
R7	Change the avatar that represents Willow
R8	Change the avatar that represents the learner
R9	Look at the learner conceptual model
R10	Look at the conceptual model of the class
R11	Log in Willow to start the course for the first time
R12	Log in Willow to keep reviewing the contents

Table 1. Recommendations defined by the educators

Due to space limitations, their descriptions in terms of the SRM have been omitted, which should include specific details for the particular object to be recommended, the applicability conditions and the semantic information. The recommendations were elicited by the educators from the e-learning scenario identified. Since there was no previous experience on an e-learning scenario, interaction data from previous blended learning scenarios were analyzed with classification and clustering techniques to suggest values for the applicability conditions, when applicable.

4.2 Experimental settings

A large scale evaluation was designed to asses our approach. The initial baseline of the research was that there was already an ACAA system, which worked as an e-learning platform. Then, on top of this ACAA system, we wanted to test if adding recommendations improves the learning experience on the ACAA system. Thus, we needed to test the ACAA system with and without the SERS. Other combinations that do not considered the ACAA did not provide functionality that could be evaluated.

Participants were collected from past learners of a course on learning to teach through the internet, which is offered as an ongoing education course to teachers at the Spanish National University for Distance Education (UNED). Two editions of the course were carried out, taking part a total of 377 learners (40% male, 60% female) whose main experience came from the face to face education. Learners were randomly assigned to the control group (which did not receive recommendations, that is, used the ACAA part of the system) and the experimental group (which could receive recommendations, that is, used the combined system with ACAA and SERS included).

We uploaded the materials into the course in Willow, created the welcome message into the forum and let the learners review the contents with Willow. It was designed to be delivered completely on-line without any face-to-face interaction.

For the evaluation, several sources of information were prepared. On the one hand, qualitative data was gathered through several questionnaires to the learners: 1) a profile questionnaire, 2) pre and post tests to control the knowledge gain, 3) satisfaction questionnaire focused on the perception in each module, and 4) global satisfaction questionnaire at the end of the course to get further details on the learning experience. Regarding quantitative information, it was collected from the data stored in the database and included: 1) active (contributions) and passive (visits) data, 2) information on the sessions, 3) information about the recommendations delivery (e.g. time when it was offered to each learner, time when it was followed by the learner, order in which was given in the list, time when the explanation information was read by the learner, feedback given by the learners). Moreover, Willow also provides some activity graphics on the reviewing process.

4.3 Findings

The improvements derived from combining the ACAA and the SERS approaches were evaluated by measuring the impact of the recommendations on the learners. For this, a set of indicators that involve usability and educational measures were defined, namely effectiveness, efficiency, satisfaction, engagement and knowledge acquisition. These indicators are dependent on the system, so specific indicators were defined for Willow.

A statistically significant analysis was carried out to assess if there was a significant impact (with confidence level set to 95%) on the indicators coming from the experimental group (i.e. learners who can receive recommendations) with respect to those from the control group (i.e. learners who do not receive recommendations).

Several findings were identified from the analysis of the responses to the questionnaires and the interaction data registered in the database.

The first finding is that there was not a significant impact on the *satisfaction* when the SERS approach was integrated in the ACAA. In both groups, learners reported very high satisfaction levels. For instance, the percentage of participants who found the course interesting was around 75% in any of the modules and groups.

In turn, the second finding showed a positive statistical impact on indicators dealing with the *engagement* measure (mainly in terms of sessions, hits, days connected and time spent), since almost all recommendations impacted on all the engagement indicators.

The third finding dealt with the *efficiency*. A positive statistical impact was found on some indicators (i.e. 'post test done within the module time frame' and 'average sessions to get to the end page of the module').

Regarding *eff*ectiveness and *knowledge* acquisition, no impact was found when comparing the recommendations as a whole. However, by analyzing the impact in each of the recommendations, some statistically significant impact was found in some of them. In particular, some recommendations impacted on the effectiveness (i.e. more questions answered by learners and more learners doing the post test) and the knowledge acquisition (i.e. more correct answers). Thus, the fifth finding suggests that the impact on the *eff*ectiveness and *knowledge* acquisition depends on the educational goal of the recommendation.

These findings showed that by extending Willow with recommendations an improvement was perceived on indicators that relate to the engagement in the course and the learning efficiency. Moreover, depending on the nature of the recommendations designed, some improvement can also be produced on indicators dealing with *eff*ectiveness and knowledge acquisition. In this way, the combined approach (ACAA and SERS) provided support to learners at home when practicing the theory given in the face to face sessions in a full e-learning context.

5 Conclusions

In order to provide intelligent functionality in any system, AI is required and usually, techniques from more than one area are combined. In this paper we have focused on how AI has been integrated and embedded to support computer-aided education. In particular, we report the evaluation of a free-text ACAA system which has been extended with recommendations that consider educational issues provided by educators. To this, the overall system integrates the ACAA and SERS approaches, supported by two previously developed running systems, which respectively involve CAA, NLP, UM and RS, HCI, DM. Their functionalities are combined into the overall system to support learners to review concepts by answering free-text questions in a full e-learning scenario guided by educational recommendations. These two approaches were previously applied in real educational settings reporting benefits by themselves. However, no evidence existed on the improvement gained due to their combination.

To assess the improvement gained by the combined approach, we have carried out a large scale evaluation with 377 learners, where 2 modules with 45 questions over 20 concepts were offered by the system to reinforce the learning of the concepts. 12 recommendations were proposed after the application of user centered design methods and data mining techniques, and delivered in a personalized way to guide learners in their interaction with the system if needed. The results reported from this evaluation showed that a positive statistical impact has been found on indicators dealing with the engagement in the course, the learning effectiveness and efficiency, as well as the knowledge acquisition. These findings suggest that the scrutability support coming from the combination of the open learner model and the explicit description of the recommendations may have strengthened the learner motivation to use the system, which is expected to impact on the learning process.

Acknowledgments

The research outcomes presented in this paper have been possible thanks to the funding obtained in several projects. For Willow, U-CAT (TIN2004-03140), and for the SERS approach, A2UN@ (TIN2008-06862-C04-01/TSI) and EU4ALL (IST-2006-034478).

References

[Aguilar and Kaijiri, 2007] Aguilar, G. and Kaijiri, K. *Design Overview of an Adaptive Computer-based Assessment System.* Interactive Educational Multimedia 14, 116-130, 2007.

[Berry, 2003] Berry, R. *Alternative assessment and assessment for learning*, in Proceedings of the 29th IAEA Conference, theme: Societies Goals and Assessment, 2003.

[Bevan, 2003] Bevan, N. *UsabilityNet Methods for user centered design.* In: Jacko, J. and Stephanidis, C. eds. Human-Computer Interaction: Theory and Practice (Part 1), Volume 1. Heraklion, Crete: Lawrence Erlbaum, p. 434-438, 2003.

[Birenbaum *et al*, 1993] Birenbaum, M., Kelly, A. and Tatsuoka, K. *Diagnosing Knowledge States in Algebra Using the Rule-Space Model*, Journal for Research in Mathematics Education 24(5), 442-459, 1993.

[Bull and Kay, 2008] Bull, S. and Kay, J. *Metacognition and Open Learner Models*, in I. Roll & V. Aleven (eds), Proceedings of Workshop on Metacognition and Self-Regulated Learning in Educational Technologies, International Conference on Intelligent Tutoring Systems, 7-20, 2008.

[Chang, 2005] Chang, M.M. Applying self-Regulated Learning Strategies in a Web-Based Instruction-an Investigation of Motivation Perception. Computer Assisted Language Learning, Vol. 18(3), 217-230, 2005.

[Felder and Silverman, 1988] Felder R. M. and Silverman L. K. *Learning and Teaching Styles In Engineering Education*, Engr. Education, 78(7), p. 674-681, 1988.

[Hummel *et al.* 2005] Hummel, H.G.K., Burgos, D., Tattersall, C., Brouns, F., Kurvers, H., Koper, R. *Encouraging contributions in learning networks using incentive mechanisms*. Journal of Computer Assisted Learning, Vol. 21, N.5, Oct. 2005, 355-365(11), 2005.

[ISO, 2010] ISO Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems. ISO 9241-210, 2010.

[Kay, 1999] Kay, J. *Ontologies for reusable and scrutable student models*. Workshop on Ontologies for Intelligent Educational Systems, 72-77, 1999.

[Manouselis *et al*., 2010] Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H. and Koper, R. *Recommender Systems in Technology Enhanced Learning*, in Kantor P., Ricci F., Rokach L., Shapira, B. (Eds.), Recommender Systems Handbook, Springer, 2010.

[Mitchell *et al.*, 2003] Mitchell, T. Aldridge, N., Williamson, W. and Broomhead, P. *Computer Based Testing of Medial Knowledge*. 7th Computer Assisted Assessment Conference, 249-267, 2003.

[Mitkov, 2003] Mitkov, R. *The Oxford Handbook of Computational Linguistics*, Oxford Univ. Press, 2003.

[Pérez-Marín *et al*., 2006] Pérez-Marín, D., Alfonseca, E. and Rodríguez, P. *On the Dynamic Adaptation of Computer Assisted Assessment of Free-Text Answers*. In proceedings of AH 2006, 2006.

[Pérez-Marín *et al.*, 2009] Pérez-Marín, D.; Pascual-Nieto, I. and Rodriguez, P. *Computer-assisted assessment of free-text answers*, The Knowledge Engineering Review 24(4), 353–374, 2009.

[Rosson and Carroll, 2001] Rosson, M. B. and Carroll, J. M. *Usability engineering: scenario-based development of human computer interaction*. Morgan Kaufmann, 2001.

[Santos and Boticario, 2010a] Santos, O.C. and Boticario, J.G. *Usability methods to elicit recommendations for Semantic Educational Recommender Systems*. IEEE Learning Technology Newsletter. Vol. 12, Issue 2, 2010.

[Santos and Boticario, 2010b] Santos, O.C. and Boticario, J.G. *Modeling recommendations for the educational domain*. Workshop Recommender Systems for Technology Enhanced Learning, pp. 2793-2800, 2010.

[Santos and Boticario 2011a] Santos, O.C. and Boticario, J.G. *TORMES methodology to elicit educational oriented recommendations*. 15th International Conference on Artificial Intelligence in Education (AIED 2011), in press.

[Santos and Boticario 2011b] Santos, O.C. and Boticario, J.G. Handbook on Educational Recommender Systems and Technologies: Practices and Challenges. IGI Publisher, in press.

[Spencer, 2009] Spencer, D. Card Sorting. Designing Usable Categories. Rosenfeld Media, 2009.

[Valenti *et al*., 2003] Valenti, S., Neri, F. and Cucchiarelli, A. *An Overview of Current Research on Automated Essay Grading*, Journal of Information Technology Education 2, 319—330, 2003.