RUSC Revista de Universidad y Sociedad del Conocimiento
Universities and Knowledge Society Journal
http://rusc.uoc.edu

ARTICLE

An Answering System for Questions Asked by Students in an e-Learning Context

Joaquim Moré
jmore@uoc.edu
Computational Linguist, Office of Learning Technologies, Open University of Catalonia (UOC)

Salvador Climent
scliment@uoc.edu
Lecturer, Arts and Humanities Department, Open University of Catalonia (UOC)

Marta Coll-Florit
mcollfl@uoc.edu
Lecturer, Arts and Humanities Department, Open University of Catalonia (UOC)

Submitted in: May 2011
Accepted in: February 2012
Published in: July 2012

Recommended citation

MORÉ, Joaquim; CLIMENT, Salvador; COLL-FLORIT, Marta (2012). "An Answering System for Questions Asked by Students in an e-Learning Context" [online article]. *Universities and Knowledge Society Journal (RUSC)*. Vol. 9, No 2, pp. 229-239 UOC. [Accessed: dd/mm/yy].
<http://rusc.uoc.edu/ojs/index.php/rusc/article/view/v9n2-more-climent-coll-florit/v9n2-more-climent-coll-florit-eng>
<http://dx.doi.org/10.7238/rusc.v9i2.1161>
ISSN 1698-580X

Abstract

In this article, we present a system that helps tutors to answer questions asked by their students at an online university: the Open University of Catalonia (UOC). Communication between UOC students and their tutors is fully online; students ask questions and tutors answer them by e-mail. The system, which is currently being developed at the UOC's Office of Learning Technologies (OLT), aims to find multilingual contexts with useful information to enable tutors to give fast, appropriate answers to students. These contexts are extracted from course learning materials, from previous messages posted on subject discussion boards, and also from articles and other sources of information available on

© Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012
© FUOC, 2012

the Internet. Apart from helping tutors to find better answers, the system is also useful for updating their knowledge and contributing to their lifelong learning.

Keywords
question-answering systems; e-learning; speech acts; tutor

Un sistema de respuestas a consultas formuladas por alumnos en un contexto de aprendizaje virtual

Resumen
En este artículo presentamos un sistema que ayuda a los docentes a responder las preguntas de sus alumnos en una universidad virtual, concretamente la Universitat Oberta de Catalunya (UOC). La comunicación entre alumno y docente se realiza de forma totalmente virtual: las preguntas y las respuestas se formulan y contestan mediante correo electrónico. El sistema, que se está desarrollando en el Área de Tecnología Educativa de la UOC, tiene como principal objetivo encontrar contextos multilingües con información útil para responder al estudiante de forma rápida y adecuada. Los contextos se extraen de los materiales del curso, los foros de participación de la asignatura, artículos y otras fuentes de información disponibles en internet. Además de ayudar a los docentes a encontrar mejores respuestas, el sistema también es útil para actualizar sus conocimientos y desarrollar su aprendizaje permanente.

Palabras clave
sistemas de pregunta-respuesta, aprendizaje virtual, actos de habla, tutor

1. Introduction

At an online university, tutors are faced with the arduous task of writing answers —containing detailed, useful information— to every query and question they receive daily from their students. The task can take up a considerable amount of the time they spend each day on teaching if the number of students is high and the tutors have a narrow timeframe in which to reply to them all.

In most question-answering systems for e-learning tutors (Hung et al., 2005; Feng et al., 2006; Wang et al., 2006; Yang, 2009), questions are answered automatically, but this approach has to overcome a number of problems. This includes the automatic recognition of a question when a student asks it implicitly rather than explicitly, with deviations from formal, normative expression. For example, spelling mistakes or typographical errors are commonplace in e-mails, as are unstable syntax and other phenomena, thus making the question more difficult to identify. Another problem is searching for an answer to a question that, despite being thematically related to the subject, does not refer to an aspect covered by course learning materials. In this case, traditional solutions based on information contained in learning materials are of no use, since such solutions rely on the extraction of information from a database with question-answer pairs, or from an annotated corpus (Feng et al., 2006; Wang et al., 2006; Yang, 2009). A solution that aims to extend beyond learning materials is the retrieval of an answer to a question asked by a student in an open online community (Bernhard

© Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012
© FUOC, 2012

& Gurevych, 2008). The problem with this, however, is that the answers returned may be absurd and very unreliable. Consequently, the system would have to learn to discriminate between good and bad answers, and this is simply too demanding for current systems.

Besides such problems, which partially account for a tutor's lack of trust in automatic question-answering engines, these systems do not contemplate an important aspect that we identified in the UOC tutors' messaging. Questions often prompt tutors to search for the latest information; by doing so, they update their knowledge. Many questions arise from a student's reflections on an exercise or on recommended reading and, as a result, tutors have to find an answer on an aspect that they had not considered or, quite simply, that they were unaware of. Students thus foster their tutors' lifelong learning.

In this article, we present a help tool for tutors, the aim of which is not to get an exact answer to a question, but rather to find contexts with useful information to enable them to give a fast, appropriate answer to students. This aim allows for the development of a method that is more flexible than the traditional one for answer search systems. The best result is obtained when there is a good direct question among the contexts that the system has found. The system cuts down on information search times, allows tutors to update their knowledge and is useful for assessing students' contribution to the acquisition of information by their fellow students and also their tutors.

This article is structured in the following manner. Section two presents the methodology, which is based on a pragmatic theory. Sections three and four describe the prototype that we have developed thus far and present an evaluation of it. The final section contains the conclusions and future work.

2. Methodology

Our system's users are the virtual classroom tutors for all of the UOC's bachelor's degrees and programmes. Accordingly, we have developed a methodology that is independent from specific thematic domains. We decided to approach the problem by positioning it within a theoretical framework, in this instance Searle's theory of speech acts (1969), which describes the bases of communication between a speaker and a hearer.

In a communicative situation where tutors and students interact by e-mail, students have objectives that they expect to be met with the tutors' help. It is therefore crucial for students to formulate speech acts whose linguistic traits clearly indicate their expectations to tutors. Conversely, speech acts that tutors formulate in their answers must contain linguistic traits confirming that they meet students' expectations.

A speech act consists of two elements. The first is the speech act expression (SAE); it is the expression by which the hearer identifies the speaker's expectations.

"I don't understand" is an example of the way in which a student expresses an expectation that someone will clarify a concept.

The second element is a speech act object (SAO); it is the key term (or terms) of the speaker's speech act. For example, if a student says "I don't understand the notion of hyponymy," then *hyponymy* is the SAO.

© Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012
© FUOC, 2012

Our hypothesis is that document segments that are useful to tutors are those containing the SAOs of a message. The question we therefore asked ourselves was this: to what extent do the SAOs of a message found in a document contribute to our system's usefulness? With this in mind, the prototype that we have developed searches for contexts in reliable sources of information in which SAOs co-occur. SAO candidates are identified automatically. However, it is up to the tutors to select the most relevant ones, since they are capable of working out students' intentions despite the fuzzy discourse relationships that are characteristic of informal e-mails.

3. Prototype

Thus far, we have developed a prototype that searches for contexts in which relevant terms contained in a student's message co-occur. In this section, we shall present the prototype procedure and the sources of information that it consults.

The prototype procedure has four stages, as shown below:

1. *Extraction of the subject of the message*

The subject of the message is extracted in order to retrieve other messages from discussion boards with the same or a similar subject.

2. *Automatic morphological analysis of the body text of the message*

The system morphologically analyses the body text of the message using the FreeLing parser (Atserias et al., 2006). Most of the messages are written in Catalan, so it is set as the default source language.

3. *Tag cloud generation*

The system presents a tag cloud with the relevant concepts found in a student's message so that tutors can select the SAOs, that is to say, the concepts that they want to focus on in order to find useful contexts and thus be able to give a good answer. The tag cloud is generated by an automatic terminology extractor. Expressions between quotation marks, verbs and nouns are extracted. Terms probably belonging to the conceptual domain of the subject of the message are highlighted. However, in order to identify these terms, a system that is independent from conceptual content has been used. The system consults the open-source Catalan-English DACCO dictionary (http://sourceforge.net/projects/dacco/), which includes information about the frequency of its entries in terms of the number of results returned by Google. If we start from the hypothesis that terms related to a specific conceptual domain return fewer results than general-vocabulary terms, then the highlighted terms are those that fall below a numeric threshold of results.

4. *Useful contexts search*

After selecting the SAOs from the tag cloud, the system searches for contexts in Catalan and English in which the denominations of the selected objects co-appear in both languages. We refer to these contexts as *useful context candidates* (UCCs). The sources of information consulted for the extraction of UCCs are the following:

© Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012
© FUOC, 2012

- Messages on subject discussion boards, written in previous semesters: the same question may have been asked in a previous semester, to which a tutor or a student may have given a good answer.
- Subject learning materials: the system uses a learning materials search engine developed by the UOC to find learning materials contexts in which the selected concepts co-appear.
- Wikipedia: links to Wikipedia in Catalan and English, where the terms selected by a tutor are explained.
- Online scholarly articles: for the prototype, the system used the Delicious (http://www.delicious.com) search engine to find articles in Catalan and English whose tags matched the concepts selected by a tutor. The system also used the search engine of CiteULike (http://www.citeulike.org/), which is a free online service that classifies scholarly publications and retrieves articles by the same method. So, the prototype showed the results pages from Delicious and CiteULike, with links to articles covering topics related to the selected terms.

Other free scholarly services can be added if tutors consider that to be necessary to obtain good answering contexts.

4. Evaluation

The prototype was evaluated to obtain two types of information: first, the prototype's usefulness in terms of enabling tutors to find and give appropriate, fast answers to students, and second, the contribution of each source of information to the prototype's usefulness. A number of elements that could be improved were also identified.

4.1. Evaluation procedure

For the evaluation, the subject General Linguistics II was chosen and two groups of evaluators were formed. The first group, called *experts*, comprised three subject consultants. The second group, called *novices*, comprised three subject specialists that had not had any experience as UOC consultants. Consequently, we were able to evaluate whether the prototype was more useful for novices than for experts, and vice versa.

Forty messages were selected for the evaluation. There are two explanations for the number of messages and the semesters they cover: first, the discussion board database only included the last two semesters, and second, we ruled out any messages that, in the opinion of at least three evaluators, were too decontextualised or unspecific to draw any useful contexts from.

A web environment was used, which was organised as follows:

- A space set aside for selecting the message that an evaluator wanted to view. This space contained a list of numbers from 1 to 40, each of them being a message reference number.
- A space set aside for displaying the message.

© Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012
© FUOC, 2012

- A space set aside for selecting the SAOs from a tag cloud of terms in the message. After selecting the SAOs, the evaluators clicked a button to start the prototype's search for relevant contexts in order to give an answer.

The prototype displayed the UCCs found in each source of information, in accordance with the terms selected from the tag cloud. The task for both expert and novice evaluators was to score the usefulness of the source of information according to the UCCs displayed. The evaluators recorded their scores on a spreadsheet organised by the message-source of information relationship. If, on a subject discussion board for example, an evaluator found that an answer given by a student to the same question was very useful, then that evaluator gave the highest score to the 'subject discussion board' source.

The usefulness of the contexts had two dimensions. The first was the usefulness of a source of information in terms of giving a good answer (UGA). The second was the usefulness of a source of information in terms of giving a fast answer (UFA). The items were scored on a scale of five values: 0 (Not useful), 1 (Not very useful), 2 (Useful), 3 (Very useful) and NC (no context). The latter was used when the system was unable to retrieve any UCCs from the sources of information. In addition, the evaluators were encouraged to write comments on the effort and difficulty involved in obtaining useful contexts. These comments provided us with very useful pointers as to how we could improve the system.

The analysis evaluation procedure was divided into two stages; a macroevaluation and a microevaluation. The macroevaluation was an analysis of the results related to the system's usefulness, that is to say, its usefulness in terms of giving a good answer and a fast answer. The purpose of the microevaluation was to determine the contribution of each source of information to the system's usefulness, thus allowing any aspect that could be improved to be identified.

4.2. Macroevaluation

We wanted to compare the experts' and novices' perceptions of usefulness in terms of giving a good answer. First of all, we calculated each group member's perception. For each message, we collected the score of the most highly valued source of information. Then we calculated the evaluator's mean score (EMS), which was the mean of the highest scores. The group members' perception of usefulness was the mean of the EMSs of the three evaluators. This was how we compared the mean of the experts' EMSs and the mean of the novices' EMSs.

We also wanted to compare the experts' and novices' perceptions of usefulness in terms of giving a good answer quickly. The EMSs of each group were calculated as explained above, but on this occasion on the scores for the system's speed.

4.3. Microevaluation

We also wanted to compare the experts' and novices' perceptions of the most useful sources of information. As was the case for the macroevaluation, we first of all calculated each group member's perception. We calculated the mean of the scores given by an evaluator for each source of information

© Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012
© FUOC, 2012

for all the messages. The result was the mean score of the usefulness (MSU) of each source according to each evaluator. By calculating the mean of the MSUs for the three evaluators of each group, we obtained the group's perception of the sources' contribution to the system's usefulness. Thus, we were able to compare the experts' and novices' perceptions. The evaluators' comments were also used for the microevaluation.

4.4. Results

The system's usefulness in terms of finding information was scored slightly higher by the experts (1.77 on a scale of 0 to 3) than by the novices (1.51). However, the experts gave lower scores for usefulness in terms of giving a fast answer (1.47) than the novices (1.75).

The group of experts considered that snippets from web pages made a better contribution to giving a good answer, a long way ahead of other sources of information. For the group of novices, snippets from web pages and Wikipedia articles scored higher, though the distance between them and other sources of information was much shorter.

According to the experts' comments, web pages were useful for giving an answer, though it meant that they had to spend a lot of time finding the most suitable context. In addition, they said that they found useful contexts after performing more than one test. In other words, they had to select different terms from the tag cloud. To some extent, this explains why the group of experts gave low scores when evaluating the system's usefulness in terms of giving a fast answer.

The contexts displayed by the learning materials search engine were the second most highly valued source of information by the group of experts, whereas the novices gave a higher score to Wikipedia. It would therefore seem that the ranking differences are due to the group of experts' greater experience of searching for information related to the subject by using search engines and to their ability to combine keywords to obtain useful results.

Articles were at the bottom of the ranking; these were scored lower than 1.5 by the experts and novices alike. According to a comment by an evaluator, this could be due to the fact that the articles found in Delicious and CiteULike deal with very specialist topics, and their target audience basically consists of lecturers and graduate students. Conversely, the topics covered in Wikipedia are better suited to the questions and reflections of undergraduate students.

Another issue we identified was the relationship between the types of student expectation and the sources of information. For example, course learning materials are useful for clarifying a concept, though students generally ask about information that is not included in the materials. Moreover, Wikipedia references and links to external online resources are useful for finding a solution to a problem or for suggesting additional reading. Wikipedia articles and web pages also provide extra information that complements learning materials and helps students to confirm that their reflections —and even their digressions— are on the right track. In addition, such sources of information are useful for updating a tutor's knowledge. If they cover topics included in subject learning materials, even scholarly articles can contribute to a tutor's lifelong learning.

Previous messages on subject discussion boards match messages posted by students asking for help. However, their usefulness depends on how recurrent the problem has been throughout

© Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012
© FUOC, 2012

the history of a subject. In addition, previous messages cannot be retrieved if, when writing them, students use under-specified references that only make sense at the time of writing.

5. Conclusions

In this article, we have presented a tutor assistant whose methodology distinguishes it from traditional approaches: the system is flexible enough to be able to deal with communicative objectives that go beyond answering students' questions. When messages contain digressions or reflections evoked by reading learning materials, the system promotes a 'learn from your students' learning process. Such a process becomes evident when a tutor finds that a student has given a good answer on a previous discussion board.

The results of the prototype evaluation are promising, bearing in mind not only the short period of time covered by the corpus evaluated, but also the fact that much depends on how recurrent a question, digression or reflection is. However, the methodology that we have developed thus far requires a lot of time to be spent on searching for information and is therefore slightly more beneficial for the group of expert tutors. For that reason, our aim is to improve the system's usefulness in terms of giving a fast answer on the one hand, and of giving a good answer without any distinguishable difference between experts and novices on the other.

We intend to improve its usefulness in terms of giving a fast answer by broadening the pragmatic approach and taking advantage of the relationship between students' expectations and sources of information. In addition, account will be taken of the relationship between a student's speech act expression and the snippet of text that best matches that student's expectation. Moreover, we intend to make the search for useful contexts easier, without any distinguishable difference between experts and novices, by expanding keywords. What we mean by this is that the terms selected by a user will activate semantically related terms, even if they are not visible in the tag cloud. Finally, we plan to integrate a search engine for scholarly articles with useful content for undergraduate students and tutors alike. Doctoral theses, and their state-of-the-art sections in particular, are interesting candidates.

Acknowledgments

This research was funded by the project KNOW2 (TIN 2009-14715-C04) of the Spanish Ministry of Science and Innovation and by the programme APLICA 2010 of the Open University of Catalonia.

References

ATSERIAS, Jordi; CASAS, Bernardino; COMELLES, Elisabet; GONZÁLEZ, Meritxell; PADRÓ, Lluís; PADRÓ, Muntsa (2006). "FreeLing 1.3: Syntactic and semantic services in an open-source NLP library". In:

© Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012
© FUOC, 2012

Proceedings of the fifth international conference on Language Resources and Evaluation (LREC 2006), ELRA. Genoa, Italy.

BERNHARD, Delphine; GUREVYCHG, Iryna (2008). "Answering Learner's Question Paraphrases from Social Q&A Sites". In: *Proceedings of the 3rd Workshop on Innovative Use of NLP for Building Educational Applications, ACL*. Columbus, Ohio. Pages 44-52.

FENG, Dpnghui; SHAW, Erin; KIM, Jihie; HOVY, Edward (2006). "An Intelligent Discussion-bot for Answering Student queries in threaded discussions". In: *Proceedings of Intelligent User Interfaces.*

HUNG, Jason C.; WANG, Ching-Sheng; YANG, Che-Yu; CHIU, Mao-Shuen; YEE, George (2005). "Applying Word Sense Disambiguation to Question-Answering System for e-Learning". In: *Proceedings of the IEEE 19th International Conference on Advanced Information Networking and Applications*. Taiwan. Pages 157-162.

SEARLE, John (1969). *Speech Acts. An Essay in the Philosophy of Language*. Cambridge.

WANG, Chun-Chia; HUNG, Jason C.; SHIH, Timothy K.; LIN, Hsiau-Wen (2006). "A Repository-based Question Answering System for Collaborative E-learning". *Journal of Computers*. Vol. 17, No 3, pages 55-58.

YANG, Che-You (2009). "A Semantic FAQ System for Online Community". *Journal of Software*. Vol. 4, No 2, pages 153-158.

About the Authors

Joaquim Moré
jmore@uoc.edu
Computational Linguist, Office of Learning Technologies, Open University of Catalonia (UOC)

Joaquim Moré holds a bachelor's degree in English and master's degree in Computational Linguistics awarded by the University of Barcelona. He currently works as a computational linguist in the Office of Learning Technologies at the UOC. In the field of education, he specialises in the retrieval of information from the Internet to facilitate teaching and learning tasks, and to ensure both lecturers' and students' lifelong learning. In other fields, he specialises in translation technologies. He is a member of the consolidated research group GRIAL (Inter-university Research Group for Linguistic Applications) and a researcher at the Internet Interdisciplinary Institute – UOC. He is doing his doctoral thesis on the evaluation of automatic translation systems.

© Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012
© FUOC, 2012

Salvador Climent
scliment@uoc.edu
Lecturer, Arts and Humanities Department, Open University of Catalonia (UOC)

Salvador Climent holds a doctorate in Romance Languages (Cognitive science and language), a master's degree in Computational Linguistics and a bachelor's degree in Spanish awarded by the University of Barcelona. He is currently a lecturer in Linguistics in the Arts and Humanities Department at the UOC, and director of the Catalan Language and Literature programme. He is a specialist in computational linguistics and cognitive linguistics, a member of the consolidated research group GRIAL (Inter-university Research Group for Linguistic Applications) and director of the Language Processing Group at the UOC.

Marta Coll-Florit
mcollfl@uoc.edu
Lecturer, Arts and Humanities Department, Open University of Catalonia (UOC)

Marta Coll-Florit holds a doctorate in the Information and Knowledge Society (Internet Interdisciplinary Institute – UOC), a master's degree in Cognitive Science and Language, and a bachelor's degree in General Linguistics (University of Barcelona). She is currently a lecturer in Linguistics in the Arts and Humanities Department at the UOC. She is a specialist in applied linguistics (computational linguistics, cognitive linguistics and psycholinguistics). She is a member of the consolidated research group GRIAL (Inter-university Research Group for Linguistic Applications) and a member of the board of SCOLA (Spanish Cognitive Linguistics Association).

Universitat Oberta de Catalunya (UOC)
Avinguda del Tibidabo, 45-47
08035 Barcelona
Spain

© *Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012*
© *FUOC, 2012*

SOME RIGHTS RESERVED

The texts published in this journal are – unless indicated otherwise – covered by the Creative Commons Spain Attribution 3.0 licence. You may copy, distribute, transmit and adapt the work, provided you attribute it (authorship, journal name, publisher) in the manner specified by the author(s) or licensor(s). The full text of the licence can be consulted here: http://creativecommons.org/licenses/by/3.0/es/deed.en.

UOC
Universitat Oberta de Catalunya

© Joaquim Moré, Salvador Climent and Marta Coll-Florit, 2012
© FUOC, 2012