Empowering the employment of Case Based Reasoning in VLE: The case were SYIM meets Moodle

Dr. Avgoustos. A. TSINAKOS

Department of Informatics
University of Macedonia
Thessaloniki-GREECE

INTRODUCTION

Asynchronous distance education was linked in the past with a number of pedagogical drawbacks. Delay responses, hazy monitoring of the students' performance progress, lack of student modeling processes etc. are commonly referred in the literature as critical problems (Wulf, 1996; Pernici and Casati, 1997; Relan and Gillani, 1997, Pritichard, 1998; Khan, 1997; Harasim, 1990; Hall, 1997b).

SYIM is a domain independent educational environment, which has been developed in order to remedy some of the educational problems, appears in asynchronous distance education (Tsinakos and Margaritis, 2001a). In its primal version, SYIM was focused on the process of providing an auto-reply to a student query posting. To accomplish this task SYIM used Case Based Reasoning techniques as part of its reasoning component (Diagnosis Process), in order to identify a similar posting (case), that might be the answer on student's original question.

According to the initial implementation schema (Tsinakos and Margaritis, 2001b; Tsinakos, 2003), when a student posts a query, the posting does not reach the tutor immediately. On the contrary two case scenarios were triggered: Scenario A of "No relevant case fount" where the Diagnosis Process fails to identify and retrieve a similar case and therefore student's posting has to be answered directly by the tutor. Scenario B of "One or more relevant cases fount" where the retrieved set of relevant cases is displayed to the student. Therefore the student can select among the listed cases, the most proper one that answers to the originally posted query.

In the current paper, the process of empowering and enriching the CBR Diagnosis Process is described. This evolvement derived by the integration of SYIM with Moodle VLE (Dougiamas and Taylor, 2003). The resulted VLE (version1.2 official release of Moodle), will be used as the national platform for Asynchronous Distance Education services of the Greek Educational Network.

IMPLEMENTATION PURPOSE

The purpose of the Questions Module (QM) is to fulfil a serious need of all educational processes: that of allowing the teacher to answer questions posed by his students. This simple educational device is always very important, since it benefits both students and teachers. The former, by obtaining a comprehensive answer to their questions can continue their learning progress, while the latter can use these questions as valuable information feedback regarding the difficulty of certain concepts, the efficiency of current teaching methods etc. All these issues are critical in asynchronous distance education, where the simple solution of "just asking the teacher" is not feasible.

There are important differences between traditional and distance education in regard to the process of asking and answering student questions. In the traditional education model, usually questions are asked and answered in front of all the students. This allows the teacher to answer any given question only once, and all the students to benefit from this one answer. In distance education, however, this is not the case. Since each student cannot interact with the whole class at the same time, questions take a "personal" character and have to be answered on a per-student basis. Of course, this is extremely inefficient for the teacher, given the fact that it is not possible to have a "canned" answer that will be understood by all the students in the same degree. Thus, the teacher not only needs to answer the same question many times, but also each time the conversation has to start from scratch and needs to be personalized. Maintaining a Knowledge Base (KB) where students' queries/answers are stored and the use of a CBR Diagnosis Process to

detect and derive a relevant "case" that maybe the answer to a student's original posting brings a viable solution to this problem of distance education.

The KB holds not only the questions and answers of the teacher, but also meta-information in the form of Keywords. This allows the system to compare the content of a newly asked question with that of existing questions (which have been answered), and thus possibly provide the student with an answer without the teacher's intervention.

An additional advantage to this solution is that the teacher can query the system regarding the questions asked by his students, specifying criteria such as student grade, topic of the question etc. This allows the teacher to obtain valuable information, for example which topics his students had the most questions on (indicating that the relevant educational material is not clear enough), which questions were asked by highly or lowly graded students (thus providing insight on the weaknesses of those students as a group).

Interface

The Questions Module has two separate interfaces: the student interface and the teacher interface. One of the two is automatically presented to the user, depending on his status in the particular course.

Student Interface

The student interface gives the student with three possible actions: a."Ask a new question" which further analysed below, b. "View unanswered questions" where all student' s questions that are still in discussion with the teacher are listed, and "View answered questions" used for reference purposes allowing the user to review the all the submitted questions.

"Ask a new question" action

The action of the "New question" is used to submit a new question to the system. In the starting screen, the user is asked to provide a summary (brief description) for his question, and also a fully expanded description of it. If the course uses topics, and there are topic names defined, then the user also has the option of selecting which topic his question falls into, via a drop down menu. Correctly setting the topic will help the user find an already existing answer, but it is not required. In the event that the Questions activity is included in such a named topic, that topic is pre-selected.
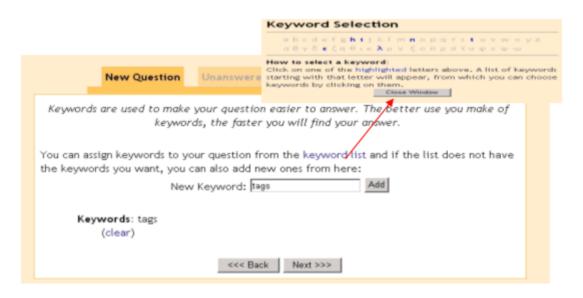

Figure 1: Ask a new question
When the user chooses Next, he is presented with the keyword

selection screen. This screen allows the user to select which Keywords best describe his question, which is very important in order to find possible answers. The can enter Keywords in two ways: either by manually typing them one at a time and selecting "Add", or by using the "keyword list"

window, which displays all existing Keywords broken down by starting letter and allows the user to automatically select keywords for his question by clicking on them.

The user can also leave the Keywords field blank and immediately click "Next". In this case, the system will do its best to guess which Keywords describe this particular question using an algorithm described in the Algorithms section.

Having the question submitted, and assuming there are no errors in the submission process, the system the Diagnosis Process in order to identify an existing question (relevant case) that may provide the answer that the student seeks. The algorithm used to find these matches is also described in the Algorithms section.

If one or more possible matches are found, the user is presented with a screen were the summaries of the 10 most probable matches are displayed (Figure 3). It is worth to notice that in the retrieved set the potential answers are ranked according the section. Thus, potential answers found in the same section of that of the original posting, are displayed first than the others found in different sections.

By clicking on question summary, the user is allowed to view the full content of that specific question in the form of a student-teacher conversation. The student while reviewing such a possible match can click on the "This answers my question" button, if the retrieve question is indeed the answer that the student is looking for. In such a case of a successful KB match, so called a "hit", the question is marked as "Answered", and the correlation between the question just asked and the one that provided the answer is noted in the database.

In case where no possible match satisfies the user, the latter can click on the "I cannot find an answer to my question" button, and the system proceeds as if there were no possible matches at all.

If no possible matches are found for the new question, the system presents an informative message that notifies the user that his question will have to be answered by a teacher. The user is then forwarded to the "View unanswered questions" screen.

Teacher interface

The teacher interface is divided into five separate options: View New Questions, View open questions, View satisfied questions, Administration, and Knowledge Base.

View New Questions The current action enables the teacher to browse at a glance all the newly posted questions that the system have failed to answer.

View open questions This screen is identical in appearance and function to the "View unanswered questions" screen displayed to students. The only difference is that since the teacher can see questions asked by all of the students, the name of the asking student is also shown next to each question thread.

View satisfied questions This is functionally a subset of the previous screen, where only the questions marked as Satisfied by students are shown. Thus, a teacher is provided easy and immediate access to those questions which have been answered and now must be Closed, in order to be no longer considered "Unanswered".

Administration This screen allows the teacher to change any and all options which globally affect the functions of the Questions Module. It displays the name of each possible setting, its current value, and a link that takes the user to a screen where he can change that setting.
The list of settings is: Rejected keywords;

- This is a list of words that will never be assumed valid
- Key words when the system automatically tries to assign
- Key words to a new question. (This happens when the user who places the question provides no Key words). This list contains common words such as "how, for, and, the" etc.

Knowledge Base This is perhaps the screen most important for the educational process. From here, the teacher can request to see a list of question threads, limited by any combination of the following criteria: Question topic, Asking user, Minimum and maximum grade of asking user, Presence or absence of the four question status flags: Public, Searchable, Satisfied, Closed. For a discussion of the four flags, (Figure 4).
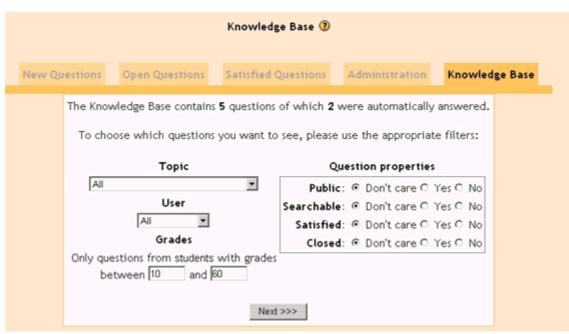
Figure 4: Search the Knowledge Base

The above figure depicts a teacher's search in the Knowledge Base,
Seeking for a list of all the questions posted by the students (User: All), who have grades between the range 10 and 60 (Grades :10-60), in every topic of the course (Topic:All). The questions properties (explained in the "Question status flags" section), as far as their flag is concerned, are set to the "Don't care" value.

The retrieve set of the above search according the conditions that have been set, appears in Figure 5.

Question status flags

The system uses four different on-off flags to categorize all questions it contains. These flags, and their impact, are:

Public Public is on by default, except on questions which have been answered by the Knowledge Base. This is because these questions do not have any educational value (in essence, they are just repetitions of existing questions). Questions with the Public flag set to off are not graded by the system as candidates for answering a new question.

Searchable At the moment, the Searchable flag has the same functionality as the Public flag. It exists for future expansion of the system's functionality.

Satisfied The status of this flag decides in which user interface screen it is displayed. Of course, its main use is denoting that a student finds the answer to his question satisfactory. The status of this flag also affects the actions that the user is allowed to take (see Chapter 4). This flag is set to off by default, except on questions answered by the Knowledge Base, where it is obviously set to on.

Closed This flag denotes that a question has been, one way or another, dealt with, and consequently will not be modified in the future. Only a teacher can manually set the Closed flag, and its status affects both the user interface screen where it is displayed and the actions that the user may take. It is set to off by default, except on questions answered by the Knowledge Base, where it is set to on.

RETRIEVAL ALGORITHMS

In the current version of Moodle the CBR Diagnosis Process is enriched with two important algorithms. The new developed Questions Module makes use of these algorithms in order to detect and derive the similar case-question that may be the answer to the original posted question of a student. These algorithms are below discussed in detail:

Automatically finding Keywords for a question
The current algorithm assigns Keywords to new question, when the student has failed to do so. Keyword assignment is achieved having the algorithm to perform the following rules in order:

- Any words in the summary that are already Keywords (for other questions) are selected as Keywords.
- Any words in the summary that also appear in the expanded text are selected as Keywords.
- Any words in the summary written entirely in capitals are selected as keywords.
- Words up to two letters long are automatically rejected.
- Words in the "Rejected Keywords" list (see section 4. B. III.) are automatically rejected.
- Rejection takes precedence over selection.

Finding possible matching answers for a new question
The retrieval algorithm is the most important feature of the Question module, since the better results it provides, the greater the efficiency of the module is. The algorithm attempts to find questions in the Knowledge Base that have similar content with a newly asked question, to provide the student with possible answers.

The algorithm functions as follows: Existing questions are graded with regard to their similarity to the new question. Grading is done by awarding points to each question, depending on what level of similarity is detected. When all awards have been made, the questions that were awarded the most points are presented in order of decreasing similarity. Questions not satisfying some minimum grading criteria internal to the algorithm are not shown at all.

Some questions may be excluded from the grading process, depending on their flag settings (i.e Public questions).
Awarding of points is done as follows:

There are four constants (point amounts) coded into the module. We will refer to these as Point Amount (P) A, B, C and D. The exact value of these constants defines the behaviour of the module, and thus they are not exposed to the user.
PA points are given to a question if it shares a Keyword with the newly asked one. Thus, if the number of shared Keywords is MA (Matches-A), the question being graded receives PA * MA points.

PB points are given to a question if it uses Keywords "related" to those of the newly asked one. Thus, if the number of related Keywords is MB, the question being graded receives PB * MB points.

"Relation" is better defined as follows: *Two Keywords are related if there is at least one question that is described by both of them, regardless of the presence or absence of any other keywords that might describe it.*

This is better understood with an example:
Let's assume that question A (QA) has the keywords *web, server*. Both of them describe QA, so they are related to each other. Let's also assume question B (QB) has the keywords *apache, server*. Finally, we suppose that a student student asks the new question QC, with the keywords *web, browser*. Of course, when searching for an answer to QC, the system will award points to QA (PA * 1) because of the common keyword web. However, QB will also be awarded points because it has the keyword server, and from QA we know that web and server are related. Thus QB will receive (PB * 1) points due to Keyword relationships.
PC points are given to a question if it is in the same course section as the one assigned to the new question. There is no multiplier.

Finally, PD points are awarded to a question if it is in the same course as the new question. Allowing for this setting gives the option of having questions in one course answered by questions asked in another course, if this is possible. However, this may lead to spurious matches being presented, so whether to use it or not is open to debate. Practically, the value of the PD constant compared to the other three will decide if this kind of behaviour is possible.
Having described all this, the number of points (P) awarded to each question (with the variables MA, MB known) is:

$$
\begin{aligned}
&\text{IF } (M_A + M_B = 0) \\
&\qquad \text{THEN } P = 0 \\
&\text{ELSE} \\
&\qquad P = P_A * M_A + P_B * M_B + P_C + P_D
\end{aligned}
$$

SUMMARY

The current paper describes the Question Module which is a new developed feature incorporated in Moodle VLE. As it has been already mentioned, the Question Module is an improvement version of the Process of Identification of Similar Question originally developed in the SYIM VLE
The retrieval Algorithms that have been integrated in the CBR Diagnosis Process aims to improve the accuracy of the findings, identified as similar cases/potential answers towards an original query.

The Diagnosis Process intend to operate even in cases where the user is unable to identify or to suggest a keyword that characterize the posted question, by guessing the potential keyword/s.

Moreover the derived set does not only appear the "exact mach" cases but also these cases that appear to have a short of general relation between the keywords attached to them.
Under the statement of such implementation, the process of auto-reply to students' queries tends to become more users friendly and accurate.

Contentment

REFERENCES

Dougiamas, M. and Taylor, P. (2003). Moodle: Using Learning Communities to Create an Open Source Course Management System, EDMEDIA 2003. Available online at http://dougiamas.com/writing/edmedia2003/

Hall, B. (1997b). Web-based training cookbook. New York: Wiley Computer Publishing.

Harasim, L. (1990). Online education: An environment for collaboration and interllecual amplification. In L. Jarasim (Ed.), Online education: Perspectives on a new environment. (pp. 39-64). New York: Praeger.

Khan, B.H. (Ed.). (1997). Web-based instruction. Englewood Cliffs, New Jersey: Educational Technology Publications, Inc.

Pernici, B. and Casati, F. (1997). The design of distance education applications based on the World Wide Web. In B. H. Khan (Ed.) Web-based instruction. (pp. 245-254). Englewood Cliffs, New Jersey: Educational Technology Publications.

Pritichard, C. L. (1998). The ASTD handbook of instructional technology. New York: MacGraw-Hill, Inc.

Relan, A. and Gillani,B. B. (1997). Web-based instruction and the traditional classroom: Similarities and differences. In B. H. Khan (Ed.) Web-based instruction. (pp. 41-46). Englewood Cliffs, New Jersey: Educational Technology Publications.

Tsinakos, A., and Margaritis, K., (2001a). See Yourself IMprove (SYIM). An educational environment for the provision of personalized distance education services and the formulation of student models", accepted in European Journal of Open and Distance Learning (EURODL), ISSN 1027-5207.

Tsinakos A. and Margaritis K., (2001b). Employing Case Based Reasoning in Asynchronous distance education, journal of Learning Technology, IEEE Computer Society Learning Technology Task Force (LTTF), ISSN 1438-0625, Vol 3, (3), July 2001.

Tsinakos A., "Asynchronous distance education: Teaching using Case Based Reasoning", Turkish Online Journal of Distance Education-TOJDE, ISSN 1302-6488 Volume:4 Number:3, July 2003.
Wulf, K. (1996). Training via Internet: Where are we? Training & Development, 50 (5), 50-55.

Biodata of author:

Dr. Avgoustos Tsinakos received the following degrees:

- **BSc in Physics (1992), Greece**
- **MSc in Applied Artificial Intelligence (1993), Aberdeen-Scotland, under the ESF scholarship.**
- **AGDDE (T) Advanced Graduate Diploma in Distance Education Technology (1999), Athabasca University of Canada,**
- **MDE Master in Distance Education (2000),Athabasca University of Canada,**
- **Ph.D. degree on "Artificial Intelligence and Internet (Asynchronous Distance Education)" (2002), University of Macedonia, Greece.**

His current occupation is, Superintendent in Secondary level Education on the filed of Computing Science. Dr. Tsinakos is a member of the coordination team of Greek Ministry of Education for a number of National Research Projects of Advanced Telematic Services (Synchronous & Asynchronous Distance Education), and a part of his scientific activities concerns the Evaluation of European Educational Projects on the field of E-learning, as being a member of EU Expert Evaluators Team (Socrates Program) Some of his main research interest concerns, Asynchronous Distance Education, Web based Distance learning, Virtual Learning Environments, Intelligent tutoring systems,Student modelling and Internet applications of AI.

Dr. Tsinakos can be reached on the following addresses:

Dr. Avgoustos Tsinakos, University of Macedonia,
Bcs, MSc, AGDDE (T), MDE, PhD, Depth of Informatics, Superintendent of Secondary Egnatia 156, 54006 level Education on ICT, Thessaloniki Greece.
E-mail: tsinakos@uom.gr
http://selene.uom.gr/~tsinakos
Tel:00302310 891890
Fax:00302310 891842

PRINT   RETURN