Expert Systems with Applications 39 (2012) 7471–7478

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A genetic fuzzy expert system for automatic question classification in a competitive learning environment

Elena Verdú, María J. Verdú *, Luisa M. Regueras, Juan P. de Castro, Ricardo García

School of Telecommunications Engineering, University of Valladolid, Paseo Belén, 15, 47011 Valladolid, Spain

ARTICLE INFO

Keywords:
Intelligent tutoring systems
Educational technology
Automatic question classification
Competitive learning
Genetic algorithms
Fuzzy systems

ABSTRACT

Intelligent tutoring systems are efficient tools to automatically adapt the learning process to the student's progress and needs. One of the possible adaptations is to apply an adaptive question sequencing system, which matches the difficulty of the questions to the student's knowledge level. In this context, it is important to correctly classify the questions to be presented to students according to their difficulty level. Many systems have been developed for estimating the difficulty of questions. However the variety in the application environments makes difficult to apply the existing solutions directly to other applications. Therefore, a specific solution has been designed in order to determine the difficulty level of open questions in an automatic and objective way. This solution can be applied to activities with special temporal and running features, as the contests developed through QUESTOURnament, which is a tool integrated into the e-learning platform Moodle. The proposed solution is a fuzzy expert system that uses a genetic algorithm in order to characterize each difficulty level. From the output of the algorithm, it defines the fuzzy rules that are used to classify the questions. Data registered from a competitive activity in a Telecommunications Engineering course have been used in order to validate the system against a group of experts. Results show that the system performs successfully. Therefore, it can be concluded that the system is able to do the questions classification labour in a competitive learning environment.

1. Introduction

During the last years, the learning process is changing substantially in order to be centred on the students and adapted to their needs and features. Different studies have shown the effectiveness of the new adaptive learning systems (Verdú, Regueras, Verdú, de Castro, & Pérez, 2008). Many of these systems attempt to be more adaptive by offering students questions with difficulty levels according to their skills and capabilities. The aim is to increase the efficiency and the level of interaction and motivation of students (Lilley, Barker, & Britton, 2004). Too difficult or too easy questions can frustrate and decrease students' motivation, while adaptive question sequencing provides a more efficient and effective learning (Wauters, Desmet, & Van den Noortgate, 2010). Moreover, according to (Lee & Heyworth, 2000), students should be able to score higher if the items or problems are arranged according to their difficulty level, since after solving easier problems, they feel more motivated to solve the harder ones.

On the other hand, the competitive learning systems, as the QUESTOURnament system, are an effective technique to capture students' interest, motivation and engagement by arousing their competitive instincts (Anderson, 2006; Philpot, Hall, Hubing & Flori, 2005). Moreover, competitive learning reduces procrastination, a common cause for students failing to complete assignments (Lawrence, 2004) and improves the learning process (Regueras et al., 2009).

QUESTOURnament is a telematic tool integrated into the e-learning platform Moodle that allows teachers to organize dynamic contests in any knowledge domain (Regueras et al., 2009). Students compete for getting the highest marks and being at the top in the ranking. They must solve exercises (known as *challenges* in QUESTOURnament) within a time limit and as soon as possible, since the scoring function varies with time.

The competitive nature of QUESTOURnament motivates students but also can provoke stress and discouragement in the worst classified students. To assign the adequate opponents and questions to a student may be an effective strategy to reduce these negative effects (Wu et al., 2007). Therefore the system should group students by knowledge level so that students with similar skills compete together and answer questions with a difficulty level suitable for them.

In this context, it is very important to correctly classify questions by difficulty level. However, it is difficult for teachers to accurately estimate the difficulty level according to the students' level of competence (Watering & Rijt, 2006). Experience helps teachers to better estimate the difficulty level of the questions, but even senior teachers sometimes fail and have to rectify when they

* Corresponding author. Address: ETSI Telecomunicación, Paseo Belén 15, 47011 Valladolid, Spain. Tel.: +34 983423707; fax: +34 983423667.

E-mail addresses: elever@tel.uva.es (E. Verdú), marver@tel.uva.es (M.J. Verdú), luireg@tel.uva.es (L.M. Regueras), jpdecastro@tel.uva.es (J.P. de Castro), ricgar@tel.uva.es (R. García).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2012.01.115

analyze the answers given by their students. An automatic estimation system could be the basis for an effective adaptation process.

A lot of systems that automatically estimate the difficulty level of items can be found in the literature (Burghof, 2001; Cheng, Shen, & Basu, 2008; Jong, Chan, Wu, & Lin, 2006; Lee, 1996; Wauters et al., 2010). However, the variety in the nature of the application environments makes difficult to apply the existing solutions directly to other applications. Therefore, a specific solution has been designed in order to turn the competitive e-learning system QUESTOURnament into an intelligent system. The objective is to make learning more effective and to mitigate some of the practical drawbacks of competitive learning.

This paper discusses the validity of an expert system that automatically estimates the difficulty level of the questions posed in the QUESTOURnament competitive learning system. Section 2 introduces the major issues about teachers' perception of difficulty and summarizes the search towards the solution. The expert system is described in Section 3. Section 4 starts with a description of the experiment developed in order to validate the system. Next, a study that analyzes the accuracy of the estimations of difficulty obtained by the intelligent system is presented. Finally, the main conclusions are stated.

2. Background

2.1. Teachers' perception of difficulty

The correct estimation of the difficulty level of learning material (questions, items...) is very important in the design and definition of assessment processes, adaptive learning systems or standard setting methods. However, there are not too many studies about the perception and estimation of difficulty level by teachers.

Estimating the difficulty level of questions is not an easy job. Several studies (Alexandrou-Leonidou & Philippou, 2005; Hadjidemetriou & Williams 2002; Lee & Heyworth, 2000; Watering & Rijt, 2006) question the ability of teachers to make accurate difficulty level estimations of learning material since teachers usually fail to identify the correct difficulty level according to the students' ability. In general terms, students' performance tends to be overestimated by teachers (Goodwin, 1999; Impara & Plake, 1998; Verhoeven, Verwijnen, Muijtjens, Scherpbier, & Van der Vleuten, 2002). Moreover, according to Watering and Rijt (2006), if the accuracy of teachers' perception of difficulty is analysed by categories, teachers tend to overestimate the difficulty of easy items and underestimate the difficulty of hard items. Impara and Plake (1998) also suggest that estimating item difficulty accurately is quite difficult; however, they do not think that teachers systematically underestimate the difficulty of hard items and overestimate the difficulty of easy items. In this respect, other contradictory results are found too. For example, Mattar (2000) states that teachers are less successful at rating very difficult or very easy items, while Zhou (2009) indicates that teachers classify better the hardest items.

In short, although there are not conclusive studies about the tendency of teachers when they classify questions by difficulty level, all researchers agree on the difficulty of doing this classification. Therefore, an automatic system that adjusts the difficulty level of questions according to the students' behaviour would be a very useful support tool and a key component for a truly adaptive learning environment.

2.2. In search of an intelligent solution for a competitive tool

There are many domain-dependent intelligent tutoring systems (ITSs) that provide students an adequate learning path through the different topics of a subject, according to the previously learnt topics. These systems are based on techniques such as Bayesian Networks (Hibou & Labat, 2004; Nouh, Karthikeyani, & Nadarajan, 2006; Vomlel, 2004) and require the previous definition of knowledge domains by using, for example, domain-specific ontologies (Colace & De Santo, 2006). Modelling these networks of knowledge components and their dependencies, generalizing them for every student, is not an easy task (Noguez, Sucar, & Ramos, 2005), especially for domain-independent systems like QUESTOURnament, which can be used for diverse subjects and levels of education.

Many domain-independent ITSs focus on presenting questions and problems adapted to the students' knowledge level. They often apply the Item Response Theory (IRT) to estimate both the characteristics of the questions, such as difficulty or guessing probability, and the knowledge level of students (Chen, Lee, & Chen, 2005; Lilley et al., 2004), independently of the knowledge domain. However, the correct application of traditional theories for tests implies some assumptions, which are not met by many examination contexts, especially when telematic tools are used for distance learning. Moreover, some of the characteristics of more specific tools, such as the competitive nature of QUESTOURnament, make the application of these theories difficult for the environment under study.

The typically used IRT models are one-dimensional, that is, they assume that the response to a question depends on a single trait, usually the knowledge level. Besides, it is also supposed that the response a student gives to a specific question does not depend on the responses given to other questions (Embretson & Reise, 2000). Therefore, using IRT entails carefully designing the tests so that these both conditions are fulfilled. Moreover, conventional IRT models only the response accuracy, ignoring response time; since it was thought to be used in pure power tests (Roskam, 1997), which assume that students have unlimited time to solve a question. Even if limited time could be assumed, at least the requirement should be that time is not a factor that affects the students' response. However, in a competitive environment as QUESTOURnament, time is very important, since only the first student who answers a challenge correctly will be able to obtain the highest score for that challenge. Therefore, there are different factors that could distort the results obtained by the IRT methods when applied to the QUESTOURnament system.

Students can apply different strategies during competition and even different personality factors can determine the students' final response to an item. Several challenges can be posed at the same time and students have to select one of them to be solved first. Many students tend to read all the different questions and select the one that seems the easiest to be solved first. Difficult challenges are usually read several times and solved after the easiest ones have been answered. On the other hand, two students with exactly the same knowledge level could respond to a same question differently, as one can be more persistent and devote more time to solve the question while another one can be more anxious with the competition and quickly respond to be the first one. Consequently, time and number of readings are important factors that should be taken into account in the model, but its modelling is dependent on the actual students' behaviour.

Moreover, when teachers pose challenges to QUESTOURnament, they do not have any restriction related to time, type of questions or skills needed to solve them. They are free to use any configuration of the system in any context. Then, there are some important factors that can vary:

- Maximum time available to submit an answer to a challenge.
- Type of questions (open questions, multiple choice questions, true/false questions, short response questions, problems, etc).
- Context surrounding students when they solve the questions: a contest may be developed in classroom or on distance during one or several days.