

Adaptive evaluation in an E-Learning System Architecture

J. López-Cuadrado*, A. Armendariz and T. A. Pérez

Department of Computer Languages and Systems, University of the Basque Country (UPV-EHU), Apto. 649, 20080 San Sebastian, Spain

This paper presents the result of the integration of an automatic evaluation module in ELSA, a complete, user-friendly and easy-to-use e-learning environment. Although this new module is already operative, it is in the marketing phase to be included both in BOGA, the commercial version of ELSA to be used on Internet, and in Hezinet, the version nowadays employed as a complementary tool in classrooms of Basque learning colleges. The system used to manage two types of tests: *predefined* tests and *dynamic* tests. Now it also offers *adaptive* tests. To achieve this objective, we started by building GenTAI, a stand-alone system for tests administration. The most remarkable advantage of this new feature comes when evaluating the skill levels of new students since this new kind of tests uses less items than other approaches. GenTAI, and thus the new module, allows the automatic and dynamic generation of Computerized Adaptive Testing (CAT) using a previously calibrated item bank according to the 3-parameter logistic Item Response Theory (IRT) model. Besides, this system has been a very significant starting point, as we have used it as a throw-away prototype, and also to define the definitive requirements in the module to be constructed and improve dialogues and the visual aspect of the interface.

Keywords e-learning environment, computerized adaptive testing, item response theory, adaptive evaluation

1. Introduction

Hezinet [1] is a pioneering commercial system for distance language learning which is nowadays being used in more than 50 Spanish and 14 South American adult Basque-schools [2]. It is a mixture of an Adaptive Hypermedia System (AHS) and an Intelligent Tutoring System (ITS), so it takes advantage of the positive aspects of each archetype and rectifies their inconveniences [3]. Concretely, hypermedia features offer flexibility when managing the navigation to didactical contents, while the tutoring system controls users and performs evaluations of the navigation to adapt the system's behaviour. This symbiosis avoids the ITS to control the instruction too strictly, so a more autonomous learning process is obtained [4]. Hezinet also benefits from the inclusion of different audio-visual media which are used by the AHS to present the contents to the student.

Contents are the smallest pedagogic units that are taught. Hezinet contains a collection of activities linked to the contents. Actually, activities are the most important pedagogical components of Hezinet, since they attract students' attention and provoke them to work with their contents. Hezinet includes up to 20 different types of activities, such as highlight mistakes in a text, multiple-choice exercises, true/false exercises, free answer exercises, ordering exercises, and compositions or essays. Figure 1 shows an example of activity that consists of several sentences in Basque whose adequate translation the student has to find. Many multimedia resources have been used to develop the set of 10231 activities that Hezinet contains, including 364 videos (4.99 Gb in MPEG format), 703 audio files (1.5 Gb in MP3-format), and 72 interactive films (500 MB in DCR format).

The system distinguishes three main user profiles: administrator, teacher and student. The administrator is the user responsible for inscribing students and teachers. Teachers have certain functions available for each of their students: they can, for instance, add or remove contents from their students' profiles as they monitor the students' learning processes. But the main user of Hezinet is the student, since the last

* Corresponding author: e-mail: javilo@si.ehu.es, Phone: +34 946 01 44 68

objective of the system is to allow her/him to acquire a certain level of knowledge of the Basque language.



Figure 1. Translation exercise.

Nowadays Hezinet works in three ways. The first version of the system appeared in 1999 as an e-learning tool that worked on demand to complement instructional lessons [5]. The second use of the system encourages the self-learning, and it was first implemented in 2000 within *Zibergelas* (Cyberclassrooms in Basque), which are situated in cultural centres, usually supported by public organisations, such as Basque houses, town halls and city libraries. Finally, the third version of the system, which was created in 2002 with the commercial name of BOGA [6], also supports the self-learning, but it is not necessarily linked to a physical place, since it works through the Internet [7].

Hezinet works with both *predefined tests* and *dynamic tests*. The former are designed in advance and they do not change as a result of interactions with the students. As these tests are previously designed, content developers can focus on the interaction and graphical design. On the other hand, dynamic tests are automatically compiled by the system attending only to those concepts that a particular student has demonstrated not to understand.

The results of Hezinet's users in the Basque Government official exams have proved the system to be very helpful in language learning [7], but at the moment the figure of the human teacher is still essential. The technology does not permit to develop a language learning system without a human teacher to advise students and to conduct the oral sessions. However, this is not necessarily applicable to some concrete but complex tasks, such as taking decisions about the starting level of new students.

Our research is dedicated to build a new generation of Hezinet by adding to it some improvements to the system. Next section will present ELSA, a new architecture for Hezinet which is allowing us to add new components to offer new features to the already functioning system. Section 3 will discuss about GenTAI, a new module that generates computerized adaptive admission tests. Finally, section 4 will be dedicated to conclusions.

2. ELSA

ELSA [8] provides the architecture for the e-learning system. It includes the same components as Hezinet, with the difference that now they are defined in a way that the system can be easily scalable by adding modules to it. We have focused on the evaluation module, one of the main strengths of Hezinet, although the architecture includes many other components. For example, ELSA already has a component to build Multiple Choice Items and include them in the item bank [9] and we are working in completing the system to extend it to some other types of items [10] such as multiple response, ordering association

(1:1), union (1:N), fill in blanks, essays, hotspots, object positioning and slices, and migrating to work with their representation following standards such as SCORM and IMS QTI [8]. We have worked on a component to provide also Computerized Adaptive Tests [11]. At the moment, we have developed a prototype of the system called GenTAI (which we will discuss in new section) and, as we will see, it will be soon available as another component for ELSA. Furthermore, we are also working on other components to calibrate items and detect which ones have become obsolete or are not valid anymore

We have identified 4 assessment-related roles to work with the system: *administrators* are responsible for the correct functioning of the whole system by tuning some parameters such as the number of items for dynamic tests. They will also include all other users and assign them their role in the system; *authors* are in charge of introducing both items and tests. They will also have to calibrate items and detect anomalies in the item bank; *tutors* are responsible for assuring that the learning process performs well for each student. They can also be in charge of detecting some obsolete items; finally, *students* use the system, for instance, to complete activities or make assessment tests.

During the implementation of ELSA we have focused on the new components rather than reimplementing those that already work. We think that all users have to share the entry page to the system. Once they identify themselves, the functionality available will depend on the roles been assigned to him or her. Figure 2 shows the entry page, which not only asks for existing users' identification, but also allows new students to register in the system. In this case, the administrator will be responsible for deciding which roles are the most convenient for them.



Figure 2. ELSA's entry page.

3. GenTAI

GenTAI [12] is a tool that allows the automatic generation of experimental Computer Adaptive Tests from an item bank calibrated by means of the 3-parameter logistic Item Response Theory model [13]. These tests are created "on the fly", that is, when the student gives a response, the system calculates which one is the item that better adapts to her or him. In basic terms, when the student answer is correct, he or she will receive a more difficult item, and if not, then the item will be easier. This process will be repeated until some condition is reached, a stop condition that is usually related to some reliability value or the number of administered items. Since the system does know the results of previous interaction, it can react to the responses given, in contrast to a previously created test, which has to deal with all didactic areas to cover the number of different situations that could occur in a test with several students.

This system is friendly, easy-to-use, intuitive, portable, scalable, highly configurable and strongly documented, and it fulfils the goals of an adaptive testing at the same time. GenTAI offers the adminis-

trator the feature to configure some characteristics of the CAT, such as the kind of CAT to generate or the way the results will be shown on the screen. Figure 3 presents a configuration form for a test.

GenTAI first looks the most adequate item in the item bank (sometimes with certain restrictions) and then shows it. Then, the user chooses a response and the system, after storing it, (re)calculates the ability of the student. Finally, it checks the stop criteria; if a stop condition has been reached, the test will be over. If not, the system will look for another item (the most adequate for the new computed ability), and so on.

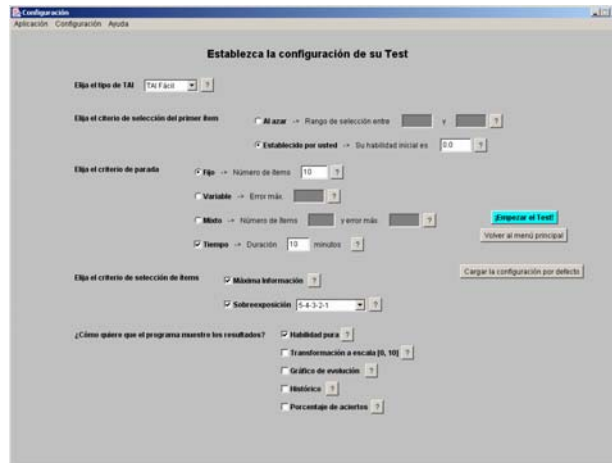


Figure 3. Computerized Adaptive Test configuration in GenTAI.

As we stated above, GenTAI is only a prototype. In order to create the final component for ELSA, we have corrected some interaction mistakes and added some other features to make it more versatile. In Figure 4 shows one of the screens to configure the behaviour of CATs.

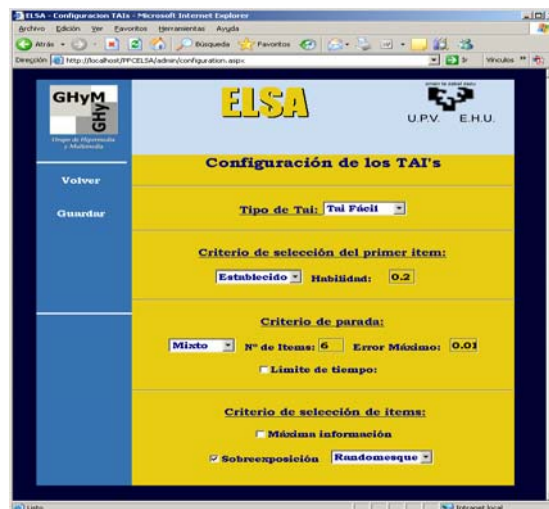


Figure 4. Computerized Adaptive Test configuration in the final module.

4. Conclusions

Hezinet and its commercial version BOGA are very powerful learning systems. However, users detect new needs that should be covered. ELSA is our means of providing the system with new modules that improve its performance. In this paper we have focused on assessments issues. The most important is the inclusion of a CAT generator called GenTAI that soon will be one of ELSA's modules.

CATs will improve the detection of the starting level for new students. A novice user, that is, the one that has not already used the system, does not know the level in which he or she should begin the interaction with the e-learning system. At this moment the student performs some predefined tests starting in the lowest level and increasing the difficulty level gradually until she or he reaches a level that is not able to complete satisfactorily. This way of determining the starting level is very common in language learning, and it is used, for instance, in some Official Languages School in Spain [14]. In order to reduce the number of items to be answered by the students, they are usually asked to estimate which is their level. However, tests administrators know that very often students tend to undervalue themselves and point out lower levels than theirs. This fact redounds in an increase of the items administered, so the whole process can last up to more than 2 hours in some cases. Computerized adaptive testing can reduce the number of items of the test to about 30-40, without losing accuracy in the results.

References

- [1] Pérez, T. A. Un hiperentorno adaptativo para el aprendizaje instructivo/ constructivo (Tesis Doctoral). San Sebastián (España), Lenguajes y Sistemas Informáticos (Universidad del País Vasco UPV-EHU): 286. (2000)
- [2] López-Cuadrado, J., Pérez, T. A., Vadillo, J. Á. y Arruabarrena, R. Integrating adaptive testing in an educational system. First International Conference on Educational Technology in Cultural Context, Joensuu (Finland), University of Joensuu (2002)
- [3] Gutiérrez, J., Pérez, T. A., Usandizaga, I. y Lopistéguy, P. Hypertutor: adapting hypermedia systems to the user. Fifth International Conference on User Modeling, UM-96, Kai-lua-Kona, Hawaii (USA). (1996).
- [4] Gutiérrez, J., Pérez, T. A., López-Cuadrado, J. y Arruabarrena, R. Evaluación en sistemas hipermedia adaptativos. Simposium sobre Test Adaptativos en Sistemas de Enseñanza y Diagnóstico Informatizados (VII Congreso de Metodología de las Ciencias Sociales y de la Salud), Madrid (España), (2001) .
- [5] González, A., Pérez, T. A., López, R., Carro, J. A. y Gutiérrez, J. Hezinet: Una alternativa telemática al tradicional aprendizaje de idiomas. VIII Conferencia de la Asociación Española para la Inteligencia Artificial / III Jornadas de Transferencia Tecnológica de Inteligencia Artificial (CAEPIA'99 / TTIA'99), Murcia (España) (1999).
- [6] Armendariz, A. J., López-Cuadrado, J., Vadillo, J. Á. y Pérez, T. A. HEUSKLEARNING: Un hiperentorno educativo para el aprendizaje del euskara. IV Encuentro Europa-América Latina sobre Formación y Cooperación Tecnológica y Profesional, Isla de Margarita (Venezuela). (2004).
- [7] Sanz-Lumbier, S., Gutiérrez, J., Pérez, T. A., Sanz-Santamaría, S., Vadillo, J. Á. y Villamañe, M. Hezinet: The hypermedia system that makes the Basque language easy to learn. 5th International Conference: Computers and Advanced Technology in Education (IASTED), Cancún (México), ACTA Press. (2002).
- [8] Armendariz, A. J., J. López-Cuadrado, A. Tapias, M. Villamañe, S. Sanz-Lumbier, and S. Sanz-Santamaría. Learning environments should follow standards: ELSA does. in World Conference on E-Learning in Corporate, Government, Healthcare, & Higher Education (E-Learn 2003). Phoenix, Arizona (USA): AACE. (2003)
- [9] López-Cuadrado, J. The CAT is out of the bank. Advances in technology-based education: towards a knowledge-based society. A. Méndez Vilas, J. A. Mesa González y J. Mesa González. Badajoz (España), Junta de Extremadura (CECT). **3**: 1832-1836. (2003).
- [10] Sanz-Santamaría, S. Standards, Adaptation & Pedagogy: Quality Assessment in e-Learning. in Proceedings on 14th International Conference on Computers in Education. (ICCE 2006). Beijing (China). (2006)
- [11] Wainer, H. (2000). Computerized adaptive testing: a primer (second edition). Mahwah, New Jersey (USA), Lawrence Erlbaum Associates.
- [12] López-Cuadrado, J. y Sánchez Fernández, J. M. "GenTAI: generador de tests adaptativos informatizados." IE Comunicaciones - Revista Iberoamericana de Informática Educativa **2**: 9-24.(2005).
- [13] van der Linden, W. J. y Hambleton, R. K. Handbook of modern item response theory. New York (USA), Springer-Verlag. (1997).
- [14] Escuela Oficial de Idiomas de Irun. www.eoirun.com (2006).