Adaptive Testing by Test++

Maria Barra, Giuseppina Palmieri, Simona Napolitano*, Vittorio Scarano, and Luca Zitarosa

Dipartimento di Informatica ed Applicazioni "R.M. Capocelli"
Università di Salerno
Baronissi (Salerno), 84081, Italy
{marbar,palmieri,vitsca}@unisa.it

Abstract. We present the adaptive features of Test++, an adaptive system for training and teaching on the Internet. The system integrates an adaptive training environment for personalized training and a cooperative environment for exams both accessible via Internet and a standard Java-enabled browser.

1 Introduction

The World Wide Web [1] explosion has strongly influenced everyday life of almost everybody in the industrialized world. Particularly it changed many things for people that, in a way or in another, are involved in the educational field. In fact, WWW now is able to efficiently deliver true *e-learning* to user by a variety of tools to support [7,8], integrate [6] and, sometimes, substitute the traditional methodologies used in schools of any kind [5,10].

Adaptive hypermedia [2,3] is one of the most effective disciplines to help tailor e-learning to each student's needs. Techniques for adaptive hypermedia, that were developed for local systems before the advent of WWW, have proven themselves extremely useful to provide an interactive and personal environment. Several systems are developed in Computer Adaptive Tests (CATs) and Item Response Theory fields (see [4] for comprehensive survey and [9] for SIETTE, an Intelligent Evaluation System). Their goal, in general, is to accurately estimate students knowledge by user-tailored systems.

We describe, here, the adaptive features of a system to support and integrate traditional teaching: Test++. The system goals are to provide supplementary personal self-training sessions and synchronous exercitations/exams to the traditional classroom and laboratory activities. The goal of the project (whose first part we describe here) is to interleave in a single session personalized testing and training. Questions are selected according to the performances of the student but they also lead to supplementary reading material.

In this paper, we focus on the adaptive training policy that is proposed to the student. We describe the strategy that is followed by the system to propose the questions, adapting the number of questions in each topic to the results

* Current affiliation: Finmatica s.p.a.

P. Brusilovsky, O. Stock, C. Strapparava (Eds.): AH 2000, LNCS 1892, pp. 264–267, 2000.

obtained by the student so that his/her preparation is accurately tested. In a certain way, since testing system is coupled with presentation of support and reading material related to questions/answers, the system will offer two-level adaptivity: a first level, described in this paper, offers a mechanism to adapt questionnaires to students so that they are challenged and (implicitely) shown their weaknesses by reiterating and insisting on some topics. A second level, not described here, will offer, on the basis of the interactions with questions, access to adapted and personalized course material.

2 Adaptivity in Test++

The system is designed to work in two different, complementary modes. The first one is the *self-training* mode that is treated throughout the rest of the paper. The system can only work in a second mode, named *virtual classroom*: students are connected through clients (Java applets or applications) and are synchronously given a whole (static) questionnaire to answer. Questionnaires appear differently to each student and the teacher is able to monitor the behaviour of students by real-time statistics (by single student and collectively).

Test++ architecture is client-server: a "fat" client (a Java applet) implements the adaptive strategy and asks server a new question to present. In this way, first, the system is more efficient than usual server-side systems (one round-trip is avoided for checking correctness of answers); moreover, it is also possible to implement the synchronous mode within the same system and, finally, the XML questions archive is safely stored on the server and not directly accessible (as a whole) by the students.

In *self-training* mode, Test++ offers a personalized questionnaire to the student. The questionnaire is built randomly from a sample of questions contained in an XML archive but according to a set of configuration rules that specify how many questions for each topic are placed into the questionnaire. The objective of the whole system is to interleave questions and material (that will also be adapted in the next phase of the project) so that the student is interactively presented material according to his/her needs.

The system offers to the teacher a wide choice of behaviours ranging from absolutely static (i.e. non adaptive at all) to absolutely "mean" (ask all the questions on the topic that is less known by the student).

We expected that the strategy could be helpful for students: their needs should be addressed and shown implicitely. If they do not study homogenously the topics then it is not possible to get high scores. Moreover, sometimes, it can be difficult for students to recognize what topics are the ones that "the teacher really care about": we expect that, in this case also, the questionnaire can be helpful since adaptive behaviour can be accurately tuned by the teacher. Some experiments (not reported here) show positive feedbacks by students.

Now we describe the adaptive mechanism for the self-training session. The goal of the mechanism is to force the student to a thorough study of all the topics

of the course. In fact, the strategy is, in a way, designed to find weaknesses of student's preparation for the exam.

We identify the knowledge domain of the adaptive strategy as consisting of T disjoint *topics*. The mechanism is aimed at ensuring that good scores are associated with a complete preparation of the student. In a way, the system wants to avoid that a student can score 15/20 (15 right answers over 20 questions) when the five wrong answers all belong to the same topic.

The questionnaire is built interactively according to configuration parameters: each time the student is presented a single question; when the answer is provided by the student the system selects the next topic and asks the server to randomly pick a test in that topic, ensuring, at the same time, that questions are not repeated in the same session.

It is well recognized that one of the problems in adaptive systems is authoring and configuring systems. To provide the teacher with a versatile mechanism, we model the self-training session with a number of *configuration parameters*:

- the total number N of questions (tests) in the session;
- for each topic i of the knowledge domain:
  - D_i the number of tests in the topic that should be presented in the session;
  - $\textsc{Min}_i$ the minimum number of right answers that the student must provide in order to pass to a different topic;
  - $\textsc{Max}_i$ the maximum number of tests that can be shown to the student about the topic (regardless of right answers).
- breadth-first round width w (more details follow).

The policy of the mechanism can be seen as a game played by the student against an adversary (Test++) that tries to understand the topics where the student has studied/understood the least. The system starts with a (randomly ordered) breadth-first round of the topics by asking w questions for each topic (w is the *width* parameter). Then, once identified "weak" topics (i.e. topics where student has shown some deficiency) the system keeps asking questions in "weak" topic i until either the student reaches the minimum number of right answers ($\textsc{Min}_i$) or the total number of questions in topic i exceeds the maximum allowed $\textsc{Max}_i$. Once all the "weak" topics are passed, then questions are picked in the topics whose D_i was not reached. It should be noticed that the role of D_i is simply to ensure that successful questionnaires (i.e. when the student is almost always correct according to the configuration) are well-balanced, anyway. The tie-breaking rule (for the choices above) was to pick topics in the definition order.

It can be seen that, assuming given N, T and w, the parameters are T triplets of integers $(D_i, \textsc{Min}_i, \textsc{Max}_i)$. The parameters allow a wide variety of behaviours (including a totally static, *a-priori* fixed questionnaire).

For example, assume that there are $T = 4$ topics and $N = 10$ questions. Then a static, non adaptive, random questionnaire Q_1 is built by specifying parameters $(3, 3, 3)$, $(3, 3, 3)$, $(2, 2, 2)$, $(2, 2, 2)$. On the other hand, an adaptive questionnaire Q_2 with configuration parameters $(3, 2, 4), (3, 2, 3), (2, 1, 3), (2, 1, 2)$ shows

a behaviour that gives a certain relevance (in this order) to each topic, by differentiating between the first and the second topic, and the third and the fourth topic by the maximum number of questions. Finally, a stronger bias toward the first topic is shown by a questionnaire Q_3 with the following configuration $(3,2,6),(3,2,3),(2,1,3),(2,1,2)$ that allows up to the 60% of the total number of questions to be about the first topic if the student does not correctly answer at least 2 questions. Notice that, given the algorithm followed by the system to challenge the student, correct configurations are such that $N = \sum_{i=1}^{T} D_i$.

It is helpful to represent configurations as surfaces in the 3D space by interpolating the points represented by the triplets. In this way static questionnaires (i.e. no adaptive behaviour) are represented by a line since triplets of the form $D_i = \text{MIN}_i = \text{MAX}_i$ adequately describe the static questionnaire[1] while adaptive questionnaires are represented by "large" surfaces that correspond to "stretching" the original (static) line into surfaces.

References

1. T. Berners-Lee, R. Cailliau, and, J.F.Groff. "The World Wide Web". Computer Networks and ISDN Systems, Nov. 1992, vol.25 (no.4-5), 45-49.
2. P. Brusilovsky, J. Eklund and E. Schwarz. "Web-based education for all: A tool for developing adaptive courseware". Computer Networks and ISDN Systems. (Proc. of Seventh International World Wide Web Conference, 14-18 April 1998) 30 (1-7), 291-300.
3. P. Brusilovsky. "Adaptive Educational Systems on the World-Wide-Web: A Review of Available Technologies". In Proc. of Workshop "WWW-Based Tutoring" at 4th International Conference on Intelligent Tutoring Systems (ITS'98), San Antonio, TX, August 16-19, 1998.
4. P. Brusilovsky and P. Miller. "Web-based testing for distance education. In: P. De Bra and J. Leggett (eds.) Proceedings of WebNet'99, World Conference of the WWW and Internet, Honolulu, HI, Oct. 24-30, 1999, AACE, pp. 149-154.
5. Cisco Academy. http://www.ciscoacademy.com
6. Paul De Bra. "Hypermedia Structures and systems" (Course 2L670) Eindhoven University of technology (http://wwwis.win.tue.nl/2L670).
7. D. Dwyer, K. Barbieri, H.M. Doerr. "*Creating a Virtual Classroom for Interactive Education on the Web*". Proc. of WWW 95, Third Int. Conf. on World Wide Web.
8. B. Ibrahim, S.D. Franklin. "*Advanced Educational Uses of the World Wide Web*". Proc. of WWW95, 3rd International Conference on World Wide Web.
9. A.Rios, J.L. Pérez de la Cruz, R.Conejo. "SIETTE: Intelligent evaluation sytem using tests for TeleEducation". Workshop "WWW-Based Tutoring" at 4th Int. Conf. on Intelligent Tutorial Systems (ITS'98).
10. W. A. Bogley, J. Dorbolo, R.O.Robson, J. A. Sechrest. "New Pedagogies and Tools for Web Based Calculus". Proc. of AACE WebNet 96, San Francisco, CA (USA), Oct. 15-19, 1996. See also http://www.peak.org for usage examples of QuestWriter.

[1] Of course, the really relevant parameters are $D_i = \text{MAX}_i$ since the MIN_i can be any number less than or equal to D_i.