Usability Evaluation of a Web-based Authoring Tool for Building Intelligent Tutoring Systems

Maria Moundridou

University of Piraeus - Department of
Technology Education and Digital
Systems
80 Karaoli & Dimitriou str., Piraeus
18534 Greece
mariam@unipi.gr

Maria Virvou

University of Piraeus - Department of
Informatics
80 Karaoli & Dimitriou str., Piraeus
18534 Greece
mvirvou@unipi.gr

Abstract

Authoring tools for Intelligent Tutoring Systems (ITSs) provide environments that can be used by authors-instructors who are not necessarily computer experts to easily develop cost-effective ITSs. However, the quality of the ITSs to be generated depends heavily on these authors-instructors who will provide the content for the courses. This means that authoring tools should be carefully designed, developed and evaluated in order to ensure that they are usable, friendly and effective. This paper reports on a study we conducted in order to evaluate the usability of WEAR, which is a Web-based ITS authoring tool. The results of the study were very encouraging. Authors were quite satisfied with the functionality of the system, they found it very friendly and easy to use and most importantly they stated that it could be really useful.

1   Introduction

The main goal of Intelligent Tutoring Systems (ITSs) is to reproduce the behaviour of a human tutor who can adapt his/her teaching to the needs and knowledge of the individual learner. As a number of successful evaluations of ITSs have shown, these systems can be educationally effective compared to traditional instruction either by reducing the amount of time it takes students to reach a particular level of achievement or by improving the achievement levels given the same time on task (Du Boulay, 2000). A common criticism for ITSs concerns the complex and time-consuming task of their construction: more than 200 hours are required to develop an ITS to cover just one hour of instruction (Woolf & Cunningham, 1987).

Authoring tools for ITSs are meant to simplify the ITS construction by providing environments that can be used by a wide range of people to easily develop cost-effective ITSs. However, it would be extremely optimistic to think that authoring tools are simply the perfect solution for the creation of ITSs. In particular, the fact that authoring tools depend heavily on the instructors for the quality of the ITSs to be generated, links these systems with various problems. For example, instructors may face several difficulties during the design process and they may become frustrated or they may provide inconsistent information to the tool that may lead to the generation of ITSs with problematic behaviour. This means that authoring tools should be carefully designed, developed and evaluated in order to ensure that they are usable, friendly and effective. The usability of authoring systems will promote the production of ITSs and improve their quality

because it will increase the instructors' degree of acceptance towards them and decrease the possibility of errors that may lead to the generation of non-effective ITSs.

This paper reports on a study we conducted in order to evaluate the usability of a Web-based ITS authoring tool that is called WEAR.

2   The system

WEAR is a Web-based authoring tool for ITSs focusing on problem construction and problem solving in Algebra-related domains (Virvou & Moundridou, 2000). However, WEAR also offers authors the facility of building adaptive textbooks in every domain even if it is not Algebra-related (Moundridou & Virvou, 2001). These textbooks offer navigation support to students, adapted to their individual needs and knowledge. In this section we will briefly describe only the way that the authoring of the adaptive textbooks is performed and not the rest of the system's operation, since the usability evaluation study that will be reported concerns this part of WEAR's functionality.

Although most of the existing authoring tools for adaptive educational textbooks approach the adaptivity issue in quite similar ways, they differ a lot in the authoring process they impose to their users (authors). For example, in Interbook (Brusilovsky, Eklund & Schwarz, 1998) the author should provide a specially structured, annotated MS-Word file. In AHA (De Bra & Calvi, 1998) the author should write annotated HTML files. MetaLinks (Murray, Shen, Piemonte, Condit & Thibedeau, 2000) on the other hand, provides a GUI interface for authoring all aspects of the electronic textbook.

In WEAR, we address authoring in a way that in its first steps resembles the simple one adopted by commercial tools like WebCT (Goldberg, Salari & Swoboda, 1996). It should be noted that this similarity between WEAR and WebCT concerns only the authoring procedure and not the resulting courses, which in the case of WebCT are not adaptive. In particular, the authoring procedure is the following: The author should prepare HTML files for the topics that would be contained in the electronic textbook. The next step is to use WEAR's facilities for uploading these files to the WEAR server. For each uploaded file the author must also specify a title, a difficulty level and the position that it should have in the topics hierarchy. Finally, the author must edit the is_prerequisite_of relationships between topics. To perform this, the author is presented with the hierarchy of topics and s/he should write next to each topic the section numbers of its prerequisite topics.

The author may also create multiple choice tests or problems and associate them with the appropriate topics. The procedure is again quite simple: the author should fill in forms stating the question, the possible answers and the right answer. Then s/he should specify the difficulty level of the exercise as well as the topic that this exercise refers to.

3   The study

Five instructors-authors were asked to work with the system to build part of an adaptive textbook for a Software Engineering course. Before working with the system the authoring and learning environments of WEAR were introduced to the participants. Authors were provided with 7 HTML files. Each of these files contained a section from the Software Engineering book chapter "Requirements specification".

The authors were requested to use WEAR's facilities to:
- upload these files to the server,
- create the topics' hierarchy,
- edit the prerequisite relationships between topics,
- insert students to their virtual class, and
- construct some multiple-choice questions.

This book chapter as well as the multiple-choice questions were also given to the participants in a hard copy so that they could easily explore the available teaching material before working with the system. The time spent by the authors to perform the requested tasks was recorded. Finally, each author was interviewed in order to gather information concerning the usability of the system.

4   The results

Concerning the time that was needed by the authors to build their course, the result was very impressive: all the authors finished their tasks in 15-20 minutes. This is considered as a very good result since the teaching material covered corresponds to more than 2 hours of classroom instruction. Furthermore, all of the authors stated that they found extremely useful to use a tool like WEAR to produce adaptive Web-based textbooks.

The answers that the five authors gave when they were interviewed were measured in the scale from 1 to 5. The mean ratings are summarised in Table 1 and indicate a high degree of usability for the system.

In particular, concerning the participants' general impression of the system, this was recorded as very good. This was also the case for the satisfaction that users felt when working with the system. The participants also stated that the system was rather motivating and easy to use.

As for the system's characteristics the authors were very pleased with the fact that the system was able to recover from errors without losing the already entered data. They were also very satisfied with the system's response time, as well as with its reliability. Users stated that the system does not inform them adequately about their errors (mean rating: 2.8) but the validity of this finding should be investigated since most of the participants did not make many errors.

Authors were quite satisfied with the amount of explanation that the system provided as well as with the clarity of the system's instruction messages. The system was also considered consistent as for the terminology it uses and the feedback messages it provides. Users also stated that the results of their actions were rather predictable.

Concerning how easy will be for a user to learn the system, the authors stated that this will be very easy and it will take minimum time. Finally, the number of actions that should be carried out in order to accomplish a task was considered by the participants as adequate and their sequence as logical.

From the participants' general comments on the system we came up with a useful finding: most of them even if they liked the system a lot, they would prefer to perform some tasks in a more visual way. For example, when authors arrange the topics in the topics' hierarchy they would prefer to do this by dragging and dropping rather than by altering the topic's position number in a text box.

Table 1: Mean ratings in questions measuring the system's usability

Questions	Mean rating (n=5)
What is your general impression of the system? (1: I disliked it… 5: I was very impressed)	4.2
How satisfied you felt when working with the system? (1: it was annoying… 5: I was very satisfied)	4.0
Was the system motivating? (1: it was boring… 5: it was very motivating)	4.0
Was the system easy to use? (1: very difficult… 5: very easy)	4.6
Was the system able to recover from errors without losing the already entered data? (1: unable… 5: very able)	5.0
How fast was the system in its responses? (1: very slow… 5: adequately fast)	4.2
Was the system reliable? (1: unreliable… 5: very reliable)	4.0
Were the system's error messages informative? (1: never… 5: very often)	2.8
Was the amount of explanation that the system provided adequate? (1: inadequate… 5: adequate)	3.8
Were the system's instruction messages clear enough? (1: they were confusing… 5: they were very clear)	4.2
How consistent did you find the terminology that the system was using? (1: inconsistent… 5: consistent)	5.0
How consistent did you find the system's feedback messages? (1: inconsistent… 5: consistent)	4.6
Could you predict the results of your actions? (1: never… 5: always)	4.4
How easy can it be to learn using the system? (1: very difficult… 5: very easy)	4.8
How long does it take to learn using the system? (1: a long time… 5: minimum time)	4.2
How did you find the number of actions that were needed to accomplish a task? (1: large… 5: adequate)	4.0
Was the sequence of actions to accomplish a task logical? (1: never… 5: always)	4.2

5   Conclusions/Discussion

The study that was conducted to evaluate the usability of the authoring environment of WEAR provided very encouraging results. Authors were quite satisfied with the functionality of the system, they found it very friendly and easy to use and most importantly they stated that it could be really useful.

However, it should be taken into account that the subjects participating in the study were computer scientists. On the one hand, this strengthens the positive results because it means that these subjects would probably have very high expectations as compared to instructors of other domains.

On the other hand, it was easier for these authors to work with the software and so the opinion of the novice computer user is not reflected in the results of the study.

It is within our future plans to run more experiments in order to confirm the positive results of the discussed study. In these subsequent evaluations, we plan to involve instructors of domains that are not computer-related. In this way, the possible problems that the non-competent computer user may face when working with WEAR will be revealed.

Finally, based on the results of this evaluation study and those to follow, we plan to redesign part of WEAR's interface in the direction that the users' opinions point. However, the results of the current study were very much in favour of the system and the only implication on redesigning the authoring environment of WEAR concerned the participants' preference to perform some of the tasks in a more visual way. Until the forthcoming evaluations, this issue is what we will be working on.

References

Brusilovsky, P., Eklund, J., & Schwarz, E. (1998). Web-based education for all: A tool for developing adaptive courseware. *Computer Networks and ISDN Systems*, 30, 291-300.

De Bra, P., & Calvi, L. (1998). AHA: a Generic Adaptive Hypermedia System. In P. Brusilovsky, & P. De Bra (Eds.), *Proceedings of 2nd Adaptive Hypertext and Hypermedia Workshop at the 9th ACM International Hypertext Conference - Hypertext'98*, Computing Science Report No. 98-12, (pp. 5-11). Eindhoven: Eindhoven University of Technology

Du Boulay, B. (2000). Can We Learn from ITSs? In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), *Intelligent Tutoring Systems, Proceedings of the 5th International Conference on Intelligent Tutoring Systems*, Lecture Notes in Computer Science No. 1839, (pp. 9-17). Berlin: Springer

Goldberg, M. W., Salari, S., & Swoboda, P. (1996). World Wide Web - Course Tool: An environment for building www-based courses. *Computer Networks and ISDN Systems*, 28, 1219-1231.

Moundridou, M., & Virvou, M. (2001). Authoring and delivering adaptive Web-based textbooks using WEAR. In T. Okamoto, R. Hartley, Kinshuk, & J. P. Klus (Eds.), *IEEE International Conference on Advanced Learning Technologies; Issues, Achievements, and Challenges - ICALT 2001*, (pp. 185-188). Los Alamitos, California: IEEE Computer Society

Murray, T., Shen, T., Piemonte, J., Condit, C., & Thibedeau, J. (2000). Adaptivity in the MetaLinks Hyper-Book Authoring Framework. In C. Peylo (Ed.), *Proceedings of the International Workshop on Adaptive and Intelligent Web-based Educational Systems* (held in Conjunction with ITS 2000), Technical Report, (pp. 61-72). Osnabrück: Institute for Semantic Information Processing

Virvou, M., & Moundridou, M. (2000). A Web-Based Authoring Tool for Algebra-Related Intelligent Tutoring Systems. *Educational Technology & Society*, 3 (2), 61-70.

Woolf, B. P., & Cunningham, P. A. (1987). Multiple knowledge sources in intelligent teaching systems. *IEEE Expert*, 2, 41-54.