An Approach to Usability Evaluation of an Intelligent Tutoring System

ANDRINA GRANIĆ and VLADO GLAVINIĆ*

Faculty of Natural Sciences, Mathematics and Education
University of Split
Nikole Tesle 12, 21000 Split
CROATIA

* Faculty of Electrical Engineering and Computing
University of Zagreb
Unska 3, 10000 Zagreb
CROATIA

Abstract: - Intelligent tutoring systems are a generation of computerized educational systems, which are intended to improve the process of learning and teaching by attempting to mimic the capabilities of human teachers. The development of application technology in the field of computerized educational systems has not been accompanied by an adequate usable, user-centered interface design, possibly resulting in scarce, unusable support for interaction and indeed in the decrease of interface's effectiveness and efficiency. In order to improve the usability of an intelligent tutoring system, nowadays perhaps the most important goal of current research in the field of human-computer interaction in general, scenario-based usability evaluation is advocated as a formal approach for intelligent tutoring system evaluation.

Key-Words: - intelligent tutoring, intelligent tutoring systems, scenario-based usability evaluation

1 Introduction

Development of information and communication technology, along with contemporary efforts in computer-supported learning and teaching, has introduced computerized educational systems with a certain level of intelligence – *intelligent tutoring systems*, ITSs. Since users can interpret these systems as user interfaces to some particular domain knowledge, the degree of their usability and efficiency should inevitably depend on the systems' usable design brought to them.

Usability is mainly concerned with making an interactive system easy to learn and easy to use, what in the particular area of intelligent tutoring means providing harmony among four components of any work situation: user, task, environment, as well as the ITS itself. However, it can be noted that starting already from the early days of development of intelligent tutoring systems inadequate consideration has been given to the communication among student, teacher and knowledge, and it still remains the Achilles heel of most systems.

The most promising approach to the realization of usable ITSs is the iteration of design and usability evaluation until a satisfactory usable solution is achieved. Consequently, an approach to ITSs' usability evaluation elaborated in this paper is based on criteria expressed in terms of objective performance measures in systems use, as well as in terms of users' subjective assessment. *Scenario-based usability evaluation*, as an operational definition of usability of an ITS, enables not only setting quantitative goals of execution before the evaluation is performed, but also enables the specification of operationally defined criteria for success. The results obtained through *scenario-based usability evaluation* are subsequently used for determining the user interface strength and weakness, hence furnishing a direction in ITS user interface redesign improvement.

2 Intelligent Tutoring Systems and Authoring Shells

Educational system is traditionally defined as a community group which includes teachers, as well as students, and their joint work in the process of learning and teaching. In such system the teacher controls the learning process. Such approach is nowadays exposed to important changes, which are also a result of the employment and support of information and communication technology, which enforced the role of computer in education altogether.

Intelligent tutoring systems, ITSs, are computerized educational systems intended to support the process of learning and teaching in an arbitrary domain knowledge by taking into consideration both teachers' and students' individuality. Depending on students' learning capabilities, intelligent tutoring systems can adjust the contents and the way of domain knowledge perception. ITSs are built on a fairly well established architecture, which relies on four interconnected modules, see Figure 1: (i) *expert module* acting as the domain knowledge unit, (ii) *student module* comprehending the generated student model based on the learning and teaching process in the domain knowledge, (iii) *teacher module* guiding the learning and teaching process and (iv) *communication module* realizing the interaction among student, teacher and knowledge [1], [2]. In most intelligent tutoring systems knowledge representation is based on semantic networks because of their ability to express the cognitive model of human memory and reasoning [3]. The basic components of semantic networks are *nodes*, used for presentation of domain knowledge objects, and *links*, which show relations between objects.

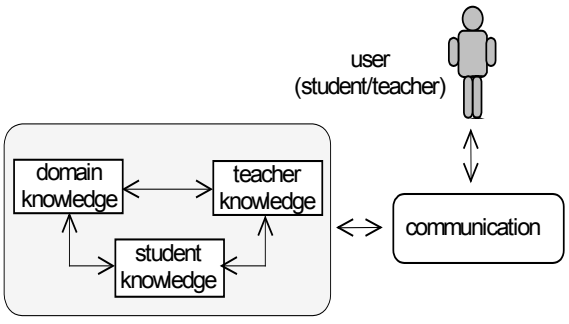

Figure 1: General ITS Architecture

As the need to cover a variety of different domain knowledge have arisen since, instead of having a number of specialized intelligent tutoring systems for the domains of interest, ITS generators were developed usually denoted as *authoring shells* [4]. These shells can be programmed for a particular domain by modifying the domain knowledge thus resulting in a specific intelligent tutoring system. In order to study the possibilities such a tool could provide in instruction/education, an intelligent hypermedia authoring shell named *Tutor-Expert System*, TEx-Sys, has been developed [5]. TEx-Sys is intended to adjust to teachers as well as to students within a hypermedia interactive learning environment by supporting teachers in development of a series of ITSs for arbitrary domain knowledge and conversely, by enabling students to learn, test themselves, consult the system for the obtained score, as well as receive from it advice for further work.

As highly interactive systems, intelligent tutoring systems and authoring shells rely upon a quality user interface, which should exhibit a number of properties like efficiency, effectiveness as well as usability. Communication between users and ITSs is inherently complex, especially when supporting student interaction because of the student's dealing with concepts - the domain knowledge - yet not understood very well [6]. A major reason for poor usability of present day ITS user interfaces is the lack of understanding of the process by which a usable interaction between users and machines is developed. Thus, the field is still open for research intended to improve the respective user interface design.

3 The Concept of Usability

Methodologies for building usable computer systems have been introduced and refined over the past fifteen or so years under the discipline of *human-computer interaction*, HCI. These methodologies are intended to provide usable and functional computer system [7] in order to achieve its almost transparency and enable end users to fully concentrate on the work. Hence, HCI principles include an early and consistent focus on end users and their tasks, empirical measurements of system usage, as well as iterative development.

As one of the key system features [8], *usability* is mainly concerned with making an interactive system easy to learn and easy to use. In particular area of intelligent tutoring systems and authoring shells this means providing

harmony among four components of any work situation: user, task, environment, as well as the ITS or the shell itself [9], [10]. A major reason for poor usability of present day ITS user interfaces, and in fact generally of most interfaces, is the lack of understanding of the process by which a usable interaction between users and machines is developed, because "… *usability is not a quality that can be spread out to cover a poor design like a thick layer of peanut butter* " [8, p. 16].

Studies show that redesigning user interfaces on the basis of user testing (namely, interaction measurement between users and computer systems) and iterating can substantially improve usability [11], because usability can only be meaningfully measured during task performance [12]. Therefore, the most promising approach to the realization of usable systems is to iterate design and usability evaluation until a satisfactory solution is achieved [11], [13], [14].

4 An Approach to Intelligent Tutoring System's Usability Evaluation

It is a well-known fact that users evaluate the usability of any interactive computer system in terms of quality of its user interface. Several different approaches at the operational level to the assessment and measurement of interaction between users and systems are known from the literature, e.g. [8], [13], while every one of them considers usability in terms of a number of operational criteria. In order to allow usability quantification those parameters "… *formalize the user behavior to be supported [and provide] usability objectives to the level at which it is to be supported* " [15, p. 15].

4.1 An Operational Definition of Usability

Usability as quality of use in context [16] should be viewed as comprising of two basic aspects - *ease of use* considered primarily as involving subjective judgements and *efficacy in use* considered primarily as involving measures of (human) performance [13]. Consequently, usability evaluation of an ITS user interface is based on criteria expressed in terms of:

- objective performance measures of *effectiveness* - the accuracy and completeness with which user achieve specified goals and *efficiency* - the resources expended in relation to the accuracy and completeness of goals achieved - in systems use, as well as
- of users' *subjective assessment* of her/his usage of the system.

An operational definition of usability of an ITS enables not only setting quantitative goals of execution before the usability evaluation is performed, but also enables the specification of operationally defined criteria for success, because it comprehends following seven measurable attributes: suitability, adaptivity, learnability, memorability, error rate, subjective satisfaction and overall subjective satisfaction, as shown in Table 1. The problem with the previous approach to usability specification, e.g. [17], is in its lack of comprehensiveness, as it does not consider explicitly the variety inherent to ITSs. Namely, additional requirements arise in the case of interfaces that are to tailor a system's interactive behavior according to specific requirements of an interaction, hence resulting in the need for adaptive interaction. Insofar as this consideration is contemplated, a supplementary usability requirement is inevitable – an adaptivity attribute as already stated before.

4.2 Scenario-based Usability Evaluation

Considering different methods of usability evaluation and the fact that we are not experienced usability specialists, having in mind that usability can only be meaningfully measured during task performance [12] and that is better to perform any kind of usability measurement than no testing at all, e.g. [18], we select an approach which comprehends formal user testing during users' walkthrough along ITS interface, guided with a set of predefined steps. Test users are tested with actual tasks under conditions that are as close to those in the actual intelligent tutoring system usage.

Scenario-based usability evaluation of an ITS interface comprehends the following steps:

- *walkthrough usability test*, composed of two parts: (i) scenario-guided task which has to be followed step by step in order to show ITS's basic functionality as well as main aspects of its interface and (ii) an arbitrary specific task which test users have to perform alone, thus enabling measures for effectiveness as well as for efficiency of following attributes: suitability, adaptivity, learnability and error rate,

Table 1: Usability Attributes of an Intelligent Tutoring System

	EFFECTIVENESS MEASUREMENT	EFFICIENCY MEASUREMENT	SATISFACTION MEASUREMENT
SUITABILITY	PERCENTAGE OF GOALS ACHIEVED DURING THE WALKTHROUGH TEST	TIME TO COMPLETE THE WALKTHROUGH TEST	RATING SCALE IN THE QUESTIONNAIRE
ADAPTIVITY			RATING SCALE IN THE QUESTIONNAIRE
LEARNABILITY	PERCENTAGE OF GOALS ACHIEVED WHEN COMPLETING A SPECIFIC TASK FROM THE WALKTHROUGH TEST	TIME TO COMPLETE A SPECIFIC TASK FROM THE WALKTHROUGH TEST	RATING SCALE IN THE QUESTIONNAIRE
ERROR RATE	RATIO BETWEEN BAD ACTIONS (ERRORS) AND GOOD ACTIONS IN A TASK FROM THE WALKTHROUGH TEST	NUMBER OF BAD ACTIONS (ERRORS) IN ATASK FROM THE WALKTHROUGH TEST	RATING SCALE IN THE QUESTIONNAIRE
MEMORABILITY	PERCENTAGE OF FUNCTIONS MEMORIZED SUCCESSFULLY DURING THE MEMO TEST	TIME TO COMPLETE THE MEMO TEST	RATING SCALE IN THE QUESTIONNAIRE
SUBJECTIVE SATISFACTION			RATING SCALE IN THE QUESTIONNAIRE
OVERALL SUBJECTIVE SATISFACTION MEASUREMENT			RATING SCALE IN THE QUESTIONNAIRE

- *memo test* which is performed after the walkthrough usability test and enables measures for effectiveness and efficiency of the interface memorability attribute by requiring the user to explain the effects of a single command or to write down command name for particular operation and
- *usability questionnaire*, which is due to measurement consistency filled out after the memo test and enables measures for user subjective satisfaction with the system's interface, as well as her/his satisfaction with its suitability, capability of adaptation, ease of learning, memorability and error rate during performance of usability test; responses are ranked on a 1 to 7 point semantic differential scale; from the standpoint of the single user the responses represent her/his subjective opinion, but as an average value taken from a number of users they indicate an objective value of intelligent tutoring system pleasantness.

4.3 Usability Attribute Specification

Since user testing/measurement, like all empirical studies, requires a theoretical framework for definitions and measures, we specify measurable usability attributes along with different ways of their measuring as shown in our previous work [9], [17]. Engineering for usability requires early specification of target levels for usability attributes [12], [13], [14]. Thus in order to state:

(i) the way a usability of an ITS is determined and

(ii) the way it serves as a measure of whether the current ITS's interface design meets the specification and in what respect it does,

all usability attributes are specified through seven items according to a formal method for specifying operationally defined criteria for success (*ibid.*). Due to space limitations only relevant tabularly defined operational criteria for the ease of ITS using and learning expressed through the usability attribute learnability is given in Table 2. Measuring concept and measuring method of the learnability's effectiveness, efficiency and satisfaction measurement, as well as of other usability attributes, express the way of ITS's usability measurement. Worst case, planned level, best case and now level serve as a measure of weather the current ITS design meets usability specification. Namely, values from now level are compared to those from worst case and if they are lower, the current version of ITS interface can not be accepted in terms of relevant usability attribute.

Table 2: Learnability attribute specification

LEARNABILITY	MEASURING CONCEPT	MEASURING METHOD	WORST CASE	PLANNED LEVEL	BEST CASE	NOW LEVEL
EFFECTIVENESS MEASUREMENT	EFFECTIVENESS IN THE EASE OF LEARNING AND USAGE	PERCENTAGE OF GOALS ACHIEVED WHEN COMPLETING A TASK FROM THE WALKTHROUGH TEST	70 %	85 %	100 %	94 %
EFFICIENCY MEASUREMENT	EFFICIENCY IN THE EASE OF LEARNING AND USAGE	TIME TO COMPLETE A SPECIFIC TASK FROM THE WALKTHROUGH TEST	30 min	20 min	10 min	14.6 min
SATISFACTION MEASUREMENT	USER'S SATISFACTION WITH THE EASE OF LEARNING	RATING SCALE IN THE QUESTIONNAIRE	5	6	7	6.1

In order to specify values for the worst case and indirectly for the planned level and best case, pilot testing was performed. Results of the *scenario-based usability evaluation* of an intelligent tutoring system show the need for the interface redesign according to the aspect of interface ability to adapt to user-machine interaction, due to unacceptable worst case values of the adaptivity attribute. The iteration of design and usability evaluation until a satisfactory usable solution is achieved is a promising approach to the realization of usable ITS interface. Consequently, identification of user's individual differences and also her/his changing knowledge and behaviour over time during the ITS's usage, as well as incorporation of adaptivity in its interface, will enable ITS's interface design to meet the specification, resulting in the improvement of intelligent tutoring system's usability.

As already stated before in order to realize a learning and teaching environment suitable for students' individual requirements, an intelligent hypermedia authoring shell *Tutor-Expert System* has been developed, providing the means for developing specialized intelligent tutoring systems for particular domains of education. Due to the fact that the iteration of design and usability evaluation is the most promising approach to the realization of usable interaction between users (teachers and students) and system, we are at the present focused on evaluation of just one part of *Tutor-Expert System*, module *Developing*. Moreover, instead of "module", more appropriate term for it is "shell", because of its main purpose of the development of an arbitrary domain knowledge base. Assessment of usability attributes performed in *scenario-based usability evaluation* show user interface strengths and weaknesses, hence furnishing a direction in interface improvement.

5 Conclusion

Improving the usability of intelligent tutoring systems, considered as emulators of human teachers in the process of learning and teaching, is perhaps the most important goal of current research in the area of intelligent tutoring. Since users can interpret these systems as user interfaces to some particular domain knowledge, the degree of their effectiveness and efficiency should inevitably depend on systems' usable design brought to them.

The paper elaborates on usability evaluation of an intelligent tutoring system user interface, which has already been in use for some time at university and high school level education. *Scenario-based usability evaluation*, as an operational definition of usability of an ITS, is targeting the improvement of the interface in order to provide usable communication, allowing users to better concentrate on the job to be done and not on the intricacies of interaction.

Acknowledgements:

This work has been carried out within project 036033 *Architectural Elements for Regional Information Infrastructures* jointly funded by the Ministry for Science and Technology of the Republic of Croatia and the Istrian County.

References:

1. Burns, H., Capps, C.: "Foundations of Intelligent Tutoring Systems: An Introduction", in M. Polson, J. Richardson (Eds.): *Foundations of Intelligent Tutoring Systems*, Lawrence Erlbaum Associates Publishers, Hillsdale, NJ, 1988, pp. 1-18.
2. Woolf, B.: "AI in Education", in I. Shapiro, C. Stuart (Eds.): *Artificial Intelligence – Encyclopedias*, John Wiley & Sons, Inc, New York, 1992, pp. 434-444.
3. Nute, D.: "Knowledge Representation", in S. Shapiro (Ed.): *Encyclopedia of Artificial Intelligence*, John Wiley & Sons, Inc, 1992, pp. 743-869.
4. Barton, M.: "Authoring Shells for Intelligent Tutoring Systems", *7th World Conference on Artificial Intelligence in Education*, Washington DC, USA, August 16-19, 1995, `http://www.pitt.edu/~al/aied/barton.html`
5. Stankov, S.: *Isomorphic Model of the System as the Basis of Teaching Control Principles in the Intelligent Tutoring System*, Ph.D., Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, 1997. (in Croatian).
6. Miller, J.: "The Role of Human-Computer Interaction in Intelligent Tutoring Systems", in M. C. Polson, J. J. Richardson (Eds.): *Foundations of Intelligent Tutoring Systems*, Lawrence Erlbaum Associates Publishers, Hillsdale, NJ, 1988, pp. 143-189.
7. Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T.: *Human-Computer Interaction,* Addison-Wesley, Wokingham, England, 1994.
8. Nielsen, J.: *Usability Engineering*, Academic Press, Boston, MA, 1993.
9. Granić, A., Glavinić, V.: "Adaptive Intelligent Tutoring Systems in the Context of Usability Requirements", 5th IEEE International Conference on *Intelligent Engineering Systems INES 2001*, Helsinki, Finland, September 16-18, 2001, pp. 231-234.
10. Stankov, S., Granić, A., Glavinić, V.: "Usability Considerations for an Intelligent Tutoring System", *Proceedings Electrical Engineering and Computer Science Conference ERK'99*, Portorož, Slovenia, September 23-25, 1999, pp. 151-154.
11. Nielsen, J.: "Iterative User-Interface Design", IEEE Computer, November 1993, pp. 32-41.
12. Bennett, J.: "Managing to meet usability requirements: establishing and meeting software development goals", in J. Bennett, D. Case, J. Sandelin, and M. Smith (Eds.): *Visual Display Terminals*, Prentice Hall, Englewood Cliffs, NJ, 1984, pp. 161-184.
13. Shackel, B.: "Usability – context, framework, design and evaluation" in B. Shackel and S. Richardson (Eds.): *Human Factors for Informatics Usability,* Cambridge University Press, Cambridge, 1991, pp. 21-38.
14. Whiteside, J., Bennett, J., Holtzblatt, K.: "Usability Engineering: Our Experience and Evolution" in M. Helander (Ed.): *Handbook of Human-Computer Interaction*, Elsevier North-Holland, Amsterdam, 1988, pp. 791-817.
15. Carroll, J., Rosson, M.: "Usability Specifications as a Tool in Iterative Development", in R. Harston (Ed.): *Advances in Human-Computer Interaction 1*, Ablex, Northwood, NJ, 1985, pp. 1-28.
16. Bevan, N., Macleod, M.: "Usability measurement in context", *Behavior & Information Technology*, 13 (1 and 2), 1994, pp. 132-145.
17. Granić, A., Glavinić, V.: "Interface Redesign Issues for Intelligent Tutoring Systems", 9th International Conference on *Human-Computer Interaction HCI International 2001*, Poster Sessions: Abridged Proceedings, New Orleans, Louisiana USA, August 5-10, 2001, pp. 133-135.
18. Nielsen, J.: "The usability engineering lifecycle", IEEE Computer, March 1992, pp. 12-22.