Designing an Adaptive Learning Module to Teach Software Testing

Rahul Agarwal, Stephen H. Edwards, and Manuel A. Pérez-Quiñones
Virginia Tech, Dept. of Computer Science
660 McBryde Hall (0106)
Blacksburg, VA 24061 USA
+1-540-231-3054

rahulaga@vt.edu, {edwards, perez}@cs.vt.edu

ABSTRACT
Adaptive learning systems aim to precisely tailor education and training to the individual needs of learners. Such systems use an internal model of a user's current knowledge to adjust the navigational affordances and presentation order of material. The user model is incrementally built and updated as the user demonstrates mastery by completing exercises and tests. Designing courses that are delivered adaptively involves addressing many complexities. This paper describes experiences designing the first adaptive module in a series intended to teach software testing skills. Experiences in using the first module and a preliminary evaluation of its effectiveness are presented.

Categories and Subject Descriptors
K.3.2 [**Computers and Education**]: Computer and Information Science Education; K.3.1 [**Computers and Education**]: Computer Uses in Education—computer-assisted Instruction; D.2.5 [**Software Engineering**]: Testing and Debugging.

General Terms
Verification

Keywords
Adaptive learning, self-paced, on-line, web-based courses, course design, adaptive evaluation, NetCoach.

1. INTRODUCTION
There is a need to teach software testing skills to computer science undergraduates **Error! Reference source not found.**. Software testing is critical for building successful software, but unfortunately it is viewed as boring and uncreative and this important aspect of software development has been traditionally neglected. To address this problem at Virginia Tech, we are investigating the strategy of adding a series of self-paced, on-line learning modules for teaching software testing to several courses spread across our curriculum. Rather than writing static content, however, we are investigating the use of *adaptive hypertext* techniques to provide modules that tailor their presentation to each individual learner's experience and prior knowledge.

SIGCSE'06, March 1–5, 2006, Houston, Texas, USA.

The adaptive modules are proposed as a supplement to current classroom teaching and learning activities. The adaptive nature of the modules helps create a self-paced learning environment for the student and encourages independent study. It also helps advanced students complete the module faster, while providing weaker students more opportunities to learn and perform better.

This paper describes our experiences with using adaptive techniques to create self-paced learning materials, and the preliminary results we have seen with our first module.

2. ADAPTIVE LEARNING SYSTEMS
Adaptive learning systems customize teaching and learning to suit each individual user. This tailoring is based upon a user model maintained by the adaptive system. The user model is incrementally built and updated as users complete various exercises and tests. The unique goals and key features of adaptive learning tools [2][19] are:

Adaptive presentation: Adapting the content based on the learner's goals, knowledge, experience and other information stored in the user model.

Curriculum sequencing (instructional planning technology): Providing an individualized "optimal path" thorough the material.

Adaptive navigation support: Displaying hyperlinks adaptively to better help navigation.

Intelligent analysis of student solutions: Providing comprehensive feedback about exercises that a student has completed and updating the user model accordingly.

Interactive problem solving support: Provide hints and intelligent help to students depending upon their performance.

Example-based problem solving: Provide intelligent help in the form of examples from previous materials or questions already covered.

The individualized content presentation makes adaptive systems flexible for the user while presenting unprecedented challenges for course designers. The advantages of using such tools are:

- Adaptive presentation and curriculum sequencing help tailor the pace of the course appropriately for novice and advanced students alike, presenting content they would enjoy.

- Adaptive navigation helps novices find their way in hyperspace and prevents them from getting lost.
- Courses adapt to a student's background and knowledge, enabling faster completion for advanced students.
- These systems are often web-based and have an interface similar to internet web sites, minimizing the effort required to interact with the materials.

The disadvantages of using such tools are:

- The course materials that instructors need to develop require more work, since the material will be adaptively presented to learners and only a subset may be viewed by each student.
- The tools are complex to develop and deploy and the technology is new and currently under research and development.
- Course setup requires more effort and preparation by the teaching staff since it is not just an online repository of slides and notes; instead, the materials, questions, course structure and content need to be carefully defined and managed.

2.1 Related Work

A number of experimental and commercial adaptive learning systems exist. SE-Coach [5] for instance bases its adaptive user model on Bayesian Networks and provides guidance for "self-explanation" by relying on a probabilistic model using explicit and implicit information gathered from the learner. AHA! [7] maintains a user model and provides adaptive link annotation, hiding, and disabling. Apart from use as an educational tool, the authors also suggest it is useful in information kiosks.

InterBook [11] is a tool that is intended for "authoring and delivering adaptive electronic textbooks", although it is used for course delivery as well. It provides adaptive guidance, navigation and help. MetaLinks [13] is similar and supports "focused and exploratory" learning. It provides active reading (high level reading, searching and problem solving), multi-level books (same book for beginners as well as experts), multi-themed books (a book that would cover different aspects and traditionally requiring multiple books), deep content representation (different ways to show content), avoiding hypermedia side effects (allowing users to jump between pages and making sure they are not lost), evolving content and student modeling.

NetCoach [14][19][20] is another adaptive learning system that is designed to enable authors to develop a course without requiring programming knowledge. It can adapt to user knowledge, goals and preferences and provides curriculum sequencing and adaptive annotation of links. The adaptive behavior of the system is specified by providing pre-requisites and inferences among various course concepts and groups of exercises or tests. Any multimedia supported by web browsers (e.g.: Applets, Animations etc) can also be used. NetCoach supports communication between students and with course staff using email, chat and discussions.

While all of these tools provide infrastructure services for creating adaptive modules, creating instructionally worthwhile materials is a challenge. There is no documented process for creating adaptive modules. However, Parker suggests a series of steps for "online course preparation" [15]: researching best practices, determining course content, determining the delivery method, developing new or adapting older learning units and tests, and determining grading criteria. Cristea [6] also outlines a systematic approach to authoring such content. Other best practices for online courses are captured in AAHE's Seven Principles [4].

3. MODULE DESIGN STRATEGY

When designing on-line modules, one must take into account the needs of both course staff and students. Instructors need to create a module that meets course objectives, has stimulating instructional content and structure, and can assess student progress and evaluate their performance. The design also affects the student's interaction with the content and their test taking capabilities. We used a five-step strategy inspired by Parker [15]:

1. Researching tools and delivery methods: There are a number of adaptive learning technologies to choose from, both experimental and commercial. A few are presented in Section 2.1 and Weber provides a good overview [20]. Choosing the most appropriate system is important and the reasons for selecting NetCoach for our implementation are presented in Section 4.

2. Determining course concepts and design: A course can be organized as a graph with concepts as nodes and interrelationships as edges. The relationships are prerequisites, where knowledge of one concept is necessary for learning another, and inferences, where mastery of one concept implies mastery of another. Creating adaptive courses can require as much as two to three times as much content as traditional courses in order to support an effective level of individualization [3]. To achieve this, an incremental model [6] is most appropriate.

Once the concept hierarchy has been created, it can be organized as an adaptive module. To make use of the adaptive capabilities within our module we adopted the following structure:

1. A Pre-test at the beginning of a module gauges prior student knowledge and builds the initial user model. This test covers representative questions from all sections of the module. Prior knowledge is an important factor in learning. Weber [21] presents quantitative evidence that students using adaptive systems save up to 80% on interaction time while performing 10-20% better on final tests compared to students working without these aids.

2. Concepts and exercises make up the bulk of the module and are presented to the learner adaptively as the student progresses through the module. We chose to group exercises into three "self-tests" capping each third of the module's content. Automatic analysis of learner performance and identification of deficiencies is used to update the user model after each exercise or self-test. Remedial concepts can then be presented to learner. This process shapes the module's navigation to fit the student's performance.

3. A Post-test is used to assess student learning once they have completed all the concepts they were guided through. This tests the student on their knowledge of the whole module.

3. Writing questions: Good questions can help students to "reflect on readings, develop problem-solving skills, form concepts, or simply practice foundational skills. Questioning focuses attention and guides study, reading, writing, communication, visualization, design, development, and other learning activities" [16]. Many adaptive learning systems support a range of question types, including True-False, Multiple-Choice,

Table 1. Assigning difficulty levels to questions

Category	Difficulty
Knowledge: recalling previously learned material (facts, theories)	0.5
Comprehension: grasp the meaning of the material (translating, interpreting material or estimating future trends, effects)	0.6, 0.7
Application: use learned material to create new situations (application of rules, methods, concepts, principles or laws)	0.8, 0.9
Analysis: break down material into its component parts and understand the organizational structure. (Identification of parts, relationships)	1.0, 1.1
Synthesis: put together parts to form a new whole (production of a new theme, plan of operations, research proposals, new patterns, …)	1.2, 1.3
Evaluation: judge the value of material (report, writing, poems, art) and define the criteria for doing so	1.4, 1.5

Testing software based purely on its expected behavior, without looking at the underlying code, is called:

◘ Glass-box Testing
◘ Clear-box Testing
◘ Black-box Testing
◘ White-box Testing
◘ Grey-box Testing

Figure 1: Example question at the "Knowledge" level

Consider the following code:

```
public String foo( int a, int b ) {
  if ( a <= b ) return foo( a + 5, b - 1 );
  else return a + " ";
}
```

Will a test case that calls foo(0,18) and expects "20 15 10 5 0 " as the result pass?

◘ Yes
◘ No

Figure 2: Example question at the "Analysis" level

Forced-Choice and Fill-in-the-Blank questions. Many also support instant feedback, explanations for correct or incorrect choices, or allow multiple attempts when students are learning.

When writing questions, the general guidelines to follow are to ensure that each question relates to a specific concept so that its correct answer implies knowledge of that concept. The questions may imply knowledge about multiple concepts as well. Other guidelines are to have questions catering to varied knowledge levels, and not purely focused on factual recall.

To update the user model, an adaptive system must know which concept a question relates to and how strongly a correct answer implies concept mastery. This usually means assigning some form of difficulty rating to questions. For instance a hard question answered correctly implies a stronger knowledge of the concept(s) it was based on than an easier question. Bloom's Taxonomy [1] presents a useful scale for rating question difficulty. The adaptive system we selected uses a 0.5 (easiest) to 1.5 (hardest) difficulty scale. We defined the rating scheme shown in Table 1 for our modules, directly relating difficulty to the level of cognitive performance required. Examples of questions at different levels are shown in Figures 1 and 2.

4. Recording conceptual relationships: In an adaptive module, a given concept may be reachable through different paths in different contexts and a hierarchical index does not capture these alternative paths [12]. Students are guided to (or away from) specific paths based on their current knowledge and the web of relationships between the concepts in the module. As a result, it is important to record correct prerequisite relationships and inference relationships between the concepts that have been defined. When a concept has one or more prerequisites, the adaptive system will only suggest it when the student has mastered the prerequisite(s). When a concept is inferred from another, the adaptive system will not steer the student toward the concept if there is implicit evidence it has already been covered.

5. Evaluating student performance: The post-test serves as the primary means of evaluating student performance. Since our adaptive modules supplement class room teaching, student performance on traditional exams will also be useful.

4. NETCOACH EXPERIENCE

NetCoach is an adaptive learning system designed to enable authors to develop a course without any programming knowledge. NetCoach provides a web interface for course developers to develop and maintain concepts, questions, and their interrelationships. NetCoach uses a four-layered user model and the these layers store whether the student has visited the concept, whether the student has completed exercises on the concept, whether the current concept is inferred from a previous completed one, and whether the student has marked the concept as completed. Each layer can be marked independently of the others.

NetCoach provides passive guidance to students. Links are annotated using different colors: Red (prerequisites not satisfied, the link is not recommended yet), Green (the student is ready to pursue this content, the link is recommended) or Grey (the student has completed this concept). However due to the passive nature of annotations, the student may visit any page at any time and may choose to disregard navigational advice.

To date, we have implemented two adaptive modules using NetCoach, with a third in the design stage and a fourth in the planning stage. Although NetCoach offers some quirks for module authors and designers, it offers a reasonably clean and easy to use interface for students. The primary effort involves planning out and writing course content and appropriate assessment questions, rather than in using the tools.

5. PRELIMINARY EVALUATION

Empirical studies on the use of adaptive hypermedia in education provide contradictory results [9][10][18]. Effectiveness of hypertext learning itself depends on many factors such as previous proficiency in the domain, the complexity of the domain and the learner's motivation. Researchers argue that it is easy to get lost in the maze of hyperlinks in an online course and the only conclusive statement is that there is no "clear link between adaptivity and an improvement in learning."

However our adaptive learning modules are designed to supplement current classroom teaching and existing student practices. Replacing classroom teaching with adaptive modules is not our goal. As a result, we are interested in evaluating how our modules as a whole will affect student learning when used this way, as well as how we can improve them.

Our first adaptive module, “Introduction to Testing”, is designed to help freshmen learn the basics of testing their own software. We regularly require freshmen to test their own work, and grade them in part on how thoroughly and correctly they test. While students are capable of performing these tasks, we felt that additional support materials on basic testing knowledge would help bring all students up to a better performance level.

With the goal of assessing the effectiveness of our supplemental materials, we designed a set of objective questions covering the range of difficulty levels on the material in our module. These questions did not appear on any of the module’s tests or self-tests. This set of questions was given to 63 students in our CS2 course at the end of the spring 2005 semester to collect a baseline for comparison before the first use of our module. Eighteen students in the summer 2005 offering of the same course were required to complete the on-line module over a period of about ten days as a regular homework assignment. The same assessment questions were then administered at the end of the summer session.

While the students taking the spring offering are regular students, students enrolled in CS2 during the summer session often include more transfer students, students repeating the course, or students who are behind schedule. Often, summer session students are somewhat weaker academically. Despite this difference, our performance comparison was encouraging.

Figure 3 shows the distribution of student scores achieved on our benchmark questions for students in the spring semester that did not use the module, and those in the summer that did use it. There is a clear shift in scores, with means that are significantly different at the 0.01 level (t-score 3.08, df = 19). This comparison provides evidence that students gain some knowledge after going through the adaptive module that they are not already getting in the regular course content or by writing software tests on their own. While this is not surprising, it does confirm that working through the module’s content has a positive effect.

Because the module’s content is presented adaptively, however, we were also interested in how much time students spent with the module, and whether or not the amount of time was related to their scores on the module’s post-test. Figure 4 shows a scatter plot of student pre- and post-test scores against the total amount of time each individual spent in the module. While the pre-test scores are scattered, the post-test scores are in a narrow band relatively close to the top. These data give some indication that spending more time on the module does not directly correlate with better performance (Pearson’s correlation 0.28).

Figure 5 directly compares pre-test and post-test score distributions for students who completed the module. The difference in the means is only significant at the 0.10 level (t-score 1.75, df = 16), perhaps due to the small sample size. Interestingly, the Pearson’s correlation between pre- and post-test scores is only 0.12, making it difficult to attribute performance to student knowledge prior to completing the module.

Figure 6 summarizes performance as students progressed through the module by comparing scores on the pre-test, the three self-

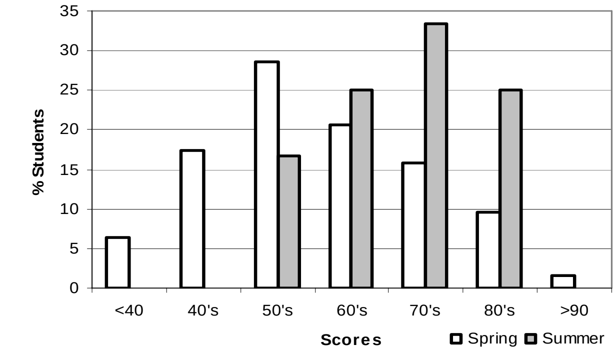

Figure 3: Comparing scores with and without the module

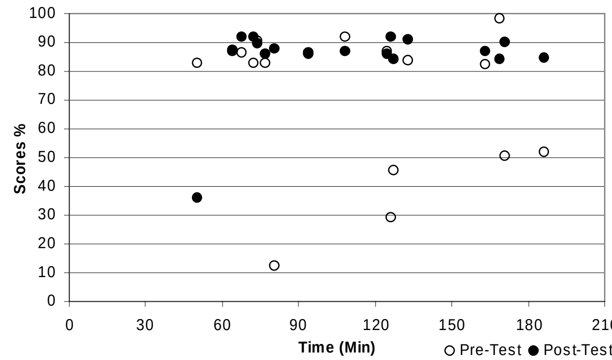

Figure 4: Pre-test and post-test scores vs. time spent

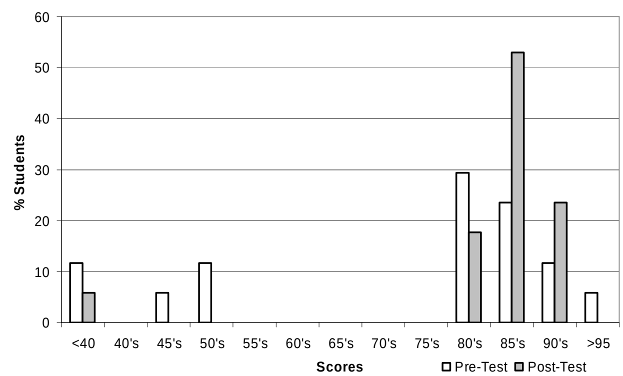

Figure 5: Score distributions on the pre-test and post-test

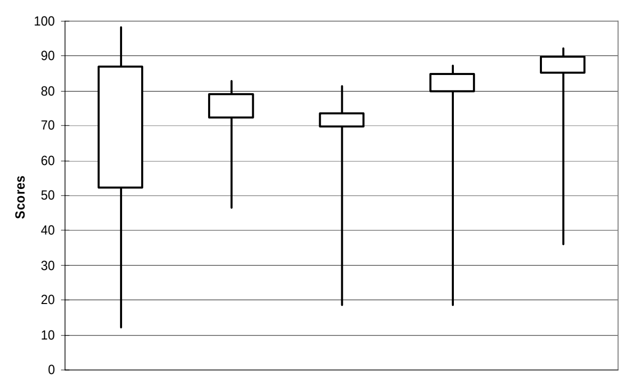

Figure 6: Minimum, maximum, 25th percentile and 75th percentile scores across all module tests

tests, and the post-test. The end-points of each candlestick show the maximum and minimum scores, while the body of each candle extends from the 25% percentile to the 75% percentile. While students begin the module with a much wider variation in knowledge, as they progress through the module, their performance converges to a relatively high level.

Another measure we have considered is a subjective student perception of the adaptive presentation. At the end of the post-test the students had the opportunity to complete a very short survey that asked questions about their perceptions of the module, content, and how their ability to write test cases changed. The results showed that a majority of the students (56%) agreed or strongly agreed that the link annotation was helpful.

Comments such as "well organized pages", "better than most online thingies I have used in the past" and "it gave me a better understanding of software testing" indicate the approach has been effective. On the other hand "A lot of questions make no sense" and "it just told me it was wrong without any reasoning" show that improvements to the test questions are still needed. A student also cautioned "don't try and do this in one day" showing that the amount of content in this module required some time to complete.

6. CONCLUSION

Adaptive modules are useful for self-paced student learning and present new challenges for course developers. This paper presents our experiences in designing a series of adaptive modules using NetCoach, together with a preliminary assessment of the effectiveness of the first module. It appears that the content without our module is successfully conveyed to students, and that students are able to proceed through the material at an individualized pace without adversely affecting their performance. Students appear to appreciate the content of the materials, although some issues remain with test questions. Continuing research is needed to evaluate and improve adaptive course materials and their impact on student learning.

7. ACKNOWLEDGMENTS

This work is supported in part by the National Science Foundation under grant DUE-0127225. Any opinions, conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF. We are also grateful to David McPherson for his assistance and to Dr. Gerhard Weber for his help with NetCoach.

8. REFERENCES

[1] Buck, D., and Stucki, D.J. Design early considered harmful: graduated exposure to complexity and structure based on levels of cognitive development. In *Proc. 31st SIGCSE Technical Symp. Computer Science Education*, ACM, 2000, pp. 75-79.

[2] Brusilovsky, P. Methods and techniques of adaptive hypermedia. *User Modeling and User-Adapted Interaction*, 6(2–3): 87–129, 1996.

[3] Brusilovsky, P. Course sequencing for static courses? Applying ITS techniques in large-scale web-based education. In *Intelligent Tutoring Systems, Proc. 5th Int'l Conf. ITS*, LNCS v.1839, Springer, Berlin, 2000, pp. 625-634.

[4] Chickering, A.W., and Ehrmann, S.C. Implementing the seven principles: Technology as lever. *AAHE Bulletin*, Oct. 1996, pp. 3-6.

[5] Conati, C., and VanLehn, K. Providing adaptive support to the understanding of instructional material. In *Proc. 6th Int'l Conf. Intelligent User Interfaces*, ACM, 2001, pp. 41-47.

[6] Cristea A., and Aroyo L. Adaptive authoring of adaptive educational hypermedia. In *Proc. 2nd Int'l Conf. Adaptive Hypermedia and Adaptive Web-Based Systems*, LNCS 2347, Springer, 2002, pp. 122-132.

[7] De Bra, P., and Calvi, L. AHA! An open adaptive hypermedia architecture. *The New Review of Hypermedia and Multimedia*, v. 4, pp. 115-140, 1998.

[8] Edwards, S.H. Improving student performance by evaluating how well students test their own programs. *J. Educational Resources in Computing*, 3(3):1-24, Sept. 2003.

[9] Eklund, J., and Brusilovsky, P. The value of adaptivity in hypermedia learning environments: A short review of empirical evidence. In *Proc. 2nd Adaptive Hypertext and Hypermedia Workshop,* Report No. 98/12, Eindhoven Univ. of Technology, Eindhoven, 1998, pp. 13-19.

[10] Eklund, J., and Sinclair, K. An empirical appraisal of the effectiveness of adaptive interfaces for instructional systems. *Educational Technology & Society*, 3(4), Oct. 2000.

[11] InterBook: http://www2.sis.pitt.edu/~peterb/InterBook.html

[12] Freire-Morán, M. *Visualization of hypermedia course structures*, Universidad Autónoma de Madrid.

[13] MetaLinks: http://ddc.hampshire.edu/metalinks

[14] NetCoach: http://art.ph-freiburg.de/www/index-e.htm

[15] Parker, B., and Hankins, J. AAHE's seven principles for good practice applied to an online literacy course. *J. Consortium for Computing in Small Colleges*, 17(4): 39-48, Mar. 2002.

[16] Pennsylvania State Univ., Writing effective questions to promote learning: http://tlt.its.psu.edu/suggestions/questionwriting/index.shtml

[17] Pilgrim, C., Leung, Y., and Grant, D. Cost effective multimedia course development. In *Proc. 2nd Conf. Integrating Technology into Computer Science Education*, ACM, 1997, pp. 45–50.

[18] Specht, M. Empirical evaluation of adaptive annotation in hypermedia. *ED428725*, 1998.

[19] Weber, G., Adaptive learning systems in the world wide web. In *Proc. 7th Int'l Conf. User Modeling*, Springer, 1999, pp. 371-377.

[20] Weber, G., Hans-Christian, K., and Weibelzahl, S. Developing adaptive internet based courses with the authoring system NetCoach. In *Revised Papers from Int'l Workshops OHS-7, SC-3, and AH-3 on Hypermedia*, LNCS 2266, Springer, 2001, pp. 226-238.

[21] Weibelzahl, S., and Weber, G. Adapting to prior knowledge of learners. In *Proc. 2nd Int'l Conf. Adaptive Hypermedia and Adaptive Web-Based Systems*, LNCS 2347, Springer, 2002, pp. 448-451.