Addictive Links: The Motivational Value of Adaptive Link Annotation in Educational Hypermedia

Peter Brusilovsky, Sergey Sosnovsky, and Michael Yudelson

School of Information Sciences
University of Pittsburgh, Pittsburgh PA 15260, USA
peterb@pitt.edu, sas15@pitt.edu, mvy3@pitt.edu

Abstract. Adaptive link annotation is a popular adaptive navigation support technology. Empirical studies of adaptive annotation in the educational context have demonstrated that it can help students to acquire knowledge faster, improve learning outcome, reduce navigation overhead, and encourage non-sequential navigation. In this paper we present our study of a rather unknown effect of adaptive annotation, its ability to significantly increase student motivation to work with non-mandatory educational content. We explored this effect and confirmed its significance in the context of two different adaptive hypermedia systems. The paper presents and discusses the results of our work.

1 Introduction

Adaptive link annotation is a popular adaptive navigation support technology. The idea of adaptive annotation is to augment links with personalized hints that inform the user about the current state of nodes behind the annotated links. Usually, these annotations are provided in the form of visual cues employing, for example, contrasting font colors, font sizes, font types for the link anchor or different icons next to the anchor [1]. Adaptive annotation is especially popular in educational hypermedia. A range of annotation approaches has been introduced and explored in adaptive educational hypermedia systems [2; 3; 7; 8; 11; 13]. Empirical studies of adaptive annotation in the educational context have demonstrated that it can help students acquire knowledge faster, improve learning outcome, reduce navigation overhead, and encourage non-sequential navigation [2; 3; 6; 9; 12]. These effects are frequently cited as the *value* of adaptive annotation.

In this paper we present our recent work on a rather unknown effect of adaptive annotation, its ability to significantly increase student motivation to work with non-mandatory educational content. This effect was first discovered during out-of-the-classroom studies of the ELM-ART system [13]. ELM-ART, an adaptive Web-based system for learning the LISP programming language was (and still is) freely available on the Web for anyone interested in learning LISP. The use of the system was not mandatory for any of the users – they worked with the system only as long as they were interested and motivated. The log analysis demonstrated that subjects who were familiar with at least one other programming language visited visibly more pages and solved more exercises and problems when they were working with adaptive link annotations. However, the effect was not statistically significant and the motivational value of adaptive annotation went unnoticed by the community.

We re-discovered the motivational value of adaptive annotation in a classroom study of QuizGuide, an adaptive hypermedia service developed to guide students to the most relevant self-assessment quizzes [5]. The use of self-assessment quizzes was not mandatory – the students were allowed to use them as much as they wanted, in order to check the level of their knowledge and to prepare for classroom quizzes and exams. The original goal of QuizGuide was to increase the quality of learning. While this goal has been achieved, the most striking effect of adaptive annotation we discovered in the study was the remarkable increase of student motivation to work with self-assessment quizzes. With QuizGuide, the students explored more questions, worked with questions more persistently, and accessed a larger diversity of questions. In some sense, adaptive annotations made the quiz work almost addictive: Once the students started a session, they stayed with the system much longer. Average session length and average number of questions attempted during a semester increased significantly [4].

Despite the obvious importance of the motivational value of adaptive annotation, a single study was not deemed sufficient to publicly announce our new respect for adaptive annotations. We spent an additional year exploring this effect and measuring its significance. To check the stability of the motivational effect, we ran another semester-long study of QuizGuide. To check its transfer, we attempted to replicate this effect using different kinds of adaptive annotations and different educational content – program examples. Our studies confirmed our original observations in both explored contexts, demonstrated the significance of the observed effect, and brought a better understanding of some underlying mechanisms. This paper reports on our exploration of the motivational effect of adaptive annotation. We start by introducing the systems that were used as the initial platform for our study, present the results of the study, discuss its importance, and chart some directions for future work.

2 QuizGuide – Adaptive Annotation for Self-Assessment Quizzes

QuizGuide [5] is an adaptive hypermedia service for personalized access to self-assessment quizzes, served by the previously developed system, QuizPACK. It informs students about their current knowledge and learning goals by giving adaptive annotations along with the links to quizzes. QuizGuide groups quizzes available to students into coarse-grained topics (Fig. 1). The link to each topic is annotated with an icon showing a target with (or without) arrows. The number of arrows (from 0 to 3) reflects the student's performance on the quizzes of that annotated topic (no arrows represent no or very little progress, three arrows representing good comprehension). The color of a target encodes the relevance of a topic to the current learning goal of the class. The topics form a prerequisite-outcome structure. Every time new topics are introduced in a lecture, they are annotated with a bright-blue target. Topics that serve as prerequisites for any of the current topics have a pale-blue target. Completed topics are assigned grey targets. Finally, topics that belong to learning goals not yet covered in class are annotated with crossed targets.

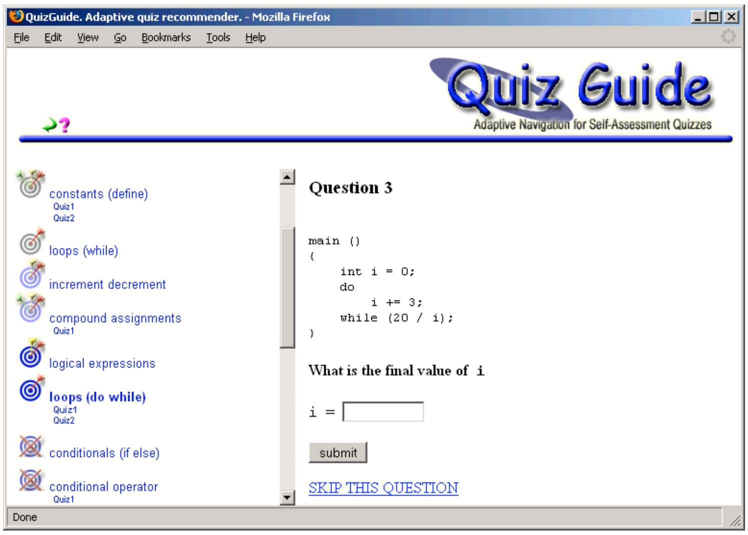

Fig. 1. The **S**tudent Interface for QuizGuide

3 NavEx – Adaptive Guidance for Annotated Code Examples

The NavEx system [14] provides adaptive access to a relatively large set of interactive programming examples. It is built as an adaptive, value-added service for a non-adaptive system, WebEx, which delivers selected examples. The added value of NavEx is that it gives adaptive visual cues for every example link. The NavEx window consists of a navigation frame (on left) and the content area (Fig. 2). Adaptive link annotations presented in the navigation frame integrate several kinds of information about examples and express it through the icon and the font type.

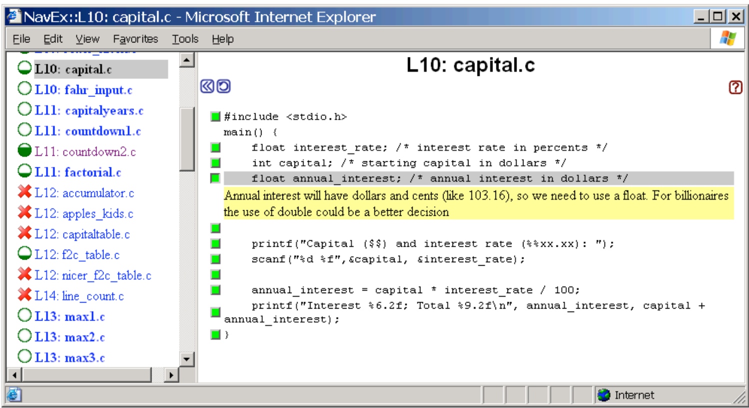

Fig. 2. The Studentf NavEx

The icon/font combination displayed by NavEx performs three functions:

1. categorize examples as ones the student is ready (annotated with green bullets) or not yet ready to explore (annotated with red X icons);
2. reflect the student's progress within each example (displayed as a gradual filling of the green bullet); and
3. emphasize with bold font the most relevant examples to explore.

4 A Study of the Motivational Effect in Two Systems

We performed a number of classroom studies of the non-adaptive systems QuizPACK and WebEx, and adaptive value-added services QuizGuide and NavEx. The studies were done in the context of an undergraduate programming course at the School of Information Sciences, University of Pittsburgh, from the Spring 2003 to the Fall 2005 semesters. The non-adaptive systems, QuizPACK and WebEx, were accessible to students starting in the Spring 2003 semester. QuizGuide was introduced in Spring 2004 and NavEx followed in Fall 2004. After students gained access to adaptive services, non-adaptive systems were still available. Students were able to access the content through the adaptively annotated links provided by new services or in the "old way" - through non-annotated links in the course portal. The set of quizzes and code examples remained to a great extent unchanged across semesters. The setup of the course remained the same across all those semesters, including lectures' content, the scope of home assignments and in-class quizzes.

The introduction of each of the adaptive services caused an impressive increase of student interaction with the supported content (quizzes for QuizGuide and examples for NavEx) as compared to their work with non-adaptive interfaces in previous semesters. Since work with this educational content continued to be non-mandatory, we refer to the observed phenomenon as the motivational effect of adaptive link annotation. The following subsections present some results derived from our study of the motivational effect in the context of QuizGuide and NavEx. Section 4.1 reports the magnitude and the significance of the increase, using several usage parameters, while section 4.2 attempts to look deeper into the process in order to uncover some mechanisms behind this effect.

4.1 Bottom-line Data and Significance

In this section, we will present our quantitative analysis of the added value that the adaptive annotations in QuizGuide and NavEx have over the non-adaptive access in QuizPACK and WebEx. The source data for the analysis were the activity logs collected by the systems. The logs recorded every user click (i.e., submitting an answer to a question or clicking on a line of example code). Data collection procedures did not differ across discussed semesters and were not dependent on method of student access to quizzes or code examples (whether via adaptive or non-adaptive systems). Student work with any of the discussed systems was included equally for user modeling. Log data gave clear evidence as to whether a student accessed quizzes or examples through the adaptive service or not.

We used three variables to parameterize student performance:

1. activity: the number of clicks on lines of code (in the case of WebEx and NavEx) or attempts to answer a quiz question (in the case of QuizPACK and QuizGuide) -- later referred to as clicks or actions;
2. quantity: the number of examples explored (WebEx and NavEx) or quizzes taken (QuizPACK and QuizGuide) --later referred to as examples or quizzes, and
3. coverage: the number of lectures that the reviewed examples or attempted quizzes were drawn from (later referred to as lectures).

Each of these variables was aggregated on two levels:

1. overall performance level – the total number of clicks made, examples/quizzes explored, and lectures covered by each user over the course of the semester; and
2. session performance level – the average number of clicks made, examples/quizzes explored, and average number of lecture topics explored per session by a user.

Our objective was to support our subjective observation that adaptive guidance does make a difference, i.e. to determine whether activity, quantity and coverage of topics were higher for students who were exposed to QuizGuide and NavEx than for those who used the non-adaptive QuizPACK and WebEx format. The results of our comparison clearly demonstrated the value of adaptive navigation support in motivating students to interact more with quizzes and examples.

In our analysis of the value of QuizGuide we compared two semesters (Spring 2003 and Fall 2003) when only QuizPACK was used to access quizzes with two semesters when both QuizGuide and QuizPACK were available to students (Spring 2004 and Fall 2004). When examining the adaptive value of NavEx, we compared two semesters when only WebEx was used by students to access annotated code examples (Fall 2003 and Spring 2004) with one semester during which students could also retrieve examples via WebEx or NavEx (Fall 2004).

During our analysis we inspected about 18,300 QuizPACK and QuizGuide user actions (question attempts) and nearly 3,400 WebEx and NavEx actions (requests for comments about lines of code). Prior to data examination, we performed outlier filtering of individual clicks according to the overall user number of clicks variable. The filtering was done by setting a plus-minus three-sigma interval around the overall number of clicks mean in each of the four semesters. The distributions of data across all our variables were severely skewed because there were a number of not very active students (in terms of clicks made, examples reviewed etc), fewer were moderately active students, and very few very active students. This, along with heterogeneity of variances, prevented us from applying parametric statistical tests to the comparison of usage data. Instead we employed Mann-Whitney tests as t-tests' substitutes.

The results of these tests have revealed that for nearly all variables and aggregation levels, users exposed to the adaptive features of QuizGuide/NavEx achieved significantly higher results (Table 1). First of all, for the QuizGuide+QuizPACK combination vs. the use of QuizPACK alone: In a comparison of the levels of overall parameters, users working with adaptive guidance were making twice as many question attempts a semester (an average of roughly 260 vs. 128), and working with almost twice as many quizzes (an average of roughly 24 vs. 13). This remarkable increase of non-mandatory activity measured by these parameters in combined QuizGuide+QuizPACK over QuizPACK alone was statistically significant. The increase of course coverage was visible, but not statistically significant. Unlike other usage parameters, course coverage has a natural maximum boundary – the total number of lectures in the course. The lack of significant difference shows that students are able to do more work with quizzes while having the same volume of quiz material available. On the level of average user session statistics across all variables, the parameters of QuizGuide+QuizPACK combined were all roughly 1.5–2 times higher then those of QuizPACK alone. All of the observed increases were significant.

Table 1. Comparing the means of the variables for semesters when adaptive systems (NavEx, QuizGuide) were used in combination with non-adaptive ones (WebEx, QuizPACK) with semesters when only non-adaptive systems were used.

		QuizPACK	QuizGuide + QuizPACK	p-value
Overall user statistics	Clicks	127.68±15.97	261.21±53.15	**0.023***
	Quizzes	13.11±1.06	23.97±1.96	**<0.001**
	Lectures	8.70±0.64	10.18±0.61	0.188
Average user session statistics	Clicks	10.48±1.32	17.19±2.03	**<0.001**
	Quizzes	1.87±0.10	3.64±0.49	**<0.001**
	Lectures	1.40±0.05	2.09±0.27	**<0.001**
		WebEx	NavEx + WebEx	p-value
Overall user statistics	Clicks	34.76±6.66	171.90±65.56	**<0.001**
	Examples	5.66±0.87	18.10±4.32	**<0.001**
	Lectures	3.52±0.42	8.20±1.23	**<0.001**
Average user session statistics	Clicks	7.85±0.87	9.49±1.28	0.122
	Examples	1.56±0.12	2.03±0.22	**0.013**
	Lectures	1.20±0.05	1.37±0.10	**0.020**

* boldface indicates p-value less then 0.05

Second, for the NavEx+WebEx vs. WebEx comparison on the level of overall parameters, users exposed to adaptive guidance were making 5 times more clicks a semester (an average of roughly 170 vs. 35), reviewing almost thrice as much code examples (an average of roughly 18 vs. 6), and covering almost twice as larger scope of lectures (an average of 8 lectures vs. 3.5). All of the mentioned advantages on the level of overall user statistics are significant. On the level of average user session statistics the number of clicks users made was higher for those users who were able to use adaptive guidance but not significantly higher. However the number of examples reviewed and number of lectures they were drawn from were significantly higher for users of adaptive guidance.

To ensure that the observed difference in motivation is not explained by differences in student population across semesters, but indicates the added value of adaptive annotation, we analyzed students' initial programming experience and gender across semesters. The results of Chi-Square test show that there was no significant difference between QuizPACK and QuizGuide groups in gender distribution (Pearson Chi-Square = 0.718, p-value = 0.397) or initial experience distribution (Chi-Square = 4.263, p-value = 0.119). Similar results were obtained for WebEx and NavEx groups for both gender (Chi-Square = 1.720, p-value = 0.268) and initial experience (Chi-Square = 0.704, p-value = 0.703).

4.2 A Deeper Look at the Usage Profile

While comparing student work with adaptive and non-adaptive versions, we noticed that typical QuizGuide and NavEx sessions are both *longer* and *more diverse* than QuizPACK and WebEx sessions alone are, correspondingly. Students attempted more questions through QuizGuide and explored a larger number of examples through NavEx. In addition, in both adaptive systems, students more frequently accessed activities corresponding to different lectures within the same session.

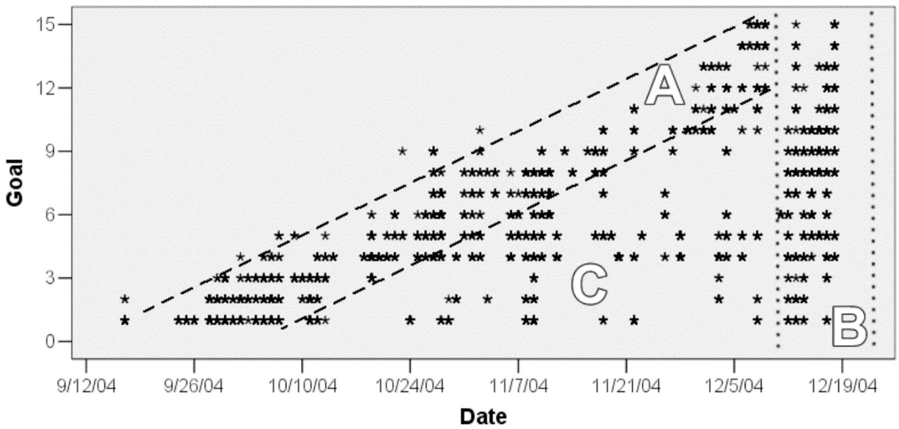

Fig. 3. Time distribution of all actions performed by students with QuizGuide and QuizPACK in Fall 2004 semester. Zone "A" – lecture *stream*, zone "B" – final exam *cut*, and zone "C" – *self-motivated* work with the material of earlier lectures.

To take a closer look at the nature of these results, we performed a deeper analysis of student activity, by taking into account lecture coverage included in all students' actions. Every selection of an example of a question was attributed to the lecture (or the learning goal) it belongs to. For example, Fig. 3 visualizes over 5,500 question-attempts performed by students using QuizGuide or QuizPACK in the Fall 2004 semester. 15 lectures form the vertical axis. The time of the action is marked on the horizontal axis. We can detect three zones of activity. The zone "A" contains all of the "current" activity that students perform along the lecture stream of the course. It is fairly broad, since home assignments and in-class quizzes introduce 1-2 weeks delay in shifting the students' focus from the previous topics. Zone "B" contains a period of preparation for the final exam. The pattern of work with our systems is totally different during this stream of time. Finally, zone "C" contains all actions that students performed during the regular part of the semester, for topics laying far from the "current" lectures. This is the zone we were particularly interested in. All actions here are not directly motivated by the "current" course situation, but rather initiated by the students themselves, possibly in an attempt to bridge the gap in their knowledge that should have been acquired earlier.

We used two measures to assess the intensity of students' *self-motivated* activity: the number of actions in the zone "C" divided by the total number of actions and the average distance between the learning goal which is current at this time and their current activity. For the calculation of the second measure, we used zones "A" and "C".

For all students we calculated an average "C"-ratio and goal distance for quizzes and examples. To evaluate the influence of QuizGuide on motivated activity, we again divided students into two groups: *adaptive* (Spring 2003 and Fall 2003) and *non-adaptive* (Spring 2004 and Fall 2004). Since assumptions of homogeneity of variance and normality were violated, we had to perform nonparametric Mann-Whitney's test. As shown by Table 2, both measures are significantly higher for the adaptive group, which means that students in the semesters when QuizGuide was available more willingly accessed non-current activities. The same analysis was performed for the WebEx/NavEx systems. Their adaptive groups included students of the Fall 2004 semester; while the non-adaptive group combined the Fall 2003 and Spring 2004 semesters. As you can see from the bottom rows of the Table 2, both measures are significantly higher for WebEx+NavEx than for WebEx alone.

Table 2. These parameters characterize the *self-motivated* activity of students with and without adaptive annotations: the "C"-ratio estimates the percentage of students' activity performed outside the current course focus, while the Goal distance assesses how broadly roams (in terms of learning goals) the voluntary interest of a student who is working with the system.

		Non-adaptive	Adaptive	p-value
Quizzes	"C" ratio	0.20±0.03	0.28±0.04	**0.025**
	Goal distance	5.89±0.84	9.56±1.61	**0.026**
Examples	"C" ratio	0.24±0.05	0.51±0.08	**0.005**
	Goal distance	8.73±1.90	17.64±2.51	**0.002**

The observed effect can be explained by the fact that progress-based and prerequisite-based adaptive annotations generated by QuizGuide and NavEx directed students' attention to the material related to earlier lectures, which they did not understand well. In QuizGuide, this guidance is very soft. The zone of recommended work is moving along *class progress* – topics are annotated as ready and current in the week they are presented in a lecture. On the other hand, topics that are prerequisite to current are clearly shown, along with their progress. Student struggling with a current topic can easily focus on prerequisite topics that were not well understood, according the arrow progress measure. The guidance in NavEx is stricter: the zone of recommended work is adapted to the student's *individual progress*. It provides a stronger push to explore insufficiently learned prerequisite concepts in previous examples. This stronger push may cause the observed higher increase of work with non-current activities in NavEx.

5 Discussion

The ability of adaptive annotations to increase student motivation to work with non-mandatory educational content is very important in the context of modern E-Learning. Over the last decade, researchers and practitioners have developed a range of advanced Web-based educational tools such as educational animations and simulations, on-line labs, tutorials, and self-assessment questions. Many developed tools have been evaluated in the lab and small-scale classroom studies and proven to be useful. However, we have now learned that the mere availability of a good tool, although known as beneficial for students, is not enough to ensure its broad educational impact [10]. An important challenge for those who research the use of computers in education is to increase the *effective* use of *student-driven* educational tools. Student-driven tools are created to assist student learning, yet their use is not required and does not count towards the student's course grade. Unlike a variety of assessment-driven tools that the students are required to use in order to complete their assignments, it is up to the students to decide to what degree and how frequently they use the student-driven tools. An instructor might work hard to provide a good set of educational tools of known benefit to the students, only to discover that these tools are dramatically underused. Our work has demonstrated that adaptive annotation can be instrumental in motivating the student to do more work with non-mandatory educational tools which is likely, in turn, to lead to better learning outcomes.

6 Conclusions and Future Work

The results of our studies reported in this paper confirm the motivational effect of adaptive link annotation in educational hypermedia, demonstrate its magnitude and significance, and shed some light on the mechanisms of this effect. We were able to demonstrate this effect in the context of two different personalized access systems. The presence of adaptive annotations caused the increase of several usage parameters. Accessing non-mandatory educational activities through adaptively annotated links, the students explored significantly more activities, worked with them more persistently, and accessed items more broadly distributed over the course lectures.

Added to the earlier report of a similar effect of adaptive annotation in the context of ELM-ART study, our results allowed us to generalize the observations and talk about the motivational value of adaptive link annotation. We consider the results we obtained as both exciting and important. First, it is always exciting to discover new value in a popular technology. Secondly, the ability to significantly increase student motivation to interact more with non-mandatory educational activities turns adaptive annotation into a technology that may become critical to the practical success of a wide range of beneficial educational technologies.

More work is required to determine the borders of the motivational effect and to master its practical use. While this effect was observed with three different kinds of adaptive annotations, the mechanisms were conceptually similar, in that they combined *appropriateness* (too early, too late, just right) and *progress-based* (how much is already done) annotations. While we argue that both mechanisms contributed to the motivational value, we do not have data to confirm it. It is also not clear whether the observed effect is specific to these two annotation mechanisms or can be generalized to other kinds of adaptive annotations (and possibly to other kinds of adaptive navigation support). To answer these questions, we intend to continue our exploration of the motivational value of adaptive link annotation.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grants No. 0310576, 042602, and 0447083.

References

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

[2] Brusilovsky, P. and Eklund, J.: A study of user-model based link annotation in educational hypermedia. Journal of Universal Computer Science **4**, 4 (1998) 429-448

[3] Brusilovsky, P. and Pesin, L.: Adaptive navigation support in educational hypermedia: An evaluation of the ISIS-Tutor. Journal of Computing and Information Technology **6**, 1 (1998) 27-38

[4] Brusilovsky, P. and Sosnovsky, S.: Engaging students to work with self-assessment questions: A study of two approaches. In: Proc. of ITiCSE'2005, Monte de Caparica, Portugal, ACM Press (2005) 251-255.

[5] Brusilovsky, P., Sosnovsky, S., and Shcherbinina, O.: QuizGuide: Increasing the Educational Value of Individualized Self-Assessment Quizzes with Adaptive Navigation Support. In: Nall, J. and Robson, R. (eds.) Proc. of World Conference on E-Learning, E-Learn 2004, Washington, DC, USA, AACE (2004) 1806-1813.

[6] Davidovic, A., Warren, J., and Trichina, E.: Learning benefits of structural example-based adaptive tutoring systems. IEEE Transactions on Education **46**, 2 (2003) 241-251

[7] De Bra, P. and Calvi, L.: AHA! An open Adaptive Hypermedia Architecture. The New Review of Hypermedia and Multimedia **4** (1998) 115-139

[8] Henze, N. and Nejdl, W.: Adaptation in open corpus hypermedia. International Journal of Artificial Intelligence in Education **12**, 4 (2001) 325-350.

[9] Masthoff, J.: Design and evaluation of a navigation agent with a mixed locus of control. In: Cerri, S. A., Gouardères, G. and Paraguaçu, F. (eds.) Intelligent Tutoring Systems. Lecture Notes in Computer Science, Vol. 2363. Springer-Verlag, Berlin (2002) 982-991

[10] Naps, T. L., Rößling, G., et al. Exploring the role of visualization and engagement in computer science education. ACM SIGCSE bulletin **35** (2003) 131–152.

[11] Papanikolaou, K. A., Grigoriadou, M., Kornilakis, H., and Magoulas, G. D.: Personalising the interaction in a Web-based Educational Hypermedia System: the case of INSPIRE. User Modeling and User Adapted Interaction **13**, 3 (2003) 213-267.

[12] Specht, M.: Empirical evaluation of adaptive annotation in hypermedia. In: Ottmann, T. and Tomek, I. (eds.) Proc. of ED-MEDIA/ED-TELECOM'98 - 10th World Conference on Educational Multimedia and Hypermedia, Freiburg, Germany, (1998) 1327-1332.

[13] Weber, G. and Brusilovsky, P.: ELM-ART: An adaptive versatile system for Web-based instruction. International Journal of Artificial Intelligence in Education **12**, 4 (2001) 351-384, available online at http://cbl.leeds.ac.uk/ijaied/abstracts/Vol_12/weber.html

[14] Yudelson, M. and Brusilovsky, P.: NavEx: Providing Navigation Support for Adaptive Browsing of Annotated Code Examples. In: Looi, C.-K., McCalla, G., Bredeweg, B. and Breuker, J. (eds.) Proceedings of AI-ED'2005. IOS, Amsterdam (2005) 710-717